Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*

T. Nicolai Siegel,[1] Doeke R. Hekstra,[2,6] Louise E. Kemp,[1,6] Luisa M. Figueiredo,[1] Joanna E. Lowell,[1,7] David Fenyo,[3] Xuning Wang,[4] Scott Dewell,[5] and George A.M. Cross[1,8]

[1]Laboratory of Molecular Parasitology, The Rockefeller University, New York, New York 10065, USA; [2]Laboratory of Living Matter, The Rockefeller University, New York, New York 10065, USA; [3]Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, New York 10065, USA; [4]Department of Information Technology, The Rockefeller University, New York, New York 10065, USA; [5]Genomics Resource Center, The Rockefeller University, New York, New York 10065, USA

Unusually for a eukaryote, genes transcribed by RNA polymerase II (pol II) in *Trypanosoma brucei* are arranged in polycistronic transcription units. With one exception, no pol II promoter motifs have been identified, and how transcription is initiated remains an enigma. *T. brucei* has four histone variants: H2AZ, H2BV, H3V, and H4V. Using chromatin immunoprecipitation (ChIP) and sequencing (ChIP-seq) to examine the genome-wide distribution of chromatin components, we show that histones H4K10ac, H2AZ, H2BV, and the bromodomain factor BDF3 are enriched up to 300-fold at probable pol II transcription start sites (TSSs). We also show that nucleosomes containing H2AZ and H2BV are less stable than canonical nucleosomes. Our analysis also identifies >60 unexpected TSS candidates and reveals the presence of long guanine runs at probable TSSs. Apparently unique to trypanosomes, additional histone variants H3V and H4V are enriched at probable pol II transcription termination sites. Our findings suggest that histone modifications and histone variants play crucial roles in transcription initiation and termination in trypanosomes and that destabilization of nucleosomes by histone variants is an evolutionarily ancient and general mechanism of transcription initiation, demonstrated in an organism in which general pol II transcription factors have been elusive.

The protozoan parasite *Trypanosoma brucei* branched early in eukaryotic evolution and is probably the most divergent well-studied eukaryote. Many discoveries of general interest have been made in *T. brucei*, emphasizing its value for understanding the evolution of core molecular processes.

Transcription of protein-coding genes by RNA polymerase II (pol II) in *T. brucei* differs in two important aspects from most other eukaryotes. First, transcription is polycistronic: Arrays of sometimes >100 genes are transcribed in polycistronic transcription units (PTUs). This organization is reminiscent of operons in prokaryotes except that there is no evidence to suggest clustering of functionally related genes in *T. brucei*. Second, mRNAs are separated post-transcriptionally by coupled splicing and polyadenylation reactions that add a 39-nucleotide (nt) "spliced-leader" to every mRNA (for review, see Liang et al. 2003). Within a PTU, genes are transcribed from the same strand, but transcription of two neighboring PTUs can either be convergent or divergent. The regions between PTUs are referred to as strand switch regions (SSRs). Strand-specific nuclear run-on assays performed in *Leishmania* (Martinez-Calvillo et al. 2003), a genus related to *T. brucei*, have shown that pol II transcription starts at SSRs between two transcriptionally divergent PTUs (divergent SSRs) and ends at SSRs between two transcriptionally convergent PTUs (convergent SSRs). Because 75% of all *Leishmania major* genes can be found in the same genomic context in *T. brucei* (El-Sayed et al. 2005), indicating a high degree of synteny, it is reasonable to hypothesize that divergent SSRs in *T. brucei* are transcription start sites (TSSs).

[6]These authors contributed equally to this work.
[7]Present address: BIO Ventures for Global Health, 1225 Eye Street, NW, Washington, DC 20005, USA.
[8]Corresponding author.
E-MAIL george.cross@rockefeller.edu; FAX (212) 327-7572.

Pol II promoters and associated transcription factors have been elusive in *T. brucei*, with the exception of the spliced-leader RNA promoter (Das et al. 2005; Schimanski et al. 2005). Hence, how most pol II transcription is initiated in *T. brucei* remains a fundamental question. Some data suggest that pol II transcription can be initiated from a site lacking an apparent promoter as long as transcription occurred at a nearby site, suggesting that an "open" chromatin structure, generated by another polymerase at a nearby site, may have been sufficient to allow pol II transcription to initiate (McAndrew et al. 1998).

Genome-wide analyses in several different organisms indicate that the extent of chromatin-mediated DNA compaction differs throughout the genome. Actively transcribed genes are generally located in less compact euchromatic regions where regulatory sequences are hypersensitive to DNase treatment (Weintraub and Groudine 1976). In contrast, silent genes are located in more compact heterochromatic regions.

The dynamic nature of chromatin packing is mediated by multiple processes (for review, see Henikoff 2008). DNA sequence can affect nucleosome positioning, keeping regulatory DNA free of nucleosomes or facilitating nucleosome eviction. Nucleosome remodeling can lead to a high turnover of histones in actively transcribed genes, and replacement of canonical histones by histone variants can create less stable nucleosomes. Genome-wide analyses in yeast, human cells, and *Drosophila* have revealed that the histone variants H2AZ and H3.3 are enriched in proximity to RNA poll II promoters (Mito et al. 2005; Raisner et al. 2005; Barski et al. 2007). Furthermore, the structure of nucleosomes containing both histone variants H2AZ and H3.3 is less stable than of nucleosomes composed of canonical histones. This finding suggests that nucleosomes containing H2AZ/H3.3 are more easily evicted from promoters during the initiation of transcription than canonical nucleosomes (Jin and Felsenfeld 2007). Finally, post-translational histone modification, especially of the N-terminal tails of histone H3 and H4, can influence nucleosome stability. For example, acetylation of the H4 tail interferes with its stabilizing interaction with adjacent nucleosomes. Besides this structural role, acetylation of the H4 tail provides a binding platform for acetyl-binding proteins, including the bromodomain-containing factors (BDFs). The function of BDFs ranges from nucleosome remodeling to recruitment of histone-modifying enzymes or transcription factors.

Although the primary structure of trypanosome core histones diverges significantly relative to other eukaryotes, trypanosomes do contain an extensively acetylated H4 tail and several putative BDFs. *T. brucei* also has one variant of each of the four core histones (Lowell and Cross 2004; Lowell et al. 2005). Given the lack of an obvious pol II promoter sequence motif, the observation that divergent SSRs contain somewhat elevated levels of H4 acetylation in *T. cruzi* (Respuela et al. 2008), and the presence of four histone variants, we hypothesized that a specific chromatin structure at divergent SSRs may compensate for the apparent lack of pol II promoter sequence motifs.

To identify factors that contribute to the chromatin structure surrounding probable pol II TSSs in *T. brucei*, we adopted chromatin immunoprecipitation (ChIP) and sequencing (ChIP-seq) technology to determine the genome-wide distribution of H4K10ac, all four histone variants, and one bromodomain protein (BDF3). Our data provide strong evidence that at least three processes associated with the generation of open chromatin—distinctive DNA sequence, incorporation of histone variants, and histone tail modification—are involved in pol II transcription initiation.

Results

Establishment of ChIP-seq for the genome-wide analysis of chromatin structure in T. brucei

We established a ChIP-seq protocol for *T. brucei*. In contrast to ChIP-on-chip, in which immunoprecipitated DNA is hybridized to oligonucleotide microarrays, the immunoprecipitated DNA in ChIP-seq is directly sequenced. Each DNA "tag" is aligned to the genome and the number of tags aligned to each base pair is summed, giving the relative enrichment of a particular factor along the genome. Reproducibility and linear range of ChIP-seq are higher than for microarray-based systems (Barski et al. 2007), and resolution is limited only by the size of the immunoprecipitated DNA fragments and not by the genomic spacing of the oligonucleotides used to create the microarray.

To ensure the robustness of the ChIP-seq approach, we optimized previously established ChIP protocols to obtain a protocol that was tolerated by antibodies specific for H3, H4, H4K10ac, HA peptides, and Ty1 peptides (Lowell and Cross 2004; Lee et al. 2006). We then established sonication conditions that yielded highly reproducible DNA fragmentation ($\sim$300-base-pair [bp] fragments) and allowed us to omit size-selection steps during preparation of the immunoprecipitated DNA for high-throughput sequencing. Partly because our protocol did not include any DNA gel extraction step, we were able to work with relatively few cells: 10^8 trypanosomes, which are equivalent in DNA content to $\sim$$10^6$ human cells.

Repetitive elements in the genome template were not masked, and hits were found using the BLAT-based algorithm (Kent 2002) for each tag. To obtain consistent results, sequence tags with multiple hits were annotated to the lowest nucleotide position on the smallest numbered chromosome that gave the best match. This annotation methodology produced artifactual peaks at the first occurrence of highly repetitive sequence motifs, even if the factor investigated was not enriched at a particular repeat region, but was unavoidable because, although the *T. brucei* genome has been almost completely sequenced, the lengths of most repetitive regions are unknown. To identify and compensate for such artifacts, we sequenced sonicated genomic DNA, DNA from

microccal nuclease-digested chromatin, and DNA immunoprecipitated with a custom-made control antibody specific for trypanosome H4 regardless of its modification state (Siegel et al. 2008a) and a commercial antibody to the C-terminal tail of H3 (Abcam 1791; see the Materials and Methods). Because DNA tags from sonicated genomic DNA should align with the genome uniformly, any observed peaks and gaps represent first and subsequent repeats of repetitive sequence elements, respectively, which is what was observed. Similarly, alignment of sequenced tags from micrococcal nuclease-treated DNA and ChIP-seq data for the H4 (see below) and H3 antibodies revealed peaks and gaps at repetitive elements.

Histone H4K10ac is enriched at probable pol II TSSs

Although the amino acid sequence of the N-terminal tail of *T. brucei* histone H4 diverges from that of mammals and even from other protozoa (Fig. 1A), numerous lysine residues, acetylated at different frequencies, have been identified (Janzen et al. 2006; Mandava et al. 2007). For example, ~80% of H4 is acetylated at K4 and ~10% is acetylated at K10. H4K4 is acetylated by the nonessential histone acetyltransferase 3 (HAT3) (Siegel et al. 2008a), but H4K10 is acetylated by HAT2, an enzyme essential for viability in *T. brucei* (Kawahara et al. 2008). The highly divergent sequence of the trypanosome histone tails prevents the use of commercially available antibodies. We therefore examined the genome-wide distribution of H4K10ac using custom-made antibodies specific to H4K10ac or to H4 regardless of its modification state. The generation and characterization of these antibodies are described elsewhere (Kawahara et al. 2008; Siegel et al. 2008a).

Peaks of H4K10ac enriched up to 300-fold over background were distributed throughout the genome, with distinct twinned peaks occurring at every divergent SSR

Figure 1. Histone modifications and gene organization in *T. brucei*. (*A*) *T. brucei* contains highly divergent histone sequences. Shown, as examples, are the N-terminal histone H4 sequences of *Tetrahymena thermophilia*, *Homo sapiens*, and *T. brucei*. Known sites of acetylation (blue) and methylation (red) are indicated. (*B*) Divergent SSRs contain pol II TSSs. Convergent SSRs contain pol II TTSs. Orange boxes represent ORFs and green arrows indicate the direction of transcription.

and no enrichment at convergent SSRs (Figs. 1B, 2A, 3), suggesting a link between H4K10 acetylation and transcription initiation. The distribution was remarkably similar between two parasite life cycle stages (bloodstream form [BF] and procyclic form [PF]), with the exception of one peak that was unique to each life cycle stage (Fig. 3, chromosomes 7 and 11, red asterisks).

We also observed 61 single H4K10ac-rich peaks at non-SSRs (Figs. 2, 3), which we suggest are unanticipated TSSs. Many of these peaks occur immediately downstream from tRNA genes (see example in Fig. 2B). Most tRNA genes are located at convergent SSRs but some are present at non-SSRs that appeared to be embedded within pol II PTUs, which raised questions about the overlap of pol III and pol II transcription. We found that all but three tRNA genes that are not associated with convergent SSRs are located upstream of H4K10ac-rich "single" peaks, but not all single peaks are associated with tRNA genes.

Two studies have reported pol II transcription initiation outside of divergent SSRs: upstream of the actin gene (Ben Amar et al. 1991) and in the HSP 70 locus (Lee 1996). While the potential actin promoter region coincides with an H4K10ac-rich region, the HSP70 promoter does not, suggesting that the HSP70 promoter was misassigned, which agrees with a previous study that failed to confirm promoter activity for the HSP70 promoter (Hausler and Clayton 1996).

The width of the H4K10-rich regions was uniform and surprisingly wide (width at half maximum enrichment 5.6 ± 0.8 kb; full width ~8.7 kb)—well above the resolution limit of 250–300 bp for our ChIP-seq analysis. Based on the total number (207) of H4K10ac peaks, their average width and the fact that ~10% of H4K10 sites are acetylated overall, we estimate that all H4 is acetylated at K10 within these peaks, which would explain their remarkably constant height.

Finally, there appear to be higher levels of H4K10ac just upstream of the first PTU and downstream from the last PTU on each chromosome (Figs. 2A, 3), but these H4K10ac-rich regions do not form distinct peaks. As for other repetitive sequences, tags were assigned to their first occurrence and therefore appear at only a few subtelomeric regions, and peak sizes cannot be corrected for the (unknown) lengths of the repeats.
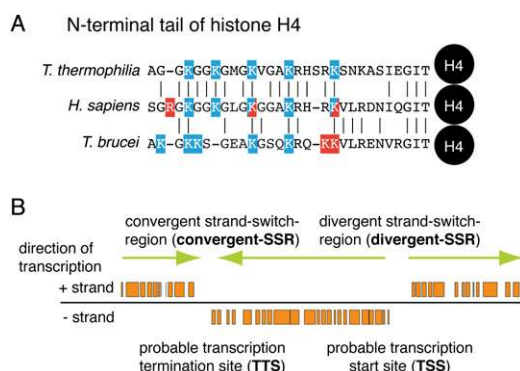
H2AZ and H2BV are enriched at probable pol II TSSs

T. brucei H2AZ is essential for viability and dimerizes exclusively with H2BV, a variant form of H2B that is also essential for viability (Lowell et al. 2005). Thus, all *T. brucei* nucleosomes that contain H2AZ also contain H2BV. Although the function of H2AZ is not understood in *T. brucei*, it has been linked to transcriptional activation in other organisms, particularly when present in a nucleosome containing a variant of histone H3 (Jin and Felsenfeld 2007). Using immunofluorescence (IF) microscopy, H4K10ac showed a punctate nuclear distribution with no cell cycle dependence (Fig. 4D; Supplemental Fig. S1A). This localization is similar to that observed previously for H2AZ and H2BV, and significantly different
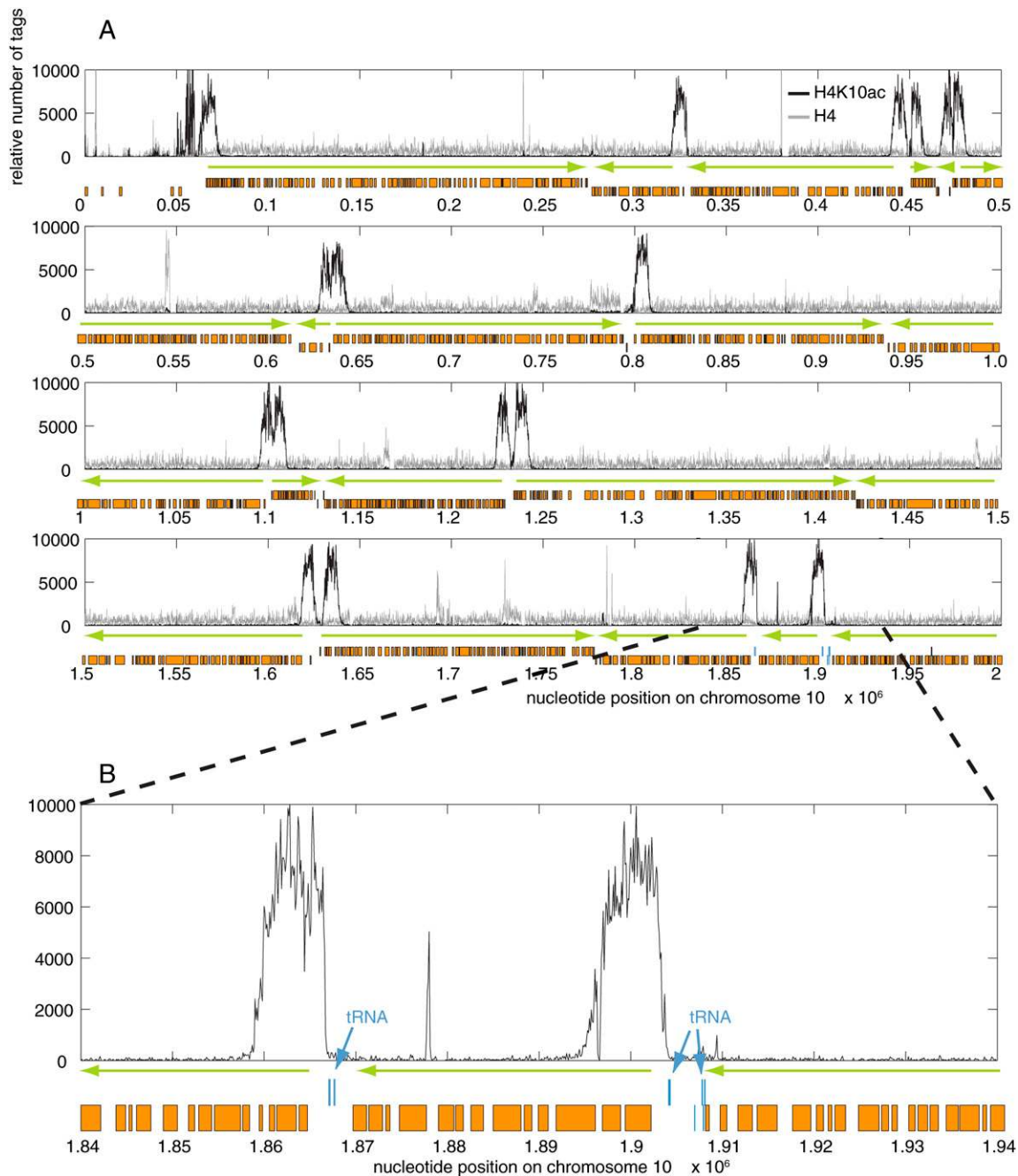
Siegel et al.

Figure 2. Histone H4K10ac is enriched at probable pol II TSSs. (*A*) For a representative region of chromosome 10, the relative number of tags for histone H4K10ac (black) and histone H4 (gray) are calculated for a window size of 100 bp. Orange boxes represent ORFs and green arrows indicate direction of transcription. (*B*) H4K10ac-rich regions at non-SSRs are often located next to tRNA genes (blue boxes).

from that of a general H4 antibody (Fig. 4D; Lowell et al. 2005). A Ty1-epitope-tagged version of H2BV also showed a high degree of colocalization with H4K10ac (Supplemental Fig. S1A). These observations led us to hypothesize that H2AZ/H2BV-containing nucleosomes might be enriched at probable pol II TSSs.

We therefore performed ChIP-seq experiments with N-terminally Ty1-tagged versions of H2AZ and H2BV. To

assure functionality of the ectopically expressed and tagged versions of the variants, we deleted both endogenous alleles of each variant. No growth defect was detected, confirming functionality of the tagged versions. Sequence analysis of the immunoprecipitated DNA indicated that the genomic distribution of H2AZ and H2BV, including the width and intensity of the peaks, was almost indistinguishable from H4K10ac (Fig. 4A). As was described for
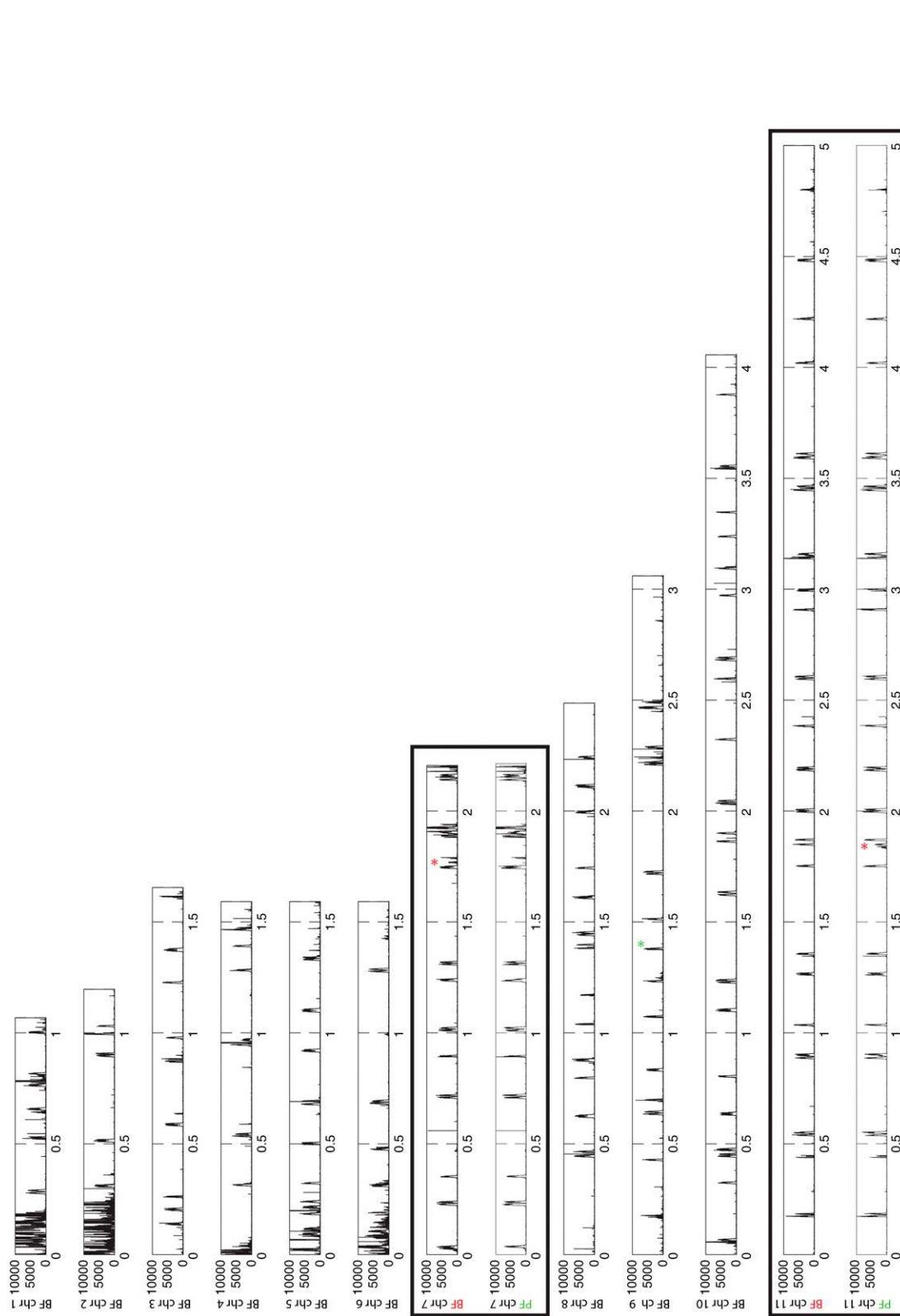
Figure 3. Genome-wide distribution of histone H4K10ac. Relative number of tags for H4K10ac in BF (chr 1–11) and PF (chr 7 and 11). The green asterisk marks the previously identified actin promoter (Ben Amar et al. 1991). The red asterisks mark life cycle-specific peaks. Window size, 100 bp.
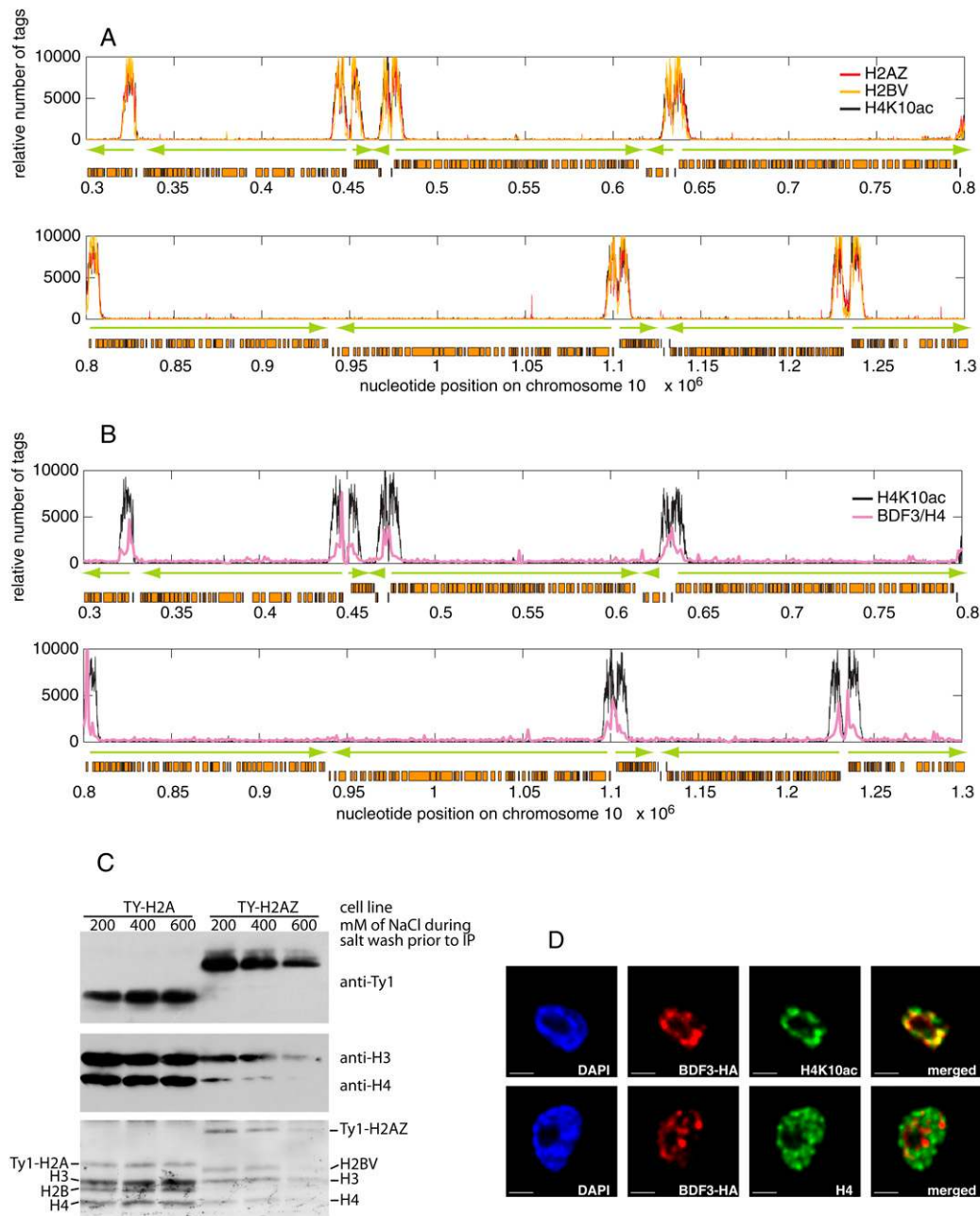
Siegel et al.

Figure 4. Histone variants H2AZ and H2BV and BDF3 are enriched at probable pol II TSSs. (*A*) For representative regions of chromosome 10, the relative number of tags for H2AZ (red), H2BV (yellow), and H4K10ac (black) are overlaid. Window size, 100 bp. Orange boxes represent ORFs and green arrows indicate direction of transcription. (*B*) For a representative region of chromosome 10, the relative number of tags for H4K10ac (black) and BDF3 (magenta) are calculated for a window size of 100 bp and 1000 bp, respectively. The number of tags from the BDF3 ChIP was normalized based on the number of tags from an H4 ChIP (not shown), to eliminate artifacts attributable to repetitive sequences. For unnormalized BDF3 data, see Supplemental Figure S4. Orange boxes represent ORFs. (*C*) Chromatin was prepared from cell lines containing Ty1-tagged H2A or H2AZ and equal amounts were washed with buffers containing increasing levels of NaCl. Chromatin was treated with micrococcal nuclease to generate mononucleosomes and immunoprecipitated with monoclonal antibody BB2. Coimmunoprecipitated histones were separated by SDS-PAGE and immuno-blotted (*top* and *middle* panels) or stained with amido black (*bottom* panel), to determine the levels of coimmunoprecipitated histones in the H2A-IP and H2AZ-IP. (*D*) A representative IF colocalization analysis of BDF3-HA (red), DAPI (blue), and H4K10ac (*top* panel, green) or H4 (*bottom* panel, green). Bar, 2 μm.

H4K10ac, two peaks of H2AZ/H2BV enrichment were observed at every divergent SSR and a single peak of H2AZ/H2BV enrichment was present at every non-SSR H4K10ac-rich region. H2AZ and H2BV were also enriched upstream of the first PTU and downstream from the last PTU on every chromosome, as was H4K10ac (above).

Next, we sought to understand the biological relevance of H2AZ and H2BV enrichment at probable pol II TSSs. In chickens, it has been shown that nucleosomes containing histone variants are destabilized relative to those containing core histones (Jin and Felsenfeld 2007). To test if H2AZ/H2BV-containing nucleosomes of *T. brucei* are less stable than nucleosomes composed of canonical histones, we performed coimmunoprecipitation (co-IP) experiments. We prepared mononucleosomes from four cell lines expressing either Ty1-tagged H2AZ or H2A or Flag-tagged H2BV or H2B, as described previously (Lowell et al. 2005) except that the chromatin pellet was washed at increasing ionic strengths. The tagged histones and histone variants were then immunoprecipitated and the extent of association with H3 and H4 was determined. A less stable nucleosome should lead to decreased co-IP; indeed, we observed significantly less association of H4 and H3 with tagged histone variants compared with tagged core histones (Fig. 4C; Supplemental Fig. S1B). Taken together, our ChIP-seq data and co-IP experiments demonstrate that probable pol II TSSs are marked by distinct and less stable nucleosomes.

BDF3 is enriched at probable pol II TSSs

Acetylated histone N termini can serve as binding platforms for effector proteins that interpret histone modification patterns and elicit downstream effects (Dhalluin et al. 1999). BDFs, which can bind to acetylated lysines (Winston and Allis 1999), are examples of effector proteins. BDFs have been shown to tether HATs or nucleosome remodeling complexes to specific chromosomal sites and to function as transcriptional coactivators (Dhalluin et al. 1999; Hassan et al. 2002). The *T. brucei* genome is predicted to contain six genes encoding BDFs that range in size from 221–660 amino acids. Sequence conservation among the putative BDFs varies (E-values 4.1×10^{-23} to 7.7×10^{-10} based on superfamily database; http://supfam.mrc-lmb.cam.ac.uk, version 1.69). Given the prominent role BDFs play in recruiting transcription machinery, we decided to test whether any trypanosome BDF colocalizes with H4K10ac at probable pol II TSSs.

We focused on the four most highly conserved candidates (BDF1–4) and a single candidate with two potential bromodomains (BDF5). These five BDFs were tagged with an HA epitope and their cellular localization determined by IF microscopy. BDF2, BDF3, and BDF5 localized to the nucleus, but only BDF3 colocalized with H4K10ac (Fig. 4D; data not shown). To substantiate the IF microscopy data, we performed ChIP-seq with a cell line containing an HA-tagged allele of BDF3, in which the second untagged allele was deleted to ensure functionality of the tagged protein. ChIP-seq demonstrated that BDF3 was highly enriched at the same sites as H4K10ac (Fig. 4B; Supplemental Fig. S4), but there were distinct differences: Whereas the signals clearly overlapped, BDF3 was concentrated toward the 5′ end of the H4K10-rich regions (Figs. 4B, 6B; Supplemental Fig. S4).

BDF3 is a 25-kDa protein containing no conserved domains other than the bromodomain. To test whether BDF3 is essential for viability, we conducted RNAi-mediated knockdown experiments. Reduction of BDF3 led to an immediate growth defect and most cells appeared to be dying after 48 h, which is consistent with BDF3 being essential (data not shown). One interpretation of BDF3 and H4K10 colocalization would be that BDF3 binds to acetylated H4K10, but we were unable to demonstrate this (see the Supplemental Material).

H3 and H4 variants are enriched at probable pol II transcription termination sites

In addition to H2AZ and H2BV, *T. brucei* contains variant forms of H3 and H4, whose functions are unknown. Past work, using IF microscopy and standard ChIP followed by dot-blot readout, revealed that H3V was enriched at telomeric repeats, but localization at additional sites could not be ruled out because of the limits of the technology employed (Lowell and Cross 2004). No phenotype could be detected for a cell line lacking both alleles of H3V (Lowell and Cross 2004). Despite H3V being nonessential under laboratory culture conditions, the conserved function of histone H3 variants in transcription regulation in other organisms suggests that H3V may have a subtle and undetected role in transcription regulation in *T. brucei.*

Because a role of H3V in transcription regulation would most likely lead to incorporation of H3V outside of telomeric and subtelomeric sites, we performed ChIP-seq using a cell line containing a Ty1-tagged copy of H3V. In agreement with previous results, we found that H3V was highly enriched at telomeric and subtelomeric regions of the genome (Fig. 5; Supplemental Fig. S5), but only ~65% of the immunoprecipitated DNA corresponded to telomeric and subtelomeric sequences. H3V was also enriched at convergent SSRs and just upstream of all H4K10ac-rich regions not associated with SSRs (Fig. 5; Supplemental Fig. S5). Taken together, these data show that H3V is enriched at the presumed ends of pol II PTUs.

T. brucei H4V has not been characterized previously. It shares 85% sequence identity with *T. brucei* H4, whereas H2AZ, H2BV, and H3V share only 37%–45% sequence identity with their core histone counterparts (Supplemental Fig. S2). To determine the cellular localization of H4V, we generated a cell line containing an ectopically expressed Ty1-tagged H4V and deleted both endogenous alleles. By IF microscopy, we observed that Ty1-H4V localized to the nucleus and displayed no apparent cell cycle regulation (Supplemental Fig. S3). Using this cell line, we determined the genomic distribution of H4V. H4V was distributed broadly throughout the genome, but we observed peaks of enrichment at the ends of pol II PTUs, similar to those observed for H3V (Fig. 5; Supplemental Fig. S5). In contrast to H3V, H4V was much less enriched at telomeric or subtelomeric sites.

We generated an H4V knockout cell line but, as for H3V, no growth defect was detected under culture conditions (data not shown). Nevertheless, their genomic distribution at the ends of PTUs strongly suggests a role for H3V and H4V in transcription regulation.
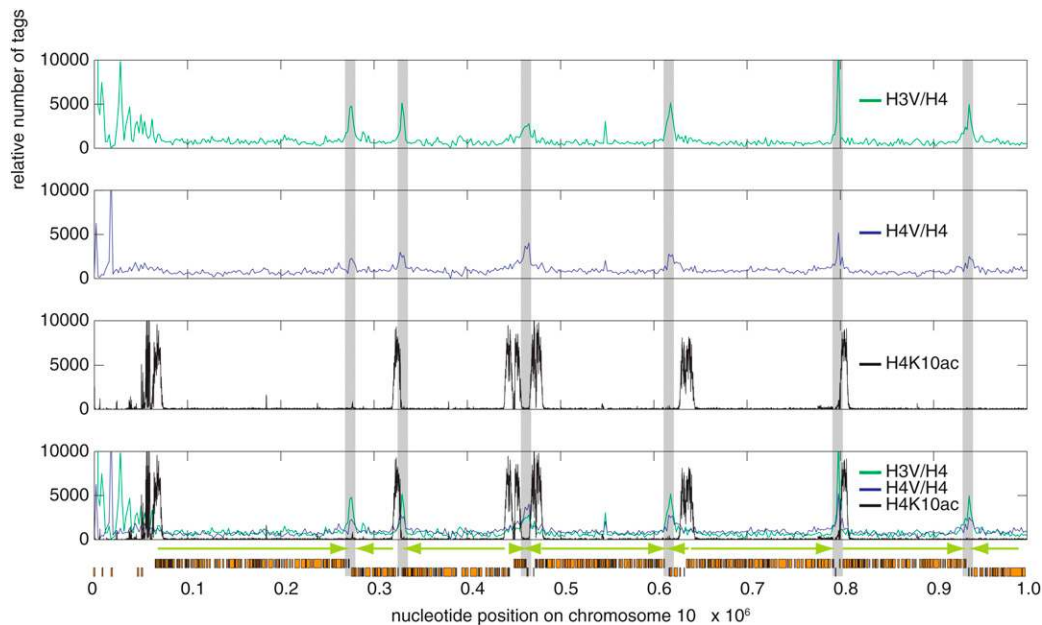
Siegel et al.

Figure 5. Histone variants H3V and H4V are enriched at probable pol II TTSs. For a representative region of chromosome 10, the relative number of tags for H3V (first panel, green) and histone H4V (second panel, blue) ChIPs were calculated for a window size of 2000 bp and normalized based on the number of tags from an H4 ChIP, to eliminate artifacts attributable to repetitive sequences. For unnormalized data, see Supplemental Figure S5. (Third panel) The relative number of tags for H4K10ac (black) is calculated for a window size of 100 bp. (Fourth panel) Overlay of H4K10ac, H3V, and H4V data. Orange boxes represent ORFs, vertical gray boxes mark pol II TTSs, and green arrows indicate direction of transcription.

Analysis of DNA sequences at probable pol II TSSs reveals long G runs

Taken together, our data suggest that chromatin immediately upstream of and downstream from PTUs has strikingly distinct features. Although it seems reasonable to speculate that DNA sequence could play a role in establishing these unique chromatin structures, no pol II promoter sequence motif has been identified in *T. brucei*. In fact, previous findings led to the hypothesis that an open chromatin structure may be sufficient to initiate pol II transcription (McAndrew et al. 1998). Nevertheless, data from the related organism *L. major* (Martinez-Calvillo et al. 2003), and the presence of two rather than one H4K10ac/H2AZ/H2BV-rich peaks at divergent SSRs, strongly suggest that pol II transcription is unidirectional and that a specific DNA feature may determine directionality.

We took several approaches to identify DNA sequence motifs that were unique to H4K10ac-rich regions. We also compared the sense strand to the complementary strand within H4K10ac-rich regions. The only motif consistently identified was a G-rich stretch in the sense strand at the upstream end of every H4K10ac-rich region (Fig. 6). Stretches of nine to 15 consecutive guanines were enriched approximately sevenfold in these regions compared with other sites (Fig. 6B). Specifically, within 91% of H4K10ac/H2AZ/H2BV-rich peaks (117 out of 128), we found stretches of at least 10 guanines, allowing for up to two mismatches, between 1000 and 5000 bp upstream of the peak center (expected: 65%, $P < 10$–12). In contrast to

the G-rich upstream part of the H4K10ac-rich region, the central 3 kb of these regions was slightly AT-rich (56%) compared with the rest of the genome (50%) (Fig. 6A). The results are modeled in Figure 6C.

Discussion

In this study, we found that the beginning and end of pol II PTUs are characterized by distinct chromatin modifications and that there may be more pol II TSSs than assumed previously, and we confirmed that probable pol II TSSs do not contain classic sequence motifs for transcription factor interactions.

Pol II transcription initiation and termination regions have distinct chromatin features

Four histone variants mark the beginning and end of pol II PTUs in *T. brucei*. H2AZ and H2BV are enriched at probable pol II TSSs. H3V and H4V are enriched at probable TTSs. We also see enrichment of H4K10ac and BDF3 at probable pol II TSSs (summarized in Fig. 7), where all H4K10 appears to be acetylated. It has been proposed before that an open chromatin structure may be sufficient for pol II transcription initiation in *T. brucei* (McAndrew et al. 1998). In the following paragraphs we propose a model for the sequence of events leading to pol II transcription initiation.

Analysis of genome-wide distribution of histone modifications has been performed in several organisms where H3K4me3 consistently localizes to promoter regions
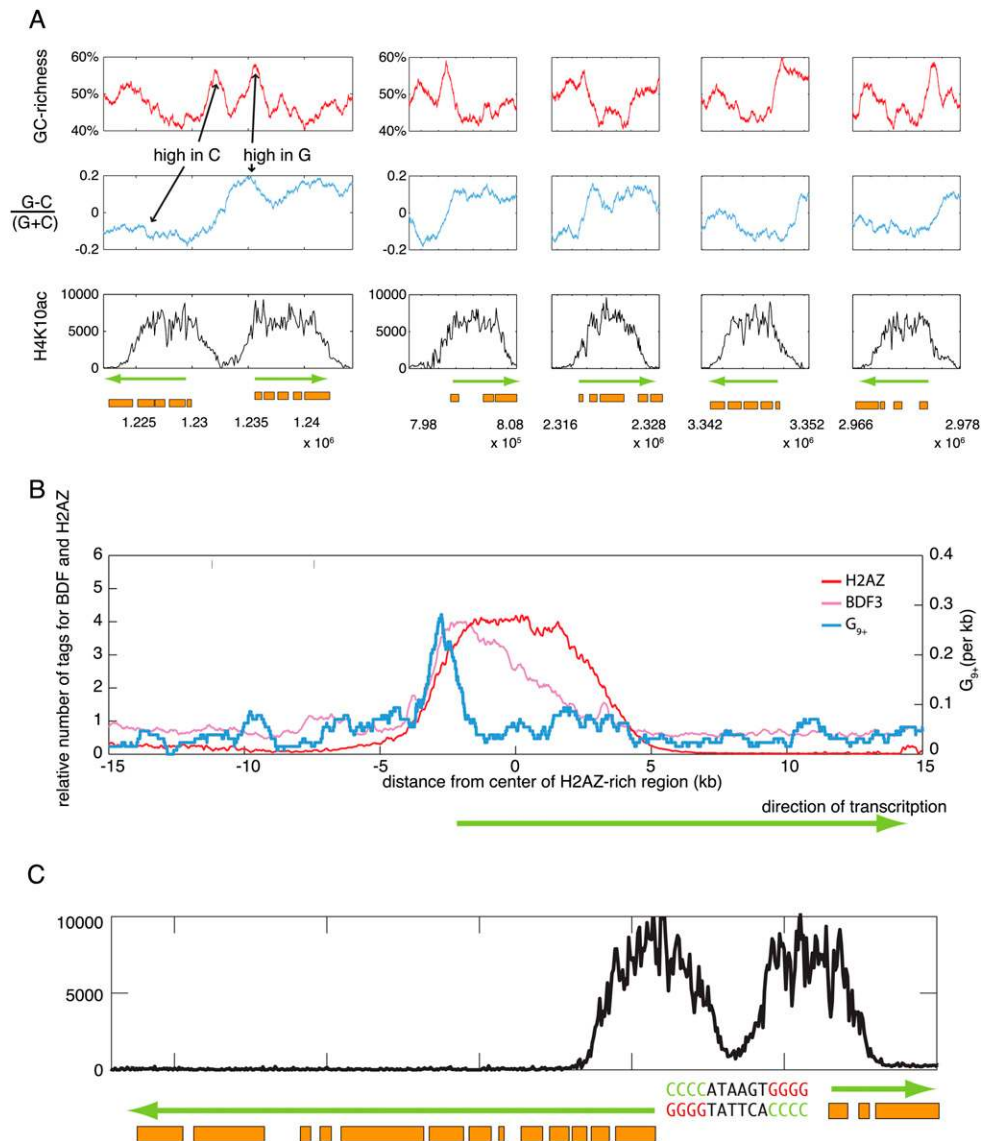
Figure 6. 5′ ends of probable pol II TSSs are marked by runs of guanine. (*A*) Divergent SSRs contain two peaks of high GC richness (red). The *left* peak indicates high levels of C in the *top* strand and the *right* peak indicates high levels of G in the *top* strand (blue). Orange boxes represent ORFs and green arrows indicate direction of transcription. Numbers on *x*-axis indicate nucleotide position on chr 10. (*B*) Average density of stretches of G ($n \geq 9$) from 128 TSSs (dark blue, *right* axis). Density shown is the mean over 38 putative TSSs at non-SSRs and 90 TSSs at divergent SSRs using a 1000-bp moving average (densities were statistically indistinguishable between the two types of TSSs; data not shown). For comparison, the mean shapes of H2AZ (red) and BDF3 (magenta) enrichment are shown at non-SSR TSSs (*left* axis, see also Supplemental Fig. S6). BDF3 data were scaled to have the same maximum as H2AZ data. TSSs centers were determined based on Gaussian fits to the H2AZ data. (*C*) Model indicating how regions upstream of probable pol II TSSs are enriched for guanine. Sequence and peak size are not drawn to scale.

(Barski et al. 2007; Shilatifard 2008). We currently lack a ChIP-grade H3K4me3 antibody for trypanosomes but, based on previous findings that H3K4me3 is enriched in H2BV-containing nucleosomes (Mandava et al. 2008), we assume that in *T. brucei* H3K4me3 is also enriched at pol II TSSs. The role of H3K4me3 in transcription initiation appears manifold, but evidence exists that H3K4me3 is required for histone tail acetylation, another chromatin feature consistently associated with active chromatin (for review, see Pillus 2008). Because HAT2 contains a methyl-binding domain (Ivens et al. 2005) and is responsible for H4K10 acetylation, it appears plausible that it is targeted to TSSs by binding to a methyl mark, possibly to H3K4me3. In *T. brucei* H3K4 trimethylation may thus be required for H4K10 acetylation. The enzyme responsible for H3K4 methylation has not yet been identified in *T. brucei*.

The role of histone variants in transcription initiation is largely unknown, although nucleosomes containing H2AZ and H3.3 variants are significantly less stable than nucleosomes containing only canonical histones (Jin and
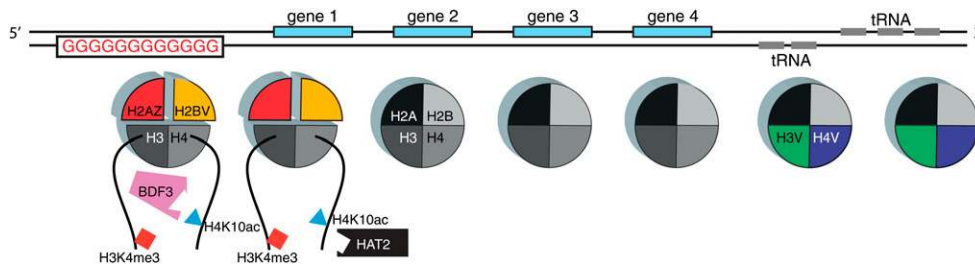
Figure 7. Model to illustrate a possible cascade of events leading to pol II transcription initiation. Histone H3K4me3 enrichment at TSSs could provide a binding site for HAT2, which contains a chromodomain and is responsible for H4K10 acetylation. H4K10ac is strongly enriched at probable pol II TSSs and might serve as a binding site for BDF3. The latter is enriched at the 5' end of probable pol II TSSs, where it may play a role in recruiting pol II transcription factors or the chromatin remodeling complex involved in H2AZ and H2BV incorporation. These two histone variants are enriched at probable pol II TSSs and form less stable nucleosomes that could facilitate binding of the transcriptional machinery. The G tract at the 5' end of probable pol II TSSs could block antisense transcription and/or provide binding sites for transcription factors. H3V and H4V variants are enriched at probable pol II TTSs, which often contain tRNA or other genes transcribed by RNA pol III.

Felsenfeld 2007). These unstable nucleosomes are often enriched at pol II TSSs, where they are likely to facilitate nucleosome eviction prior to transcription initiation (Guillemette et al. 2005; Mito et al. 2005; Zhang et al. 2005). In this study, we show that *T. brucei* histone variants H2AZ and H2BV are also enriched around probable pol II TSSs, and nucleosomes containing H2AZ and H2BV are less stable than nucleosomes containing the corresponding core histones. These observations strongly support the hypothesis that the presence of H2AZ and H2BV at *T. brucei* pol II TSSs contributes to a more open chromatin structure.

In yeast, the Swr1 chromatin remodeling complex is responsible for H2AZ incorporation. Targeting of this complex to the correct genomic location appears to be mediated by a member of the Swr1 complex, the bromodomain-containing protein BDF1 (for review, see Korber and Horz 2004). Thus, yeast BDF1, may target the chromatin remodeling complex to sites containing acetylated histones. Further evidence that H2AZ incorporation may depend on histone acetylation stems from the finding that H4K16ac is required for H2AZ incorporation in subtelomeric regions (Shia et al. 2006). Our data indicate that *T. brucei* BDF3 is localized to a relatively narrow region at TSSs compared with H2AZ. Therefore, it appears plausible that BDF3 plays a role in targeting a chromatin remodeling complex to TSSs in *T. brucei*, a role similar to BDF1 in yeast. Because we were unable to demonstrate an interaction between H4K10ac and BDF3 in vitro, we cannot exclude the possibility that a different acetyl mark may serve as binding site for BDF3.

The *T. brucei* H3 and H4 variants mark the end of probable pol II PTUs. This further emphasizes the notion that pol II transcription is regulated by chromatin structure. However, the near complete acetylation of H4K10 at all probable pol II TSSs suggests this mark plays no role in regulating the level of transcription initiation: It simply may be required to turn on transcription. Trypanosomes appear to regulate pol II-driven gene expression post-transcriptionally: The efficiency of *trans*-splicing individual mRNAs differs significantly, as do mRNA stability, translation efficiency, and protein stability (for review, see Clayton 2002).

Polyguanine tracts could guide directional pol II transcription

The presence of unstable, noncanonical nucleosomes at probable pol II TSSs leads to a more open chromatin structure that may be sufficient for transcription initiation. However, we also identified an enrichment of polyguanine motifs at probable pol II TSSs, a common feature of promoter regions (Huppert and Balasubramanian 2007). The presence of short GC-rich tracts at SSRs was previously recognized in *Leishmania* (Martinez-Calvillo et al. 2003). Two lines of evidence placed TSSs on chromosome I quite precisely on either side of a GC-rich tract, but this SSR was very short compared with those in *T. brucei*. There are other significant differences in gene transcription between *Leishmania* and *T. brucei*, so it is not possible to make unqualified extrapolations from one organism to the other. It will be interesting to determine the distribution of histone modifications and variants within the *Leishmania* genome.

Polyguanine motifs have the potential to form G quadruplexes, a stable secondary structure, but the function of this sequence element in *T. brucei* remains to be determined. One possibility is that the G-rich DNA adopts a secondary structure that functions as transcription factor-binding site. Transcription factors specifically binding to G quadruplexes have been identified in humans (Etzioni et al. 2005). Alternatively, the polyguanine motif could govern the unidirectionality of transcription in *T. brucei*. It has been shown that stretches of guanines located on the mRNA sense strand can hinder the progression of pol II (Tornaletti et al. 2008). If an open chromatin structure is sufficient to initiate transcription in *T. brucei* and if transcription were to start within the H4K10ac-rich region and proceed in both directions (divergent transcription), it would encounter the G-rich stretch only if it went the "wrong" way (Figs. 6C, 7). The polyguanine motif could thus hinder pol II from proceeding, and thereby promote unidirectional transcription. Divergent transcription initiation followed by termination of antisense transcription may be a common phenomenon. A recent study reports divergent transcription

initiation from >40% of all protein-encoding gene promoters in mice. In mice, however, transcription termination appears to correlate with the lack of H3K79me2 (Seila et al. 2008).

Novel pol II TSSs

Whereas it has generally been assumed that pol II transcription starts at divergent SSRs, we also found many putative pol II TSSs at non-SSRs, often downstream from tRNA genes. As tRNA genes can inhibit pol II transcription in yeast (Hull et al. 1994), a similar mechanism for tRNA-mediated transcriptional silencing was proposed for *T. brucei* (Marchetti et al. 1998). It has been suggested that the subnuclear localization of the tRNA genes may be important for tRNA-mediated silencing (Kendall et al. 2000). More recently, it has been shown in yeast that TFIIIC, which is required for pol III transcription complex assembly, tethers tRNA genes to distinct loci at the nuclear periphery (Noma et al. 2006). This could generate pol III transcription clusters from which pol II is largely excluded. TFIIIC binds to the B-box, one of two highly conserved promoter elements required for tRNA transcription in most eukaryotes, including *T. brucei* (Nakaar et al. 1994). As in yeast, the *T. brucei* B-box is only essential for tRNA transcription in vivo, not in vitro (Margottin et al. 1991; Nakaar et al. 1997), so its role in subnuclear clustering of tRNA may be conserved in trypanosomes, although a homolog of TFIIIC has not been identified in *T. brucei*. Such subnuclear clustering of tRNA genes into pol III transcription clusters could explain why we were unable to initiate pol II transcription by placing long stretches of DNA sequence from divergent SSRs into convergent SSRs (data not shown).

If tRNA genes indeed interrupt pol II transcription, reinitiation would be required downstream from the tRNA "roadblock." We found that all but three tRNA genes located outside of convergent SSRs are surrounded by high levels of H3V and H4V and followed by high levels of H4K10ac/H2AZ/H2BV/BDF3, suggesting that tRNA indeed represents a boundary element and that pol II transcription has to reinitiate downstream from this boundary. The three tRNA genes not located at convergent SSRs and not next to a region of high H4K10ac are isolated and are not in clusters like most other tRNAs. Individual tRNA genes with modest TFIIIC binding are unable to function as boundary elements in yeast (Noma et al. 2006).

If all of the putative pol II TSSs at non-SSRs represent reinitiation sites, and because only some of these sites are located downstream from tRNA genes, what is the nature of the non-tRNA roadblocks? Could these regions contain unrecognized RNA genes that are also transcribed by pol I or pol III?

Conclusion

ChIP-seq technology offers new insight into the long-standing problem of transcription initiation in *T. brucei*, although a mechanistic description of the role of the different chromatin components awaits detailed biochemical studies. The apparent lack of conserved pol II promoter motifs makes trypanosomes an interesting model system in which to study the role of chromatin composition and modification in the control of transcription initiation and termination. Being evolutionarily divergent, *T. brucei* should also shed light on the evolution of chromatin-mediated epigenetic regulation in eukaryotes.

Materials and methods

Trypanosome culture

All *T.brucei* BF cell lines were produced from wild-type Lister 427 (MITat 1.2, clone 221a) or a derivative "single marker" line, which expresses T7 RNA polymerase and the Tet repressor (Wirtz et al. 1999). Cells were cultured in HMI-9 medium (Hirumi and Hirumi 1989) with appropriate drug selection (http://tryps.rockefeller.edu/trypsru2_genetics.html). PFs of wild-type Lister 427 were grown in SDM-79 (Brun and Schonenberger 1979) containing 10% fetal bovine serum and 0.25% hemin. Transfections were performed using a Nucleofector (Amaxa) as described previously (Scahill et al. 2008).

ChIP-seq and data analysis

BF or PF (10^8) were collected for each ChIP and lysed as described elsewhere (Lee et al. 2006). DNA was sonicated for 10 cycles (30 sec on/30 sec off) using a Bioruptor (Wolf Laboratories Limited). Immunoprecipitations (IPs) were performed overnight with 10 μg of Ab coupled to M-280 magnetic beads (Dynal). Beads were then washed with RIPA buffer (50 mM HEPES-KOH, pKa 7.55, 500 mM LiCl, 1 mM EDTA, 1.0% NP-40, 0.7% Na-Deoxycholate) and eluted DNA was prepared for sequencing as recommended by Illumina. The following antibodies were used: rabbit anti-H4K10ac and anti-H4 (Cross laboratory), BB2 (kind gift of Keith Gull) (Bastin et al. 1996), HA clone 12cA5, H3 (Abcam, 1791).

Immunoprecipitated DNA was sequenced using a Solexa sequencer (Illumina). The sequenced DNA tags (32 bp or 36 bp) were annotated based on the *T. brucei* genome (version 4) using the Blast-Like Alignment Tool (BLAT) (Kent 2002) allowing ≤2 mismatches. Default parameters were used except for tileSize (sets the size of match that triggers an alignment) = 10 bp, stepSize (spacing between tiles) = 5, and minScore (sets the minimum score and consists of matches minus mismatches minus a gap penalty) = 15. The number of hits per nucleotide position was determined using custom PERL scripts and displayed using MATLAB (The MathWorks).

DNA sequence analysis

Sequences under double peaks of H2AZ (divergent SSRs) and single peaks (novel putative TSSs) were extracted from the *T. brucei* genome. Peaks close to repetitive sequence, mainly in subtelomeric and centromeric regions, and poorly annotated regions were excluded, retaining 128 peaks for further analysis. Each peak was fitted to a Gaussian basis function using nonlinear regression, and its peak width, height, and location recorded. Extracted sequences were analyzed by exhaustive enumeration (up to length 7) and submitted to Trawler, a deterministic optimization algorithm (Ettwiller et al. 2007; methods reviewed in D'haeseleer 2006; Tompa et al. 2005). Exhaustive enumeration

yielded a number of G-rich motifs, and Trawler yielded a G_n cluster as the only statistically significant motif distributed over many peaks. In both methods any genome sequence for which the H2AZ density was <10% of its typical TSS peak density served as the reference sequence.

Plasmid construction and transgenic cell lines

H2AZ, H2BV, and H3V ChIP-seq analyses were performed using previously published Ty1-tagged cell lines (Lowell and Cross 2004; Lowell et al. 2005). For H4V ChIP-seq, a cell line containing an ectopic copy of Ty1-tagged H4V (cloned into pLew82) was generated and both endogenous H4V alleles were deleted using a PCR-based approach (Gaud et al. 1997). Cell lines containing C-terminally triple-HA-tagged BDF1-5 were generated using a PCR-based approach (Oberholzer et al. 2006). For the BDF3 ChIP-seq, we used the triple-HA-tagged BDF3 cell line described above and deleted the untagged BDF3 allele using a PCR-based approach. The cell line for inducible RNAi-mediated depletion of BDF3 was generated as described previously (Alibu et al. 2005). An ~450-bp fragment of BDF3 was PCR-amplified and ligated into p2T7^{TAblue}, and the vector was transfected into a cell line containing a triple-HA-tagged BDF3. Efficient RNAi-mediated BDF3 depletion was monitored over 27 h by Western blot (data not shown).

IF microscopy

IF microscopy and evaluation of epitope colocalization was performed as described elsewhere (Siegel et al. 2008b). Antibodies were used at the following concentrations: BB2 (Bastin et al. 1996), 1:2000; HA clone 16b12 (Covance), 1:1000; H4K10ac 1:4000; H4, 1:4000.

Co-IP

IP of tagged Ty1-tagged H2AZ or H2A or Flag-tagged H2BV or H2B was performed as described elsewhere (Lowell et al. 2005), except that prior to the IP the chromatin was washed with buffer containing increased concentrations of NaCl. IPs were performed using monoclonal BB2 or polyclonal rabbit anti-Flag antibody (F7425, Sigma). Coimmunoprecipitated histones were separated by SDS-PAGE and stained with amido black or immunoblotted using the following antibodies: BB2, anti-Flag, anti-H4, anti-H3.

Accession numbers

TrypDB (http://www.genedb.org/genedb/tryp) reference numbers for BDFs discussed in this study are Tb10.6k15.2370 (BDF1), Tb10.6k15.3240 (BDF2), Tb11.01.1830 (BDF3), Tb927.7.4380 (BDF4), and Tb11.01.5000 (BDF5).

Acknowledgments

References

Alibu, V.P., Storm, L., Haile, S., Clayton, C., and Horn, D. 2005. A doubly inducible system for RNA interference and rapid RNAi plasmid construction in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **139:** 75–82.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I. and K., Z. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* **129:** 823–837.

Bastin, P., Bagherzadeh, A., Matthews, K.R., and Gull, K. 1996. A novel epitope tag system to study protein targeting and organelle biogenesis in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **77:** 235–239.

Ben Amar, M.F., Jefferies, D., Pays, A., Bakalara, N., Kendall, G. and E,P. 1991. The actin gene promoter of *Trypanosoma brucei*. *Nucleic Acids Res.* **19:** 5857–5862.

Brun, R. and Schonenberger, M. 1979. Cultivation and in vitro cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Short communication. *Acta Trop.* **36:** 289–292.

Clayton, C.E. 2002. Life without transcriptional control? From fly to man and back again. *EMBO J.* **21:** 1881–1888.

Das, A., Zhang, Q., Palenchar, J.B., Chatterjee, B., Cross, G.A.M., and Bellofatto, V. 2005. Trypanosomal TBP functions with the multisubunit transcription factor tSNAP to direct spliced-leader RNA gene expression. *Mol. Cell. Biol.* **25:** 7314–7322.

D'haeseleer, P. 2006. How does DNA sequence motif discovery work? *Nat. Biotechnol.* **24:** 959–961.

Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K., and Zhou, M.M. 1999. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399:** 491–496.

El-Sayed, N.M., Myler, P.J., Blandin, G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renauld, H., Worthey, E.A., Hertz-Fowler, C., et al. 2005. Comparative genomics of trypanosomatid parasitic protozoa. *Science* **309:** 404–409.

Ettwiller, L., Paten, B., Ramialison, M., Birney, E., and Wittbrodt, J. 2007. Trawler: De novo regulatory motif discovery pipeline for chromatin immunoprecipitation. *Nat. Methods* **4:** 563–565.

Etzioni, S., Yafe, A., Khateb, S., Weisman-Shomer, P., Bengal, E., and Fry, M. 2005. Homodimeric MyoD preferentially binds tetraplex structures of regulatory sequences of muscle-specific genes. *J. Biol. Chem.* **280:** 26805–26812.

Gaud, A., Carrington, M., Deshusses, J., and Schaller, D.R. 1997. Polymerase chain reaction-based gene disruption in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **87:** 113–115.

Guillemette, B., Bataille, A.R., Gevry, N., Adam, M., Blanchette, M., Robert, F., and Gaudreau, L. 2005. Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. *PLoS Biol.* **3:** e384. doi: 10.1371/journal.pbio.0030304.

Hassan, A.H., Prochasson, P., Neely, K.E., Galasinski, S.C., Chandy, M., Carrozza, M.J., and Workman, J.L. 2002. Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell* **111:** 369–379.

Hausler, T. and Clayton, C. 1996. Post-transcriptional control of hsp70 mRNA in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **76:** 57–71.

Henikoff, S. 2008. Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat. Rev. Genet.* **9:** 15–26.

Hirumi, H. and Hirumi, K. 1989. Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *J. Parasitol.* **75:** 985–989.

Hull, M.W., Erickson, J., Johnston, M., and Engelke, D.R. 1994. tRNA genes as transcriptional repressor elements. *Mol. Cell. Biol.* **14:** 1266–1277.

Huppert, J.L. and Balasubramanian, S. 2007. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* **35:** 406–413.

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

Janzen, C.J., Fernandez, J.P., Deng, H., Diaz, R., Hake, S.B., and Cross, G.A.M. 2006. Unusual histone modifications in *Trypanosoma brucei. FEBS Lett.* **580:** 2306–2310.

Jin, C. and Felsenfeld, G. 2007. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes & Dev.* **21:** 1519–1529.

Kawahara, T., Siegel, T.N., Ingram, A.K., Alsford, S., Cross, G.A.M., and Horn, D. 2008. Two essential MYST-family proteins display distinct roles in histone H4K10 acetylation and telomeric silencing in trypanosomes. *Mol. Microbiol.* **69:** 1054–1068.

Kendall, A., Hull, M.W., Bertrand, E., Good, P.D., Singer, R.H., and Engelke, D.R. 2000. A CBF5 mutation that disrupts nucleolar localization of early tRNA biosynthesis in yeast also suppresses tRNA gene-mediated transcriptional silencing. *Proc. Natl. Acad. Sci.* **97:** 13108–13113.

Kent, W.J. 2002. BLAT—The BLAST-like alignment tool. *Genome Res.* **12:** 656–664.

Korber, P. and Horz, W. 2004. SWRred not shaken; Mixing the histones. *Cell* **117:** 5–7.

Lee, M.G. 1996. An RNA polymerase II promoter in the hsp70 locus of *Trypanosoma brucei. Mol. Cell. Biol.* **16:** 1220–1230.

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

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

Lowell, J.E. and Cross, G.A.M. 2004. A variant histone H3 is enriched at telomeres in *Trypanosoma brucei. J. Cell Sci.* **117:** 5937–5947.

Lowell, J.E., Kaiser, F., Janzen, C.J., and Cross, G.A.M. 2005. Histone H2AZ dimerizes with a novel variant H2B and is enriched at repetitive DNA in *Trypanosoma brucei. J. Cell Sci.* **118:** 5721–5730.

Mandava, V., Fernandez, J.P., Deng, H., Janzen, C.J., Hake, S.B., and Cross, G.A.M. 2007. Histone modifications in *Trypanosoma brucei. Mol. Biochem. Parasitol.* **156:** 41–50.

Mandava, V., Janzen, C.J., and Cross, G.A.M. 2008. Trypanosome H2Bv replaces H2B in nucleosomes enriched for H3 K4 and K76 trimethylation. *Biochem. Biophys. Res. Commun.* **368:** 846–851.

Marchetti, M.A., Tschudi, C., Silva, E., and Ullu, E. 1998. Physical and transcriptional analysis of the *Trypanosoma brucei* genome reveals a typical eukaryotic arrangement with close interspersionof RNA polymerase II- and III-transcribed genes. *Nucleic Acids Res.* **26:** 3591–3598.

Margottin, F., Dujardin, G., Gerard, M., Egly, J.M., Huet, J., and Sentenac, A. 1991. Participation of the TATA factor in transcription of the yeast U6 gene by RNA polymerase C. *Science* **251:** 424–426.

Martinez-Calvillo, S., Yan, S., Nguyen, D., Fox, M., Stuart, K., and Myler, P.J. 2003. Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. *Mol. Cell* **11:** 1291–1299.

McAndrew, M., Graham, S., Hartmann, C., and Clayton, C. 1998. Testing promoter activity in the trypanosome genome: Isolation of a metacyclic-type VSG promoter, and unexpected insights into RNA polymerase II transcription. *Exp. Parasitol.* **90:** 65–76.

Mito, Y., Henikoff, J.G., and Henikoff, S. 2005. Genome-scale profiling of histone H3.3 replacement patterns. *Nat. Genet.* **37:** 1090–1097.

Nakaar, V., Dare, A.O., Hong, D., Ullu, E., and Tschudi, C. 1994. Upstream tRNA genes are essential for expression of small nuclear and cytoplasmic RNA genes in trypanosomes. *Mol. Cell. Biol.* **14:** 6736–6742.

Nakaar, V., Gunzl, A., Ullu, E., and Tschudi, C. 1997. Structure of the *Trypanosoma brucei* U6 snRNA gene promoter. *Mol. Biochem. Parasitol.* **88:** 13–23.

Noma, K., Cam, H.P., Maraia, R.J., and Grewal, S.I. 2006. A role for TFIIIC transcription factor complex in genome organization. *Cell* **125:** 859–872.

Oberholzer, M., Morand, S., Kunz, S., and Seebeck, T. 2006. A vector series for rapid PCR-mediated C-terminal in situ tagging of *Trypanosoma brucei* genes. *Mol. Biochem. Parasitol.* **145:** 117–120.

Pillus, L. 2008. MYSTs mark chromatin for chromosomal functions. *Curr. Opin. Cell Biol.* **20:** 326–333.

Raisner, R.M., Hartley, P.D., Meneghini, M.D., Bao, M.Z., Liu, C.L., Schreiber, S.L., Rando, O.J., and Madhani, H.D. 2005. Histone variant H2A.Z marks the 5′ ends of both active and inactive genes in euchromatin. *Cell* **123:** 233–248.

Respuela, P., Ferella, M., Rada-Iglesias, A., and Aslund, L. 2008. Histone acetylation and methylation at sites initiating divergent polycistronic transcription in *Trypanosoma cruzi. J. Biol. Chem.* **283:** 15884–15892.

Scahill, M.D., Pastar, I., and Cross, G.A.M. 2008. CRE recombinase-based positive-negative selection systems for genetic manipulation in *Trypanosoma brucei. Mol. Biochem. Parasitol.* **157:** 73–82.

Schimanski, B., Nguyen, T.N., and Gunzl, A. 2005. Characterization of a multisubunit transcription factor complex essential for spliced-leader RNA gene transcription in *Trypanosoma brucei. Mol. Cell. Biol.* **25:** 7303–7313.

Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. 2008. Divergent transcription from active promoters. *Science* **322:** 1849–1851.

Shia, W.J., Li, B., and Workman, J.L. 2006. SAS-mediated acetylation of histone H4 Lys 16 is required for H2A.Z incorporation at subtelomeric regions in *Saccharomyces cerevisiae. Genes & Dev.* **20:** 2507–2512.

Shilatifard, A. 2008. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Curr. Opin. Cell Biol.* **20:** 341–348.

Siegel, T.N., Kawahara, T., Degrasse, J.A., Janzen, C.J., Horn, D., and Cross, G.A.M. 2008a. Acetylation of histone H4K4 is cell cycle regulated and mediated by HAT3 in *Trypanosoma brucei. Mol. Microbiol.* **67:** 762–771.

Siegel, T.N., Hekstra, D.R., and Cross, G.A.M. 2008b. Analysis of the *Trypanosoma brucei* cell cycle by quantitative DAPI imaging. *Mol. Biochem. Parasitol.* **160:** 171–174.

Tompa, M., Li, N., Bailey, T.L., Church, G.M., De Moor, B., Eskin, E., Favorov, A.V., Frith, M.C., Fu, Y., Kent, W.J., et al. 2005. Assessing computational tools for the discovery of transcription factor binding sites. *Nat. Biotechnol.* **23:** 137–144.

Tornaletti, S., Park-Snyder, S., and Hanawalt, P.C. 2008. G4-forming sequences in the non-transcribed DNA strand pose blocks to T7 RNA polymerase and mammalian RNA polymerase II. *J. Biol. Chem.* **283:** 12756–12762.

Weintraub, H. and Groudine, M. 1976. Chromosomal subunits in active genes have an altered conformation. *Science* **193:** 848–856.

Winston, F. and Allis, C.D. 1999. The bromodomain: A chromatin-targeting module? *Nat. Struct. Biol.* **6:** 601–604.

Wirtz, E., Leal, S., Ochatt, C., and Cross, G.A.M. 1999. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei. Mol. Biochem. Parasitol.* **99:** 89–101.

Zhang, H., Roberts, D.N., and Cairns, B.R. 2005. Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* **123:** 219–231.

Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*

T. Nicolai Siegel, Doeke R. Hekstra, Louise E. Kemp, et al.

Supplemental Material	http://genesdev.cshlp.org/content/suppl/2009/04/16/gad.1790409.DC1
Related Content	**Chromatin-based transcriptional punctuation** Paul B. Talbert and Steven Henikoff Genes Dev. May , 2009 23: 1037-1041
References	This article cites 59 articles, 21 of which can be accessed free at: http://genesdev.cshlp.org/content/23/9/1063.full.html#ref-list-1 Articles cited in: http://genesdev.cshlp.org/content/23/9/1063.full.html#related-urls
License	
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**