Molecular Biology

**TbRAP1 has an unusual duplex DNA binding activity required for its telomere localization and VSG silencing**

Marjia Afrin1*, Amit Kumar Gaurav1*, Xian Yang2, Xuehua Pan2, Yanxiang Zhao2t, Bibo Li1,3,4†

Localization of Repressor Activator Protein 1 (RAP1) to the telomere is essential for its telomeric functions. RAP1 homologs either directly bind the duplex telomere DNA or interact with telomere-binding proteins. We find that *Trypanosoma brucei* RAP1 relies on a unique double-stranded DNA (dsDNA) binding activity to achieve this goal. *T. brucei* causes human sleeping sickness and regularly switches its major surface antigen, variant surface glycoprotein (VSG), to evade the host immune response. VSGs are monoallelically expressed from subtelomeres, and *TbRAP1* is essential for VSG regulation. We identify dsDNA and single-stranded DNA binding activities in *TbRAP1*, which require positively charged RKR residues that overlap with *TbRAP1*’s nuclear localization signal in the MybLike domain. Both DNA binding activities are electrostatics-based and sequence nonspecific. The dsDNA binding activity can be substantially diminished by phosphorylation of two RKR residues adjacent S residues and is essential for *TbRAP1*’s telomere localization, VSG silencing, telomere integrity, and cell proliferation.

**INTRODUCTION**

Telomeres, the nucleoprotein complex at chromosome ends, can form a specialized heterochromatic structure that suppresses expression of genes located at the subtelomere, which is known as telomere position effect or telomeric silencing (1). Among known telomere core components, Repressor Activator Protein 1 (RAP1) is one of the most conserved, with homologs identified from protozoa to mammals (2–6). Although RAP1 homologs do not all have the same functions, most have been shown to play key roles to protect the chromosome end, maintain stable telomere length, and establish/maintain the telomeric silencing (7).

The telomere function of RAP1 homologs depends on their localization at the telomere. Most RAP1 homologs do not have any direct DNA binding activity, despite the presence of a Myb domain that typically has DNA binding activities (8). Instead, these RAP1s are tethered to the telomere through interaction with other telomere-binding proteins, such as TTAGGG repeat-binding factor 2 (TRF2) in humans (2) and Taz1 in *Schizosaccharomyces pombe* (4). So far, budding yeast RAP1s, including *Saccharomyces cerevisiae* RAPI, are the only ones that bind the duplex telomere DNA, recognizing a consensus sequence 5′ ACACCCAYACAYY 3′ (where Y represents a pyrimidine) (9) using both its Myb and MybLike domains (10). The DNA binding domain of ScRAP1 is the only region essential for cell viability (11).

We have identified a RAP1 homolog in *Trypanosoma brucei*, a protozoan parasite that causes human African trypanosomiasis. The bloodstream form (BF) *T. brucei* proliferates in the extracellular space of the mammalian host and regularly switches its major surface antigen, variant surface glycoprotein (VSG), to evade the host immune response (12). There are ~2500 VSG genes and pseudogenes in the *T. brucei* genome (13). However, VSG is expressed exclusively from subtelomeric VSG expression sites (ESSs), in which VSG is located within 2 kb from the telomere repeats (14). *T. brucei* has multiple BF ESs (14), but only one is fully transcribed at any time, presenting a single type of VSG on the cell surface (15). Monoallelic VSG expression is regulated by multiple factors, such as chromatin structure, subnuclear localization of the VSG transcription site, inositol phosphate pathway, a subtelomere and VSG-associated VSG exclusion (VEX) complex, and telomeric silencing (6, 16–19). VSG switching has two major pathways (20). A coupled silencing of the active ES and derepression of a silent ES leads to an in situ switch, and a silent VSG gene can be recombined into the active ES to replace the originally active VSG. Telomere proteins (21–23) and many factors important for homologous recombination, DNA damage repair, and DNA replication have been shown to influence VSG switching frequencies (24).

We identified *TbRAP1* as a *TbTRF*-interaction factor (6), while *TbTRF* binds the duplex TTAGGG repeats of *T. brucei* telomere (25). *TbRAP1* associates with the telomere chromatin and is partially co-localized with *TbTRF* (6). We previously showed that depletion of *TbRAP1* leads to a marked derepression of all ES-linked VSGs (6, 19). The *TbRAP1*-mediated silencing has a stronger effect on telomere proximal genes than those located further away, suggesting that localization of *TbRAP1* at the telomere is essential for this silencing (6). Our previous study showed that association of *TbRAP1* with the telomere chromatin helps suppress the expression of telomeric transcript [telomeric repeat-containing RNA (TERRA)], while an increased amount of TERRA and telomeric R-loops in *TbRAP1*-depleted cells leads to more telomeric/subtelomeric DNA damage (23). These findings suggest that the telomere localization of *TbRAP1* is also a prerequisite for telomere integrity. However, the mechanism of recruiting *TbRAP1* to the telomere was unknown.

*TbRAP1* has both a Myb and a MybLike domain (6), but their roles in targeting *TbRAP1* to the telomere have not been investigated, even though the Myb domain is a common structural motif for DNA binding activities (8). In this study, we find that the telomere...
localization of \( T\text{bRAP1} \) is independent of its Myb domain and \( T\text{bTRF} \). Unexpectedly, within the \( T\text{bRAP1} \) MybLike domain and overlapping with nuclear localization signal (NLS), we identify \( 73\text{RKKRRR}_{74} \), a group of positively charged residues, to have both double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA) binding activities. We show that \( 73\text{RKKRRR}_{74} \) is required for the association of \( T\text{bRAP1} \) with the telomere chromatin, VSG silencing, telomere integrity, and normal cell growth. Phosphorylation of the R/K patch–adjacent S742 and S744 was detected in \( T. \text{brucei} \) cells (26, 27). We further demonstrate that the phospho-mimicking mutation of these S residues substantially diminishes \( T\text{bRAP1} \)’s dsDNA binding activity in vitro and abolishes \( T\text{bRAP1} \)’s telomere localization, causes VSG derepression, accumulates telomere/subtelomere DNA damage, and leads to cell growth arrest. Our results indicate that the dsDNA binding activity of \( T\text{bRAP1} \) is essential and can be regulated by postranslational modification of \( T\text{bRAP1} \).

RESULTS

Localization of \( T\text{bRAP1} \) to the telomere is independent of \( T\text{bTRF} \) and the Myb domain

\( T\text{bRAP1} \) is an intrinsic component of the \( T. \text{brucei} \) telomere complex (6), but how \( T\text{bRAP1} \) is located to the telomere was unknown. Because \( T\text{bTRF} \) binds the duplex telomere DNA and interacts with \( T\text{bRAP1} \) (6, 25), we tested whether \( T\text{bRAP1} \) is recruited to the telomere by \( T\text{bTRF} \). A hemagglutinin (HA) monoclonal antibody 12CA5 was used to chromatin immunoprecipitate (ChIP) the FLAG-HA-HA (F2H)–tagged \( T\text{bRAP1} \) in \( T\text{bRAP1}^{12H+/−} \) \( T\text{bTRF} \) RNA interference (RNAi) cells (table S1 lists all strains used in this study). Before and after the depletion of \( T\text{bTRF} \) (Fig. 1A), \( T\text{bRAP1} \) associated with the telomere chromatin at nearly the same level (Fig. 1B). As a control, \( T\text{bTRF} \) was no longer at the telomere after the RNAi induction (Fig. 1B). Therefore, the localization of \( T\text{bRAP1} \) at the telomere is \( T\text{bTRF} \) independent.

\( T\text{bRAP1} \) has a putative Myb domain (Fig. 1C) (6). While Myb motifs frequently bind DNA, the role of \( T\text{bRAP1} \)’s Myb domain in localizing \( T\text{bRAP1} \) to the telomere was unknown. We have established the \( T\text{bRAP1}^{12\Delta M} \) strain, in which one \( T\text{bRAP1} \) allele was replaced with the \( \Delta \text{Myb} \) (\( \Delta M \)) mutant, and the other (\( \text{F} \) allele) was flanked by loxP repeats so that it can be deleted when Cre is expressed (28). F2H-\( T\text{bRAP1}^{12H+/−} \) and Cre induction depleted the wild-type (WT) \( T\text{bRAP1} \) protein within 30 hours (28). In \( T\text{bRAP1}^{12H+/−} \) and Cre-expressing \( T\text{bRAP1}^{12\Delta M} \) cells, CHiP using the HA antibody 12CA5 and a rabbit \( T\text{bTRF} \) antibody (25) showed that F2H-\( T\text{bRAP1}^{12H+/−} \) and \( T\text{bTRF} \) associated with the telomere chromatin in both cells (Fig. 1D). Therefore, the Myb domain is not necessary for localizing \( T\text{bRAP1} \) to the telomere.

\( T\text{bRAP1} \) has electrostatics-based DNA binding activities that rely on \( 73\text{RKKRRR}_{74} \)

\( T\text{bRAP1} \) also has a MybLike domain [amino acids (aa) 639 to 761] (Fig. 1C) (6), which contains a positively charged \( 73\text{RKKRRR}_{74} \) patch. To test whether this domain has any DNA binding activity, we partially purified the recombinant TrxA-His6–tagged \( T\text{bRAP1}^{639−761} \), \( T\text{bRAP1}^{639−733} \), and \( T\text{bRAP1}^{734−761} \) from \( E. \text{coli} \) (fig. S1A; table S2 lists all recombinant proteins used in this study) and performed electrophoretic mobility shift assay (EMSA). \( T\text{bRAP1}^{639−761} \) and \( T\text{bRAP1}^{734−761} \) bound a dsDNA containing (TTAGGG)$_{12}$, a 100–base pair (bp) dsDNA and a 100-nucleotide (nt) ssDNA containing a random sequence (Fig. 2, A to C), while \( T\text{bRAP1}^{639−733} \) or TrxA-His6 did not bind these DNA substrates (Fig. 3E and fig. S1, B to E) (table S3 lists the sequences of all EMSA substrates used in this study). In addition, dsDNA with either (TTAGGG)$_{12}$ or a random sequence competed for \( T\text{bRAP1}^{639−761} \)’s binding when a (TTAGGG)$_{12}$–containing dsDNA was used as the probe (fig. S1M), indicating that the DNA binding activity is sequence nonspecific. We also tested the DNA binding activity of glutathione S-transferase (GST)–tagged \( T\text{bRAP1}^{114−485} \) (fig. S1A) and got the same results (fig. S1, F to H). Therefore, \( T\text{bRAP1} \) has sequence-nonspecific dsDNA and ssDNA binding activities in the aa 734 to 761 region, which we named DNA binding (DB; Fig. 1C).

To pinpoint which residues are critical for \( T\text{bRAP1} \)’s DNA binding activities, we used nuclear magnetic resonance (NMR) to analyze the heteronuclear single-quantum correlation (HSQC) spectrum of \( ^{15}\text{N} \)-labeled \( T\text{bRAP1}^{639−761} \) in the presence and the absence of DNA substrates (Fig. 2D and fig. S1I). The HSQC spectra showed that positively charged residues K738 and R741 within the \( 73\text{RKKRRR}_{74} \) patch (the R/K patch) underwent notable chemical shifts when a dsDNA with a random sequence was added (Fig. 2D). Similar patterns of chemical shifts were observed when ds(TTAGGG)$_{3}$ was added (fig. S1I). These data suggest that the R/K patch is directly responsible for the \( T\text{bRAP1} \)’s dsDNA binding. Notably, the HSQC signal for many residues in the MybLike domain preceding the R/K patch disappeared after the addition of DNA substrates (Fig. 2D and fig. S1I). This is likely due to the broadening of the HSQC signal for these residues after the formation of a larger-sized \( T\text{bRAP1} \)–DNA complex rather than direct interaction between these residues and the DNA substrate.

To confirm the importance of the R/K patch in DNA binding, we generated a \( T\text{bRAP1}^{639−761} \)5A mutant with all five R and K residues replaced by A (fig. S1A). No chemical shifts were observed when dsDNA of either random or telomeric sequence was added to the \( ^{15}\text{N} \)-labeled 5A mutant (Fig. 2E and fig. S1J). Therefore, \( 73\text{RKKRRR}_{74} \) is directly responsible for \( T\text{bRAP1} \)’s DNA binding activities. This finding was further corroborated by EMSA results, where \( T\text{bRAP1}^{639−761} \)5A did not bind dsDNA or ssDNA (Fig. 2, A to C), indicating that these DNA binding activities are based on electrostatic attraction between positively charged \( 73\text{RKKRRR}_{74} \) and the DNA substrates. We further determined \( K_d \) (equilibrium dissociation constant) values of \( T\text{bRAP1}^{639−761} \) binding to a 100-bp dsDNA and a 100-nt ssDNA with a random sequence to be 21.5 M and 310 nM, respectively (fig. S1, K and L).

\( T\text{bRAP1}^{639−761} \) bound dsDNA substrates with an apparent increasing affinity when the substrate length increased from 60 to 150 bp (Fig. 3A). In addition, using an 80-bp random-sequence dsDNA as the substrate, longer dsDNAs competed better than shorter ones (500 > 200 > 100 > 60 bp) (Fig. 3B). We observed the same preference of \( T\text{bRAP1}^{639−761} \) binding to shorter DNA oligos with weaker affinity (40 < 60 < 80 < 100 nt) (Fig. 3C). Therefore, \( T\text{bRAP1} \) is different from its vertebrate and fission yeast homologs in that it has DNA binding activities. \( T\text{bRAP1} \) is also different from its budding yeast homologs in that its DNA binding activities are electrostatics based, sequence nonspecific, and substrate size dependent. Furthermore, the R/K patch is within \( T\text{bRAP1} \)’s NLS (aa 727 to 741; Fig. 1C) (28). Hence, \( T\text{bRAP1} \) is unique among its homologs in that its NLS has dual roles in nuclear import and DNA binding.
Phospho-mimicking mutations of the R/K patch–adjacent S residues significantly affect TbRAP1’s dsDNA binding activity

TbRAP1’s DNA binding activities rely on the R/K patch and are apparently electrostatics based. Phospho-proteomics analyses showed that S742 and S744 of TbRAP1 are phosphorylated in T. brucei cells (26, 27). Because phosphorylation adds negative charges to the local environment, we speculate that phosphorylation of S742 and S744 may interfere with DNA binding. To test this, we expressed S742AS744A (2SA) and S742DS744D (2SD) mutants (fig. S1A) of TbRAP1639–761 to mimic nonphosphorylated and phosphorylated states, respectively. EMSA showed that TbRAP1639–7612SD lost nearly all the dsDNA binding activity, while TbRAP1639–7612SA still retained most of it (Fig. 3D). Therefore, TbRAP1’s dsDNA binding activity is likely sensitive to phosphorylation of both S residues. Both 2SA and 2SD mutants still bound the ssDNA, although 2SD has a slightly weaker activity than WT TbRAP1639–761 (Fig. 3E). With specific and substantial reduction in the dsDNA binding yet minimal impact on the ssDNA binding, the 2SD mutant allows us to differentiate the functional significance of these two DNA binding activities in vivo. In addition, single mutants TbRAP1639–761S742D and TbRAP1639–761S744D (fig. S1B) bound both dsDNA and ssDNA the same as WT TbRAP1 (fig. S1O), suggesting that phosphorylation of both S residues is necessary to exert a detectable effect on dsDNA binding.

In vivo telomere localization of TbRAP1 requires the R/K patch and is disrupted by phospho-mimicking mutation of adjacent S residues

We have established a Cre-loxP–mediated conditional deletion system for TbRAP1 (28). In TbRAP1F2H+/−, we replaced the WT TbRAP1 allele with F2H-tagged DB domain mutants to generate TbRAP1F2Hmut strains (fig. S2A). For mutants missing the functional TbRAP1 NLS (aa 727 to 741), we added an N-terminal SV40 NLS, which is sufficient for nuclear import of TbRAP1 (28). Genotypes of TbRAP1F2Hmut strains were confirmed by Southern (fig. S2B) and sequencing analyses.

To examine whether F2H-tagged TbRAP1 mutants are associated with the telomere chromatin, we performed ChIP using the HA antibody 12CA5 in TbRAP1F2Hmut cells. TbRAP1 self-interacts through its BRCA1 C-terminus (BRCT) domain (Fig. 1C) (28). Hence, we removed the WT TbRAP1 allele by Cre [confirmed by polymerase chain reaction (PCR); fig. S2C] to specifically examine TbRAP1 mutants’ behavior without the influence from the WT protein. As a control, ChIP was done in TbRAP1F2H+/− using the 12CA5 antibody. All TbRAP1 mutants were expressed at the same level as F2H-TbRAP1, except ∆MybLike (∆ML) at a subtly lower level (Fig. 4A). Only a residual amount of TbRAP1∆ML (Fig. 4B), ∆DB (Fig. 4C), and 5A (Fig. 4D) mutants associated with the telomere chromatin, which was much lower than F2H-TbRAP1. TbTRF was still at the telomere in these mutants (Fig. 4, B to D). Immunofluorescence (IF) analyses were done to examine the subnuclear localization of TbRAP1 point mutants. In TbRAP1F2H+/− cells, F2H-TbRAP1 partially colocalized with TbTRF (Fig. 4E, top), the same as we reported previously (6). However, TbRAP1-5A was no longer colocalized with TbTRF, even though it was imported into the nucleus via the SV40 NLS (Fig. 4E, bottom). Therefore, the R/K patch is required for the telomere localization of TbRAP1.

In addition, significantly less TbRAP1-2SD was associated with the telomere chromatin than F2H-TbRAP1 (Fig. 4F). IF also showed that TbRAP1-2SD did not colocalize with TbTRF (Fig. 4G), indicating that the TbRAP1’s dsDNA binding activity is critical for its localization at the telomere. In contrast, mutation of TbRAP1’s phosphorylation sites to A did not affect its telomere localization. Five TbRAP1 residues were found to be phosphorylated in T. brucei cells (26, 27), and the TbRAP1F2H/5SA strain was established previously with all five phosphorylation sites mutated to A (S265AS586AS742AS744AT752A). To specifically investigate the function of S742 and S744, we also established TbRAP1F2H/5SA with only S742AS744A mutations.
Both TbRAP1-5SA and 2SA mutants associated with the telomere chromatin (Fig. 4H and fig. S2D). IF showed that 5SA and 2SA both partially colocalized with TbTRF (Fig. 4I and fig. S2E). Therefore, TbRAP1 with unphosphorylated S742 and S744 is localized at the telomere, while phosphorylation of both S742 and S744 can remove TbRAP1 from the telomere.

The telomere localization of TbRAP1 is essential for normal cell growth

We recently showed that the MybLike domain is essential for normal cell proliferation (28). To determine the functions of TbRAP1’s DNA binding activities specifically, we first examined cell growth in TbRAP1 F/mut cells carrying DB mutations after induction of Cre. F2H-tagged TbRAP1 mutants were detected by the HA Probe antibody. The expression of total TbRAP1 was also examined by a rabbit antibody (6) that recognizes the MybLike domain (28) and, specifically, TbRAP1639–761 but not TbRAP1 734–761 (fig. S2F).

Western analysis showed the deletion of WT TbRAP1 and a persistent expression of F2H-NLS–tagged TbRAP1-∆DB (fig. S2G), 5A (Fig. 5A), and 2SD (Fig. 5C) mutants and F2H-tagged 5SA (fig. S2I) and 2SA (Fig. 5E) mutants in various TbRAP1 F/mut cells upon Cre induction. TbRAP1 F/∆DB (fig. S2H) and TbRAP1 F/5SA (Fig. 5B) cells exhibited a growth arrest after inducing Cre for 24 to 30 hours, indicating that TbRAP1’s DNA binding activities are essential for normal cell growth. TbRAP1 F/2SD cells showed a growth arrest upon Cre induction (Fig. 5D). Since the 2SD mutant disrupted most of
the dsDNA binding without affecting ssDNA binding significantly (Fig. 3, D and E), this observation indicates that TbRAP1's dsDNA binding is essential for normal cell proliferation. In contrast, TbRAP1<sup>S742A S744A</sup> (fig. S2J) and TbRAP1<sup>S742D S744D</sup> (Fig. 5F) cells only grew mildly slower after the WT TbRAP1 allele was deleted. Therefore, all TbRAP1 mutants that were not localized at the telomere also experienced cell growth arrest, while those that were still at the telomere kept proliferating. 5SA and 2SA are so far the only TbRAP1 mutants capable

![Fig. 3. TbRAP1 has higher affinity for longer DNA substrates.](http://advances.sciencemag.org/)
RAP1 homologs are known to regulate expression of both subtelomeric and nontelomeric genes (3, 29, 30). We suspect that TbRAP1 may affect expression of genes other than VSGs. To examine TbRAP1’s function in global gene expression, we performed RNA sequencing (RNAseq) analysis in TbRAP1 point mutants that did not bind DNA. As a control, we first compared the gene expression profile in TbRAP1+/– and TbRAP1+/+ cells after both were induced for Cre expression for 30 hours. Compared to TbRAP1+/+ cells, >7200 genes were up-regulated in TbRAP1+/– cells (fig. S4A), among which >2500 were VSG genes and pseudogenes (fig. S4B). There are ~2500 VSG genes and pseudogenes in our T. brucei strain (13), suggesting that nearly all VSG genes were derepressed in TbRAP1+/– cells. All BF ES–linked VSGs were derepressed (fig. S5). Some ES-linked ESAGs were up-regulated, some were not affected significantly, and others were down-regulated (fig. S5), indicating that VSGs and ESAGs can be regulated differently by TbRAP1 even when they are located in the same ESs. In addition, >2500 genes were down-regulated upon TbRAP1 deletion, including a number of ribosomal protein genes (fig. S4, A and B). In consistence, Gene Ontology term analysis showed that genes involved in immune evasion were up-regulated, and genes involved in protein synthesis were down-regulated (fig. S4E). However, much fewer genes were down-regulated than up-regulated, and the change in gene expression level is much stronger for up-regulated genes than that for down-regulated ones (fig. S4A), suggesting that TbRAP1 is mainly a repressor. Compared to TbRAP1+/+ cells, in Cre-induced TbRAP1−/– cells, nearly 5300 genes were up-regulated (Fig. 5G), including 2119 VSG genes (fig. S4C), while >1400 genes were down-regulated (Fig. S5), including 66 ribosomal protein genes (fig. S4C). All ES-linked VSGs were derepressed although at various levels (fig. S6). In addition, ES-linked ESAGs were up-regulated, unaffected, or down-regulated by the 5A mutation (fig. S6). Therefore, TbRAP1+/– has a similar transcriptome profile as TbRAP1+/–, indicating that the R/K patch is essential for TbRAP1’s role in gene expression regulation. We observed a similar phenotype in TbRAP1+/2SA cells. Upon Cre induction, nearly 7000 genes were up-regulated (Fig. 5H), including 2574 VSG genes (fig. S4D), and >2100 genes were down-regulated (Fig. 5H), including 59 ribosomal protein genes (fig. S4D). All BF ES–linked VSGs were up-regulated although at various levels (fig. S7), and ESAGs were similarly affected as in TbRAP1+/– and TbRAP1+/SA cells (fig. S7). Therefore, TbRAP1’s dsDNA binding activity is required for its role in gene expression regulation.

In contrast, 5A and 2SA mutants exhibited only mild VSG derepression. In TbRAP1+/5SA and TbRAP1+/2SA cells, a number of silent VSGs were derepressed up to several 10-fold when analyzed by qRT-PCR after Cre induction for 30 hours (Fig. 5I and fig. S3C). Furthermore, IF analysis using a VSG3 mouse antibody and a VSG6 rabbit antibody showed that these initially silent VSGs were both expressed in the same individual cells upon Cre induction, although a large fraction of the proteins was not deposited on the cell surface (Fig. 5K and fig. S3E), indicating that 5A and 2SA caused VSG derepression. Since these mutants continued to proliferate, TbRAP1−/– 5SA and TbRAP1−/2SA cells (after deleting the WT TbRAP1 allele by Cre) were cultured continuously. qRT-PCR showed that silent VSGs were expressed at a higher level in these cells than in TbRAP1+/+ cells (fig. 5J and fig. S3D), indicating that VSG derepression is not just a transient phenotype in these mutants. Although 5SA and 2SA mutants are located at the telomere (Fig. 4, H and I, and fig. S2, D and E), 2SA has a mildly weaker dsDNA binding activity than the...
Fig. 5. TbRAP1’s dsDNA binding activity is required for VSG silencing and normal cell growth. (A, C, and E) Western analysis of protein extracted from TbRAP1F/5A (A), TbRAP1F/2SD (C), and TbRAP1F/2SA (E) cells before and after induction of Cre for 30 hours using the HA monoclonal antibody HA Probe (Santa Cruz Biotechnologies), a TbRAP1 rabbit antibody (6), and the tubulin antibody TAT-1 (top three rows). To separate TbRAP1 mutants and the endogenous TbRAP1, proteins were run on a 7.5% tris polyacrylamide gel for 7 hours and detected by the TbRAP1 rabbit antibody (the fourth row). The associated loading control (tubulin blot) is shown at the bottom. (B, D, and F) Growth curves of TbRAP1F/5A (B), TbRAP1F/2SD (D), and TbRAP1F/2SA (F) cells with and without Cre. Average values from three independent experiments were calculated. (G and H) A volcano plot of genes up-regulated and down-regulated in TbRAP1F/5A (G) and TbRAP1F/2SD (H) cells compared to TbRAP1F/+ cells 30 hours after Cre induction. (I and J) qRT-PCR of mRNA levels of the active VSG2, several silent ES-linked VSGs, and chromosome internal TbTERT and ribosomal RNA (rRNA) in TbRAP1F/2SA cells after 30 hours of Cre induction (I) and in TbRAP1F/−/−SA cells (J). The fold changes in mRNA level are shown. Average was calculated from three to six independent experiments. (K) VSG6 and VSG3 expression was examined in TbRAP1F/2SA cells before and after Cre induction (left) and in TbRAP1F/−/−SA cells (right) by IF analyses. A VSG6 rabbit antibody and a VSG3 mouse antibody were used. DNA was stained by DAPI.
WT protein (Fig. 3D). Therefore, these observations further indicate that the dsDNA binding activity is critical for VSG silencing.

The telomere localization of TbRAP1 is required for subtelomere/telomere integrity

We cannot estimate the VSG switching frequency in mutants that exhibited growth arrest, because recovering switchers relies on cell proliferation, and the Cre-loxP–mediated TbRAP1 deletion is not reversible. However, subtelomeric DNA damage, particularly that in the active ES, is a potent inducer of VSG switching (31, 32), and we previously showed that TbRAP1 suppresses VSG switching by maintaining telomere/subtelomere integrity (23). Therefore, we examined DNA damage in all TbRAP1 mutants, using γH2A as an indicator (33). Western blotting showed that the γH2A level increased upon Cre induction in TbRAP1^F/∆ML (Fig. 6A), TbRAP1^F/ΔDB (Fig. 6B), TbRAP1^/F/5A (Fig. 6B), and TbRAP1^/F/25D (Fig. 6C) cells. We also performed γH2A IF in TbRAP1^F/5A and TbRAP1^/F/25D cells before and after Cre induction. There were only few γH2A-positive nuclei (~7%) in both cells before adding Cre (Fig. 6D), and the γH2A signal was faint. In contrast, after Cre induction, more than 90% of nuclei were γH2A positive (Fig. 6D), and the γH2A signal was very bright (Fig. 6E). However, γH2A gave a punctate staining pattern in both induced TbRAP1 mutant cells, suggesting that the increase in DNA damage is not throughout the whole genome. In addition, we detected TbTF in IF as a marker for the telomere and found that γH2A is partially colocalized with TbTF (Fig. 6E), suggesting that some of the DNA damage is at the telomere vicinity. We subsequently performed γH2A ChIP in the 5A and 2SD mutants. Southern hybridization using telomere and tubulin probes following γH2A hybridization using telomere and tubulin probes frequently performed detected DNA damage is not throughout the whole genome. In addition, we detected TbRAP1's NLS in the MybLike domain. We showed that the dsDNA binding activity is essential for TbRAP1's association with the telomere chromatin, VSG silencing, suppressing VSG switching, telomere integrity, and normal cell proliferation. Our study revealed a key and previous unidentified mechanism underlying the essential functions of TbRAP1.

An unusual electrostatics-based DNA binding activity required for targeting TbRAP1 to the telomere

Targeting TbRAP1 to the telomere chromatin requires a unique mechanism among known RAP1s. First, although TbRAP1 interacts with the duplex telomere DNA binding factor TbTRF (6), its localization at the telomere is TbTRF independent, which is different from the scenarios in vertebrates and fission yeast (10, 34). Second, TbRAP1's DNA binding activities are unusual, because it binds to both dsDNA and ssDNA substrates, while RAP1s in budding yeasts bind dsDNA (5, 10). In addition, TbRAP1's DNA binding activities are independent of its Myb domain, while most telomere DNA binding factors—including mammalian TRF1/2, fission yeast Taz1, budding yeast RAP1s, a number of plant telomere repeat–binding factors, and TbTRF—use a Myb-type helix-turn-helix motif to recognize the telomere DNA (25, 35, 36). Furthermore, the TbRAP1's DNA binding activities are electrostatics based and sequence non-specific. In contrast, ScRAP1 recognizes a substrate with the consensus sequence 5′ ACACCAYACAYY 3′ (9).

Mutation of 737 RKRRR 741 to AAAA disrupts the association of TbRAP1 and the telomere chromatin in vivo. However, because 737 RKRRR 741 has both the dsDNA and ssDNA binding activities, it was unclear whether both activities are required for the telomere localization of TbRAP1. Fortunately, the S742DS744D mutation disrupts most dsDNA binding, retains nearly all ssDNA binding, and is not localized at the telomere, indicating that the dsDNA binding activity is the key for localization of TbRAP1 at the telomere. It is interesting that the TbRAP1's DNA binding activities are sequence non-specific, yet ChIP and IF analyses indicate that TbRAP1 is enriched at the telomere much more than at other chromosome loci such as 50-bp repeats upstream of BF ESs. Therefore, additional mechanism(s) is (are) necessary to help TbRAP1 to discriminate different loci. ScRAP1 is well known to bind the promoter of a number of genes and several silencers in addition to the telomere (3), and interaction with different protein partners apparently plays a key role in targeting ScRAP1 to different chromatin loci (7). A similar scenario may apply to TbRAP1, although further investigation is necessary to illustrate the underlying mechanisms.

The dsDNA binding activity of TbRAP1 can be regulated by posttranslational modification

We showed that TbRAP1's DNA binding activities are mediated by electrostatically favorable interaction between the positively charged 737 RKRRR 741 patch and the negatively charged DNA substrate. In addition, the phospho-mimicking S742DS744D mutant substantially disrupts the dsDNA binding activity, presumably due to unfavorable electrostatic interaction between the D residues and the DNA substrate. Phosphoproteomic studies detected phosphorylated TbRAP1 at S742 and S744 residues in both the infectious and insect stages of T. brucei (26, 27). Therefore, phosphorylation of both S742 and S744 in vivo, hence decreasing the amount of positive charges adjacent the R/K patch, can serve as an important mechanism to regulate TbRAP1's dsDNA binding activity and, subsequently, VSG silencing. This regulatory mechanism can be a key for achieving monoallelic VSG expression. So far, it is unclear how TbRAP1 silences all

DISCUSSION

In this study, we identified both dsDNA and ssDNA binding activities of TbRAP1 in a positively charged R/K patch that overlaps with TbRAP1's NLS in the MybLike domain. We showed that the dsDNA binding activity is essential for TbRAP1's association with the telomere chromatin, VSG silencing, suppressing VSG switching, telomere integrity, and normal cell proliferation. Our study revealed a key and previous unidentified mechanism underlying the essential functions of TbRAP1.
subtelomeric VSGs except the one in the active ES. With a means to regulate the dsDNA binding activity, it is possible that TbRAP1 may be prevented from interacting with the active ES DNA and establishing local silencing, thus leaving the active VSG fully transcribed. However, little is known how regulation of TbRAP1’s dsDNA binding activity is achieved in vivo. Quantitative proteomic analysis suggested that phosphorylation of the S residues may be at a very subtly higher level (~1.4-fold) in BF cells than in cells at the insect stage (27). However, it is unknown whether these phosphorylations are cell cycle regulated. Our current study revealed potential physiological significance of these posttranslational modifications of TbRAP1. Further studies, such as identification of the kinase that phosphorylates S742 and S744 and the signals that activate the phosphorylation of these S residues, are key to understand the mechanism of this regulation.

Possible cross-talk between telomere localization and nuclear import of TbRAP1

It is interesting that the R/K patch overlaps with the TbRAP1 NLS (28), indicating that this short peptide has at least two functions. Nuclear import per se mediated by SV40 NLS is not sufficient to ensure telomere localization of TbRAP1. Rather, the dsDNA binding activity is required. The fact that DNA binding activities of TbRAP1 rely on the same peptide that signals for nuclear import suggests that deposition of TbRAP1 to the telomere chromatin may...
be regulated. A recent structural study showed that importin-9 binds histone H2A-H2B and functions more like a storage chaperon (37). After H2A-H2B is transported into the nucleus, RanGTP does not directly dissociate the importin-9 and H2A-H2B interaction. Rather, the RanGTP•importin-9•H2A-H2B complex helps modulate the dissociation of importin-9 from H2A-H2B by DNA and assemble the histones into a nucleosome. It is possible that the interaction of importin α and TbrAP1 NLS may be disrupted only after TbrAP1 binds DNA. However, more detailed structural and functional analyses are necessary to test this hypothesis.

**TbrAP1’s functions in VSG silencing and normal cell proliferation are independent**

TbrAP1 is essential for VSG silencing and normal cell proliferation (6, 19). However, whether the two functions are linked or independent was unknown. Most TbrAP1 mutants simultaneously show defects in both VSG silencing and cell growth, preventing us from independently dissecting the two functions. In addition, many other genes that are involved in VSG silencing are also essential, such as key factors in the inositol phosphate pathway, origin recognition complex 1 (ORC1) and minichromosome maintenance protein complex-binding protein (MCM-BP), the VEX complex, etc. (38).

5SA and 2SA from this study are the only two TbrAP1 mutants capable of cell proliferation, even though they also caused VSG derepression, indicating that VSG silencing and cell growth are not tightly coupled. A recent study showed similar uncoupling between VSG silencing and cell proliferation, where overexpression of Trypanosome DNA-binding Protein 1 (TDP1) leads to simultaneous multiple VSG expression without significantly affecting cell growth (39). Therefore, loss of VSG monoallelic expression per se is not lethal for T. brucei. Nevertheless, TbrAP1 mutants that have defective dsDNA binding are also defective in VSG silencing. For 5SA and 2SA, mildly weaker dsDNA binding activity is most likely the reason for the mild VSG derepression phenotype.

We found that all TbrAP1 mutants that lose the telomere association are defective in cell proliferation and have an increased amount of DNA damage at the telomere vicinity, suggesting that an increased amount of telomere damage contributes to the cell growth defects. ScRAP1’s DNA binding domain is also essential for cell viability (11). It is well known that ScRAP1 is required for transcribing a number of genes encoding ribosomal proteins, and the DNA binding domain is critical for ScRAP1’s transcription activation function (3). Therefore, defective ribosomal protein gene expression may be a major reason why the ScRAP1 null mutant is lethal. In TbrAP1 conditional knockout cells and mutants defective in telomere localization, 1400 to 2600 genes were down-regulated for cell growth and likely another reason why TbrAP1 is essential for cell proliferation. However, further studies are necessary to verify the transcription activation function of TbrAP1.

In this study, we demonstrate that TbrAP1’s dsDNA binding activity is essential for localization of TbrAP1 to the telomere, cell proliferation, VSG silencing, and telomere/subtelomere integrity. This activity depends on a short stretch of positively charged residues that overlaps with TbrAP1’s NLS, which is unique among all known telomere binding factors. Our findings provide a molecular basis that the DNA binding domain of TbrAP1 may serve as a good target for antiparasite agents, as targeting this site can inactivate essential TbrAP1 functions by blocking its nuclear import and disrupting its DNA binding simultaneously.

**MATERIALS AND METHODS**

**T. brucei strains and plasmids**

All T. brucei strains used in this study (listed in table S1) are derived from BF Lister 427 cells that express VSG22 and express the T7 polymerase and the Tet repressor (also known as single marker or SM) (40). All BF T. brucei cells were cultured in HMI-9 medium supplemented with 10% fetal bovine serum and appropriate antibiotics. TbrAP1ΔDB was established previously and described in (28).

All TbrAP1F/+ strains were established using the same strategy. N-terminal F2H-tagged and SV40 NLS-tagged TbrAP1ΔML (deletion of the MybLike domain), TbrAP1ΔDB, TbrAP1-5A, TbrAP1-2SD, and F2H-tagged TbrAP1-5SA and TbrAP1-2SA mutants flanked by sequences upstream and downstream of the TbrAP1 gene, together with a PUR marker, were cloned into pBlueScript SK to generate respective targeting constructs. All mutant-targeting plasmids were digested with Sac I before transfectedainenting the TbrAP1F/+ cells to generate respective TbrAP1F/+ strains. All mutant strains were confirmed by Western and Southern analyses. Point mutations were also validated by sequencing PCR-amplified genomic DNA fragments (one PCR primer is specific to the F2H tag).

For examination of TbrAP1’s association with the telomere chromatin in the presence and absence of TbrTRF, a TbrTRF RNAi strain was first established by transfecting the Not I–digested pZ-JMβ-TbrTRF-Mid1 RNAi construct (25) into SM cells. Subsequently, one endogenous TbrAP1 allele was tagged with an N-terminal F2H tag by transfecting a Sac II–digested pSK-PUR-F2H-TbrAP1-tar2 construct into the TbrTRF RNAi cells. The other TbrAP1 allele was replaced by Hygromycin resistance (HYG) to establish the TbrAP1F2H+/− TbrTRF RNAi strain. Bacterial expression plasmids used in this study are listed in table S2.

**Quantitative reverse transcription polymerase chain reaction**

qRT-PCR experiments were performed, as described in (22).

**Chromatin immunoprecipitation**

Two hundred million cells were cross-linked by 1% formaldehyde for 20 min at room temperature with constant mixing, and the cross-linking was stopped by 0.1 M glycine. Chromatin was sonicated by a Bioruptor for six cycles (each 30-s sonication and 30-s rest) at the high level to get DNA fragments of ~500 bp on average. After saving a small amount of sonicated sample as the input fraction, the sample was equally divided into three fractions, each incubating with 12CAS, TbrTRF antibody, or immunoglobulin G (IgG) conjugated with Dynabeads Protein G (Thermo Fisher Scientific) for 3 hours at 4°C. In γH2A ChIP, the total lysate was equally divided into two fractions, each incubating with the γH2A antibody or IgG conjugated with Dynabeads Protein G. After washing, immunoprecipitated products were eluted from the beads, and DNA was isolated from the products followed by Southern slot blot hybridization or qPCR analysis.

**Recombinant protein expression and purification**

Recombinant protein expression constructs were transformed into various E. coli strains for optimum expression (table S2). Protein
samples used for EMSA studies were expressed in standard LB media. Proteins used for acquiring \( ^{15} \text{N} \) HSQC NMR spectrum were expressed in M9 minimal media, with \( ^{15} \text{N} \)-labeled ammonium chloride (\( ^{15} \text{N}, 98\% + \)) (Cambridge Isotope Laboratories Inc.) as nitrogen source and \( \delta \)-glucose (Cambridge Isotope Laboratories Inc.) as carbon source. Protein expression was induced by isopropyl-\( \beta \)-d-thiogalactopyranoside. TrxA-His\( _{6} \)-tagged proteins were purified with a His\( _{6} \)bind resin (Millipore) according to the manufacturer’s protocol. GST-tagged proteins were purified with Glutathione Sepharose 4 Fast Flow beads (GE Healthcare) according to the manufacturer’s protocol. For protein samples used for acquiring the \( ^{15} \text{N} \) HSQC NMR spectrum, the fusion tag was removed by 3C protease. Purified proteins were dialyzed in dialysis buffer [20 mM Hepes (pH 7.9), 100 mM KCl, 0.1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 15% glycerol, and 1 mM dithiothreitol (DTT)] at 4°C overnight.

**Electrophoretic mobility shift assay**

Partially purified recombinant proteins were incubated with 0.5 ng of radiolabeled DNA probe [except in Fig. 3 (D and E) and fig. S1 (K, L, and O), where 0.25 ng of the probe was used] in 15 µl of 1× DNA EMSA buffer [20 mM Hepes (pH 7.9), 100 mM KCl, 1 mM MgCl\(_2\), 0.1 mM EDTA, bovine serum albumin (BSA; 100 ng/µl), 5% glycerol, and 1 mM DTT] at room temperature for 30 min. EMSA loading dye (1.5 µl) (50% glycerol, 0.1% bromophenol blue, and 0.1% xylene cyanol) was added to each sample before it was electrophoresed in a 0.6% agarose gel in 0.5× tris-borate EDTA running buffer. Gels were dried and exposed to a phosphorimager.

In EMSA competition assays, unlabeled competitor was added to the reaction with the labeled probe in 1× DNA EMSA buffer first followed by adding recombinant proteins and incubation at room temperature for 30 min. Sequences of all probes used in this study are listed in table S3.

**DNA probe preparation for EMSA**

One hundred fifty nanograms of double-stranded linear DNA was radiolabeled using the Klenow fragment [New England Biolabs (NEB)] and \( ^{32} \text{P} \) alpha 2′-deoxycytidine 5′-triphosphate (dCTP) in a radiolabeled using the Klenow fragment [New England Biolabs DNA probe preparation for are listed in table S3. temperature for 30 min. Sequences of all probes used in this study buffer. Gels were dried and exposed to a phosphorimager. phosphorised in a 0.6% agarose gel in 0.5× tris-borate EDTA running buffer. Gels were dried and exposed to a phosphorimager.

In EMSA competition assays, unlabeled competitor was added to the reaction with the labeled probe in 1× DNA EMSA buffer first followed by adding recombinant proteins and incubation at room temperature for 30 min. Sequences of all probes used in this study are listed in table S3.

**DNA oligo (100 pmol) was radiolabeled using the T4 Polynucleotide Kinase (NEB) and \( ^{32} \text{P} \) gamma adenosine triphosphate in 30 µl of reaction [1× PNK (T4 polynucleotide kinase) buffer and 4.8% poly(ethylene glycol)] at 37°C for 60 min. The radiolabeled probe was purified by a QiaQuick nucleotide removal kit (Qiagen) according to the manufacturer’s protocol. Radiolabeled probes were size purified from 10% urea polyacrylamide gel and eluted in 400 µl of 10 mM tris•Cl/1 mM EDTA (pH 8.0). Labeled oligo was precipitated overnight in 0.2 M sodium acetate (pH 5.5)/ethanol followed by washes with 70% ethanol and was resuspended in 40 µl of ddH\(_2\)O.

**K\(_d\) calculation**

Densitometry data from fig. S1 (K and L) were obtained from ImageQuant (GE Healthcare). Titration curves were generated by plotting protein concentration (nanomolar) versus percentage shift of the radio-labeled probe in Prism GraphPad. K\(_d\) was calculated using the following equation: Y = Bmax*X/(K\(_d\) + X) + NS*X, where Bmax is the maximum specific binding of radiolabeled probe, K\(_d\) is the equilibrium dissociation constant, and NS is the slope of nonspecific binding.

**NMR titration assay**

All \( ^{15} \text{N} \) HSQC NMR experiments were performed at 298 K. Spectra were processed and analyzed using an AVANCE III 700 NMR spectrometer (Bruker). \( ^{15} \text{N} \) HSQC experiments were acquired with samples in 20 mM sodium sulfate (pH 6.5), 150 mM NaCl, 1 mM EDTA, and 1 mM DTT. Concentration of \( ^{15} \text{N} \)-labeled TBRAP1\(_{639-761}\) and TBRAP1\(_{639-761}^{\text{A}}\) is 0.1 mM. Two probes were used in NMR titration (purchased from Integrated DNA Technologies): random dsDNA (5′ ‘TTGTTAGGAGGTGTGAT 3′/3′ ATCACACACTTCCACA 3′) and telomeric dsDNA (5′ ‘TTAGGGTTAGGGTTAGGG 3′/3′ CCCTAACCTAAACCTAA 3′).

**IF analyses**

IF analyses were done the same way as described in (6).

**VSG switching assay**

VSG switching assay was performed, as described in (41). Detailed protocol can be found in the Supplementary Materials.

**RNAseq and data analysis**

The Cre expression was induced by doxycycline in TBRAP1\(_{F/+}\), TBRAP1\(_{F/+}\), TBRAP1\(_{F/-}\), and TBRAP1\(_{A/c}\) cells for 30 hours before total RNA was isolated and purified through RNeasy columns (Qiagen). All RNA samples were run on a Bioanalyzer 2100 (Agilent Technologies) using the Agilent RNA 6000 Nano Kit to verify the RNA quality before they were sent to Novogene for library preparation and RNA high-throughput sequencing followed by bioinformatic analysis. Details can be found in the Supplementary Materials.
Afrin et al., Sci. Adv. 2020; 6 : eabc4065     18 September 2020
SCIENCE ADVANCES | RESEARCH ARTICLE

ACKNOWLEDGMENTS: We thank G. A. M. Cross for the VSG6 and VSG2 rabbit antibodies and K. Gull for the TAT-1 antibody. We thank R. Sandhu for generating the SSA mutant construct. We thank A. Severson and G. V. Boerner and the Li lab members for the comments on the manuscript. Funding: This work is supported by an NIH R01 grant AI066095 (P. D. T. B. L.), an NIH S10 grant 51OD025252 (P. D. T. B. L.), a Research Grants Council GRF grant PolyU 151060/18M (P. D. T. B. L.), and a Research Grants Council AoE grant AoE/M-09/12 (P. D. T. B. L.). The publication cost is partly supported by GRHD at CSU. Author contributions: Conceptualization, B. L. and Y. Z.; methodology, B. L., Y. Z., and A. K. G.; investigation, all authors; writing—original draft, B. L.; writing—review and editing, B. L., Y. Z., M. A., and A. K. G.; funding acquisition, B. L. and Y. Z.; resources, B. L. and Y. Z.; supervision, B. L. and Y. Z. Competing interests: The authors declare that they have no competing interests. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors. Reagents generated in this study will be made available upon request, but we may require a payment and a completed materials transfer agreement if there is potential for commercial application. Data and code availability: The RNAseq data generated during this study are available at GEO repository (GSE147734).

Submitted 22 April 2020
Accepted 31 July 2020
Published 18 September 2020
10.1126/sciadv.abc4065

Citation: Afrin et al., African trypanosome telomeres exhibit a novel dependency of gene expression, with an alternative DNA binding activity required for its telomere localization and VSG silencing. Sci. Adv. 6, eabc4065 (2020).
TbRAP1 has an unusual duplex DNA binding activity required for its telomere localization and VSG silencing
Marjia Afrin, Amit Kumar Gaurav, Xian Yang, Xuehua Pan, Yanxiang Zhao and Bibo Li

Sci Adv 6 (38), eabc4065.
DOI: 10.1126/sciadv.abc4065