TbISWI and its role in transcriptional control in *Trypanosoma brucei*

Manish Kushwaha

A thesis submitted for the degree of

Doctor of Philosophy

to the Department of Biochemistry, University of Oxford

Wolfson College

2010

Abstract

ISWI is a member of a versatile family of ATP-dependent chromatin remodelling complexes involved not only in transcription regulation (initiation, elongation and termination), but also in other cellular functions like maintenance of higher order chromatin structure and DNA replication. TbISWI, a novel ATPase of the ISWI family in *Trypanosoma brucei*, is involved in the transcriptional repression of silent *VSG* expression sites (ESs) in both bloodstream form (BF) and procyclic form (PF) life cycle stages of the parasite.

Using *in silico* analysis, I have found that TbISWI is well conserved across the eukaryotic lineage, including those members of the order *Kinetoplastida* that do not exhibit antigenic variation. Compared to the ISWIs of higher eukaryotes, TbISWI has greater representation of random coils within its structure, an indicator of more structural fluidity and flexibility of interaction with multiple protein partners.

Using an *eGFP* reporter based assay, I have studied the role of TbISWI in transcriptional repression of silent areas of the *T. brucei* genome. TbISWI was found to be involved in preventing inappropriate transcription of the silent VSG repertoires. TbISWI was also found to downregulate transcription in RNA pol I, but not pol II, transcription units. These results argue for the presence of at least two functionally distinct TbISWI complexes in *T. brucei*.

Using DNA staining and fluorescence *in situ* hybridisation (FISH), I have investigated the potential effect of TbISWI depletion on cell cycle progression and minichromosome segregation. I did not find any evidence for the role of TbISWI in the maintenance of centromeric heterochromatin in *T. brucei*.

Contents

Abstract 2

Contents 3

Acknowledgements 5

1) Introduction

1.1 Chromatin Remodelling as a way of regulating gene expression 7

1.2 Types of Chromatin Remodelling proteins 11

1.3 Multiple roles of ISWI complexes 18

1.4 Trypanosome biology and gene regulation 22

1.5 VSG Expression Sites and their regulation 33

1.6 TbISWI and Expression Site regulation 38

2) Understanding TbISWI- an *in silico* analysis

2.1 ISWI homologs in other kinetoplastids

Introduction *41*

Results *43*

Discussion *46*

2.2 TbISWI domains

Introduction *50*

Results *50*

Discussion *54*

3) Role of TbISWI in transcriptional silencing of VSG genes

3.1 Minichromosomes

Introduction *61*

Results *64*

Discussion *70*

3.2 VSG basic copy arrays

Introduction *75*

Results *75*

Discussion *82*

3.3 Derepression of telomeric VSGs

Introduction *85*

Results *85*

Discussion *86*

4) Role of TbISWI in transcriptional regulation of Pol-I and Pol-II transcription units

4.1 rDNA spacers

Introduction *89*

Results *90*
Discussion *99*

4.2 Procyclin EP1 locus
Introduction *103*
Results *104*
Discussion *108*

4.3 Tubulin locus
Introduction *110*
Results *111*
Discussion *115*

5) Role of TbISWI in chromosome segregation
5.1 Cell Cycle effects
Introduction *118*
Results *118*
Discussion *119*

5.2 Minichromosome segregation
Introduction *122*
Results *124*
Discussion *128*

6) Conclusions **132**
7) Materials and Methods **141**
8) References **162**

Appendices
A1: Opposing T7 promoters 181
A2: TbISWI binds the silent repeat regions 185
A3: TbISWI is an essential protein 186
A4: Laboratory names of cell lines 189

Acknowledgements

I would like to begin by thanking my DPhil supervisor, Dr Gloria Rudenko, for the support, guidance and constant supervision she has made available throughout this project, even during the days of her ill health. I am grateful to Prof Jane Mellor, Prof Rodney Phillips and Dr Paul Klenerman for allowing a generous use of their lab equipment, and to Dr Jay Taylor for his advice with the phylogenetic trees.

I thank Dr Bill Wickstead for being the eternal trouble-shooter, and Dr Klaus Ersfeld and Mohamed Bessat for sharing their expertise with fluorescence *in situ* hybridisation (FISH).

Thanks to Tara Stanne, Alexander Fullbrook, Matthew Wand and Viola Denninger for sharing their unpublished results and figures. Thanks also to all the past and present lab members, Mani Shankar Narayanan, Nadina Vasileva and Megan Povelones included, for being the immense support they have always been. Thanks to Anjana, David, Deepti, Kaushal, Neha, Surya, Susmita and Tahir, all of whom have cushioned me at some point or the other when bumpy road was all that was in sight.

Special thanks to my parents, Rekha and Satish Kushwaha, for their unshakable faith in me, and for always letting me choose my own path.

This research was funded by the Wellcome Trust, and I was supported by the Inlaks Foundation and the Clarendon Fund.

Chapter 1

Introduction

1.1. Chromatin Remodelling as a way of regulating Gene expression

DNA packaging

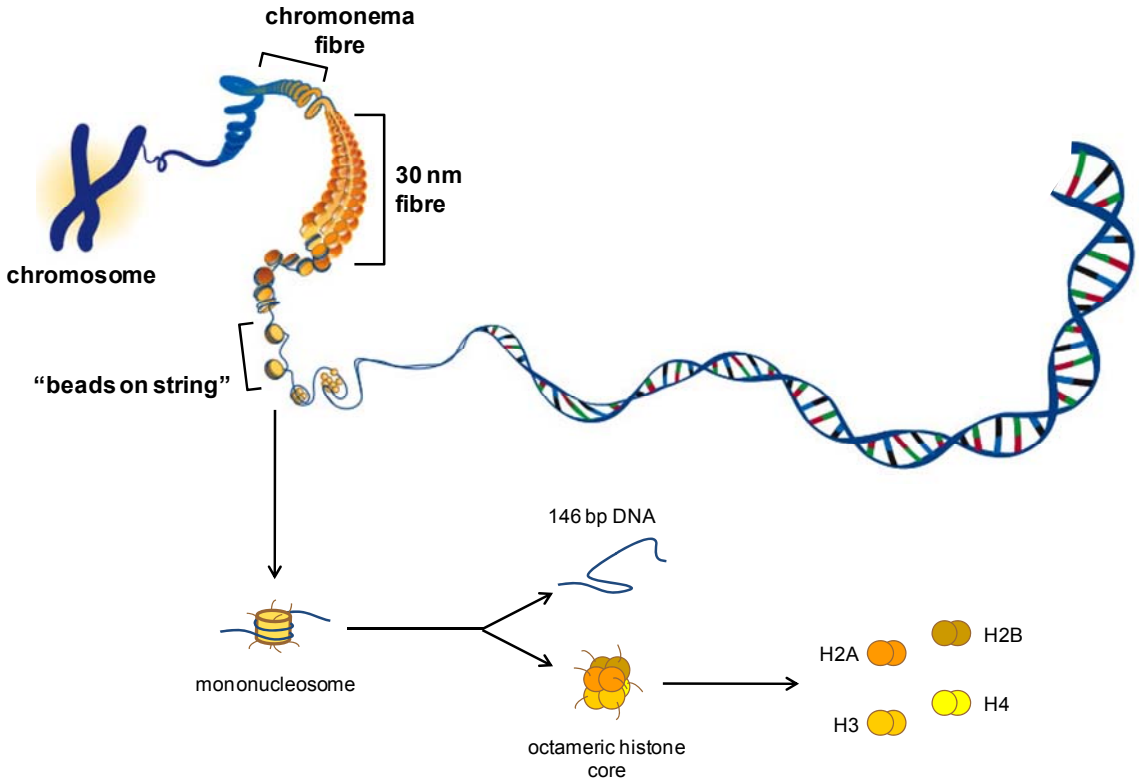


Figure 1.1.1 Multiple levels of chromatin organisation. DNA compaction in chromosome (depicted on the left) of higher eukaryotes occurs through a hierarchy of histone-dependent interactions that can be subdivided into primary, secondary, and tertiary levels of structure. "Beads on string" form of nucleosome organisation comprises the primary structural unit depicted here. Formation of 30-nm fibres through histone tail-mediated nucleosome-nucleosome interactions provides a secondary level of compaction, whereas tail-mediated association of individual fibres produces tertiary structures (such as chromonema fibres). (Adapted from an illustration from The National Human Genome Research Institute made freely available at *http://www.genome.gov/pressDisplay.cfm?photoID=20150*)

DNA of higher eukaryotes is packaged within the nucleus in the form of highly compact chromosomes achieved by association with proteins that fold DNA into successively higher levels of organisation. The schematic in Figure 1.1.1 illustrates the many levels of organisation involved in the packaging of double stranded DNA into condensed chromosomes in a typical eukaryotic nucleus. At the heart of the chromosomal organisation lies the nucleosome, the structural unit of all chromatin (Kornberg 1977).

The nucleosome contains 147 base pairs (bp) of negatively charged DNA wrapped in a left-handed superhelix 1.7 times around a core positively charged histone octamer. The core histone octamer, in turn, consists of two copies each of histones H2A, H2B, H3, and H4 (Luger et al. 1997) . Each histone contains a signature "histone fold" motif sufficient for histone-histone and histone-DNA contacts within the nucleosome, and an N-terminal "tail" domain that contain sites for posttranslational modifications (such as acetylation, methylation, phosphorylation, and ubiquitination). These histone tails appear to emanate radially from the nucleosome, positioned to associate with "linker" DNA or with adjacent nucleosomes (Hansen 2002) .

Chromatin organisation in trypanosomes, although similar, is not identical to that seen in higher eukaryotes. It differs in some important aspects of compaction and histone structure and composition and chromosomes are not condensed (Hecker 1995). These aspects will be discussed in Section 1.4.

Chromatin structure in gene regulation

By organising DNA into higher order chromatin, eukaryotes can effectively compact it into a small nuclear compartment. However, this also restricts the access of DNA-binding transcription factors and requires the cell to strike a balance between effective genome packaging and accessibility (Kornberg & Lorch 1991). This is achieved by a dynamic competition between nucleosomes and transcription factors for important *cis*-regulatory sequences in gene promoters.

This competition is influenced by protein complexes that either "modify" the nucleosomes, or "remodel" them on the DNA. While the chromatin "modifying" complexes covalently alter the histone tails, resulting in differential binding by regulatory proteins and transcription factors, the chromatin "remodelling" complexes reposition,

reconfigure, eject or even reconstitute nucleosomes (Narlikar et al. 2002; Saha et al. 2006). Together these complexes help create promoter architectures with specific density, distribution and composition of nucleosomes with respect to important *cis*-regulatory sites (Cairns 2009). Recent advances in chromatin immunoprecipitation (ChIP) sequencing technology have helped draw genome-wide high resolution (up to 5 bp) consensus maps of nucleosome distribution in yeast, providing a more holistic perspective on chromatin organisation (Jiang & Pugh 2009).

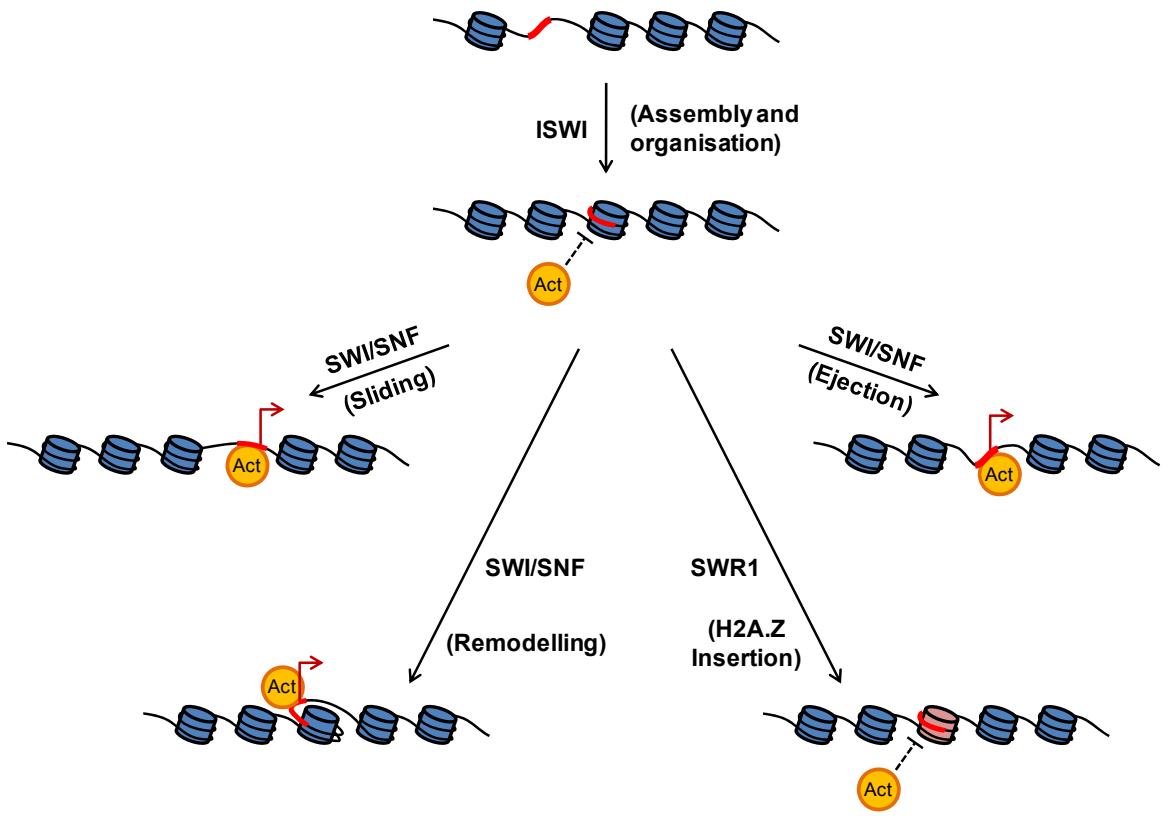


Figure 1.1.2 Functions of chromatin remodellers in facilitating nucleosome dynamics. Chromatin remodelling proteins use energy derived from ATP hydrolysis to drive nucleosome remodelling and are specialized for certain tasks. Most remodellers of the ISWI family help organise the chromatin by providing uniform spacing of nucleosomes. This organisation may cover a binding site (red) for a transcriptional activator (Act). The SWI/SNF family of remodelling proteins provide access to binding sites in nucleosomal DNA, mainly through nucleosome movement, remodelling histone/ DNA contacts or nucleosome ejection. The Swr1 family of remodellers reconstruct nucleosomes by inserting the histone variant H2A.Z into nucleosomes. This can create an unstable nucleosome and might lead to ejection, sliding or reconstruction at promoters. (Adapted from Cairns 2009 and Mohrmann & Verrijzer 2005)

Chromatin remodelling complexes play an important role in constructing the initial chromatin states and then catalysing transitions to alternative chromatin states (Figure 1.1.2), using energy from ATP hydrolysis (Cairns 2009). Different remodeller families influence chromatin configuration differently. The SWI/SNF family of remodellers helps move or eject nucleosomes, usually increasing access to the transcriptional machinery (Martens & Winston 2003). The ISWI remodellers help conduct chromatin assembly and organisation resulting in uniformly spaced nucleosomes and, with some exceptions, cause transcriptional repression (Corona & Tamkun 2004). The Ino80/ Swr1 remodellers can reconstruct nucleosomes by inserting histone variant H2A.Z into nucleosomes thereby specialising their composition (Mizuguchi et al. 2004).

Every chromatin remodelling complex is made up of many protein subunits and results in a different profile of nucleosome positioning at transcription sites (Saha et al. 2006). However, a particular positioning of nucleosomes may not have the same effect at all loci, resulting in the same remodelling complex causing repression at some locus and activation at another (Cairns 2009). The four main classes of chromatin remodelling complexes will be reviewed in the next section.

1.2. Types of Chromatin Remodelling proteins

Nucleosomes are increasingly being viewed as dynamic structural units that have a role in many chromosomal processes, including transcription, replication, DNA repair, kinetochore and centromere construction, as well as telomere maintenance (Kornberg & Lorch 1999; Zlatanova et al. 2009). The structure and positioning of nucleosomes on DNA determines the tightness of packing of nuclear chromatin, which in turn influences the accessibility of DNA to polymerases and other nuclear factors involved in the various cellular processes.

Despite their wide ranging influence, nucleosomes themselves are only structural entities with no inherent catalytic function. Their dynamic properties are due to the action of other proteins like the histone modifying enzymes or the chromatin remodelling complexes (Saha et al. 2006). By covalently altering histone tails, histone modifying enzymes can create markers in the language of a "histone code" that can be read by other regulatory proteins in the cell (Jenuwein & Allis 2001). Chromatin remodelling complexes, on the other hand, restructure, mobilise and eject nucleosomes to directly regulate access to DNA (Owen-Hughes 2003). The coordinated action of the histone modifying enzymes and the chromatin remodelling complexes is important for a variety of biological processes (Geiman & Robertson 2002).

Chromatin remodelling complexes require ATP hydrolysis to alter the contacts between histones and DNA, and these contain an ATPase domain very similar to that found in DNA translocases (Mohrmann & Verrijzer 2005; Saha et al. 2006) . As shown in Fig. 1.1.2., remodellers can mediate (A) nucleosome sliding, in which the position of a nucleosome on the DNA changes, (B) the creation of a remodelled state, in which the DNA becomes more accessible but histones remain bound, (C) the complete dissociation

of DNA and histones, or (D) histone replacement with a variant histone (Mohrmann & Verrijzer 2005). All of the above processes work in concert to establish or alter the properties of chromatin, although the relative importance and order of these processes vary from one genomic locus to another.

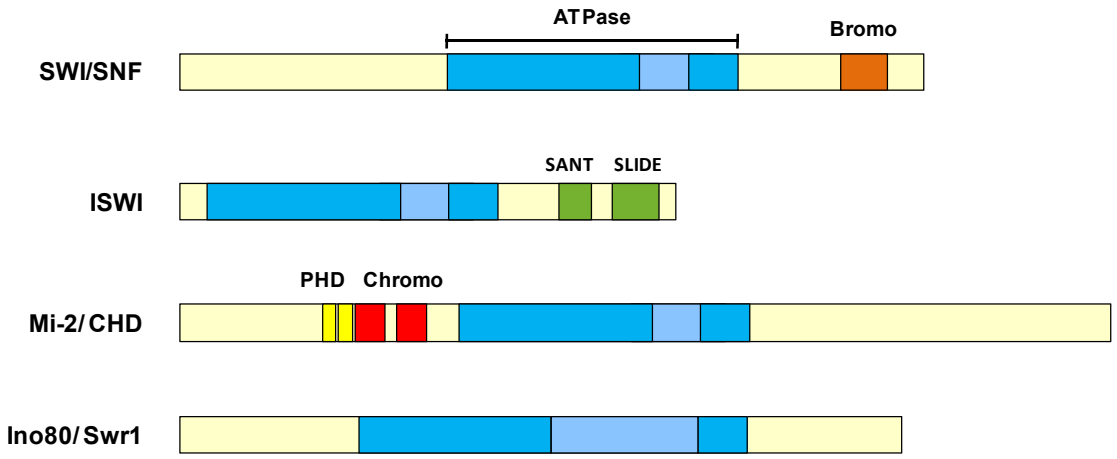


Figure 1.2.1 ATPases of the four main families of ATP-dependent SWI2/SNF2-related chromatin remodelling complexes: SWI/SNF, ISWI, Mi-2/ CHD and Ino80/ Swr1. Each family is defined by the presence of a conserved SWI2/SNF2 ATPase with one or more other signature domains, and a unique subunit composition. (Adapted from Mohrmann & Verrijzer 2005)

On the basis of the presence of a distinct SWI2/SNF2 ATPase with other accompanying signature domains, and a unique subunit composition, eukaryotic chromatin remodelling complexes can be broadly grouped into four different classes: SWI/SNF, ISWI, Mi-2/ CHD and Ino80/ Swr1 (Figure 1.2.1) (Mohrmann & Verrijzer 2005). The constituent subunits of these chromatin remodelling complexes share conserved motifs like chromodomain, bromodomain, plant homeodomain and SANT domain that help target them to specific aspects of chromatin, like the histone tail modifications (Taverna et al. 2007). Sometimes these targeting motifs are present on the ATPase subunit itself and at other times they are present on other subunits of the complex.

Table 1.2.1 Roles of SWI2/SNF2 chromatin remodelling complexes in different species	
SWI2/SNF2-family remodellers	***Biological functions***
SWI/SNF	Pol-II transcription activation and elongation; Transcription activator targeting; DSB repair; Cell signalling; Spindle-assembly checkpoint; Chromosome/ plasmid segregation; Cohesion; Cell-cycle progression; Splicing; Development; Tumour suppressor; Differentiation; Development; Elongation; Signalling
ISWI	Transcription Elongation; Pol II repression; Replication; X-chromosome regulation; Cohesion; Embryonic development and differentiation; Chromatin assembly; Nucleosome spacing; Transcriptional activation
Mi-2/ CHD	Transcription repression and silencing
INO80/ SWR1	DNA repair; Pol-II activation; Homologous recombination; Gene transcription; non-canonical histone deposition; DNA repair
(modified from Saha et al. 2006)	

These chromatin remodelling complexes together mediate a vast array of biological processes (Table 1.2.1) within the cell and have been described below in brief.

SWI/SNF remodellers

The SWI/SNF remodelling complexes were one of the first chromatin-remodelling complexes to be described (Winston & Carlson 1992). These were identified while screening for yeast mutations affecting either mating type switching or growth on sucrose; hence, the names Switching defective (SWI) and Sucrose nonfermenting (SNF) (Stern et al. 1984; Neigeborn & Carlson 1984). When it was recognised that many genes identified as SWI were identical to the SNF genes, the family came to be known as the SWI/SNF family. The SWI/SNF family of remodellers is highly conserved and homologous complexes have been identified in *Drosophila*, mammals and other eukaryotes (Sawa et al. 2000; Tamkun et al. 1992; Wang 2003).

The SWI/SNF family of remodelling complexes is characterised by the presence of a conserved ATPase subunit and a set of additional, conserved core proteins that together define the family (Vignali et al. 2000). For example, the yeast RSC (Remodel the Structure of Chromatin) complex of the SWI/SNF family contains the ATPase subunit

Sth1 and the core proteins Rsc8, Rsc6, Sfh1, Arp7 and Arp9 (Cairns et al. 1996). In addition to the signature ATPase domain belonging to the SWI2/SNF2 ATPase superfamily, the conserved ATPase subunit of this family (Sth1 in the case of the RSC complex) also contains other unique domains like the bromodomain near the C-terminal end that help target the complex to chromatin rich in acetylated lysines (Figure 1.2.1) (Winston & Allis 1999).

SWI/SNF complexes are generally involved in facilitating transcription by disordering uniformly spaced nucleosomes to allow greater access for the transcriptional machinery (Martens & Winston 2003). Consistent with the disordering function, when mononucleosomes are exposed to SWI/SNF complexes, numerous nucleosome products are generated, with the DNA at many different translational positions along the octamer (Whitehouse et al. 1999). However, SWI/SNF complexes have also been shown to have a repressive role in one genome-wide study (Holstege et al. 1998). The documented role of SWI/SNF remodellers in both transcription activation and repression suggests that they do not act as singular monolithic complexes. Instead, each subunit brings unique properties to the multiprotein complex, and combinations of individual subunits form many different kinds of protein complexes with different functions.

ISWI remodellers

The first Imitation Switch (ISWI) complex was identified on the basis of homology between its ATPase subunit and that of the *Drosophila* SWI/SNF BRM remodelling complex (Elfring et al. 1994). Although *Drosophila* has only one ATPase of the ISWI class, it forms at least three different ISWI remodelling complexes- NURF (nucleosome remodelling factor), ACF (ATP-utilizing chromatin assembly and remodelling factor) and CHRAC (chromatin-accessibility complex) with different sets of

partners (Figure 1.3.1) (Bouazoune & Brehm 2006). ISWI remodelling complexes have also been identified in many other organisms, including yeast and humans (Corona & Tamkun 2004).

The ISWI ATPase and the SWI/SNF ATPase subunits share strong homology in their ATPase domains. However, the ISWI ATPase is characterised by the presence of additional domains at its C-terminal end- the closely spaced SANT and SLIDE domains- and unique ISWI-associated subunits (Aasland et al. 1996; Grüne et al. 2003). The SANT (SWI3, ADA2, N-CoR and TFIIB) domain was originally identified based on its homology to the DNA-binding c-Myb domain (Aasland et al. 1996), and the SLIDE (SANT-like ISWI domain) domain was identified as a special case of the SANT domain (Grüne et al. 2003). It has recently been argued that the specific recognition targets (histone tails or linker DNA) of these domains may differ from one species to another, depending on the charge distribution on the two domains (Pinskaya et al. 2009).

ISWI complexes are generally involved in chromatin assembly by ordering and spacing of nucleosomes following DNA replication, facilitating transcriptional repression in both RNA polymerase I and RNA polymerase II transcribed loci (Corona & Tamkun 2004). However, there are also chromatin loci where ISWI complexes promote transcriptional elongation (Morillon et al. 2003). Like the SWI/SNF complexes, it appears the ISWI complexes too are capable of both facilitating and repressing transcription depending on the constituent subunits and the chromatin context. ISWI remodellers will be discussed in more detail in Section 1.3.

Mi-2/ CHD remodellers

The first CHD remodelling ATPase was identified and cloned from mouse as containing a characteristic chromodomain (Chromatin organisation modifier domain)

(Delmas et al. 1993). Following this, related proteins were also found in *Drosophila* and later the human Mi-2 (dermatomyositis-specific antigen recognised by the patient Mitchell autoimmune antibodies 2) antigen was also identified as being related to the CHD ATPases (Woodage et al. 1997).

The ATPase subunit of the CHD family of remodelling proteins is characterised by the presence of a pair of chromodomains at its N-terminus, a highly conserved ATPase domain and a less conserved C-terminal DNA-binding domain. Additional features differentiate the subfamilies or individual proteins. Some ATPase subunits have one or two PHD (plant homeodomain) zinc finger domains at the N-terminal end and no DNA-binding domain at the C-terminus. Other ATPases contain a SANT domain as the C-terminal DNA-binding domain (Hall & Georgel 2007).

The CHD1 complex colocalises extensively with active forms of RNA polymerase II in *Drosophila* (Srinivasan et al. 2005) and its chromodomains have been found to interact with transcription activating H3K4Me tails in humans (Sims et al. 2005). However, CHD1 has also been found to associate with NCoR, a transcription repressor, as well as with HDAC1 and HDAC2 (Tai et al. 2003). These observations suggest that like the SWI/SNF and ISWI remodelling complexes, the CHD family is also involved as both a positive and negative regulator of gene expression. An additional feature of the CHD complexes is their tissue specificity (Thompson et al. 2003) and involvement in developmental regulation (Shur et al. 2006).

Ino80/ Swr1

The Ino80 ATPase was originally found in yeast as a gene required for transcription activation of the INO1 gene, a gene induced by the absence of inositol (Ebbert et al. 1999). It is the most recent addition to the chromatin remodelling

complexes, and its homologues have already been identified in yeast, flies, plants and mammals (Morrison & Shen 2009).

The Ino80 ATPase is believed to represent a new class of ATPases due to the presence of a unique ATPase domain that is split into two subdomains with a spacer region that is important for association with other subunits of the complex (Wu et al. 2005). On the N-terminus is present the HSA (helicase-SANT-associated) domain and next to it is the PTH (post-HSA) domain. The HSA and PTH domains interact with Actin and actin-related proteins Arp4p and Arp8p, all of which are subunits of the Ino80 complex. (Conaway & Conaway 2009)

The Ino80 complexes notably are unique among the other remodellers in exhibiting ATP-dependent helicase activity *in vitro*. The Ino80 ATPase is also stimulated both by DNA and nucleosomes *in vitro*. (Shen et al. 2003; Shen et al. 2000) *In vivo*, Ino80 influences a rather dramatic nucleosome eviction, facilitating transcriptional activation (van Attikum et al. 2007; Tsukuda et al. 2005). Another remarkable feature of the Ino80 complexes is their ability to exchange canonical histones with histone-variants in chromatin (Krogan et al. 2003; Mizuguchi et al. 2004). Ino80 complexes have also been found to assist in double-strand break repair, influence checkpoint pathways and chromosome segregation (Morrison & Shen 2009).

1.3. Multiple roles of ISWI complexes

Imitation Switch (ISWI) ATPases are one of the most highly conserved chromatin remodelling ATPases, and can be distinguished from other SWI2/SNF2 related DNA-dependent ATPases by the presence of the SANT and SLIDE signature domains (Corona & Tamkun 2004). Since their first identification in *Drosophila* (Elfring et al. 1994), many have been found in various other organisms including yeast and humans (Figure 1.3.1).

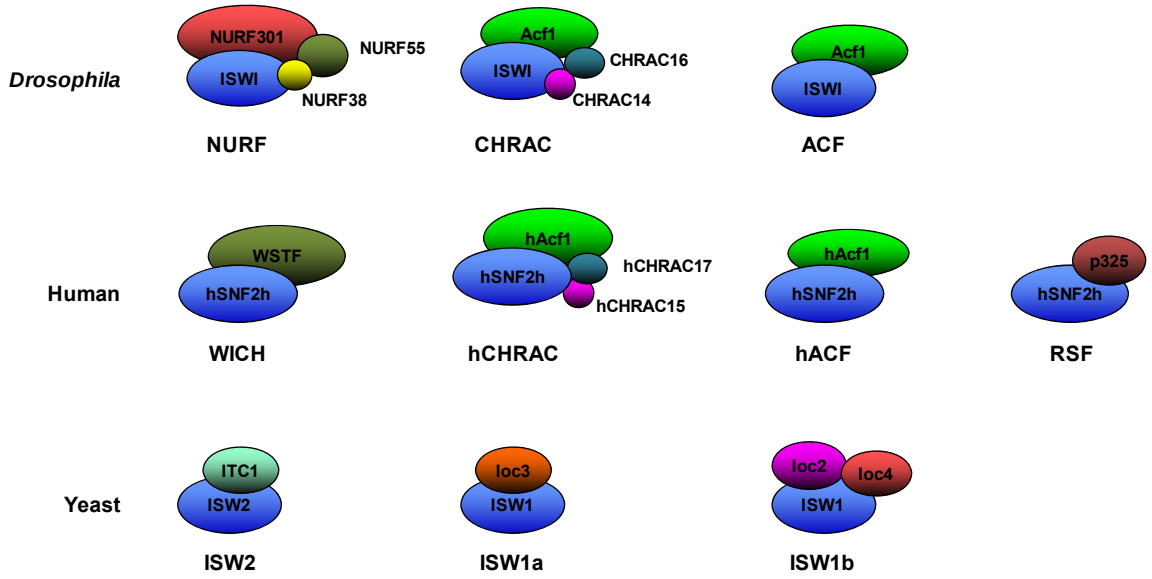


Figure 1.3.1 ISWI family of chromatin-remodelling complexes. ISWI ATPase in these complexes is represented by the blue ball, while the other protein partners are indicated with other coloured balls. The same ISWI ATPase containing complexes can be involved both in transcriptional repression, as well as activation, depending on the other partners in the complex. (Adapted from Corona & Tamkun 2004)

In vitro studies have demonstrated that the ATPase activity of ISWI is stimulated by the presence of nucleosomal DNA, but not by free DNA or free histones (Corona et al. 1999). Although it can facilitate chromatin assembly by sliding mediated nucleosome spacing, the chromatin remodelling activity of the ISWI ATPase is quite limited by itself (Längst & Becker 2001). Therefore, ISWI, like other chromatin remodelling ATPases, functions in multiprotein complexes whose partner composition determines the specific role of the ISWI complex. The same ISWI ATPase can associate differently to give rise

to multiple ISWI complexes with specialised roles- Isw1a and Isw1b complexes of yeast have different roles in transcriptional regulation despite sharing a common ATPase (ISW1) subunit (Figure 1.3.2) (Mellor & Morillon 2004).

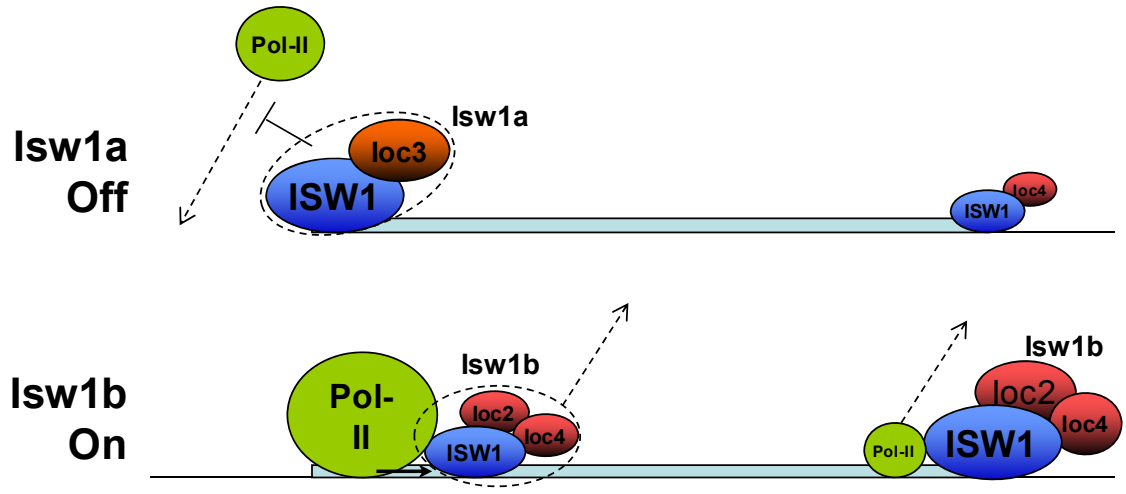


Figure 1.3.2 Role of Isw1a and Isw1b in transcription initiation, elongation and termination. The Isw1a complex (ISW1/Ioc3p) is associated predominantly at the TATA region and excludes RNA Pol-II and the basal transcription machinery from the promoter. In contrast, the Isw1b complex (ISW1/Ioc2p/Ioc4p) regulates transcription elongation from the RNA Pol-II that has initiated transcription at the promoter at the early coding region and also facilitates transcription termination at the 3′ end of the gene. The size of the ISW1 oval with respect to that of the Pol-II represents the relative amount of ISW1 and RNA Pol-II at the early coding region and the 3' end of genes. (Adapted from Mellor & Morillon 2004)

The two characteristic N-terminal domains of the ISWI ATPase, SANT and SLIDE, have been shown to interact with DNA and nucleosomes in the chromatin. However, the choice of interaction seems to be dependent on the charge distribution on the two domains. While the SANT domain is negatively charged in *Drosophila* and interacts with the nucleosomes, in yeast it is positively charged and interacts with DNA. Conversely, the SLIDE domain is positively charged in *Drosophila* and interacts with the DNA, while it is negatively charged in yeast and interacts with the nucleosomes (Pinskaya et al. 2009).

ISWI complexes have a crucial role in chromatin assembly following DNA replication (Corona & Tamkun 2004). Their role as transcriptional regulators is also

defined by their ability to organise and order nucleosomes on the DNA (Deuring et al. 2000). In some chromatin contexts, the ordering role of ISWI results in basal level repression of the chromatin whereas at other loci, nucleosome ordering may expose particular sites to transcriptional activators thereby facilitating transcription (Saha et al. 2006).

Drosophila ISWI (DmISWI) forms three different chromatin remodelling complexes- NURF, ACF and CHRAC- whose subunit composition is shown in Figure 1.3.1. While NURF facilitates a GAGA factor mediated transcription activation of Ultrabithorax and engrailed gene loci (Badenhorst et al. 2002), ACF and CHRAC can mediate both transcription repression and activation (Ito et al. 1997; Varga-Weisz et al. 1997). Similarly, yeast ISWI Isw1a and Isw1b complexes behave differently with respect to their role in transcription regulation despite having a common ISW1 ATPase. While the association of the Isw1a complex (ISW1/Ioc3) excludes RNAP-II and basal transcription machinery from transcription initiation at the promoter, the Isw1b complex (ISW1/Ioc2/Ioc4) regulates transcription elongation at the early coding region and also controls transcription termination at the 3' end of genes (Figure 1.3.2) (Mellor & Morillon 2004).

Apart from transcription regulation, ISWI has also been found to be involved in the maintenance of higher order chromatin structure as well as in replication. ISWI deletion results in defects in male X chromosome organisation in *Drosophila* (Deuring et al. 2000), and in mammalian cells, ISWI complexes facilitate efficient progression of DNA replication through condensed chromatin (Collins et al. 2002). Recent work has also discovered functional coordination between ISWI and SWI/SNF chromatin remodelling complexes (Erkina et al. 2010).

ISWI chromatin remodelling complexes are versatile and can perform a variety of chromatin related functions in the cell, depending on the context of the chromatin in question and its assistant protein partners. In this respect, the function of the ISWI ATPase may be viewed as that of an energy source for driving the remodelling reaction while the protein partners bring a specialised function to the complex. TbISWI, the subject matter of this study, has been recently identified in *Trypanosoma brucei* and its protein partners are hitherto unknown. This work aims to understand the role of TbISWI in transcription regulation in *T. brucei*.

1.4. Trypanosome biology and gene regulation

Sleeping Sickness and trypanosomes

The African trypanosome, *Trypanosoma brucei*, is a parasitic protozoan endemic to sub-Saharan Africa that causes sleeping sickness in humans and nagana in livestock. It is transmitted between its mammalian hosts by the tsetse fly vector (*Glossina spp.*). Sleeping sickness affects 70 million people and about 50,000 cases are reported annually (Beaglehole et al. 2004). If untreated, the disease is fatal in humans and, depending on the subspecies, can cause an acute infection (*T. b. rhodesiense*) lasting a few months or a chronic one (*T. b. gambiense*) lasting many years (World Health Organization 1998).

T. brucei shows extreme adaptation to its environment, differentiating into different life cycle stages with discrete morphologies and patterns of gene expression (Matthews 1999). Two widely studied stages of the *T. brucei* life cycle are those that occur in the blood of the mammalian host, called the slender bloodstream form stage (BF), and in the midgut of the insect vector, called the procyclic form stage (PF). Both of these environments are extremely hostile, and *T. brucei* has evolved specialised strategies to deal with them. In the mammalian bloodstream the parasite is faced with constant attack by the host immune system- both innate and adaptive (Baral 2010). In the insect gut, *T. brucei* must protect itself against secreted proteases, lectins, trypanolysins and other anti-microbial peptides (Stiles et al. 1990; Osir et al. 1995).

GPI-anchored surface coat proteins play a critical role in allowing BF and PF trypanosomes to endure extreme environments. In BF trypanosomes, the cell surface is covered by a protective coat of variant surface glycoprotein (VSG) monolayer of a specific antigenic type. VSG coats are continuously switched to a different antigenic type, allowing the trypanosome to escape destruction by the host immune response (Baral

2010). In PF trypanosomes, the protective layer formed on the cell surface by procyclin coat proteins of the EP and the GPEET types is resistant to the proteases of the tsetse fly (Pearson 2001).

Early evolutionary divergence

After many years of contradictory evidence, it is now clear that African trypanosomes form a monophyletic genus (Lukes et al. 1997; Simpson et al. 2006). They belong to the order *Kinetoplastida*, one of the earliest diverging eukaryotic lineages. Such early divergence and the selective pressures experienced by the organism in evolving a parasitic lifestyle have resulted in their becoming unique in many aspects of sub-cellular structure and function. These novel aspects of their cellular biology can be of particular use in developing therapies that exploit the differences between them and their hosts.

Differential gene expression

Multiple studies have shown that differentiation of trypanosomes from one life cycle stage to another involves significant changes in protein expression (Vertommen et al. 2008; Parodi-Talice et al. 2004). However, microarray analyses have identified comparatively few transcript mRNAs (2-6% of the total) that show changes in abundance and many that do, exhibit very modest changes (<10 fold) (Brems et al. 2005; Duncan et al. 2004; Diehl et al. 2002; Koumandou et al. 2008). These observations suggest that in trypanosomes, unlike in most other eukaryotes, gene regulation is not mediated primarily by transcriptional regulation. Instead, trypanosomes control their genes by post-transcriptionally regulating mRNA processing, stability and degradation, as well as by translational control (Clayton 2002). This explains why very few eukaryotic transcription factors and several RNA-binding proteins have been discovered in trypanosomes using bioinformatic methods (Iyer et al. 2008; De Gaudenzi et al. 2003).

While most genes in trypanosomes are regulated post-transcriptionally, some are regulated at the level of transcription (Günzl et al. 2003; Clayton 2002). These mainly include life cycle specific variant and invariant surface coat proteins in the different life cycle stages, re-emphasising the life cycle specific importance of surface proteins for these organisms.

Genome structure and gene organisation

The trypanosome nuclear genome is about 26 Mb and contains a few hundred minichromosomes (50 to 150 kb each), a variable number of intermediate chromosomes (200 to 900 kb each) and eleven diploid large chromosomes (1.1 to 5.4 Mb each) (Figure 1.4.1.) (Wickstead et al. 2004; Melville et al. 1998; Berriman et al. 2005). All of these chromosomes are linear and have protective telomeres at their ends (El-Sayed et al. 2000; Donelson 2003).

Minichromosomes constitute about 10% of all nuclear DNA and consist of a palindromic core of arrays of 177 bp tandem repeat sequences flanked on either end by silent VSG genes at the telomeres (Wickstead et al. 2004; Weiden et al. 1991). They are transcriptionally silent, and serve as reservoirs of silent VSG genes. These VSGs are preferred for antigenic switching events early in parasitemia (Robinson et al. 1999). The intermediate chromosomes have not been characterised in great detail, but some of them seem to show a basic organisation similar to the minichromosomes, with larger and more complex sub-telomeric regions (Wickstead et al. 2004).

The large chromosomes constitute the bulk of the genome and are predicted to encode about 9000 genes (Berriman et al. 2005). Genes are arranged on these chromosomes in the form of long polycistronic arrays where adjacent genes are in the same transcriptional orientation, an organisation reminiscent of bacterial operons

(Clayton 2002; Campbell et al. 2003). However, unlike in bacteria, the neighbouring genes are not functionally related to each other. The transcription units are non-overlapping, and the genes themselves are almost entirely devoid of introns. The large polycistronic precursor transcript is cleaved into individual mRNAs in a trans-splicing reaction, whereby a 39-nucleotide capped SL RNA is added to the 5'-ends and polyadenylation simultaneously occurs at the 3'-ends (Gilinger & Bellofatto 2001; Campbell et al. 2000).

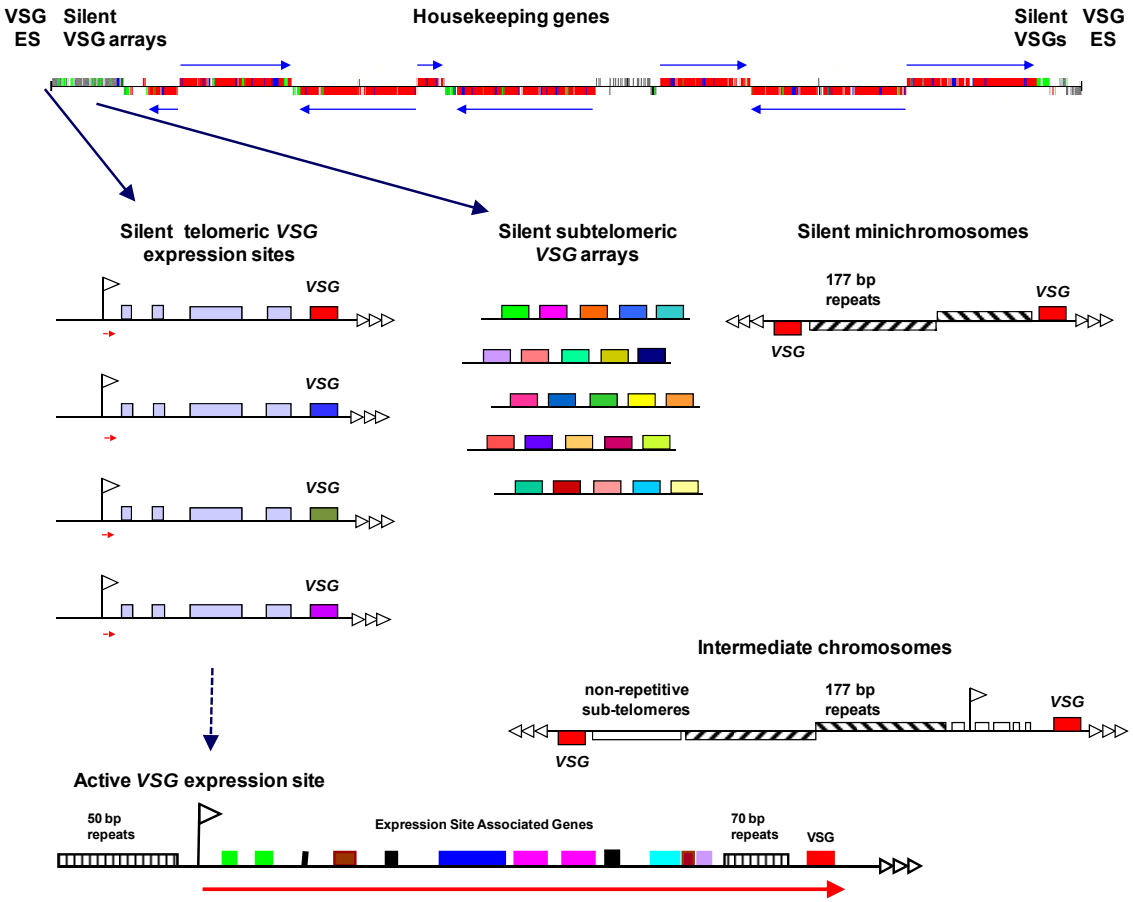


Figure 1.4.1 Schematic diagram of the three size classes of nuclear chromosomes and their molecular anatomy. The topmost panel shows the typical architecture of a large chromosome, with non-overlapping large polycistronic arrays in the central core, silent *VSG* genes in the subtelomeric arrays and the *VSG* expression sites (ES) at the telomeres. The middle panel shows four silent telomeric *VSG* ESs on the left, with the white flags representing the *VSG* ES promoter, the light purple boxes representing the different ESAG genes, four differently coloured boxes representing silent telomeric VSGs, triangles representing the telomere repeats, and the small arrows under the ES promoter representing non-processive transcription. Arrays of differently coloured boxes represent silent subtelomeric VSG genes and pseudogenes in the central middle panel. The middle right panel shows silent mini- and intermediate- chromosomes with the diagonally striped core representing the 177 bp repeat arrays, white boxes representing non-repetitive sub-telomeres, white flag representing an ES promoter, the red boxes representing

the VSG genes, and the triangles at both ends representing the telomere repeats. The bottommost panel shows an active *VSG* ES with the 50 bp and 70 bp repeats shown by vertically striped boxes, the ES promoter by a flag, the ESAG genes by differently coloured boxes, the *VSG* gene by a red box, the telomere repeats by triangles and the processive transcription by a long arrow under the ES.

The housekeeping genes reside in the central core of the large chromosomes, large basic copy arrays of silent VSG genes and pseudogenes are present in arrays at sub-telomeric loci, and the VSG expression sites (ESs) are located at the telomeres (Figure 1.4.1). In contrast to the VSG genes in the telomeric ESs, most of the VSGs in the basic copy arrays are oriented away from the telomeres with 70-bp repeats located upstream (Berriman et al. 2005).

RNA polymerases

Eukaryotes have three different RNA polymerases (pol I, II and III), each dedicated to a defined set of genes. RNA pol I transcribes the ribosomal RNA (rRNA) genes, pol II transcribes the protein coding genes as well as many small nuclear RNAs (snRNA), and pol III transcribes the tRNA genes, the 5S rRNA and some snRNAs. Homologues of all three RNA polymerases have been found in African trypanosomes, although their functions are slightly different (Palenchar & Bellofatto 2006).

T. brucei RNA pol II transcribes all protein coding genes, with the exception of those encoding the surface coat proteins and their associated genes. Pol II also transcribes the "spliced leader" (SL) RNA which is added to the 5'-ends of those mRNAs. As is found in other eukaryotes, *T. brucei* RNA pol I transcribes the rDNA genes into rRNA that forms an important component of ribosomes. In addition, trypanosome RNA pol I also transcribes other transcription units including the surface coat proteins, VSG and procyclin, and their associated genes (Lee & Van Der Ploeg 1997). Trypanosomes are the only organisms known to transcribe protein coding mRNAs using RNA pol I. They are

able to do this because, unlike in other eukaryotes, mRNA transcription in trypanosomes is uncoupled from its capping. Pre-mRNAs synthesised by pol I are effectively capped by addition of a previously capped SL RNA onto their 5'-ends, in a trans-splicing reaction (Günzl et al. 2003). This enables even exogenously introduced bacteriophage polymerases to transcribe protein coding genes in trypanosomes (Wirtz et al. 1994).

Two properties of RNA pol I may be responsible for its selection over pol II for transcribing the surface coat proteins. First, pol II transcription does not appear to be regulated in trypanosomes (Clayton 2002). This makes it unsuitable for transcription of VSG and procyclin, since both the surface proteins need to be tightly regulated in a life cycle specific fashion. Secondly, pol I promoters in trypanosomes appear to be about an order of magnitude stronger than pol II promoters (Biebinger et al. 1996). Since about 10% of all cellular protein in a BF cell is VSG (Böhme & Cross 2002) and there is only one VSG gene transcriptionally active at a time, *T. brucei* appears to have recruited its strongest promoter for transcription of the VSG expression site. Regulation of abundance of other structural proteins that are pol II transcribed appears to be primarily mediated through allele copy number amplification or transcript stability (Jackson 2007; Clayton 2002).

Promoter architecture

RNA pol I in trypanosomes transcribes the surface coat proteins VSG and procyclin in addition to rRNA (Günzl et al. 2003). Hence, pol I promoters are found not only at the rDNA loci but also at VSG and procyclin expression sites. The four structural domains of the trypanosome rDNA and procyclin promoters show functional conservation with those of the *S. cerevisiae* rDNA promoter- the essential domain I marking the transcription initiation site and constituting the core promoter, upstream

domains II and III forming a bipartite element that facilitates transcription factor (TF) binding, and the further upstream domain IV that is involved in chromatin remodelling (Fig. 1.4.2.) (Zomerdijk et al. 1991; Janz & Clayton 1994; Laufer & Günzl 2001). In contrast, the VSG expression site (ES) pol I promoter that drives the transcription of ES associated genes (ESAGs) and VSG is quite different from the rDNA promoter. It has only two small structural domains that form the core promoter and both are essential for transcription. It lacks any recognisable transcription factor (TF) binding or chromatin remodelling element (Zomerdijk et al. 1990; Vanhamme et al. 1995; Pham et al. 1996). The ES promoter lacking a chromatin remodelling element is more sensitive to the chromatin structure of its genomic context. This has been confirmed by experiments where the ES promoter was found to be more sensitive to its locus of integration than an rDNA promoter when placed in the transcriptionally silent rDNA spacer and VSG basic copy arrays (Horn & Cross 1997).

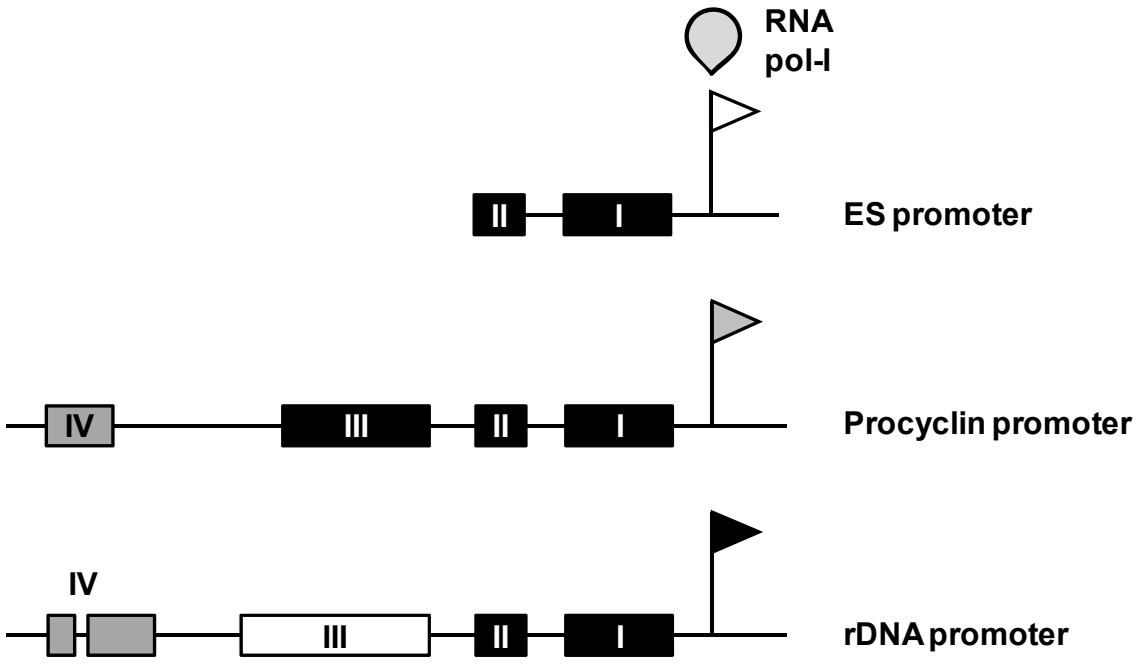


Figure 1.4.2 ***T. brucei*** **RNA pol I promoters.** A schematic of structural domain organisation in the three pol-I promoters of T. brucei is shown above. Domain I marks the transcription initiation site; domains II and III form a bipartite element that facilitates transcription factor (TF) binding; domain IV is involved in chromatin remodelling. (Adapted from Günzl et al. 2007)

The SL RNA promoter is the only RNA pol II promoter that is well characterised in *T. brucei* (Günzl et al. 1997). It shows high sequence conservation with SL promoters from other trypanosomatids and consists of an initiator element at the transcription initiation site and a bipartite upstream sequence element (USE) (Luo et al. 1999). Attempts to identify RNA pol II promoters for protein coding genes have not been very successful in trypanosomes. Studies of the strand-switch (SS) regions between divergent polycistronic transcription units have not found any conserved sequence motifs and reporter assays have suggested only weak bi-directional promoter activity making the results unreliable (Liniger et al. 2001). Recently emerged evidence suggests the pol II promoters in trypanosomes may not be defined by specific sequences, but instead by epigenetic marks. Putative pol II transcription start sites have been found to be enriched in modified histone H4K10ac and histone variants H2AZ and H2BV, while putative termination sites are enriched in histone variants H3V and H4V (Siegel et al. 2009). This could be another departure of trypanosomes from standard eukaryotic biology.

Regulation of pol I transcription

Most of the *T. brucei* genome is transcribed constitutively as polycistronic arrays (Clayton 2002). The only confirmed examples of transcription regulation include the RNA pol I transcribed genes encoding the surface coat proteins (Günzl et al. 2003). In the insect vector, trypanosomes differentiate to the procyclic stage where they express a GPI-anchored invariant surface protein called procyclin. *T. brucei* has about eight to ten EP and GPEET procyclin genes present in clusters of two or three in RNA pol I transcription units (Vassella et al. 2001). Exposure of these invariant proteins to the immune system in BF trypanosomes, would lead to their clearance. To avoid exposure of invariant antigens, these procyclin genes are kept downregulated by multiple layers of control in BF cells.

While transcription downregulation accounts for about ten fold repression (Hotz et al. 1998), another ten fold downregulation is achieved by rapid degradation of any transcript produced (Furger et al. 1997; Hotz et al. 1997).

Apart from the surface proteins, RNA pol I also transcribes rRNA genes in the trypanosome genome. 50-100 rRNA genes, depending on the species, are arranged in multiple pol I transcription units located on several large chromosomes (Castro et al. 1981; Berriman et al. 2005). These units encode the canonical eukaryotic 18S, 5.8S and 28S rRNA (subsequently processed into 28Sα and 28Sβ), and additionally four smaller rRNA genes. Eight of these rRNA transcription units are arranged in tandem arrays separated by untranscribed spacer regions (Berriman et al. 2005).

Indirect evidence from expression of ectopically inserted genes suggests that there are differences in expression from one rDNA unit to another (Alsford et al. 2005). Studies have also shown that the expression of rRNA genes in trypanosomes drops by 4-7 fold as the cells differentiate from BF to PF stage, indicating a life cycle stage dependent transcriptional level control of rRNA gene regulation (Vanhamme, Berberof, et al. 1995). However, details of regulatory mechanisms have remained unclear, in part because of the repetitive nature of the nearly identical ribosomal DNA (rDNA) transcription units. Evidence from other eukaryotes suggests that only half of all rRNA repeats are transcriptionally active at any time (Grummt 2007; Dammann et al. 1995) and the silencing of the inactive rDNA units involves chromatin remodelling complexes like the ISWI complex NoRC (Li et al. 2005; Santoro et al. 2002). Transcriptionally active rDNA units have been found to be largely devoid of nucleosomes in yeast (Merz et al. 2008), an observation similar to that recently made for active VSG expression sites in trypanosomes

(Figueiredo et al. 2010; Stanne et al. 2010). Taken together, these data point to a chromatin-mediated transcriptional regulation of pol I transcribed rRNA genes.

Chromatin structure

In trypanosomes, individual chromosomes are not visible at mitosis due to the lack of their compact condensation (Hecker et al. 1985). However, to accommodate a ~20 cm genome in a 2 μm nucleus their DNA still needs to be considerably packed even during interphase (Renauld et al. 2007). While evidence exists for "beads-on-string" organisation of trypanosome chromatin, no 30 nm fibres are formed (Hecker et al. 1994). Trypanosome DNA is organised around histones in a manner similar to other eukaryotes although its histones are highly divergent, particularly at their N-terminal tails (Alsford et al. 2004). Apart from containing the four canonical histones (H2A, H2B, H3 and H4) and the linker histone H1, four variant histones (H2Az, H2Bv, H3v and H4v) have also been found in trypanosomes (Lowell et al. 2005; Lowell et al. 2004; Alsford et al. 2004). These variant histones have been found to be enriched at putative RNA pol II transcription initiation and termination sites (Siegel et al. 2009).

Histones are not only important for DNA compaction, but also as epigenetic markers that regulate gene expression. Studies on histone N-terminal tail modifications suggest that many modifications otherwise well conserved in eukaryotes are not found in trypanosomes (Janzen, Fernandez, et al. 2006). Some trypanosome specific histone modifications have also been found- for example methylation of N-terminal alanine of H2A, H2B and H4; and acetylation or methylation of the C-terminus of H2A (Mandava et al. 2007; Janzen et al. 2006). While some of these modifications mark boundaries of transcription units, none has yet been shown to be specific to transcriptionally active or repressed regions (Siegel et al. 2009).

Some recent studies have shown differential nucleosome occupancy and histone binding on active versus silent VSG expression sites (Figueiredo et al. 2010; Stanne et al. 2010). While these results provide the first direct link between chromatin structure and gene regulation in trypanosomes, more work is needed for a careful dissection of the mechanisms involved.

1.5. VSG expression sites and their regulation

In the bloodstream of the mammalian host, trypanosomes are covered by a homogeneous coat of variant surface glycoprotein (VSG), a GPI-anchored protein (Cross 1975). The trypanosome genome has an extensive repertoire of more than 1200 VSG genes and pseudogenes (Berriman et al. 2005), of which only one is expressed at a time. Continuous switching of the VSG coat proteins helps the trypanosome evade elimination by the host immune system. The active VSG gene is expressed from one of about 20 telomeric expression sites (ES) (Becker et al. 2004; Young et al. 2008).

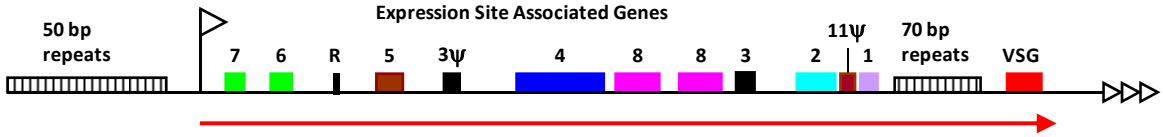


Figure 1.5.1 Schematic representation of an active bloodstream form *VSG* Expression Site (ES). A white flag indicates the ES promoter; transcription is indicated with a red arrow, Expression Site associated genes (ESAGs) by numbered boxes, and *VSG* by a red box. The triangles indicate the telomere. Upstream of the *VSG* gene are the 70 bp repeats and upstream of the ES promoter are the 50 bp repeat arrays.

The VSG ESs are polycistronic transcription units of 30 to 60 kb located at trypanosome telomeres (Figure 1.5.1). A typical ES is characterised by an upstream array of 50 bp repeats (20-50 kb) that insulates the ES from transcription in the rest of the chromosome (Sheader et al. 2003). Downstream of the 50 bp repeats are the ES promoter, and many ES associated genes (ESAGs) that are thought to play a role in host adaptation (Pays et al. 2001; Bitter et al. 1998). At the distal end of the ES is the VSG gene, preceded upstream by 70bp repeats that have a role in recombination at the time of switching (Liu et al. 1983).

While low levels of non-processive transcription has been observed simultaneously from the inactive ES promoters in both bloodstream and insect form

trypanosomes, only one ES is stably activated at a time in bloodstream form trypanosomes (Ulbert et al. 2002; Chaves et al. 1999). The switch from one VSG to another can occur either by *in situ* transcriptional switching to a different ES, or by a gene conversion event resulting in the replacement of the VSG in the already active ES, the latter mechanism being used more frequently (Robinson et al. 1999).

Telomeric positioning

As seen in many other pathogens- bacterial and eukaryotic- trypanosome contingency genes (the VSGs) are found to be associated with telomeres (Barry et al. 2003). Following the reporting of the "telomere position effect" (TPE) in yeast, which involves a localised silencing of genes inserted adjacent to telomeric repeats (Gottschling et al. 1990), the silencing effect of telomeres was also investigated in trypanosomes (Horn et al. 1997; Horn et al. 1995). It was established that a localised gradient of silencing operated on pol I promoters inserted very close to the telomere end (Glover et al. 2006). However, this localised gradient of silencing could be disrupted by cleaving off the telomere, while the silencing operating on the VSG ES remained intact. This suggested that TPE has no role in the monoallelic exclusion of VSG ESs (Glover et al. 2007). Similarly, trypanosome homologues of other proteins known to be associated with TPE in yeast, like SIR2 or the KU70/80 dimer, were found to be involved in the localised silencing immediately around the telomere only, with no role observed in VSG ES regulation (Janzen et al. 2004; Alsford et al. 2007).

However, recent evidence shows that TbRAP1, a telomere binding protein, is critical for the silencing of the inactive VSG ESs. Knocking down TbRAP1 resulted in a derepression of all silent ESs, with a stronger derepression observed in genes closer to the telomeres (Yang et al. 2009). None of the silent ESs was fully derepressed. In yeast,

RAP1 protein has been shown to be involved in regulating telomere length, activating and repressing transcription as well as controlling chromatin boundaries (Morse 2000).

Taken together, these observations suggest that the telomeric location may have some role in VSG ES regulation, although SIR2 and TbRAP1 seem to be only components of a multi-layered regulatory mechanism.

DNA modifications

Trypanosomes contain a modified base: β-D-glucopyranosyloxymethyluracil or base J (Borst et al. 2008; Gommers-Ampt et al. 1993), which is found mainly at the silent VSG expression sites, the telomere repeats, as well as at the 50 bp repeats upstream of the VSG ES promoter. Although present in the 50 bp repeats upstream of an active ES as well as in the telomere repeats downstream, base J is absent from the active VSG ES itself. This observation gave rise to the speculation that base J may be involved in antigenic variation (van Leeuwen et al. 1997). However, the fact that *JBP1* and *JBP2* double knockout (both involved in its biosynthesis) did not disrupt VSG ES repression disproved this idea. At the moment, the role of base J remains unclear, although it has been suggested that J may be involved in repressing homologous recombination between repetitive sequences (Borst et al. 2008).

5-methylcytosine, a common DNA modification found in eukaryotes, has also been found in both BF and PF trypanosomes (Militello et al. 2008). The modification is associated with strong transcriptional repression and impaired gene expression (Klose et al. 2006), although its genomic distribution in trypanosomes is not yet known.

Sub-nuclear localisation

Sub-nuclear localisation of a gene can have dramatic effects on its expression. In *Saccharomyces cerevisiae*, transcriptional silencing of telomeric genes is associated with nuclear envelope positioning (Taddei et al. 2004). In trypanosomes, studies using fluorescence *in situ* hybridisation have shown that telomeric and minichromosome sequences localise in clusters at the periphery of the nucleus (Chung et al. 1990).

It has been argued that in bloodstream form trypanosomes the active ES is housed in an extra-nucleolar subnuclear location, referred to as the Expression Site Body (ESB). The ESB is specific to BF trypanosomes, accommodates a single ES, and is thought to contain the transcription and RNA processing factories playing an important role in the stable expression of the genes within the active ES (Navarro et al. 2001). The ESB was found to remain intact after DNAse I digestion of the nuclei, arguing that it is a sub-nuclear structure positioned outside the nucleolus rather than simply a protein complex binding the active ES. In experiments attempting to force trypanosomes to transcribe two ESs simultaneously, it was found that the distance between the two VSG ESs was smaller than that between control loci. This suggested that the ESs were both competing for the single ESB, and hence rapidly switching back and forth from the expression of one ES to another (Chaves et al. 1999).

While the active ES is located in the ESB in BF cells, the active procyclin genes in PF cells are transcribed at the nucleolar periphery. Upon differentiation of the BF cell to the PF cell, the ESB disappears and the active VSG promoter selectively undergoes a rapid and dramatic repositioning to the nuclear envelope and is subsequently downregulated (Landeira et al. 2007).

Chromatin structure in ES regulation

The importance of chromatin structure in eukaryotic gene regulation has been discussed above in 1.1 and 1.4. Early studies of chromatin structure in trypanosomes focussed on comparing nucleosome phasing on the active versus silent VSG ESs. While it was reported that the active ES was more susceptible to strand specific endonucleases, no substantial differences were found in nucleosomal distribution assessed using micrococcal nuclease (Greaves et al. 1987). Exogenous T7 RNA promoter and polymerase were also used to probe for ES chromatin accessibility. These experiments indicated that there is chromatin mediated silencing in PF trypanosomes but did not find any evidence for this in BF trypanosomes (Navarro et al. 1999). However, recent evidence suggests that the chromatin structure of the active ES is highly depleted of nucleosomes compared to that of the silent ES (Figueiredo et al. 2010; Stanne et al. 2010). These experiments were performed using chromatin immunoprecipitation (ChIP), formaldehyde-assisted isolation of regulatory elements (FAIRE) and qPCR techniques that are far more sensitive than the Southern blots previously used.

New chromatin remodelling and chromatin modifying enzymes have been recently discovered in trypanosomes. DOT1A and DOT1B are chromatin modifying enzymes with H3K76 (the equivalent of H3K79 in yeast and mammals) dimethylation and trimethylation activities, respectively (Janzen, Hake, et al. 2006). Deleting DOT1B results in partial derepression of silent ESs and slowing down of the rate of *in situ* activation to another ES (Figueiredo et al. 2008). TbISWI, a member of the ISWI family of SWI2/SNF2 related chromatin remodelling complexes, has also been shown to be involved in transcriptional repression of the silent ESs both in BF and PF trypanosomes (Hughes et al. 2007). TbISWI will be discussed in the next section.

1.6. TbISWI and Expression Site regulation

TbISWI was originally isolated from procyclic form cell lysate using a DNA affinity chromatography screen as binding to the 177 bp minichromosome repeats and was identified by mass spectrometry (Tilston and Ersfeld, unpublished). It is one of the 13 genes in the *T. brucei* genome that have a putative SNF2 domain (Figueiredo et al. 2009), and was the first member of the SWI2/SNF2 related chromatin remodelling ATPases that was shown to have a role in VSG Expression Site (ES) silencing (Hughes et al. 2007). Depletion of TbISWI causes growth reduction after 2 days of RNAi induction in bloodstream form *T. brucei* and 6 days in procyclic form *T. brucei*. It localises to the nucleus in both life cycle stages.

TbISWI was classified as a member of the ISWI family of SWI2/SNF2 related ATPases based on strong sequence homology in the N-terminal SNF2_N and Helicase_C domains as well as structural homology with the C-terminal Myb-like DNA binding domain. However, SANT and SLIDE domains, which are specialised forms of the Myb-domain and form a part of the signature of ISWI ATPases in other eukaryotes could not be identified initially. Nonetheless, TbISWI was classified as a "novel" ISWI based on the available evidence.

In order to study the role of TbISWI in ES silencing, the investigators integrated a reporter gene downstream of a silent ES promoter in bloodstream form and procyclic form cells and monitored its expression after RNAi-mediated TbISWI depletion. Reporter gene expression was found to increase by 30-60 fold in BF (day 3 of induction) and 10-17 in PF (day 6 of induction) *T. brucei*. The authors argued that this derepression effect was not a stress-induced secondary effect as knockdown of other essential genes did not lead to similar derepression.

However, no evidence was found for the activation of the telomeric silent VSGs after TbISWI depletion. In addition, even in its most derepressed state, the promoter proximal reporter gene expression remained ~10 fold lower than a similarly integrated reporter in an active ES, suggesting that TbISWI formed only one component of a multi-layered regulatory mechanism that ensures monoallelic expression of the VSG ES.

The authors speculated that TbISWI has a role in keeping the silent VSG repertoire downregulated by preventing any fortuitous initiation of transcription. Following on from the previous work, TbISWI protein has been analysed in greater detail using *in silico* methods and its role in downregulation of other silent areas of the genome investigated and presented in this thesis.

Chapter 2

Understanding TbISWI- an *in silico* analysis

2.1. ISWI homologues in other kinetoplastids

Introduction

As mentioned in the last section, TbISWI, a member of the ISWI family of SWI2/SNF2-related chromatin remodelling complexes has been shown to be involved in the repression of silent VSG expression sites (ES) in both bloodstream form (BF) and procyclic form (PF) trypanosomes (Hughes et al. 2007). Although the protein was provisionally named TbISWI due to its overall sequence homology to ISWIs from other organisms, the characteristic C-terminal SANT/ SLIDE domains of the ISWI family could not be identified. Instead, a myb-like DNA-binding domain was identified by structural homology. This had led the authors to propose that TbISWI was a "novel" ISWI.

A comparative analysis of other kinetoplastid ISWIs was performed with the aim of answering some relevant questions. Has the rate of ISWI evolution in these early diverging eukaryotes been significantly different from that of other housekeeping genes? How different is *T. brucei* ISWI from other eukaryotic ISWIs? Do the differences between ISWI proteins from *T. brucei* and other eukaryotes extend to other kinetoplastids as well? Is the *T. brucei* ISWI different from the ISWIs of other kinetoplastids that do not show antigenic variation (like *T. cruzi* and *L. major*) of the surface coat protein?

In order to understand the evolutionary relationship between kinetoplastid and other eukaryotic ISWIs, ISWI proteins from sixteen different eukaryotes were used to construct a phylogenetic tree. These organisms were chosen from three groups: (a) kinetoplastid protozoa- *Trypanosoma brucei*, *Trypanosoma brucei gambiense*, *Trypanosoma congolense*, *Trypanosoma vivax*, *Trypanosoma cruzi*, *Leishmania braziliensis*, *Leishmania infantum* and *Leishmania major*, (b) other parasitic

protozoa- *Plasmodium falciparum* and *Toxoplasma gondii*, and (c) other eukaryotic model organisms- *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Mus musculus* and *Homo sapiens*. The diplomonad *Giardia lamblia* was taken along as an outgroup eukaryote for rooting of the phylogenetic tree.

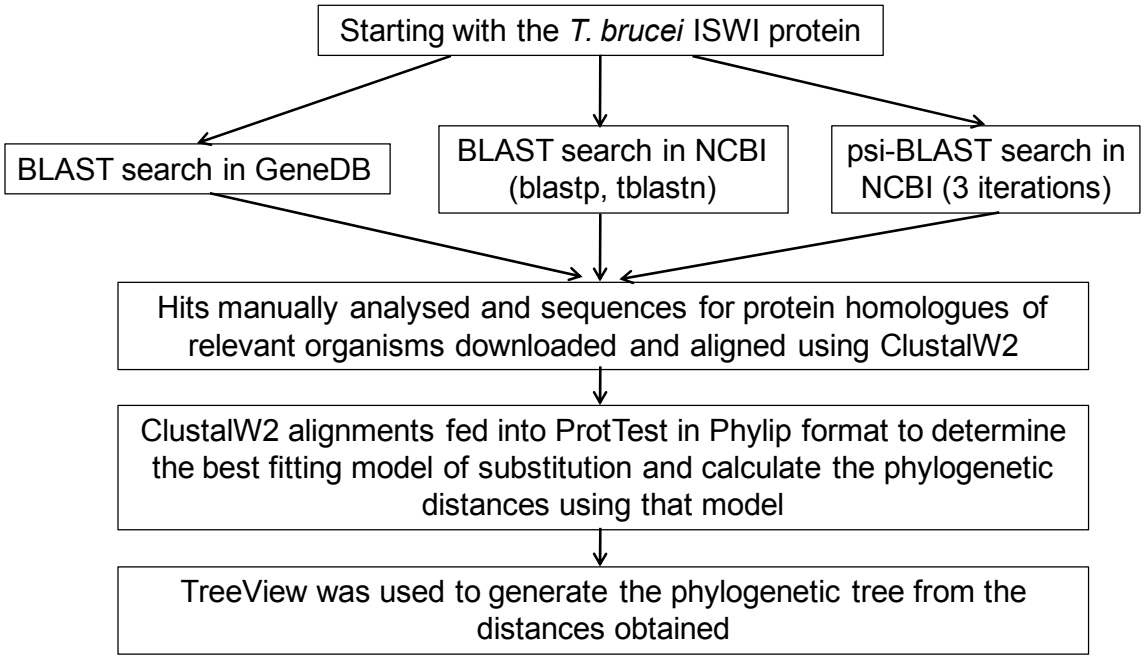


Figure 2.1.1 Flowchart for construction of protein phylogenetic tree. The flowchart shows a step-by-step description of the strategy used for data mining and the construction of phylogenetic trees for ISWI and RPB1 (RNA pol II large subunit) proteins.

To compare the evolution of ISWI across the divergent eukaryotes listed above, it is useful to have an idea of a reference tree that depicts the evolutionary relationship between these organisms. Essential housekeeping genes evolve slowly and can be used as an effective measure of organism level evolutionary relationship for such comparative studies. The large subunit of RNA polymerase II (RPB1), which has previously been used for investigation of eukaryotic phylogeny (Sidow et al. 1994; Friedlander et al. 1994), was used to generate a reference phylogenetic tree for these organisms. Subsequently, the kinetoplastid ISWIs were also scanned for conserved protein domains.

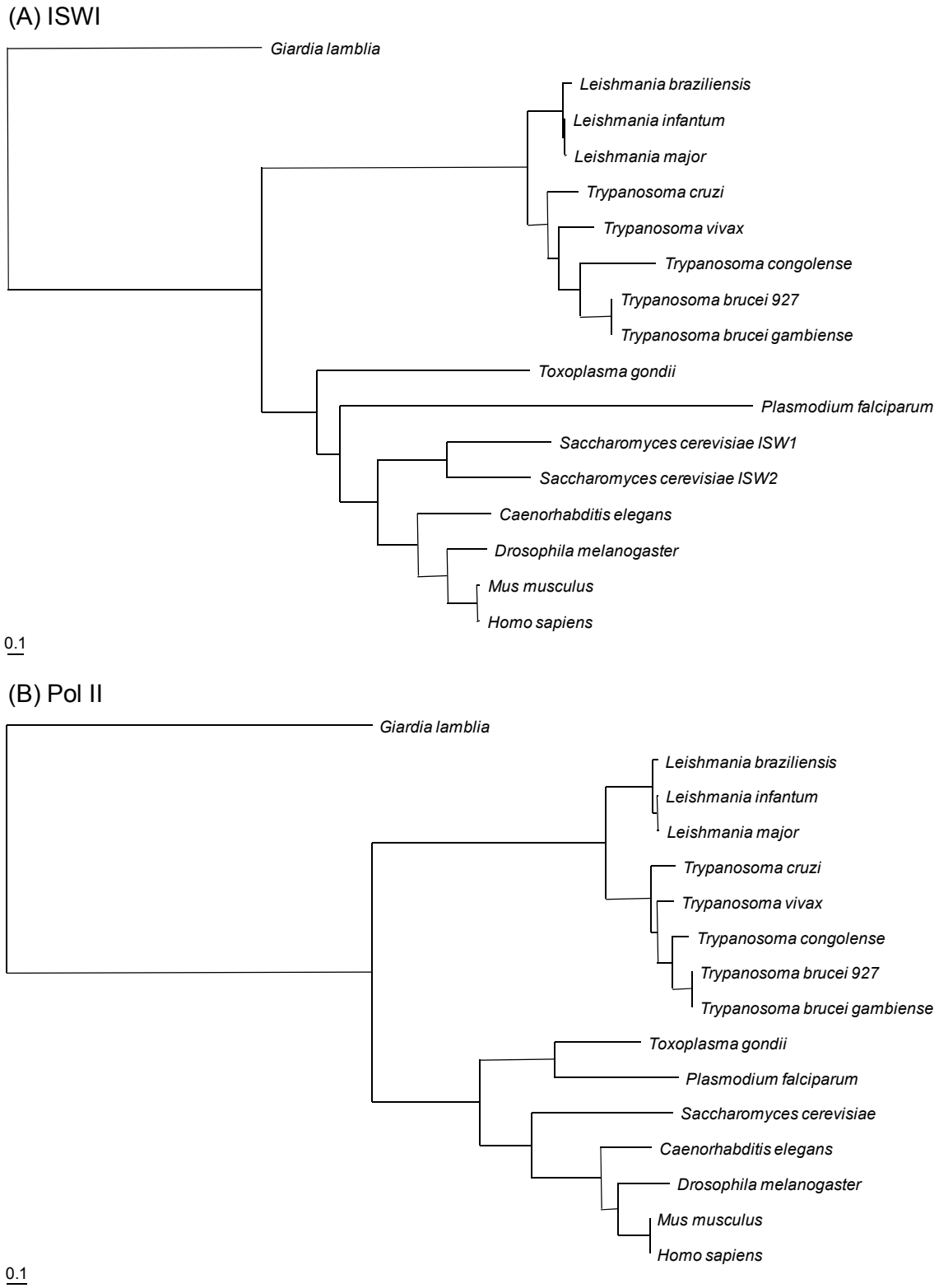


Figure 2.1.2 Evolution of (A) ISWI ATPase and (B) RNA polymerase II large subunit (RPB1). The maximum-likelihood phylogenetic trees show the evolutionary relationship between the ISWI ATPase and RPB1 proteins of organisms from different eukaryotic groups- (a) kinetoplastids- *T. brucei*, *T. b. gambiense*, *T. congolense*, *T. vivax*, *T. cruzi*, *L. braziliensis*,

L. infantum and *L. major*, (b) other parasitic protozoa- *P. falciparum* and *Toxoplasma gondii*, and (c) other eukaryotic model organisms- *S. cerevisiae*, *C. elegans*, *D. melanogaster*, *M. musculus* and *H. sapiens*. *G. lamblia* was used as an outgroup for rooting. Accession numbers of the proteins have been listed in the Materials and Methods.

ISWI homologues were identified in other kinetoplastids, parasitic protozoa and eukaryotic model organisms by BLAST searches using the TbISWI amino acid sequence as query. Most of the hits obtained were non-ISWI proteins showing strong local homology with the conserved SWI2/SNF2 ATPase domain of TbISWI. Using multiple criteria described in the Materials and Methods, the BLAST hits were screened and the relevant proteins were downloaded (Figure 2.1.1). Similar to *T. brucei*, each of the kinetoplastids studied was found to encode only one ISWI protein. RPB1 homologues from the same organisms were also downloaded.

The evolutionary relatedness of ISWIs from the aforementioned organisms was compared with that of RPB1 protein (RNA polymerase II large subunit) as a control. Phylogenetic trees of ISWI and RPB1 protein were constructed using ProtTest (Figure 2.1.1). After comparing the two trees with each other it was found that, with the exception of the *Plasmodium falciparum* branch, the topology of the ISWI phylogenetic tree was very similar to that of RPB1 (Figures 2.1.2 A&B).

After analysing TbISWI, other kinetoplastid ISWIs, and yeast ISW1 and ISW2 with InterProScan (Quevillon et al. 2005), it was found that the SNF2_N and the Helicase_C subdomains that constitute the SWI2/SNF2 ATPase domain of ISWI are very well conserved in all of the ISWI proteins (Figure 2.1.3). A well-conserved HAND domain was also found in all ISWIs, with the exception of *T. congolense* ISWI. The SLIDE domain, one of the two characteristic ISWI C-terminal domains, was also found to

be well conserved. However, InterProScan could identify the SANT domain only in the yeast ISWIs.

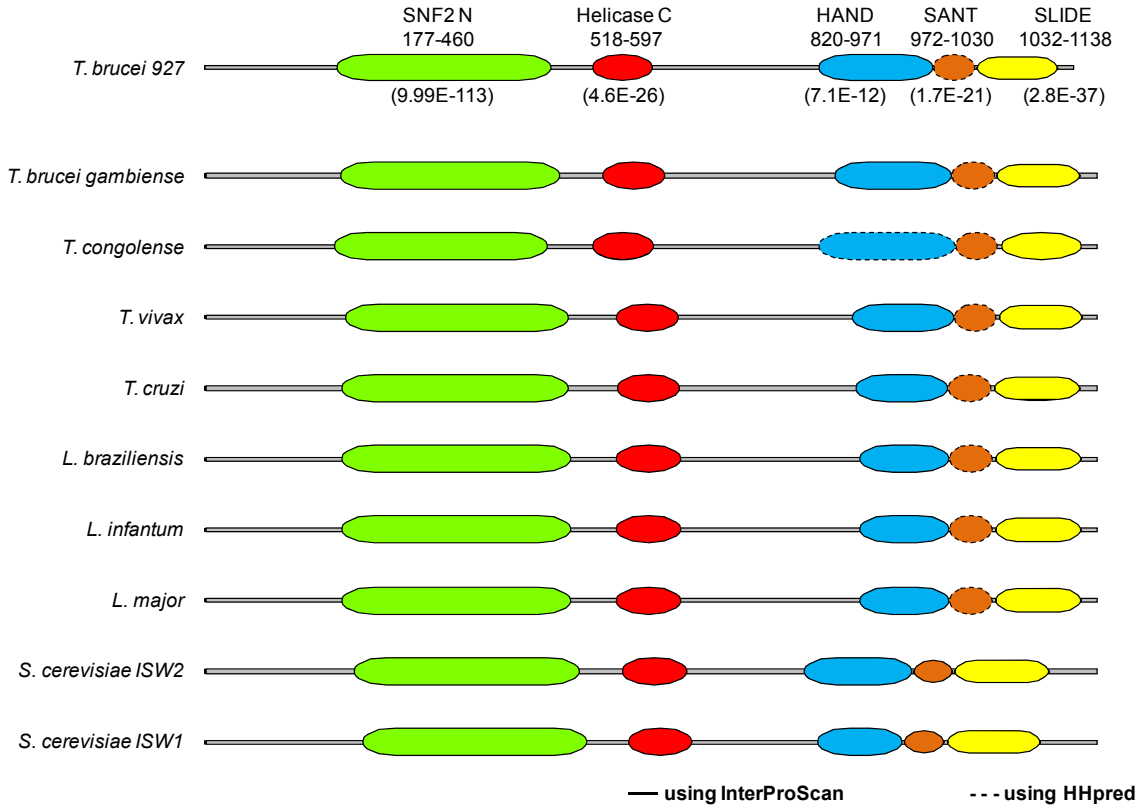


Figure 2.1.3 Domain architecture of kinetoplastid ISWI homologues. The figure shows the predicted domains found in the ISWI ATPases of kinetoplastids and yeast. The green ovals depict the SNF2_N subdomain and the red ovals depict the Helicase_C subdomain of the ISWI SWI2/SNF2 ATPase domain. The blue, orange and yellow ovals depict the HAND, SANT and SLIDE domains. Domains depicted with solid outlines were predicted using InterProScan, while those depicted using dashed outlines were predicted using HHpred (a remote homology and secondary structure based modelling tool). The grey bar depicts the rest of the protein sequence. For TbISWI, the start and end points for each predicted domain, as well as the E-value of prediction obtained from the prediction tool, have also been mentioned.

Further analysis of the C-terminal domains of the ISWI proteins was performed using HHpred, a remote homologue search and secondary structure prediction algorithm (Söding et al. 2005). This identified not only the missing HAND domain of *T. congolense* but also the SANT domain in all the remaining ISWI proteins (Figure 2.1.3).

Discussion

Although housekeeping proteins can be used to build reference phylogenetic trees for comparison, different selection pressures operating on different proteins during evolution mean that the models of amino acid substitution that apply to them are also different. To overcome this problem, the best-fit substitution models were determined separately for ISWI and RPB1 using ProtTest (Abascal et al. 2005). Both ISWIs and RPB1s were found to best fit the LG+I+G+F model (Le et al. 2008; Reeves 1992; Yang 1993; Cao et al. 1994), which was used for constructing maximum-likelihood phylogenetic trees.

It was found that the ISWI phylogenetic tree was very similar in its topology to the RPB1 phylogenetic tree. The only exception was *P. falciparum* ISWI that branched off from the rest of the higher eukaryotic lineage after *T. gondii* ISWI. In contrast, *P. falciparum* RPB1 branched off from the lineage together with *T. gondii* RPB1. This minor difference notwithstanding, the comparative study shows that ISWI protein seems to have evolved at rates of evolution comparable to that of the overall genetic drift across the eukaryotic lineage. No obvious difference was seen between the ISWI and the RPB1 phylogenetic tree in the kinetoplastid branch indicating that the rate of ISWI evolution has not been significantly different between kinetoplastids that exhibit antigenic variation and those that do not.

While some eukaryotes like *Drosophila* and *C. elegans* have a single ISWI ATPase, some others like *S. cerevisiae* have two. Although, the two ISWI proteins (ISW1 and ISW2) in yeast are individually non-essential, triple deletion mutants of *isw1*, *isw2* and *chd1* are temperature and formamide sensitive (Giaever et al. 2002). In ISWI null mutants of *Drosophila* that grow up to late larval stages and then die, the X chromosome

appears to be largely deformed (Deuring et al. 2000). RNAi silencing of the *C. elegans* ISW-1 results in embryonic lethality (Kamath et al. 2003; Sönnichsen et al. 2005). It appears that organisms having a single ISWI gene are entirely dependent on its activity, and hence a deletion mutation of their ISWI renders them non-viable. Since the search for ISWI in kinetoplastids yielded only one ISWI per genome, this suggests that ISWI would also be essential in other kinetoplastids, as was shown for TbISWI (Hughes et al. 2007). Since only one ISWI was found in each of the kinetoplastids, an early branching eukaryotic group, this study also suggests that multiple ISWI paralogs, as seen in yeast, are a result of a more recent duplication event generating greater diversity of function.

The kinetoplastid ISWIs were subsequently scanned for conserved protein domains using InterProScan, a domain signature recognition tool (Quevillon et al. 2005), and HHpred, a remote homology and secondary structure based modelling tool (Söding et al. 2005). The SNF2_N and Helicase_C subdomains were found to be highly conserved. The SLIDE domain was also identified in all ISWIs. The HAND domain too was identified in all ISWIs, although its sequence conservation in *T. congolense* was relatively poor. The characteristic SANT domain was identified in yeast ISWIs by sequence homology and in the other organisms by secondary structural homology.

The SWI2/SNF2 family of ATPases contains a large number of proteins with characteristic ATPase domains that are distinct from other DNA-dependent ATPases (Carlson et al. 1994; Eisen et al. 1995). This family includes proteins involved in multiple functions like DNA repair (RAD5, RAD16), recombination (RAD54), transcriptional control (MOT1, ISWI) and cell cycle control (STH1) (Richmond et al. 1996). ISWI protein forms the ATPase subunit of the ISWI family of SWI2/SNF2 related chromatin remodelling complexes. Since SNF2_N and Helicase_C subdomains together carry the

ATPase/ Helicase activity of ISWI, the high conservation, and hence easy identification, of these domains was quite expected.

The HAND domain had been originally identified in *Drosophila* ISWI as a novel domain, made up of four alpha-helices, that was closely packed with the adjacent SANT domain in the DmISWI (Grüne et al. 2003). Later, it was shown to interact with the DNA entry/ exit site of the nucleosome in the yeast ISW2 complex (Dang et al. 2007). Its good conservation across all eukaryotic ISWIs studied hints at its functional relevance.

The SANT domain was originally identified based on its homology to the DNA binding c-Myb domain, and is one of the two signature domains of a typical ISWI protein (Aasland et al. 1996). In the ISWI structure, the SANT domain is closely juxtaposed against the HAND domain on one side and onto a connecting spacer helix on another (see Figure 2.2.2) (Grüne et al. 2003). Such positioning allows limited surface exposure to the SANT domain, a possible reason why its overall poor sequence conservation may not adversely affect its function.

The C-terminal SLIDE domain was first identified as a special case of the SANT domain in *Drosophila* ISWI, and is one of the two signature domains of a typical ISWI (Grüne et al. 2003). The SLIDE domain was readily identified by InterProScan across all ISWI proteins studied, and was found to be strongly conserved. The strong conservation suggests a functional relevance for the domain.

Results from the above analyses suggest that TbISWI fits the eukaryotic prototype of ISWI proteins, both in terms of its evolutionary history as well as in its domain composition. ISWI proteins can have many different functions depending on their partners in different chromatin remodelling complexes (Mohrmann et al. 2005). Also, the role of the SANT/ SLIDE domains of ISWI can vary from one organism to another

(Pinskaya et al. 2009). In the next section, attempts will be made to understand the role of the SANT/ SLIDE domains of TbISWI.

2.2. TbISWI domains

Introduction

The SNF2_N and the Helicase_C subdomains together form the ATPase/ Helicase domain of the SWI2/SNF2-related ATPase superfamily. Highly conserved and distinct from other DNA-dependent ATPases, this ATPase/ Helicase domain contains the seven sequence motifs characteristic of the DEAD/H superfamily of ATPases (Carlson et al. 1994; Eisen et al. 1995). As seen in Figure 2.2.1, the seven helices are numbered I, IA, II, III, IV, V and VI respectively. All SWI2/SNF2 related chromatin remodelling complexes contain one such SWI2/SNF2 domain containing ATPase protein that provides energy for the chromatin remodelling process, together with other protein partners that determine the specific activity of that complex.

TbISWI ATPase/ Helicase domain shows strong conservation with the ATPase/ Helicase domains of other eukaryotic ISWIs (Hughes et al. 2007). In this section of the thesis, the ATPase/ Helicase domain conservation of other kinetoplastid ISWIs was studied. In addition, the surface charge distribution of C-terminal SANT/ SLIDE domains of TbISWI was compared with that of yeast and *Drosophila* ISWIs.

Results

The conservation of the SNF2_N and the Helicase_C subdomains of ISWI was studied by aligning the kinetoplastid ISWI set (from the previous section) across the seven characteristic conserved ATPase/helicase motifs. The kinetoplastid ISWIs were found to be very well conserved in these domains, including the essential amino acids marked with a star (Figure 2.2.1).

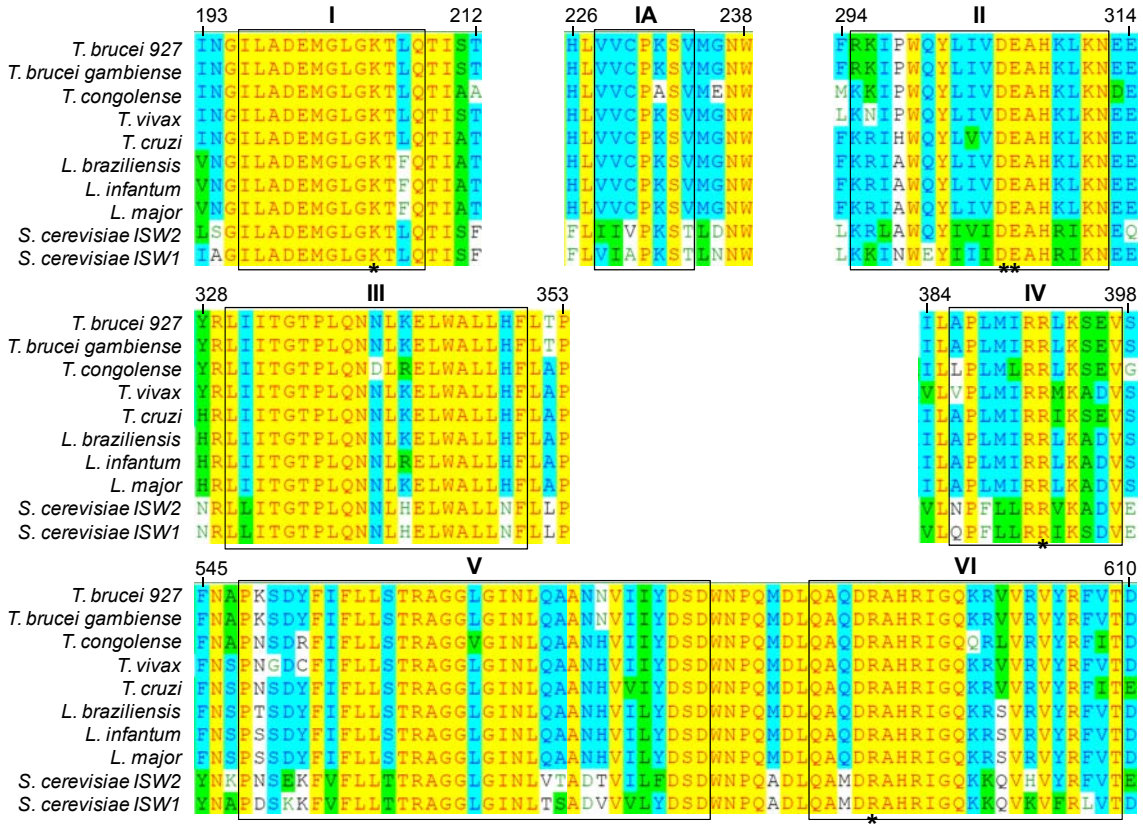


Figure 2.2.1 Alignment of the ATPase/Helicase domain of kinetoplastid ISWI homologues. (based on Richmond and Peterson, 1996) The figure shows an alignment of the ATPase/Helicase domain of ISWI homologues of kinetoplastid protozoa and yeast. The seven conserved motifs are indicated with boxes numbered in bold. The numbers above each alignment block on both the sides indicate the amino acid locus of TbISWI. Residues conserved in all homologues have been indicated with red-on-yellow letters, similar or conservatively substituted residues have been indicated by blue-on-cyan or black-on-green letters, weakly similar residues have been indicated by green-on-white letters and dissimilar residues by black-on-white letters. Stars indicate the essential amino acids.

In order to analyse the structure and function of the C-terminal TbISWI domains more closely, the region of the protein from the HAND to the SLIDE domain as determined in the previous section, was analysed using Phyre (Kelley et al. 2009) for structural modelling along with the corresponding parts of DmISWI and ScISW1. The models obtained suggest that DmISWI and ScISW1 C-termini are very similar to each other, to the extent of being almost superimposable. The TbISWI C-terminus, although similar, is not entirely superimposable, and has larger stretches of unstructured coils and shorter α-helices (Figure 2.2.2).

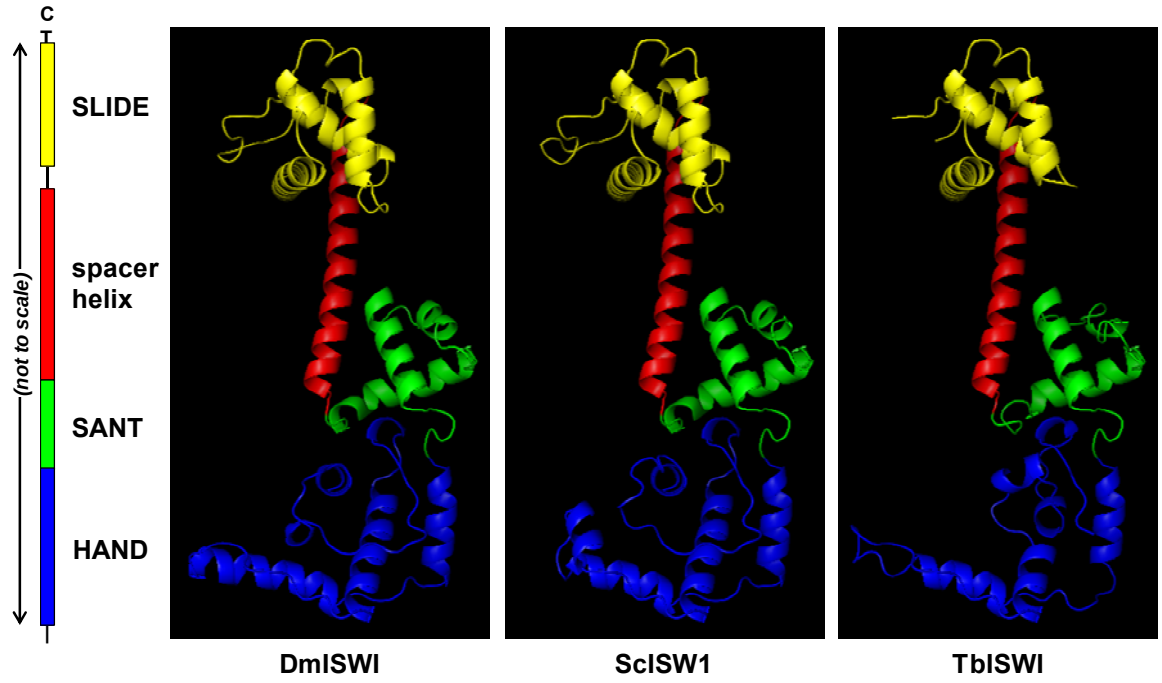


Figure 2.2.2 Cartoon depictions of the structure of C-terminal ISWIs of *Drosophila*, yeast and *T. brucei*. The cartoon depictions were made by modelling the region from HAND to SLIDE domains for these three ISWI proteins using Phyre and then visualising the structure in PyMOL software. As indicated on the left, the HAND domain has been indicated in blue, the SANT domain in green, the spacer helix in red and the SLIDE domain in yellow.

In order to understand the role of the TbISWI SANT and SLIDE domains in comparison to those of DmISWI and ScISW1, the predicted structural models of the individual domains were also obtained using Phyre and their surface charge distribution was determined using APBS tools in PyMOL software (Baker et al. 2001; DeLano 2009).

A typical SANT domain consists of three α-helices (numbered SA1 to SA3 in Figure 2.2.3) separated by short stretches of unstructured coils. Although the TbISWI SANT domain is structurally conserved, its SA1 domain is shorter by two amino acids and the SA2 domain displaced by three amino acids in the alignment. In addition, the coils connecting the three α-helices are longer than those of the other two. The most striking difference between the SANT domains of the three proteins lies in their surface charge. While DmSANT is negatively charged (acidic) on its surface, the ScSANT is

quite positively charged (basic), and the TbSANT surface is made up of a mosaic of negatively and positively charged patches resulting in an average neutral charge.

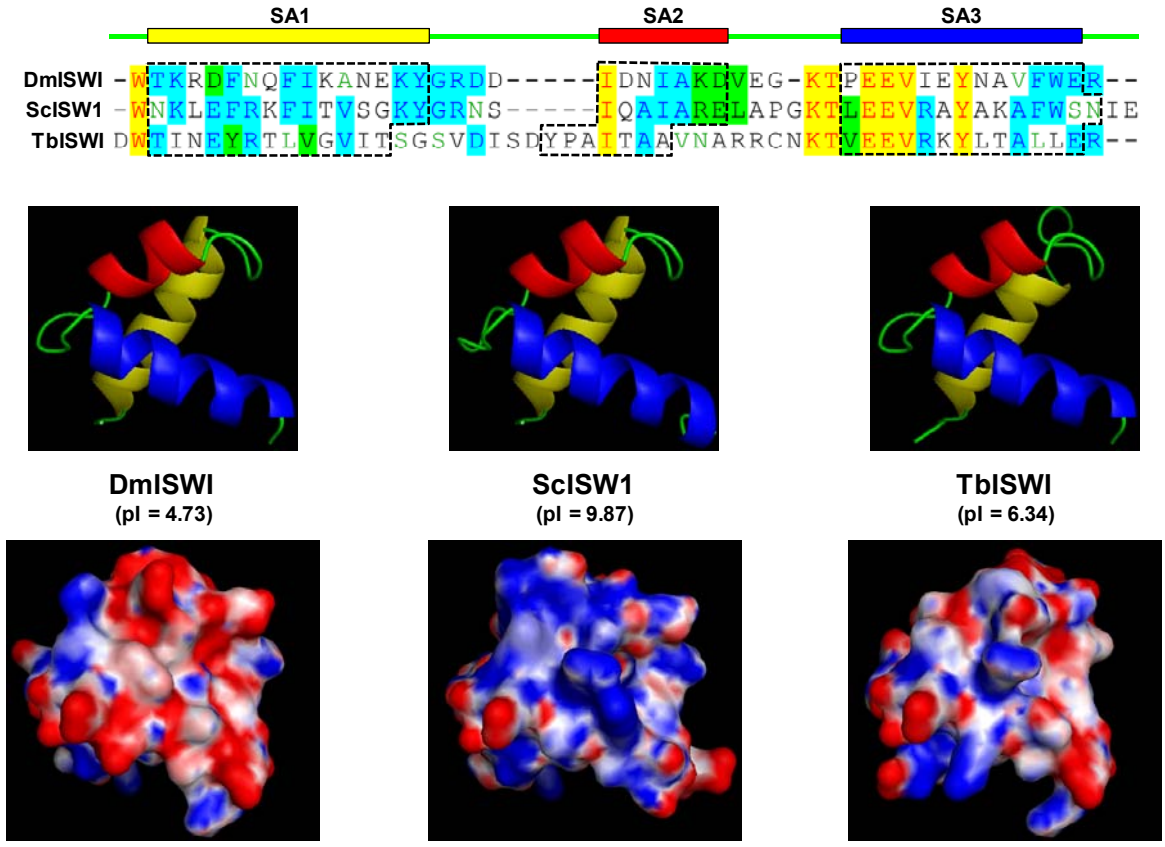


Figure 2.2.3 Comparing the charge distribution on the SANT domains of DmISWI, ScISW1 and TbISWI. (based on Pinskaya et al, 2009) The top panel of the figure shows an alignment of the predicted SANT domains of the three ISWI proteins- DmISWI, ScISW1 and TbISWI and blocks indicating the three α-helices SA1-SA3. The middle panel shows a cartoon depiction of the spatial arrangement of the three helices. The lower panel shows the charge distribution on the surface of the domain- red indicating negative charge (acidic) and blue indicating positive charge (basic). Between the middle and the lower panels, the calculated pI values of the domains have also been indicated.

A typical SLIDE domain consists of four α-helices (numbered SL1 to SL4 in Figure 2.2.4) separated by short stretches of unstructured coils. Although the TbISWI SLIDE domain is structurally conserved, its SL1 domain is shorter by one amino acid and SL2 by three. Also, the coil connecting SL1 and SL2 is shorter than the other two, while that connecting SL2 and SL3 is longer. The most striking difference between the SANT domains of the three proteins lies in their surface charge. While DmSLIDE is positively

charged in its surface charge, the ScSLIDE is quite negatively charged, and the TbSANT surface is made up of a mosaic of negatively and positively charged patches giving it an overall neutral charge.

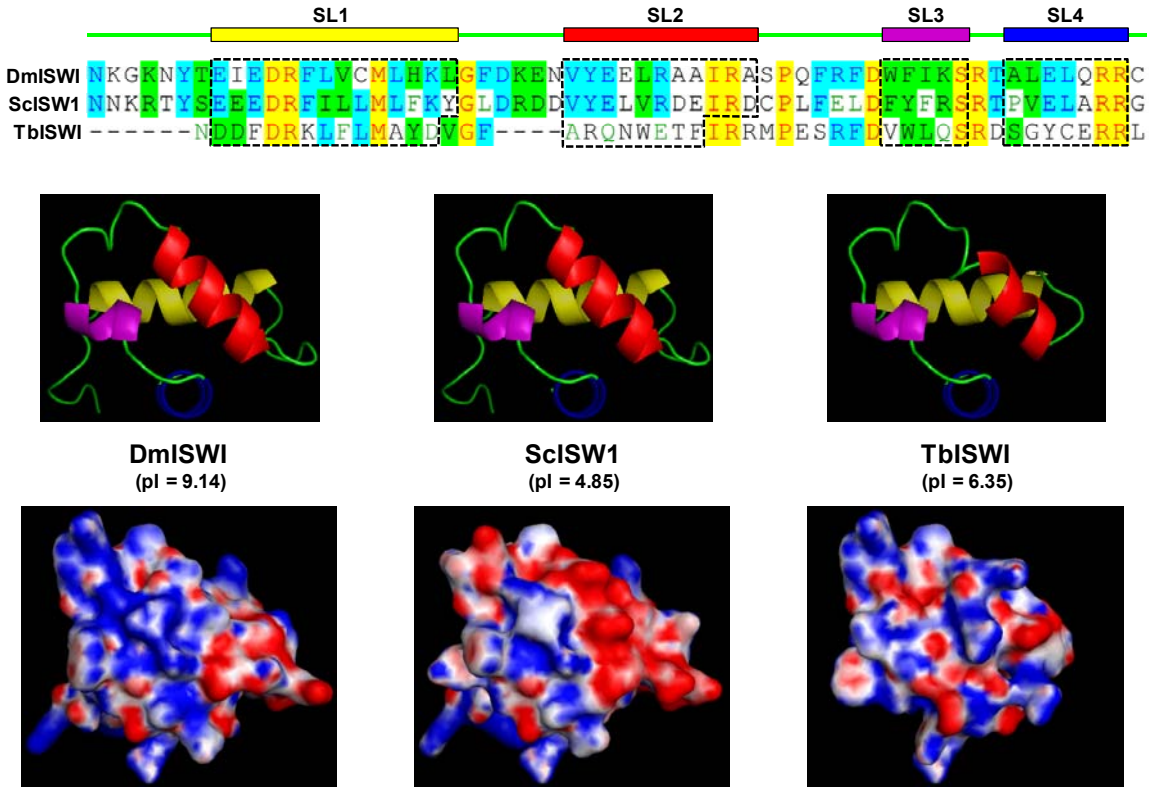


Figure 2.2.4 Comparing the charge distribution on the SLIDE domains of DmISWI, ScISW1 and TbISWI. (based on Pinskaya et al, 2009) The top panel of the figure shows an alignment of the predicted SLIDE domains of the three ISWI proteins- DmISWI, ScISW1 and TbISWI and blocks indicating the four α-helices SL1-SL4. The middle panel shows a cartoon depiction of the spatial arrangement of the three helices. The lower panel shows the charge distribution on the surface of the domain- red indicating negative charge (acidic) and blue indicating positive charge (basic). Between the middle and the lower panels, the calculated pI values of the domains have also been indicated.

Discussion

In order to compare the SWI2/SNF2 domains of the various kinetoplastid ISWIs, all of them were aligned using the AlignX program. As seen in Figure 2.2.1, large parts of the seven helices are well conserved, and amino acids that do vary seem to be conserved within the kinetoplastid protozoa. The critical amino acids marked with stars are all well conserved across all of the ISWIs. Such strong conservation of the ATPase domain is

important for the functionality of the protein. Also, the strong N-terminal conservation suggests that most of the protein divergence seen in the phylogenetic tree of ISWI is contributed by the C-terminal domain.

To explore this further, the TbISWI C-terminus was compared with that of two extensively studied ISWIs from model organisms *Drosophila* and yeast. Using Phyre (a structural modelling program) tertiary structural models were made for the C-termini (from HAND to SLIDE domain) of all three proteins- DmISWI, ScISW1 and TbISWI (Figure 2.2.2). With the exception of some parts of the HAND domain, DmISWI and ScISW1 show very strong structural conservation and can be almost entirely superimposed using PyMOL software. Despite the phylogenetic tree constructed using ISWI sequence (Figure 2.1.2A) which suggests significant divergence between TbISWI and the DmISWI/ ScISWI pair, its C-terminus appears to be fairly conserved in structure. However, TbISWI's structural conservation is not as strong as seen between DmISWI and ScISW1; it has comparatively longer unstructured coils and shorter α-helices throughout the length of the C-terminus. This moderate structural homology may explain why it had been originally difficult to identify the various C-terminal domains of TbISWI.

While the C-terminus is thought to be involved in the chromatin remodelling activity of ISWI, different interpretations exist about the role of the different C-terminal domains. In DmISWI, the SLIDE domain has been proposed to be responsible for the DNA binding function and the SANT domain for interaction with the nucleosomes (Grüne et al. 2003). In *Xenopus* ISWI too, the SANT domain was proposed to be involved in nucleosome binding (Horton et al. 2007). However, in yeast ISW2, the ATPase domain was proposed to directly interact with the nucleosomes, while the HAND

and SLIDE containing C-terminus was proposed to be interact with DNA (Dang et al. 2007). A recent study has shown that the properties of the DmISWI SANT and SLIDE domains are quite different from those of the ScISW1 SANT and SLIDE domains, with the differences being attributable to the surface charges of the two domains in the two proteins (Pinskaya et al. 2009). In DmISWI, the SANT domain is negatively charged and appears to interact with nucleosomes, while in ScISW1, the SANT domain is positively charged and interacts with DNA. The opposite is true for the SLIDE domains in the two proteins- it is positively charged in DmISWI and interacts with DNA, while in ScISW1, it is negatively charged and interacts with nucleosomes.

To get an insight into the potential role of the TbISWI SANT/ SLIDE domains in DNA/ nucleosome interaction, the sequence alignments of these domains and their surface charge distribution were studied. The sequences of the SANT and SLIDE domains of the three proteins were aligned, modelled by Phyre, and their surface charges plotted using the APBS tools in PyMOL software (DeLano 2009; Baker et al. 2001; Kelley et al. 2009).

Compared to the first two ISWI proteins, the first α-helix of the SANT domain (SA1) in TbISWI was shorter by 2 amino acids (aa) and the second helix (SA2) was displaced in the alignment by 3 aa (Figure 2.2.3). The connecting loops between the α-helices were also longer in TbISWI, and the overall sequence alignment was very poor. These observations explain why the TbISWI SANT domain could not be identified by sequence homology. The charge distribution on the SANT domains of DmISWI and ScISW1 confirmed the earlier observations (Pinskaya et al. 2009). While the surface of the DmISWI SANT domain was negatively charged, that of ScISW1 was positive, and this was reflected in their pI values too. However, the surface of the TbISWI SANT

domain consisted of a patchwork of negatively and positively changed regions, resulting in a near-neutral pI value, and not allowing any inference about the interaction partners of this domain.

As expected from a relatively easier identification of the SLIDE domain in TbISWI, stronger sequence conservation was seen in the SLIDE domains of the three proteins (Figure 2.2.4). However, the lengths of the unstructured coils were variable. The charge distribution on the SLIDE domains of DmISWI and ScISW1 confirmed the earlier observations (Pinskaya et al. 2009). While the surface of DmISWI SLIDE domain was positively charged, that of ScISW1 was negative, reflected in their pI values too. However, the surface of the TbISWI SLIDE domain, similar to that of the TbISWI SANT domain, consisted of a patchwork of negatively and positively changed regions, resulting in a near-neutral pI value, and not allowing any inference about the interactions of this domain.

The above observations from the surface charge distribution on the TbISWI domains seem to suggest that neither the SANT nor the SLIDE domain is likely to interact with DNA or nucleosomes for chromatin remodelling in *T. brucei*. However, ISWI chromatin remodelling complexes are known to have many different functions depending on the proteins the ISWI ATPases partner with (Corona et al. 2004). Also, conserved DNA-binding motifs such as chromodomains and bromodomains are often found on the non-ATPase subunits of ISWI complexes (Mohrmann et al. 2005). It is therefore possible that in the *T. brucei* ISWI chromatin remodelling complex(es), the primary role of the ATPase subunit is to supply energy for the remodelling reaction while the physical interaction with the chromatin is taken care of by one or more of the other subunits.

It has been argued that the proteolytic sensitivity of the ISWI spacer helix that separates the SANT and the SLIDE domains (Figure 2.2.2) may indicate its susceptibility to becoming unstructured, allowing greater movement for the domains on its either side (Horton et al. 2007). Such conformational changes could also expose surfaces otherwise considered hidden in a static view of the protein, and promote more dynamic interactions with chromatin substrates or with other protein partners. A recent study that analysed a range of chromatin remodelling proteins has argued that the extraordinary repertoire of interaction partners of chromatin remodelling proteins can be explained by the abundant regions of intrinsic disorder in them, which imparts flexibility and functional versatility to them (Sandhu 2009). In this regard, it is noteworthy that TbISWI has longer stretches of unstructured loops (Figure 2.2.2), making it structurally more flexible and more capable of interaction with partners.

Affinity maturation is a process whereby proteins evolve in a manner so as to optimise their binding potential (Yang et al. 2003). *T. brucei* chromatin organisation is quite different compared to higher eukaryotes. Its chromosomes do not condense during mitosis and its histones are highly divergent, particularly at their N-termini (Hecker et al. 1985; Alsford et al. 2004; Thatcher et al. 1994). Several studies have documented the interaction of ISWI complexes with the N-terminal tails of histones H3 and H4 (Clapier et al. 2002; Corona et al. 2002; Grüne et al. 2003; Dang et al. 2006). Following from the above, it is possible that the mottled surface charge distribution of the TbISWI SANT and SLIDE domains has evolved to optimise interaction with the highly divergent histone tails of *T. brucei*.

It is important to be cautious and not over-interpret the results from the predicted models of the SANT and the SLIDE domains. These models could also be non-

representative due to the limitations of the prediction algorithm. An important concern in this regard relates to the availability of only two solved crystal structures of ISWI C-terminal domains that can be used as templates for sequence threading- *Drosophila* and *Xenopus* ISWIs (Grüne et al. 2003; Horton et al. 2007). Furthermore, structures obtained by X-ray crystallography are only static instances of proteins and may not represent the protein in its dynamic form.

More insight into the role of TbISWI will come from experiments that assay the DNA binding and nucleosome remodelling activities of the protein, in full or with parts selectively deleted.

Chapter 3

Role of TbISWI in transcriptional silencing of VSG genes

3.1. Minichromosomes

Introduction

TbISWI depletion by RNAi has been shown to cause derepression of a fluorescent reporter gene inserted downstream of the ES promoter in a silent expression site. While ES derepression was found to be 30-60 fold in BF cells, it was only 10-17 fold in PF cells (Hughes et al. 2007). However, this study could not find evidence for the derepressed transcription proceeding all the way to the distal VSG gene of the silent ES. This led the authors to propose that although TbISWI plays an important role in ES downregulation, fully processive ES activation is a multilayered process also involving other players.

In order to further understand the role of TbISWI in transcriptional repression in *T. brucei*, it is important to identify transcriptionally silent regions in the trypanosome genome that can be similarly tested for derepression after TbISWI depletion. However, as most of the *T. brucei* genome is transcribed constitutively (Clayton 2002), there are only a few genomic regions that can be employed for such testing. An experimental strategy was designed whereby an eGFP reporter gene was inserted, with and without an upstream ES promoter, into the silent genomic locus being tested and its expression monitored after the RNAi-mediated depletion of TbISWI.

The reporter construct designed to target the transcriptionally silent locus contained the eGFP gene with a tubulin splice signal at its 5'-end and a tubulin polyadenylation signal at its 3'-end (Figure 3.1.1A). The eGFP gene in the construct was followed downstream by an rDNA promoter (indicated with a black flag in Figure 3.1.1A), which drove the transcription of a Blasticidin resistance gene flanked by a 5'-tubulin splice signal and a 3'-actin polyadenylation signal. The stepwise building of this construct, called peGFP_RPtubBlast, is described in the Materials and Methods chapter.

The construct forms the –ESp reporter backbone that was modified to target it to different silent genomic loci.

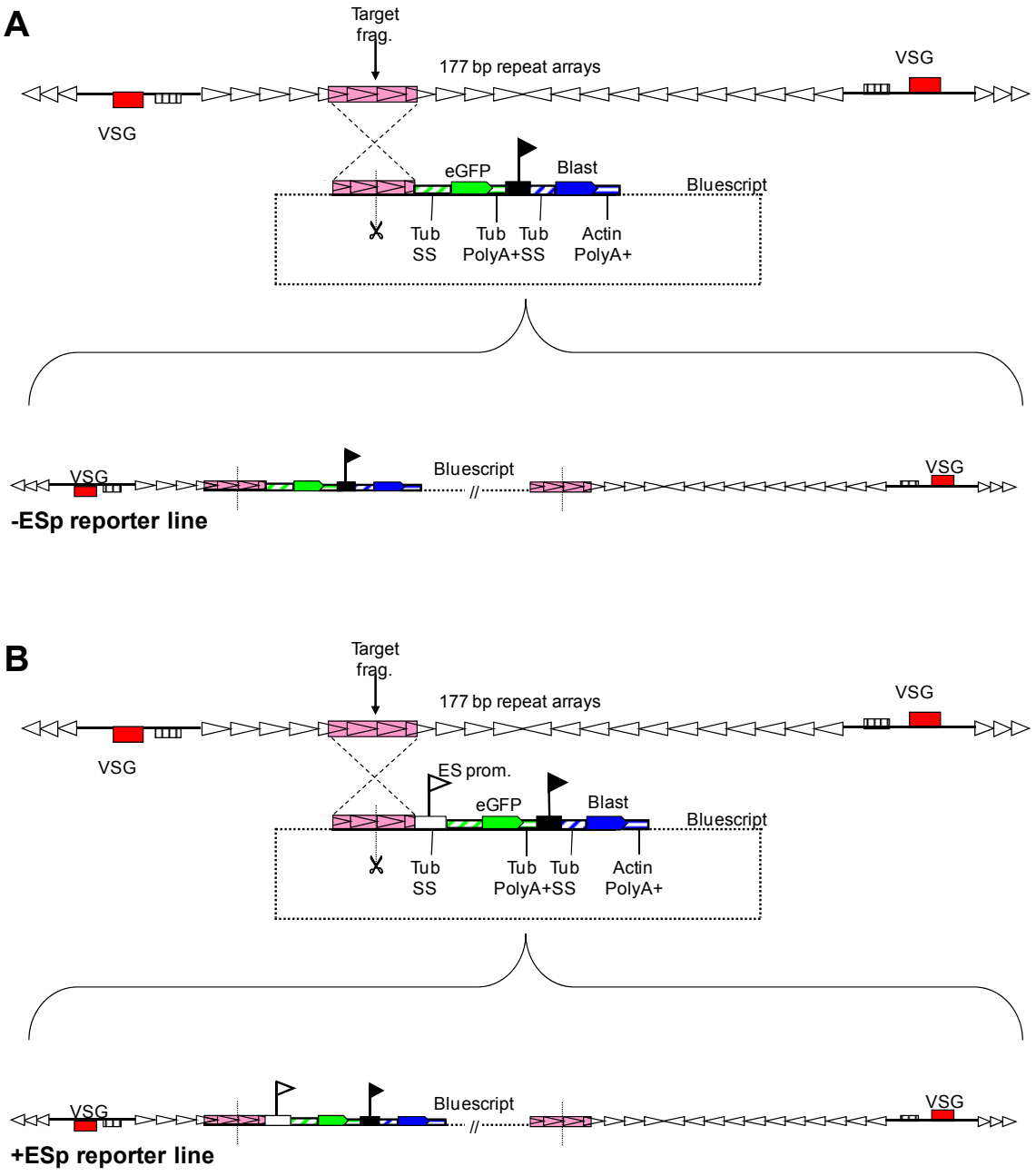


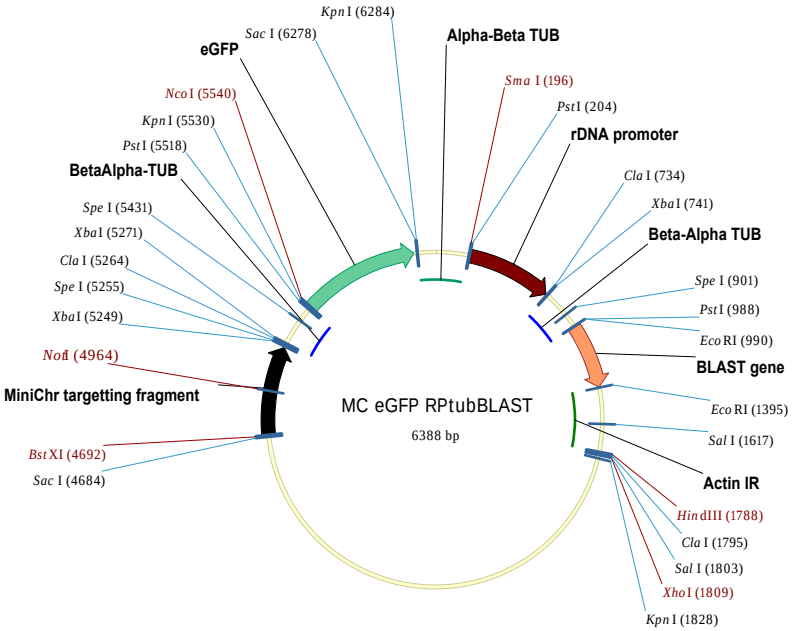
Figure 3.1.1 Plan of reporter *eGFP* integration into the minichromosomes. The schematics show integration of an *eGFP* reporter gene into the minichromosomal 177bp repeats. Panel (A) shows a –ESp reporter construct without a promoter sequence upstream of *eGFP*. Panel (B) shows a +ESp reporter construct with an Expression Site promoter upstream of *eGFP*. In both variants of the reporter constructs, the *eGFP* gene with flanking transcript processing signals (tubulin intergenic regions) is followed downstream by an rDNA promoter and a *Blasticidin* resistance gene cassette. The linearised (indicated by scissors) constructs integrate into the minichromosomes by a single-crossover facilitated homologous recombination.

Since previous reports have shown that an ES promoter is more sensitive to its genomic context than an rDNA promoter (Horn et al. 1997), and that TbISWI depletion results in its derepression (Hughes et al. 2007), another version of the reporter constructs was made with an ES promoter (indicated with a white flag in Figure 3.1.1B) upstream of the eGFP gene cassette to study the effect of TbISWI depletion on the ES promoter placed in a particular genomic context. The context sensitive ES promoter served as a "probe" for the chromatin state of the genomic locus into which the reporter construct was integrated. The stepwise building of this construct, called pESeGFP_RPtubBlast, is described in the Materials and Methods chapter. The construct forms the +ESp reporter backbone that can be modified to target it to a different genomic locus.

In this section of the thesis, an attempt has been made to understand the role of TbISWI in transcriptional downregulation of the minichromosomes. Minichromosomes are a transcriptionally silent class of trypanosome chromosomes that are thought to serve as reservoirs of silent telomeric VSG genes preferred for antigenic switching events early in a parasitemia (Robinson et al. 1999). There are about 100-150 minichromosomes in *T. brucei*, each ranging in size from 50 to 150 kb and making up about 10% of the nuclear genome (Berriman et al. 2005; Weiden et al. 1991). A typical minichromosome consists of 177 bp tandem repeats, which are flanked on either side by silent VSG genes at the telomere ends. Although most of the 177 bp repeats are confined to the minichromosomes, some have also been found in the sub-telomeres of large chromosomes (Wickstead et al. 2004).

Results

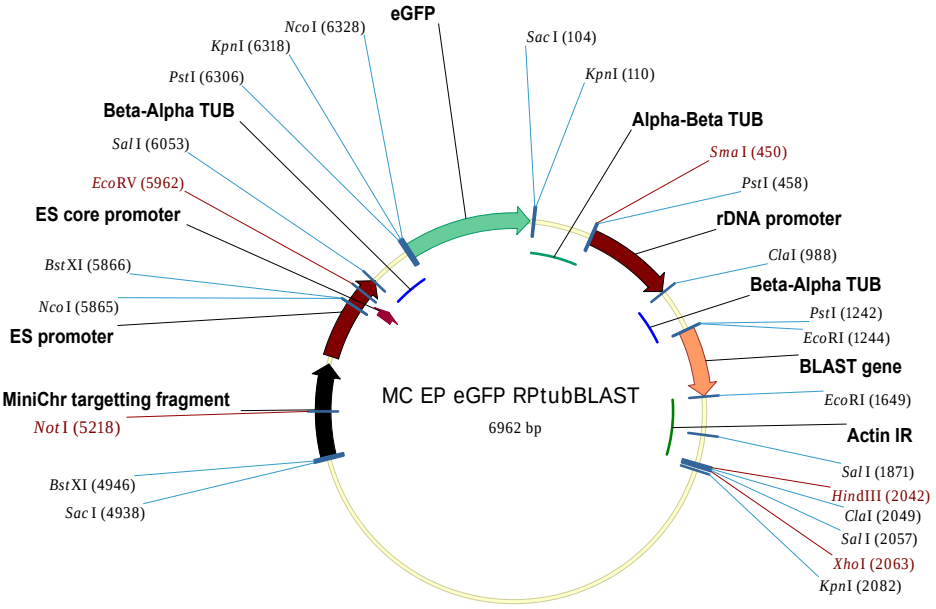
A



STEPS:

1. The Minichromosome Targeting Fragment was isolated from p2T7_177_hygro (Wickstead et al, 2002) by cutting with KpnI+SacI and then blunted.
2. peGFP_RPtubBLAST was linearised with NotI (+blunted) and the blunt-ended Minichromosome Targeting Fragment from Step 1 was ligated into it. This construct was called MC_eGFP_RPtubBLAST.
3. MC_eGFP_RPtubBLAST was linearised with NotI for targeting into the trypanosome genome.

B



STEPS:

1. The Minichromosome Targeting Fragment was isolated from p2T7_177_hygro (Wickstead et al, 2002) by cutting with KpnI+SacI and then blunted.
2. pESP_tubeGFPtub_RPtubBLAST was linearised with NotI (+blunted) and the blunt-ended Minichromosome Targeting Fragment from Step 1 was ligated into it. This construct was called MC_EPeGFP_RPtubBLAST.
3. MC_EPeGFP_RPtubBLAST was linearised with NotI for targeting into the trypanosome genome.

Figure 3.1.2 Maps of the minichromosome targeting reporter constructs. The two reporter constructs were obtained by inserting a minichromosome targeting fragment into the respective reporter construct backbones. While (A) MCeGFP_RPtubBlast was derived from the –ESp reporter backbone of peGFP_RPtubBlast, (B) MCEPeGFP_RPtubBlast was derived from the

+ESp reporter backbone of pEPeGFP_RPtubBlast. Both the constructs were linearised for transfection by NotI digestion.

The minichromosomal targeting –ESp eGFP reporter construct, MCeGFP_RPtubBlast, was made by inserting a minichromosome targeting fragment (MC^{177}), consisting of three 177 bp repeats, into the NotI linearised –ESp eGFP reporter backbone as detailed in Figure 3.1.2A. The construct was verified by diagnostic digests and sequencing. Similarly, the minichromosome targeting +ESp eGFP reporter construct, MCEPeGFP_RPtubBlast, was made by inserting the 177 bp repeat targeting fragment into the NotI linearised +ESp eGFP reporter backbone as detailed in Figure 3.1.2B, and verified by diagnostic digests and sequencing.

Targeting was achieved by linearising MCeGFP_RPtubBlast and MCEPeGFP_RPtubBlast constructs with NotI restriction in the middle of the MC targeting sequence, and generating a double strand break inside a homologous region to facilitate recombination (Clayton 1999). This strategy results in the integration of the reporter construct by a single crossover event that also integrates the Bluescript vector backbone into the minichromosomes (see Figure 3.1.1). Bloodstream form (BF) S16.221 and procyclic form (PF) 29-13 trypanosomes were transfected with the –ESp eGFP and the +ESp eGFP reporter constructs. Both of these parental cell lines encode constitutively expressed Tet repressor and T7 RNA polymerase, a setup that can be effectively used for tetracycline inducible RNAi (Wirtz et al. 1999).

To verify that the reporter constructs had integrated into the minichromosomes, the reporter cell lines were analysed by separating their chromosomes on an agarose gel using pulsed-field gel electrophoresis (PFGE) and probing for eGFP by Southern blot analysis. Analysis of BF cell lines found that 6 out of 8 integration events using the –ESp

construct had resulted in successful targeting to the minichromosomes, while only 3 of 4 integration events using the +ESp construct had targeted a minichromosome. All of the four PF cell lines analysed, two each for –ESp and +ESp constructs, had constructs integrated into the minichromosomes. Figure 3.1.3 shows PFGE chromosome separation and Southern blot analysis for eight representative cell lines.

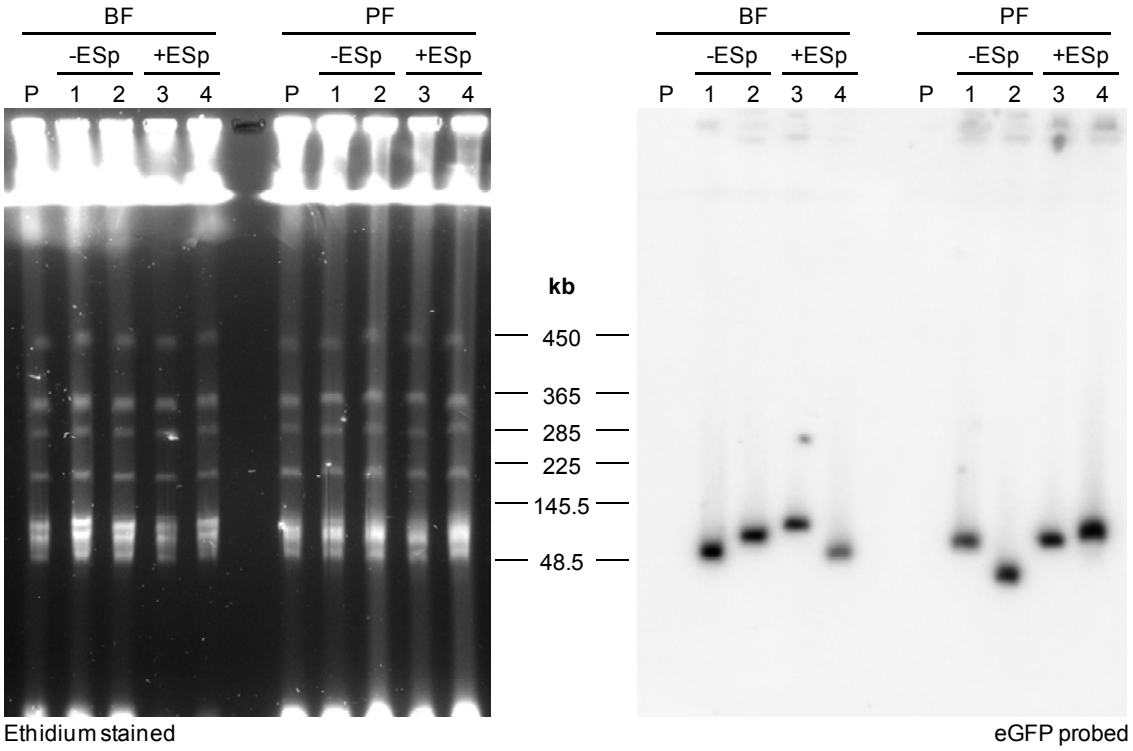


Figure 3.1.3 Integration of a reporter *eGFP* construct into the *T. brucei* minichromosomes. Bloodstream (BF) and procyclic (PF) form *T. brucei* minichromosome reporter transformants were analysed using Pulsed-field gel electrophoresis (PFGE), where chromosomes from parental cell lines (lane P) were separated on a gel alongside those from the –ESp (lanes 1&2) and the +ESp (lanes 3&4) transformants for comparison. An ethidium bromide (EtBr) stained gel is shown on the left, the *eGFP* probed Southern blot is shown on the right, and the size markers are indicated in kilobases in between. For the BF lanes, P:221Pur_A3, 1: BF_MCe, 2:BF_MCe^{b}, 3:BF_MCEP, 4:BF_MCEP^{b}. For the PF lanes, P:29-13, 1:PF_MCe, 2:PF_MCe^{b}, 3:PF_MCEP, 4:PF_MCEP^{b}.

After verification of successful reporter integration into the minichromosomes, four reporter cell lines- BF_MCe, BF_MCEP, PF_MCe and PF_MCEP- were transfected with the MC^{177} TbISWI-A RNAi construct (Hughes et al. 2007) and two RNAi cell lines per reporter cell line were obtained. The RNAi cell lines obtained were called BF_MCe1,

BF_MCe2, BF_MCEP1, BF_MCEP2, PF_MCe1, PF_MCe2, PF_MCEP1 and PF_MCEP2. These were further used for the planned experiments.

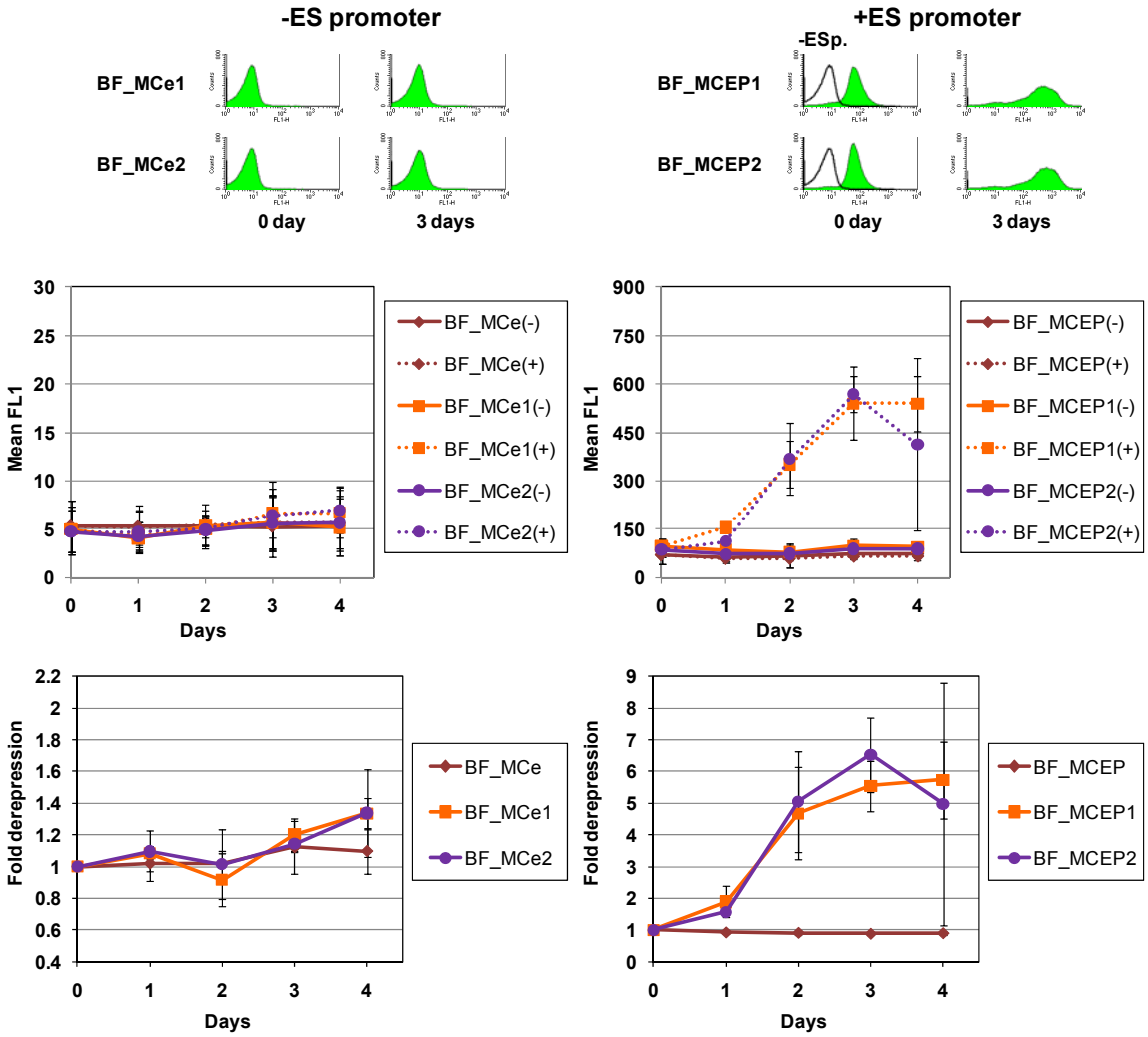


Figure 3.1.4 Monitoring eGFP derepression after induction of RNAi against *TbISWI* in bloodstream form minichromosome reporter cell lines. Levels of expression of the *eGFP* gene were measured by flow cytometry for both the –ESp (left panel) and the +ESp (right panel) reporter lines. Top: representative flow cytometry traces for two cell lines at two different time-points are shown (day 0 and day 3), with eGFP fluorescence monitored in the FL1 channel (x-axis). The +ESp traces from day 0 also show an overlaid –ESp trace from the same time-point for comparison. Middle: the graph shows the mean fluorescence (FL1 channel) of eGFP in tetracycline induced (+) and uninduced (-) cell lines for different durations of induction; 100,000 cells were analysed per sample. Bottom: the graph shows the fold derepression of the *eGFP* gene for different durations of tetracycline induced *TbISWI* RNAi. Fold derepression was calculated as the ratio of mean fluorescence of induced cells to that of the uninduced cells. The graph shows the average of three independent experiments with the error bars indicating the standard deviation. BF_MCe1 and BF_MCe2 were derived from BF_MCe by transfection with the TbISWI_RNAi construct. BF_MCEP1 and BF_MCEP2 were derived from BF_MCEP by transfection with the TbISWI_RNAi construct.

In order to study the effect of TbISWI depletion on reporter eGFP expression, the RNAi cell lines were induced using 750 ng ml^{-1} tetracycline and their fluorescence measured in the FL1 channel using flow cytometry as described in the Materials and Methods. Figures 3.1.4 and 3.1.6 show representative FACS traces of eGFP fluorescence of different cell lines in the top panel, the mean FL1 fluorescence for different durations of induction in the middle panel, and the fold derepression of the eGFP gene, expressed as the ratio of fluorescence of induced versus uninduced cells, through time in the bottom panel. The depletion of the TbISWI protein can be seen in Figures 3.1.5 and 3.1.7.

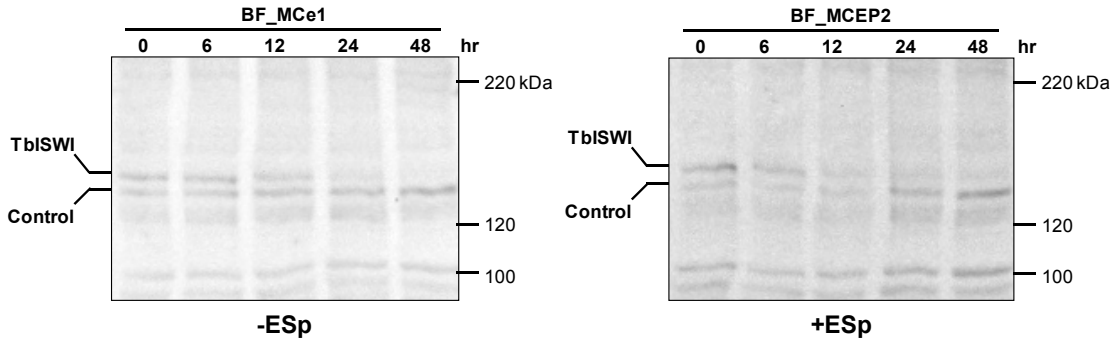


Figure 3.1.5 Depletion of TbISWI in bloodstream form minichromosome reporter cell lines after induction of RNAi. Protein lysates from –ESp (left) and +ESp (right) cell lines were prepared after the induction of RNAi against *TbISWI* with tetracycline for 0, 6, 12, 24 and 48 hours. TbISWI and a cross-reactive control band are indicated on the left and the size markers in kiloDaltons (kDa) on the right of the Western blot.

In the BF RNAi cell lines, TbISWI depletion did not result in significant derepression of the –ESp eGFP (Figure 3.1.4). The degree of fluorescence observed in the FL1 channel remained low even after day 3 of *TbISWI* RNAi, despite TbISWI protein becoming undetectable on a Western blot after 48 hours of tetracycline induction (Figure 3.1.5).

However, in the +ESp RNAi BF cell lines, the fluorescence observed in the FL1 channel was found to be increased compared to that of the –ESp RNAi lines even in their uninduced states (Figure 3.1.4). After TbISWI depletion, the +ESp eGFP cell line was

derepressed by 6 fold, with fluorescence increasing even further. Western blot analysis confirmed that TbISWI protein becomes undetectable after 48 hours of tetracycline induction (Figure 3.1.5).

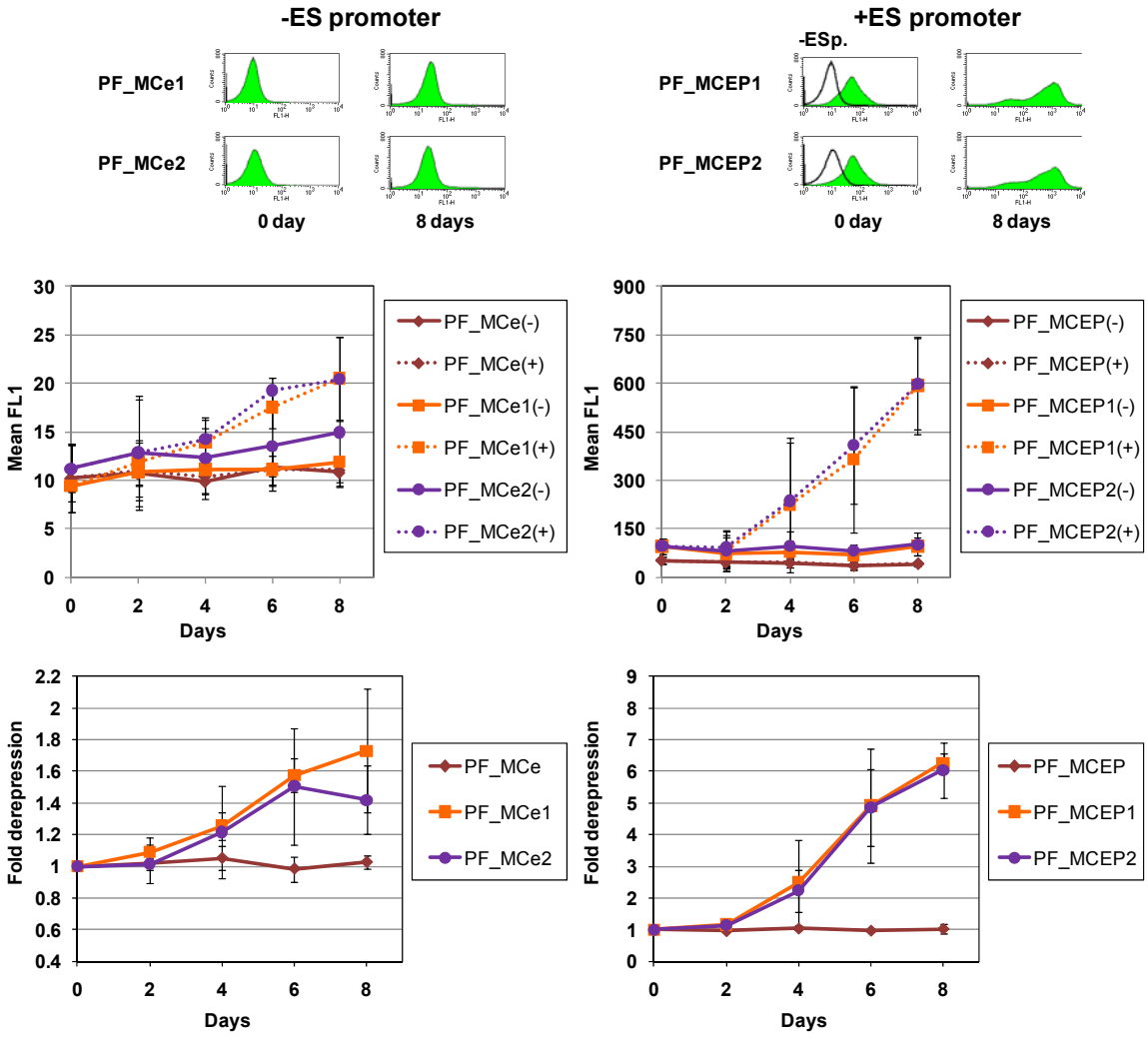


Figure 3.1.6 Monitoring eGFP derepression after induction of RNAi against *TbISWI* in procyclic form minichromosome reporter cell lines. Levels of expression of the *eGFP* gene were measured by flow cytometry for both the –ESp (left panel) and the +ESp (right panel) reporter lines. Top: representative flow cytometry traces for two cell lines at two different time-points are shown (day 0 and day 8), with eGFP fluorescence monitored in the FL1 channel (x-axis). The +ESp traces from day 0 also show an overlaid –ESp trace from the same time-point for comparison. Middle: the graph shows the mean fluorescence (FL1 channel) of eGFP in tetracycline induced (+) and uninduced (-) cell lines for different durations of induction; 100,000 cells were analysed per sample. Bottom: the graph shows the fold derepression of the *eGFP* gene for different durations of tetracycline induced *TbISWI* RNAi. Fold derepression was calculated as the ratio of mean fluorescence of induced cells to that of the uninduced cells. The graph shows the average of three independent experiments with the error bars indicating the standard deviation. PF_MCe1 and PF_MCe2 were derived from PF_MCe by transfection with the TbISWI_RNAi construct. PF_MCEP1 and PF_MCEP2 were derived from PF_MCEP by transfection with the TbISWI_RNAi construct.

In the PF RNAi cell lines, TbISWI depletion resulted in 1.4-1.7 fold derepression of the –ESp eGFP (Figure 3.1.6). The fluorescence observed in the FL1 channel increased slightly by day 8 of *TbISWI* RNAi. In the +ESp RNAi PF cell lines, however, more eGFP fluorescence was seen even in the uninduced state (Figure 3.1.6). As tetracycline induced TbISWI depletion proceeded, the +ESp eGFP was derepressed by 6 fold, with fluorescence increasing further. After 6 days of tetracycline induction, TbISWI protein became almost undetectable on Western blots for both RNAi cell lines (Figure 3.1.7).

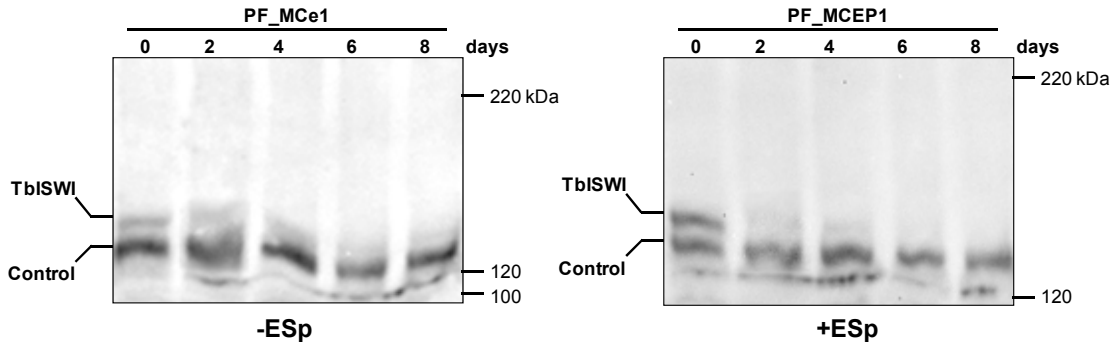


Figure 3.1.7 Depletion of TbISWI in Procyclic form minichromosome reporter cell lines after induction of RNAi. Protein lysates from –ESp (left) and +ESp (right) cell lines were prepared after induction of RNAi against *TbISWI* with tetracycline for 0, 2, 4, 6 and 8 days. TbISWI and a cross-reactive control band are indicated on the left and the size markers in kiloDaltons (kDa) on the right of the Western blot.

Discussion

As the 177 bp repeats used for the targeting of the reporter constructs are also found in genomic regions other than the minichromosomes (Wickstead et al. 2004), it was important to ascertain if the constructs had correctly integrated into the minichromosomes. Pulsed-field gel electrophoresis was used to separate *T. brucei* chromosomes using conditions suited for the resolution of intermediate chromosomes so that the three chromosomal size classes- mini, intermediate and large/ megabase chromosomes- could be unambiguously identified. In all of the 16 integration events (12

BF and 4 PF) analysed, the reporter construct integrated into the intermediate chromosomes only once. It integrated into the large chromosomes in only two cases and into the minichromosomes on the remaining 13 instances- a minichromosome targeting rate of 81.25%, which is very close to the 80% targeting rate reported in a previous study (Wickstead et al. 2002). In addition, in 1 of the 15 cell lines analysed two integration events in two different chromosomes (one intermediate and one mini-) were also seen. The cell line is expected to have been clonal as it was one of the two positive wells picked out of a 24 well plate after transfection. However, the case for polyclonality was not investigated any further. A previous study showed that ~15% of all constructs targeted to the 177 bp repeats integrated into the intermediate chromosomes and only ~3% to the large chromosomes (Wickstead et al. 2002). However, a larger sample of 38 integration events was used in that study and targeting efficiencies may also be influenced by the particular architecture of the construct in question.

The results from the –ESp reporter experiments in BF *T. brucei* suggest that depletion of TbISWI does not result in fortuitous initiation of transcription from cryptic promoters upstream of the eGFP gene inserted in a silent minichromosome. When an ES promoter is inserted before the *eGFP* gene in the same location, TbISWI depletion results in 6 fold increased transcription from the ES promoter in BF cells. These results suggest that TbISWI plays a role in transcriptional repression of an ES promoter inserted in the minichromosomes of BF trypanosomes.

Depletion of TbISWI results in 1.4-1.7 fold derepression of the –ESp eGFP inserted into a silent minichromosome in PF *T. brucei*. Although these derepression values are small, they are statistically significant ($P \leq 0.05$) as the gap between the SD (standard deviation) error bar of each RNAi cell line and that of the parental cell line is

greater than 16% of the sum of the lengths of the two error bars. While non-overlap of SD error bars is not a direct proof of statistical significance in all cases, the knowledge of the number of experiments (n = 3) makes it possible to infer how much SD gap between them would correspond to 2x SE (standard error) gap, a direct measure of statistical significance (Cumming et al. 2007). These results suggest that depletion of TbISWI causes low but significant transcription of the –ESp eGFP gene inserted into a silent minichromosome. Unlike in the BF *T. brucei*, there does appear to be some transcriptional activity from cryptic promoters in the minichromosome repeats in PF. The differences between the two lifecycle stages are possibly a consequence of a more stringent control over transcription factor (TF) binding in the BF cells as compared to the PF cells.

Results from the +ESp reporter experiments in PF cell lines show that similar to the results from BF trypanosomes, TbISWI depletion causes 6 fold derepression of *eGFP* expression. This supports the view that TbISWI is involved in transcriptional repression of minichromosomes in PF trypanosomes as well.

From the above results, it can be inferred that TbISWI plays a role in the silencing of minichromosomes in both BF and PF *T. brucei*. However, there appears to be stricter control of cryptic/ fortuitous transcription in BF cells, as seen from the –ESp *eGFP* results, presumably a mechanism to prevent transcription of the *VSG* genes located at the silent minichromosomes in BF life cycle stage where exposure of VSG variants to the host bloodstream must be tightly controlled to avoid premature exhaustion of the repertoire. A *GFP* gene inserted downstream of the ES promoter in an active expression site showed a mean FL1 fluorescence in the range of 4500-5500 units (data not shown), which is about 8-10 fold higher than that from the derepressed +ESp eGFP located in a

minichromosome. As suggested for ES regulation (Hughes et al. 2007), there may be multiple layers of transcriptional control operating on the minichromosomes too, and some of them may act to reduce initiation of transcription from cryptic promoters specifically in the BF life cycle stage.

The observation that the mean FL1 fluorescence values for the uninduced –ESp reporter cell lines lie in the range of 4-5 units for BF cells and in the range of 10-12 units for PF cells further confirms the view that fortuitous initiation of transcription is more strictly downregulated in BF cells. In contrast, the mean FL1 fluorescence values for the uninduced +ESp reporter cell lines are higher in the BF cells (60-75 units) in comparison to the PF cells (35-55 units). When comparing the +ESp eGFP reporter cell lines after TbISWI knockdown, the BF cells exhibit derepression comparable to their PF counterparts. This appears to be different from the observations made in the –ESp eGFP reporter cell lines where transcription was more tightly controlled in the BF cells. A possible explanation could be that parallel transcriptional regulatory mechanisms keep fortuitous transcription in check in both life cycle stages, but are more active in the BF. However, when a certain basal level of transcription has been achieved due the presence of an ES promoter, these regulatory mechanisms become less important.

An important point is that the RNA processing signals used for the eGFP reporter gene are the same in both the –ESp and the +ESp constructs. Therefore, the relative stability of the transcript in these different cell lines should to be comparable, validating the use of protein fluorescence as an indirect measure of transcript abundance. However, some caution needs to be exercised in interpreting fluorescence values across different life cycle stages because relative metabolic differences may result in different levels of protein even for comparable levels of the transcript. Also, biological variability may

result in different levels of fluorescence from the same cells across different experiments. Therefore, the degree of derepression caused by TbISWI depletion is a better measure of its role in transcriptional regulation than absolute fluorescence values.

As *TbISWI* RNAi leads to a lethal phenotype, prolonged periods of tetracycline induction may select for cells that are defective in RNAi or tetracycline inducibility. The average observed derepression of the whole population may therefore be lower than the derepression observed in the responsive population. These unresponsive populations are visible as small humps/ trails on the left in the fluorescence traces of induced BF (day 3) and PF (day 8) cells of the +ESp reporter lines (Figures 3.1.4 and 3.1.6).

3.2. VSG basic copy arrays

Introduction

VSG basic copy arrays (BCA) are located at the sub-telomeric region of the large chromosomes and serve as reservoirs of silent VSG genes to be used for antigenic switching. Studies have shown that exogenous Pol-I promoters (rDNA and ES) inserted into the VSG 118 BCA locus are repressed in bloodstream form trypanosomes. In procyclic form trypanosomes, while the ES promoter is repressed, the rDNA promoter retains about 10% of its maximal activity (Cross et al. 1998; Horn et al. 1997). These results indicate that the VSG 118 BCA is largely a transcriptionally repressed genomic region.

To understand the role of TbISWI in transcriptional repression at this silent genomic locus, the experimental strategy used in Section 3.1 above for the minichromosomes was adapted for this purpose (Figure 3.2.1). An eGFP reporter gene was inserted, with and without an upstream ES promoter, into the silent VSG 118 BCA and its expression monitored after RNAi-mediated TbISWI depletion. The –ESp and +ESp reporter backbone constructs described previously, peGFP_RPtubBlast and pESeGFP_RPtubBlast, were modified to target them to the VSG 118 BCA (Figure 3.2.2).

Results

In order to target the reporter constructs to the VSG 118 BCA, a ~1 kb targeting fragment was amplified by PCR (see Materials and Methods) and added to both the –ESp and the +ESp reporter backbones to obtain VBeGFP_RPtubBlast and VBEPeGFP_RPtubBlast, respectively (see Figure 3.2.2). Both constructs were verified by diagnostic digests and sequencing. Targeting into the genome was achieved by linearising

VBeGFP_RPtubBlast and VBEPeGFP_RPtubBlast constructs with NarI restriction in the middle of the VSG 118 BCA targeting sequence. This generated a double strand break inside the homologous region to facilitate recombination (Clayton 1999). As before, this strategy results in the integration of the reporter construct by a single crossover event that also integrates the Bluescript vector backbone into the target site, VSG 118 BCA in this case.

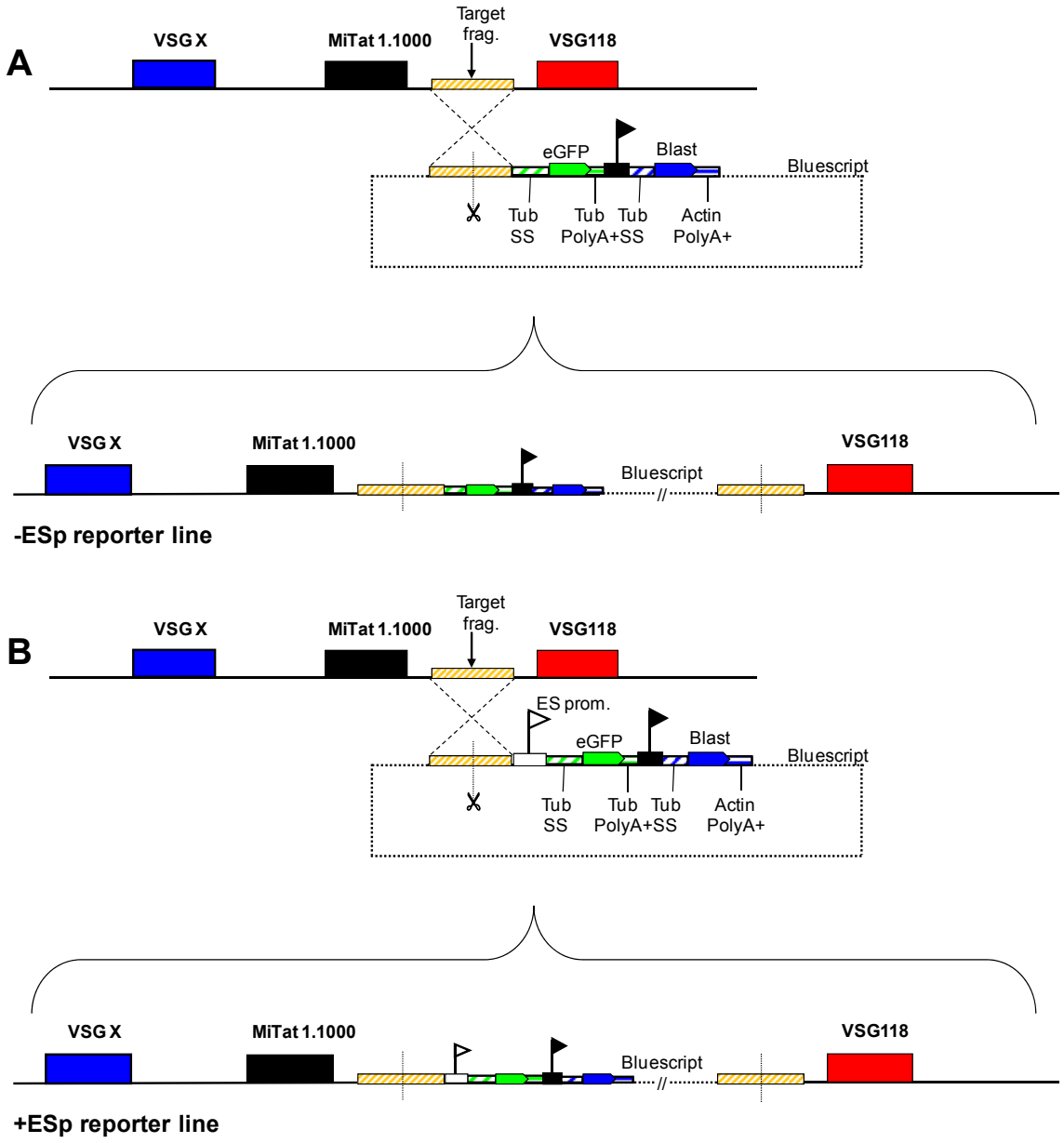
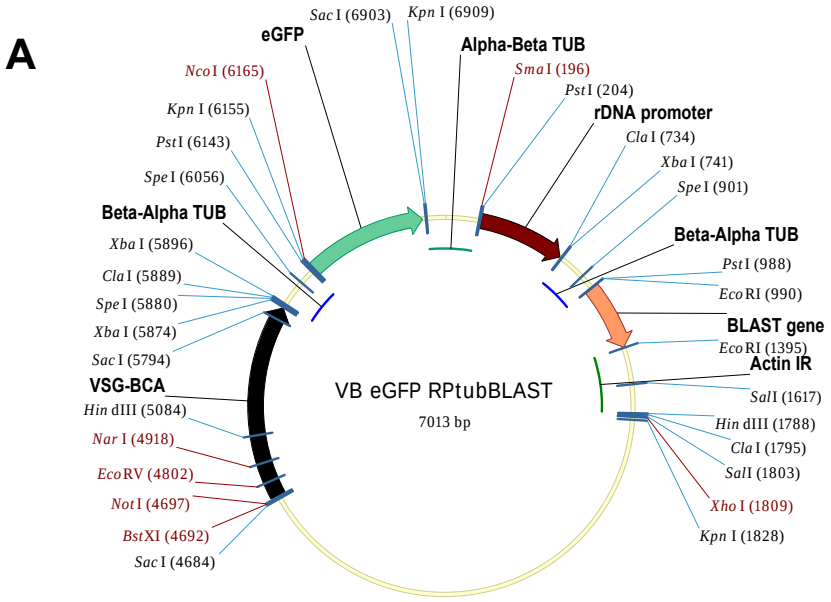


Figure 3.2.1 Plan of reporter *eGFP* integration into the VSG basic copy array. The schematic shows integration of an *eGFP* reporter gene into the VSG 118 basic copy array. Panel (A) shows a –ESp reporter construct without any promoter sequence upstream of *eGFP*. Panel

(B) shows a +ESp reporter construct with an Expression Site promoter upstream of *eGFP*. In both variants of the reporter constructs, the *eGFP* gene with flanking transcript processing signals (tubulin intergenic regions) is followed downstream by an rDNA promoter and a *Blasticidin* resistance gene cassette. The linearised (indicated by scissors) constructs integrate into the VSG basic copy array by a single-crossover facilitated homologous recombination.

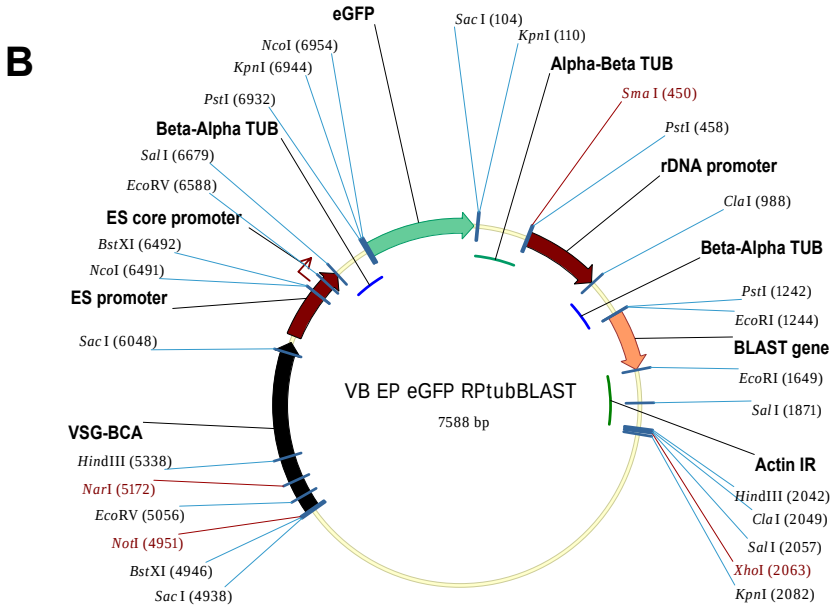
BF S16.221, a derivative of the S16 single marker cell line, and the PF 29-13 cell lines were used for transfection of –ESp and +ESp reporter constructs (Wirtz et al. 1999). Both these cell lines encode constitutively expressed Tet repressor and T7 RNA polymerase, allowing tetracycline inducible RNAi.

To verify correct targeting of the reporter constructs, genomic DNA was isolated from the reporter cell lines and used as a template for PCR linking reactions. The reporter construct could be linked to the genomic locus in all the cell lines tested. Figure 3.2.3 shows a schematic of the primers used and the PCR products for four representative cell lines.



STEPS:

1. The VSG Basic Copy Array Targeting Fragment was PCR amplified (Pwo polymerase) from HN1VO2 gDNA using primers BCVSG118-863s and BCVSG118-2025as.
2. peGFP_RPtubBLAST was linearised with NotI (+blunted) and the blunt-ended VSG Basic Copy Array Targeting Fragment from Step 1 was ligated into it. Correct ligation of the construct in the forward orientation was verified by PCRs linking over the junctions and restriction digestions of the obtained construct. This construct was called VB_eGFP_RPtubBLAST.
3. VB_eGFP_RPtubBLAST was linearised with NarI for targeting into the trypanosome genome.



STEPS:

1. The VSG Basic Copy Array Targeting Fragment was PCR amplified (Pwo polymerase) from HN1VO2 gDNA using primers BCVSG118-863s and BCVSG118-2025as.
2. pESP_tubeGFPtub_RPtubBLAST was linearised with NotI (+blunted) and the blunt-ended VSG Basic Copy Array Targeting Fragment from Step 1 was ligated into it. Correct ligation of the construct in the forward orientation was verified by PCRs linking over the junctions and restriction digestions of the obtained construct. This construct was called VB_EPeGFP_RPtubBLAST.
3. VB_EPeGFP_RPtubBLAST was linearised with NarI for targeting into the trypanosome genome.

Figure 3.2.2 Maps of the VSG basic copy array targeting reporter constructs. The two reporter constructs were obtained by inserting a VSG 118 basic copy array targeting fragment into the respective reporter construct backbones. While (A) VBeGFP_RPtubBlast was derived from the –ESp reporter backbone of peGFP_RPtubBlast, (B) VBEPeGFP_RPtubBlast was derived from the +ESp reporter backbone of pEPeGFP_RPtubBlast. Both the constructs were linearised for transfection by NarI digestion.

After verification of successful reporter construct integration into the VSG 118 BCA, four reporter cells lines- BF_VBe, BF_VBEP, PF_VBe and PF_VBEP- were transfected with the MC[177] TbISWI-A RNAi construct (Hughes et al. 2007). Two RNAi cell lines were obtained from each of these reporter lines and used for the planned experiments. The RNAi cell lines obtained were called BF_VBe1, BF_VBe2, BF_VBEP1, BF_VBEP2, PF_VBe1, PF_VBe2, PF_VBEP1 and PF_VBEP2.

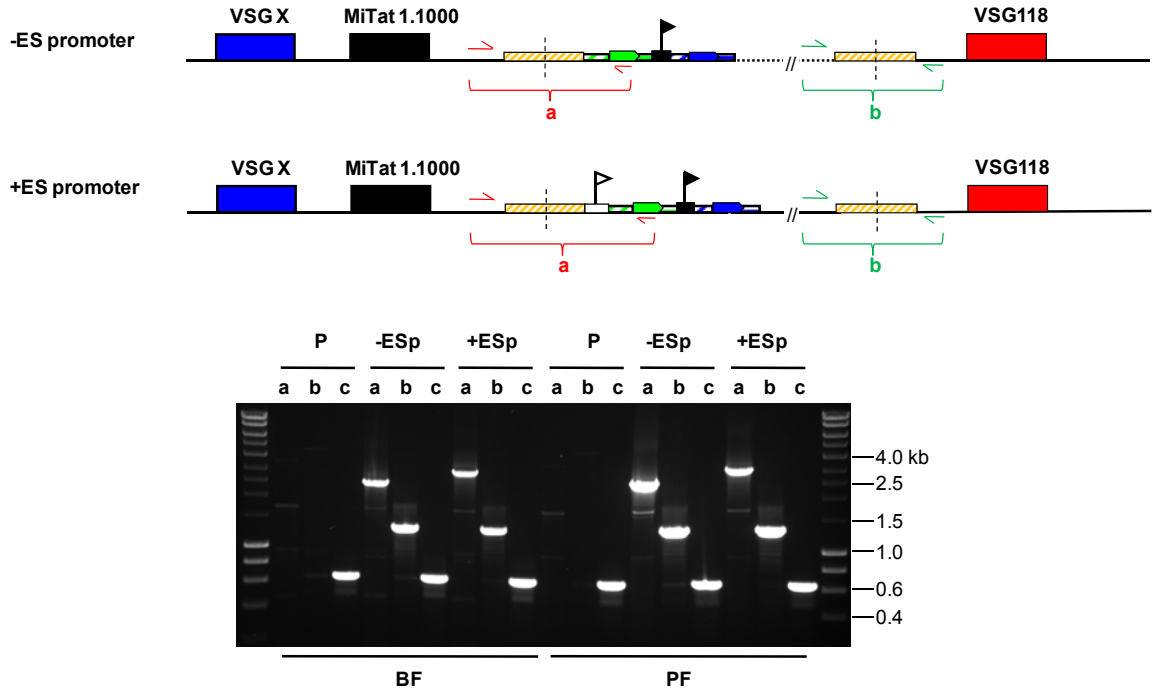


Figure 3.2.3 Integration of the reporter *eGFP* construct into the VSG basic copy array. PCR amplification using genomic DNA isolated from bloodstream (BF) and procyclic (PF) form *T. brucei* VSG basic copy array reporter transformants was used to confirm integration into the correct genomic locus. The top panel shows a schematic of primer sets used for the purpose. Primer sets 'a' and 'b' were designed to amplify over the upstream and the downstream reporter construct integration points (indicated by dotted vertical lines), respectively. Primer set 'c' was used as a positive control for amplification of genomic DNA. The bottom panel shows the products of the PCR reactions using genomic DNA from various bloodstream and procyclic form cell lines, with the size markers in kilobases (kb) indicated on the right of the gel picture. For the BF lanes, P:221Pur_A3, –ESp:BF_VBe, +ESp:BF_VBEP. For the PF lanes, P:29-13, –ESp:PF_VBe, +ESp:PF_VBEP. Primer sets used: 'a' (BCVSG118-42s & eGFP-183as), 'b' (T3 & BCVSG118-2100as) and 'c' (Pol I-484s & Pol I-1123as).

To study the effect of RNAi-mediated depletion of TbISWI on reporter eGFP expression, the RNAi cell lines were grown in the presence of 750 ng ml^{-1} tetracycline and the mean fluorescence of cells was measured in the FL1 channel using flow cytometry as previously described. Figures 3.2.4 and 3.2.6 show representative FACS traces of eGFP fluorescence of different cell lines in the top panel, the mean fluorescence of both induced and uninduced cells at different time points in the experiment in the middle panel, and the fold derepression of the eGFP gene, expressed as the ratio of fluorescence of induced versus uninduced cells, at different durations of induction in the bottom panel. The depletion of the TbISWI protein can be seen in Figures 3.2.5 and 3.2.7.

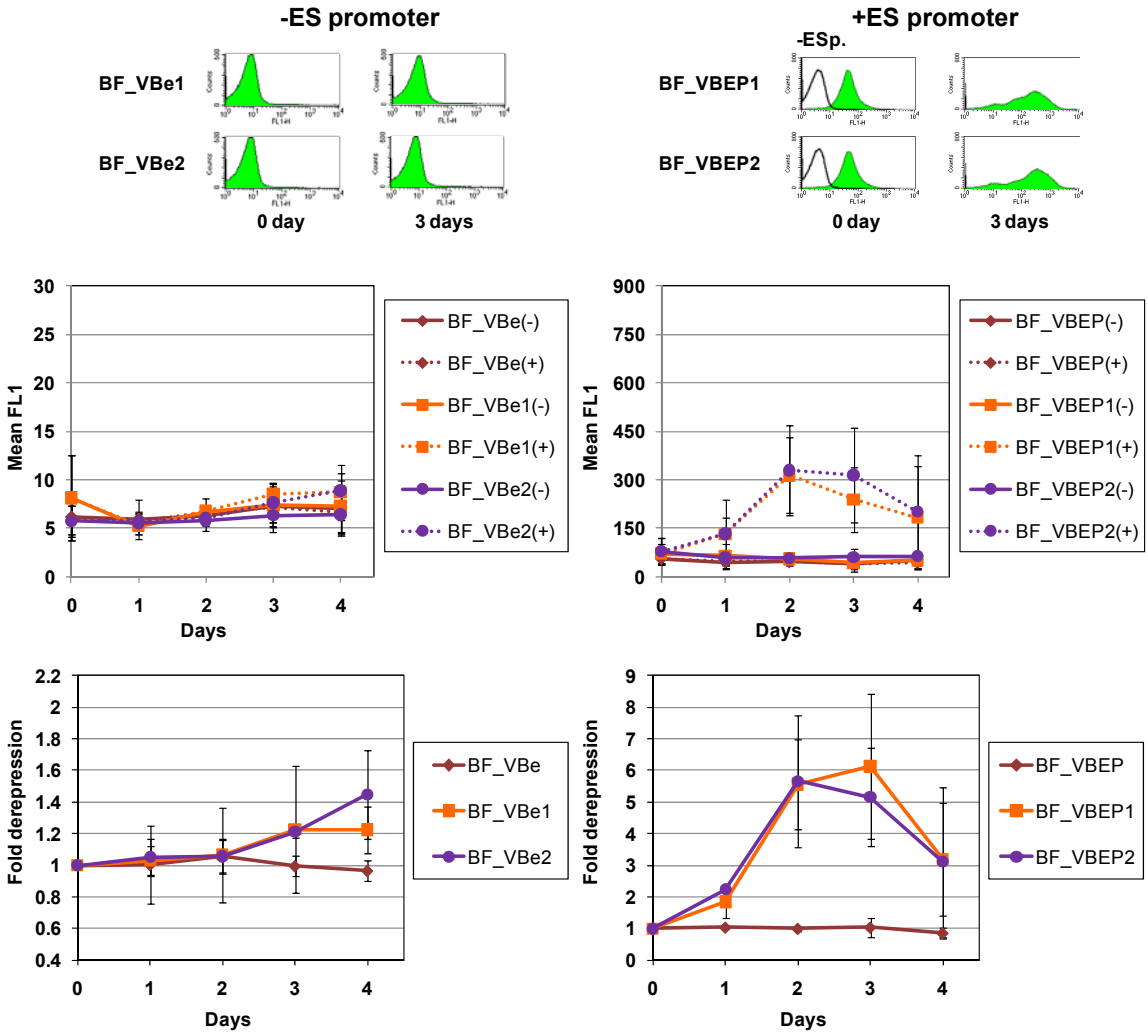


Figure 3.2.4 Monitoring eGFP derepression after induction of RNAi against *TbISWI* in bloodstream form VSG basic copy array reporter cell lines. Levels of expression of the *eGFP* gene were measured by flow cytometry for both the –ESp (left panel) and the +ESp (right panel) reporter lines. Top: representative flow cytometry traces for two cell lines at two different time-points are shown (day 0 and day 3), with eGFP fluorescence monitored in the FL1 channel (x-axis). The +ESp traces from day 0 also show an overlaid –ESp trace from the same time-point for comparison. Middle: the graph shows the mean fluorescence (FL1 channel) of eGFP in tetracycline induced (+) and uninduced (-) cell lines for different durations of induction; 100,000 cells were analysed per sample. Bottom: the graph shows the fold derepression of the *eGFP* gene for different durations of tetracycline induced *TbISWI* RNAi. Fold derepression was calculated as the ratio of mean fluorescence of induced cells to that of the uninduced cells. The graph shows the average of three independent experiments with the error bars indicating the standard deviation. The BF_VBe1 and BF_VBe2 cell lines were derived from BF_VBe after transfection with the TbISWI_RNAi construct. The BF_VBEP1 and BF_VBEP2 cell lines were derived from BF_VBEP after transfection with the TbISWI_RNAi construct.

In the BF RNAi cell lines, TbISWI depletion resulted in 1.2-1.4 fold derepression derepression of the –ESp eGFP (Figure 3.2.4). Although small, this change represents the average of three independent experiments and the error bars on the plot show that the

results are statistically significant. The fluorescence observed in the FL1 channel increased slightly after day 3 of *TbISWI* RNAi. In the BF +ESp RNAi cell lines, *TbISWI* RNAi resulted in 5-6 fold derepression of the +ESp eGFP after 3 days of induction. Almost complete knockdown of TbISWI was observed after 48 hours of tetracycline induction using Western blot analysis in both cell lines (Figure 3.2.5).

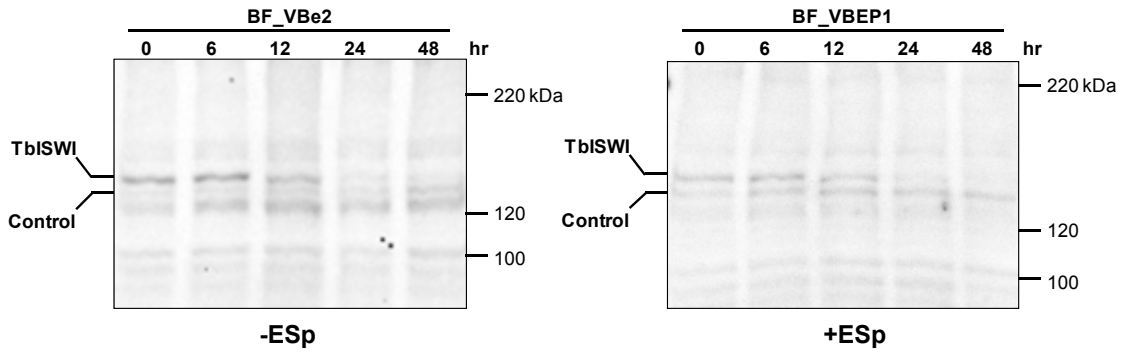


Figure 3.2.5 Depletion of TbISWI in bloodstream form *T. brucei* VSG basic copy array reporter cell lines after induction of RNAi. Protein lysates from –ESp (left) and +ESp (right) cell lines were prepared after the induction of RNAi against *TbISWI* with tetracycline for 0, 6, 12, 24 and 48 hours. TbISWI and a cross-reactive control band are indicated on the left and the size markers in kiloDaltons (kDa) on the right of the Western blot.

In the PF RNAi cell lines, TbISWI depletion resulted in 1.3 fold higher expression of –ESp eGFP (Figure 3.2.6). The fluorescence observed in the FL1 channel increased slightly by day 8 of *TbISWI* RNAi, while TbISWI protein had become undetectable on a Western Blot by 4 days of tetracycline induction (Figure 3.2.7).

As expected, the presence of an ES promoter in the +ESp RNAi PF cell lines resulted in the higher fluorescence observed in the FL1 channel compared to that of the –ESp RNAi lines even in their uninduced states (Figure 3.2.6). After TbISWI depletion, the +ESp eGFP was derepressed by ~4 fold, with fluorescence increasing further. Western blot analysis confirmed TbISWI knockdown to undetectable levels after 4 days of tetracycline induction (Figure 3.2.7).

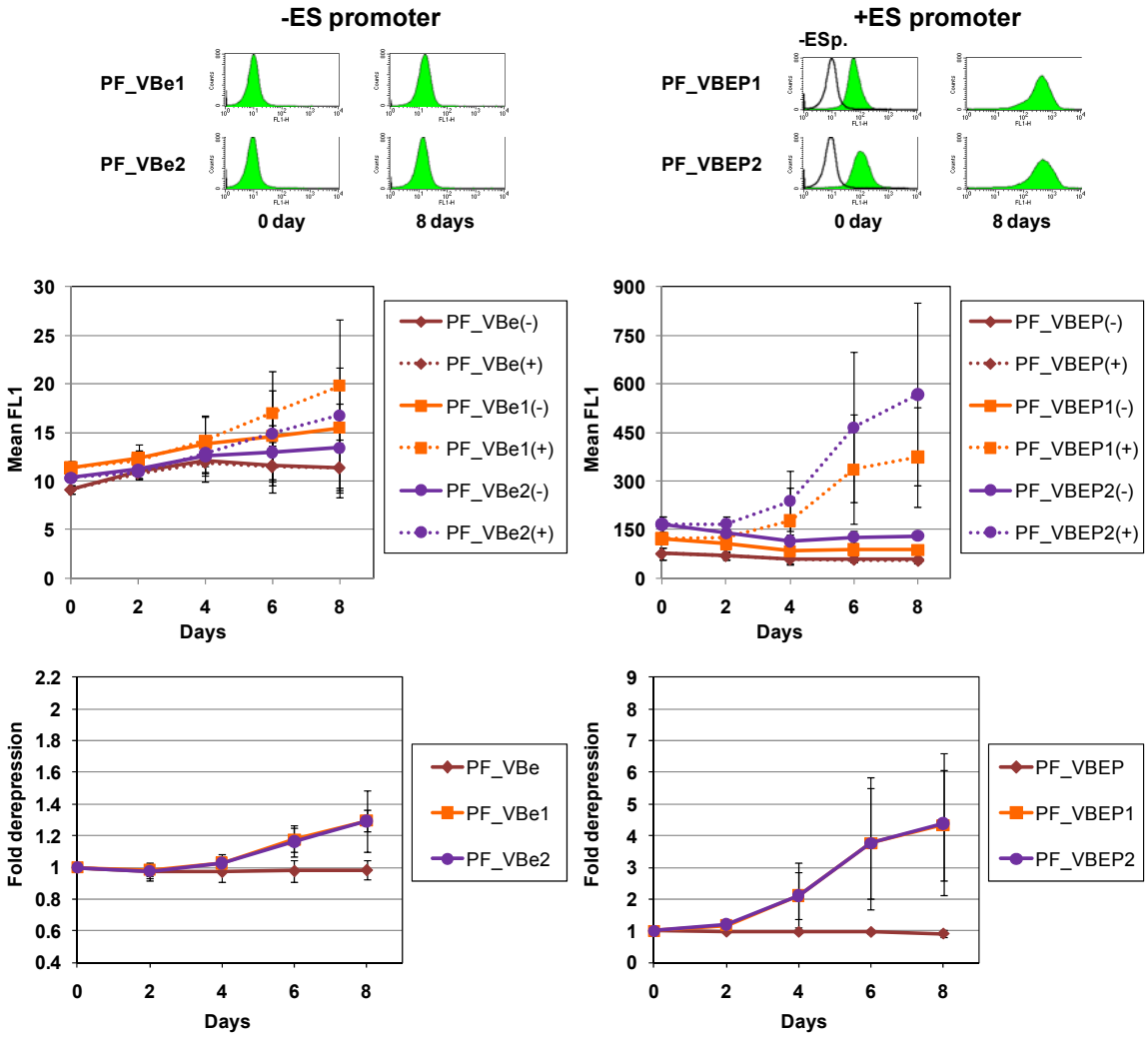


Figure 3.2.6 Monitoring eGFP derepression after induction of RNAi against *TbISWI* in procyclic form *T. brucei* VSG basic copy array reporter cell lines. Levels of expression of the *eGFP* gene were measured by flow cytometry for both the –ESp (left panel) and the +ESp (right panel) reporter lines. Top: representative flow cytometry traces for two cell lines at two different time-points are shown (day 0 and day 8), with eGFP fluorescence monitored in the FL1 channel (x-axis). The +ESp traces from day 0 also show an overlaid –ESp trace from the same time-point for comparison. Middle: the graph shows the mean fluorescence (FL1 channel) of eGFP in tetracycline induced (+) and uninduced (-) cell lines for different durations of RNAi induction; 100,000 cells were analysed per sample. Bottom: the graph shows the fold derepression of the *eGFP* gene for different durations of tetracycline induced TbISWI RNAi. Fold derepression was calculated as the ratio of mean fluorescence of induced cells to that of the uninduced cells. The graph shows the average of three independent experiments with the error bars indicating the standard deviation. The PF_VBe1 and PF_VBe2 cell lines were derived from PF_VBe after transfection with the TbISWI_RNAi construct. The PF_VBEP1 and PF_VBEP2 cell lines were derived from PF_VBEP after transfection with the TbISWI_RNAi construct.

Discussion

These results suggest that in both BF and PF trypanosomes, TbISWI may play a role in downregulating the fortuitous initiation of transcription from cryptic promoters in

the silent VSG 118 BCA. Alternatively, because the VSG 118 BCA is located on a large chromosome, the transcription observed after TbISWI depletion could be a result of impaired termination of transcription from an upstream pol II transcription unit. Unlike in the minichromosomes, the increase in transcription from the VSG BCA after TbISWI knockdown is comparable in the two life cycle stages.

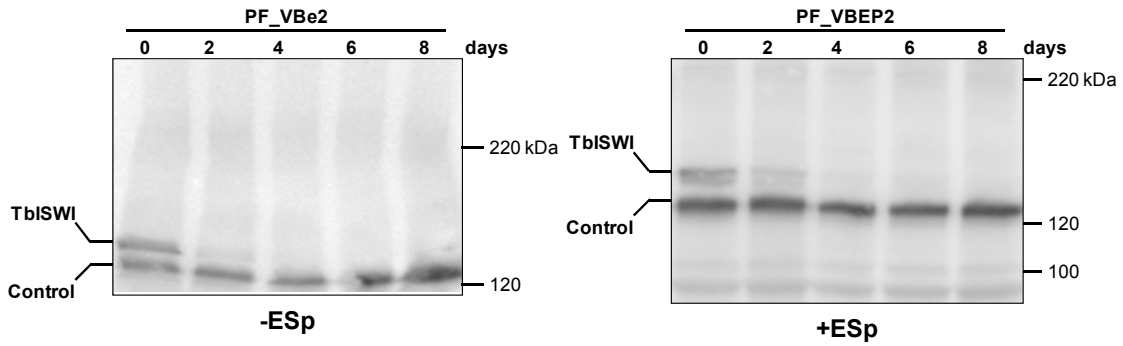


Figure 3.2.7 Depletion of TbISWI in procyclic form *T. brucei* VSG basic copy array reporter cell lines after induction of RNAi. Protein lysates from –ESp (top) and +ESp (bottom) cell lines were prepared after the induction of RNAi against *TbISWI* with tetracycline for 0, 2, 4, 6 and 8 days. TbISWI and a cross-reactive control band are indicated on the left and the size markers in kiloDaltons (kDa) on the right of the Western blot.

It is expected for genomic regions functioning as a repository for silent VSG genes, like the minichromosomes and the VSG BCA, to be tightly downregulated in the BF stage. However, these results suggest that TbISWI has a greater role in silencing the VSG BCA rather than the minichromosomes of BF cells. This is possibly to counter the greater chance of read-through into the BCA, a genomic locus next to constitutively active polycistronic transcription units. Alternatively, TbISWI depletion is able to result in more transcription from the VSG BCA because the greater sequence complexity of this region, compared to that of the simple sequence repeats of the minichromosomes, increases the chance of encoding cryptic promoters.

In the +ESp eGFP reporter cell lines, the BF cells exhibit more derepression than their PF counterparts. This is similar to results from previous experiments where the ES

promoters of the silent ESs were more derepressed in the BF than in the PF *T. brucei* (Hughes et al. 2007), and suggests that the repressive function of TbISWI is more pronounced in BF.

The observation that the mean FL1 fluorescence values for the uninduced –ESp reporter cell lines lie in the range of 5-6 units for BF cells and in the range of 10-15 units for PF cells suggests that basal levels of fortuitous initiation of transcription are lower in BF cells. However, unlike in the minichromosomes, in the VSG BCA the mean FL1 fluorescence values for the uninduced +ESp reporter cell lines in the BF cells (35-50 units) are quite comparable to the PF cells (40-60 units), with the PF values being slightly higher. The mean FL1 fluorescence values for the induced +ESp reporter cell lines are also quite comparable in the BF cells and the PF cells, both in the range of 350-400 units. Therefore, the higher derepression observed in the BF cell lines is probably an outcome of a slightly lower level of basal transcription than in the PF cells. Previous experiments have shown that exogenous genes inserted in the minichromosomes show less transcriptional variability than those in the large chromosomes; possibly because there is a greater probability of leaky read-through transcription from adjacent transcription units in the latter case (Wickstead et al. 2002).

As before, transcript processing signals used for the reporter genes are the same for both the –ESp and the +ESp constructs validating the use of protein fluorescence as an indirect measure of transcript abundance. As biological variability may give rise to different levels of fluorescence from the same cells across different experiments, the degree of derepression caused by TbISWI depletion is a better measure of its role in transcriptional regulation than the absolute fluorescence values.

3.3. Derepression of telomeric VSGs

Introduction

Sections 3.1 and 3.2 have shown that TbISWI is involved in repression of the transcriptionally silent areas of the genome where inactive *VSG* genes are known to reside- the minichromosomes and the VSG basic copy arrays. TbISWI depletion has also been shown to result in derepression of promoter proximal regions of silent Expression Sites in both bloodstream form and procyclic form *T. brucei*. However, no convincing evidence was found for derepression of the telomeric VSGs in those ESs (Hughes et al. 2007). However, all these results are based on reporter gene expression from the silent genomic regions. In this section, the effect of TbISWI depletion on transcription from some silent genomic loci is investigated at the level of the RNA transcript.

Results

TbISWI was depleted from bloodstream cell line T3-SA1 (Hughes et al. 2007) by tetracycline inducible RNAi and total RNA was isolated after different durations of induction. The RNA was then used to make cDNA (complementary DNA) by reverse transcription, and the transcript levels from different genomic regions and silent telomeric VSGs were quantified.

After levels of the *TbISWI* transcript were reduced to about half, an increase in other transcripts was observed (Figure 3.3.1B-D). The *eGFP* transcript located in a silent VSG221 ES increased by about 200 fold after 2 days of tetracycline induction of *TbISWI* RNAi (Figure 3.3.1C), while the rDNA spacer, the VSG118 basic copy array and the procyclin transcripts went up by 2-2.5 fold. The silent telomeric VSGs studied (VSG121,

VSG1.8 and VSGJS1) were derepressed by 4-8 fold (Figure 3.3.1D). These results confirm that TbISWI depletion results in derepression of the various silent loci.

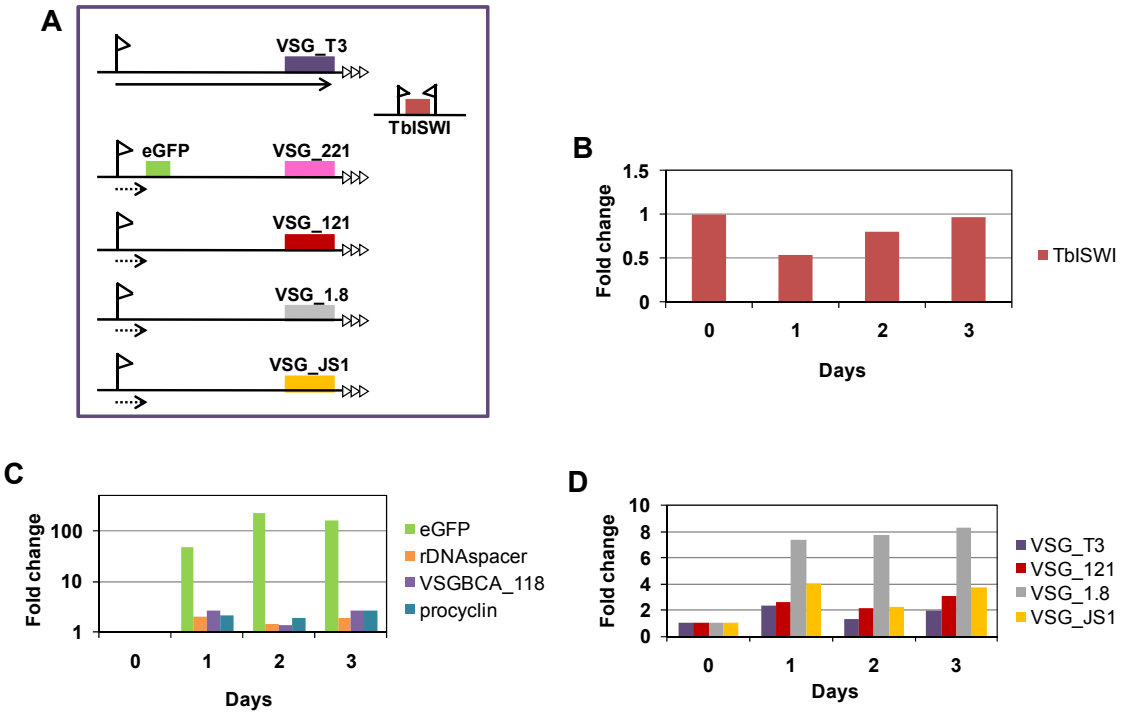


Figure 3.3.1 Depletion of TbISWI in bloodstream from *T. brucei* results in derepression of the telomeric silent VSGs. (A) The schematic shows a box representing a trypanosome with an active VSGT3 ES (the solid black arrow indicates active transcription) and other silent ESs (the dotted arrows indicate non-processive transcription). The silent 221 ES contains an eGFP reporter gene downstream of the ES promoter. The cell also contains a *TbISWI* RNAi construct. qPCR using total cDNA was performed to quantitate the abundance of transcripts and the concentration of each transcript was normalised against that of *γ-tubulin*. Fold change was then calculated by dividing the normalised concentration of the transcript on a particular day of induction by that of its day 0 concentration. (B) shows the depletion of the *TbISWI* transcript during RNAi induction. (C) shows the derepression of transcripts from various silent loci during RNAi induction. (D) shows the transcript abundance of the telomeric VSGs after *TbISWI* RNAi induction.

Discussion

While earlier results had shown the derepression of an ES promoter proximal reporter gene (Hughes et al. 2007), results from this section suggest that derepression mediated by TbISWI depletion extends to the telomeric VSGs, with the degree of derepression decreasing along the length of the ES. The promoter proximal reporter eGFP

gene in a silent ES is derepressed by about 200 fold, while the telomeric VSGs are derepressed by only ~4-8 fold.

These results also demonstrate that the degree of derepression of the silent loci as measured by transcript abundance is higher than that measured by reporter fluorescence. While the degree of derepression of *eGFP* as determined by fluorescence from a silent ES was found to be 30-60 fold in the BF *T.brucei* (Hughes et al. 2007), at transcript level it is seen to be ~200 fold derepression (Figure 3.3.1C). Similarly, the degree of increase in transcripts from the VSG BCA (as monitored by qPCR; Figure 3.3.1) was found to be ~2.5 fold, while the increase in expression of the –ESp eGFP reporter gene (as monitored by fluorescence in Chapter 3) inserted in them was only about 1.5 fold.

Chapter 4

Role of TbISWI in transcriptional regulation of Pol-I and Pol-II transcription units

4.1. rDNA spacers

Introduction

Trypanosomes encode about 100 rRNA genes arranged in multiple pol-I transcription units clustered on several large chromosomes (Wagner et al. 1990; Castro et al. 1981). Eight of these rRNA transcription units are arranged in tandem arrays separated by non-transcribed spacer regions of about 8 kb each (Berriman et al. 2005; White et al. 1986). The non-transcribed rDNA spacer sequences can also be found dispersed on several chromosomes at a total copy number of 15-20 (Wickstead et al. 2003). Indirect evidence from expression of ectopically inserted genes suggests that there are differences in expression from one rDNA spacer to another (Biebinger et al. 1996; Alsford et al. 2005).

An attempt was made to understand the role of TbISWI in transcriptional repression operating within the rDNA spacer. The experimental strategy used in Chapter 3 above for the minichromosomes and the VSG basic copy array was adapted for this purpose (Figure 4.1.1). A construct containing eGFP reporter gene was inserted, with and without an upstream ES promoter, into an rDNA spacer and its expression monitored after RNAi-mediated TbISWI depletion. The previously used –ESp and +ESp reporter backbone constructs, peGFP_RPtubBlast and pESeGFP_RPtubBlast, were modified to target them into the rDNA spacer (Figure 4.1.2).

ISWI complexes in yeast have been shown to play a role in the regulation of transcription initiation, elongation, as well as termination (Mellor et al. 2004; Morillon et al. 2003; Alén et al. 2002). Since the rDNA spacer targeting sequence is also found in other loci in the genome, it would be difficult to interpret any increased transcription from the reporter constructs if they integrated at sites that are not in between two rDNA

transcription units. Therefore, the reporter cell lines obtained were tested by Southern Blot analyses for reporter construct integration in between two rDNA units, so that readthrough from an adjacent Pol II transcription unit can be avoided and any possible increase in transcription interpreted unambiguously.

Results

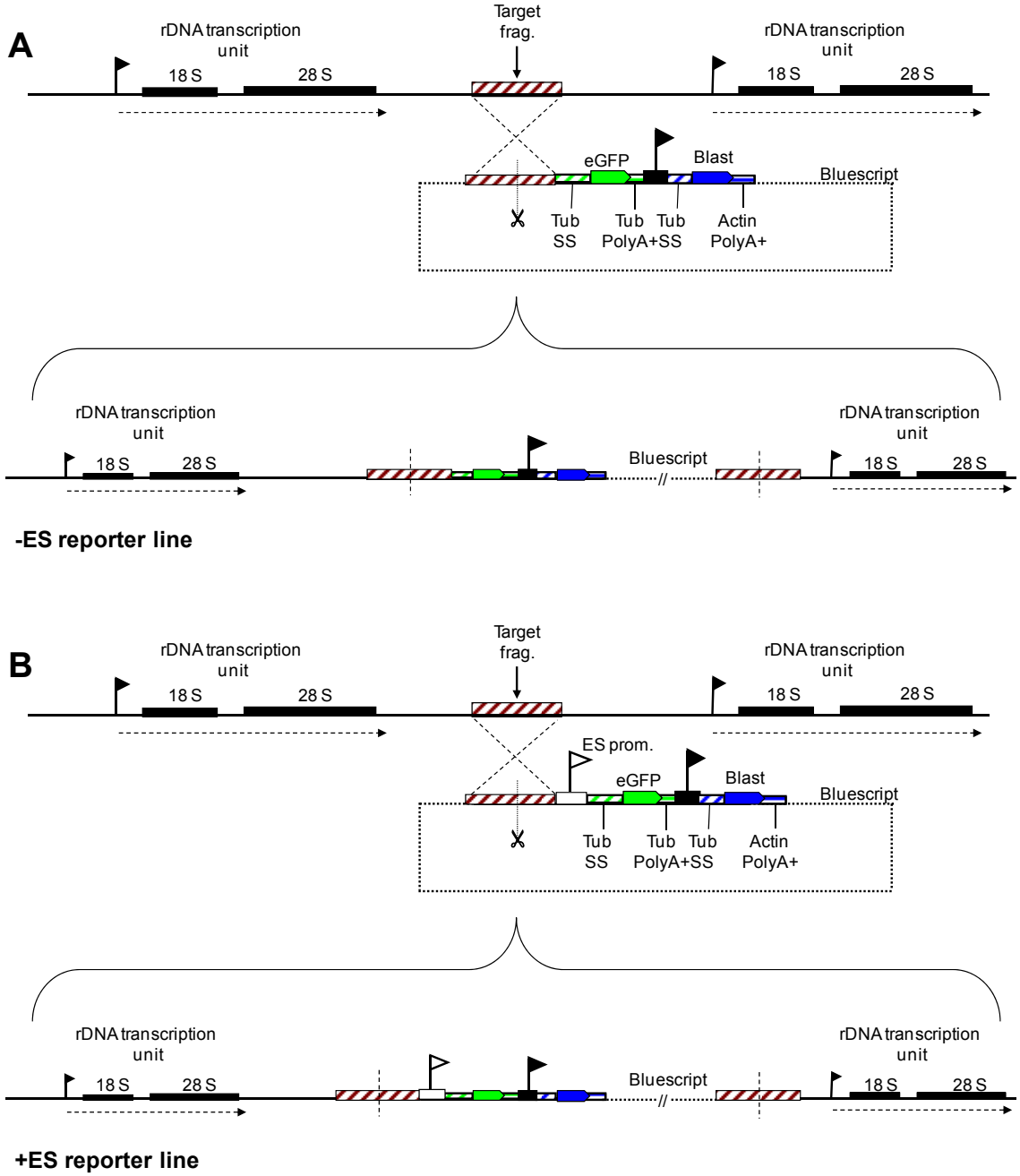
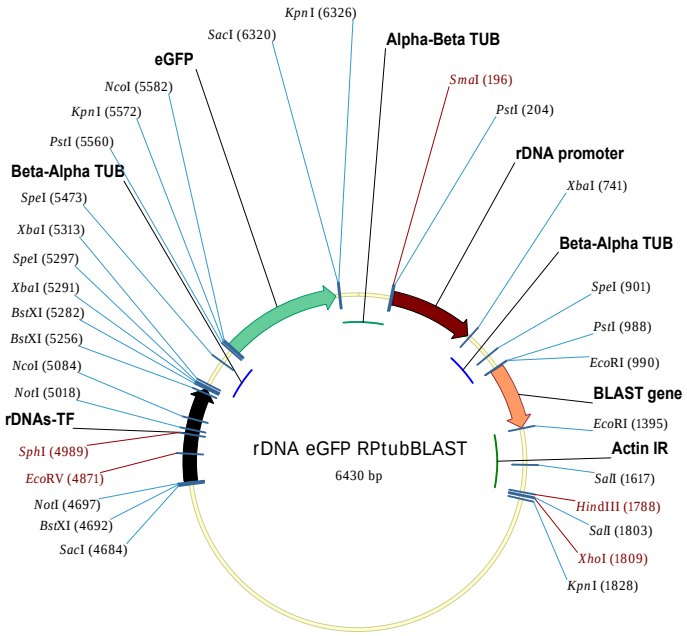


Figure 4.1.1 Strategy for reporter *eGFP* integration into the rDNA spacer. The schematics show integration of an *eGFP* reporter gene into the spacer region between two rDNA transcription

units. While (A) shows a –ESp reporter construct without a promoter sequence upstream of *eGFP*, (B) shows a +ESp reporter construct with an Expression Site (ES) promoter upstream of *eGFP*. In both variants of the reporter constructs, the *eGFP* gene with flanking transcript processing signals (tubulin intergenic regions) is followed downstream by an rDNA promoter and a *Blasticidin* resistance gene cassette. Linearised where indicated by the scissors, the constructs integrate into the rDNA spacer by a single-crossover homologous recombination event.

The rDNA spacer targeting –ESp eGFP reporter construct, rDeGFP_RPtubBlast, was made by inserting the rDNA spacer targeting fragment from p2T7^{Ti}A/GFP (LaCount et al. 2002) into the NotI linearised –ESp eGFP reporter backbone as detailed in Figure 4.1.2A. The construct was verified by diagnostic digests and sequencing. Similarly, the rDNA spacer targeting +ESp eGFP reporter construct, rDEPeGFP_RPtubBlast, was made by inserting the rDNA spacer targeting fragment into the NotI linearised +ESp eGFP reporter backbone as detailed in Figure 4.1.2B, and verified by diagnostic digests and sequencing.

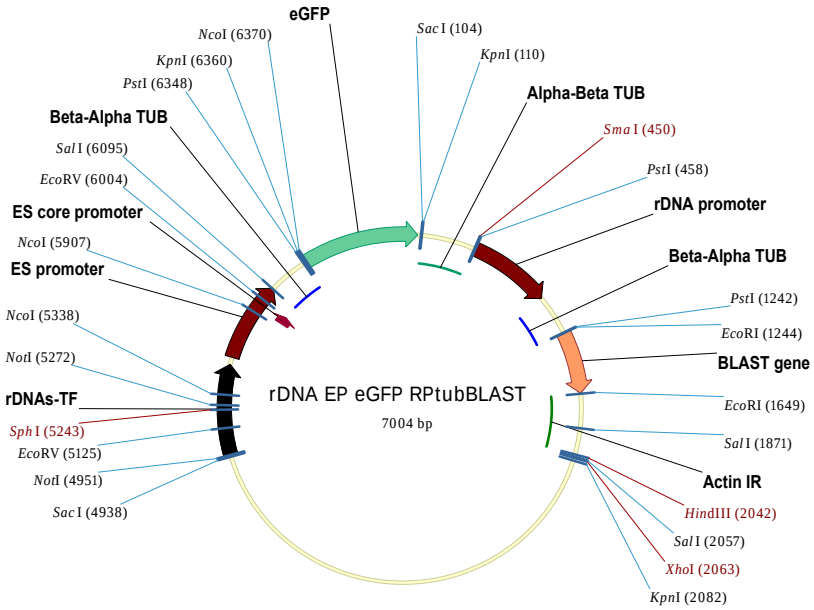
A



STEPS:

1. The rDNA spacer Targeting Fragment was isolated from p2T7tiA_GFP_phleo (LaCount et al, 2002) by cutting with SacI+PstI and then blunted.
2. peGFP_RPtubBLAST was linearised with NotI (+blunted) and the blunt-ended rDNA spacer Targeting Fragment from Step 1 was ligated into it. Correct ligation of the construct in the forward orientation was verified by PCRs linking over the junctions and restriction digestions of the obtained construct. This construct was called rDNA_eGFP_RPtubBLAST.
3. rDNA_eGFP_RPtubBLAST was linearised with SphI for targeting into the trypanosome genome.

B



STEPS:

1. The rDNA spacer Targeting Fragment was isolated from p2T7tiA_GFP_phleo (LaCount et al, 2002) by cutting with SacI+PstI and then blunted.
2. pESP_tubeGFPtub_RPtubBLAST was linearised with NotI (+blunted) and the blunt-ended rDNA spacer Targeting Fragment from Step 1 was ligated into it. Correct ligation of the construct in the forward orientation was verified by PCRs linking over the junctions and restriction digestions of the obtained construct. This construct was called rDNA_EPeGFP_RPtubBLAST.
3. rDNA_EPeGFP_RPtubBLAST was linearised with SphI for targeting into the trypanosome genome.

Figure 4.1.2 Maps of the rDNA spacer targeting reporter constructs. The two reporter constructs were obtained by inserting an rDNA spacer targeting fragment into the respective reporter construct backbones. While (A) rDeGFP_RPtubBlast was derived from the –ESp reporter backbone of peGFP_RPtubBlast, (B) rDEPeGFP_RPtubBlast was derived from the +ESp reporter backbone of pEPeGFP_RPtubBlast. Both constructs were linearised for transfection by SphI digestion.

In order to obtain the eGFP reporter cell lines, bloodstream form (BF) S16.221 and procyclic form (PF) 29-13 *T. brucei* were transfected with the –ESp eGFP and the +ESp eGFP reporter constructs described above. Both of these parental cell lines contain genes encoding constitutively expressed Tet repressor and T7 RNA polymerase, allowing tetracycline inducible RNAi (Wirtz et al. 1999). The targeting was achieved by linearising rDeGFP_RPtubBlast and rDEPeGFP_RPtubBlast constructs with SphI restriction in the middle of the rDNA spacer targeting sequence, generating a double strand break inside a homologous region to facilitate recombination (Clayton 1999). A Bluescript vector backbone also gets integrated into the target site using this strategy.

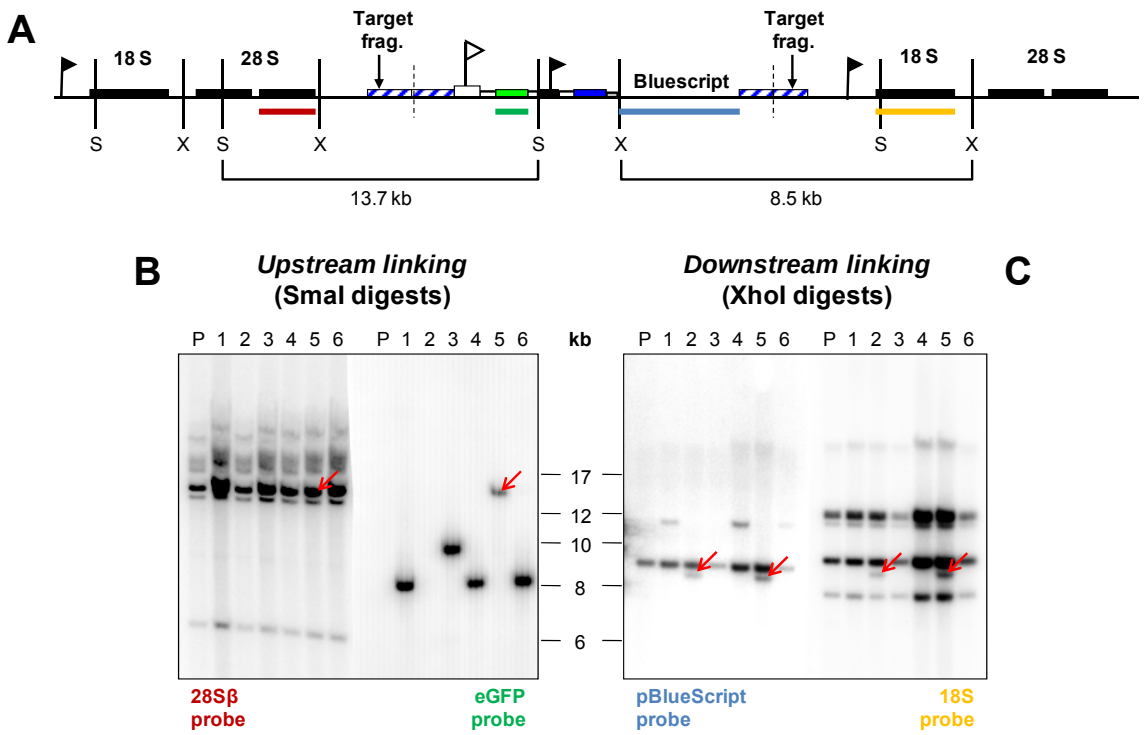


Figure 4.1.3 Integration of reporter *eGFP* construct into the rDNA spacer. In order to verify the integration of the *eGFP* reporter construct in between two rDNA units, Southern blot analysis was performed on *T. brucei* rDNA spacer transformants. (A) A schematic of a *T. brucei* rDNA spacer locus with an integrated reporter construct is shown (not to scale). The points of integration are indicated with dotted vertical lines, the restriction sites of SmaI (S) and XhoI (X) with solid vertical lines, and the probe fragments as coloured bars under the locus map. (B) SmaI digested genomic DNA from different cell lines was run on a gel in two separate panels. The panels were probed for *28Sβ* and *eGFP* to show linkage between the digested DNA fragments. Bands that match up are marked with red arrows. (C) XhoI digested genomic DNA from different cell lines was run on a gel in two separate panels. The panels were probed for *18S* and pBlueScript to show linkage between the digested DNA fragments. Bands that match up are marked with red arrows. Linking bands are marked with red arrows. Cell Lines used- P: parental (29-13), 1: PF_rDEP^{b}, 2: PF_rDEP^{c}, 3: PF_rDEP^{d}, 4: PF_rDEP^{e}, 5: PF_rDEP, 6: PF_rDEP^{f}.

To verify the integration of the reporter constructs into a spacer region in between two rDNA transcription units, the reporter cell lines were analysed by restriction enzyme digestion mapping around the integrated construct and probing for an upstream 28Sβ rDNA gene and a downstream 18S rDNA gene using Southern blot analysis (Figure 4.1.3). Only in 1 of 12 BF lines (data not shown) and in 1 of 12 PF lines (Figure 4.1.3) tested was the reporter construct found to have integrated in between two rDNA transcription units.

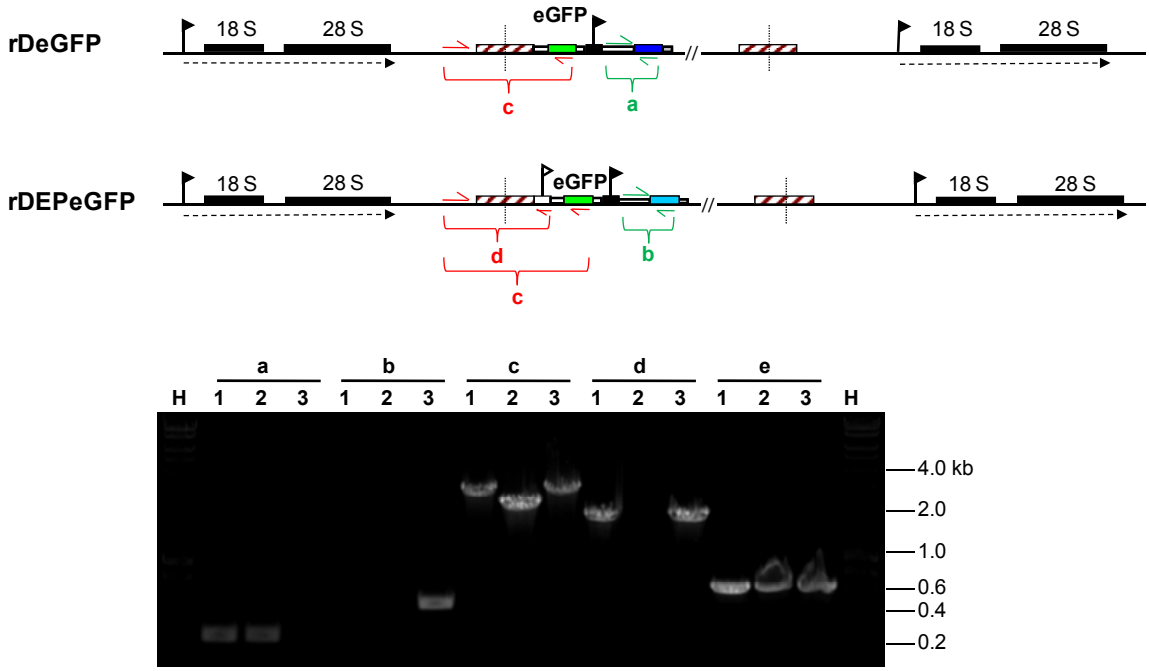


Figure 4.1.4 Replacement of a –ESp rDNA reporter construct with a +ESp reporter construct in BF *T. brucei*. In order to obtain a +ESp rDNA spacer reporter cell line, the *Blasticidin* gene containing –ESp reporter construct in BF_rDe was replaced by a *Phleomycin* gene containing +ESp rDNA targeting reporter construct. PCR amplification to test correct construct integration was performed using genomic DNA isolated from (1) PF_rDEP- a control PF cell line with an integrated +ESp reporter construct, (2) BF_rDe- a BF reporter line with an integrated –ESp reporter construct and (3) the BF_rDEP transformant was used to confirm integration into the correct genomic locus. The top panel shows a schematic of primer sets used for this purpose. Primer sets 'a' and 'b' were used to confirm the presence of the drug resistance genes *Blasticidin* and *Phleomycin*, respectively. Primer set 'c' was used to amplify over the upstream construct integration point (indicated by dotted vertical line) producing differently sized amplification products for the –ESp and the +ESp reporter cell lines. Primer set 'd' was used to confirm the presence of an Expression Site promoter. Primer set 'e' amplified a segment of the RNA polymerase I large subunit, and was used as a positive control for PCR amplification from the genomic DNA. The bottom panel shows the products of the PCR reactions using genomic DNA from the above mentioned cell lines, with the size markers indicated on the right of the gel pictures in kilobases (kb). Primer sets used: 'a' (batub-1835s & BSR-138as), 'b' (batub-1835s & PhleoII-372as), 'c' (rDNASTF-17610s & eGFP-183as), 'd' (rDNASTF-17610s & DES-1291as) and 'e' (PolI-484s & PolI-1123as).

As the only BF cell line where the reporter construct had integrated between two rDNA transcription units was a –ESp line, the equivalent +ESp line was obtained by replacing the –ESp construct with a +ESp construct using transfection with a SacII linearised rDEPeGFP_RPtubPhleo and verified by PCR linking (Figure 4.1.4). Similarly, the only PF cell line with reporter construct integration between two rDNA transcription units was a +ESp cell line. Therefore, the equivalent –ESp line was obtained by replacing the +ESp construct with a –ESp construct after transfection with a SacII linearised

rDeGFP_RPtubPhleo and verified by PCR linking (Figure 4.1.5). This would also enable comparison between results from the –ESp and the +ESp cell lines since the reporter construct would be present at the same genomic locus in both of them.

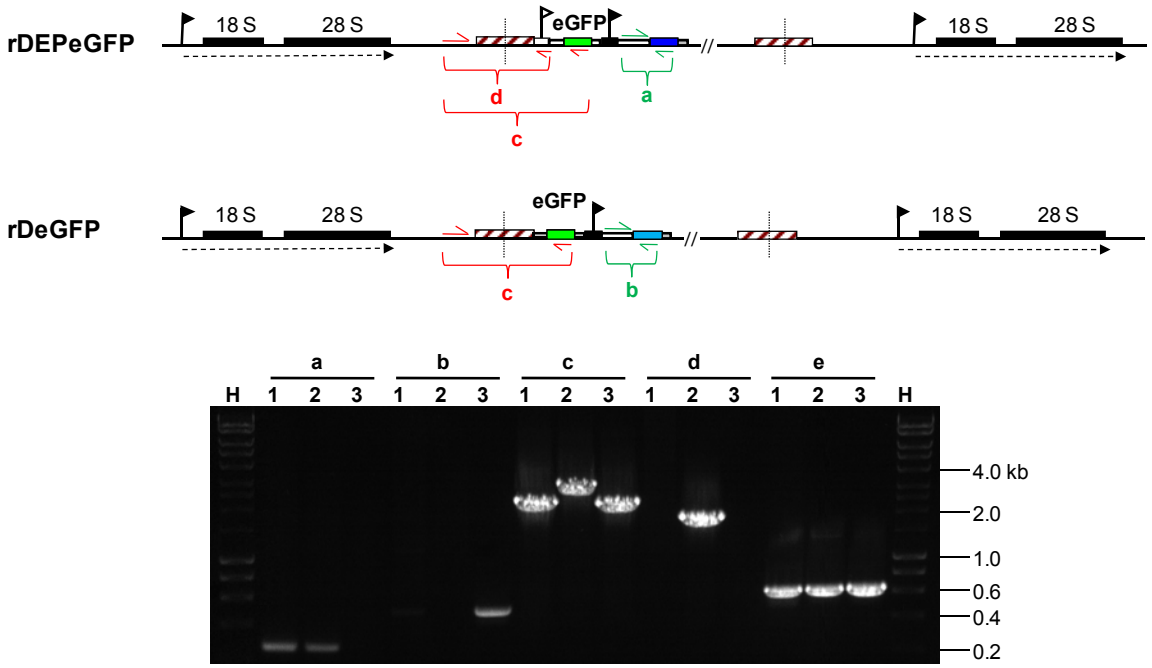


Figure 4.1.5 Replacement of a +ESp rDNA reporter construct by a –ESp reporter construct in PF *T. brucei*. In order to obtain a –ESp rDNA spacer reporter cell line, the *Blasticidin* gene containing +ESp reporter construct in PF_rDEP was replaced by a *Phleomycin* gene containing –ESp rDNA targeting reporter construct. PCR amplification to test correct construct integration was performed using genomic DNA isolated from (1) BF_rDe- a control BF cell line with an integrated –ESp reporter construct, (2) PF_rDEP- a PF reporter line with an integrated +ESp reporter construct and (3) the PF_rDe transformant was used to confirm integration into the correct genomic locus. The top panel shows a schematic of primer sets used for this purpose. Primer sets 'a' and 'b' were used to confirm the presence of the drug resistance genes *Blasticidin* and *Phleomycin*, respectively. Primer set 'c' was used to amplify over the upstream construct integration point (indicated by dotted vertical line) producing differently sized amplification products for the –ESp and the +ESp reporter cell lines. Primer set 'd' was used to confirm the presence of an Expression Site promoter. Primer set 'e' amplified a segment of the RNA polymerase I large subunit, and was used as a positive control for PCR amplification from the genomic DNA. Primer sets used: 'a' (batub-1835s & BSR-138as), 'b' (batub-1835s & PhleoII-372as), 'c' (rDNASTF-17610s & eGFP-183as), 'd' (rDNASTF-17610s & DES-1291as) and 'e' (PolI-484s & PolI-1123as).

After verification of correct integration of the reporter constructs, reporter cell lines BF_rDe, BF_rDEP, PF_rDe and PF_rDEP were transfected with the MC[177] TbISWI-A RNAi construct (Hughes et al. 2007) and two RNAi cell lines per reporter cell line were obtained. The *T. brucei* RNAi cell lines obtained were called

BF_rDe1, BF_rDe2, BF_rDEP1, BF_rDEP2, PF_rDe1, PF_rDe2, PF_rDEP1 and PF_rDEP2.

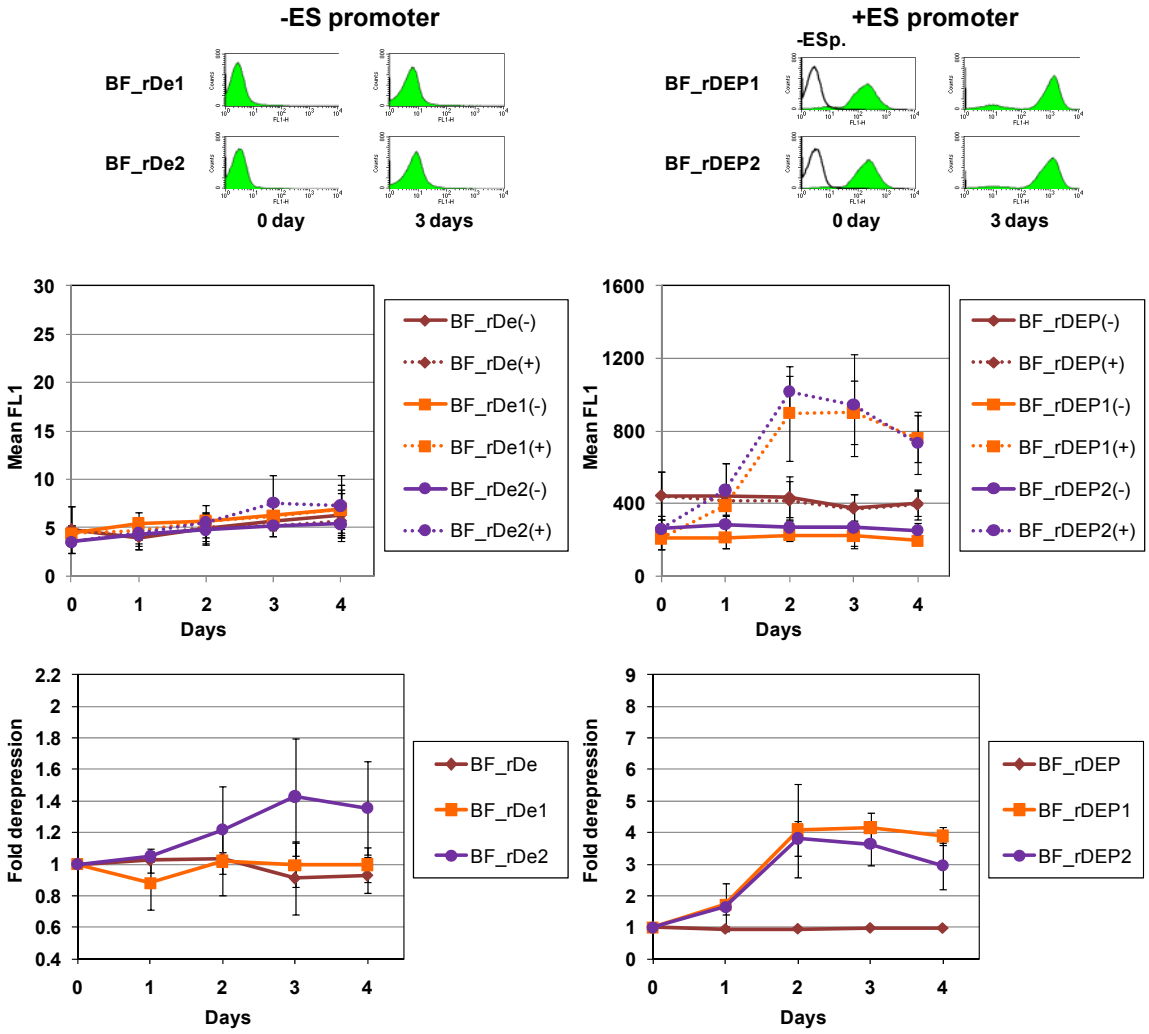


Figure 4.1.6 Monitoring derepression of eGFP after induction of RNAi against *TbISWI* in bloodstream form rDNA spacer reporter cell lines. Levels of expression of the *eGFP* gene were measured by flow cytometry for both the –ESp (left panel) and the +ESp (right panel) reporter lines. Top: representative flow cytometry traces for two cell lines at two different time-points are shown (day 0 and day 3), with eGFP fluorescence monitored in the FL1 channel (x-axis). The +ESp traces from day 0 also show an overlaid –ESp trace from the same time-point for comparison. Middle: the graph shows the mean fluorescence (FL1 channel) of eGFP in tetracycline induced (+) and uninduced (-) cell lines for different durations of induction; 100,000 cells were analysed per sample. Bottom: the graph shows the degree of derepression of the *eGFP* gene for different durations of tetracycline induced *TbISWI* RNAi. The degree of derepression was calculated as the ratio of mean fluorescence of induced cells to that of the uninduced cells. The graph shows the average of three independent experiments with the error bars indicating the standard deviation. The BF_rDe1 and BF_rDe2 cell lines were derived from BF_rDe by transfection with the TbISWI_RNAi construct. The BF_rDEP1 and BF_rDEP2 cell lines were derived from BF_rDEP by transfection with the TbISWI_RNAi construct.

In order to study the effect of TbISWI depletion on reporter eGFP expression, the RNAi cell lines were induced using 750 ng ml^{-1} tetracycline and fluorescence was measured in the FL1 channel using flow cytometry as described in the Materials and Methods. Figures 4.1.6 and 4.1.8 show representative FACS traces of eGFP fluorescence of different cell lines in the top panel, the mean FL1 fluorescence values in the middle panel, and the degree of derepression of the eGFP gene, expressed as the ratio of fluorescence of induced versus uninduced cells, at different durations of RNAi induction in the bottom panel. The depletion of the TbISWI protein can be seen in Figures 4.1.7 and 4.1.9.

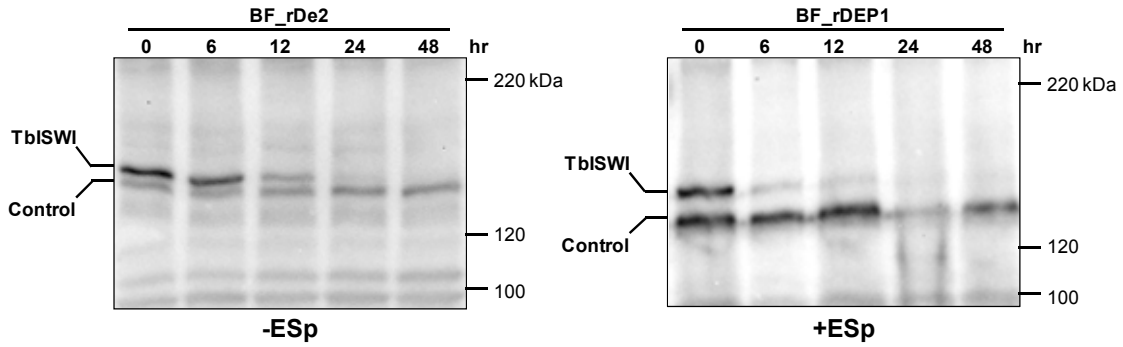


Figure 4.1.7 Depletion of TbISWI in bloodstream form rDNA spacer reporter cell lines after induction of *TbISWI* RNAi. Protein lysates from –ESp (left) and +ESp (right) cell lines were prepared after the induction of RNAi against *TbISWI* with tetracycline for 0, 6, 12, 24 and 48 hours. TbISWI and a cross-reactive control band are indicated on the left and the size markers in kiloDaltons (kDa) on the right of the Western blot.

In the BF RNAi cell lines, TbISWI depletion did not result in significant transcription of the –ESp eGFP, although some leaky transcription with overlapping error bars (~1.4 fold derepression) was observed in the BF_rDe2 cell line (Figure 4.1.6). The degree of fluorescence observed in the FL1 channel did not change much even on day 3 after *TbISWI* RNAi, despite TbISWI protein becoming undetectable on a Western blot after 48 hours of tetracycline induction (Figure 4.1.7).

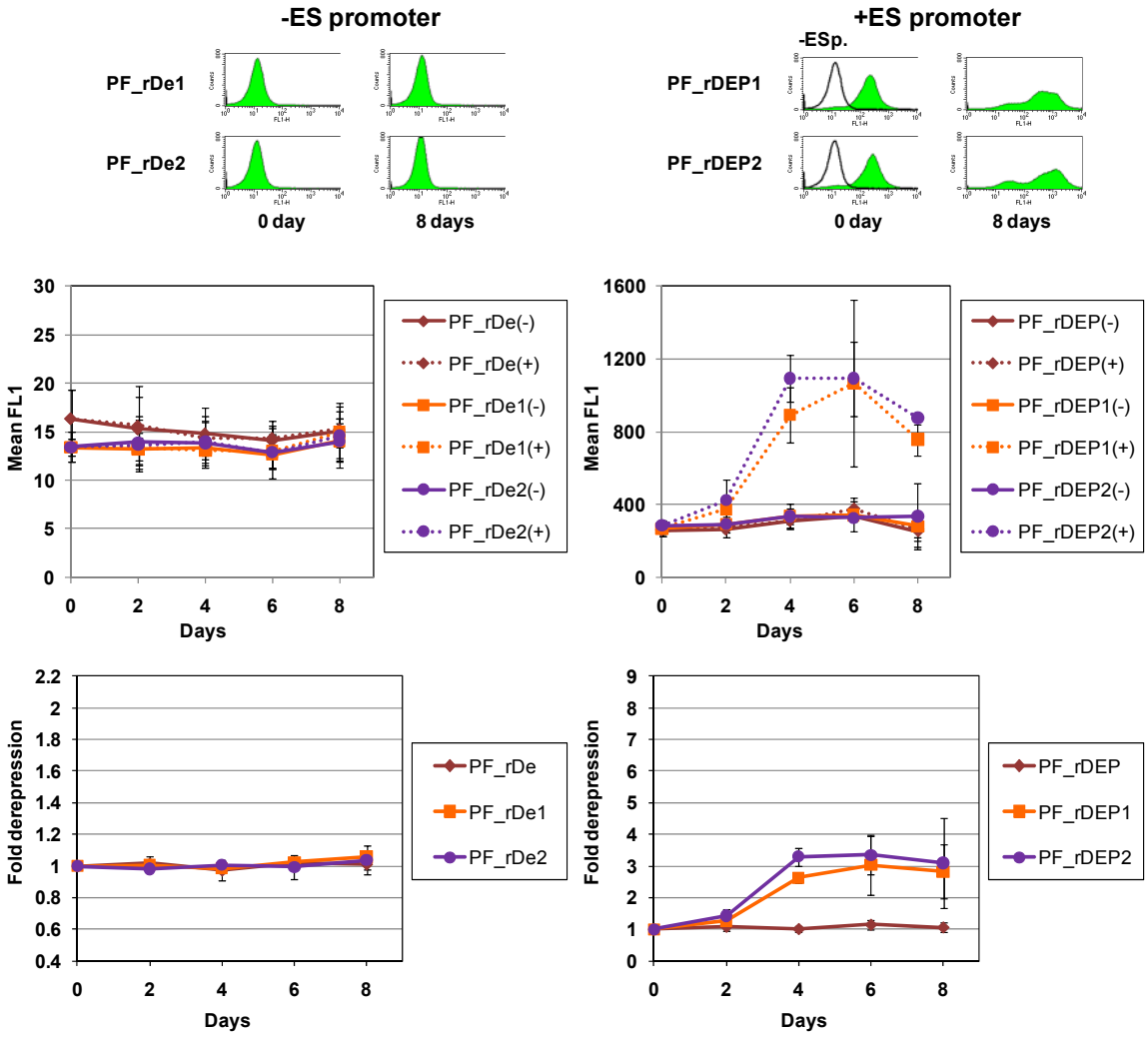


Figure 4.1.8 Monitoring eGFP derepression after induction of RNAi against *TbISWI* in Procyclic form rDNA spacer reporter cell lines. Levels of expression of the *eGFP* gene were measured by flow cytometry for both the –ESp (left panel) and the +ESp (right panel) reporter lines. Top: representative flow cytometry traces for two cell lines at two different time-points are shown (day 0 and day 8), with eGFP fluorescence monitored in the FL1 channel (x-axis). The +ESp traces from day 0 also show an overlaid –ESp trace from the same time-point for comparison. Middle: the graph shows the mean fluorescence (FL1 channel) of eGFP in tetracycline induced (+) and uninduced (-) cell lines for different durations of induction; 100,000 cells were analysed per sample. Bottom: the graph shows the fold derepression of the *eGFP* gene after the induction of *TbISWI* RNAi for different time points. Fold derepression was calculated as the ratio of mean fluorescence of induced cells to that of the induced cells. The graph shows the average of three independent experiments with the error bars indicating the standard deviation. PF_rDe1 and PF_rDe2 were derived from PF_rDe by transfection with the TbISWI_RNAi construct. PF_rDEP1 and PF_rDEP2 were derived from PF_rDEP by transfection with the TbISWI_RNAi construct.

In the +ESp RNAi BF cell lines, the fluorescence observed in the FL1 channel was increased compared to that in –ESp RNAi lines even in their uninduced states (Figure 4.1.6). After TbISWI depletion, the +ESp eGFP cell line showed ~4 fold

derepression. Western blot analysis confirmed that TbISWI protein is highly depleted after 48 hours of tetracycline induction (Figure 4.1.7).

In the PF RNAi cell lines, TbISWI depletion did not result in derepression of the –ESp eGFP (Figure 4.1.8). The degree of fluorescence observed in the FL1 channel changed very little even on day 8 of *TbISWI* RNAi, despite TbISWI protein becoming highly depleted after 6 days of tetracycline induction (Figure 4.1.9).

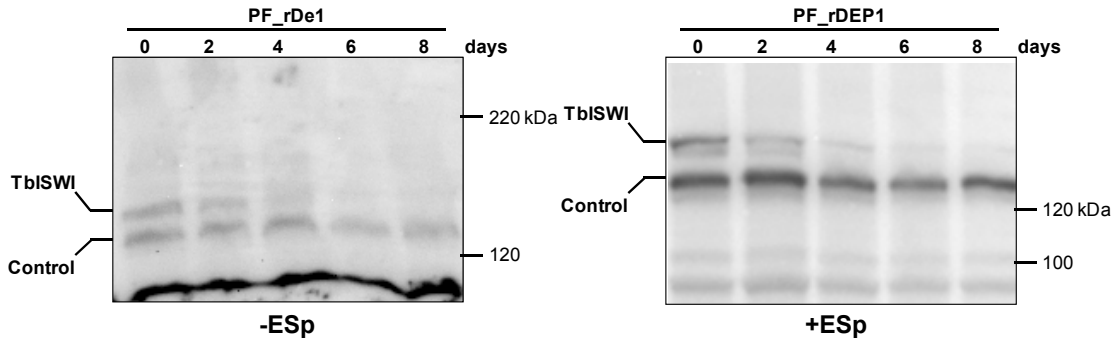


Figure 4.1.9 Depletion of TbISWI in procyclic form rDNA spacer reporter cell lines after the induction of RNAi. Protein lysates from –ESp (left) and +ESp (right) cell lines were prepared after the induction of RNAi against *TbISWI* with tetracycline for 0, 2, 4, 6 and 8 days. TbISWI and a cross-reactive control band are indicated on the left and the size markers on the right of the Western blot.

In the +ESp RNAi PF cell lines, fluorescence observed in the FL1 channel was increased compared to that of the –ESp RNAi lines even in their uninduced states (Figure 4.1.8). After TbISWI depletion, the +ESp eGFP was derepressed by ~3 fold. Western blot analysis confirmed that TbISWI protein is barely detectable after 6 days of tetracycline induction (Figure 4.1.9).

Discussion

Reporter cell lines where the reporter construct has integrated in between two rDNA transcription units could provide useful information about the role of TbISWI in rDNA transcription termination. Additionally, this would also obviate misinterpretation

of transcriptional read-through from an adjacent Pol-II transcription unit on a large chromosome. Southern blot analysis was therefore performed to verify the correct integration of the eGFP reporter construct downstream of a 28Sβ rRNA gene and upstream of an 18S rRNA gene, establishing its correct integration in between two rDNA units.

A BLAST search of the *T. brucei* 927 genome revealed that there are at least four rDNA spacer regions that are flanked by two rDNA transcription units, suggesting a theoretical integration rate of 4/15. However, only 1 of 12 cell lines had the reporter construct integrating at the required spacer region. This lower success rate could be because tightly downregulated non-transcribed regions may be less accessible for DNA recombination. Alternatively, this could be due to mismatches between the fragment used for construct targeting and the targeted locus, which could introduce a bias for where the constructs eventually integrate (Blundell et al. 1996).

Since only one BF –ESp reporter cell line was found to have the construct integrated correctly, the rDNA spacer targeting +ESp reporter cell line was obtained by replacing the integrated –ESp report construct (Blasticidin based) with a +ESp reporter construct (Phleomycin based). Similarly, the single PF +ESp reporter cell line was used to obtain the –ESp reporter cell line by replacing the integrated +ESp report construct (Blasticidin based) with a –ESp reporter construct (Phleomycin based). These construct replacements were confirmed using PCR linking over the part of the construct in between the targeting fragment and the eGFP reporter gene.

From the above results of the FACS analysis of RNAi cell lines, it is evident that TbISWI plays a role in the silencing of the rDNA spacer in both BF and PF *T. brucei*.

Similar to the minichromosome and VSG basic copy array (BCA) targeting constructs, the basal mean FL1 values are lower in the BF cells (5-7 units) than in the PF cells (14-16 units), suggesting stricter transcriptional repression in BF. However, unlike in these two loci, where there appeared to be stricter control of background levels of transcription in BF cells even after TbISWI depletion, the –ESp eGFP in the rDNA spacer appears to be more strictly controlled in the PF cells, with almost no leaky transcription after TbISWI RNAi.

One of the –ESp eGFP BF RNAi lines shows ~1.4 fold derepression, but the error bars are large and overlap with the parental, suggesting that this result is insignificant. Both of the RNAi clones analysed are derived from the same reporter cell line, and therefore have the reporter gene in the same rDNA locus. It is noteworthy here that in comparison to minichromosomes, an exogenous promoter inserted into the rDNA spacer has previously been shown to be >100 fold less regulatable in BF cells and >20 fold less regulatable in PF cells (Wickstead et al. 2002). Also, the rDNA spacer was shown to be 10-25 fold less repressed than the VSG 118 basic copy array in BF (Horn et al. 1997). The variability in the behaviour between the two clones is in agreement with this.

Among the +ESp eGFP reporter cell lines, the BF cells exhibit greater derepression compared to their PF counterparts, as was the case with the silent ESs, another pol-I transcribed locus (Hughes et al. 2007). Since the mean FL1 values of the derepressed cells are quite comparable in the BF and the PF cell lines, this difference is largely attributable to the lower starting point in the BF cells. Even in its most derepressed state, the activity of the ES promoter in the rDNA spacer is about 4-5 fold less than that in a fully active ES. As suggested in a discussion of ES regulation (Hughes et al. 2007), there may be multiple layers of transcriptional control operating on the

rDNA spacer too. These observations also imply that when TbISWI is depleted, transcription from the rDNA locus is comparable in both lifecycle stages, and any other parallel repressive mechanisms acting on the spacers are probably acting equally well in both BF and PF.

A curious observation can be made in the BF +ESp and the PF –ESp cell lines. In both of these cell lines, the reporter line has a slightly higher mean fluorescence (FL1 channel) than its derivative RNAi lines (Figures 4.1.6 and 4.1.8). One common link between these lines is the fact that they were both derived by replacing a *Blasticidin* resistance gene based reporter construct with the other *Phleomycin* resistance gene based counterpart. Phleomycin acts on cells by binding to and causing DNA damage, which has been shown to cause promoter derepression in trypanosomes (Pietsch et al. 1968; Iwata et al. 1971; Sheader et al. 2004). Even if the presence of phleomycin is causing increased derepression of trypanosome promoters, it is unclear why their derivative RNAi lines do not have similarly derepressed chromatin in the uninduced states.

As before, transcript processing signals used for the *eGFP* reporter gene are the same for both the –ESp and the +ESp reporter constructs, allowing comparison with results from the other cell lines used in Chapter 3.

4.2. Procyclin EP1 locus

Introduction

Procyclin, the GPI-anchored invariant surface protein expressed by insect form trypanosomes is present on the cell surface thereby protecting the trypanosome from the tsetse fly proteases (Pearson 2001). There are about eight to ten EP and GPEET procyclin genes present in clusters of two or three in RNA pol-I transcription units (Vassella et al. 2001). If the BF trypanosomes were to expose these invariant proteins to the mammalian bloodstream, they would be cleared by the immune system. To avoid this from happening, the procyclin genes are kept downregulated by multiple layers of control in BF cells. While transcriptional downregulation of procyclin accounts for about ten fold repression of this locus (Hotz et al. 1998), another ten fold downregulation is achieved by the rapid degradation of any procyclin transcript produced in BF cells (Furger et al. 1997; Hotz et al. 1997).

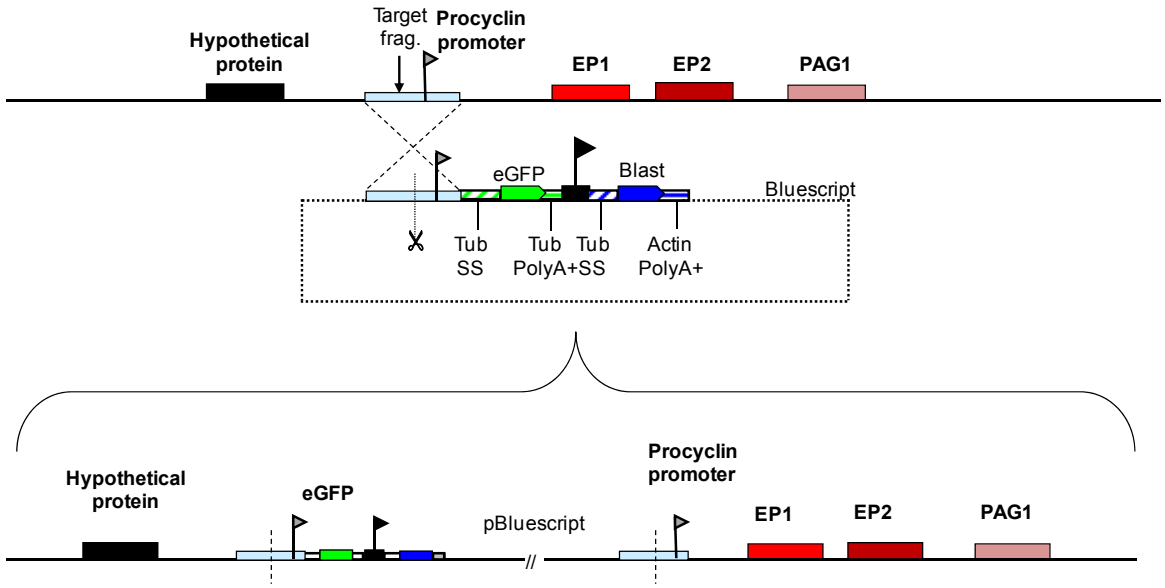


Figure 4.2.1 Reporter *eGFP* integration into the EP1 procyclin locus. The schematic shows integration of an *eGFP* reporter gene into the EP1 procyclin locus. In the reporter construct, the *eGFP* gene with flanking transcript processing signals (tubulin intergenic regions) is followed downstream by an rDNA promoter and a *Blasticidin* resistance gene cassette. The targeting fragment is present upstream of *eGFP* and includes the EP1 core promoter. The

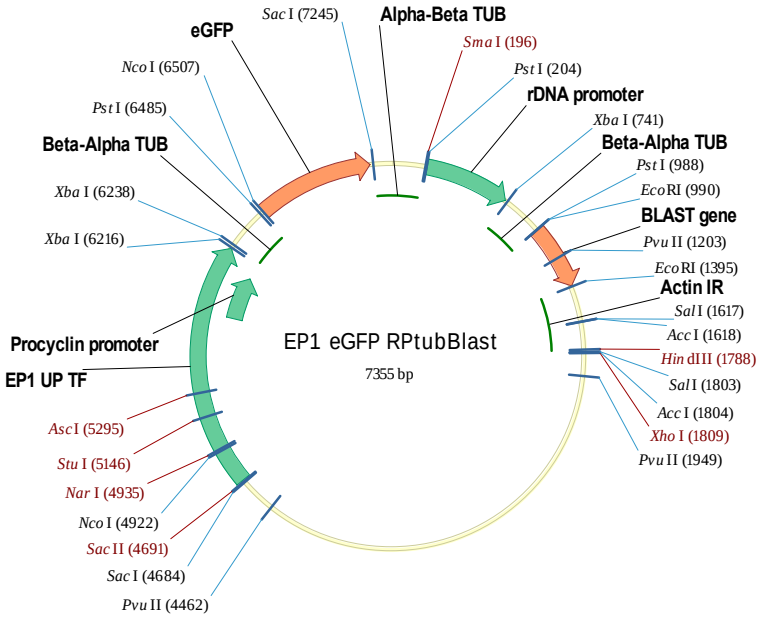
linearised (indicated with scissors) construct integrates into the procyclin locus by a single-crossover homologous recombination event.

In order to understand the role of TbISWI in transcriptional repression, another transcriptionally inactive genomic locus, the procyclin locus in BF trypanosomes, was tested for possible derepression after TbISWI depletion. The experimental strategy used in Chapter 3 above for the minichromosomes and the VSG basic copy array was adapted for this purpose (Figure 4.2.1). A construct containing an *eGFP* reporter gene was inserted downstream of the endogenous EP1 promoter in the EP1 procyclin locus and its expression monitored after RNAi-mediated TbISWI depletion. The same –ESp reporter backbone construct, peGFP_RPtubBlast, was modified to target this construct into the procyclin locus (Figure 4.2.2).

Results

As with the other –ESp reporter constructs, the EP1 procyclin targeted eGFP reporter construct, EP1eGFP_RPtubBlast, was made by inserting the PCR amplified ~1.5 kb targeting fragment (see Materials and Methods) into the NotI linearised –ESp eGFP reporter backbone as detailed in the Materials and Methods. The construct obtained, EP1eGFP_RPtubBlast, was verified by diagnostic digests (Figure 4.2.2). Since the targeting fragment contained the EP1 promoter within it, this strategy results in the duplication of the EP1 promoter (see Figure 4.2.1). In order to obtain the eGFP reporter cell lines, bloodstream form (BF) S16.221 and procyclic form (PF) 29-13 trypanosomes were transfected with the eGFP reporter construct. Both of these cell lines encode constitutively expressed Tet repressor and T7 RNA polymerase genes, which allow for tetracycline inducible RNAi (Wirtz et al. 1999).

Targeting of the construct was achieved by linearising the reporter construct with AscI restriction in the middle of the procyclin targeting sequence, generating a double strand break inside a homologous region to facilitate recombination (Clayton 1999). This strategy results in the integration of the reporter construct by a single crossover event that also integrates the Bluescript vector backbone into the procyclin locus.



STEPS:

1. The EP1 Targeting Fragment was PCR amplified with Pwo polymerase by using primers EP1_UP_2480s & EP1_UP_3991as and S16 genomic DNA as template. The fragment was then phophorylated using T4 polynucleotide kinase.
2. peGFP_RPtubBLAST was linearised with NotI (+blunted) and the blunt-ended EP1 procyclin Targeting Fragment from Step 1 was ligated into it. Correct ligation of the construct in the forward orientation was verified by PCRs linking over the junctions and restriction digestions of the obtained construct. This construct was called EP1_eGFP_RPtubBLAST.
3. EP1_eGFP_RPtubBLAST was linearised with AscI for targeting into the trypanosome genome.

Figure 4.2.2 Map of the EP1 procyclin locus targeting reporter construct. EP1eGFP_RPtubBlast was derived from the –ESp reporter backbone of peGFP_RPtubBlast by inserting into it the EP1 targeting fragment. The construct was linearised for transfection by AscI digestion.

PCR reactions using genomic DNA from the reporter cell lines as template were used to verify correct targeting of the reporter constructs. The reporter construct could be linked to the genomic locus in all the cell lines tested. Figure 4.2.3 shows a schematic of the primers used and the PCR products for four representative cell lines. After verification of the correct integration of the reporter constructs, reporter cell lines BF_Pro

and PF_Pro were transfected with the MC[177] TbISWI-A RNAi construct (Hughes et al. 2007) and two RNAi cell lines per reporter line were obtained. The RNAi cell lines obtained were called BF_Pro1, BF_Pro2, PF_Pro1 and PF_Pro2.

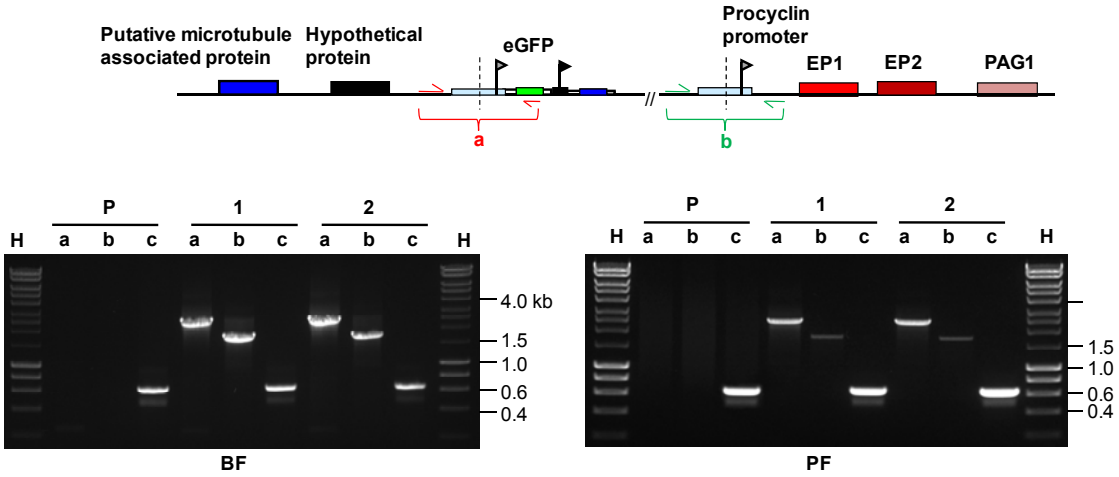


Figure 4.2.3 Integration of the reporter eGFP construct into the EP1 procyclin locus. PCR amplification using genomic DNA isolated from bloodstream (BF) and procyclic (PF) form *T. brucei* EP1 procyclin reporter transformants was used to confirm integration into the correct genomic locus. The top panel shows a schematic of primer sets used for the purpose. Primer sets 'a' and 'b' were designed to amplify over the upstream and the downstream reporter construct integration points (indicated by dotted vertical lines), respectively. Primer set 'c' amplified a segment of the RNA polymerase I large subunit, and was used as a positive control for PCR amplification from the genomic DNA. The bottom panel shows the products of the PCR reactions using genomic DNA from various bloodstream and procyclic form cell lines, with the size markers indicated on the right of the gel pictures. For the BF lanes, P:221Pur_A3, 1:BF_Pro, 2:BF_Pro^{b}. For the PF lanes, P:29-13, 1:PF_Pro, 2:PF_Pro^{b}. Primer sets used: 'a' (EP1-2082s & eGFP-183as), 'b' (T3 & EP1-4177as) and 'c' (Pol I-484s & Pol I-1123as).

In order to investigate the role of TbISWI in transcriptional repression of the procyclin locus, the RNAi cell lines were grown in the presence of 750 ng ml^{-1} tetracycline and fluorescence was measured in the FL1 channel using flow cytometry as described in the Materials and Methods. Figures 4.2.4 shows representative FACS traces of eGFP fluorescence of the cell lines in the top panel, mean FL1 fluorescence values in the middle panel, and the fold derepression of the eGFP gene, expressed as the ratio of fluorescence of induced versus uninduced cells, at different durations of induction in the bottom panel. The depletion of the TbISWI protein can be seen in Figure 4.2.5.

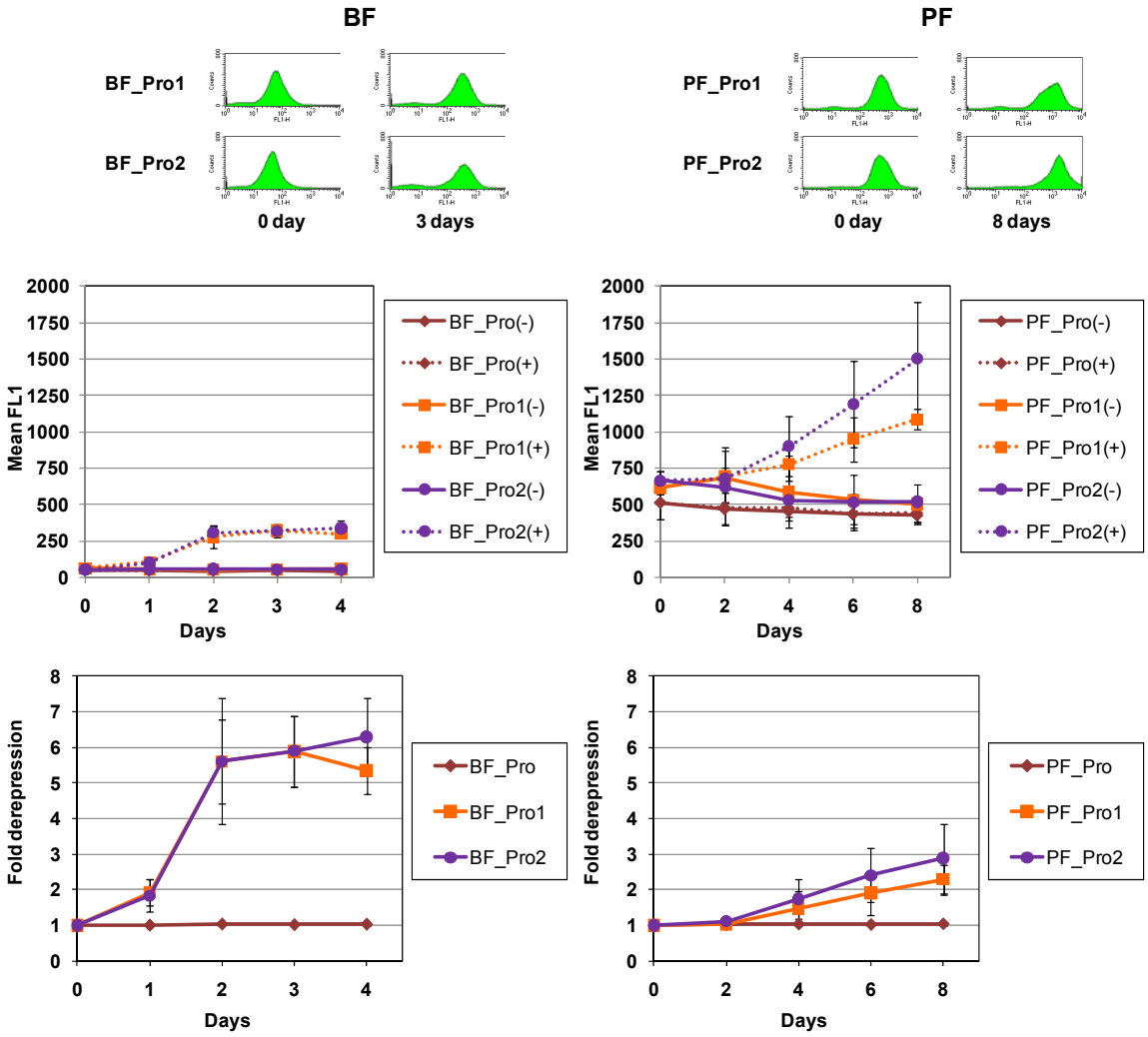


Figure 4.2.4 Monitoring eGFP derepression after induction of RNAi against *TbISWI* in bloodstream and procyclic form EP1 procyclin reporter cell lines. Levels of expression of the *eGFP* gene were measured by flow cytometry for both the bloodstream form (left panel) and the procyclic form (right panel) reporter lines. Top: representative flow cytometry traces for two cell lines at two different time-points are shown (day 0 and day 3 for BF; day 0 and day 8 for PF), with eGFP fluorescence monitored in the FL1 channel (x-axis). Middle: the graph shows the mean fluorescence (FL1 channel) of eGFP in tetracycline induced (+) and uninduced (-) cell lines for different durations of induction; 100,000 cells were analysed per sample. Bottom: the graph shows the degree of derepression of the *eGFP* gene after different durations of tetracycline induced *TbISWI* RNAi. Fold derepression was calculated as the ratio of mean fluorescence of induced cells to that of the uninduced cells. The graph shows the average of three independent experiments with the error bars indicating the standard deviation. The BF_Pro1 and BF_Pro2 cell lines were derived from BF_Pro by transfection with the TbISWI_RNAi construct. The PF_Pro1 and PF_Pro2 cell lines were derived from PF_Pro by transfection with the TbISWI_RNAi construct.

In the BF RNAi cell lines, TbISWI depletion resulted in ~6 fold derepression of the eGFP reporter gene whereas ~3 fold derepression was seen in the PF cell lines (Figure 4.2.4). Mean fluorescence observed in the FL1 channel increased during the course of the

induction of *TbISWI* RNAi. Western blot analysis confirmed that TbISWI was highly depleted after 48 hours of tetracycline induction in the BF cells and by 4 days in the PF cells (Figure 4.2.5).

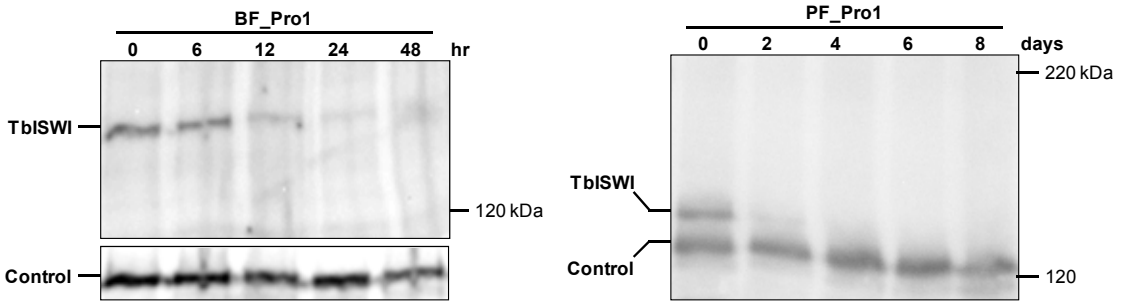


Figure 4.2.5 Depletion of TbISWI in bloodstream and procyclic form EP1 procyclin reporter cell lines after induction of RNAi. Protein lysates from BF (left) cell lines were prepared after the induction of RNAi against *TbISWI* with tetracycline for 0, 6, 12, 24 and 48 hours. TbISWI and an SMC3 control band are indicated on the left and the size markers on the right. Protein lysates from PF (right) cell lines were prepared after the induction of RNAi against *TbISWI* with tetracycline for 0, 2, 4, 6 and 8 days. TbISWI and a cross-reactive control band are indicated on the left and the size markers on the right.

Discussion

These results suggest that TbISWI plays a role in the repression of the procyclin locus in both BF and PF trypanosomes, its depletion resulting in increased transcription from the eGFP reporter construct. Procyclin genes are known to be transcriptionally downregulated in BF cells by about 10 fold (Hotz et al. 1998). Therefore, it was not unexpected to find the basal mean FL1 values to be lower in the BF reporter line (~50 units) by about 9 fold compared to the PF reporter line (~450 units). Compared to the PF cells, lower basal fluorescence in the BF cells had also been observed using reporter constructs in the minichromosome, the VSG basic copy array and the rDNA spacer, although the differences were not as drastic. In addition, lower fluorescence of the BF reporter lines at the other silent loci was observed using constructs that lacked a promoter. In the case of the procyclin locus, this difference was seen despite the presence of the endogenous EP1 procyclin promoter. It is also noteworthy that even when maximally

derepressed by about 6 fold after TbISWI depletion, the BF eGFP fluorescence (~350 units) remained lower than the basal level of fluorescence (~450 units) in the PF. These observations are presumably a consequence of the need to downregulate the procyclin gene in the BF trypanosomes, where exposure of an invariant surface molecule can be hazardous for the trypanosomes.

In addition, it is surprising that there is some derepression, although less than 3 fold, of the procyclin locus in the PF cells. This suggests that the procyclin locus is not active to its full potential in the native state and that TbISWI plays a role in keeping it silent. As suggested for ES regulation (Hughes et al. 2007), there may be multiple layers of transcriptional control operating on the procyclin locus too, and other downregulating mechanisms may be more active in the BF cells in comparison with the PF cells.

In this case, comparisons between the derepression of the procyclin locus and that observed from the other +ESp reporter cell lines from previous experiments must be made with caution because the EP1 promoter is very different from the ES promoter. While the ES promoter lacks any recognisable transcription factor (TF) binding or chromatin remodelling elements, the procyclin and the rDNA promoters contain both those elements (Zomerdijk et al. 1990; Vanhamme et al. 1995; Pham et al. 1996). These differences possibly render the ES promoter more sensitive to its genomic context (Horn et al. 1997).

4.2. Tubulin locus

Introduction

While studying transcription from the procyclin locus, it was found that even in the PF cells where procyclin is expressed abundantly in the native state, TbISWI depletion led to increased procyclin expression. This suggested that TbISWI may also be downregulating other actively expressed genomic loci. All other genomic regions studied earlier were either promoter-less or pol-I promoter proximal regions.

An important feature of the trypanosome genome includes the polycistronic organisation of large numbers of RNA pol-II transcribed constitutively expressed gene clusters on the large chromosomes (Clayton 2002). Some housekeeping genes are found to be duplicated in these gene clusters, presumably to allow for increased expression levels, as is the case with some tandemly arrayed genes like the α/β tubulin genes (Eid et al. 1991; Jackson 2007).

In order to understand the role of TbISWI in transcriptional regulation of a constitutively active pol-II transcription unit, an experiment was planned to test for changes in transcription of the α/β tubulin array after TbISWI depletion. The experimental strategy used in Chapter 3 above for the minichromosomes and the VSG basic copy array was adapted for this purpose (Figure 4.3.1). An eGFP reporter gene was inserted inside the tubulin array, targeted to β-tubulin, and its expression monitored after RNAi-mediated TbISWI depletion.

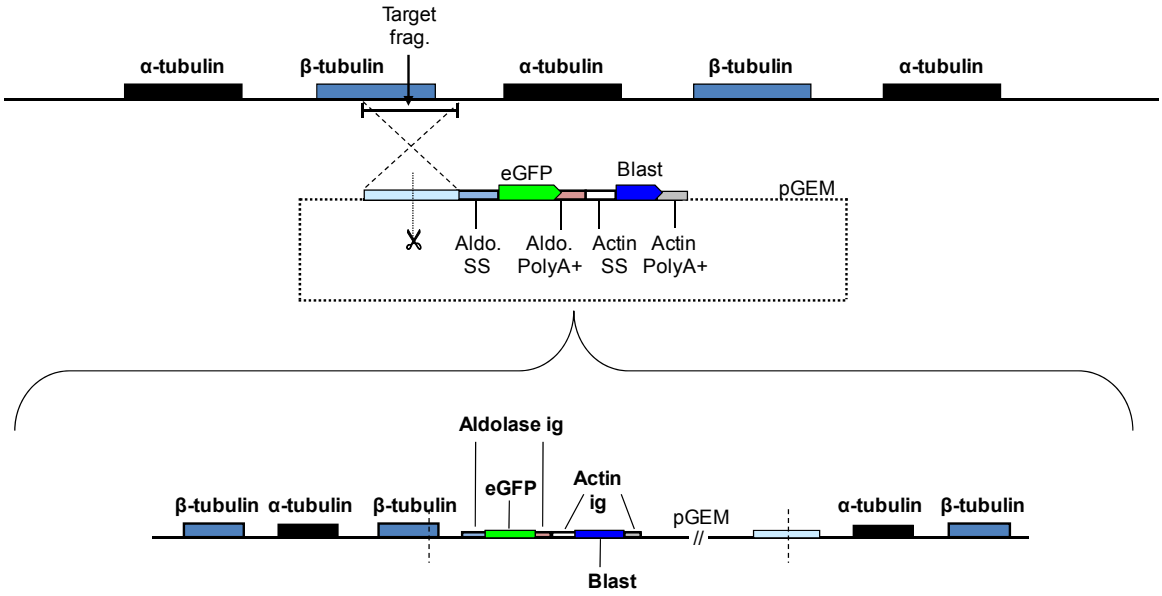
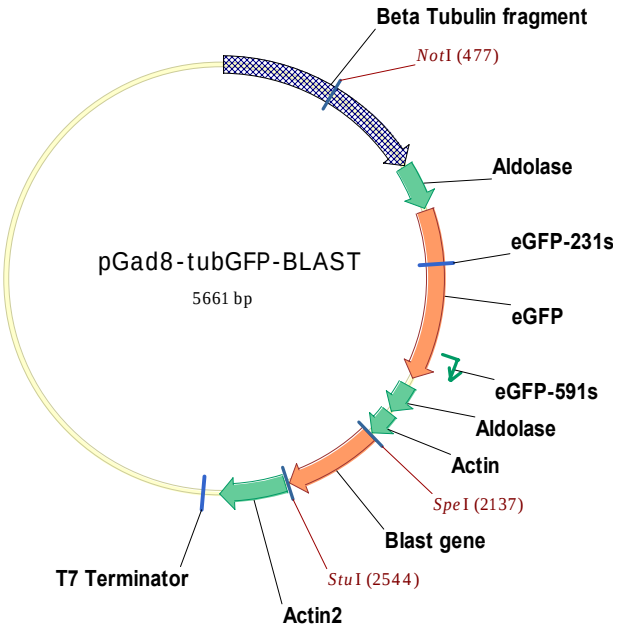


Figure 4.3.1 Strategy for reporter *eGFP* integration into the tubulin locus. The schematic shows integration of an *eGFP* reporter gene into the β-tubulin locus. In the reporter construct, the *eGFP* gene with flanking transcript processing signals (aldolase intergenic regions) is followed downstream by a *Blasticidin* resistance gene cassette. The targeting sequence consists of a β-tubulin fragment present upstream of the *eGFP*. The linearised (indicated by the scissors) construct integrates into the tubulin array by a single-crossover homologous recombination event.

Results

The β-tubulin locus targeted eGFP reporter construct, pGad8-tubGFP-Blast, was made by replacing the Hygro^r gene in pGad8-tubulin (Wickstead et al. 2003) with a Blasticidin resistance gene as detailed in Figure 4.3.2, and the construct was verified by diagnostic digests. Targeted insertion of the construct was achieved by linearising the pGad8-tubGFP-Blast construct by NotI restriction in the middle of the β-tubulin targeting sequence, generating a double strand break inside a homologous region to facilitate recombination (Clayton 1999). This strategy results in the integration of the reporter construct by a single crossover event that also integrates the pGEM vector backbone into the tubulin locus. The BF S16.221 cell line, a derivative of the S16 single marker cell line, and the PF 29-13 cell lines were used for transfection of the tubGFP reporter construct (Wirtz et al. 1999). Both of these cell lines encode constitutively expressed Tet

repressor and T7 RNA polymerase genes, which allow for the induction of tetracycline inducible RNAi.



STEPS:

1. pGad8-tubulin (Wickstead et al, 2003) was cut with StuI + SpeI to excise out the HygroR gene. The construct backbone was isolated.
2. Blasticidine resistance gene was amplified with Pfu polymerase by using primers BSR1s-SpeI & BSR399as-StuI and MC_eGFP_RPtubBLAST construct as template, and digested with StuI + SpeI.
3. The sticky-ended BlasticidineR gene from Step 2 was ligated into the sticky-ended pGad8-tubulin backbone from Step 1 and verified by PCRs linking over the junctions and restriction digestions of the obtained construct. The construct was called pGad8-tubGFP-BLAST.
4. pGad8-tubGFP-BLAST was linearised with NotI for targeting into the trypanosome genome.

Figure 4.3.2 Map of the tubulin locus targeting reporter construct. pGad8_tubGFP_Blast was derived from pGad8-tubulin (Wickstead et al. 2003) by replacing the Hygromycin resistance gene with a Blasticidin resistance gene. The construct was linearised for transfection by NotI digestion.

In order to verify correct targeting of the reporter constructs, genomic DNA was isolated from the reporter cell lines and used as a template for PCR. The reporter construct could be linked to the tubulin locus in all the cell lines tested. Figure 4.3.3 shows a schematic of the primers used and the PCR products for four representative cell lines. After verification of successful reporter construct integration into the procyclin locus, the reporter cell lines- BF_tub and PF_tub- were transfected with the MC[177] TbISWI-A RNAi construct (Hughes et al. 2007). Two RNAi cell lines were

obtained from each of these reporter lines and used for the planned experiments. The RNAi cell lines were called BF_tub1, BF_tub2, PF_tub1 and PF_tub2.

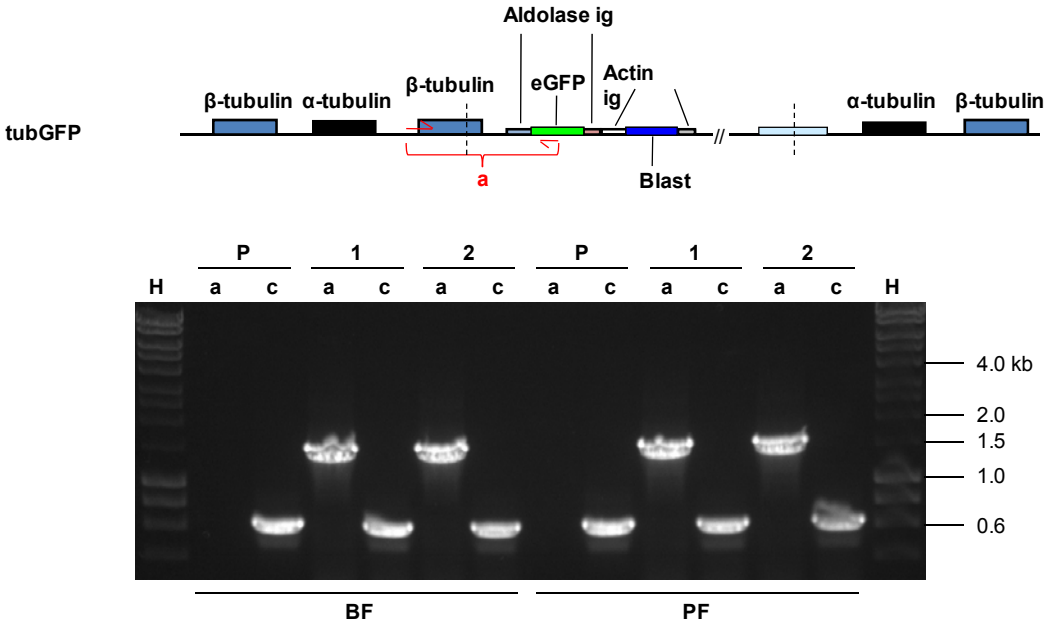


Figure 4.3.3 Integration of the reporter eGFP construct into the tubulin locus. PCR amplification using genomic DNA isolated from bloodstream (BF) and procyclic (PF) form *T. brucei* tubulin reporter transformants was used to confirm integration into the correct genomic locus. The top panel shows a schematic of primer sets used for this purpose. Primer set 'a' was designed to amplify over the upstream reporter construct integration point (indicated by dotted vertical lines). Primer set 'c' amplified a segment of the RNA polymerase I large subunit, and was used as a positive control for PCR amplification from the genomic DNA. The bottom panel shows the products of the PCR reactions using genomic DNA from various bloodstream and procyclic form cell lines, with the size markers indicated on the right. For the BF lanes, P:221Pur_A3, 1:BF_tub, 2:BF_tub^{b}. For the PF lanes, P:29-13, 1:PF_tub, 2:PF_tub^{b}. Primer sets used: 'a' (Btub-261s & eGFP-183as) and 'c' (Pol I-484s & Pol I-1123as).

As before, the RNAi cell lines were induced with tetracycline and the degree of fluorescence measured in the FL1 channel using flow cytometry. Figures 4.3.4 shows representative FACS traces of eGFP fluorescence of the cell lines in the top panel, mean FL1 fluorescence values for both induced and uninduced cells in the middle panel, and the degree of derepression of the eGFP gene, expressed as the ratio of fluorescence of induced versus uninduced cells, at different durations of induction of TbISWI RNAi in the bottom panel. The depletion of the TbISWI protein can be seen in Figure 4.3.5.

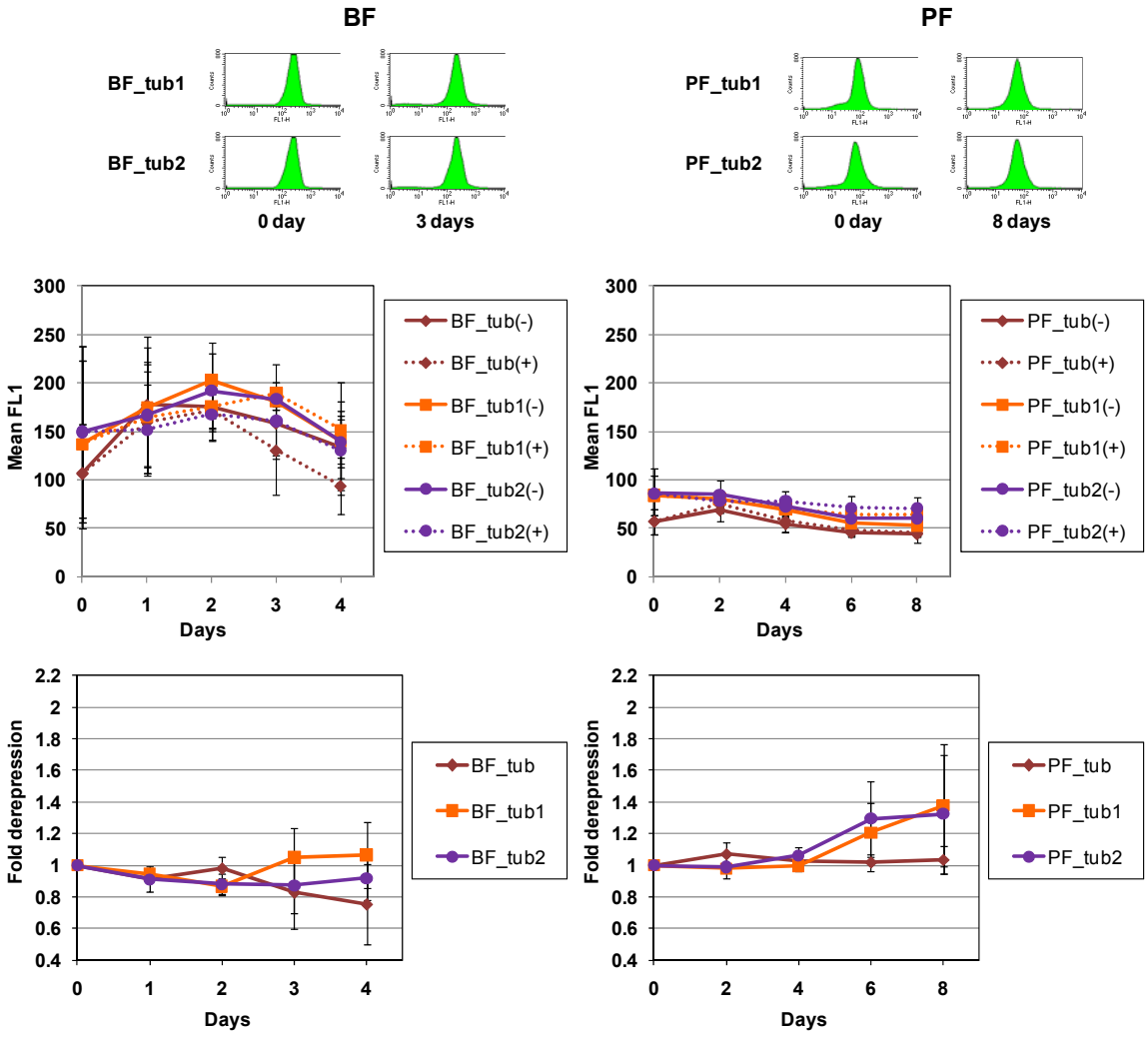


Figure 4.3.4 Derepression of eGFP after induction of RNAi against *TbISWI* in bloodstream and procyclic form tubulin reporter cell lines. Levels of expression of the *eGFP* gene were measured by flow cytometry for both the bloodstream form (left panel) and the procyclic form (right panel) reporter lines. Top: representative flow cytometry traces for two cell lines at two different time-points are shown (day 0 and day 3 for BF; day 0 and day 8 for PF), with eGFP fluorescence monitored in the FL1 channel (x-axis). Middle: the graph shows the mean fluorescence (FL1 channel) of eGFP in tetracycline induced (+) and uninduced (-) cell lines for different durations of induction; 100,000 cells were analysed per sample. Bottom: the graph shows the degree of derepression of the *eGFP* gene for different durations of tetracycline induced TbISWI RNAi. The degree of derepression was calculated as the ratio of mean fluorescence of induced cells to that of the uninduced cells. The graph shows the average of the results of three independent experiments with the error bars indicating the standard deviation. BF_tub1 and BF_tub2 were derived from BF_tub by transfection with the TbISWI_RNAi construct. PF_tub1 and PF_tub2 were derived from PF_tub by transfection with the TbISWI_RNAi construct.

In both the BF and the PF RNAi cell lines, TbISWI depletion resulted in no significant derepression of the eGFP reporter gene within the tubulin array, although there was some insignificant derepression seen in the PF cell lines (Figure 4.3.4). The

fluorescence observed in the FL1 channel remained unchanged during the course of *TbISWI* RNAi. Western blot analysis confirmed that TbISWI had become almost undetectable after 48 hours of tetracycline induction in the BF cells and after 4 days in the PF cells (Figure 4.3.5).

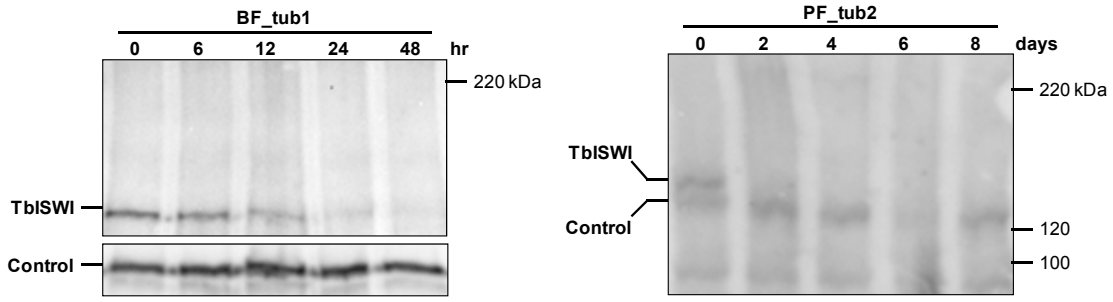


Figure 4.3.5 Depletion of TbISWI in Bloodstream and Procyclic form tubulin reporter cell lines after induction of RNAi. (A) Protein lysates from BF (left) cell lines were prepared after the induction of RNAi against *TbISWI* with tetracycline for 0, 6, 12, 24 and 48 hours. TbISWI and an SMC3 control band are indicated on the left and the size markers on the right of the Western blot. (B) Protein lysates from PF (right) cell lines were prepared after the induction of RNAi against *TbISWI* with tetracycline for 0, 2, 4, 6 and 8 days. TbISWI and a cross-reactive control band are indicated on the left and the size markers on the right of the Western blot.

Discussion

These results suggest that TbISWI does not play a role in the downregulation of a constitutively active and pol-II transcribed tubulin array in either BF or PF *T. brucei*. It is possible this locus is transcribed at its maximal potential constitutively, without any downregulation. This serves as a control for reporter expression from other genomic loci, arguing that the derepression seen in those regions after TbISWI depletion is significant.

Unlike other genomic loci tested thus far- the minichromosomes, the VSG basic copy array, the rDNA spacer and the procyclin locus- the basal mean FL1 values were found to be higher in the BF reporter line (~100 units) as compared to the PF reporter line (~50 units). This could show that there is a higher level of tubulin expression in BF trypanosomes. Alternatively, and more likely, this could be a result of the aldolase

transcript processing signals that flank the GFP reporter gene. Aldolase mRNA is 6 fold more abundant in BF compared with PF trypanosomes, an effect of its 3'UTR that is highly stable in the bloodstream form cells (Clayton 1985; Hug et al. 1993). The use of aldolase processing signals for the *eGFP* gene inserted in the β-tubulin locus also serves as a complication while comparing expression from that from the eGFP reporter constructs, which have tubulin processing signals for the *eGFP* gene. The difference in processing signals could cause the transcripts to have different stabilities and hence invalidate fluorescence measurement for comparing transcript abundance across the different reporter cell lines used in these studies.

Chapter 5

Role of TbISWI in chromosome segregation

5.1. Cell cycle effects

Introduction

As discussed in Chapter 1, chromatin remodelling complexes have a wide variety of functions in eukaryotic cells. They are involved not only in transcription regulation (initiation, elongation and termination), but also in other cellular functions like DNA replication and repair, kinetochore and centromere construction, chromosome segregation and telomere maintenance (Saha et al. 2006). SWI2/SNF2 related chromatin remodelling complexes play important roles in cell cycle regulation (Cao et al. 1997; Muchardt et al. 1999), both by maintaining centromeric heterochromatin for kinetochore attachment (Grewal et al. 2007), as well as by loading the cohesin complex to hold the sister chromatids together (Hakimi et al. 2002).

Recent evidence suggests that cohesin is also involved in maintaining VSG monoallelic expression in *T. brucei* (Landeira et al. 2009). Given the role of TbISWI in the downregulation of silent VSG ESs (Hughes et al. 2007), and other areas of the genome (as seen in Chapters 3 and 4), the effect of TbISWI depletion on cell cycle progression was monitored by counting the number of nuclei and kinetoplasts in treated and untreated cells. This was achieved by DAPI (4',6-diamidino-2-phenylindole) and PI (propidium iodide) staining of the DNA of these organelles in both the bloodstream and procyclic form cells. While the BF cells were induced with tetracycline for 2 days, the PF cells were induced for 4 days.

Results

Figure 5.1.1 shows data from the quantification of the DAPI stained cells where the percentage of cells with different nuclei/ kinetoplast configuration (400-500 cells

counted per sample) was determined in both the induced and the uninduced populations of BF/ PF *T. brucei*. Figure 5.1.2 shows data from the flow cytometry based quantification of cellular DNA content from both the induced and the uninduced BF/ PF cells (50,000 cells per sample analysed by FACS).

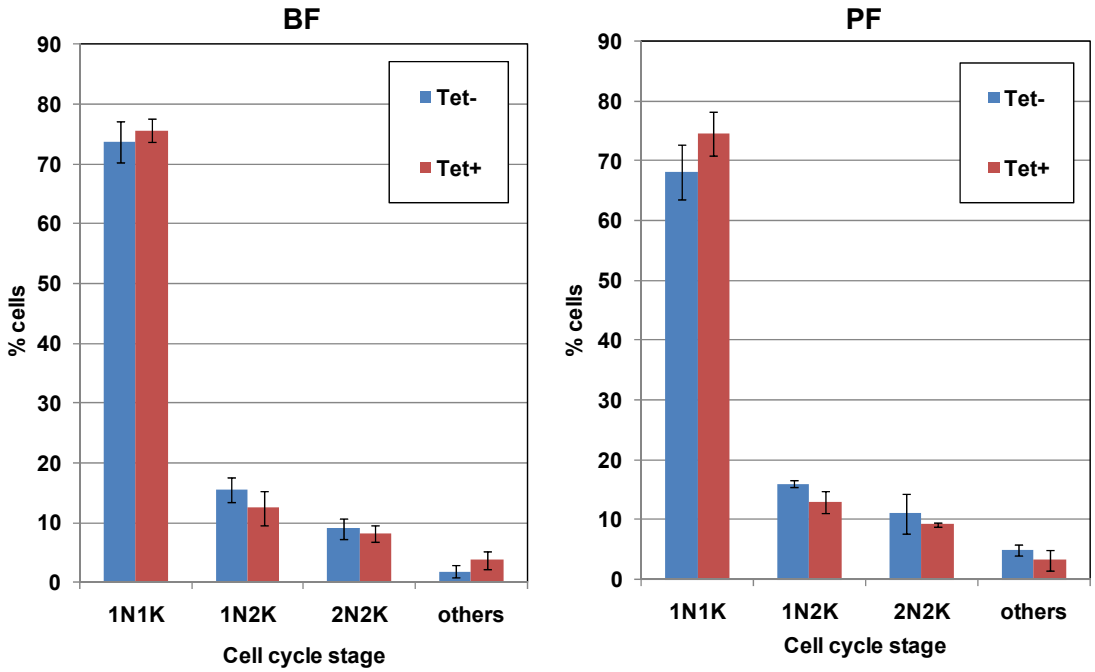


Figure 5.1.1 Effect of TbISWI depletion on progression through the *T. brucei* cell cycle. Microscopic analysis of DAPI stained *T. brucei* cells was performed after induction of *TbISWI* RNAi. Bloodstream form (BF) cells were induced for 2 days and procyclic cells (PF) for 4 days. Cells were categorised according to the number of nuclei (N) or kinetoplasts (K) present and the percentage of cells in the different categories was plotted for both the induced (Tet+) and the induced (Tet-) cells. The average of three independent experiments is plotted with standard deviation indicated by error bars. Cell lines used for the experiment were *T. brucei* T3-SA1 (BF) and D3-SA1 (PF) (Hughes et al. 2007).

Comparing the results of the induced versus uninduced cells, there does not appear to be a significant effect of TbISWI knockdown on cell cycle progression.

Discussion

TbISWI is involved in transcriptional repression of silent regions of the genome (as seen in Chapters 3 and 4). It binds ubiquitously throughout the genome and is

particularly enriched in BF centromeric CIR147 repeats as compared to the PF (Tara Stanne's unpublished results). It could therefore have a role in the maintenance of heterochromatin at the centromeres and its knockdown might impair kinetochore attachment and, in turn, chromosome segregation. Defects in chromosome segregation would be detectable in the form of accumulation of cells at particular cell cycle stages.

Observations from the experiments involving DAPI and PI staining suggest that both in the BF and the PF cells TbISWI depletion does not lead to a significant change in the distribution of cells in the cell cycle stages analysed. Therefore, it does not appear to have a role in maintaining centromeric heterochromatin.

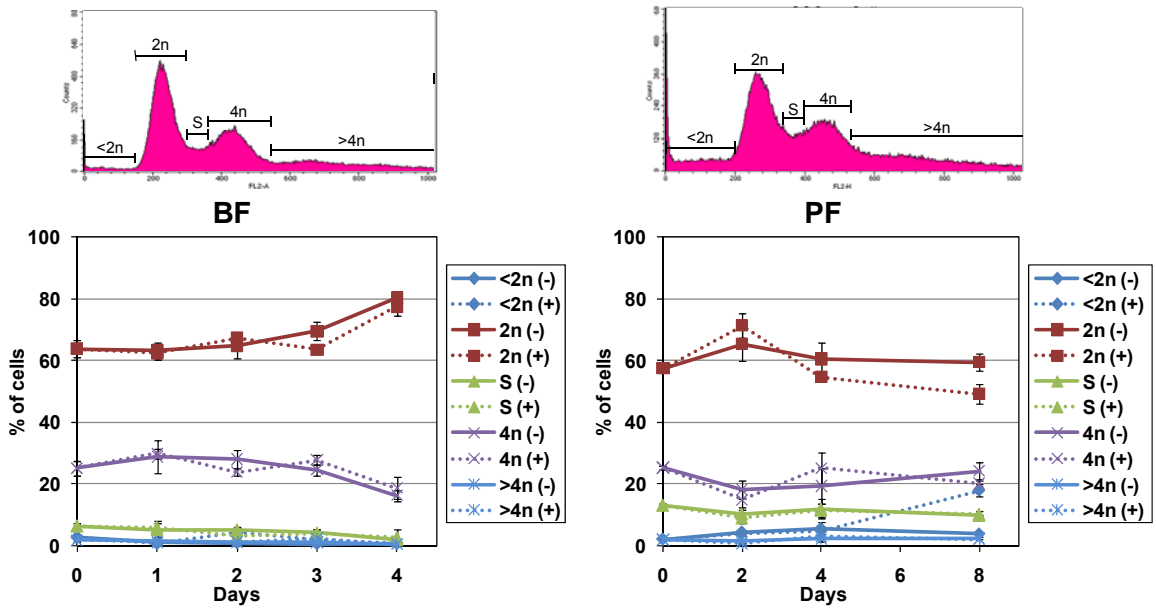


Figure 5.1.2 Monitoring DNA content of *T. brucei* cells after the induction of *TbISWI* RNAi. *T. brucei* cells were stained with propidium iodide (PI) and were then analysed using flow cytometry for their DNA content. The analysis was performed over 4 days of tetracycline induction for bloodstream form (BF) cells and over 8 days of induction for procyclic form (PF) cells. The top panel shows representative flow cytometry traces, with PI fluorescence (corresponding to the DNA content) monitored in the FL2 channel (x-axis). The gates indicate different categories of DNA content, with 'n' representing haploid DNA amount and S representing the DNA synthesis stage of cell cycle. The plots in the lower panel show a change in the percentage distribution of DNA content over the duration of induction for both the induced (+) and the uninduced (-) cell populations. The average of three independent experiments is plotted with standard deviation indicated by error bars. Cell lines used for the experiment were *T. brucei* T3-SA1 (BF) and D3-SA1 (PF) (Hughes et al. 2007).

While there was no accumulation of cells at any particular cell cycle stage after TbISWI depletion, the role of TbISWI in silencing minichromosomes (as seen in Chapter 3) suggests that TbISWI depletion may still affect their sub-nuclear localisation and/ or segregation (impair/ delay). This issue will be explored further in the next section.

5.2. Minichromosome segregation

Introduction

T. brucei has about hundred minichromosomes (50-150 kb each) that make up about 10% of the trypanosome nuclear genome and are predominantly composed of transcriptionally silent 177 bp repeats, with a silent VSG gene at the telomeres (Wickstead et al. 2004). Although the minichromosomes contain only a small fraction of the total silent VSG repertoire, they are the preferred donors for VSG switching events early in parasitemia (Robinson et al. 1999). Despite lacking any recognisable centromeric sequences, they are faithfully segregated to the daughter cells during mitosis possibly by laterally stacking onto spindle microtubules (Ersfeld et al. 1997). Studies using fluorescent *in situ* hybridisation (FISH) have shown that minichromosomes localise asymmetrically in the interphase nucleus, either as a crescent shaped cluster at the nuclear periphery or distributed into several clusters in one half of the nucleus (Chung et al. 1990).

Transcriptionally silent heterochromatin has been shown to be tethered to the nuclear envelope in many eukaryotes, while active genes have a specific sub-nuclear localisation (Taddei et al. 2004). In the bloodstream form trypanosomes, the active VSG gene is localised to an extra-nucleolar area of the nucleus called the Expression Site Body (ESB) (Navarro et al. 2001). These observations strongly point to the role of sub-nuclear localisation in determining or maintaining chromatin state of a genetic locus and, in turn, its transcriptional activity. Additionally, recent studies point to the role of chromatin remodelling complexes in centromere function and appropriate mitotic segregation of chromosomes (Hur et al. 2010; Morrison et al. 2009).

In Chapter 3, it was shown that TbISWI depletion results in derepression of a reporter construct integrated into the minichromosomes (Figures 3.1.4 and 3.1.6). In order to investigate whether TbISWI knockdown also results in a relocation and/ or impaired segregation of the minichromosomes, an experiment was designed to visualise them using fluorescently labelled 177 bp repeat probes in FISH experiments, both before and after tetracycline induction.

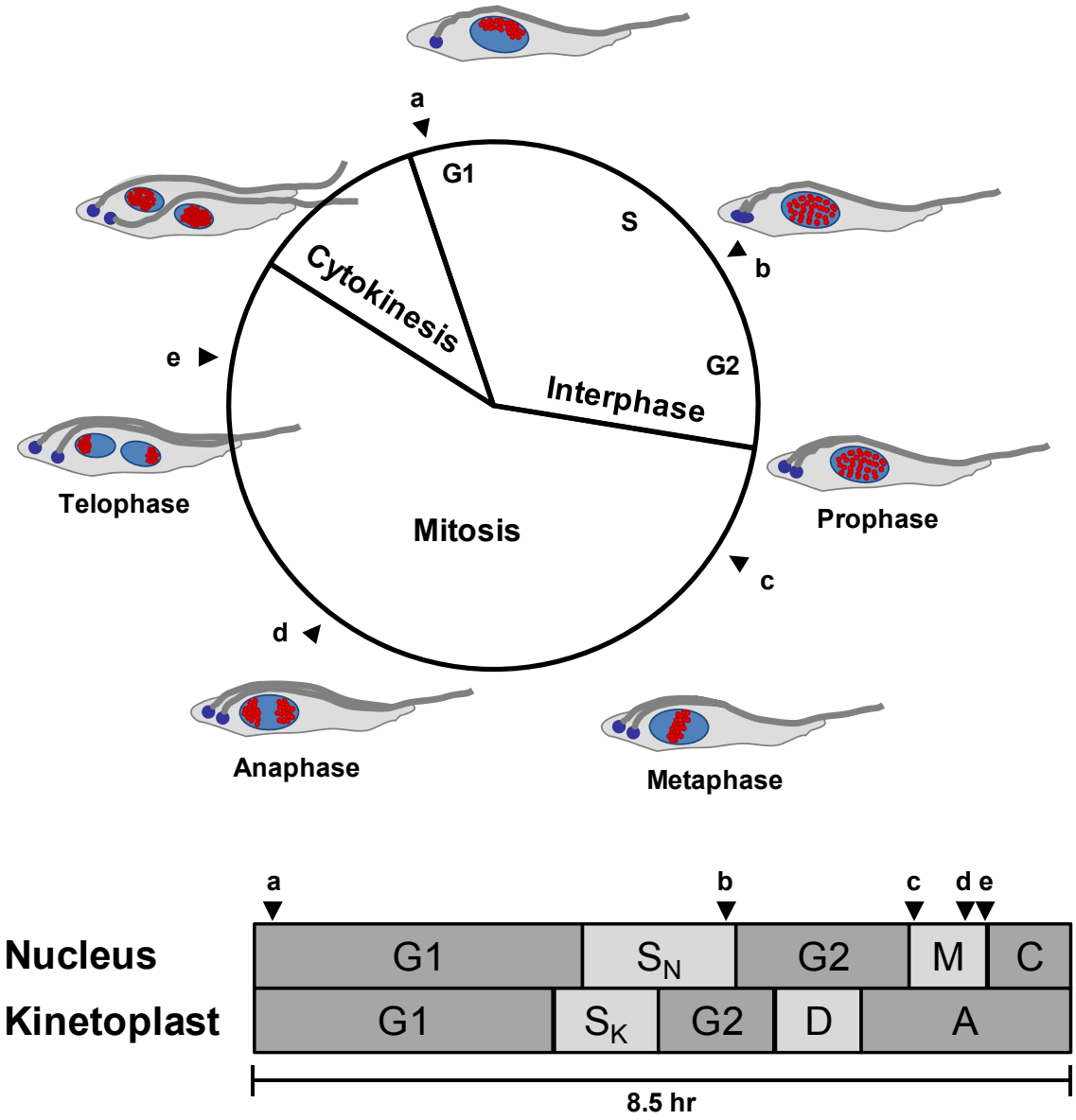


Figure 5.2.1 Minichromosome segregation during the *T. brucei* cell cycle. The upper panel shows a schematic of the *T. brucei* cell cycle (not to scale) with the nucleus in blue and the minichromosomes in red (adapted from Denninger et al. 2010). The kinetoplasts are shown as dark blue dots. The lower panel shows the relative timing (to scale) of the nucleus and kinetoplast replication cycles for procyclic trypanosomes (adapted from McKean 2003). G1 and G2: gap

phases; S_N and S_K: DNA synthesis phases for the nucleus and kinetoplast respectively; M: mitosis; C: cytokinesis; D: kinetoplast division; A: kinetoplast segregation. The approximate positions of the five cell cycle stages (a-e) used for classifying cells in Figures 5.2.2 and 5.2.3 are shown in both upper and lower panels.

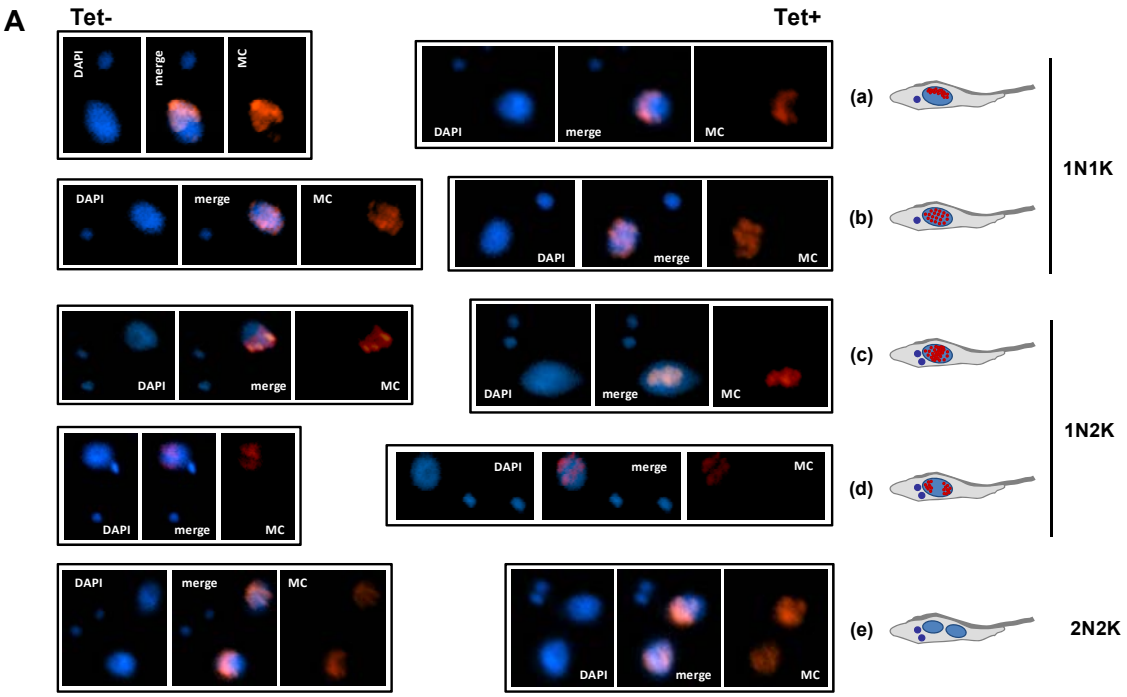
For purposes of quantification, the trypanosome cell cycle modelled in Ersfeld & Gull (1997) is sub-divided into five categories distinguished on the basis of the nucleus/ kinetoplast (NK) ratio and the sub-nuclear localisation of the minichromosomes (Figure 5.2.1)-

a) G1/S phase- characterised by 1N1K configuration and asymmetrical (crescent/ hemispherical) clustering of the minichromosomes,
b) S_N phase- characterised by 1N1K configuration and minichromosomes scattered throughout the nucleus,
c) Prophase/ metaphase- characterised by 1N2K configuration and the minichromosomes either scattered throughout the nucleus or stacked along the metaphase plate,
d) Anaphase- characterised by 1N2K configuration and minichromosomes split away from the metaphase plate, either progressing towards or already close to the polar ends,
e) Telophase/ karyokinesis- characterised by 2N2K configuration.

Results

On the basis of the above criteria, the minichromosomes do not appear to occupy any unusual sub-nuclear positions as a result of TbISWI depletion. Some representative microscopic images are shown in Figures 5.2.2A and 5.2.3A. The NK configuration and the minichromosome localisation allowed the identification of cell cycle stages in both BF and PF cells. As expected, a difference was seen in the organisation of the 2N2K

configuration of the cells in the different life cycle stages. It was possible to identify NKKN configuration in the BF cells, and NKNK in the PF cells, an outcome of differences in cell cycle regulation between the two life cycle stages (McKean 2003).



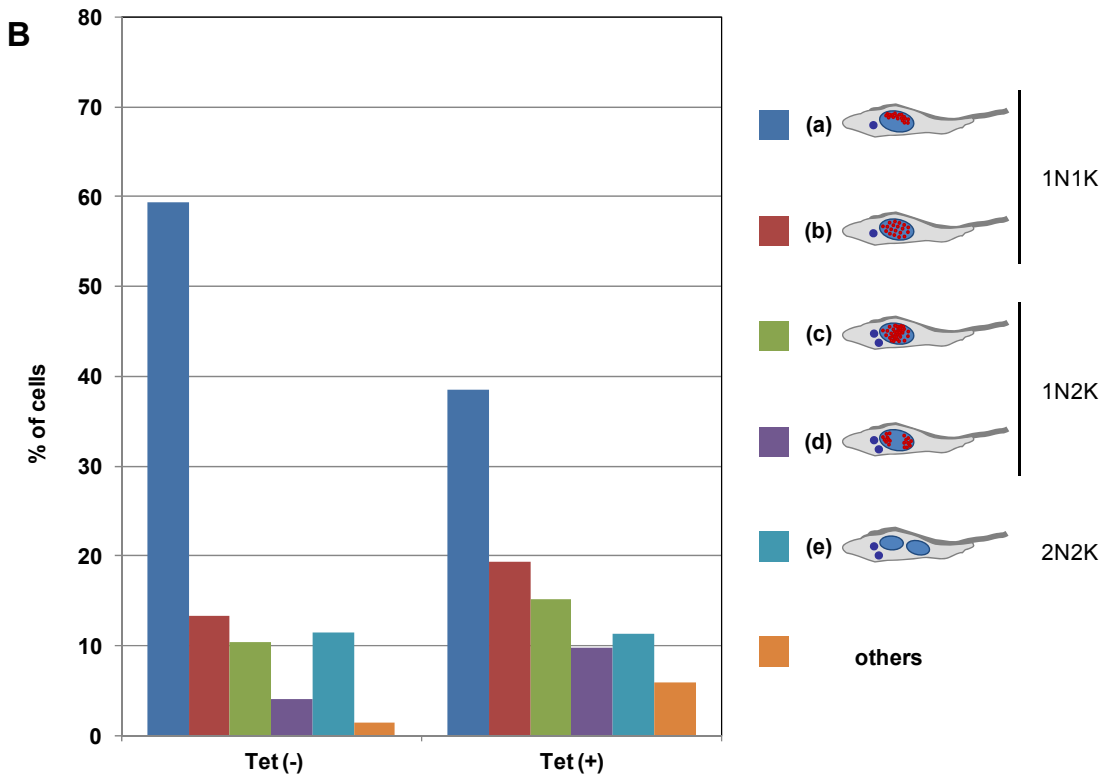


Figure 5.2.2 Sub-nuclear localisation and segregation of minichromosomes in bloodstream form *T. brucei* after TbISWI depletion. Fluorescent *in situ* hybridisation (FISH) was used to visualise *T. brucei* minichromosomes using biotin labelled 177 bp repeat probes. The BF_VBEP2 RNAi cell line was induced with tetracycline for 2 days and the results compared with those from the uninduced cells. Figure (A) shows one cell each from the tetracycline induced (Tet+) and the uninduced (Tet-) population for five different cell cycle stages (a-e). The cell cycle stage is described with a schematic on the right. Red signal represents the minichromosomes while the blue signal shows the DAPI stained DNA. (B) The induced (Tet+) and uninduced (Tet-) cells were classified according to nucleus (N)/ kinetoplast (K) number and the distribution of the minichromosomes and quantified; 'others' include any intact cells that did not fit any of the five cell cycle categories (a-e).

The effect of TbISWI depletion on the segregation of minichromosomes was investigated using fluorescent *in situ* hybridisation (FISH) in both the bloodstream (BF) and procyclic form (PF) cells as detailed in the Materials and Methods chapter. Figures 5.2.2A and 5.2.3A show fluorescence microscopy images that show the DAPI (4',6-diamidino-2-phenylindole) stained DNA in blue and the sub-nuclear distribution of the minichromosomes in red. Five stages in the cell cycle are shown in the figures, with two representative cells each for the different categories of the cell cycle stages. A total of 300-400 cells per slide were counted for both BF and PF cell cycle stages.

Differences were observed between the tetracycline induced and uninduced cells in the BF stage (Figure 5.2.2B), but these differences were always less than 3 fold for every cell cycle stage. While tetracycline induction for two days decreases the percentage of G1/S phase cells to ~0.65 fold compared to the uninduced population, S_N and prophase/ metaphase cells increase to ~1.5 fold and anaphase increases to ~2.4 fold. The percentage of cells in the telophase remained unchanged. The others, including the polynucleate cells and the zoids, increase to ~4 fold. If the 1N1K stages G1/S and S_N are considered together, tetracycline induction results in their decrease to ~0.8 fold. Similarly grouped 1N2K stages prophase/ metaphase and anaphase increase to ~1.7 fold.

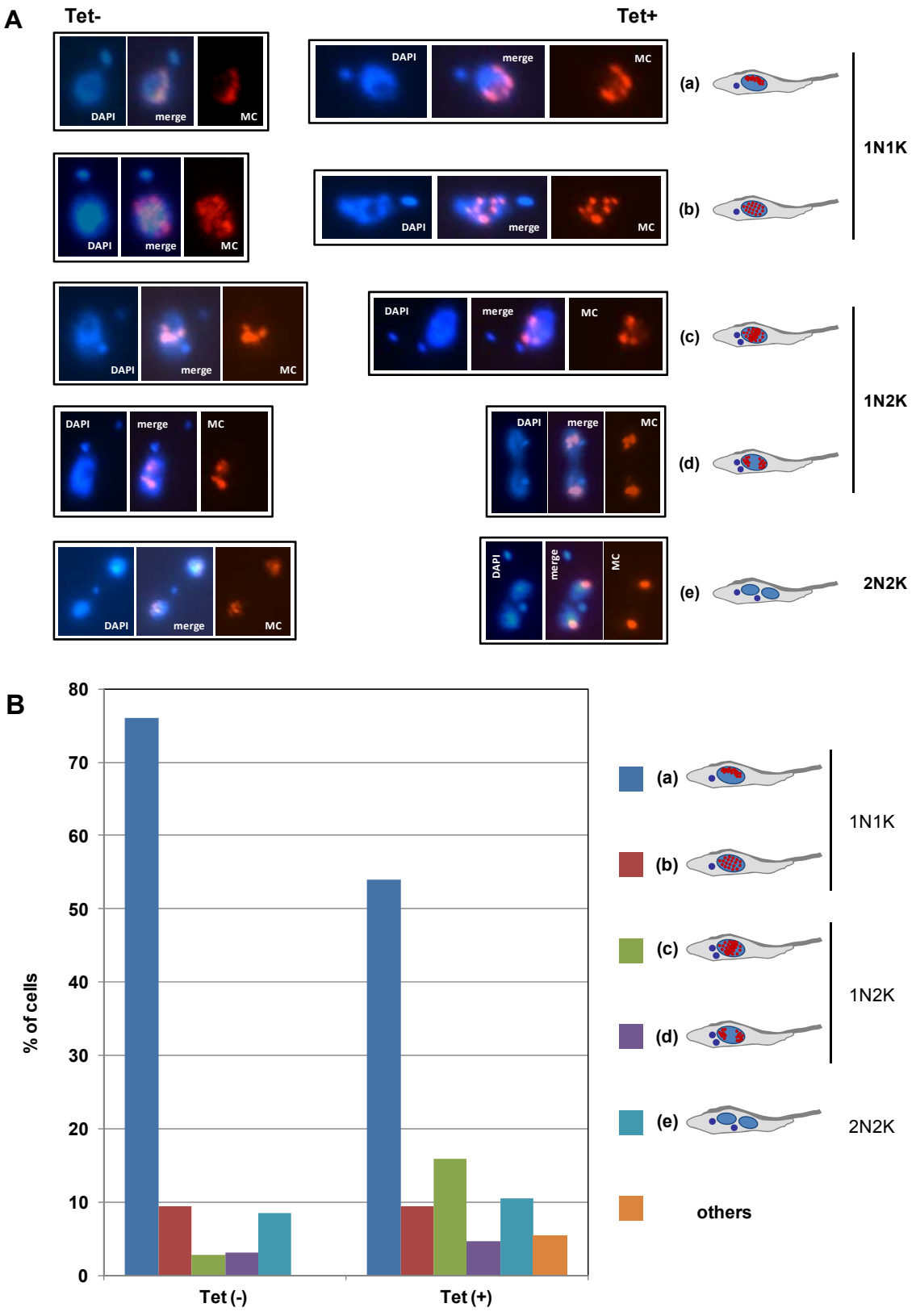


Figure 5.2.3 Sub-nuclear localisation and segregation of minichromosomes in procyclic form *T. brucei* after TbISWI depletion. Fluorescent *in situ* hybridisation (FISH) was used to visualise *T. brucei* minichromosomes using biotin labelled 177 bp repeat probes. The D3-SA1 (Hughes et al. 2007) RNAi cell line was induced with tetracycline for 4 days and the results compared with those of the uninduced cells. Figure (A) shows one cell each from the tetracycline

induced (Tet+) and the uninduced (Tet-) population for five different cell cycle stages (a-e). The cell cycle stage is described with a schematic on the right. Red signal represents the minichromosomes while the blue signal shows the DAPI stained DNA. (B) The induced (Tet+) and uninduced (Tet-) cells were classified according to nucleus (N)/ kinetoplast (K) number and the distribution of the minichromosomes and quantified; 'others' include any intact cells that did not fit any of the five cell cycle categories (a-e).

Differences were observed in the PF cells as well (Figure 5.2.3B). G1/S phase cells were reduced to ~0.7 fold after four days of tetracycline induction, while S_N phase cells remained unaltered in their relative distribution. Prophase/ metaphase cells increased to ~6 fold, while anaphase and telophase showed a modest increase of 1.2-1.4 fold. Other cells that could not be identified as any of the categories 'a'-'e', showed an increase from 0% to 5.4% after tetracycline induction of RNAi. Stages G1/S and S_N considered together as 1N1K show a decrease to ~0.7 fold after tetracycline induction, while stages prophase/ metaphase and anaphase grouped as 1N2K increases to ~3.4 fold.

Discussion

BF cells were induced with tetracycline for 2 days, sufficient to deplete most of the TbISWI protein as seen in Chapters 3 and 4. Tetracycline induction for this duration resulted in a decrease in the G1/S phase to about 65% of the uninduced cells. The S_N phase as well as the prophase/ metaphase cells showed a ~0.5 fold increase, while the anaphase cells showed a ~1.4 fold increase. The small increase (by 0.5x) in the S_N as well as the prophase/ metaphase cells, followed by a slightly higher (by 1.4x) increase in the anaphase cells suggests that TbISWI depletion does not appear to affect minichromosome segregation although it may result in the delaying of telophase a bit. However, these differences are minor and come from a single experiment making their significance difficult to gauge.

PF cells were induced with tetracycline for 4 days before they were analysed for minichromosome segregation/ localisation defects. Based on the observations from Chapters 3 and 4, most TbISWI gets depleted by this time. While the percentage of G1/S phase cells declined to ~70% of the uninduced cells, the prophase/ metaphase cells increased to ~6 fold, followed by modest (~1.2-1.4x) change in anaphase and cytokinesis cells. The results showing the accumulation of prophase/ metaphase cells suggest that TbISWI depletion adversely affects minichromosome segregation. The fact that this effect is visible in PF cells is consistent with the observation that compared to the BF, the PF cells showed a greater increase in fortuitous initiation of transcription of the minichromosomes after TbISWI depletion (Figures 3.1.4 and 3.1.6). However, as these are the results from a single experiment, their significance cannot be adequately estimated.

The percentage of the 1N1K cells appears to decrease and that of the 1N2K cells appears to increase in the FISH experiments as a result of TbISWI depletion, in both life cycle stages (Figures 5.2.2 and 5.2.3). While the significance of these observations is not clear, they are in disagreement with the more significant results seen in the last section (Figures 5.1.1 and 5.1.2) and are therefore unlikely to represent the true distribution of cells in the different life cycle stages on TbISWI depletion.

TbISWI is an essential trypanosome protein whose depletion results in growth arrest in both life cycle stages (Hughes et al. 2007). With the exception of the Pol II transcribed tubulin locus, TbISWI depletion resulted in derepression of all other areas of the *T. brucei* genome investigated in Chapters 3 and 4. Also, the observations from chromatin immunoprecipitation (ChIP) results in Appendix A2 and from those of Tara Stanne (unpublished) suggest that TbISWI binds ubiquitously in the *T. brucei* genome. It

is therefore possible that TbISWI depletion results in loss/ impairment of transcriptional regulation in many parts of the genome affecting different life cycle stages differently. In that respect, the differences in the cell cycle phenotypes between the induced and uninduced cells may be an artefact of the procedure used, which might lead to increased rupturing of cells from one or more cell cycle stages preferentially in the induced population.

Chapter 6

Conclusions

At the outset of this study, TbISWI was recognised as a "novel" ATPase of the ISWI family of SWI2/SNF2 related chromatin remodelling complexes. It was known as an essential nuclear protein involved in transcriptional repression of silent VSG Expression Sites (ES) in both bloodstream (BF) and procyclic form (PF) *T. brucei*, and its depletion was known to cause derepression of silent ES both in BF and in PF cells (Hughes et al. 2007).

Using *in silico* analysis in this study, I have found that TbISWI is very similar to other ISWI ATPases from kinetoplastid protozoa as well as other eukaryotes, its phylogenetic distances from the ISWIs of other organisms closely reflecting the phylogenetic relationship with the organisms investigated. Only one ISWI ATPase per genome was found in all the kinetoplastid protozoa studied, suggesting that it must be similarly essential in all of them. When comparing its domain composition and architecture, I found TbISWI to be very similar to other ISWIs - with a highly conserved SWI2/SNF2 ATPase domain at its N-terminus and the characteristic SANT and SLIDE domains at its C-terminus. Comparing TbISWI C-terminal structure with that of ScISW1 (*S. cerevisiae* ISW1) and DmISWI (*Drosophila melanogaster* ISWI), I found that the length of random coils was greater in TbISWI, possibly a difference that is a consequence of the early evolutionary divergence of *T. brucei*. Greater representation of unstructured coils hints towards a greater fluidity of structure and greater capacity for interaction with many different protein partners (Sandhu 2009).

Also, after comparing the charge distribution on SANT and SLIDE domains in TbISWI, I found it to be unusual in comparison to other eukaryotic ISWIs. By virtue of their positive charge, the SANT domain of ScISW1 and the SLIDE domain of DmISWI are likely to have DNA binding activity (Pinskaya et al. 2009). Conversely, the negatively

charged SLIDE domain of ScISW1 and the SANT domain of DmISWI are likely to interact with nucleosomes. The TbISWI SANT and SLIDE domains were found to have a variegated charge distribution, making it unlikely that there is direct interaction with DNA as well as nucleosomes. With such variegation of charge distribution, it is unclear what role these domains of TbISWI have when interacting with the chromatin. It is possible that *T. brucei* ISWI complex(es) have other protein partners with DNA binding/ nucleosome binding motifs that make up for the lack of such domains on the ATPase moiety itself.

ISWI chromatin-remodelling complexes form the largest and highly abundant family of SWI2/SNF2 related chromatin remodelling complexes (Dirscherl et al. 2004; Deuring et al. 2000). In yeast, an ISWI dependent chromatin structure occurs every 5 kb of the genome (Mellor et al. 2004). The same ISWI ATPase can form multiple ISWI complexes with a different composition of protein partners, each complex taking on different roles. For example, the NURF complex, which activates transcription, and the ACF and CHRAC complexes, which can activate as well as repress transcription, in *Drosophila* contain the same ISWI ATPase (Bouazoune et al. 2006). Similarly, the yeast ISW1 and ISW2 form at least three different ISWI complexes (ISW1a, ISW1b and ISW2) that are involved in regulating transcriptional initiation, elongation as well as termination (Mellor et al. 2004).

All of the above suggest that ISWI complexes have wide-ranging effects on eukaryotic transcription. While these effects can, for reasons of simplicity, be often explained as outcomes of ISWI-driven differential nucleosomal positioning at the genomic locus in question (Cairns 2009), evidence also exists for more direct roles of ISWI in transcription regulation (Mellor et al. 2004). For example, yeast ISW1 represses

PHO8 expression by displacing the TATA-binding protein (TBP) from the promoter (Moreau et al. 2003), an effect that has also been noticed in *Xenopus* (Kikyo et al. 2000). Even if it is difficult to establish whether TBP exclusion or ISWI driven nucleosome remodelling happens first, the overall effect of ISWI on transcriptional regulation can still be established by transcriptional assays.

By studying the expression of reporter genes inserted into the silent minichromosomes and VSG basic copy arrays (BCA) (in the MC and VB –ESp reporter lines), which together house the bulk of the silent *VSG* repertoire, I found that the mean eGFP fluorescence was consistently lower in the BF cells as compared to the PF cells. This could suggest that these loci are more tightly downregulated in BF, possibly in order to prevent premature expression of silent VSGs that would be detrimental for immune evasion. Alternatively, the difference in fluorescence could be an outcome of differential quenching in the two lifecycle stages.

By depleting TbISWI in the above reporter lines, I found that it is involved in preventing fortuitous transcription from these genomic loci. When TbISWI was depleted, leaky transcription was observed from the minichromosomes as well as the VSG BCA. Both of these silent regions after TbISWI knockdown were found to be more permissive for transcription in the procyclic form (PF) cells than in the bloodstream form (BF) cells, suggesting that additional mechanisms exist in the BF to prevent premature expression of silent VSGs for reasons noted earlier. In addition, when comparing the two loci, the VSG BCA appeared to be more permissive for transcription after TbISWI knockdown. This suggests a greater role for TbISWI in preventing fortuitous transcription from the VSG BCA. This is possibly because its proximity to the constitutively active pol II transcription units increases the chance of readthrough into this locus. These results imply

that at least one TbISWI complex is involved in the downregulation of transcription initiation from these loci.

The above results are particularly significant in the light of there being very few defined promoter sequences in *T. brucei*, and the presence of constitutive transcription of almost the entire genome (Clayton 2002). In yeast, ADR1 transcription factor (TF) has been shown to bind non-specific DNA sequences with affinities that are only ~10 fold lower than the native promoter motif (Schaufler et al. 2003). It is therefore possible that low levels of transcription can be generated from almost any genomic sequence in *T. brucei,* provided that adequate access to DNA is available. In such a scenario, the higher sequence complexity of the VSG BCA locus would allow a greater potential for encoding cryptic promoters than the simple sequence repeats of the minichromosomes, possibly another reason why TbISWI depletion results in greater derepression of the VSG BCA. Transcription from the silent VSG repertoires after TbISWI knockdown using tetracycline inducible RNAi can, therefore, be interpreted both as a result of fortuitous TF binding to nucleosome-depleted DNA as well as a result of less stringent TF exclusion in the absence of TbISWI.

From previous studies it is known that the ES promoter lacks the chromatin remodelling domain IV that is found in the structural architecture of other *T. brucei* pol I promoters like the rDNA and procyclin promoters (Zomerdijk et al. 1991; Pham et al. 1996; Laufer et al. 2001). Deletion of domain IV of the procyclin promoter resulted in significant loss of transcription from chromatin templates, while *in vitro* transcription was unaffected (Brown et al. 1992; Laufer et al. 2001). These results could explain why an ES promoter is more sensitive to its genomic context than other *T. brucei* pol I promoters (Rudenko et al. 1994; Horn et al. 1997). The sensitivity of the ES promoter can therefore

be used to "probe" for the chromatin state of a genomic locus it is inserted into. The ES promoter has also been shown to be derepressed in response to TbISWI depletion, suggesting that TbISWI has a role in maintaining the silent ESs in an inactive state (Hughes et al. 2007). I therefore made the reporter cell lines as before, albeit this time with an ES promoter upstream of the eGFP gene.

When TbISWI was depleted from the +ESp reporter cell lines, derepression of the eGFP was observed in both of the minichromosome and the VSG BCA locus. This argues that TbISWI is involved in maintaining the repressed state of the silent VSG repertoires. Contrary to the observations from the –ESp constructs, in this case the BF cell lines showed a higher derepression at both the loci. This suggests that the repressive role of TbISWI is greater in BF than in PF *T. brucei*, whereby its loss causes greater derepression in BF as a result. This also agrees with the previous study where BF silent ESs were more derepressed than PF silent ESs after TbISWI depletion (Hughes et al. 2007). As well as influencing transcription initiation, I propose that TbISWI may be involved in downregulating pol I transcription elongation. After TbISWI depletion, the downregulation of pol I elongation gets relieved, resulting in more transcription from stalled polymerases, and therefore, higher transcriptional derepression of these silent loci.

Analysing the rDNA spacer region, another transcriptionally silent locus in the *T. brucei* genome, I found that the –ESp reporter constructs in the rDNA locus were very tightly downregulated (with the exception of one cell line) in both lifecycle stages and TbISWI depletion had little effect on reporter expression. However, the +ESp reporter constructs did show a derepression effect after TbISWI depletion, but this was lower than that observed in the minichromosome and the VSG BCA loci. The observations from

both the –ESp and the +ESp cell lines suggest that the role of TbISWI in repressing the rDNA spacer is less pronounced than that in repressing the other loci.

It has been shown in the past that silent *VSG* ESs in both BF and PF trypanosomes exhibit low levels of non-processive transcription at their promoters (Ansorge et al. 1999; Rudenko et al. 1994; Vanhamme et al. 2000). This has led to the suggestion that VSG ES regulation is controlled at the level of transcription elongation (Vanhamme et al. 1998). A similar situation is also seen in the procyclin loci that are only weakly active in the bloodstream form *T. brucei* (Pays et al. 1990; Vanhamme et al. 1995). When TbISWI is knocked down using tetracycline inducible RNAi, both of these pol I loci get derepressed. While the promoter proximal genes in the silent expression sites are derepressed by 30-60 fold in BF and 10-17 in PF *T. brucei* (Hughes et al. 2007), the distal *VSGs* are derepressed by ~4-8 fold; and, the derepression of a silent procyclin locus in BF cells is ~6 fold. Unexpectedly, even the active procyclin locus in the PF cells gets derepressed by 2.5 fold after TbISWI depletion, suggesting that the *EP1* is not active to its maximal capacity in the native PF state due to a TbISWI mediated repressive mechanism. Taken together, these observations suggest that TbISWI could be involved in downregulation of transcription elongation in pol I transcription units in *T. brucei*. Alternatively, this could also be explained by an increase in transcription initiation events from these loci on TbISWI depletion, a fraction of which can escape the promoter-distal silencing in the ES.

Even when TbISWI was depleted from BF cell lines which had the *eGFP* reporter in the procyclin locus, the mean fluorescence at its most derepressed state was found to be still lower than the basal fluorescence in the PF cells. These observations suggest that transcription from the *EP1* locus is under additional repressive mechanisms in BF *T. brucei*.

The possible dual roles of TbISWI in preventing fortuitous initiation of transcription and regulating transcription elongation suggests the presence of two different TbISWI complexes in *T. brucei*. The role of TbISWI observed here both in transcription repression and elongation is similar to that seen for the yeast ISW1 at pol II loci, where the ISW1a complex is involved in repressing transcription initiation while ISW1b complex regulates elongation (Ioc4 being a positive regulator and Ioc2 being a negative regulator in the ISW1b complex) (Mellor et al. 2004).

Maintenance of centromere structure is important for kinetochore attachment and chromosome segregation, and therefore, for the orderly progression through the cell cycle. In many organisms, chromatin remodelling complexes have been shown to be involved in the maintenance of centromeric heterochromatin and cell cycle regulation (Okada et al. 2009; Grewal et al. 2007; Cao et al. 1997; Muchardt et al. 1999). Recent studies in *T. brucei* have found that disruption of the chromatin remodelling FACT complex results in a G2/early-M cell cycle arrest and impaired minichromosome segregation along with derepression of the silent ESs (Denninger et al. 2010). My experiments enquiring into the potential effect of TbISWI depletion on cell cycle progression and minichromosome segregation in *T. brucei* did not find any evidence for this. This also suggests that TbISWI is not involved in the maintenance of centromeric heterochromatin in *T. brucei*.

In summary, results from this study suggest that TbISWI, like yeast ISW1, possibly form two functionally distinct chromatin remodelling complexes. While one of them is involved in preventing fortuitous initiation of transcription, the other regulates transcriptional elongation at pol I transcription units. The higher representation of unstructured coils in the TbISWI structure suggests that it has a more flexible structure

and may be able to interact with many protein partners. Also, the lack of clearly defined charge zones on its C-terminal domains suggests that other protein partners in the complex may be involved in interaction with the DNA/ nucleosome.

In the future, it will be important to dissect the role of the individual C-terminal domains of TbISWI and the various binding partners of the ATPase. Cell lines will have to be made that express TbISWI with deletion/ mutation of specific C-terminal domains and the effect on chromatin remodelling activity monitored. In addition, TAP-tagged TbISWI can be used to pull down and identify other binding partners of the various TbISWI complexes. Investigation of the binding of TbISWI to various genomic loci, as performed for some repeat arrays in Appendix A2, will also help us understand its functional relevance. The effect of TbISWI mediated chromatin remodelling can also be studied by studying differences in nucleosome/ RNA polymerase distribution before and after TbISWI knockdown- either by micrococcal nuclease (MNase) digestion (for nucleosomes) or chromatin immunoprecipitation (ChIP) (for both nucleosomes and polymerases). Together, these experiments will be able to throw additional light on the role of TbISWI in transcriptional regulation in *T. brucei*.

Chapter 7

Material and Methods

Trypanosoma brucei cell lines and culturing conditions

Bloodstream form T. brucei reporter lines

All Bloodstream form (BF) *T. brucei* reporter cell lines used in this study were derived from S16.221 (lab name 221PurA3), a derivative of the *T. brucei* S16 "single marker cell line" (Wirtz et al. 1999). *T. brucei* S16.221 was obtained by transfecting *T. brucei* S16 with the pHNES221Pur1.6 construct (Tomoko Isobe, unpublished) that targets a puromycin resistance gene downstream of the active 221 expression site (ES) promoter. This cell line contains the genes encoding T7 RNA polymerase and the TetR (tetracycline repressor) present in the parental S16 line.

The *T. brucei* S16.221 cell line and its derivative reporter lines were cultured in HMI-9 medium with 10% v/v foetal calf serum (FCS) and 10% v/v Serum Plus (JRH Biosciences) (Hirumi et al. 1989) at 37 ˚C and 5% CO_2 in the presence of appropriate drugs. S16.221 trypanosomes were cultured in the presence of 2 μg ml^{-1} G418 (Invitrogen) and 0.2 μg ml^{-1} Puromycin (Sigma). Its derivative reporter cell lines were cultured in the presence of additional drugs depending on the drug selection marker present in either the reporter constructs or the RNAi constructs.

Procyclic form T. brucei reporter lines

All Procyclic form (PF) *T. brucei* reporter cell lines used in this study were derived from the *T. brucei* 29-13 cell line (Wirtz et al. 1999). *T. brucei* 29-13 contains genes encoding T7 RNA polymerase and TetR.

The *T. brucei* 29-13 cell line and its derivative reporter lines were cultured in SDM-79 medium (Brun et al. 1979) with the addition of 10% v/v FCS and 5 mg l^{-1} haemin at 27 ˚C in the presence of the appropriate drugs. 29-13 trypanosomes were

cultured in the presence of 20 μg ml^{-1} G418 (Invitrogen) and 25 μg ml^{-1} Hygromycin (Roche). Its derivative reporter cell lines were cultured in the presence of additional drugs depending on the drug selection marker in either the reporter constructs or the RNAi constructs.

Bloodstream form cell transfections

Transfections of the bloodstream form trypanosomes were performed by using either the BTX ECM 830 electroporator or the Amaxa Nucleofector® according to the following methods.

Method 1

This method was used essentially as described by Blundell *et al*. (Blundell et al. 1996). 5-10 μg of sterile DNA, prepared using phenol:chloroform extraction and ethanol precipitation, was electroporated using a BTX ECM 830 electroporator to deliver 3x 100 μs and 1.5 kV pulses at an interval of 100 ms. After 6 hr of recovery in HMI-9 medium at 37 ˚C/ 5% CO_2, appropriate drug selection was applied to the cells and they were plated out in 24 well plates and re-incubated. The plates were observed regularly for about 5-7 days when all parental-type cells had died and the transformants had grown out. A variation of the protocol (George Cross, *http://tryps.rockefeller.edu/trypsru2_protocols_index.html*) was also used whereby 3x 100 μs and 1.7 kV pulses were delivered at an interval of 200 ms. The cells were immediately plated with subsequent overnight recovery in the parental selection medium.

Method 2

Transfections using this method were performed with an Amaxa Nucleofactor® using its pre-defined X-001 program and the accompanying Human T Cells solution.

3×10^7 cells were transfected with 10 µg of sterile DNA, prepared using phenol:chloroform extraction and ethanol precipitation.

Procyclic form *T. brucei* transfections

Procyclic form cells were transfected using a protocol modified from Blundell et al. (Blundell et al. 1996). $2x10^7$ cells were electroporated with 5-10 µg of sterile DNA (prepared using phenol:chloroform extraction and ethanol precipitation) using a BTX ECM 830 electroporator to deliver 3x 100 µs and 1.7 kV pulses at an interval of 200 ms. After overnight recovery in SDM-79 medium at 27 ˚C, the cells were serially diluted in 20% v/v conditioned SDM-79 medium with appropriate drug selection and plated out in 96-well plates and re-incubated. The plates were observed for about 10-12 days after which all parental-type cells had died and the transformants had grown out. If parental cell death was not apparent after 7 days, the plates were replica-plated with two fold dilution in fresh 20% v/v conditioned SDM-79 medium with appropriate drug selection. The conditioned medium for the purpose was prepared by centrifuging a wild type PF cell culture of density $1x10^7$ cells ml^{-1} and saving the supernatant.

Genomic DNA extraction from *T. brucei* cell cultures

Genomic DNA (gDNA) was prepared from the *T. brucei* cell cultures by using the DNeasy Blood and Tissue kit (QIAGEN). Trypanosomes were centrifuged at 1000 rcf (relative centrifugal force) for 10 min. BF cells were washed in PSG (60 mM Na_2HPO_4, 3 mM NaH_2PO_4, 44 mM NaCl, 1% w/v glucose) and PF cells washed in PBS (7.7 mM Na_2HPO_4, 8.7 mM KCl, 2.3 mM KH_2PO_4, 120 mM NaCl, pH 7.4), and then resuspended in ATL buffer (DNeasy kit). After this step, the kit protocol for DNA extraction from yeast cells was followed.

Protein extraction from *T. brucei* cell cultures

Protein lysates were made using one of the two methods outlined below:

Method 1

Cells were centrifuged, washed once in ice cold buffer (BF cells in PSG and PF cells in PBS), and then resuspended in ice cold lysis buffer (50 mM HEPES pH 7.5, 10% v/v glycerol, 1% v/v Triton X-100, 1.5 mM $MgCl_2$, 1 mM EGTA and Roche complete protease inhibitors) at 10^9 cells ml^{-1}. After incubation at 4 °C for 20 min with gentle rotation, 4x Sample Buffer+DTT mix (0.19 M Tris pH 6.8, 6% w/v SDS, 30% v/v glycerol, 0.0075% w/v bromophenol blue, 0.25 M DTT) was added to the lysates and boiled for 10 min at 100 ˚C.

Method 2

Cells were centrifuged, washed once in ice cold buffer (BF cells in PSG and PF cells in PBS), and then resuspended in boiling 1x Sample Buffer+DTT mix (0.047 M Tris pH 6.8, 1.5% w/v SDS, 7.5% v/v glycerol, 0.0019% w/v bromophenol blue, 0.25 M DTT) and boiled for 10 min at 100 ˚C.

Flow Cytometry

For quantifying GFP fluorescence

Trypanosome cells were spun down at 1000 rcf for 10 min, washed (BF cells in PSG and PF cells in PBS), and resuspended in the respective buffer to an end density of $2x10^6$ cells ml^{-1}. Green florescence of the cells was measured in the FL1 channel using a Becton Dickinson FACSCalibur machine and CellQuest (BD) software.

Quantification of DNA content in BF cells

Trypanosome BF cells were centrifuged at 1000 rcf for 10 min, washed once in PSG, and then resuspended in 100 µl PSG. 100 µl of 4% v/v paraformaldehyde was added to the cell suspension and mixed; cells were then fixed for 2 hr at room temperature (RT) on a rotating wheel. After fixation, the cells were spun down at 1000 rcf for 3 min, washed, and permeablised by incubation at RT for 10 min in 0.1% v/v Triton X-100 (in PBS). After permeablisation, the cells were spun down at 1000 rcf for 3 min, washed once in PBS and resuspended in 500 µl PBS containing RNase A (final concentration 10 µg ml^{-1}) and propidium iodide (final concentration 10 µg ml^{-1}). After a 45 min incubation at 37 °C, red fluorescence of the cells was measured in the FL2 channel using a Becton Dickinson FACSCalibur machine and CellQuest (BD) software.

Quantification of DNA content in PF cells

DNA content in PF cells was measured using propidium iodide staining as described in Bessat & Ersfeld (2009). Trypanosome PF cells were spun down at 1000 rcf for 10 min, washed once in ice cold PBS, and resuspended in 200 µl of 0.5% v/v paraformaldehyde (in PBS; ice-cold and freshly made) and incubated on ice for 5 min. 2 ml of 70% v/v ethanol was then added drop wise to the cell suspension, and gently mixed. Cells were incubated at 4 ˚C for 1 hr and then centrifuged at 1000 rcf for 5 min and resuspended in 1 ml PBS containing RNase A (final concentration 50 µg ml^{-1}) and propidium iodide (final concentration 50 µg ml^{-1}). After a 30 min incubation at 37 °C, red florescence of the cells was measured in the FL2 channel using a Becton Dickinson FACSCalibur machine and CellQuest (BD) software.

Fluorescent *in situ* hybridisation (FISH)

FISH was performed on trypanosome cells as described in Bessat & Ersfeld (2009). About $5x10^6$ trypanosomes were induced for various time periods with 750 ng ml^{-1} tetracycline. Cells were centrifuged at 500 rcf / 4 ˚C for 10 min, washed twice in 1 ml ice cold buffer (BF cells in PSG and PF cells in PBS) and then resuspended in 100 µl of the respective buffer, and diluted to 1 ml of end concentration 3.6% v/v formaldehyde for fixation. After fixation for 15 min at room temperature (RT) on a rotating wheel, the cells were centrifuged, washed in PBS and allowed to settle in a 1x1 cm square delineated using a lipid pen on a Colorfrost® Plus Gold slide (Thermo Scientific) in a humid chamber at RT for 1 hr. After settling of the cells, the slide was washed with PBS and the cells were permeablised by incubating them with 100 µl of 0.1% w/v NP40 (in PBS) for 10 min for BF cells or for 5 min for PF cells. After permeablisation, the cells were washed twice for 5 min each in PBS and subsequently treated with 50 µg ml^{-1} RNase A (in PBS) for 1 hr at 37 °C. After RNase A treatment, the cells were washed twice with PBS and then covered with pre-hybridisation mix (2x SSC, 10% w/v dextran sulphate, 50 mM sodium phosphate buffer, 50% v/v formamide) and incubated in a humid chamber at RT for 1 hr. After pre-hybridisation, the hybridisation mix was carefully removed and a Geneframe® (ABgene) was stuck to the slide to create a gas-tight hybridisation chamber. Following this, the probe containing hybridisation buffer (pre-hybridisation mix with 1 mg ml^{-1} salmon sperm DNA, 1 mg ml^{-1} yeast tRNA, and the biotin labelled probe for 177 bp minichromosome repeats (gift from Klaus Ersfeld) was added to the hybridisation chamber and the chamber was sealed. The slide was heated to 85 °C for 5 min to denature the probe and the chromosomes and then incubated at 37 °C overnight for hybridisation. The slide was washed with 2x SSC/ 50% v/v formamide at 37 °C for 30 min, 2x SSC at 50 °C for 10 min, 0.2x SSC at 50 °C for 30 min, and 4x SSC at RT for 10 min to remove unbound probe. The cells were then incubated with Streptavidin, Alexa Fluor® 594

(Molecular Probes, dilution 1:500) for 1 hr at 37 °C in a humid chamber, washed twice with 0.01% v/v Tween 20/ TBS (25 mM Tris-HCl, 2.7 mM KCl, 137 mM NaCl, pH 7.4) and twice with PBS. The cells were then mounted in a drop of VECTASHIELD® Mounting Medium with 40 ng ml^{-1} DAPI (Vector Labs). Cells were observed using a fluorescence microscope (Zeiss Axioplan 2 imaging E) and pictures taken with a CoolSnap HQ camera using MetaMorph® software.

Cell Cycle analysis by microscopy

About $2x10^6$ trypanosomes were induced for various time periods with 750 ng ml^{-1} tetracycline. Cells were centrifuged at 1000 rcf for 10 min, washed (BF cells in PSG and PF cells in PBS) and resuspended in 100 µl of the respective buffer. The suspension was then diluted with an equal volume of 4% v/v paraformaldehyde by gentle mixing. The final suspension was incubated at RT for 20 min on a rotating wheel for fixation. After fixation, the cells were allowed to settle in a 1x1 cm square delineated using a lipid pen on a Colorfrost® Plus Gold slide (Thermo Scientific). After the cells were allowed 30 min of settling onto the slides, the excess suspension was blotted away with a tissue paper. The slide was then washed once in PBS, blotted, air dried and then mounted in a drop of the VECTASHIELD® Mounting Medium with 40 ng ml^{-1} DAPI (Vector Labs). Cells were observed using a fluorescence microscope (Zeiss Axioplan 2 imaging E) and pictures taken with a CoolSnap HQ camera using MetaMorph® software.

Analysis of *T. brucei* genomic DNA

Pulsed-field gel electrophoresis

Pulsed-field Gel Electrophoresis (PFGE) was used to analyse whether the reporter constructs had correctly integrated into the minichromosomes. Trypanosomes were

embedded in blocks of 0.5% w/v low melting point agarose (in PSG), at a concentration of $2x10^7$ trypanosomes per 75 µl block, and digested with proteinase K for 24 hr at 56 ˚C (modified from (Melville et al. 1998)). Chromosome separation was performed using a BioRad CHEF DRIII electrophoresis system in a 1% w/v agarose gel (routine use; Sigma-Aldrich) in 0.5x $TBE_{0.1}$ (45 mM Tris-borate, 0.1 mM EDTA, pH 8.0). The chromosomes were electrophoresed at 6 V cm^{-1} for 22 hr with a switching time of 30 sec, a reorientation angle of 120°, and at a constant temperature of 14 °C maintained by buffer circulation. The gel was stained with ethidium bromide (aqueous solution, 0.5 µg ml^{-1}) and imaged using a UV transilluminator.

Restriction enzyme mapping

Restriction enzyme mapping was performed to analyse the correct integration of the reporter constructs into the rDNA spacers in the genome. 2 µg of *T. brucei* genomic DNA (gDNA) was digested for 3-4 hr using the appropriate restriction enzyme/s (New England Biolabs) in five fold excess. 1 µg digested gDNA was loaded per lane onto a 0.6% w/v agarose gel / 0.5x TBE (45 mM Tris-borate, 1 mM EDTA, pH 8.0), and electrophoresed overnight at 30 V_{const}. The gel was stained with ethidium bromide (aqueous solution, 0.5 µg ml^{-1}) and imaged using a UV transilluminator.

Southern Blot analysis

Restriction enzyme digested gDNA or chromosomes separated on an agarose gel were depurinated in 0.25 M HCl for 20 min, denatured and transferred onto Hybond XL membranes (Amersham) using standard procedures (Sambrook et al. 2001). Probes were prepared from 25 ng DNA by radiolabelling with α-^{32}P-dCTP by random priming, using the Amersham Ready-To-Go DNA Labeling Beads (-dCTP) and purified with Illustra

ProbeQuant G50 Micro columns using the manufacturers' protocols. Membranes were hybridised with the radioactive probes overnight at 42 °C in Hybridisation mix (0.05 M NaPi pH 7.4, 0.9 M NaCl, 5× Denhardt solution, 10% w/v Dextran Sulphate, 5 mM EDTA pH 8.0, 0.1% w/v SDS, 40% v/v formamide, 0.05 mg ml^{-1} salmon sperm DNA). Membranes were washed to an end stringency of 0.1x SSC, 0.1% w/v SDS at 65 °C for 45 min. Hybridised probe was detected by exposing a phosphor screen (Bio-Rad Imaging Screen-K) with the probed membrane and scanning the screen using a phosphoimager (Bio-Rad Personal Molecular Imager System) using Quantity One software (Bio-Rad).

Southern blots were hybridised with probes for the eGFP gene, pBlueScript plasmid, 18s rRNA gene, and 28sβ rRNA gene. The eGFP probe (738 bp) spanned the full length of the gene and was isolated from peGFP plasmid (Daniëlle te Vruchte, unpublished) by EcoRI / NcoI digestion. The pBlueScript probe (2856 bp) consisted of the full length of the pBluescript backbone and was obtained from the pBlueScript KS plasmid by excising out the multiple cloning site (MCS) from the plasmid using SacI / KpnI digestion. The 18S rDNA probe (2195 bp) spanned the full length of the gene and was amplified using primers 18s-probe-3249s (5'-GCCAGTAGTCATATGCTTGTTTCAAGGACTTAGCCATGC-3') and 18s-probe-5443as (5'- GACTTTTGCTTCCTCTATTGAAGCAATATCGGTGA-3'). The 28Sβ rDNA probe (1501 bp) spanned most of the length of the gene and was amplified using primers 28sB-probe-8881s (5'-TCCCAACTGCAGACCGTACTCATCACCGCATCAGG-3') and 28sB-prb-10381as (5'-AAAAAAACGGACTGCCTCCGCTTTGCGCGG-3').

Western Blot analysis

Protein lysates from 7.5-10x10^6 cells (with the same number used with each gel) were loaded per lane on a 6% SDS-polyacrylamide gel and electrophoresed using standard procedures (Sambrook et al. 2001). The proteins from the gel were transferred onto Hybond-P (Amersham) membrane using the Mini Trans-Blot Cell (Bio-Rad). For Western blot analysis the membrane was blocked with 5% w/v milk in PBST (phosphate buffered saline with 0.02% v/v Tween-20) for 1 hr, rinsed twice in PBST for 2 min each and then incubated for 1 hr with the primary antibody, appropriately diluted in PBST. After incubation with the primary antibody, the membrane was rinsed once and washed twice for 15 min in PBST. The signal on the membrane was then incubated for 1 hr with the secondary antibody coupled to horseradish peroxidase (HRP), also diluted appropriately in PBST. After incubation with the secondary antibody, the membrane was rinsed once and washed twice for 15 min each in PBST. The membrane was then visualised using the Western Lightning Plus ECL system (PerkinElmer). The membrane was visualised using the Fluor-S MultiImager (Bio-Rad) and Quantity One software (Bio-Rad). If re-probing was required, membranes were stripped by incubating them at 60 °C for 30 min in a Stripping Solution (100 mM β-mercaptoethanol, 2% w/v SDS, 62.5 mM Tris-Cl, pH 6.8). The stripped membranes were then washed three times with PBST for 5 min before re-blocking in 5% w/v milk.

For all Western blots showing TbISWI depletion, a rabbit polyclonal antibody raised against the C-terminal 207 aa of the protein was used at 1:1000 dilution (Hughes et al. 2007). A rabbit affinity purified anti-SMC3 antibody (gift from Klaus Ersfeld) was used at 1:500 dilution to confirm equal loading. A mouse monoclonal anti-Ty BB2 antibody (gift from Keith Gull) was used at 1:100 dilution to probe for the expression of Ty-tagged TbISWI.

For membranes probed with rabbit primary antibodies, ECL Rabbit IgG HRP-linked Whole Ab (from donkey; Amersham) was used as a secondary antibody at 1:10000 dilution. For membranes probed with rabbit primary antibodies, ECL Mouse IgG HRP-linked Whole Ab (from sheep; Amersham) was used as a secondary antibody at 1:5000 dilution.

Plasmid constructs

All plasmid constructs were amplified in *E. coli* DH5α (Stratagene) cells. Standard protocols were used for making competent bacterial cells using calcium chloride and their subsequent transformation with DNA (Sambrook et al. 2001).

QIAGEN Miniprep or Midiprep kits were used for DNA isolation from bacterial cells and MinElute Gel Extraction Kit (QIAGEN) for elution of DNA fragments from agarose gels. T4 DNA ligase (Roche) was used for DNA ligation reactions, T4 DNA polymerase (New England Biolabs) for blunting, Shrimp Alkaline Phosphatase (Roche) for dephosphorylation of DNA ends, and T4 polynucleotide kinase (New England Biolabs) for phosphorylation of DNA ends. All the DNA isolation and modification procedures were carried out as per the manufacturer's instructions.

Reporter constructs

All reporter constructs used in this study contain a Blasticidin resistance (BlastR) gene, flanked by appropriate transcript processing signals, downstream of an rDNA promoter. The BlastR gene acts as a marker for the selection of *T. brucei* transformants. The constructs also contain an enhanced GFP (eGFP) reporter gene also flanked by appropriate processing signals. The reporter constructs were made in two sets. The first

set (–ESp) did not have an expression site (ES) promoter. The second set (+ESp) had an ES promoter upstream of the eGFP gene cassette.

All reporter constructs of the –ES set were derived from peGFP_RPtubBLAST, by inserting a specific genome targeting fragment upstream of the eGFP gene cassette. Steps involved in the construction of peGFP_RPtubBLAST are detailed below:

1) The eGFP fragment was isolated from peGFP (Daniëlle te Vruchte, unpublished) by excision with NotI/SmaI and then blunted.

2) ΔpBAN (Gloria Rudenko, unpublished) was linearised by EcoRI and blunted, and the blunt-ended eGFP fragment from Step 1 ligated into it. The new construct was called ΔpBAN_eGFP.

3) The tub-eGFP-tub cassette was isolated from ΔpBAN_eGFP by excision with BamHI/HindIII and then blunted.

4) RPtubBLAST (Marcus Gould, unpublished) was linearised with BamHI and blunted, and the blunt ended tub-eGFP-tub from Step 3 was ligated into it. The new construct was called peGFP_RPtubBLAST.

5) peGFP_RPtubBLAST was linearised with NotI and blunted, and the blunt-ended genome targeting fragments for different genomic loci were ligated into it.

All reporter constructs of the +ES set were derived from pESeGFP_RPtubBLAST, by inserting a specific genome targeting fragment upstream of the ES promoter. Steps involved in the construction of pESeGFP_RPtubBLAST are detailed below:

1. The ES promoter fragment was excised from rDES1 (Rudenko et al. 1994) by digestion with XbaI.

2. ΔpBAN_eGFP (described above) was linearised with XbaI and the ES promoter fragment with compatible ends from Step 3 was ligated into it. The new construct was called pESp_tub-eGFP-tub.

3. The ESp-tub-eGFP-tub fragment was isolated from pESp_tub-eGFP-tub by digesting with NotI/HindIII and then blunted.

4. RPtubBLAST (Marcus Gould, unpublished) was linearised with BamHI and blunted, and the blunt ended ESp-tub-eGFP-tub fragment from Step 5 was ligated into it. The new construct was called pESeGFP_RPtubBLAST.

5. pESeGFP_RPtubBLAST was linearised with NotI and blunted, and the blunt-ended genome targeting fragments for different genomic loci were ligated into it.

To amplify the VSG BCA targeting fragment from *T. brucei* 427 genomic DNA, the following primers were used: BCVSG118-863s (5'-tatggatccCTTAAGTGAGCCAACACACCAG-3') and BCVSG118-2025as (5'-ctggaattcCTGACGAACCTCTAGTTACCG-3'). To amplify the EP1 procyclin targeting fragment from genomic DNA, the following primers were used: EP1_UP_2480s (5'-tcagtggttgtgtggtag-3') and EP1_UP_3991as (5'-tacttcttggtgtaattgaagt-3').

TbISWI knockout construct

Two TbISWI knockout constructs, pSpot5KOHA57 and pSpot5KOPhleo, were made with the aim of replacing TbISWI with drug resistance genes, HygroR and PhleoR respectively. Steps involved in the construction of these plasmids are detailed below:

1) The upstream region of the TbISWI was amplified from *T. brucei* 427 genomic DNA using primers Spot5-957s-UPF (5'-tatctagaTACCAGTGCTGTTGGGCCACCTTC-3') and Spot5-1455as-US (5'-atggatccCGGTACCTTATCTGGTTCTATG-3'), the amplified DNA digested with XbaI/BamHI and inserted into a pBlueScript vector linearised with XbaI/BamHI. The plasmid was called pSpot5UP.

2) The downstream region of the TbISWI was amplified from *T. brucei* 427 genomic DNA using primers Spot5-4935s-DSF (5'-taaagcttTAATCATATTGAAGGGGGATCC-3') and Spot5-5520as-DS (5'-tactcgagGCACCTTGATGTTCCTATGC-3'), the amplified DNA digested with HindIII/XhoI and inserted into a pBlueScript vector linearised with HindIII/XhoI. The plasmid was called pSpot5DN.

3) pSpot5UP was cut open with BamHI/HindIII. The Hygromycin resistance gene was cut out from pG3HA57 with BamHI/HindIII. The Phleomycin resistance gene was amplified using primers Phleo31s (5'-ATGGGCGAAATGACCGACCA-3') and Phleo420as (5'-TCATGAGATGCCTGCAAGCA-3') and then digested with BamHI/HindIII. In two separate reactions, the Hygromycin and the Phleomycin genes were inserted into the linearised pSpot5UP construct resulting in plasmids pSpot5UPHA57 and pSpot5UPPhleo.

4) The downstream region of TbISWI was isolated from pSpot5DN by cutting with HindIII/XhoI. pSpot5UPHA57 and pSpot5UPPhleo were linearised with HindIII/XhoI and the TbISWI downstream region inserted into them. The resulting plasmids were called pSpot5KOHA57 and pSpot5KOPhleo, respectively.

Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) was performed as described in Stanne & Rudenko (2010). Trypanosome cultures were fixed in 1% v/v formaldehyde for 20 min at RT. The reaction was stopped by adding glycine to a final concentration of 125 mM. Chromatin was sonicated to an average DNA size of about 200-500 bp using a BioRuptor (Cosmo Bio), and clarified by centrifugation at 15000 rcf for 5 min at 4 °C. The sonicated extract was precleared using Protein A Sepharose CL-4B beads (GE Healthcare) and incubated for 16 hr either with or without the anti-Myc antibody (Sigma 9E10 clone, M5546). Chromatin from $7x10^7$ cells equivalents were used per IP. The protein-DNA complexes were incubated with Protein A Sepharose CL-4B beads (GE Healthcare) for 2 hours and eluted from the beads after washing in 1% w/v SDS/ 0.1 M NaHCO3 preheated to 65 ºC. The cross-links were reversed by adding NaCl to a final concentration of 325 mM and incubated at 65 ºC overnight. Following RNAse A and proteinase K treatment, DNA was purified using QIAquick PCR purification kit (QIAGEN). The results were analysed by hybridisation of slot-blots with radiolabelled probes followed by signal quantification using Quantity One software (Bio-Rad).

Quantitative PCR

Quantitative PCR (qPCR) was used to quantify different transcripts in trypanosome cells depleted of TbISWI by induction of *TbISWI* RNAi with 750 ng ml^{-1} tetracycline for various time periods. The cell lines used for the experiment were *T. brucei* T3-SA1 (BF) and D3-SA1 (PF) (Hughes et al. 2007).

After the induction of RNAi, total RNA was isolated from the cells using the RNeasy Kit (QIAGEN), according to the manufacturer's protocol. To remove any gDNA contamination, the extracted RNA was DNase-treated using the TURBO DNA-free Kit

(Ambion). 1 µg RNA per sample was used as template for complementary DNA (cDNA) synthesis using the Omniscript RT Kit (QIAGEN). All qPCR samples were amplified in triplicate in a 20 µl total reaction volume containing 19 µl of qPCR reaction mix (LightCycler480 SYBR green Master Mix; Roche), 0.5 µM (each) forward and reverse primers, and 1 µl of the cDNA template. Reaction mixtures were amplified using a LightCycler480 Real-Time PCR System (Roche) according to the manufacturer's instructions.

The qPCR reactions were performed for 45 cycles of denaturation (95 ˚C for 15s), annealing (variable temperature for 15s) and elongation (72 ˚C for variable time length). All qPCR primers and conditions used are listed in Table 7.1. Primers were optimised by Tara Stanne (unpublished) to allow the production of a single amplification product of the correct size.

Table 7.1 Primers used for Quantitative PCR to determine transcript abundance in total cellular RNA obtained after RNAi mediated TbISWI depletion.

Transcript of interest	Annealing temperature (ºC)	Elongation time (s)	Primer name	Primer Sequence (5'→3')
TbISWI	60	15	ISWI_3100s	cgtgtagaggagaggatac
			ISWI_3267as	gtacgccataagaaataact
eGFP	63	10	eGFP_184s	gtgaccaccctgacctac
			eGFP_324as	cttgtagttgccgtcgtc
rDNA spacer	60	10	rDNAspacer_178s	attttctctacccctctctt
			rDNAspacer_323as	atcatcgtatcattttcatc
VSG 118 BCA	63	10	VSGBCA118_108s	ctgcaagatgtctaaggaac
			VSGBCA118_193as	caagttctaggatctccgtta
Procyclin	60	10	Procyclin_8s	ctcgttccctttatctactc
			Procyclin_102as	cttgtcttctggtccttc
VSG T3	64	15	VSGT3_675s	aaacgacgcaacctctatc
			VSGT3_904as	ctatgtcttctgggcttga
VSG 121	61	10	VSG121_953s	tacagaaggaaggcaaatac
			VSG121_1022as	aatatcttgtcgaggttctt
VSG 1.8	64	10	VSG1.8_753s	gcagcaaacaataacggact

			VSG1.8_834as	tgacgactgaaagagggtgt
VSG JS1	64	15	VSGJS1_70s	actgtaagggtgttgcttct
			VSGJS1_234as	ttgtgctgttaggctcttct
γ-tubulin	57	10	gammaTub_491s	aatcaccgtatttggacacaag
			gammaTub_583as	gtttgagtgttcgctgaaagag

PCR using genomic DNA as template

For trypanosome cell lines containing reporter constructs targeted to the VSG118 basic copy arrays (BCA), the procyclin EP1 locus and the tubulin locus, PCR amplification on gDNA extracted from these lines was used to establish that the construct has integrated correctly. In addition, PCR amplification using gDNA template was also used to verify replacement of a –ESp reporter construct by a +ESp reporter construct and *vice versa* in the rDNA spacer targeting *eGFP* reporter cell lines. All PCR primers used for gDNA linking and accompanying control reactions are listed in Table 7.2.

Each 20 µl PCR reaction contained the template gDNA (1.5-10 ng μl^{-1}), forward and reverse primers (0.5 mM each), dNTPs (0.5 mM per base), 1x ThermoPol Reaction buffer (New England Biolabs), Taq DNA polymerase (0.125 U μl^{-1}; New England Biolabs). The PCR reactions were performed for 35 cycles of denaturation (94 ˚C), annealing (57 ˚C) and elongation (72 ˚C).

Table 7.2 Primers used for linking PCR using genomic DNA as template to verify correct integration of the reporter construct.

Construct targeting locus	**Direction of linking**	**Primer name**	**Primer Sequence (5'→3')**
VSG 118 basic copy array	Upstream	BCVSG118-42s	tacggatccgcaactaagtgtgcaagacaag
		eGFP-183as	taggtcagggtggtcacg
	Downstream	T3	aattaaccctcactaaaggg
		BCVSG118-2100as	ctggcgtgtcgctgactgc
EP1	Upstream	EP1-2082s	ggatggtagcctttccacaa

procyclin		eGFP-183as	As above
	Downstream	T3	As above
		EP1-4177as	gtcagtgccattggtatcgtc
Tubulin	Upstream	Btub-261s	cgacaacttcatctttggacag
		eGFP-183as	As above
rDNA spacer	Upstream (with eGFP)	rDNASTF-17610s	atacccagcactttcagatttctc
		eGFP-183as	As above
	Upstream (with ES promoter)	rDNASTF-17610s	As above
		DES-1291as	accgaaatatcgctctgcat
Blasticidin resistance gene	Upstream	batub-1835s	gcggtgctgatttctgacag
		BSR-138as	gacaccagtgaagatgcggc
Phleomycin resistance gene	Upstream	batub-1835s	As above
		PhleoII-372as	gtcctgctcctcggccacgaa
RNA polymerase I	Internal (control)	PolI-484s	caggaggatcgttcggcaccttggc
		PolI-1123as	cactgaatgatcttagatggctaatcg

Protein phylogenetics

Using the *T. brucei* proteins ISWI (Tb927.2.1810) and RPB1 (Tb927.4.5020; RNA polymerase II large subunit) as a starting point, homologous proteins from other organisms were identified using a combination of GeneDB BLAST (http://www.genedb.org/), NCBI BLASTp and tBLASTn, and NCBI psi-BLAST for three iterations (http://blast.ncbi.nlm.nih.gov/). The results were manually analysed on the basis of multiple criteria: (a) E-value of BLAST hit, (b) domains predicted using InterProScan at http://www.ebi.ac.uk/Tools/InterProScan (Quevillon et al. 2005), (c) synteny conservation as annotated in TriTrypDB (Aslett et al. 2010), (d) orthologous group membership in OrthoMCL-DB (Chen et al. 2006). The relevant protein sequences were downloaded (accession numbers in Table 7.3) and processed further.

The *T. brucei* protein and its kinetoplastid homologues were aligned using ClustalW2 (Larkin et al. 2007) at http://www.ebi.ac.uk/Tools/clustalw2/ and the

alignments downloaded in phylip format. The alignments were subsequently processed using locally downloaded ProtTest (Abascal et al. 2005) to determine the best-fit amino acid substitution model for the protein and calculate the phylogenetic distances using that model. The results were downloaded in Newick format and visualised using TreeView (Page 1996).

Table 7.3 Accession numbers of proteins used for phylogenetic analysis

Organism	**ISWI homologue**	**RNA pol II large subunit**
Trypanosoma brucei	Tb927.2.1810	Tb927.4.5020
Trypanosoma brucei gambiense	Tbg.972.2.360	Tbg972.4.5210
Trypanosoma congolense	TcIL3000.2.50	TcIL3000.4.4270
Trypanosoma vivax	TvY486_0200200	TvY486_0806890
Trypanosoma cruzi	Tc00.1047053509213.170	Tc00.1047053509011.20
Leishmania braziliensis	LbrM33_V2.1970	LbrM31_V2.2950
Leishmania infantum	LinJ33_V3.1800	LinJ31_V3.2680
Leishmania major	LmjF33.1700	LmjF31.2610
Plasmodium falciparum	XP_001347729	XP_001351252
Toxoplasma gondii	XP_002371685	XP_002366199
Saccharomyces cerevisiae	NP_009804, NP_014948	NP_010141
Caenorhabditis elegans	AAA50636	P16356
Drosophila melanogaster	NP_725204	P04052
Mus musculus	AAK52454	P08775
Homo sapiens	AAH23144	P24928
Giardia lamblia	GLP15_3376	GLP15_1808

a) All *Trypanosoma* and *Leishmania* accession numbers are from GeneDB *www.genedb.org*
b) *Giardia* accession numbers are from GiardiaDB *www.giardiadb.org*
c) All other accession numbers are from the NCBI Protein database *www.ncbi.nlm.nih.gov/protein/*

Domain prediction

ISWI domains were predicted using a combination of InterProScan (Quevillon et al. 2005), a tool that uses different protein signature recognition methods and HHpred

(Söding et al. 2005), a remote homologue search and secondary structure prediction algorithm.

Homology modelling

Amino acid sequences of *Saccharomyces cerevisiae* ISW1, *Drosophila melanogaster* ISWI and *Trypanosoma brucei* ISWI were aligned using AlignX and their C-termini (from HAND to SLIDE domain as determined using the prediction methods above) modelled using Phyre (version 0.2) (Kelley et al. 2009). The SANT and SLIDE domains from these ISWIs were also modelled separately. The structural PDB files from Phyre were processed using PDB2PQR (version 1.6) at *http://www.poissonboltzmann.org/pdb2pqr* to calculate the Poisson-Boltzmann electrostatics of the modelled surfaces (Dolinsky et al. 2004). The PDB and PQR files, obtained from Phyre and PDB2PQR respectively, were then used to plot the calculated charges onto the modelled domain surface using APBS tools in PyMOL software (DeLano 2009; Baker et al. 2001).

Chapter 8

References

Aasland, R., Stewart, A. F., and Gibson, T. (1996). The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. Trends in Biochemical Sciences 21, 87-8.

Abascal, F., Zardoya, R., and Posada, D. (2005). ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21, 2104-5.

Alsford, S., Kawahara, T., Glover, L., and Horn, D. (2005). Tagging a T. brucei RRNA locus improves stable transfection efficiency and circumvents inducible expression position effects. Molecular and Biochemical Parasitology 144, 142-8.

Alsford, S., Kawahara, T., Isamah, C., and Horn, D. (2007). A sirtuin in the African trypanosome is involved in both DNA repair and telomeric gene silencing but is not required for antigenic variation. Molecular Microbiology 63, 724-36.

Alsford, S., and Horn, D. (2004). Trypanosomatid histones. Molecular Microbiology 53, 365-72.

Alén, C., Kent, N. A., Jones, H. S., O'Sullivan, J., Aranda, A., and Proudfoot, N. J. (2002). A role for chromatin remodeling in transcriptional termination by RNA polymerase II. Molecular Cell 10, 1441-52.

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

Aslett, M. et al. (2010). TriTrypDB: a functional genomic resource for the Trypanosomatidae. Nucleic Acids Research 38, D457-62.

Badenhorst, P., Voas, M., Rebay, I., and Wu, C. (2002). Biological functions of the ISWI chromatin remodeling complex NURF. Genes & Development 16, 3186-98.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proceedings of the National Academy of Sciences of the United States of America 98, 10037-41.

Baral, T. N. (2010). Immunobiology of african trypanosomes: need of alternative interventions. Journal of Biomedicine & Biotechnology 2010, 389153.

Barry, J. D., Ginger, M. L., Burton, P., and McCulloch, R. (2003). Why are parasite contingency genes often associated with telomeres? International Journal for Parasitology 33, 29-45.

Beaglehole, R., Irwin, A., and Prentice, T. (2004). The World Health Report 2004- changing history (World Health Organization).

Becker, M., Aitcheson, N., Byles, E., Wickstead, B., Louis, E., and Rudenko, G. (2004). Isolation of the repertoire of VSG expression site containing telomeres of

Trypanosoma brucei 427 using transformation-associated recombination in yeast. Genome Research 14, 2319-29.

Berriman, M. et al. (2005). The genome of the African trypanosome Trypanosoma brucei. Science 309, 416-22.

Bessat, M., and Ersfeld, K. (2009). Functional characterization of cohesin SMC3 and separase and their roles in the segregation of large and minichromosomes in Trypanosoma brucei. Molecular Microbiology 71, 1371-85.

Biebinger, S., Rettenmaier, S., Flaspohler, J., Hartmann, C., Peña-Diaz, J., Wirtz, L. E., Hotz, H. R., Barry, J. D., and Clayton, C. (1996). The PARP promoter of Trypanosoma brucei is developmentally regulated in a chromosomal context. Nucleic Acids Research 24, 1202-11.

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

Blundell, P. A., Rudenko, G., and Borst, P. (1996). Targeting of exogenous DNA into Trypanosoma brucei requires a high degree of homology between donor and target DNA. Molecular and Biochemical Parasitology 76, 215-29.

Borst, P., and Sabatini, R. (2008). Base J: discovery, biosynthesis, and possible functions. Annual Review of Microbiology 62, 235-51.

Bouazoune, K., and Brehm, A. (2006). ATP-dependent chromatin remodeling complexes in Drosophila. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology 14, 433-49.

Brems, S., Guilbride, D. L., Gundlesdodjir-Planck, D., Busold, C., Luu, V., Schanne, M., Hoheisel, J., and Clayton, C. (2005). The transcriptomes of Trypanosoma brucei Lister 427 and TREU927 bloodstream and procyclic trypomastigotes. Molecular and Biochemical Parasitology 139, 163-72.

Brown, S. D., Huang, J., and Van Der Ploeg, L. H. (1992). The promoter for the procyclic acidic repetitive protein (PARP) genes of Trypanosoma brucei shares features with RNA polymerase I promoters. Molecular and Cellular Biology 12, 2644-52.

Brun, R., and Schönenberger (1979). Cultivation and in vitro cloning or procyclic culture forms of Trypanosoma brucei in a semi-defined medium. Short communication. Acta Tropica 36, 289-92.

Böhme, U., and Cross, G. A. (2002). Mutational analysis of the variant surface glycoprotein GPI-anchor signal sequence in Trypanosoma brucei. Journal of Cell Science 115, 805-16.

Cairns, B. R. (2009). The logic of chromatin architecture and remodelling at promoters. Nature 461, 193-8.

Cairns, B. R. et al. (1996). RSC, an essential, abundant chromatin-remodeling complex. Cell 87, 1249-60.

Campbell, D. A., Sturm, N. R., and Yu, M. C. (2000). Transcription of the kinetoplastid spliced leader RNA gene. Parasitology Today 16, 78-82.

Campbell, D. A., Thomas, S., and Sturm, N. R. (2003). Transcription in kinetoplastid protozoa: why be normal? Microbes and Infection 5, 1231-40.

Cao, Y., Adachi, J., Janke, A., Pääbo, S., and Hasegawa, M. (1994). Phylogenetic relationships among eutherian orders estimated from inferred sequences of mitochondrial proteins: instability of a tree based on a single gene. Journal of Molecular Evolution 39, 519-27.

Cao, Y., Cairns, B. R., Kornberg, R. D., and Laurent, B. C. (1997). Sfh1p, a component of a novel chromatin-remodeling complex, is required for cell cycle progression. Molecular and Cellular Biology 17, 3323-34.

Carlson, M., and Laurent, B. C. (1994). The SNF/SWI family of global transcriptional activators. Current Opinion in Cell Biology 6, 396-402.

Castro, C., Hernández, R., and Castañeda, M. (1981). Trypanosoma cruzi ribosomal RNA: internal break in the large-molecular-mass species and number of genes. Molecular and Biochemical Parasitology 2, 219-33.

Chaves, I., Rudenko, G., Dirks-Mulder, A., Cross, M., and Borst, P. (1999). Control of variant surface glycoprotein gene-expression sites in Trypanosoma brucei. The EMBO Journal 18, 4846-55.

Chen, F., Mackey, A. J., Stoeckert, C. J., and Roos, D. S. (2006). OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. Nucleic Acids Research 34, D363-8.

Chung, H. M., Shea, C., Fields, S., Taub, R. N., Van Der Ploeg, L. H., and Tse, D. B. (1990). Architectural organization in the interphase nucleus of the protozoan Trypanosoma brucei: location of telomeres and mini-chromosomes. The EMBO Journal 9, 2611-9.

Clapier, C. R., Nightingale, K. P., and Becker, P. B. (2002). A critical epitope for substrate recognition by the nucleosome remodeling ATPase ISWI. Nucleic Acids Research 30, 649-55.

Clayton, C. E. (1999). Genetic manipulation of kinetoplastida. Parasitology Today 15, 372-8.

Clayton, C. E. (2002). Life without transcriptional control? From fly to man and back again. The EMBO Journal 21, 1881-8.

Clayton, C. E. (1985). Structure and regulated expression of genes encoding fructose biphosphate aldolase in Trypanosoma brucei. The EMBO Journal 4, 2997-3003.

Collins, N., Poot, R. A., Kukimoto, I., García-Jiménez, C., Dellaire, G., and Varga-Weisz, P. D. (2002). An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. Nature Genetics 32, 627-32.

Conaway, R. C., and Conaway, J. W. (2009). The INO80 chromatin remodeling complex in transcription, replication and repair. Trends in Biochemical Sciences 34, 71-7.

Corona, D. F., Clapier, C. R., Becker, P. B., and Tamkun, J. W. (2002). Modulation of ISWI function by site-specific histone acetylation. EMBO reports 3, 242-7.

Corona, D. F., Längst, G., Clapier, C. R., Bonte, E. J., Ferrari, S., Tamkun, J. W., and Becker, P. B. (1999). ISWI is an ATP-dependent nucleosome remodeling factor. Molecular Cell 3, 239-45.

Corona, D. F., and Tamkun, J. W. (2004). Multiple roles for ISWI in transcription, chromosome organization and DNA replication. Biochimica et Biophysica Acta 1677, 113-9.

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

Cross, G. A., Wirtz, L. E., and Navarro, M. (1998). Regulation of vsg expression site transcription and switching in Trypanosoma brucei. Molecular and Biochemical Parasitology 91, 77-91.

Cumming, G., Fidler, F., and Vaux, D. L. (2007). Error bars in experimental biology. The Journal of Cell Biology 177, 7-11.

Dammann, R., Lucchini, R., Koller, T., and Sogo, J. M. (1995). Transcription in the yeast rRNA gene locus: distribution of the active gene copies and chromatin structure of their flanking regulatory sequences. Molecular and Cellular Biology 15, 5294-303.

Dang, W., Kagalwala, M. N., and Bartholomew, B. (2006). Regulation of ISW2 by concerted action of histone H4 tail and extranucleosomal DNA. Molecular and Cellular Biology 26, 7388-96.

Dang, W., and Bartholomew, B. (2007). Domain architecture of the catalytic subunit in the ISW2-nucleosome complex. Molecular and Cellular Biology 27, 8306-17.

De Gaudenzi, J. G., D'Orso, I., and Frasch, A. C. (2003). RNA recognition motif-type RNA-binding proteins in Trypanosoma cruzi form a family involved in the interaction with specific transcripts in vivo. The Journal of Biological Chemistry 278, 18884-94.

DeLano, W. (2009). The PyMOL Molecular Graphics System.

Delmas, V., Stokes, D. G., and Perry, R. P. (1993). A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-like helicase domain.

Proceedings of the National Academy of Sciences of the United States of America 90, 2414-8.

Denninger, V., Fullbrook, A., Bessat, M., Ersfeld, K., and Rudenko, G. (2010). FACT is involved in cell cycle specific control of VSG expression sites in Trypanosoma brucei. Manuscript submitted for publication.

Deuring, R. et al. (2000). The ISWI chromatin-remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo. Molecular Cell 5, 355-65.

Diehl, S., Diehl, F., El-Sayed, N. M., Clayton, C., and Hoheisel, J. D. (2002). Analysis of stage-specific gene expression in the bloodstream and the procyclic form of Trypanosoma brucei using a genomic DNA-microarray. Molecular and Biochemical Parasitology 123, 115-23.

Dirscherl, S. S., and Krebs, J. E. (2004). Functional diversity of ISWI complexes. Biochemistry and Cell Biology 82, 482-9.

Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004). PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Research 32, W665-7.

Donelson, J. E. (2003). Antigenic variation and the African trypanosome genome. Acta Tropica 85, 391-404.

Duncan, R. C., Salotra, P., Goyal, N., Akopyants, N. S., Beverley, S. M., and Nakhasi, H. L. (2004). The application of gene expression microarray technology to kinetoplastid research. Current Molecular Medicine 4, 611-21.

Ebbert, R., Birkmann, A., and Schüller, H. J. (1999). The product of the SNF2/SWI2 paralogue INO80 of Saccharomyces cerevisiae required for efficient expression of various yeast structural genes is part of a high-molecular-weight protein complex. Molecular Microbiology 32, 741-51.

Eid, J. E., and Sollner-Webb, B. (1991). Homologous recombination in the tandem calmodulin genes of Trypanosoma brucei yields multiple products: compensation for deleterious deletions by gene amplification. Genes & Development 5, 2024-32.

Eisen, J. A., Sweder, K. S., and Hanawalt, P. C. (1995). Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. Nucleic Acids Research 23, 2715-23.

El-Sayed, N. M., Hegde, P., Quackenbush, J., Melville, S. E., and Donelson, J. E. (2000). The African trypanosome genome. International Journal for Parasitology 30, 329-45.

Elfring, L. K., Deuring, R., McCallum, C. M., Peterson, C. L., and Tamkun, J. W. (1994). Identification and characterization of Drosophila relatives of the yeast transcriptional activator SNF2/SWI2. Molecular and Cellular Biology 14, 2225-34.

Erkina, T. Y., Zou, Y., Freeling, S., Vorobyev, V. I., and Erkine, A. M. (2010). Functional interplay between chromatin remodeling complexes RSC, SWI/SNF and ISWI in regulation of yeast heat shock genes. Nucleic Acids Research 38, 1441-9.

Ersfeld, K., and Gull, K. (1997). Partitioning of large and minichromosomes in Trypanosoma brucei. Science 276, 611-4.

Figueiredo, L. M., Cross, G. A., and Janzen, C. J. (2009). Epigenetic regulation in African trypanosomes: a new kid on the block. Nature Reviews. Microbiology 7, 504-13.

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

Figueiredo, L. M., and Cross, G. A. (2010). Nucleosomes are depleted at the VSG expression site transcribed by RNA polymerase I in African trypanosomes. Eukaryotic Cell 9, 148-54.

Friedlander, T. P., Regier, J. C., and Mitter, C. (1994). Phylogenetic Information Content of Five Nuclear Gene Sequences in Animals: Initial Assessment of Character Sets from Concordance and Divergence Studies. Systematic Biology 43, 511-525.

Furger, A., Schürch, N., Kurath, U., and Roditi, I. (1997). Elements in the 3' untranslated region of procyclin mRNA regulate expression in insect forms of Trypanosoma brucei by modulating RNA stability and translation. Molecular and Cellular Biology 17, 4372-80.

Geiman, T. M., and Robertson, K. D. (2002). Chromatin remodeling, histone modifications, and DNA methylation-how does it all fit together? Journal of Cellular Biochemistry 87, 117-25.

Giaever, G. et al. (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature 418, 387-91.

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

Glover, L., Alsford, S., Beattie, C., and Horn, D. (2007). Deletion of a trypanosome telomere leads to loss of silencing and progressive loss of terminal DNA in the absence of cell cycle arrest. Nucleic Acids Research 35, 872-80.

Glover, L., and Horn, D. (2006). Repression of polymerase I-mediated gene expression at Trypanosoma brucei telomeres. EMBO reports 7, 93-9.

Gommers-Ampt, J. H., Van Leeuwen, F., de Beer, A. L., Vliegenthart, J. F., Dizdaroglu, M., Kowalak, J. A., Crain, P. F., and Borst, P. (1993). beta-D-glucosyl-hydroxymethyluracil: a novel modified base present in the DNA of the parasitic protozoan T. brucei. Cell 75, 1129-36.

Gottschling, D. E., Aparicio, O. M., Billington, B. L., and Zakian, V. A. (1990). Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. Cell 63, 751-62.

Greaves, D. R., and Borst, P. (1987). Trypanosoma brucei variant-specific glycoprotein gene chromatin is sensitive to single-strand-specific endonuclease digestion. Journal of Molecular Biology 197, 471-83.

Grewal, S. I., and Jia, S. (2007). Heterochromatin revisited. Nature Reviews. Genetics 8, 35-46.

Grummt, I. (2007). Different epigenetic layers engage in complex crosstalk to define the epigenetic state of mammalian rRNA genes. Human Molecular Genetics 16 Spec No, R21-7.

Grüne, T., Brzeski, J., Eberharter, A., Clapier, C. R., Corona, D. F., Becker, P. B., and Müller, C. W. (2003). Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI. Molecular Cell 12, 449-60.

Günzl, A. et al. (1997). Transcription of the Trypanosoma brucei spliced leader RNA gene is dependent only on the presence of upstream regulatory elements. Molecular and Biochemical Parasitology 85, 67-76.

Günzl, A., Bruderer, T., Laufer, G., Schimanski, B., Tu, L., Chung, H., Lee, P., and Lee, M. G. (2003). RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in Trypanosoma brucei. Eukaryotic Cell 2, 542-51.

Günzl, A., Vanhamme, L., and Myler, P. J. (2007). Transcription in Trypanosomes: a Different Means to the End. In Trypanosomes: After The Genome D. Barry, R. McCulloch, J. Mottram, and A. Acosta-Serrano (Wymondham, UK: Horizon Bioscience), p. 186.

Hakimi, M., Bochar, D. A., Schmiesing, J. A., Dong, Y., Barak, O. G., Speicher, D. W., Yokomori, K., and Shiekhattar, R. (2002). A chromatin remodelling complex that loads cohesin onto human chromosomes. Nature 418, 994-8.

Hall, J. A., and Georgel, P. T. (2007). CHD proteins: a diverse family with strong ties. Biochemistry and Cell Biology 85, 463-76.

Hansen, J. C. (2002). Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. Annual Review of Biophysics and Biomolecular Structure 31, 361-92.

Hecker, H. (1995). Functional morphology of trypanosome chromatin. Parasitology Today 11, 79-83.

Hecker, H., Betschart, B., Bender, K., Burri, M., and Schlimme, W. (1994). The chromatin of trypanosomes. International Journal for Parasitology 24, 809-19.

Hecker, H., and Gander, E. S. (1985). The compaction pattern of the chromatin of trypanosomes. Biology of the cell / under the auspices of the European Cell Biology Organization 53, 199-208.

Hirumi, H., and Hirumi, K. (1989). Continuous cultivation of Trypanosoma brucei blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. The Journal of Parasitology 75, 985-9.

Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S., and Young, R. A. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. Cell 95, 717-28.

Horn, D., and Cross, G. A. (1995). A developmentally regulated position effect at a telomeric locus in Trypanosoma brucei. Cell 83, 555-61.

Horn, D., and Cross, G. A. (1997). Position-dependent and promoter-specific regulation of gene expression in Trypanosoma brucei. The EMBO Journal 16, 7422-31.

Horton, J. R., Elgar, S. J., Khan, S. I., Zhang, X., Wade, P. A., and Cheng, X. (2007). Structure of the SANT domain from the Xenopus chromatin remodeling factor ISWI. Proteins 67, 1198-202.

Hotz, H. R., Biebinger, S., Flaspohler, J., and Clayton, C. (1998). PARP gene expression: control at many levels. Molecular and Biochemical Parasitology 91, 131-43.

Hotz, H. R., Hartmann, C., Huober, K., Hug, M., and Clayton, C. (1997). Mechanisms of developmental regulation in Trypanosoma brucei: a polypyrimidine tract in the 3'-untranslated region of a surface protein mRNA affects RNA abundance and translation. Nucleic Acids Research 25, 3017-26.

Hug, M., Carruthers, V. B., Hartmann, C., Sherman, D. S., Cross, G. A., and Clayton, C. (1993). A possible role for the 3'-untranslated region in developmental regulation in Trypanosoma brucei. Molecular and Biochemical Parasitology 61, 87-95.

Hughes, K., Wand, M., Foulston, L., Young, R., Harley, K., Terry, S., Ersfeld, K., and Rudenko, G. (2007). A novel ISWI is involved in VSG expression site downregulation in African trypanosomes. The EMBO Journal 26, 2400-10.

Hur, S., Park, E., Han, J., Kim, Y., Kim, J., Kang, D., and Kwon, J. (2010). Roles of human INO80 chromatin remodeling enzyme in DNA replication and chromosome segregation suppress genome instability. Cellular and Molecular Life Sciences.

Ito, T., Bulger, M., Pazin, M. J., Kobayashi, R., and Kadonaga, J. T. (1997). ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. Cell 90, 145-55.

Iwata, A., and Consigli, R. A. (1971). Effect of phleomycin on polyoma virus synthesis in mouse embryo cells. Journal of Virology 7, 29-40.

Iyer, L. M., Anantharaman, V., Wolf, M. Y., and Aravind, L. (2008). Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. International Journal for Parasitology 38, 1-31.

Jackson, A. P. (2007). Tandem gene arrays in Trypanosoma brucei: comparative phylogenomic analysis of duplicate sequence variation. BMC Evolutionary Biology 7, 54.

Janz, L., and Clayton, C. (1994). The PARP and rRNA promoters of Trypanosoma brucei are composed of dissimilar sequence elements that are functionally interchangeable. Molecular and Cellular Biology 14, 5804-11.

Janzen, C. J., Fernandez, J. P., Deng, H., Diaz, R., Hake, S. B., and Cross, G. A. (2006). Unusual histone modifications in Trypanosoma brucei. FEBS Letters 580, 2306-10.

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

Janzen, C. J., Lander, F., Dreesen, O., and Cross, G. A. (2004). Telomere length regulation and transcriptional silencing in KU80-deficient Trypanosoma brucei. Nucleic Acids Research 32, 6575-84.

Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. Science 293, 1074-80.

Jiang, C., and Pugh, B. F. (2009). Nucleosome positioning and gene regulation: advances through genomics. Nature Reviews. Genetics 10, 161-72.

Kamath, R. S. et al. (2003). Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature 421, 231-7.

Kelley, L. A., and Sternberg, M. J. (2009). Protein structure prediction on the Web: a case study using the Phyre server. Nature Protocols 4, 363-71.

Kikyo, N., Wade, P. A., Guschin, D., Ge, H., and Wolffe, A. P. (2000). Active remodeling of somatic nuclei in egg cytoplasm by the nucleosomal ATPase ISWI. Science 289, 2360-2.

Klose, R. J., and Bird, A. P. (2006). Genomic DNA methylation: the mark and its mediators. Trends in Biochemical Sciences 31, 89-97.

Kornberg, R. D. (1977). Structure of chromatin. Annual Review of Biochemistry 46, 931-54.

Kornberg, R. D., and Lorch, Y. (1991). Irresistible force meets immovable object: transcription and the nucleosome. Cell 67, 833-6.

Kornberg, R. D., and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell 98, 285-94.

Koumandou, V. L., Natesan, S. K., Sergeenko, T., and Field, M. C. (2008). The trypanosome transcriptome is remodelled during differentiation but displays limited responsiveness within life stages. BMC Genomics 9, 298.

Krogan, N. J. et al. (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Molecular Cell 12, 1565-76.

LaCount, D. J., Barrett, B., and Donelson, J. E. (2002). Trypanosoma brucei FLA1 is required for flagellum attachment and cytokinesis. The Journal of Biological Chemistry 277, 17580-8.

Landeira, D., Bart, J., Van Tyne, D., and Navarro, M. (2009). Cohesin regulates VSG monoallelic expression in trypanosomes. The Journal of Cell Biology 186, 243-54.

Landeira, D., and Navarro, M. (2007). Nuclear repositioning of the VSG promoter during developmental silencing in Trypanosoma brucei. The Journal of Cell Biology 176, 133-9.

Larkin, M. A. et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947-8.

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

Le, S. Q., and Gascuel, O. (2008). An improved general amino acid replacement matrix. Molecular Biology and Evolution 25, 1307-20.

Lee, M. G., and Van Der Ploeg, L. H. (1997). Transcription of protein-coding genes in trypanosomes by RNA polymerase I. Annual Review of Microbiology 51, 463-89.

Li, J., Santoro, R., Koberna, K., and Grummt, I. (2005). The chromatin remodeling complex NoRC controls replication timing of rRNA genes. The EMBO Journal 24, 120-7.

Liniger, M., Bodenmüller, K., Pays, E., Gallati, S., and Roditi, I. (2001). Overlapping sense and antisense transcription units in Trypanosoma brucei. Molecular Microbiology 40, 869-78.

Liu, A. Y., Van Der Ploeg, L. H., Rijsewijk, F. A., and Borst, P. (1983). The transposition unit of variant surface glycoprotein gene 118 of Trypanosoma brucei. Presence of repeated elements at its border and absence of promoter-associated sequences. Journal of Molecular Biology 167, 57-75.

Lowell, J. E., Kaiser, F., Janzen, C. J., and Cross, G. A. (2005). Histone H2AZ dimerizes with a novel variant H2B and is enriched at repetitive DNA in Trypanosoma brucei. Journal of Cell Science 118, 5721-30.

Lowell, J. E., and Cross, G. A. (2004). A variant histone H3 is enriched at telomeres in Trypanosoma brucei. Journal of Cell Science 117, 5937-47.

Luger, K., Mäder, 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, 251-60.

Lukes, J., Jirkû, M., Dolezel, D., Kral'ová, I., Hollar, L., and Maslov, D. A. (1997). Analysis of ribosomal RNA genes suggests that trypanosomes are monophyletic. Journal of Molecular Evolution 44, 521-7.

Luo, H., Gilinger, G., Mukherjee, D., and Bellofatto, V. (1999). Transcription initiation at the TATA-less spliced leader RNA gene promoter requires at least two DNA-binding proteins and a tripartite architecture that includes an initiator element. The Journal of Biological Chemistry 274, 31947-54.

Längst, G., and Becker, P. B. (2001). Nucleosome mobilization and positioning by ISWI-containing chromatin-remodeling factors. Journal of Cell Science 114, 2561-8.

Mandava, V., Fernandez, J. P., Deng, H., Janzen, C. J., Hake, S. B., and Cross, G. A. (2007). Histone modifications in Trypanosoma brucei. Molecular and Biochemical Parasitology 156, 41-50.

Martens, J. A., and Winston, F. (2003). Recent advances in understanding chromatin remodeling by Swi/Snf complexes. Current Opinion in Genetics & Development 13, 136-42.

Matthews, K. R. (1999). Developments in the differentiation of Trypanosoma brucei. Parasitology Today 15, 76-80.

McKean, P. (2003). Coordination of cell cycle and cytokinesis in Trypanosoma brucei. Current Opinion in Microbiology 6, 600-607.

Mellor, J., and Morillon, A. (2004). ISWI complexes in Saccharomyces cerevisiae. Biochimica et Biophysica Acta 1677, 100-12.

Melville, S. E., Leech, V., Gerrard, C. S., Tait, A., and Blackwell, J. M. (1998). The molecular karyotype of the megabase chromosomes of Trypanosoma brucei and the assignment of chromosome markers. Molecular and Biochemical Parasitology 94, 155-73.

Merz, K., Hondele, M., Goetze, H., Gmelch, K., Stoeckl, U., and Griesenbeck, J. (2008). 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 & Development 22, 1190-204.

Militello, K. T., Wang, P., Jayakar, S. K., Pietrasik, R. L., Dupont, C. D., Dodd, K., King, A. M., and Valenti, P. R. (2008). African trypanosomes contain 5-methylcytosine in nuclear DNA. Eukaryotic Cell 7, 2012-6.

Mizuguchi, G., Shen, X., Landry, J., Wu, W., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science 303, 343-8.

Mohrmann, L., and Verrijzer, C. P. (2005). Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. Biochimica et Biophysica Acta 1681, 59-73.

Moreau, J., Lee, M., Mahachi, N., Vary, J., Mellor, J., Tsukiyama, T., and Goding, C. R. (2003). Regulated displacement of TBP from the PHO8 promoter in vivo requires Cbf1 and the Isw1 chromatin remodeling complex. Molecular Cell 11, 1609-20.

Morillon, A., Karabetsou, N., O'Sullivan, J., Kent, N., Proudfoot, N., and Mellor, J. (2003). Isw1 chromatin remodeling ATPase coordinates transcription elongation and termination by RNA polymerase II. Cell 115, 425-35.

Morrison, A. J., and Shen, X. (2009). Chromatin remodelling beyond transcription: the INO80 and SWR1 complexes. Nature Reviews. Molecular Cell Biology 10, 373-84.

Morse, R. H. (2000). RAP, RAP, open up! New wrinkles for RAP1 in yeast. Trends in Genetics 16, 51-3.

Muchardt, C., and Yaniv, M. (1999). ATP-dependent chromatin remodelling: SWI/SNF and Co. are on the job. Journal of Molecular Biology 293, 187-98.

Narlikar, G. J., Fan, H., and Kingston, R. E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. Cell 108, 475-87.

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. The EMBO Journal 18, 2265-72.

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

Neigeborn, L., and Carlson, M. (1984). Genes affecting the regulation of SUC2 gene expression by glucose repression in Saccharomyces cerevisiae. Genetics 108, 845-58.

Okada, M., Okawa, K., Isobe, T., and Fukagawa, T. (2009). CENP-H-containing complex facilitates centromere deposition of CENP-A in cooperation with FACT and CHD1. Molecular Biology of the Cell 20, 3986-95.

Osir, E. O., Abubakar, L., and Imbuga, M. O. (1995). Purification and characterization of a midgut lectin-trypsin complex from the tsetse fly Glossina longipennis. Parasitology research 81, 276-81.

Owen-Hughes, T. (2003). Colworth memorial lecture. Pathways for remodelling chromatin. Biochemical Society Transactions 31, 893-905.

Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences : CABIOS 12, 357-8.

Palenchar, J. B., and Bellofatto, V. (2006). Gene transcription in trypanosomes. Molecular and Biochemical Parasitology 146, 135-41.

Parodi-Talice, A., Durán, R., Arrambide, N., Prieto, V., Piñeyro, M. D., Pritsch, O., Cayota, A., Cerveñansky, C., and Robello, C. (2004). Proteome analysis of the causative agent of Chagas disease: Trypanosoma cruzi. International Journal for Parasitology 34, 881-6.

Pays, E., Coquelet, H., Tebabi, P., Pays, A., Jefferies, D., Steinert, M., Koenig, E., Williams, R. O., and Roditi, I. (1990). Trypanosoma brucei: constitutive activity of the VSG and procyclin gene promoters. The EMBO Journal 9, 3145-51.

Pays, E., Lips, S., Nolan, D., Vanhamme, L., and Pérez-Morga, D. (2001). The VSG expression sites of Trypanosoma brucei: multipurpose tools for the adaptation of the parasite to mammalian hosts. Molecular and Biochemical Parasitology 114, 1-16.

Pearson, T. (2001). Procyclins, proteases and proteomics: dissecting trypanosomes in the tsetse fly. Trends in Microbiology 9, 299-301.

Pham, V. P., Qi, C. C., and Gottesdiener, K. M. (1996). A detailed mutational analysis of the VSG gene expression site promoter. Molecular and Biochemical Parasitology 75, 241-54.

Pietsch, P., and Garrett, H. (1968). Primary site of reaction in the in vitro complex of phleomycin in DNA. Nature 219, 488-9.

Pinskaya, M., Nair, A., Clynes, D., Morillon, A., and Mellor, J. (2009). Nucleosome remodeling and transcriptional repression are distinct functions of Isw1 in Saccharomyces cerevisiae. Molecular and Cellular Biology 29, 2419-30.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., and Lopez, R. (2005). InterProScan: protein domains identifier. Nucleic Acids Research 33, W116-20.

Reeves, J. H. (1992). Heterogeneity in the substitution process of amino acid sites of proteins coded for by mitochondrial DNA. Journal of Molecular Evolution 35, 17-31.

Renauld, H., Kelly, J. M., and Horn, D. (2007). Chromosome Structure and Dynamics. In Trypanosomes: After The Genome D. Barry, R. McCulloch, J. Mottram, and A. Acosta-Serrano (Wymondham, UK: Horizon Bioscience), p. 116.

Richmond, E., and Peterson, C. L. (1996). Functional analysis of the DNA-stimulated ATPase domain of yeast SWI2/SNF2. Nucleic Acids Research 24, 3685-92.

Robinson, N. P., Burman, N., Melville, S. E., and Barry, J. D. (1999). Predominance of duplicative VSG gene conversion in antigenic variation in African trypanosomes. Molecular and Cellular Biology 19, 5839-46.

Rudenko, G., Blundell, P. A., Taylor, M. C., Kieft, R., and Borst, P. (1994). VSG gene expression site control in insect form Trypanosoma brucei. The EMBO Journal 13, 5470-82.

Saha, A., Wittmeyer, J., and Cairns, B. R. (2006). Chromatin remodelling: the industrial revolution of DNA around histones. Nature Reviews. Molecular Cell Biology 7, 437-47.

Sambrook, J., and Russell, D. (2001). Molecular Cloning: A Laboratory Manual 3 ed. (Cold Spring Harbor Laboratory Press, New York, U.S.A.).

Sandhu, K. S. (2009). Intrinsic disorder explains diverse nuclear roles of chromatin remodeling proteins. Journal of Molecular Recognition 22, 1-8.

Santoro, R., Li, J., and Grummt, I. (2002). The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. Nature Genetics 32, 393-6.

Sawa, H., Kouike, H., and Okano, H. (2000). Components of the SWI/SNF complex are required for asymmetric cell division in C. elegans. Molecular Cell 6, 617-24.

Schaufler, L. E., and Klevit, R. E. (2003). Mechanism of DNA binding by the ADR1 zinc finger transcription factor as determined by SPR. Journal of Molecular Biology 329, 931-9.

Sheader, K., Berberof, M., Isobe, T., Borst, P., and Rudenko, G. (2003). Delineation of the regulated Variant Surface Glycoprotein gene expression site domain of Trypanosoma brucei. Molecular and Biochemical Parasitology 128, 147-56.

Sheader, K., te Vruchte, D., and Rudenko, G. (2004). Bloodstream form-specific up-regulation of silent vsg expression sites and procyclin in Trypanosoma brucei after inhibition of DNA synthesis or DNA damage. The Journal of Biological Chemistry 279, 13363-74.

Shen, X., Mizuguchi, G., Hamiche, A., and Wu, C. (2000). A chromatin remodelling complex involved in transcription and DNA processing. Nature 406, 541-4.

Shen, X., Ranallo, R., Choi, E., and Wu, C. (2003). Involvement of actin-related proteins in ATP-dependent chromatin remodeling. Molecular Cell 12, 147-55.

Shur, I., Socher, R., and Benayahu, D. (2006). In vivo association of CReMM/CHD9 with promoters in osteogenic cells. Journal of Cellular Physiology 207, 374-8.

Sidow, A., and Thomas, W. K. (1994). A molecular evolutionary framework for eukaryotic model organisms. Current Biology 4, 596-603.

Siegel, T. N., Hekstra, D. R., Kemp, L. E., Figueiredo, L. M., Lowell, J. E., Fenyo, D., Wang, X., Dewell, S., and Cross, G. A. (2009). Four histone variants mark the

boundaries of polycistronic transcription units in Trypanosoma brucei. Genes & Development 23, 1063-76.

Simpson, A. G., Stevens, J. R., and Lukes, J. (2006). The evolution and diversity of kinetoplastid flagellates. Trends in Parasitology 22, 168-74.

Sims, R. J., Chen, C., Santos-Rosa, H., Kouzarides, T., Patel, S. S., and Reinberg, D. (2005). Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. The Journal of Biological Chemistry 280, 41789-92.

Srinivasan, S., Armstrong, J. A., Deuring, R., Dahlsveen, I. K., McNeill, H., and Tamkun, J. W. (2005). The Drosophila trithorax group protein Kismet facilitates an early step in transcriptional elongation by RNA Polymerase II. Development 132, 1623-35.

Stanne, T. M., and Rudenko, G. (2010). Active VSG expression sites in Trypanosoma brucei are depleted of nucleosomes. Eukaryotic Cell 9, 136-47.

Stern, M., Jensen, R., and Herskowitz, I. (1984). Five SWI genes are required for expression of the HO gene in yeast. Journal of Molecular Biology 178, 853-68.

Stiles, J. K., Ingram, G. A., Wallbanks, K. R., Molyneux, D. H., Maudlin, I., and Welburn, S. (1990). Identification of midgut trypanolysin and trypanoagglutinin in Glossina palpalis sspp. (Diptera: Glossinidae). Parasitology 101 Pt 3, 369-76.

Söding, J., Biegert, A., and Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Research 33, W244-8.

Sönnichsen, B. et al. (2005). Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. Nature 434, 462-9.

Taddei, A., Hediger, F., Neumann, F. R., and Gasser, S. M. (2004). The function of nuclear architecture: a genetic approach. Annual Review of Genetics 38, 305-45.

Tai, H. H., Geisterfer, M., Bell, J. C., Moniwa, M., Davie, J. R., Boucher, L., and McBurney, M. W. (2003). CHD1 associates with NCoR and histone deacetylase as well as with RNA splicing proteins. Biochemical and Biophysical Research Communications 308, 170-6.

Tamkun, J. W., Deuring, R., Scott, M. P., Kissinger, M., Pattatucci, A. M., Kaufman, T. C., and Kennison, J. A. (1992). brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. Cell 68, 561-72.

Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., and Patel, D. J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nature Structural & Molecular Biology 14, 1025-40.

Thatcher, T. H., and Gorovsky, M. A. (1994). Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. Nucleic Acids Research 22, 174-9.

Thompson, P. M., Gotoh, T., Kok, M., White, P. S., and Brodeur, G. M. (2003). CHD5, a new member of the chromodomain gene family, is preferentially expressed in the nervous system. Oncogene 22, 1002-11.

Tsukuda, T., Fleming, A. B., Nickoloff, J. A., and Osley, M. A. (2005). Chromatin remodelling at a DNA double-strand break site in Saccharomyces cerevisiae. Nature 438, 379-83.

Ulbert, S., Chaves, I., and Borst, P. (2002). Expression site activation in Trypanosoma brucei with three marked variant surface glycoprotein gene expression sites. Molecular and Biochemical Parasitology 120, 225-35.

Vanhamme, L., Berberof, M., Le Ray, D., and Pays, E. (1995). Stimuli of differentiation regulate RNA elongation in the transcription units for the major stage-specific antigens of Trypanosoma brucei. Nucleic Acids Research 23, 1862-9.

Vanhamme, L., Pays, A., Tebabi, P., Alexandre, S., and Pays, E. (1995). Specific binding of proteins to the noncoding strand of a crucial element of the variant surface glycoprotein, procyclin, and ribosomal promoters of trypanosoma brucei. Molecular and Cellular Biology 15, 5598-606.

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. Molecular Microbiology 36, 328-40.

Vanhamme, L., and Pays, E. (1998). Controls of the expression of the Vsg in Trypanosoma brucei. Molecular and Biochemical Parasitology 91, 107-16.

Varga-Weisz, P. D., Wilm, M., Bonte, E., Dumas, K., Mann, M., and Becker, P. B. (1997). Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. Nature 388, 598-602.

Vassella, E., Acosta-Serrano, A., Studer, E., Lee, S. H., Englund, P. T., and Roditi, I. (2001). Multiple procyclin isoforms are expressed differentially during the development of insect forms of Trypanosoma brucei. Journal of Molecular Biology 312, 597-607.

Vertommen, D., Van Roy, J., Szikora, J., Rider, M. H., Michels, P. A., and Opperdoes, F. R. (2008). Differential expression of glycosomal and mitochondrial proteins in the two major life-cycle stages of Trypanosoma brucei. Molecular and Biochemical Parasitology 158, 189-201.

Vignali, M., Hassan, A. H., Neely, K. E., and Workman, J. L. (2000). ATP-dependent chromatin-remodeling complexes. Molecular and Cellular Biology 20, 1899-910.

Wagner, W., and So, M. (1990). Genomic variation of Trypanosoma cruzi: involvement of multicopy genes. Infection and immunity 58, 3217-24.

Wang, W. (2003). The SWI/SNF family of ATP-dependent chromatin remodelers: similar mechanisms for diverse functions. Current Topics in Microbiology and Immunology 274, 143-69.

Weiden, M., Osheim, Y. N., Beyer, A. L., and Van Der Ploeg, L. H. (1991). Chromosome structure: DNA nucleotide sequence elements of a subset of the minichromosomes of the protozoan Trypanosoma brucei. Molecular and Cellular Biology 11, 3823-34.

White, T. C., Rudenko, G., and Borst, P. (1986). Three small RNAs within the 10 kb trypanosome rRNA transcription unit are analogous to domain VII of other eukaryotic 28S rRNAs. Nucleic Acids Research 14, 9471-89.

Whitehouse, I., Flaus, A., Cairns, B. R., White, M. F., Workman, J. L., and Owen-Hughes, T. (1999). Nucleosome mobilization catalysed by the yeast SWI/SNF complex. Nature 400, 784-7.

Wickstead, B., Ersfeld, K., and Gull, K. (2002). Targeting of a tetracycline-inducible expression system to the transcriptionally silent minichromosomes of Trypanosoma brucei. Molecular and Biochemical Parasitology 125, 211-6.

Wickstead, B., Ersfeld, K., and Gull, K. (2003). The frequency of gene targeting in Trypanosoma brucei is independent of target site copy number. Nucleic Acids Research 31, 3993-4000.

Wickstead, B., Ersfeld, K., and Gull, K. (2004). The small chromosomes of Trypanosoma brucei involved in antigenic variation are constructed around repetitive palindromes. Genome Research 14, 1014-24.

Winston, F., and Allis, C. D. (1999). The bromodomain: a chromatin-targeting module? Nature Structural Biology 6, 601-4.

Winston, F., and Carlson, M. (1992). Yeast SNF/SWI transcriptional activators and the SPT/SIN chromatin connection. Trends in Genetics 8, 387-91.

Wirtz, E., Hartmann, C., and Clayton, C. (1994). Gene expression mediated by bacteriophage T3 and T7 RNA polymerases in transgenic trypanosomes. Nucleic Acids Research 22, 3887-94.

Wirtz, E., Leal, S., Ochatt, C., and Cross, G. A. (1999). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. Molecular and Biochemical Parasitology 99, 89-101.

Woodage, T., Basrai, M. A., Baxevanis, A. D., Hieter, P., and Collins, F. S. (1997). Characterization of the CHD family of proteins. Proceedings of the National Academy of Sciences of the United States of America 94, 11472-7.

World Health Organization (1998). Control and Surveillance of African Trypanosomiasis 881 ed. (Technical Report Series).

Wu, W., Alami, S., Luk, E., Wu, C., Sen, S., Mizuguchi, G., Wei, D., and Wu, C. (2005). Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. Nature Structural & Molecular Biology 12, 1064-71.

Yang, J., Swaminathan, C. P., Huang, Y., Guan, R., Cho, S., Kieke, M. C., Kranz, D. M., Mariuzza, R. A., and Sundberg, E. J. (2003). Dissecting cooperative and additive binding energetics in the affinity maturation pathway of a protein-protein interface. The Journal of Biological Chemistry 278, 50412-21.

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.

Yang, Z. (1993). Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. Molecular Biology and Evolution 10, 1396-401.

Young, R., Taylor, J. E., Kurioka, A., Becker, M., Louis, E. J., and Rudenko, G. (2008). Isolation and analysis of the genetic diversity of repertoires of VSG expression site containing telomeres from Trypanosoma brucei gambiense, T. b. brucei and T. equiperdum. BMC Genomics 9, 385.

Zlatanova, J., Bishop, T. C., Victor, J., Jackson, V., and van Holde, K. (2009). The nucleosome family: dynamic and growing. Structure 17, 160-71.

Zomerdijk, J. C., Kieft, R., Shiels, P. G., and Borst, P. (1991). Alpha-amanitin-resistant transcription units in trypanosomes: a comparison of promoter sequences for a VSG gene expression site and for the ribosomal RNA genes. Nucleic Acids Research 19, 5153-8.

Zomerdijk, J. C., Ouellette, M., ten Asbroek, A. L., Kieft, R., Bommer, A. M., Clayton, C. E., and Borst, P. (1990). The promoter for a variant surface glycoprotein gene expression site in Trypanosoma brucei. The EMBO Journal 9, 2791-801.

van Attikum, H., Fritsch, O., and Gasser, S. M. (2007). Distinct roles for SWR1 and INO80 chromatin remodeling complexes at chromosomal double-strand breaks. The EMBO Journal 26, 4113-25.

van Leeuwen, F., Wijsman, E. R., Kieft, R., van Der Marel, G. A., van Boom, J. H., and Borst, P. (1997). Localization of the modified base J in telomeric VSG gene expression sites of Trypanosoma brucei. Genes & Development 11, 3232-41.

Appendices

A1. Opposing T7 promoters

While planning the reporter constructs for the various reporter cell lines integrating into the minichromosomes, the VSG basic copy arrays, the rDNA spacer and the *EP1* procyclin locus (see Chapters 3 and 4), the presence of a T7 promoter in the Bluescript plasmid backbone was overlooked. Figure A1.1 shows the location of the T7 promoter in the reporter plasmid and the reporter construct integrated in the VSG BCA.

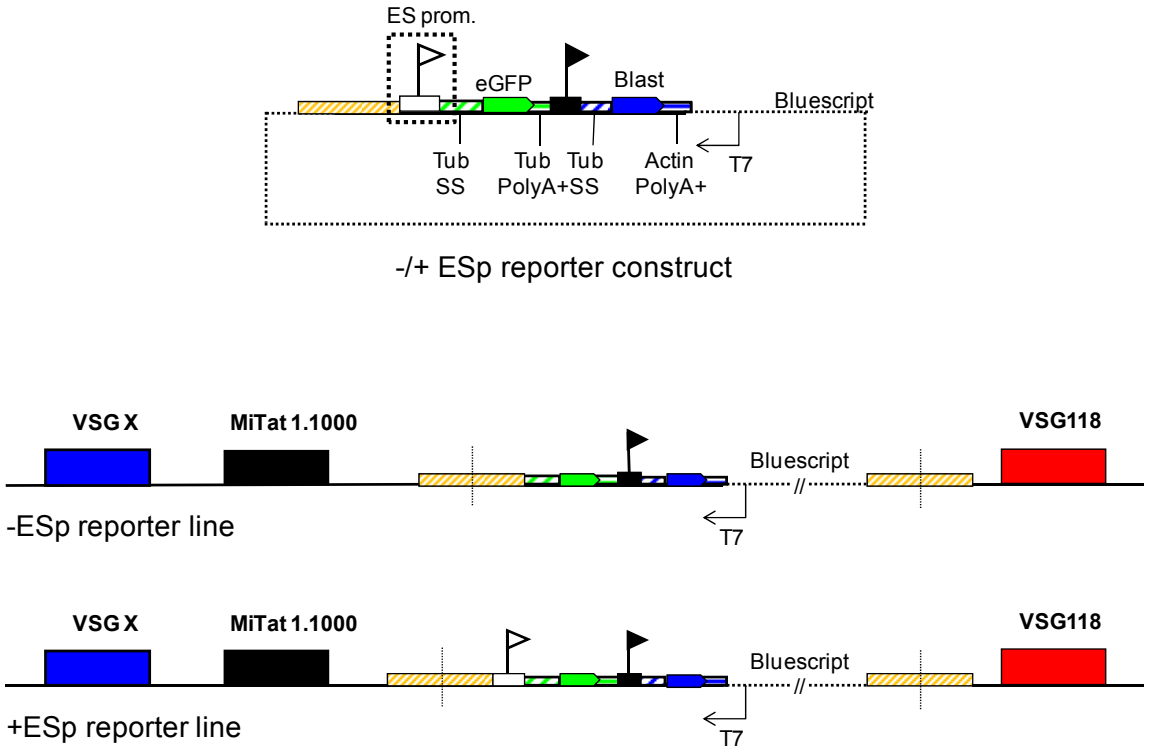


Figure A1.1 Location of the Opposing T7 promoter in VSG basic copy array targeting reporter constructs. Top panel: shows a schematic of the VSG basic copy array targeting reporter construct (see Section 3.2) indicating the position of the opposing T7 promoter. Bottom panel: shows the location of the opposing T7 promoter in the –ESp and the +ESp reporter constructs integrated into the VSG BCA.

As the parental cell lines used for reporter construct integration, S16.221 and 29-13, have a constitutively expressed T7 RNA polymerase, the Bluescript T7 promoter can be expected to constitutively transcribe in the reverse direction, possibly causing

antisense transcription of the *Blasticidin* gene and possibly even the *eGFP* gene since there is no T7 terminator in the reporter construct.

In order to test whether this was interfering with the results, alternative reporter constructs targeting the VSG BCA were made. The new strategy involved the insertion of a downstream VSG BCA targeting fragment into the original reporter construct (Figure 3.2.1) downstream of the *Blasticidin* resistance gene cassette and then targeting the construct by a double crossover event (Figure A1.2). The strategy was adopted similarly for the +ESp reporter construct. Using the new strategy did not result in the integration of the Bluescript backbone into the targeted locus.

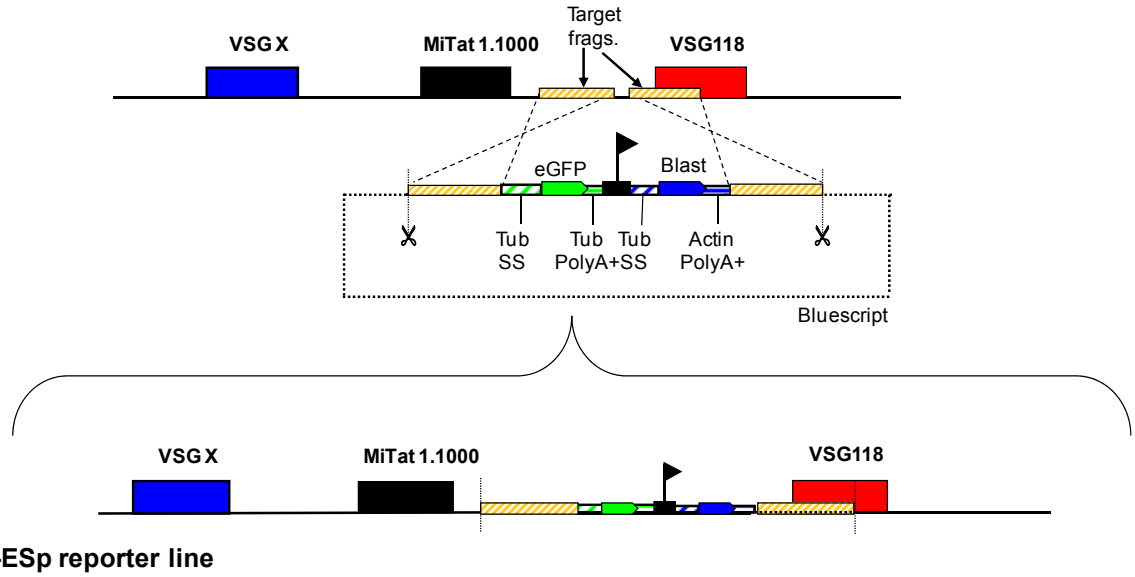


Figure A1.2 Alternative strategy for reporter construct integration into the VSG basic copy arrays. The schematic shows the alternative strategy for the integration of the –ESp eGFP reporter construct into the VSG basic copy array (compare with Figure 3.2.1). The alternative strategy requires the use of two targeting fragments that facilitate the integration of the construct by a double crossover event. The construct was digested at the sites marked by scissors before transfection.

The correct integration of the reporter constructs was verified by PCR reactions (data not shown). RNAi constructs were transfected into the reporter cell line, and the effect of TbISWI depletion on reporter expression compared with that from cells made

using the original strategy. The left panel in Figure A1.3 compares the results from the RNAi cell line BF_VBe1 obtained using the original strategy and RNAi cell lines BF_Vbe-T7.1 and BF_Vbe-T7.2 obtained using the alternative new strategy. Similarly, the right panel in Figure A1.3 compares the results from the RNAi cell line BF_VBEP1 obtained using the original strategy and RNAi cell lines BF_VBEP-T7.1 and BF_VBEP-T7.2 obtained using the alternative new strategy. Results for the PF cells are compared in Figure A1.4.

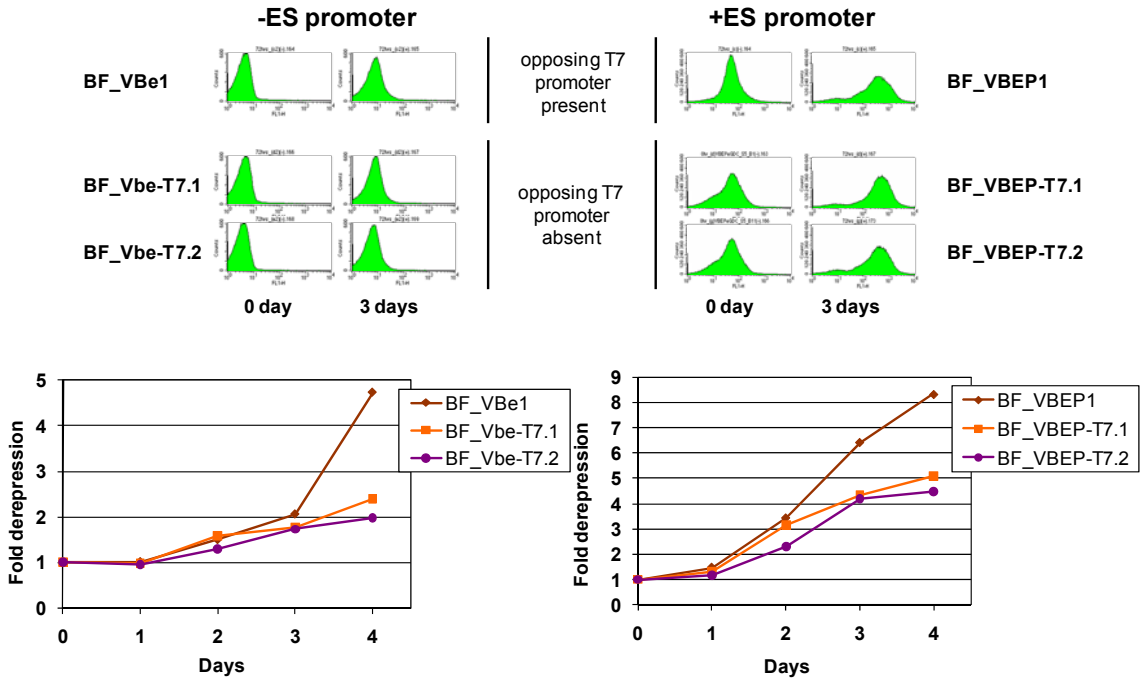


Figure A1.3 Comparing the two strategies of reporter integration by monitoring eGFP derepression after induction of RNAi against TbISWI in bloodstream form VSG basic copy array reporter cell lines. Top: representative flow cytometry traces for three cell lines at two different time-points are shown (day 0 and day 3), with eGFP fluorescence monitored in the FL1 channel (x-axis). Bottom: the graph shows the degree of derepression of the *eGFP* gene for different durations of tetracycline induced *TbISWI* RNAi. The degree of derepression was calculated as the ratio of mean fluorescence of induced cells to that of the uninduced cells.

Comparing the results from RNAi lines using the old strategy versus the new strategy reveals that although the general trend of derepression is similar in both –ESp and +ESp reporter lines in both life cycle stages, the degree of derepression is higher using the original strategy that had an oppositely directed T7 promoter. Looking at the

FACS traces from the RNAi lines it appears that the presence of the T7 promoter had been reducing low levels of transcription possibly through the production of antisense transcripts of the *eGFP*, resulting in its degradation by RNAi.

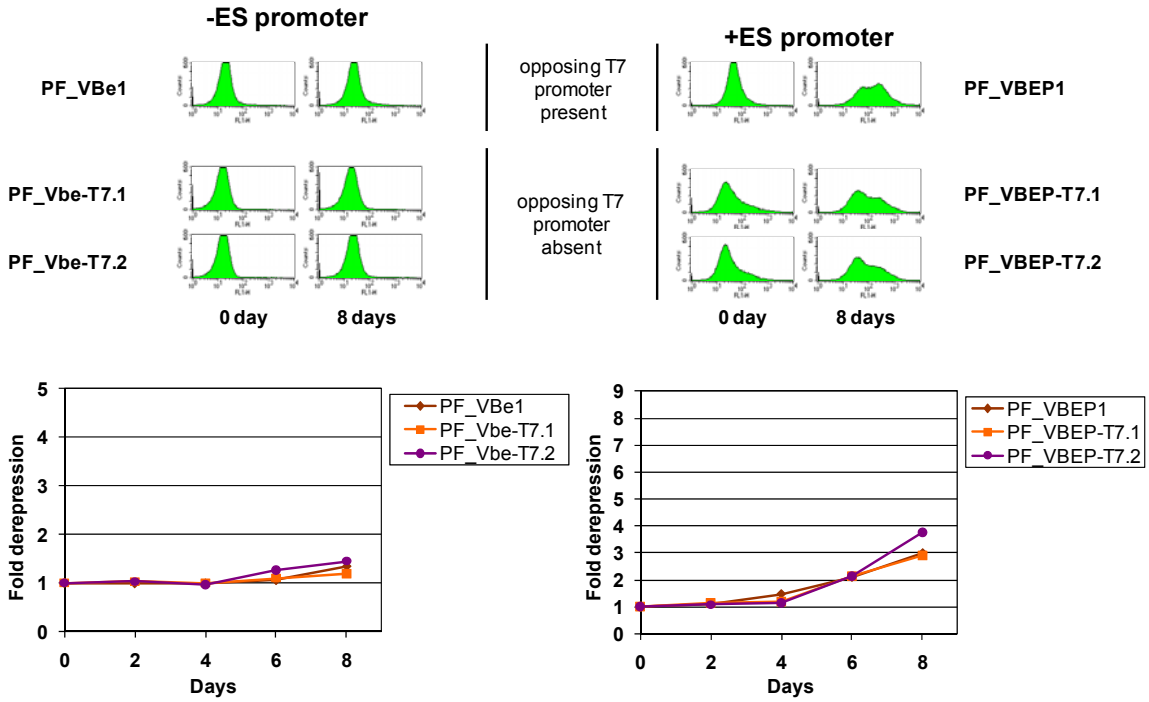


Figure A1.4 Comparing the two strategies of reporter integration by monitoring eGFP derepression after induction of RNAi against TbISWI in procyclic form VSG basic copy array reporter cell lines. Top: representative flow cytometry traces for three cell lines at two different time-points are shown (day 0 and day 8), with eGFP fluorescence monitored in the FL1 channel (x-axis). Bottom: the graph shows the degree of derepression of the *eGFP* gene for different durations of tetracycline induced *TbISWI* RNAi. Fold derepression was calculated as the ratio of mean fluorescence of induced cells to that of the uninduced cells.

In order to modify the minichromosome targeting reporter constructs, unique restriction enzymes were used to cut on either side of the Bluescript T7 promoter to excise it out. The digested ends of the reporter construct were then blunted and resealed. As before, the construct was linearised for transfection by using NotI. Results (data not shown; obtained with help from Alexander Fullbrook) from the RNAi lines obtained using the two strategies were compared. Similar to the observations made in the VSG BCA, the trend of derepression was similar using both the strategies, while the degree of derepression was higher using the original strategy.

A2. TbISWI binds the silent repeat regions

Chromatin immunoprecipitation (ChIP) was performed on procyclic *T. brucei* cell lines expressing wild type and myc-tagged TbISWI as described in the Materials and Methods. The densities of the slotblot signals was quantified by Quantity One software and plotted on the graph as % of input immunoprecipitated.

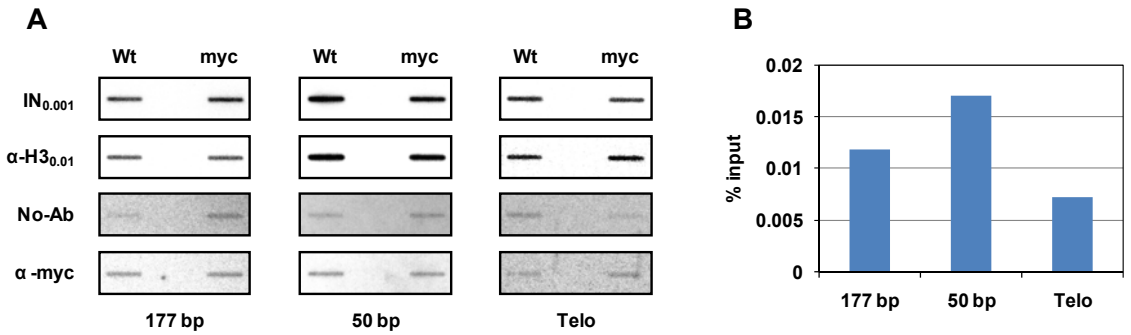


Figure A2.1 TbISWI binds silent repeat regions. Chromatin immunoprecipitation (ChIP) was performed on procyclic form cells expressing wild type (Wt) and myc-tagged (myc) TbISWI. (A) For both the Wt and myc cell lines, slot blots were loaded with 0.01% of the total input material ($IN_{0.001}$), 0.1% of the total material precipitated by α-H3 antibody ($\alpha\text{-}H3_{0.01}$), the material precipitated using no antibody (No-Ab) and the material precipitated by α-myc antibody. The blots obtained by slot blotting were probed with radioactive probes for the 177 bp repeats, the 50 bp repeats and the telomeric repeats. (B) Quantification of the signal intensity from the radioactive blots was used to determine the % input immunoprecipitated by first normalising the intensity values to total input in the two cell lines (Wt and myc) and then subtracting the Wt signal from the myc signal.

The results show that TbISWI binds all of the minichromosomeal 177 bp repeats, the 50 bp repeats and the telomeric repeats in PF *T. brucei*. The abundance of TbISWI in all these silent regions is very low as compared to that seen for histone H3 (Stanne et al. 2010). However, it is comparable across the three repeat regions, suggesting that TbISWI may have a role in silencing the other repeat regions similar to the minichromosomes. Results obtained independently by Tara Stanne (unpublished data) confirm that TbISWI binds ubiquitously in the *T. brucei* genome, including the silent repeat arrays probed for above.

A3. TbISWI is an essential protein

RNAi mediated depletion of TbISWI results in derepression of silent expression sites (Hughes et al. 2007). This downregulation of TbISWI results in a growth reduction but it was not clear if TbISWI was completely essential. In order to determine if TbISWI is an essential protein, constructs were made to knock out TbISWI by replacing the native gene with a drug resistance marker. To provide convincing evidence that TbISIWI is non-essential, both alleles need to be targeted. Therefore, two different knockout constructs, one with a hygromycin resistance gene (pSpot5KOHA57) and the other with a phleomycin resistance gene (pSpot5KOPhleo) resistance gene were made (Figures A3.1 and A3.2).

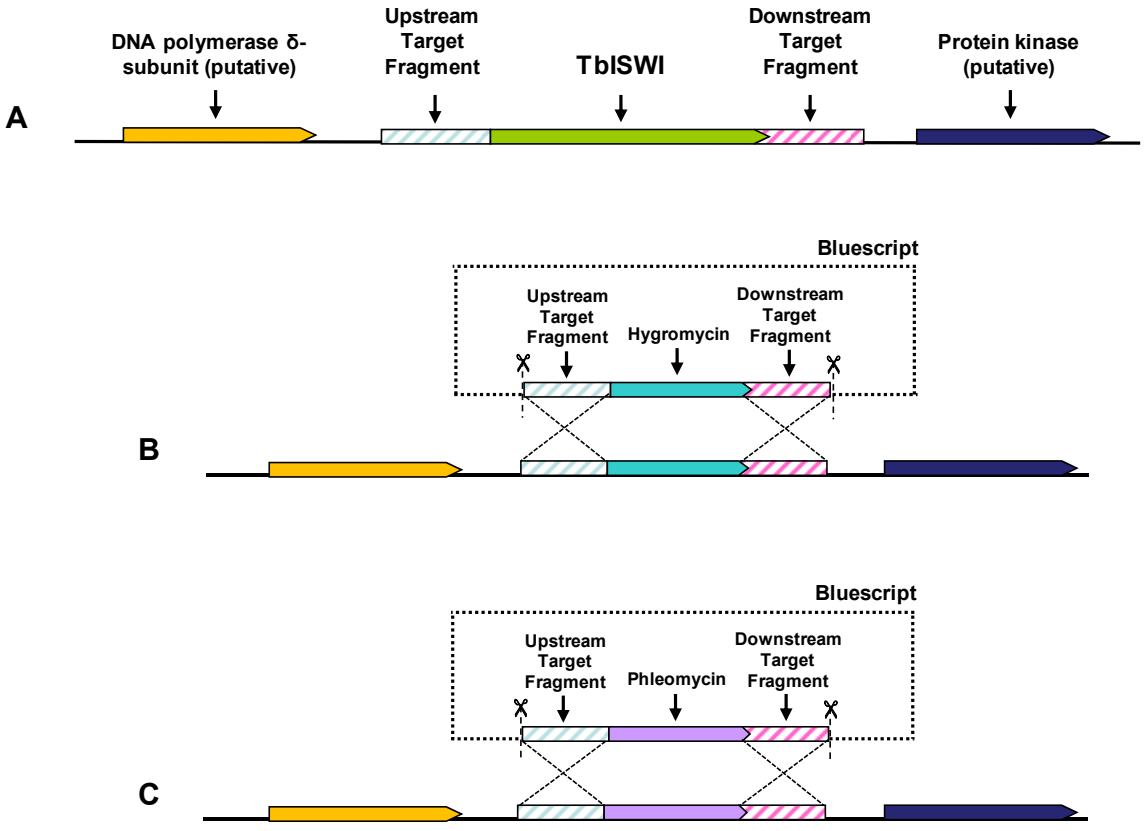


Figure A3.1 Making single-allele genetic knockouts of TbISWI. Schematic representations of (A) the native *TbISWI* locus, (B) the TbISWI locus after replacement with a *Hygromycin* resistance gene using the pSpot5KOHA57 knockout construct, (C) the TbISWI locus after replacement with a *Phleomycin* resistance gene using the pSpot5KOPhleo knockout construct. The sites of digestion of the knockout constructs are indicated by scissors.

The knockout constructs were then transfected into BF *T. brucei* 221GP1(VO2+) (Sheader et al. 2004) in order to obtain single-allele knockouts of *TbISWI* as schematically represented in Figure A3.1. Genomic DNA from the resulting cell lines (TbISWI_KO_Hyg and TbISWI_KO_Phleo) was then used to verify the correct integration of the constructs by PCR (data not shown). This proved that single-allele knockouts of TbISWI are obtainable, suggesting that the expression of the protein from a single locus is sufficient for normal cell functioning.

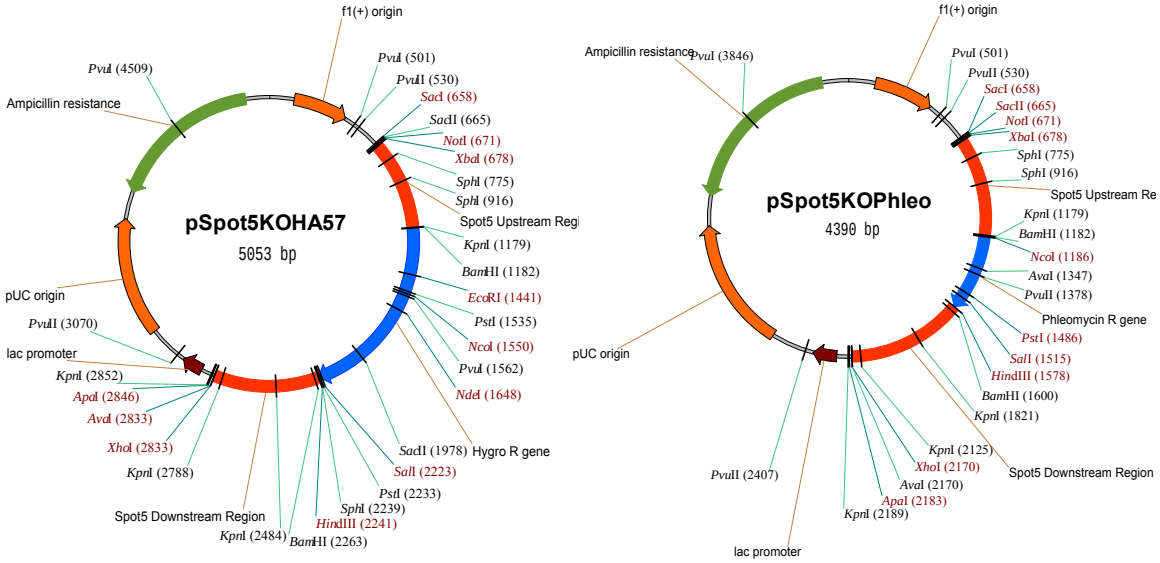


Figure A3.2 Maps of the *TbISWI* knockout constructs. (A) pSpot5KOHA57 was used to replace the *TbISWI* gene with a *Hygromycin* resistance gene. (B) pSpot5KOPhleo was used to replace the *TbISWI* gene with a *Phleomycin* resistance gene.

Subsequently, Matthew Wand attempted to obtain double knockouts of TbISWI by transfecting the pSpot5KOHA57 knockout construct into the TbISWI_KO_Phleo cell line. Although, stable transformants appearing to be double-allele knockouts were obtained, further analysis revealed that they were not true double-knockouts. Restriction digestion of the genomic DNA and its Southern blotting identified an extra allele of *TbISIWI* (data not shown). Such duplication of one allele of an essential gene has been

observed earlier by several other laboratories attempting to knock out both alleles of an essential gene in *T. brucei*. This indicates that *TbISWI* is indeed an essential gene in *T. brucei*.

A4. Laboratory names of cell lines

Table A4.1 Laboratory names of the various cell lines described in the thesis.

Thesis Name	Lab Name
S16.221	221Pur_A3
BF_MCe	MCeG_A1
BF_MCe1	MCeG_S5_A1
BF_MCe2	MCeG_S5_B1
BF_MCEP	MCEPeG_D1
BF_MCEP1	MCEPeG_S5_F1
BF_MCEP2	MCEPeG_S5_G2
BF_VBe	VBeG_A3
BF_VBe1	VBeG_S5_A2
BF_VBe2	VBeG_S5_B2
BF_VBEP	VBEPeG_B1
BF_VBEP1	VBEPeG_S5_A2
BF_VBEP2	VBEPeG_S5_B2
BF_rDe	rDeG_1.A4(2)
BF_rDe1	rDeG_S5_B2.2
BF_rDe2	rDeG_S5_B5.2
BF_rDEP	rDeGxEPeG_B5
BF_rDEP1	rDeGxEPeG_S5_B1
BF_rDEP2	rDeGxEPeG_S5_B9
BF_Pro	EP1eGFP_B4
BF_Pro1	EP1eGFP_S5_B2
BF_Pro2	EP1eGFP_S5_B3
BF_tub	tubGFP_B6R
BF_tub1	tubGFP_S5_B1
BF_tub2	tubGFP_S5_B2
PF_MCe	PF_MCeG_1.4
PF_MCe1	PF_MCeG_S5_1.2a
PF_MCe2	PF_MCeG_S5_1.3a
PF_MCEP	PF_MCEPeG_2.5
PF_MCEP1	PF_MCEPeG_S5_1.6a
PF_MCEP2	PF_MCEPeG_S5_2.4a
PF_VBe	PF_VBeG_1.2
PF_VBe1	PF_VBeG_S5_1.3a
PF_VBe2	PF_VBeG_S5_2.3a

PF_VBEP	PF_VBEPeG_1.2
PF_VBEP1	PF_VBEPeG_S5_1.2a
PF_VBEP2	PF_VBEPeG_S5_1.2c
PF_rDe	PF_rDEPxeG_1.3b
PF_rDe1	PF_rDEPxeG_S5_1.2a
PF_rDe2	PF_rDEPxeG_S5_2.2b
PF_rDEP	PF_rDEPeG_2.5
PF_rDEP1	PF_rDEPeG_S5_1.4b
PF_rDEP2	PF_rDEPeG_S5_2.3a
PF_Pro	PF_EP1eGFP_1.3b
PF_Pro1	PF_EP1eGFP_S5_1a
PF_Pro2	PF_EP1eGFP_S5_1b
PF_tub	PF_tubGFP_1.6b
PF_tub1	PF_tubGFP_S5_1.3b
PF_tub2	PF_tubGFP_S5_1.3c
TbISWI_KO_Hyg	KS717_Spot5_Hygro_A1
TbISWI_KO_Phleo	BF221GPI(VO2+)Spot5KOPhleo_A1