MOLECULAR BIOLOGY

*Tb*RAP1 has an unusual duplex DNA binding activity required for its telomere localization and VSG silencing

Marjia Afrin[1]*, Amit Kumar Gaurav[1]*, Xian Yang[2], Xuehua Pan[2], Yanxiang Zhao[2†], Bibo Li[1,3,4†]

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 *Tb*RAP1 is essential for VSG regulation. We identify dsDNA and single-stranded DNA binding activities in *Tb*RAP1, which require positively charged $_{737}$RKRRR$_{741}$ residues that overlap with *Tb*RAP1'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 $_{737}$RKRRR$_{741}$-adjacent S residues and is essential for *Tb*RAP1'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* RAP1, 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 *Sc*RAP1 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 (ESs), 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 *Tb*RAP1 as a *Tb*TRF-interaction factor (*6*), while *Tb*TRF binds the duplex TTAGGG repeats of *T. brucei* telomere (*25*). *Tb*RAP1 associates with the telomere chromatin and is partially colocalized with *Tb*TRF (*6*). We previously showed that depletion of *Tb*RAP1 leads to a marked derepression of all ES-linked *VSG*s (*6*, *19*). The *Tb*RAP1-mediated silencing has a stronger effect on telomere proximal genes than those located further away, suggesting that localization of *Tb*RAP1 at the telomere is essential for this silencing (*6*). Our previous study showed that association of *Tb*RAP1 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 *Tb*RAP1-depleted cells leads to more telomeric/subtelomeric DNA damage (*23*). These findings suggest that the telomere localization of *Tb*RAP1 is also a prerequisite for telomere integrity. However, the mechanism of recruiting *Tb*RAP1 to the telomere was unknown.

*Tb*RAP1 has both a Myb and a MybLike domain (*6*), but their roles in targeting *Tb*RAP1 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

[1]Center for Gene Regulation in Health and Disease, Department of Biological, Geological, and Environmental Sciences, College of Sciences and Health Professions, Cleveland State University, 2121 Euclid Avenue, Cleveland, OH 44115, USA. [2]Department of Applied Biology and Chemical Technology, State Key Laboratory of Chirosciences, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, People's Republic of China. [3]Case Comprehensive Cancer Center, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA. [4]Department of Inflammation and Immunity, Lerner Research Institute, Cleveland Clinic, 9500 Euclid Avenue, Cleveland, OH 44195, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: b.li37@csuohio.edu (B.L.); yanxiang.zhao@polyu.edu.hk (Y.Z.)

localization of *Tb*RAP1 is independent of its Myb domain and *Tb*TRF. Unexpectedly, within the *Tb*RAP1 MybLike domain and overlapping with nuclear localization signal (NLS), we identify $_{737}$RKRRR$_{741}$, a group of positively charged residues, to have both double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA) binding activities. We show that $_{737}$RKRRR$_{741}$ is required for the association of *Tb*RAP1 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. brucei* cells (*26*, *27*). We further demonstrate that the phospho-mimicking mutation of these S residues substantially diminishes *Tb*RAP1's dsDNA binding activity in vitro and abolishes *Tb*RAP1'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 *Tb*RAP1 is essential and can be regulated by posttranslational modification of *Tb*RAP1.

RESULTS

Localization of *Tb*RAP1 to the telomere is independent of *Tb*TRF and the Myb domain

*Tb*RAP1 is an intrinsic component of the *T. brucei* telomere complex (*6*), but how *Tb*RAP1 is located to the telomere was unknown. Because *Tb*TRF binds the duplex telomere DNA and interacts with *Tb*RAP1 (*6*, *25*), we tested whether *Tb*RAP1 is recruited to the telomere by *Tb*TRF. A hemagglutinin (HA) monoclonal antibody 12CA5 was used to chromatin immunoprecipitate (ChIP) the FLAG-HA-HA (F2H)–tagged *Tb*RAP1 in *TbRAP1*$^{F2H+/-}$ *Tb*TRF RNA interference (RNAi) cells (table S1 lists all strains used in this study). Before and after the depletion of *Tb*TRF (Fig. 1A), *Tb*RAP1 associated with the telomere chromatin at nearly the same level (Fig. 1B). As a control, *Tb*TRF was no longer at the telomere after the RNAi induction (Fig. 1B). Therefore, the localization of *Tb*RAP1 at the telomere is *Tb*TRF independent.

*Tb*RAP1 has a putative Myb domain (Fig. 1C) (*6*). While Myb motifs frequently bind DNA, the role of *Tb*RAP1's Myb domain in localizing *Tb*RAP1 to the telomere was unknown. We have established the *TbRAP1*$^{F/\Delta M}$ strain, in which one *TbRAP1* allele was replaced with the ΔMyb (ΔM) mutant, and the other (F allele) was flanked by loxP repeats so that it can be deleted when Cre is expressed (*28*). F2H-*Tb*RAP1ΔM expressed at the same level as F2H-*Tb*RAP1, and Cre induction depleted the wild-type (WT) *Tb*RAP1 protein within 30 hours (*28*). In *TbRAP1*$^{F2H+/-}$ and Cre-expressing *TbRAP1*$^{F/\Delta M}$ cells, ChIP using the HA antibody 12CA5 and a rabbit *Tb*TRF antibody (*25*) showed that F2H-*Tb*RAP1ΔM was located at the telomere the same as F2H-*Tb*RAP1, and *Tb*TRF associated with the telomere chromatin in both cells (Fig. 1D). Therefore, the Myb domain is not necessary for localizing *Tb*RAP1 to the telomere.

*Tb*RAP1 has electrostatics-based DNA binding activities that rely on $_{737}$RKRRR$_{741}$

*Tb*RAP1 also has a MybLike domain [amino acids (aa) 639 to 761] (Fig. 1C) (*6*), which contains a positively charged $_{737}$RKRRR$_{741}$ patch. To test whether this domain has any DNA binding activity, we partially purified the recombinant TrxA-His$_6$–tagged *Tb*RAP1$_{639–761}$, *Tb*RAP1$_{639–733}$, and *Tb*RAP1$_{734–761}$ from *Escherichia coli* (fig. S1A; table S2 lists all recombinant proteins used in this study) and performed electrophoretic mobility shift assay (EMSA). *Tb*RAP1$_{639–761}$ and *Tb*RAP1$_{734–761}$ 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 *Tb*RAP1$_{639–733}$ or TrxA-His$_6$ 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 *Tb*RAP1$_{639–761}$'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 *Tb*RAP1$_{414–855}$ (fig. S1A) and got the same results (fig. S1, F to H). Therefore, *Tb*RAP1 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 *Tb*RAP1's DNA binding activities, we used nuclear magnetic resonance (NMR) to analyze the heteronuclear single-quantum correlation (HSQC) spectrum of ^{15}N-labeled *Tb*RAP1$_{639–761}$ 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 $_{737}$RKRRR$_{741}$ 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 obtained when ds(TTAGGG)$_3$ was added (fig. S1I). These data suggest that the R/K patch is directly responsible for the *Tb*RAP1'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 *Tb*RAP1-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 *Tb*RAP1$_{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}N-labeled 5A mutant (Fig. 2E and fig. S1J). Therefore, $_{737}$RKRRR$_{741}$ is directly responsible for *Tb*RAP1's DNA binding activities. This finding is further corroborated by EMSA results, where *Tb*RAP1$_{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 $_{737}$RKRRR$_{741}$ and the DNA substrates. We further determined K_d (equilibrium dissociation constant) values of *Tb*RAP1$_{639–761}$ binding to a 100-bp dsDNA and a 100-nt ssDNA with a random sequence to be 21.5 μM and 310 nM, respectively (fig. S1, K and L).

*Tb*RAP1$_{639–761}$ 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 *Tb*RAP1's ssDNA binding activity. *Tb*RAP1$_{639–761}$ bound shorter DNA oligos with weaker affinity (40 < 60 < 80 < 100 nt) (Fig. 3C). Therefore, *Tb*RAP1 is different from its vertebrate and fission yeast homologs in that it has DNA binding activities. *Tb*RAP1 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 *Tb*RAP1's NLS (aa 727 to 741; Fig. 1C) (*28*). Hence, *Tb*RAP1 is unique among its homologs in that its NLS has dual roles in nuclear import and DNA binding.

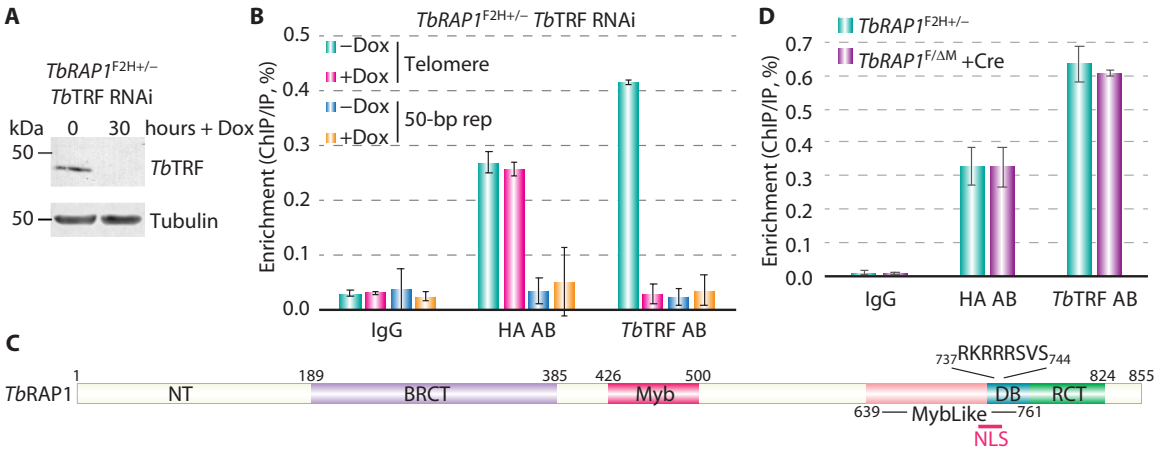

Fig. 1. Telomere localization of *Tb*RAP1 is independent of *Tb*TRF and the Myb domain. (**A**) Protein lysates from *TbRAP1*$^{F2H+/-}$ *Tb*TRF RNAi cells before and after adding doxycycline (Dox) were analyzed by Western blotting. A rabbit anti-*Tb*TRF antibody (*25*) and the tubulin antibody TAT-1 (*42*) were used. In this and many other strains, *Tb*RAP1 has an N-terminal FLAG-HA-HA (F2H) tag. (**B**) ChIP analysis using an HA monoclonal antibody (AB) 12CA5 (Memorial Sloan Kettering Cancer Center monoclonal AB Core) and a rabbit *Tb*TRF antibody in *TbRAP1*$^{F2H+/-}$ *Tb*TRF RNAi before and after depletion of *Tb*TRF. DNA isolated from the ChIP products were hybridized with a telomere probe and a 50-bp repeat probe in Southern blotting. Blots were exposed to a phosphorimager, and results were quantified by ImageQuant (GE Healthcare). Averages from three independent experiments were calculated. In this and other figures, error bars represent SD. (**C**) Domain structure of *Tb*RAP1. The BRCT, Myb, MybLike, and RCT domains and the NLS (aa 727 to 741) were identified previously (*6*, *28*). (**D**) ChIP analysis using the HA antibody 12CA5 and a rabbit *Tb*TRF antibody in *TbRAP1*$^{F2H+/-}$ and the Cre-induced *TbRAP1*$^{F/\Delta M}$ cells (*28*). ChIP products were analyzed the same way as in (B).

Phospho-mimicking mutations of the R/K patch–adjacent S residues significantly affect *Tb*RAP1's dsDNA binding activity

*Tb*RAP1's DNA binding activities rely on the R/K patch and are apparently electrostatics based. Phosphoproteomic analyses showed that S742 and S744 of *Tb*RAP1 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 *Tb*RAP1$_{639–761}$ to mimic nonphosphorylated and phosphorylated states, respectively. EMSA showed that *Tb*RAP1$_{639–761}$2SD lost nearly all the dsDNA binding activity, while *Tb*RAP1$_{639–761}$2SA still retained most of it (Fig. 3D). Therefore, *Tb*RAP1's dsDNA binding activity is likely sensitive to phosphorylation of both S residues. Both 2SA and 2SD mutants still bound the ssDNA, although 2SD has a slightly weaker activity than WT *Tb*RAP1$_{639–761}$ (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 *Tb*RAP1$_{639–761}$S742D and *Tb*RAP1$_{639–761}$S744D (fig. S1N) bound both dsDNA and ssDNA the same as WT *Tb*RAP1 (fig. S1O), suggesting that phosphorylation of both S residues is necessary to exert a detectable effect on dsDNA binding.

In vivo telomere localization of *Tb*RAP1 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 *Tb*RAP1 (*28*). In *TbRAP1*$^{F/+}$, we replaced the WT *TbRAP1* allele with F2H-tagged DB domain mutants to generate *TbRAP1*$^{F/mut}$ strains (fig. S2A). For mutants missing the functional *Tb*RAP1 NLS (aa 727 to 741), we added an N-terminal SV40 NLS, which is sufficient for nuclear import of *Tb*RAP1 (*28*). Genotypes of *TbRAP1*$^{F/mut}$ strains were confirmed by Southern (fig. S2B) and sequencing analyses.

To examine whether F2H-tagged *Tb*RAP1 mutants are associated with the telomere chromatin, we performed ChIP using the HA antibody 12CA5 in *TbRAP1*$^{F/mut}$ cells. *Tb*RAP1 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 *Tb*RAP1 mutants' behavior without the influence from the WT protein. As a control, ChIP was done in *TbRAP1*$^{F2H+/-}$ using the 12CA5 antibody. All *Tb*RAP1 mutants were expressed at the same level as F2H-*Tb*RAP1, except ΔMybLike (ΔML) at a subtly lower level (Fig. 4A). Only a residual amount of *Tb*RAP1ΔML (Fig. 4B), ΔDB (Fig. 4C), and 5A (Fig. 4D) mutants associated with the telomere chromatin, which was much lower than F2H-*Tb*RAP1. *Tb*TRF was still at the telomere in these mutants (Fig. 4, B to D). Immunofluorescence (IF) analyses were done to examine the subnuclear localization of *Tb*RAP1 point mutants. In *TbRAP1*$^{F2H+/-}$ cells, F2H-*Tb*RAP1 partially colocalized with *Tb*TRF (Fig. 4E, top), the same as we reported previously (*6*). However, *Tb*RAP1-5A was no longer colocalized with *Tb*TRF, 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 *Tb*RAP1.

In addition, significantly less *Tb*RAP1-2SD was associated with the telomere chromatin than F2H-*Tb*RAP1 (Fig. 4F). IF also showed that *Tb*RAP1-2SD did not colocalize with *Tb*TRF (Fig. 4G), indicating that the *Tb*RAP1's dsDNA binding activity is critical for its localization at the telomere. In contrast, mutation of *Tb*RAP1's phosphorylation sites to A did not affect its telomere localization. Five *Tb*RAP1 residues were found to be phosphorylated in *T. brucei* cells (*26*, *27*), and the *TbRAP1*$^{F/5SA}$ 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 *TbRAP1*$^{F/2SA}$ with only S742AS744A mutations.

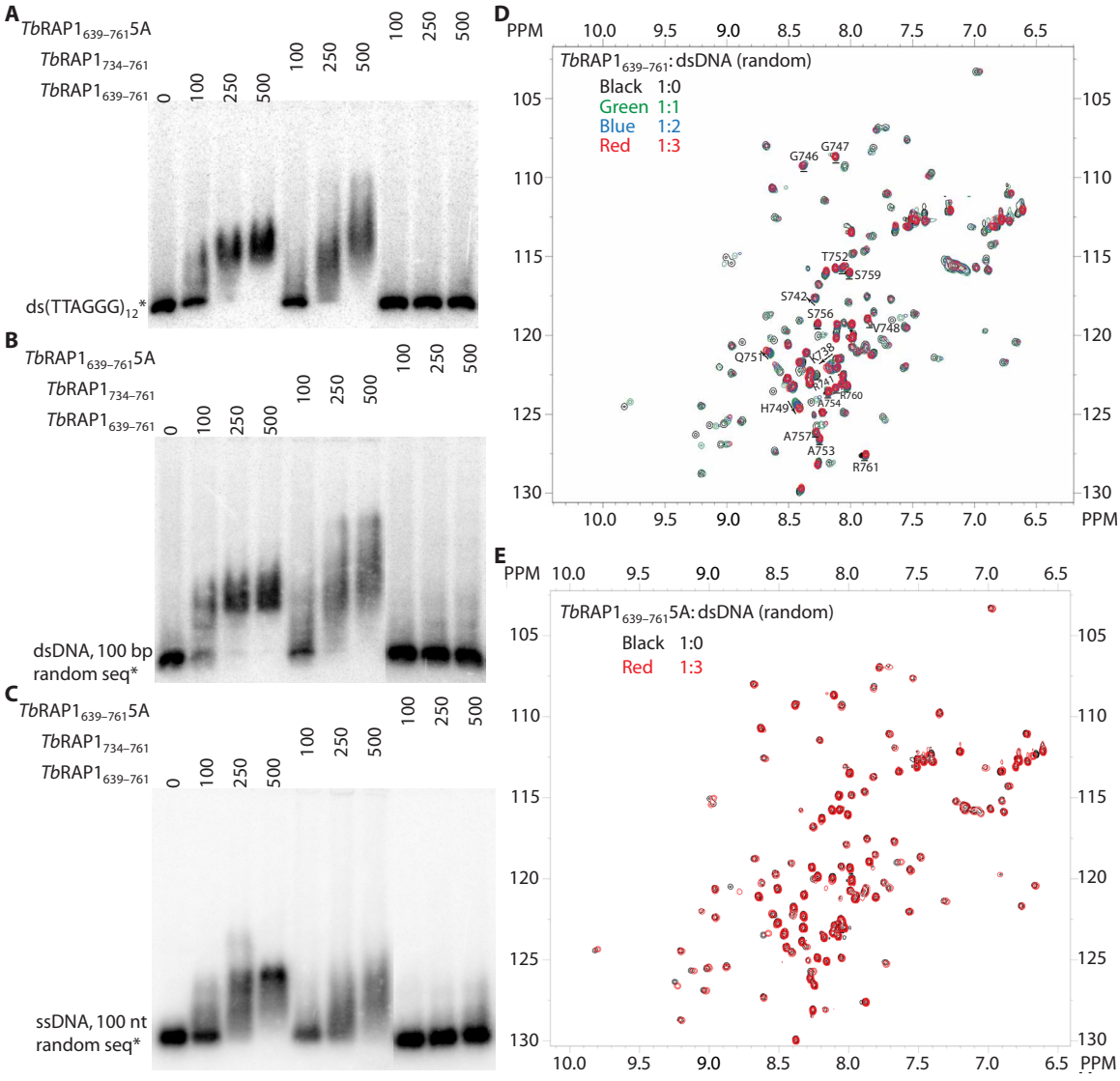

Fig. 2. *Tb*RAP1 has dsDNA and ssDNA binding activities that depend on $_{737}$RKRRR$_{741}$ (the R/K patch). (**A** to **C**) *Tb*RAP1$_{734–761}$ has DNA binding activities (numbers indicate amino acid positions). EMSA experiments were done using TrxA-His$_6$–tagged *Tb*RAP1$_{639–761}$, *Tb*RAP1$_{734–761}$, and *Tb*RAP1$_{639–761}$5A ($_{737}$RKRRR$_{741}$ mutated to $_{737}$AAAAA$_{741}$) and a radiolabeled dsDNA containing (TTAGGG)$_{12}$ (A), a dsDNA with a random sequence (B), or a DNA oligo with a random sequence (C) as the probe. The amount (nanogram) of protein used is indicated on top of each lane. (**D**) Four overlapped HSQC NMR spectra of ^{15}N-labeled *Tb*RAP1$_{639–761}$ in the absence (black) and presence of random dsDNA in 1× (green), 2× (blue), and 3× (red) molar excess. Labeled residues are within the R/K patch. Arrows indicate substantial chemical shifts after adding the DNA substrate. (**E**) Two overlapped HSQC NMR spectra of ^{15}N-labeled *Tb*RAP1$_{639–761}$5A in the absence (black) and presence of random dsDNA in 3× (red) molar excess. No detectable chemical shifts were observed when the DNA substrate was added. PPM, parts per million.

Both *Tb*RAP1-5SA and 2SA mutants associated with the telomere chromatin (Fig. 4H and fig. S2D). IF showed that 5SA and 2SA both partially colocalized with *Tb*TRF (Fig. 4I and fig. S2E). Therefore, *Tb*RAP1 with unphosphorylated S742 and S744 is localized at the telomere, while phosphorylation of both S742 and S744 can remove *Tb*RAP1 from the telomere.

The telomere localization of *Tb*RAP1 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 *Tb*RAP1's DNA binding activities specifically, we first examined cell growth in *TbRAP1*$^{F/mut}$ cells carrying DB mutations after induction of Cre. F2H-tagged *Tb*RAP1 mutants were detected by the HA Probe antibody. The expression of total *Tb*RAP1 was also examined by a rabbit antibody (*6*) that recognizes the MybLike domain (*28*) and, specifically, *Tb*RAP1$_{639–733}$ but not *Tb*RAP1$_{734–761}$ (fig. S2F).

Western analysis showed the deletion of WT *Tb*RAP1 and a persistent expression of F2H-NLS–tagged *Tb*RAP1-ΔDB (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*$^{F/\Delta DB}$ (fig. S2H) and *TbRAP1*$^{F/5A}$ (Fig. 5B) cells exhibited a growth arrest after inducing Cre for 24 to 30 hours, indicating that *Tb*RAP1'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

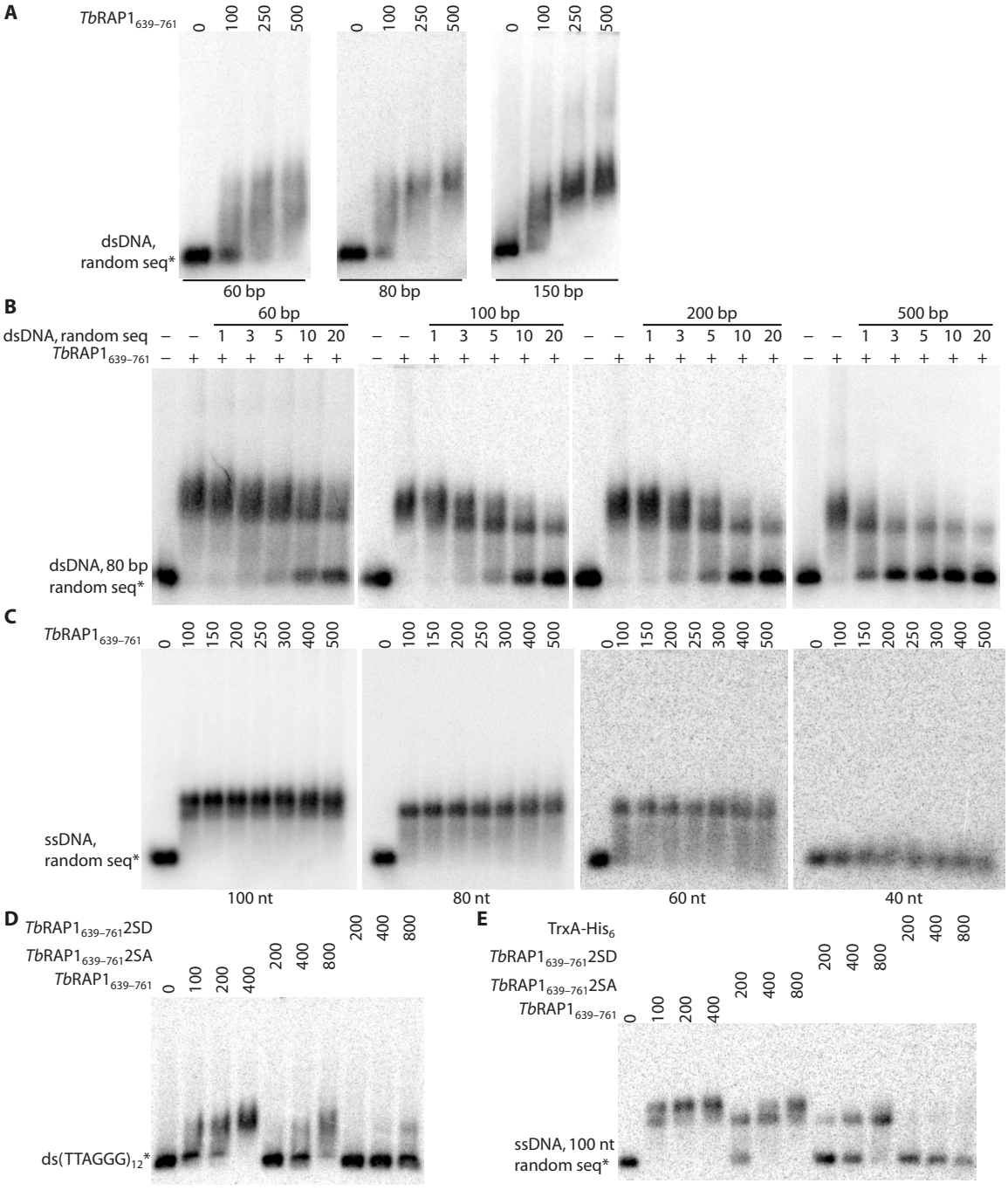

Fig. 3. *Tb*RAP1 has higher affinity for longer DNA substrates. (**A**) EMSA using TrxA-His$_6$-*Tb*RAP1$_{639–761}$ and radiolabeled dsDNAs (of 60, 80, or 150 bp) containing a random sequence as the substrate. (**B**) EMSA using 250 ng of TrxA-His$_6$-*Tb*RAP1$_{639–761}$ and a radiolabeled 80-bp dsDNA with a random sequence as the substrate. Non-radiolabeled dsDNAs with a random sequence of 60, 100, 200, or 500 bp were used as competitors. The amounts of competitors are indicated as molar folds of the probe. (**C**) EMSA experiments were done using TrxA-His$_6$-*Tb*RAP1$_{639–761}$ and a radiolabeled 100-, 80-, 60-, or 40-nt DNA oligo containing a random sequence as the substrate. (**D** and **E**) *Tb*RAP1 S742 and S744 residues are important for the dsDNA binding activity. EMSA using TrxA-His$_6$–tagged *Tb*RAP1$_{639–761}$, *Tb*RAP1$_{639–761}$2SA (S742AS744A), and *Tb*RAP1$_{639–761}$2SD (S742DS744D) and a radiolabeled dsDNA containing (TTAGGG)$_{12}$ (D) or a 100-nt ssDNA with a random sequence (E) as the substrate. TrxA-His$_6$ was used as a negative control in (E).

the dsDNA binding without affecting ssDNA binding significantly (Fig. 3, D and E), this observation indicates that *Tb*RAP1's dsDNA binding is essential for normal cell proliferation. In contrast, *TbRAP1*$^{F/5SA}$ (fig. S2J) and *TbRAP1*$^{F/2SA}$ (Fig. 5F) cells only grew mildly slower after the WT *TbRAP1* allele was deleted. Therefore, all *Tb*RAP1 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 *Tb*RAP1 mutants capable

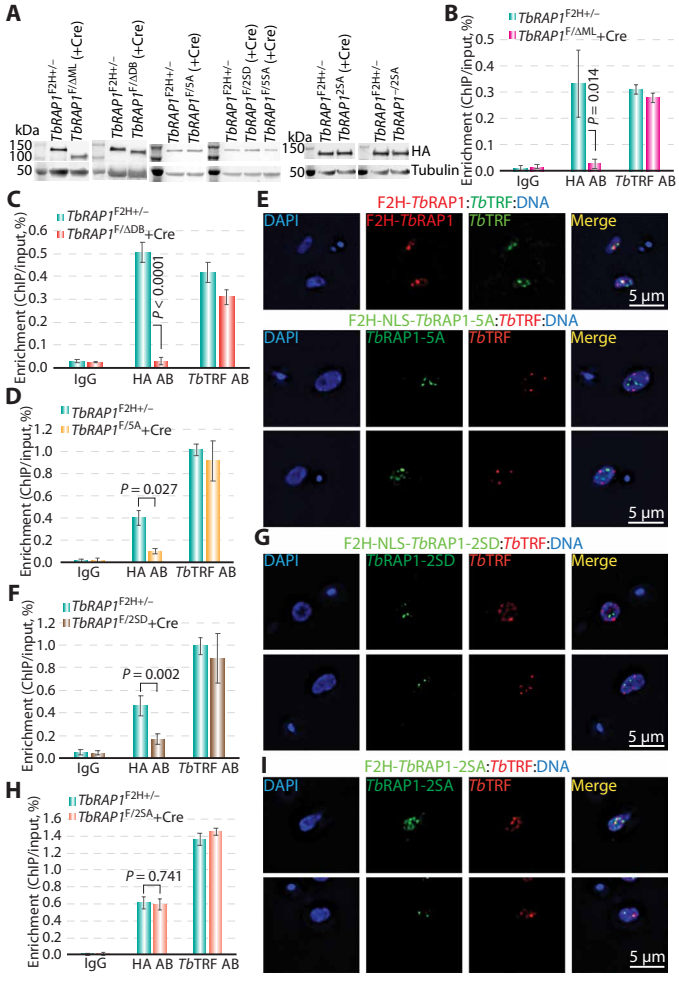

Fig. 4. The R/K patch is important for the telomere localization of *Tb*RAP1. (**A**) Western analyses to compare expression of F2H-tagged *Tb*RAP1 proteins. Whole-cell lysates were prepared from the indicated strains. The HA antibody 12CA5 and tubulin antibody TAT-1 were used. ChIP experiments using the HA antibody 12CA5 and a *Tb*TRF rabbit antibody were done in *TbRAP1*$^{F2H+/-}$ cells and Cre-induced (for 30 hours) *TbRAP1*$^{F/\Delta ML}$ (**B**), *TbRAP1*$^{F/\Delta DB}$ (**C**), *TbRAP1*$^{F/5A}$ (**D**), *TbRAP1*$^{F/2SD}$ (**F**), and *TbRAP1*$^{F/2SA}$ (**H**) cells. IF analyses were done in *TbRAP1*$^{F2H+/-}$ (**E**, top), *TbRAP1*$^{F/5A}$ (E, bottom), *TbRAP1*$^{F/2SD}$ (**G**), and *TbRAP1*$^{F/2SA}$ (**I**) cells. 12CA5 and a *Tb*TRF chicken antibody (*6*) were used. DNA was stained by 4′,6-diamidino-2-phenylindole (DAPI). In this and other IF images, all images are of the same scale, and size bars are shown in one of the images in each panel.

of cell proliferation, providing useful genetic tools for further investigation of *Tb*RAP1's functions in the future.

The R/K patch–mediated dsDNA binding activity of *Tb*RAP1 is essential for VSG silencing

We previously hypothesized that the telomere localization of *Tb*RAP1 is a prerequisite of normal VSG silencing (*6*). To test this, we examined mRNA levels of several ES-linked *VSG*s in domain deletion mutants and transcriptomic profiles in point mutations that lost *Tb*RAP1's dsDNA binding activity. By quantitative reverse transcription PCR (qRT-PCR), we found that multiple ES-linked silent *VSG*s were derepressed several hundred-folds to thousand-folds in *TbRAP1*$^{F/\Delta ML}$ (fig. S3A) and *TbRAP1*$^{F/\Delta DB}$ (fig. S3B) cells after induction of Cre for 30 hours, indicating that the DB region is essential for VSG silencing.

RAP1 homologs are known to regulate expression of both subtelomeric and nontelomeric genes (*3*, *29*, *30*). We suspect that *Tb*RAP1 may affect expression of genes other than *VSG*s. To examine *Tb*RAP1's function in global gene expression, we performed RNA sequencing (RNAseq) analysis in *Tb*RAP1 point mutants that did not bind DNA. As a control, we first compared the gene expression profile in *TbRAP1*$^{F/-}$ and *TbRAP1*$^{F/+}$ cells after both were induced for Cre expression for 30 hours. Compared to *TbRAP1*$^{F/+}$ cells, >7200 genes were up-regulated in *TbRAP1*$^{F/-}$ 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*$^{F/-}$ cells. All BF ES–linked *VSG*s were derepressed (fig. S5). Some ES-linked *ESAG*s were up-regulated, some were not affected significantly, and others were down-regulated (fig. S5), indicating that *VSG*s and *ESAG*s can be regulated differently by *Tb*RAP1 even when they are located in the same ESs. In addition, >2500 genes were down-regulated upon *Tb*RAP1 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 *Tb*RAP1 is mainly a repressor. Compared to *TbRAP1*$^{F/+}$ cells, in Cre-induced *TbRAP1*$^{F/5A}$ cells, nearly 5300 genes were up-regulated (Fig. 5G), including 2119 *VSG* genes (fig. S4C), while >1400 genes were down-regulated (Fig. 5G), including 66 ribosomal protein genes (fig. S4C). All ES-linked *VSG*s were derepressed although at various levels (fig. S6). In addition, ES-linked *ESAG*s were up-regulated, unaffected, or down-regulated by the 5A mutation (fig. S6). Therefore, *TbRAP1*$^{F/5A}$ has a similar transcriptome profile as *TbRAP1*$^{F/-}$, indicating that the R/K patch is essential for *Tb*RAP1's role in gene expression regulation. We observed a similar phenotype in *TbRAP1*$^{F/2SD}$ 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 *VSG*s were up-regulated although at various levels (fig. S7), and *ESAG*s were similarly affected as in *TbRAP1*$^{F/-}$ and *TbRAP1*$^{F/5A}$ cells (fig. S7). Therefore, *Tb*RAP1's dsDNA binding activity is required for its role in gene expression regulation.

In contrast, 5SA and 2SA mutants exhibited only mild VSG derepression. In *TbRAP1*$^{F/5SA}$ and *TbRAP1*$^{F/2SA}$ cells, a number of silent *VSG*s 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 5SA 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 *VSG*s 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

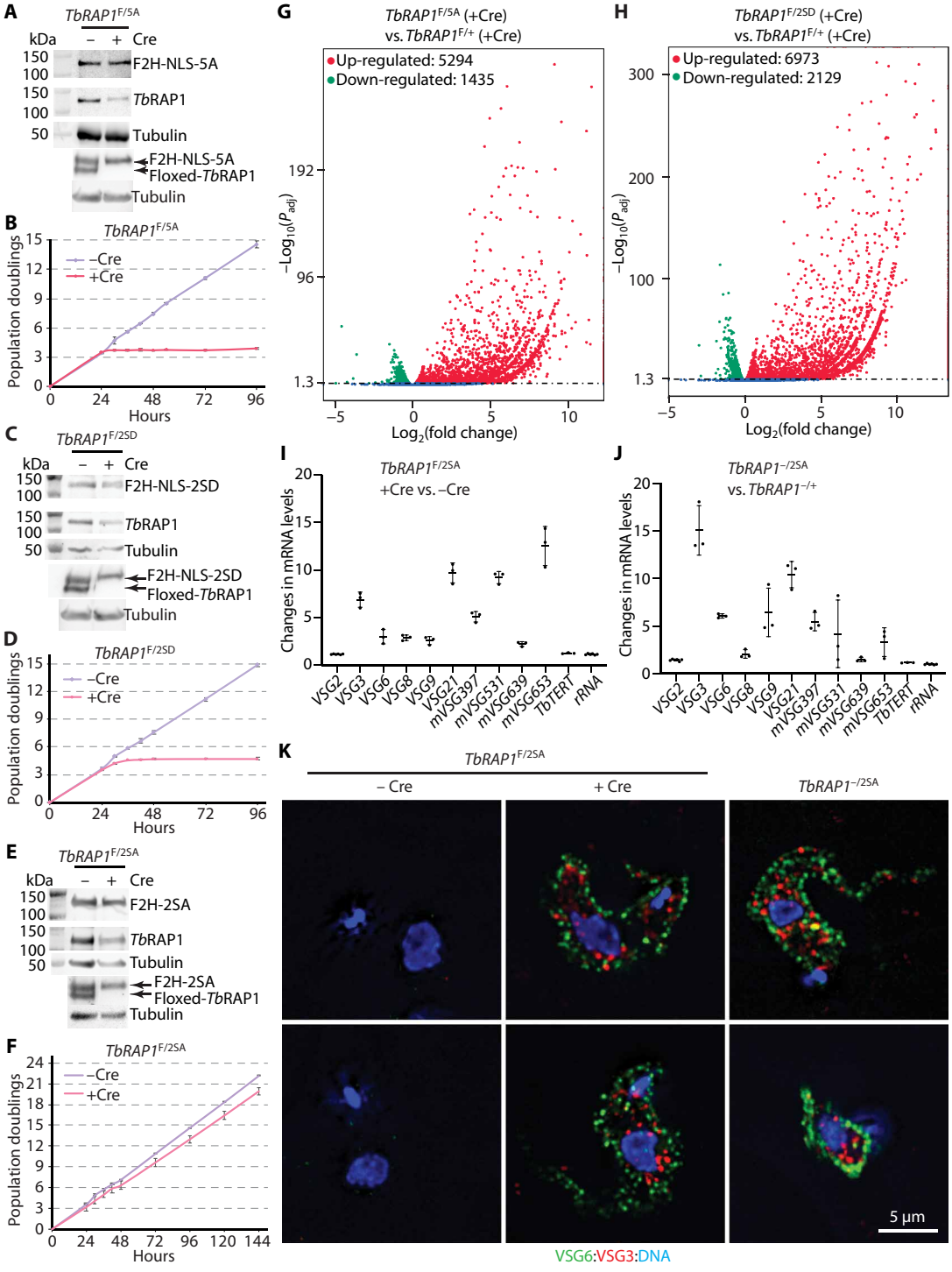

Fig. 5. *Tb*RAP1's dsDNA binding activity is required for VSG silencing and normal cell growth. (**A**, **C**, and **E**) Western analysis of protein extracted from *TbRAP1*$^{F/5A}$ (A), *TbRAP1*$^{F/2SD}$ (C), and *TbRAP1*$^{F/2SA}$ (E) cells before and after induction of Cre for 30 hours using the HA monoclonal antibody HA Probe (Santa Cruz Biotechnologies), a *Tb*RAP1 rabbit antibody (*6*), and the tubulin antibody TAT-1 (top three rows). To separate *Tb*RAP1 mutants and the endogenous *Tb*RAP1, proteins were run on a 7.5% tris polyacrylamide gel for 7 hours and detected by the *Tb*RAP1 rabbit antibody (the fourth row). The associated loading control (tubulin blot) is shown at the bottom. (**B**, **D**, and **F**) Growth curves of *TbRAP1*$^{F/5A}$ (B), *TbRAP1*$^{F/2SD}$ (D), and *TbRAP1*$^{F/2SA}$ (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 *TbRAP1*$^{F/5A}$ (G) and *TbRAP1*$^{F/2SD}$ (H) cells compared to *TbRAP1*$^{F/+}$ cells 30 hours after Cre induction. (**I** and **J**) qRT-PCR of mRNA levels of the active *VSG2*, several silent ES-linked *VSG*s, and chromosome internal *Tb*TERT and ribosomal RNA (rRNA) in *TbRAP1*$^{F/2SA}$ cells after 30 hours of Cre induction (I) and in *TbRAP1*$^{-/2SA}$ 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 *TbRAP1*$^{F/2SA}$ cells before and after Cre induction (left) and in *TbRAP1*$^{-/2SA}$ 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 *Tb*RAP1 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 *Tb*RAP1 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 *Tb*RAP1 suppresses VSG switching by maintaining telomere/subtelomere integrity (*23*). Therefore, we examined DNA damage in all *Tb*RAP1 mutants, using γH2A as an indicator (*33*). Western blotting showed that the γH2A level increased upon Cre induction in *TbRAP1*$^{F/\Delta ML}$ (Fig. 6A), *TbRAP1*$^{F/\Delta DB}$ (Fig. 6B), *TbRAP1*$^{F/5A}$ (Fig. 6B), and *TbRAP1*$^{F/2SD}$ (Fig. 6C) cells. We also performed γH2A IF in *TbRAP1*$^{F/5A}$ and *TbRAP1*$^{F/2SD}$ 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 *Tb*RAP1 mutant cells, suggesting that the increase in DNA damage is not throughout the whole genome. In addition, we detected *Tb*TRF in IF as a marker for the telomere and found that γH2A is partially colocalized with *Tb*TRF (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 ChIP showed that more γH2A associated with the telomere chromatin but not with the tubulin chromatin in *TbRAP1*$^{F/5A}$ and *TbRAP1*$^{F/2SD}$ than in *TbRAP1*$^{F/+}$ cells after Cre induction (Fig. 6F). γH2A ChIP followed by qPCR in *TbRAP1*$^{F/5A}$ (Fig. 6G) and *TbRAP1*$^{F/2SD}$ (Fig. 6H) cells also showed significantly more γH2A at the active ES, including *VSG2* and 70-bp repeats loci upon Cre induction. Therefore, there is an increased amount of DNA damage at the telomere and in the active ES when *Tb*RAP1 is no longer located at the telomere. In contrast, 5SA and 2SA did not increase the γH2A level (Fig. 6I and fig. S3F) and were still localized at the telomere (Fig. 4, H and I, and fig. S2, D and E), further supporting this conclusion.

In IF analysis, few Cre-expressing *TbRAP1*$^{F/5SA}$ cells (fig. S3E) and *TbRAP1*$^{-/2SA}$ cells (Fig. 5K) expressed a relatively high level of VSG6 on the cell surface, suggesting that these cells may have gone through a VSG switching and that 5SA and 2SA mutants may affect VSG switching frequency. We estimated the VSG switching rate in *TbRAP1*$^{-/5SA}$, *TbRAP1*$^{-/2SA}$, and *TbRAP1*$^{-/+}$ cells—all of which initially expressed VSG2. VSG switchers were enriched by passing cells through a magnetic-activated cell sorting column coupled with a VSG2 monoclonal antibody and verified by Western slot blot using a VSG2 rabbit antibody (without the cross-reaction portion). Compared to *TbRAP1*$^{-/+}$ cells, *TbRAP1*$^{-/2SA}$ and *TbRAP1*$^{-/5SA}$ cells had a twofold and four- to fivefold higher VSG switching rate, respectively (Fig. 6J). Therefore, 5SA and 2SA are also mildly defective in suppression of VSG switching.

DISCUSSION

In this study, we identified both dsDNA and ssDNA binding activities of *Tb*RAP1 in a positively charged R/K patch that overlaps with *Tb*RAP1's NLS in the MybLike domain. We showed that the dsDNA binding activity is essential for *Tb*RAP1'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 *Tb*RAP1.

An unusual electrostatics-based DNA binding activity required for targeting *Tb*RAP1 to the telomere

Targeting *Tb*RAP1 to the telomere chromatin requires a unique mechanism among known RAP1s. First, although *Tb*RAP1 interacts with the duplex telomere DNA binding factor *Tb*TRF (*6*), its localization at the telomere is *Tb*TRF independent, which is different from the scenarios in vertebrates and fission yeast (*10*, *34*). Second, *Tb*RAP1'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, *Tb*RAP1'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 *Tb*TRF—use a Myb-type helix-turn-helix motif to recognize the telomere DNA (*25*, *35*, *36*). Furthermore, the *Tb*RAP1's DNA binding activities are electrostatics based and sequence nonspecific. In contrast, *Sc*RAP1 recognizes a substrate with the consensus sequence 5′ ACACCCAYACAYY 3′ (*9*).

Mutation of $_{737}$RKRRR$_{741}$ to AAAAA disrupts the association of *Tb*RAP1 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 *Tb*RAP1. 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 *Tb*RAP1 at the telomere. It is interesting that the *Tb*RAP1's DNA binding activities are sequence nonspecific, yet ChIP and IF analyses indicate that *Tb*RAP1 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 *Tb*RAP1 to discriminate different loci. *Sc*RAP1 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 *Sc*RAP1 to different chromatin loci (*7*). A similar scenario may apply to *Tb*RAP1, although further investigation is necessary to illustrate the underlying mechanisms.

The dsDNA binding activity of *Tb*RAP1 can be regulated by posttranslational modification

We showed that *Tb*RAP1'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 *Tb*RAP1 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 *Tb*RAP1'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 *Tb*RAP1 silences all

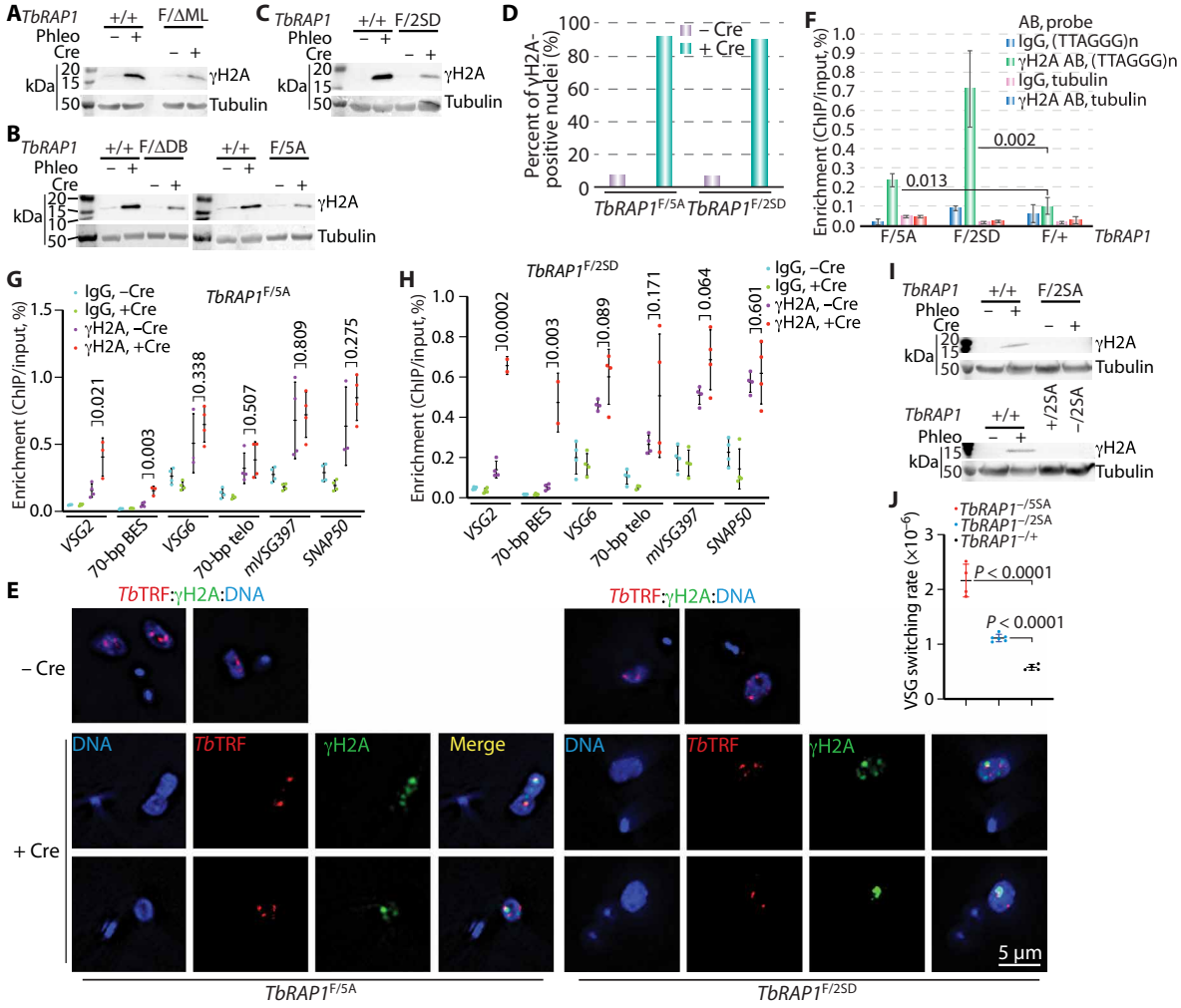

Fig. 6. Nontelomere-localized *Tb*RAP1 mutants have an increased amount of DNA damage at the telomere/subtelomere. (**A** to **C** and **I**) Western analysis to examine the γH2A protein level in WT cells before and after phleomycin treatment (as a positive control) and in *TbRAP1*$^{F/\Delta ML}$ (A), *TbRAP1*$^{F/\Delta DB}$ (B), *TbRAP1*$^{F/5A}$ (B), *TbRAP1*$^{F/2SD}$ (C), and *TbRAP1*$^{F/2SA}$ (I) cells before and after Cre induction and in *TbRAP1*$^{-/2SA}$ and uninduced *TbRAP1*$^{F/2SA}$ cells (I). A γH2A rabbit antibody (*23*) and the tubulin antibody TAT-1 were used. (**D**) Percent of γH2A-positive nuclei in *TbRAP1*$^{F/5A}$ and *TbRAP1*$^{F/2SD}$ cells before and after Cre induction. (**E**) IF analysis using a *Tb*TRF rabbit antibody (*25*) and a γH2A rabbit antibody (*23*) in *TbRAP1*$^{F/5A}$ and *TbRAP1*$^{F/2SD}$ cells before and after Cre induction. DNA was stained by DAPI. Merged images with signals from three channels are shown for –Cre cells (top). Images showing signals from individual and merged channels are shown in +Cre cells (bottom). (**F**) ChIP using the γH2A rabbit antibody and immunoglobulin G (IgG) in *TbRAP1*$^{F/5A}$, *TbRAP1*$^{F/2SD}$, and *TbRAP1*$^{F/+}$ cells after 30 hours of Cre induction followed by Southern blotting using a telomere and a tubulin probe. Blots were exposed to a phosphorimager. Images were quantified using ImageQuant, and average values were calculated from three independent experiments. *P* values of unpaired *t* tests (mutant versus control cells) are shown. (**G** and **H**) ChIP using a γH2A rabbit antibody and IgG in *TbRAP1*$^{F/5A}$ (G) and *TbRAP1*$^{F/2SD}$ (H) cells followed by qPCR using primers specific to the indicated active and silent ES loci. *SNAP50* is a chromosome internal gene. Average enrichment (ChIP/input) was calculated from three to six independent experiments. *P* values of unpaired *t* tests (γH2A ChIP products, +Cre versus –Cre) are shown. (**J**) VSG switching rates in *TbRAP1*$^{-/5SA}$, *TbRAP1*$^{-/2SA}$, and *TbRAP1*$^{-/+}$ cells.

subtelomeric *VSG*s except the one in the active ES. With a means to regulate the dsDNA binding activity, it is possible that *Tb*RAP1 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 *Tb*RAP1'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 *Tb*RAP1. 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 *Tb*RAP1

It is interesting that the R/K patch overlaps with the *Tb*RAP1 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 *Tb*RAP1. Rather, the dsDNA binding activity is required. The fact that DNA binding activities of *Tb*RAP1 rely on the same peptide that signals for nuclear import suggests that deposition of *Tb*RAP1 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 *Tb*RAP1 NLS may be disrupted only after *Tb*RAP1 binds DNA. However, more detailed structural and functional analyses are necessary to test this hypothesis.

*Tb*RAP1's functions in VSG silencing and normal cell proliferation are independent

*Tb*RAP1 is essential for VSG silencing and normal cell proliferation (*6*, *19*). However, whether the two functions are linked or independent was unknown. Most *Tb*RAP1 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 maitenance protein complex-binding protein (MCM-BP), the VEX complex, etc. (*38*).

5SA and 2SA from this study are the only two *Tb*RAP1 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, *Tb*RAP1 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 *Tb*RAP1 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. *Sc*RAP1's DNA binding domain is also essential for cell viability (*11*). It is well known that *Sc*RAP1 is required for transcribing a number of genes encoding ribosomal proteins, and the DNA binding domain is critical for *Sc*RAP1's transcription activation function (*3*). Therefore, defective ribosomal protein gene expression may be a major reason why the *Sc*RAP1 null mutant is lethal. In *Tb*RAP1 conditional knockout cells and mutants defective in telomere localization, 1400 to 2600 genes were down-regulated for their expression, although the change in their mRNA levels is mild. Among the affected are a number of ribosomal protein genes. The down-regulation of ribosomal protein genes would be deleterious to cell growth and likely another reason why *Tb*RAP1 is essential for cell proliferation. However, further studies are necessary to verify the transcription activation function of *Tb*RAP1.

In this study, we demonstrate that *Tb*RAP1's dsDNA binding activity is essential for localization of *Tb*RAP1 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 *Tb*RAP1's NLS, which is unique among all known telomere binding factors. Our findings provide a molecular basis that the DNA binding domain of *Tb*RAP1 may serve as a good target for antiparasite agents, as targeting this site can inactivate essential *Tb*RAP1 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 VSG2 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.

TbRAP1$^{F/+}$ was established previously and described in (*28*). All *TbRAP1*$^{F/mut}$ strains were established using the same strategy. N-terminal F2H-tagged and SV40 NLS-tagged *Tb*RAP1ΔML (deletion of the MybLike domain), *Tb*RAP1ΔDB, *Tb*RAP1-5A, *Tb*RAP1-2SD, and F2H-tagged *Tb*RAP1-5SA and *Tb*RAP1-2SA 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 II before transfecting the *TbRAP1*$^{F/+}$ cells to generate respective *TbRAP1*$^{F/mut}$ 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 *Tb*RAP1's association with the telomere chromatin in the presence and absence of *Tb*TRF, a *Tb*TRF RNAi strain was first established by transfecting the Not I–digested pZJMβ-*Tb*TRF-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-*Tb*RAP1-tar2 construct into the *Tb*TRF RNAi cells. The other *Tb*RAP1 allele was replaced by *Hygromycin resistance* (*HYG*) to establish the *TbRAP1*$^{F2H+/-}$ *Tb*TRF 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 12CA5, *Tb*TRF 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}N HSQC NMR spectrum were expressed in M9 minimal media, with ^{15}N-labeled ammonium chloride (^{15}N, 98%+) (Cambridge Isotope Laboratories Inc.) as nitrogen source and D-glucose (Cambridge Isotope Laboratories Inc.) as carbon source. Protein expression was induced by isopropyl-β-D-thiogalactopyranoside. TrxA-His$_6$–tagged proteins were purified with a His•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}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}P alpha 2′-deoxycytidine 5′-triphosphate (dCTP) in a 50 μl of reaction [50 mM tris (pH 6.8), 10 mM magnesium acetate, 0.1 mM DTT, BSA (0.05 mg/ml), and 0.6 mM deoxynucleotide triphosphates without dCTP] at room temperature for 60 min. The radiolabeled probe was purified by 3 ml of Sephadex G-50 column and precipitated overnight in 0.2 M sodium acetate (pH 5.5)/ethanol followed by washes with 70% ethanol and resuspension in 50 μl of ddH_2O.

DNA oligo (100 pmol) was radiolabeled using the T4 Polynucleotide Kinase (NEB) and ^{32}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_2O.

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 radiolabeled probe in Prism GraphPad. K_d was calculated using the following equation: $Y = B\text{max}*X/(K_d + X) + \text{NS}*X$, 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}N HSQC NMR experiments were performed at 298 K. Spectra were processed and analyzed using an AVANCE III 700 NMR spectrometer (Bruker). ^{15}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}N-labeled *Tb*RAP1$_{639–761}$ and *Tb*RAP1$_{639–761}$5A is 0.1 mM. Two probes were used in NMR titration (purchased from Integrated DNA Technologies): random dsDNA (5′ TGTTGAGGAGGTGGTGAT 3′/5′ATCACCACCTCCTCAACA 3′) and telomeric dsDNA (5′ TTAGGGTTAGGGTTAGGG 3′/5′ CCCTAACCCTAACCCTAA 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*$^{F/5A}$, and *TbRAP1*$^{F/2SD}$ 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.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/38/eabc4065/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. D. E. Gottschling, O. M. Aparicio, B. L. Billington, V. A. Zakian, Position effect at *S. cerevisiae* telomeres: Reversible repression of pol II transcription. *Cell* **63**, 751–762 (1990).
2. B. Li, S. Oestreich, T. de Lange, Identification of human Rap1: Implications for telomere evolution. *Cell* **101**, 471–483 (2000).
3. D. Shore, K. Nasmyth, Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements. *Cell* **51**, 721–732 (1987).
4. M. J. Park, Y. K. Jang, E. S. Choi, H. S. Kim, S. D. Park, Fission yeast Rap1 homolog is a telomere-specific silencing factor and interacts with Taz1p. *Mol. Cells* **13**, 327–333 (2002).
5. E. Y. Yu, W.-F. Yen, O. Steinberg-Neifach, N. F. Lue, Rap1 in *Candida albicans*: An unusual structural organization and a critical function in suppressing telomere recombination. *Mol. Cell. Biol.* **30**, 1254–1268 (2010).
6. X. Yang, L. M. Figueiredo, A. Espinal, E. Okubo, B. Li, RAP1 is essential for silencing telomeric variant surface glycoprotein genes in *Trypanosoma brucei*. *Cell* **137**, 99–109 (2009).
7. B. Piña, J. Fernández-Larrea, N. García-Reyero, F.-Z. Idrissi, The different (sur)faces of Rap1p. *Mol. Genet. Genomics* **268**, 791–798 (2003).
8. J. S. Lipsick, One billion years of Myb. *Oncogene* **13**, 223–235 (1996).
9. I. R. Graham, A. Chambers, Use of a selection technique to identify the diversity of binding sites for the yeast RAP1 transcription factor. *Nucleic Acids Res.* **22**, 124–130 (1994).
10. P. Konig, R. Giraldo, L. Chapman, D. Rhodes, The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA. *Cell* **85**, 125–136 (1996).

11. I. R. Graham, R. A. Haw, K. G. Spink, K. A. Halden, A. Chambers, In vivo analysis of functional regions within yeast Rap1p. *Mol. Cell. Biol.* **19**, 7481–7490 (1999).
12. J. D. Barry, R. McCulloch, Antigenic variation in trypanosomes: Enhanced phenotypic variation in a eukaryotic parasite. *Adv. Parasitol.* **49**, 1–70 (2001).
13. G. A. M. Cross, H. S. Kim, B. Wickstead, Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427. *Mol. Biochem. Parasitol.* **195**, 59–73 (2014).
14. L. S. M. Müller, R. O. Cosentino, K. U. Förstner, J. Guizetti, C. Wedel, N. Kaplan, C. J. Janzen, P. Arampatzi, J. Vogel, S. Steinbiss, T. D. Otto, A.-E. Saliba, R. P. Sebra, T. N. Siegel, Genome organization and DNA accessibility control antigenic variation in trypanosomes. *Nature* **563**, 121–125 (2018).
15. G. A. M. Cross, Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei*. *Parasitology* **71**, 393–417 (1975).
16. I. Cestari, K. Stuart, Transcriptional regulation of telomeric expression sites and antigenic variation in trypanosomes. *Curr. Genomics* **19**, 119–132 (2018).
17. A. Günzl, J. K. Kirkham, T. N. Nguyen, N. Badjatia, S. H. Park, Mono-allelic VSG expression by RNA polymerase I in *Trypanosoma brucei*: Expression site control from both ends? *Gene* **556**, 68–73 (2015).
18. J. Faria, L. Glover, S. Hutchinson, C. Boehm, M. C. Field, D. Horn, Monoallelic expression and epigenetic inheritance sustained by a *Trypanosoma brucei* variant surface glycoprotein exclusion complex. *Nat. Commun.* **10**, 3023 (2019).
19. U. M. Pandya, R. Sandhu, B. Li, Silencing subtelomeric VSGs by *Trypanosoma brucei* RAP1 at the insect stage involves chromatin structure changes. *Nucleic Acids Res.* **41**, 7673–7682 (2013).
20. P. Myler, R. G. Nelson, N. Agabian, K. Stuart, Two mechanisms of expression of a variant antigen gene of *Trypanosoma brucei*. *Nature* **309**, 282–284 (1984).
21. S. E. Jehi, X. Li, R. Sandhu, F. Ye, I. Benmerzouga, M. Zhang, Y. Zhao, B. Li, Suppression of subtelomeric VSG switching by *Trypanosoma brucei* TRF requires its TTAGGG repeat-binding activity. *Nucleic Acids Res.* **42**, 12899–12911 (2014).
22. S. E. Jehi, F. Wu, B. Li, *Trypanosoma brucei* TIF2 suppresses VSG switching by maintaining subtelomere integrity. *Cell Res.* **24**, 870–885 (2014).
23. V. Nanavaty, R. Sandhu, S. E. Jehi, U. M. Pandya, B. Li, *Trypanosoma brucei* RAP1 maintains telomere and subtelomere integrity by suppressing TERRA and telomeric RNA:DNA hybrids. *Nucleic Acids Res.* **45**, 5785–5796 (2017).
24. R. McCulloch, L. J. Morrison, J. P. J. Hall, DNA recombination strategies during antigenic variation in the African trypanosome. *Microbiol. Spectr.* **3**, MDNA3–0016–2014, (2015).
25. B. Li, A. Espinal, G. A. M. Cross, Trypanosome telomeres are protected by a homologue of mammalian TRF2. *Mol. Cell. Biol.* **25**, 5011–5021 (2005).
26. I. R. E. Nett, D. M. A. Martin, D. Miranda-Saavedra, D. Lamont, J. D. Barber, A. Mehlert, M. A. J. Ferguson, The phosphoproteome of bloodstream form *Trypanosoma brucei*, causative agent of African sleeping sickness. *Mol. Cell. Proteomics* **8**, 1527–1538 (2009).
27. M. D. Urbaniak, D. M. A. Martin, M. A. J. Ferguson, Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of *Trypanosoma brucei*. *J. Proteome Res.* **12**, 2233–2244 (2013).
28. M. Afrin, H. Kishmiri, R. Sandhu, M. A. G. Rabbani, B. Li, *Trypanosoma brucei* RAP1 has essential functional domains that are required for different protein interactions. *mSphere* **5**, e00027–e00020 (2020).
29. P. Martinez, M. Thanasoula, A. R. Carlos, G. Gómez-López, A. M. Tejera, S. Schoeftner, O. Dominguez, D. G. Pisano, M. Tarsounas, M. A. Blasco, Mammalian Rap1 controls telomere function and gene expression through binding to telomeric and extratelomeric sites. *Nat. Cell Biol.* **12**, 768–780 (2010).
30. H. Teo, S. Ghosh, H. Luesch, A. Ghosh, E. T. Wong, N. Malik, A. Orth, P. de Jesus, A. S. Perry, J. D. Oliver, N. L. Tran, L. J. Speiser, M. Wong, E. Saez, P. Schultz, S. K. Chanda, I. M. Verma, V. Tergaonkar, Telomere-independent Rap1 is an IKK adaptor and regulates NF-κB-dependent gene expression. *Nat. Cell Biol.* **12**, 758–767 (2010).
31. C. E. Boothroyd, O. Dreesen, T. Leonova, K. I. Ly, L. M. Figueiredo, G. A. M. Cross, F. N. Papavasiliou, A yeast-endonuclease-generated DNA break induces antigenic switching in *Trypanosoma brucei*. *Nature* **459**, 278–281 (2009).
32. L. Glover, S. Alsford, D. Horn, DNA break site at fragile subtelomeres determines probability and mechanism of antigenic variation in African trypanosomes. *PLOS Pathog.* **9**, e1003260 (2013).
33. L. Glover, D. Horn, Trypanosomal histone γH2A and the DNA damage response. *Mol. Biochem. Parasitol.* **183**, 78–83 (2012).
34. J. Kanoh, F. Ishikawa, spRap1 and spRif1, recruited to telomeres by Taz1, are essential for telomere function in fission yeast. *Curr. Biol.* **11**, 1624–1630 (2001).
35. P. König, D. Rhodes, Recognition of telomeric DNA. *Trends Biochem. Sci.* **22**, 43–47 (1997).
36. V. Peška, P. P. Schrumpfová, J. Fajkus, Using the telobox to search for plant telomere binding proteins. *Curr. Protein Pept. Sci.* **12**, 75–83 (2011).
37. A. Padavannil, P. Sarkar, S. J. Kim, T. Cagatay, J. Jiou, C. A. Brautigam, D. R. Tomchick, A. Sali, S. D'Arcy, Y. M. Chook, Importin-9 wraps around the H2A-H2B core to act as nuclear importer and histone chaperone. *eLife* **8**, e43630 (2019).
38. A. Saha, V. P. Nanavaty, B. Li, Telomere and subtelomere R-loops and antigenic variation in trypanosomes. *J. Mol. Biol.* **432**, 4767–4185 (2020).
39. F. Aresta-Branco, M. Sanches-Vaz, F. Bento, J. A. Rodrigues, L. M. Figueiredo, African trypanosomes expressing multiple VSGs are rapidly eliminated by the host immune system. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 20725–20735 (2019).
40. E. Wirtz, S. Leal, C. Ochatt, G. A. M. Cross, A tightly regulated inducible expression system for dominant negative approaches in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **99**, 89–101 (1999).
41. G. A. Hovel-Miner, C. E. Boothroyd, M. Mugnier, O. Dreesen, G. A. M. Cross, F. N. Papavasiliou, Telomere length affects the frequency and mechanism of antigenic variation in *Trypanosoma brucei*. *PLOS Pathog.* **8**, e1002900 (2012).
42. A. Woods, T. Sherwin, R. Sasse, T. H. MacRae, A. J. Baines, K. Gull, Definition of individual components within the cytoskeleton of Trypanosoma brucei by a library of monoclonal antibodies. *J. Cell Sci.* **93**, 491–500 (1989).

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 5SA 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 (PI, to B.L.), an NIH S10 grant S10OD025252 (PI, to B.L.), a Research Grants Council GRF grant PolyU 151062/18M (PI, to Y.Z.), and a Research Grants Council AoE grant AoE/M-09/12 (PI, to Y.Z.). 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: M. Afrin, A. K. Gaurav, X. Yang, X. Pan, Y. Zhao, B. Li, *Tb*RAP1 has an unusual duplex DNA binding activity required for its telomere localization and VSG silencing. *Sci. Adv.* **6**, eabc4065 (2020).

Science Advances

*Tb*RAP1 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