PfGBP2 is a novel G-quadruplex binding protein in Plasmodium falciparum

Pratima Gurung, Ana Rita Gomes, Rafael Martins, Stefan Juranek, Patrizia Alberti, Diane-Ethna Mbang-Benet, S Urbach, Elodie Gazanion, Vincent Guitard, Katrin Paeschke, et al.

cellular microbiology

PfGBP2 is a novel G-quadruplex binding protein in *Plasmodium falciparum*

Journal:	*Cellular Microbiology*
Manuscript ID	CMI-20-0187
Manuscript Type:	Research article
Date Submitted by the Author:	10-Jul-2020
Complete List of Authors:	Gurung, Pratima; University of Montpellier, Laboratory of Pathogen-Host Interactions (LPHI) Gomes, Ana Rita; University of Montpellier, Laboratory of Pathogen-Host Interactions (LPHI) Martins, Rafael; University of Montpellier, Laboratory of Pathogen-Host Interactions (LPHI) Juranek, Stefan; University Hospital Bonn, Department of Oncology, Hematology and Rheumatology Alberti, Patrizia; Muséum National d'Histoire Naturelle, Nucleic Acid Structures, Telomeres and Evolution Mbang-Benet, Diane-Ethna ; University of Montpellier, Laboratory of Pathogen-Host Interactions (LPHI) Urbach, Serge; BioCampus Montpellier, CNRS UMR 5203, IGF,34090 Gazanion, Elodie; MIVEGEC, Laboratory of Parasitology and Mycology, University of Montpellier, CNRS UMR 5290 Guitard, Vincent; University of Montpellier, Laboratory of Pathogen-Host Interactions (LPHI) Paeschke, Katrin; University Hospital Bonn, Department of Oncology, Hematology and Rheumatology Lopez-Rubio, Jose-Juan; University of Montpellier, Laboratory of Pathogen-Host Interactions (LPHI)
Key Words:	Antigenic variations, Diseases, Genetic, Molecular genetic, Protozoa

PfGBP2 is a novel G-quadruplex binding protein in *Plasmodium falciparum*

Authors: Pratima Gurung[a,b], Ana Rita Gomes[a,b], Rafael M. Martins[a] , Stefan A. Juranek[c], Patrizia Alberti[d], Diane-Ethna Mbang-Benet[a,b], Serge Urbach[e], Elodie Gazanion[b], Vincent Guitard[a,b], Katrin Paeschke[c], Jose-Juan Lopez-Rubio[a,b*]

[a] *LPHI, UMR 5235, CNRS, University of Montpellier, 2 Place E Bataillon-Bat 24 cc107, 34095 Montpellier, France.*

[b]*Laboratory of Parasitology and Mycology, University of Montpellier, CNRS UMR 5290 MIVEGEC, 39 Avenue Charles Flahault, 34295 Montpellier, France*

[c] *Department of Oncology, Hematology and Rheumatology, University Hospital Bonn, 53127, Bonn, Germany*

[d]*Nucleic Acid Structures, Telomeres and Evolution, MNHN, CNRS UMR 7196, INSERM U1154, Sorbonne University, 43 rue Cuvier, 75231 Paris, France*

[e]*BioCampus Montpellier, CNRS UMR 5203, IGF,34090 Montpellier, France.*

** e-mail: jose-juan.lopez-rubio@inserm.fr*

Abstract

Guanine-quadruplexes (G4s) are non-canonical DNA structures that can regulate key biological processes such as transcription, replication and telomere maintenance in several organisms including eukaryotes, prokaryotes, and viruses. Recent reports have identified the presence of G4s within the AT-rich genome of *Plasmodium falciparum*, the protozoan parasite causing malaria. In *Plasmodium*, potential G4-forming sequence (G4FS) are enriched in the telomeric and sub-telomeric regions of the genome where they are associated with telomere maintenance and recombination events within virulence genes. However, there is a

little understanding about the biological role of G4s and G4-binding proteins. Here, we provide the first snapshot of G4-interactome in *P. falciparum* using DNA pull-down assay followed with LC-MS/MS. Interestingly, we identified ~29 potential G4-binding proteins (G4-BP) that binds to a stable G4FS (AP2_G4). Furthermore, we characterized the role of G-strand binding protein 2 (PfGBP2), a putative telomere binding protein in *P. falciparum*. We validated the interaction of PfGBP2 with G4 *in vitro* as well as *in vivo*. PfGBP2 is expressed throughout the intra-erythrocytic developmental cycle and is essential for the parasites in the presence of G4-stabilizing ligand, Pyridostatin. Gene knockout studies showed the role of PfGBP2 in expression of *var* genes. Taken together, this study suggests that PfGBP2 is a *bona fide* G4-binding protein, which are likely to be involved in the regulation of G4-related functions in these malarial parasites. In addition, this study sheds light on this understudied G4 biology in *P. falciparum*.

Introduction

G-quadruplexes (G4s) are non-canonical secondary nucleic acid structures that can form in guanine-rich nucleic acid sequences. These structures act as enhancers, repressors, and/or as blockades in multiple biological processes (Rhodes and Lipps, 2015). In *Plasmodium falciparum*, an etiological agent of malaria, several putative G4-forming sequences (G4FS) are identified within the key regulatory regions of this AT-biased parasite genome. Moreover, these G4FS are highly enriched in telomeric and sub-telomeric regions including virulence gene families, particularly *var* genes(Smargiasso *et al.*, 2009; Calvo and Wasserman, 2016; Stanton *et al.*, 2016; Harris *et al.*, 2018; Marsico *et al.*, 2019; Gazanion *et al.*, 2020). The *var* gene family

encodes for erythrocyte membrane protein 1 (PfEMP1), which is an exported surface protein on infected erythrocytes. PfEMP1 plays a crucial role in malaria pathogenesis through cytoadherence and antigenic variation (Scherf *et al.*, 2008). The G4 are associated with telomeric maintenance and mitotic recombination events amongst *var* genes of these parasites (De Cian *et al.*, 2008; Calvo and Wasserman, 2016; Stanton *et al.*, 2016; Gazanion *et al.*, 2020). Several G4 ligands such as naphthalene di-imide, TMPyP4 and telomestatin affect the parasite growth, thus targeting G4-associated processes might represent new strategies for malaria intervention (Calvo and Wasserman, 2016; Anas *et al.*, 2017; Belmonte-Reche *et al.*, 2018; Harris *et al.*, 2018).

Over the past decade, an increasing number of studies reported the identification of G4 binding proteins (G4-BPs) and their role in modulating the functions of G4 in numerous organisms including yeast and mammalian cells (Sissi *et al.*, 2011; Wang *et al.*, 2012; Brázda *et al.*, 2014; Von Hacht *et al.*, 2014; Mcrae *et al.*, 2017; Brázda *et al.*, 2018). Most of the G4-BPs were identified by homology-based search and/or DNA pull-down based approaches (González et al., 2009),(De Silanes et al., 2010; Martadinata and Phan, 2013). For instance, the affinity-based approach revealed that Nucleolin stabilizes DNA G4 in the promoter of c-myc proto-oncogene and reduces the promoter activity (González et al., 2009). Some RNA binding proteins including hnRNPs are shown to interact with telomeric repeat-containing RNA (TERRA) that forms G4 (De Silanes et al., 2010; Martadinata and Phan, 2013). However, despite the availability of vast information on G4-BPs in different organisms, only RecQ helicase (particularly PfWRN) has been reported to influence G4-related phenotypes such as a change in *var* gene recombination patterns and high sensitivity against G4 ligands in *P. falciparum* (Stanton *et al.*, 2016; Claessens *et al.*, 2018).

Hence, to identify a putative G4-BPs in *P. falciparum*, we performed a DNA pull-down assay followed by LC-MS/MS using a validated G4 (AP2_G4) as a bait. We next characterized PfGBP2, one of the potential G4-BPs obtained in our study. A previous study showed that PfGBP2 contains two RNA recognition motifs (RRM) and binds telomeric DNA *in vitro* (Calvo and Wasserman, 2015). We showed that PfGBP2 binds to the G4 formed by the AP2_G4 motif *in vitro* and it also interacts with the G4FS of *P. falciparum*. Thus, suggesting that PfGBP2 is a *bona fide* G4-BP in *P. falciparum*. The ChIP-Seq data also confirmed that PfGBP2 binds to the telomeres *in vivo*. Besides, we showed that the PfGBP2 is essential for the intra-erythrocytic cycle of the malarial parasites in the presence of G4-stabilizing ligand, Pyridostatin. The loss of PfGBP2 does not affect the telomere length up to 30 generations of culturing while it showed modest derepression of several *var* gene expression.

Results

Biophysical characterization of the AP2_G4 motif

Here, we chose a putative G4FS (AP2_G4), to identify the proteins that interact with the G4. This AP2_G4 motif is located at 18 bp upstream, on the non-coding strand, of the transcription start site (TSS) of a dynamically expressing putative AP2 transcription factor (Pf3D7_0934400) (Figure. 1a) (Aurrecoechea *et al.*, 2009). This motif is predicted as potential G4FS with a high propensity score of 2.55 by G4Hunter (Gazanion *et al.*, 2020). G4Hunter is a G4 prediction tool that gives the quadruplex propensity score of a given nucleic acid sequence based on G-richness and G-skewness of a sequence (Bedrat *et al.*, 2016). Next, we determined the ability of the AP2_G4 to form G4 using different biophysical experiments including thermal differential spectra (TDS), UV melting profile, and circular dichroism (CD). In

addition, we tested a control sequence (mut_G4), where the central Guanine (G) of AP2_G4 is substituted by Adenine (A) to disrupt the G4 formation (Figure 1a). The TDS of AP2_G4 displayed two positive peaks at 240 nm and 275 nm, and a negative peak at 295 nm, corresponding to the TDS signature feature of G4 formation (Figure 1b) (Mergny *et al.*, 2005). In contrast, mut_G4 did not display any significant peaks of G4 formation. Similarly, using UV melting profiles of the probes at 295 nm, we observed that AP2_G4 displays an inverted transition with a Tm of 68 ± 1 °C, characteristics of stable G4 while mut_G4 does not form any G4 (Figure S1) (Mergny *et al.*, 1998). Having confirmed that AP2_G4 forms a G4, we next evaluated the G4 topology using CD spectra recordings at 20°C with 100 mM KCl. The CD spectra of the AP2_G4 showed a positive peak at 260 nm and a negative peak at 240 nm, which are distinctive of parallel G4 (Figure 1c) (del Villar-Guerra *et al.*, 2018). Taken together, the data showed that AP2_G4 forms a stable parallel G4 whereas mut_G4 does not form a G4 under same conditions.

Identification of PfGBP2 as G4 binding protein by DNA pull-down based approach

We next used the biotinylated G4-forming AP2_G4 or control mut_G4 oligonucleotides as a bait in DNA pull-down based approach to identify G4-BPs from nuclear lysate of intra-erythrocytic parasites (Figure S2a). We identified 29 potential G4-BPs based on the label-free quantification (LFQ) intensity score (Supplementary Table 1). Among these candidates, we identified p1/s1 nuclease and Replication A1 protein, which interacts with G4 in other organisms, hence reinforcing the value of our approach (Qureshi *et al.*, 2012; Zhou *et al.*, 2013). The gene ontology (GO) analysis of the obtained candidates revealed most of these proteins are involved in nuclear-based processes such as translation, transcription, replication, and telomere maintenance

(Figure S2b). Since a majority of G4FS are identified in telomeric regions of Plasmodium genome that can form stable G4 *in vitro* (De Cian *et al.*, 2008; Smargiasso *et al.*, 2009; Calvo and Wasserman, 2016; Stanton *et al.*, 2016; Gazanion *et al.*, 2020), hence, we focused on one of the obtained candidates PfGBP2, shown to bind to a telomeric sequence of *P. falciparum in vitro* (Calvo and Wasserman, 2015).

Domain characterization of PfGBP2

PfGBP2 is a putative single-strand telomere-binding protein, which was named after its homolog *Cryptosporidium parvum* CpGBP (Calvo and Wasserman, 2015). Multiple sequence alignment of PfGBP2 revealed that PfGBP2 contains two highly conserved RNA recognition motifs (RRMs) of approximately 80 amino acid residues long, which binds to single-stranded RNAs (Keene and Query, 1991). However, unlike its homologs, PfGBP2 contains a long arginine-rich (30 %) hinge region between the RRM domains. This hinge region contains two RGG motifs, which interact with nucleic acids and protein (FigureS3a) (Thandapani *et al.*, 2013). We next sought to examine the evolutionary conservation of PfGBP2 across closely related Plasmodium species, called *Laverania* species. The *Laverania* species are a subgenus of Plasmodium that can infect apes and cause malaria. The bioinformatics analysis showed that the PfGBP2 is highly conserved among closely related species except for *P. billcollinsi G01*, suggesting the crucial role of GBP2 across the *Laverania* species (Figure S3b).

Characterization of *in vitro* interaction between recombinant PfGBP2 and AP2_G4

We expressed and purified N-terminal His-tagged PfGBP2 protein from *Escherichia coli* to confirm *in vitro* interaction between PfGBP2 and G4-forming AP2_G4 (Figure S4a). Using standard filter-binding assays (Wong and Lohman, 1993), we showed

that PfGBP2 binds to AP2_G4 (K_d = 0.7698 ± 0.62 μM), which is 23-fold tighter than binding to mut_G4 (K_d =17.72 ± 14.45 μM) (Figure 2a and S4b). The selective binding of PfGBP2 to AP2_G4 was further confirmed by electrophoretic mobility shift assay (EMSA)(Figure S4c). Altogether, these binding studies confirmed that PfGBP2 preferentially binds to the AP2_G4 *in vitro*.

Genome-wide identification of PfGBP2 binding sites in *P. falciparum*

To study the role of PfGBP2 *in vivo*, we successfully generated endogenously expressing HA-tagged PfGBP2 parasite line (iKO-PfGBP2) based on CRISPR-Cas9 gene editing (Figure S5a-c) (Knuepfer *et al.*, 2017). The immunofluorescence assay (IFA) showed that PfGBP2 is expressed in all the intra-erythrocytic stages (rings, trophozoites, and schizonts) of parasites (Figure 3a). Moreover, PfGBP2 displays a diffuse cytoplasmic pattern along with a punctuate pattern in the nuclei, predominantly in schizonts.

To determine the genome-wide binding sites of the PfGBP2, we performed two independent chromatin immunoprecipitation (ChIP) assays followed by high throughput sequencing (ChIP-Seq) from HA-tagged PfGBP2 expressing parasite line (iKO-PfGBP2) with α-HA or IgG antibodies. We used the overlapping peaks with logQ >50 obtained from both the experiments in our further studies (Figure S6a and b) (Supplementary Table 2). The ChIP-Seq data revealed that PfGBP2 binds to multiple sites throughout the genome of the parasite with the enriched peaks localized on the edges of the chromosomes (Figure 3b), which are enriched with G4FS in *P. falciparum* (Smargiasso *et al.*, 2009; Stanton *et al.*, 2016; Gazanion *et al.*, 2020). We next searched for putative G4FS within the PfGBP2 binding sites by the G4Hunter web application with a window size of 25 nucleotides and a threshold score of 1.2

(Brázda *et al.*, 2019). The G4 prediction analysis showed that our dataset contains 1,949 putative G4FS (Supplementary Table 3), while the whole genome of *P. falciparum*, with a GC content of 19.3%, harbors 6,033 putative G4FS. Therefore, the data showed that PfGBP2 binding sites have an increased GC content of 26.8% and 15-fold enrichment of G4FS as compared to the *P. falciparum* genome (hypergeometric p-value = 2,994 e-8), indicating *in vivo* binding of PfGBP2 to G4 or G4FS. To further explore the binding specificity of PfGBP2, we looked for consensus binding motif of PfGBP2 among its binding regions using the MEME suite tool with a window size of 12 nucleotides wide (Machanick and Bailey, 2011). As anticipated, the obtained top motif corresponds to the telomeric repeat of *P. falciparum* (GGGTTYA, where Y is T or C) (Vernick and McCutchan, 1988; Scherf, 1996), confirming that PfGBP2 is *bona fide* telomere binding protein (Figure 3c).

Effect of deletion of the PfGBP2 gene on parasite proliferation, telomere length and *var* gene expression

To study the physiological role of PfGBP2 on parasite viability, we excised the PfGBP2 gene in the iKO-PfGBP2 parasite line with rapamycin treatment via DiCre/Lop recombination system, thus resulting into PfGBP2-knockout (△PfGBP2) parasite line (Figure S7a-c). The growth phenotype analysis showed that there is a subtle effect on the growth of the △PfGBP2 parasite line to control parasite lines over two consecutive cycles (Figure S7d), thus highlighting the dispensability of PfGBP2 for the intra-erythrocytic cycle of the parasite. As we have confirmed that PfGBP2 interacts with G4FS, we next examined whether the presence of the highly selective G4-stabilizing ligand pyridostatin (PDS) would affect the growth of parasites lacking PfGBP2. We observed that △PfGBP2 parasite line did not survive as compared to the

control parasites when cultured with PDS (Figure 4a), thus indicating that parasites are highly sensitive to PDS in the absence of PfGBP2.

Having previously shown that PfGBP2 binds to the telomere, we further investigated whether PfGBP2 depletion affects telomere length of the parasites. To address this, we isolated the genomic DNA from △PfGBP2 parasite line to perform telomere restriction fragment (TRF) length using Southern blot analysis with a biotinylated probe against the telomeric repeats of *P. falciparum*. The data demonstrated that the TRF length profile is comparable in both △PfGBP2 and control parasite lines (Figure 4b). Moreover, the mean telomere length of these parasites is consistent with the average length of telomeres (~ 1.5 kb long) in *P. falciparum* (Figueiredo *et al.*, 2002), hence, suggesting that deletion of PfGBP2 does not affect telomere length for up to 30 generations.

Previous studies have shown that several G4s are present in the regulatory regions of *var* genes and are associated in the regulation of *var* gene expression (Smargiasso *et al.*, 2009; Stanton *et al.*, 2016; Claessens *et al.*, 2018; Gazanion *et al.*, 2020). Mutually exclusive expression of *var* genes contribute to antigenic virulence and immune evasion in *P. falciparum*. Therefore, we examined the effect of loss of PfGBP2 protein on the expression of *var* genes by qRT-PCR. The analysis showed that the same *var* gene (PF3D7_0711700) was expressed at a similar level in both △PfGBP2 and control parasite lines even after 15 generations (Figure 4c). This indicates that a mutually exclusive *var* gene expression is retained in the parasites lacking PfGBP2. In contrast, we observed slight derepression in the expression of remaining *var* genes in the △PfGBP2 parasite line, although the overall level of expression is low (Figure 4c). Altogether, the results indicate that PfGBP2 could play a role in the regulation of *var* gene expression.

Discussion

Several computational and experimental studies unveiled the existence of G4 in the genomes of *P. falciparum* (De Cian *et al.*, 2008; Smargiasso *et al.*, 2009; Calvo and Wasserman, 2016; Stanton *et al.*, 2016; Anas *et al.*, 2017; Harris *et al.*, 2018; Marsico *et al.*, 2019; Gazanion *et al.*, 2020). However, there is a lack of information on the proteins that can interact or support the G4 formation within these parasites.

Here, we report the first comprehensive study on the identification of G4-BPs that can interact with G4s and modulate their function. Using a DNA pull-down based approach, we identified 29 potential G4-BPs in *P. falciparum*. The gene ontology (GO) analysis of the obtained candidates revealed that the most of these proteins are involved in nuclear-based processes such as translation, transcription, replication, and telomere maintenance.

A previous study has shown that PfGBP2 binds to the telomeric sequence of *P. falciparum in vitro*, however, no functional implications have been elucidated (Calvo and Wasserman, 2015). Based on ChIP-Seq data, we confirmed that PfGBP2 is a bona fide telomere binding protein of *P. falciparum*. Besides, we showed that PfGBP2 binding sites are enriched with putative G4FS and PfGBP2 binds to G4 forming AP2_G4 motif *in vitro*, indicating that PfGBP2 is also a *bona fide* G4 binding protein in *P. falciparum.*

In a recent study, Gazanion *et al* showed that PDS-treated parasites displays large scale pertubations in the expression of plasmodial genes (Gazanion *et al.*, 2020). Interestingly, the data showed that there is an upregulation in the expression of PfGBP2 in PDS-treated parasites. In addition, we showed that the parasites lacking PfGBP2 are highly sensitive to PDS treatment. Altogether, the data suggest that PDS-treated parasites upregulates the expression of PfGBP2 that can recognize and bind to these folded G4 to cope with the

increased G4 formation while lacking PfGBP2 are unable to survive. Thus, we speculate that PfGBP2 plays a critical role in modulating G4-related phenotypes.

Several telomere-associated proteins are shown to interact with telomeric G4 and regulate telomere maintenance (Wang *et al.*, 2012; Pagano *et al.*, 2015). For instance, a variant of heterogeneous nuclear ribonucleoprotein A2 (hnRNP A2 *) unfolds the telomeric G4 and promotes telomere extension by providing access to telomerase while even though human POT1 unfolds the telomeric G4, it inhibits telomere extension by blocking the access of telomerase (Kelleher *et al.*, 2005). In our study, we observed that telomere binding protein PfGBP2 interacts with G4, but the loss of protein does not affect telomere length for up to 30 generations. Likewise, the deletion of Rlf6p/GBP2 , the yeast homolog of PfGBP2, did not affect the telomere length in *S. cerevisiae* (Konkel *et al.*, 1995). This led us to speculate that PfGBP2 might contribute to telomere structure complex or the subtle effect of PfGBP2-KO on telomere length could be due to presence of additional proteins that have partially or completely redundant telomere functions with PfGBP2.

In *Plasmodium falciparum*, telomere-associated proteins such as histone deacetylases PfSir2A and PfSir2B are shown to regulate the transcription of *var* genes in *P. falciparum* (Figueiredo and Scherf, 2005; Tonkin *et al.*, 2009). The knocking out of both PfSir2A and PfSir2B affected the expression of *var* gene while the absence of PfSir2A, but not PfSir2B affected the telomere lengths. Likewise, we observed that the loss of PfGBP2, modestly derepress the expression of several *var* genes after 15 generations. Given that several G4FS are present in the regulatory regions of *var* genes (Stanton *et al.*, 2016; Claessens *et al.*, 2018; Gazanion *et al.*, 2020), it could be possible that the PfGBP2 regulates the expression of *var* genes by interacting with G4. Gazanion et al showed that G4 found in the non-coding strand of var promoters modulates reporter gene expression (Gazanion *et al.*, 2020). Thus, it would

be interesting to re-examine the effect of PfGBP2 on *var* gene expression by culturing these PfGBP2 lacking parasites for a longer duration or with PDS.

On the side note, PfGBP2 might be involved in RNA metabolism due to the presence of the RNA recognition motif (RRM) region (Maris *et al.*, 2005) and its localization in the cytoplasmic region. In fact, the homologs in yeast and human have been described to regulate RNA metabolic processes (Weighardt *et al.*, 1996; Windgassen and Krebber, 2003), thus further studies are need to be carried out to study the role of PfGBP2 in RNA related functions.

In conclusion, we here report the first snapshot of the G4-interactome in *P. falciparum*. We identified several G4-BPs and further investigated one of the candidates, PfGBP2, during the intra-erythrocytic development of *P. falciparum*. We showed that PfGBP2 is a *bona fide* G4-binding protein and essential for parasites when cultured in G4- stabilizing conditions. However, the mechanism of G4 binding needs to be addressed in future studies. Moreover, our study further supports the existence of G4 in *P. falciparum* and opens new avenues for dissecting the molecular mechanisms underlying the G4 functions in these malarial parasites.

Experimental procedures

Parasite culturing

Parasites were cultured in A^+ human erythrocytes in RPMI 1640 medium (Gibco Life Technologies, 52400 RPMI 1640, HEPES) supplemented with 5% human serum and 0.5% Albumax II, 0.2 mM hypoxanthine (C.C.Pro GmbH) and 25 µg/mL gentamicin (Sigma). The cultures were kept at 37°C under a controlled trigaz atmosphere (3% CO_2, 5% O_2, and 92% N_2). Synchronization of parasites was done by a sequential combination of Percoll (Saul *et*

al., 1982) and sorbitol treatment. (Lambros and Vanderberg, 1979) Parasite development was monitored by Giemsa staining. (Moll *et al.*, 2013)

Construction of plasmid and Transfection

To generate an inducible Knockout parasite line, we have employed the combined approach of CRISPR-Cas9 and DiCre/Loxp recombination system.(Knuepfer *et al.*, 2017) The modified pLN vector (consisting of multiple cloning site and HA_3loxp region) was used as a template for donor plasmid (pLN-don-HA_3loxp) where 5' homology region, modified gene coding region, and 3' homology regions were cloned using infusion cloning. PCR amplified 400 bp upstream of the start codon to the coding region of the respective gene until the insertion site of loxpintron was cloned into ApaI and EcoRV; Gblock consisting of loxpintron and recodonized gene until stop codon was synthesized by geneart gene synthesis (ThermoFisher) and cloned into BseR1 and MluI; PCR amplified 3' HR containing 3'UTR of the gene was cloned into SpeI and XhoI. All the PCR reactions were performed using PfuUltra II fusion DNA polymerase and resulting plasmids were verified using Sanger sequencing. The gRNA was selected for the respective gene using the CHOPCHOP tool and was cloned at the BbsI site of pDC- Cas9 U6-hdhfr/yfcu vector (gifted by Ellen knuepfer).

Synchronized ring stage of the Pf3D7 p230pDiCre line was transfected with the plasmid DNA mixture consists of 60 µg of pDC2 Cas9/gRNA/hDHFR/yFCU plasmid and 60 µg of linearized pLN-loxp donor plasmid (linearized with *Apa*I and *Xho*I) by electroporation. Electroporation was performed using a Bio-Rad Gene Pulser at settings of 310 V, 950 µF, and 200 Q in 0.2 cm cuvette. The obtained time constants were 9-12 ms. These electroporated samples were immediately mixed with culture media and grown in agitation. After 4 hours of post-transfection, the culture medium was replaced with fresh media supplemented with 2.5 nM of WR99210, which was withdrawn after 5 days. Once the parasites were observed on

Giemsa staining, limiting dilution cloning was performed to obtain the individual clones with desired endogenous locus modification. Parasites were genotyped to check for correct modification at the gene locus using primers. All the primers used in the construction of the plasmids and genotyping are listed in this ESI.

In order to induce the DiCre mediated knockout of the gene, early ring-stage parasites were treated with rapamycin (20 nM) upto 12 h. After treatment, parasites were washed and returned to culture.

DNA Pull-down assay and Mass Spectrometry

The biotinylated oligonucleotides, AP2_G4 (bio-GGGATTTGGGAGGGGGGG) and Mut_G4 (bio-GAGATTTGAGAGGGAGGG) were synthesized by Eurofins Genomics. These oligos (1200 pmol /500 µL) were heated at 95°C, followed by cooling at RT for overnight in the presence of 150 mM KCl. The nuclear lysate was prepared from *P. falciparum NF54* as previously described. (Chêne *et al.*, 2012)

The subtractive-based pull-down assay was performed where the nuclear lysate was first pre-cleared with mut_G4 and then incubated with AP2_G4. Further steps were carried out as described in the published protocol. (Jutras *et al.*, 2012) Eluted fractions were dialyzed against the dialysis buffer (25mM Tris-Cl, 6mM DTT, and 1mM iodoacetamide) using Vivaspin 10MWCO tubes.

Thereafter, these eluted proteins were digested in solution. Briefly, each sample was diluted (final volume 100 µL) in TEAB 100 mM. One microliter of DTT 1 M was added and incubation was performed for 30 min at 60 °C. A volume of 10 µL of IAA 0.5 M was added (incubation for 30 min in the dark). Enzymatic digestion was performed by the addition of 1 µg trypsin (Gold, Promega, Madison USA) in TEAB 100 mM and incubation overnight at

30 °C. After digestion, peptides were purified and concentrated using OMIX (Agilent Technologies Inc.) according to the manufacturer's specifications. Peptides were dehydrated in a vacuum centrifuge. After resuspension in formic acid (0.1%, buffer A) samples were loaded onto a 50 cm reversed-phase column (75 mm inner diameter, Acclaim Pepmap 100® C18, Thermo Fisher Scientific) and separated with an Ultimate 3000 RSLC system (Thermo Fisher Scientific) coupled to a Q Exactive HF (Thermo Fisher Scientific) via a nano-electrospray source, using a 125 min gradient of 2 to 40% of buffer B (80% ACN, 0.1% formic acid) and a flow rate of 300 nL/min.

MS/MS analyses were performed in a data-dependent mode. Full scans (375-1,500 m/z) were acquired in the Orbitrap mass analyzer with a 60,000 resolution at 200 m/z. For the full scans, 3e6 ions were accumulated within a maximum injection time of 60 ms and detected in the Orbitrap analyzer. The twelve most intense ions with charge state ≥ 2 were sequentially isolated (1e5) with a maximum injection time of 45 ms and fragmented by HCD (Higher-energy collisional dissociation) in the collision cell (normalized collision energy of 28%) and detected in the Orbitrap analyzer at 30,000 resolutions. Raw spectra were processed using the MaxQuant environment (Cox and Mann, 2008) and Andromeda for database search with label-free quantification (LFQ) and match between runs (Cox *et al.*, 2011). The MS/MS spectra were matched against the UniProt Reference proteomes (Proteome ID UP000001450 and UP000005640) of *Plasmodium falciparum* and Human and 250 frequently observed contaminants as well as reversed sequences of all entries (MaxQuant contaminant database). Enzyme specificity was set to trypsin/P, and the search included cysteine carbamidomethylation as a fixed modification and oxidation of methionine and acetylation (protein N-term) as variable modifications. Up to two missed cleavages were allowed for protease digestion. FDR was set at 0.01 for peptides and proteins and the minimal peptide length at 7.

Biophysical confirmation of G4 folding of oligonucleotides

Oligonucleotides were purchased from Eurogentec (Belgium) and were resuspended at a strand concentration of 4 μM, unless stated, in the cacodylic acid buffer (10mM cacodylic acid, pH 7.2 adjusted with LiOH), containing 100mM KCl, followed by carrying out the measurements for biophysical experiments in 1cm quartz cells.

Absorbance Spectroscopy

Absorption spectra and thermal denaturation profiles (absorbance as a function of temperature) were acquired on a UVIKON XS spectrophotometer. For thermal denaturation profiles, the absorbance was recorded at different wavelengths (such as 245, 260, 273, 295, and 335 nm), while cooling and then heating the samples between 5°C and 95°C at a rate of 0.2°C/min. Melting temperatures (T_m) were graphically calculated as the intercept between the melting curves and the median lines between low and high-temperature absorbance linear baselines. Thermal difference spectra (TDS) was obtained by subtracting the absorption spectra at 2°C (low temperature) from the one at 90°C (high temperature), the spectrum were recorded after annealing from 90°C to 2°C at 0.2°C/min.(Mergny *et al.*, 2005) UV/Vis spectra were recorded on a SAFAS spectrometer.

Circular Dichroism

Circular dichroism (CD) spectra were recorded by using a Jasco J-810 spectropolarimeter at 20°C or 5°C, after annealing from 90°C to 2°C at 0.2°C/min. The oligonucleotide was dissolved at a strand concentration of 3 or 6 μM in the cacodylic acid buffer (as mentioned above).

Bioinformatic analysis

The fasta sequence of PfGBP2 (Q8IJX3) and its homologs including *Chlamydomonas reinhardtii* CrGBP1p (Q39568), *Cryptosporidium parvum* CpGBP (Q86PS0), *Saccharomyces cerevisiae* ScGBP2 (P25555), and *Homo sapiens* HnRNP-A2 (P22626) and Laverania species: P. *billcollinsi G01* (PBILCG01_1003800), *P. blacklocki G01* (PBLACG01_1004900), *P. praefalciparum strain G01* (PPRFG01_1008000), *P. reichenowi CDC* (PRCDC_1006200), *P. adleri G01* (PADL01_1005300) and *P. gaboni G01* (PGABG01_1004800) were obtained from PSI-BLAST, Plasmodb, and UniProt. Multiple sequence alignment of PfGBP2 and its homologs were carried out using the default MUSCLE algorithm at www.phylogeny.fr.

Cloning, expression, and purification of recombinant protein

The recodonized PfGBP2 coding region was synthesized and cloned into the NdeI and BamHI sites of a pET15b expression vector (Novagen, Germany) using infusion cloning. The resulting construct was confirmed by restriction digestion and sequencing. The N-terminal 6× His tagged protein was expressed in BL21 (DE3) pLysS cells grown in medium supplemented with 100 µg/mL Ampicillin and 25 µg/mL of chloramphenicol. Expression was induced at an optical density of 0.7-0.9 at 600 nm, with 0.5- 1 mM isopropyl β-D-thiogalactoside (IPTG) at 25 °C for 5h.

All the purification steps were carried out at 4 °C. Cell lysis was performed in lysis buffer (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 5% *(v/v)* glycerol and 20 mM of imidazole) by sonication (0.5s ON, 0.5s OFF at 60% pulse intensity, 3 min) using a Branson Digital Sonifier 450-D. After centrifugation of the lysate (14000 rpm, 30 min), the supernatant was loaded onto HisPur Cobalt Spin column and purification was carried out according to the manufacturer's instruction. The eluates were run through Superdex 200 16/600 (10 – 60kDa) size exclusion chromatography. The positive fraction was combined and concentrated using

Amicon Ultra-15 centrifugal filters. Expression and purification of His-tagged protein were confirmed by coomassie staining and western blot analysis with the α-his antibody.

Double filter binding assay

Double filter binding assay was performed as described earlier with some modifications.(Götz *et al.*, 2019) DNA (20 pmol) was 5' labeled with 25 µCi [γ-32P] ATP by T4 polynucleotide kinase (NEB, Ipswich,UK), followed by purification of AP2_G4 and mut_G4 structures from 7 % SDS-PAGE. DNA protein binding was analyzed by double-filter binding assays (Wong and Lohman, 1993) using a 96-well Bio-Dot SF apparatus (Bio-Rad, Hercules, CA, USA). 0.1 µM DNA in binding buffer (50 mM Tris/HCl pH 8.0, 125 mM KCl, 5 mM DTT, 10% (v/v) glycerol (Paeschke *et al.*, 2005) was incubated with the increased concentration of recombinant His-tagged PfGBP2 from 0 to 25 µM for 20 min at 21°C. After incubation, the reactions were filtered through nitrocellulose and a positively charged nylon membrane, followed by three washing steps with binding buffer and drying of the membranes. The membranes were analyzed by phosphorimaging on a Typhoon FLA 7000 (GE Healthcare). The DNA bound percentage values of PfGBP2 were determined using ImageQuant and the graph was plotted curve fitting using linear regression on GraphPad Prism.

EMSA

Electromobility shift assay (EMSA) was performed using biotinylated oligos and expressed recombinant protein as described earlier with some modifications (Takahama *et al.*, 2011). Previously used biotin-labeled oligonucleotides (AP2_G4 and Mut_G4) were diluted to 100 nM in 10 mM Tris-HCl, pH 7.5 with or without 150 mM KCl. The quadruplex formation was performed by heating the samples at 95 °C and then cooling it at room temperature overnight in the presence of potassium ions. DNA-protein binding reaction was carried out in a final

volume of 15 uL using 100 nM oligonucleotide with a varying concentration of recombinant protein in a binding buffer (50 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 0.5 mM DTT, 0.1 mg/mL bovine serum albumin, and 100 mM KCl). After the samples were incubated for 1h at 25 °C, they were resolved on a 10% native PAGE gel with or without 20 mM KCl. The native PAGE gel was pre-ran at 100 V for 1h at 4 °C. Electrophoresis was performed at 100V for 1h at 4 °C in 0.5× TBE buffer supplemented with or without 20 mM KCl. The DNA bound proteins were transferred to the Amersham Hybond- N^+ nylon membrane by electroblotting at 130 mA for 30 min using Trans-Blot Semi-dry blot transfer cell (Biorad). The interactions were detected using a Thermo Scientific™ Pierce™ LightShift chemiluminescent EMSA kit following the manufacturer's guidelines.

ChIP-Seq

Infected RBCs were cross-linked using 1% formaldehyde for 10 min at 37ºC. 0.125 M glycine was added for quenching the cross-linking reaction. The samples were washed thrice using 1X PBS (chilled) before proceeding with saponin lysis. Followed by the steps as mentioned in the previously published ChIP-Seq protocol with some modifications (Lopez-rubio *et al.*, 2013). The lysed samples were sonicated in Bioruptor (Bioruptor® Plus , Diagenode) with settings (24 cycles of sonication 10 sec on/off) to obtain the chromatin size of 200-400 bp. Two independent experiments were carried out, where in the first experiment (rep1), ChIP samples underwent a pre-clearing step of adding IgG prior to HA pull-down step whereas in the second experiment (rep2), an IgG control was prepared independently from the IP sample.

ChIP-sequencing libraries for all the samples were prepared from purified DNA using TruSeq ChIP Library Preparation Kit (Illumina), as described by the manufacturer. Libraries were then quality control checked by a bioanalyzer and qPCR prior to pooling and sequencing on

an Illumina NovaSeq 6000 system, flowcell SP (ID AHJHNGDRXX). Sequencing of 150 bp paired-ends yielded on average $70x10^6$ reads per library. Fastq files were obtained by demultiplexing the data using bcl2fastq software (Illumina), prior to downstream analysis. Mapping was performed using BWA mem and peaks were calculated using the command "callpeak" from MACS2 and the input as "control" sample. The intersection of peaks from the different replicate samples using Bedtools produced the final dataset. Motif surveys were performed using MEME (version 5.1.1 (Release date: Wed Jan 29 15:00:42 2020 -0800) The R package ChipQC and ggplots2 were used to plot the various datasets.

Growth phenotype assay

This assay was performed in trophozoites stage of synchronized parasite culture with an initial parasitaemia of 0.01% in 5% hematocrit. The parasites were treated with or without 20 nM rapamycin, followed by culturing for 2 consecutive cycles (for 6 days), where media was changed, and slides were prepared every 48 h. The assay was done in triplicates and the parasitemia was counted under microscopy. Data were represented as mean and SD of triplicates. To analyze the effect of Pyridostatin on the growth of iKO-PfGBP2 parasites line, ring-staged parasites were treated for 12 h with 1 µM Pyridostatin. Gazanion et al showed that PDS is toxic for *P. falciparum* with IC50 =5.2 ± 0.9 µM, (Gazanion *et al.*, 2020) thus we chose 1 µM PDS that has moderate effect on parasite growth for our growth phenotype analysis.

Telomere Restriction Fragment (TRF) Southern Blotting

To perform TRF southern blotting, the iKO-PfGBP2 parasite line was cultured and collected after 30 and 60 days of with or without rapamycin treatment. The genomic DNA was extracted from saponin-lysed parasites using the MN NucleoSpin Blood QuickPure kit according to the manufacturer's guidelines.

Two micrograms of genomic DNA was digested with 10 units of each enzyme: AluI, DdeI, MboII, RsaI overnight at 37°C. Digestions were precipitated and run in 1% agarose gels at 70 V for 5 h and transferred to a Hybond N+ membrane overnight through capillarity (Sambrook et al. Molecular Cloning). The DNA was then cross-linked to the membrane in a UV oven for 1 min and pre-hybridized in 6X SSC 0.1% SDS 2.5% skimmed milk for 1 h at 42°C. 10 pmol of biotinylated telomeric probes (PG184-5' GGGTTTAGGGTTTAGGGTTTAGGGTTTA 3' and PG185-5'GGGTTCAGGGTTCAGGGTTCAGGGTTCA 3') were added and incubated overnight, followed by four times 30 min washes of 6X SSC with 0.1% SDS. The biotin-labeled probes were detected using the LightShift Chemiluminescent EMSA Kit and imaging in a BioRad ChemiDoc.

Quantitative reverse transcription-PCR (qRT-PCR) on *var* genes

To perform qRT-PCR, samples were prepared from the iKO-PfGBP2 parasite line that is treated with or without rapamycin for 30 days. Using saponin treatment, parasites were harvested at 14-17h, where peak expression of *var* genes is observed. The obtained parasite pellet was resuspended in Trizol to extract the RNA, as described in the published protocol.(Moll *et al.*, 2013)

Quantitative PCR was performed on cDNA using specific primers for each *var* gene, as previously described (Salanti *et al.*, 2003) with few modifications from Dzikowski *et al* (Dzikowski *et al.*, 2006). DNase-treated RNA samples were reverse-transcribed into cDNA, using SuperScript III first-strand synthesis SuperMix (Invitrogen), according to manufacturer's instructions. Alternatively, RNA samples were run without RT enzyme, to check for genomic DNA contamination. Target genes in cDNA samples were quantified using PowerUp SYBR Green Master Mix (Applied Biosystems) and normalized using the

housekeeping gene-Serine-tRNA-ligase. The results were expressed as relative copy number with SD.

Western blotting

Saponin harvested parasites (equivalent to 10^8 parasites per lane) were lysed using 1× Laemmli Buffer supplemented with 1× protease inhibitors and 50 mM DTT. The proteins were run on a 10 % SDS-PAGE gel and transferred onto a PVDF membrane. The membrane was blocked with 3 % BSA/TBST for 1h at RT. The immunoblot was probed with primary antibody (rat α-HA 3F10) in 3% BSA/TBST for 1h at RT, followed by incubation with HRP-conjugated secondary antibody. The immunoblot was developed with Clarity Western ECL substrate (Biorad).

Immunofluorescence assay

Immunofluorescence assay was performed on smears of infected erythrocytes. The smear was fixed with 4% PFA for 15 min at RT, followed by neutralization with 0.1M glycine/PBS for 10 min. Cells were permeabilized with 0.1% Triton-X-100/PBS for 15 min. Washing with 1× PBS was performed after every step. Cells were blocked with 1.5 % BSA/PBS for 60 min, followed by incubation with primary antibody (rat α-HA 3F10) in 0.15% BSA/PBS for either 1h at RT or overnight at 4 °C. After three washes, cells were incubated with Alexa conjugated secondary antibodies, diluted in 0.15% BSA/PBS for 1h at RT. Slides were washed to remove the unbound antibodies and were stained with Hoechst (diluted in 1: 4000 in 1× PBS) for 10 min. The cells were mounted with ProLong™ Gold antifade reagent (Invitrogen, P10144) and a coverslip was placed. Images were captured and processed by Zen Blue software (Zeiss).

Acknowledgements

We acknowledge the MRI Imaging facility of the University of Montpellier for providing access to their microscopy imaging facility. This work has been supported by the Atip-Avenir program, ANR « investissements d´avenir » program (ANR-11-LABX-0024-01 « PARAFRAP ») and by the Fondation pour la Recherche Médicale (ARF20150934098).

Conflict of Interest

There are no conflicts to declare

References

Anas, M., Sharma, R., Dhamodharan, V., Pradeepkumar, P.I., Manhas, A., Srivastava, K., *et al.* (2017) Investigating Pharmacological Targeting of G-Quadruplexes in the Human Malaria Parasite. *Biochemistry* **56**: 6691–6699.

Aurrecoechea, C., Brestelli, J., Brunk, B.P., Dommer, J., Fischer, S., Gajria, B., *et al.* (2009) PlasmoDB: A functional genomic database for malaria parasites. *Nucleic Acids Res* **37**: 539–543.

Bedrat, A., Lacroix, L., and Mergny, J.-L. (2016) Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res* **44**: 1746–1759.

Belmonte-Reche, E., Martínez-García, M., Guédin, A., Zuffo, M., Arévalo-Ruiz, M., Doria, F., *et al.* (2018) G-Quadruplex Identification in the Genome of Protozoan Parasites Points to Naphthalene Diimide Ligands as New Antiparasitic Agents. *J Med Chem* **61**: 1231–1240.

Brázda, V., Cerveň, J., Bartas, M., Mikysková, N., Coufal, J., and Pečinka, P. (2018) The

amino acid composition of quadruplex binding proteins reveals a shared motif and predicts new potential quadruplex interactors. *Molecules* **23**: 2341–2357.

Brázda, V., Hároníková, L., Liao, J.C.C., and Fojta, M. (2014) DNA and RNA quadruplex-binding proteins. *Int J Mol Sci* **15**: 17493–17517.

Brázda, V., Kolomazník, J., Lýsek, J., Bartas, M., Fojta, M., Šťastný, J., *et al.* (2019) G4Hunter web application: A web server for G-quadruplex prediction. *Bioinformatics* **35**: 3493–3495.

Calvo, E.P., and Wasserman, M. (2015) PfGBP : A Plasmodium falciparum telomere binding protein. *Quim Org y Bioq* **44**: 5–10.

Calvo, E.P., and Wasserman, M. (2016) G-Quadruplex ligands: Potent inhibitors of telomerase activity and cell proliferation in Plasmodium falciparum. *Mol Biochem Parasitol* **207**: 33–38.

Chêne, A., Vembar, S.S., Rivière, L., Lopez-Rubio, J.J., Claes, A., Siegel, T.N., *et al.* (2012) PfAlbas constitute a new eukaryotic DNA/RNA-binding protein family in malaria parasites. *Nucleic Acids Res* **40**: 3066–3077.

Cian, A. De, Grellier, P., Mouray, E., Depoix, D., Bertrand, H., Monchaud, D., *et al.* (2008) Plasmodium telomeric sequences: structure, stability and quadruplex targeting by small compounds. *Chembiochem* **9**: 2730–2739.

Claessens, A., Harris, L.M., Stanojcic, S., Chappell, L., Stanton, A., Kuk, N., *et al.* (2018) RecQ helicases in the malaria parasite Plasmodium falciparum affect genome stability, gene expression patterns and DNA replication dynamics. *PLoS Genet* **14**: 1–32.

Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**: 1367–1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J. V., and Mann, M. (2011) Andromeda: A peptide search engine integrated into the MaxQuant environment. *J Proteome Res* **10**: 1794–1805.

Dzikowski, R., Frank, M., and Deitsch, K. (2006) Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. *PLoS Pathog* **2**: 0184–0194.

Figueiredo, L., and Scherf, A. (2005) Plasmodium telomeres and telomerase: The usual actors in an unusual scenario. *Chromosom Res* **13**: 517–524.

Figueiredo, L.M., Freitas-Junior, L.H., Bottius, E., Olivo-Marin, J.C., and Scherf, A. (2002) A central role for Plasmodium falciparum subtelomeric regions in spatial positioning and telomere length regulation. *EMBO J* **21**: 815–824.

Gazanion, E., Lacroix, L., Alberti, P., Gurung, P., Wein, S., Cheng, M., *et al.* (2020) Genome wide distribution of G-quadruplexes and their impact on gene expression in malaria parasites. *PLoS Genet* **16**: e1008917.

González, V., Guo, K., Hurley, L., and Sun, D. (2009) Identification and characterization of nucleolin as a c-myc G-quadruplex-binding protein. *J Biol Chem* **284**: 23622–23635.

Götz, S., Pandey, S., Bartsch, S., Juranek, S., and Paeschke, K. (2019) A novel G-quadruplex binding protein in yeast-Slx9. *Molecules* **24**: 14–18.

Hacht, A. Von, Seifert, O., Menger, M., Schütze, T., Arora, A., Konthur, Z., *et al.* (2014) Identification and characterization of RNA guanine-quadruplex binding proteins. *Nucleic Acids Res* **42**: 6630–6644.

Harris, L.M., Monsell, K.R., Noulin, F., Toyin Famodimu, M., Smargiasso, N., Damblon, C., *et al.* (2018) G-quadruplex DNA motifs in the malaria parasite plasmodium falciparum and their potential as novel antimalarial drug targets. *Antimicrob Agents Chemother* **62**: 1–32.

Jutras, B.L., Verma, A., and Stevenson, B. (2012) Identification of Novel DNA-Binding proteins Using DNA-Affinity Chromatography/Pull Down. *Curr Protoc Microbiol* 1–13.

Keene, D., and Query, C. (1991) Nuclear RNA-binding Proteins. *PNAS* **41**: 179–202.

Kelleher, C., Kurth, I., and Lingner, J. (2005) Human Protection of Telomeres 1 (POT1) Is a Negative Regulator of Telomerase Activity In Vitro. *Mol Cell Biol* **25**: 808–818.

Knuepfer, E., Napiorkowska, M., Ooij, C. Van, and Holder, A.A. (2017) Generating conditional gene knockouts in Plasmodium - A toolkit to produce stable DiCre recombinase-expressing parasite lines using CRISPR/Cas9. *Sci Rep* **7**: 3881–3893.

Konkel, L.M.C., Enomoto, S., Chamberlain, E.M., Mccune-Zierath, P., Iyadurai, S.J.P., and Berman, J. (1995) A class of single-stranded telomeric DNA-binding proteins required for Rap1p localization in yeast nuclei. *Proc Natl Acad Sci U S A* **92**: 5558–5562.

Lambros, C., and Vanderberg, J.P. (1979) Synchronization of Plasmodium falciparum Erythrocytic Stages in Culture. *J Parasitol* **65**: 418.

Lopez-rubio, J., Siegel, T.N., and Artur (2013) *Genome-wide Chromatin*

Immunoprecipitation-Sequencing in Plasmodium. .

Machanick, P., and Bailey, T.L. (2011) MEME-ChIP: Motif analysis of large DNA datasets. *Bioinformatics* **27**: 1696–1697.

Maris, C., Dominguez, C., and Allain, F.H.T. (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J* **272**: 2118–2131.

Marsico, G., Chambers, V.S., Sahakyan, A.B., McCauley, P., Boutell, J.M., Antonio, M. Di, and Balasubramanian, S. (2019) Whole genome experimental maps of DNA G-quadruplexes in multiple species. *Nucleic Acids Res* **47**: 3862–3874.

Martadinata, H., and Phan, A.T. (2013) Structure of human telomeric RNA (TERRA): Stacking of two G-quadruplex blocks in K+ solution. *Biochemistry* **52**: 2176–2183.

Mcrae, E.K.S., Booy, E.P., Padilla-Meier, G.P., and Mckenna, S.A. (2017) On Characterizing the Interactions between Proteins and Guanine Quadruplex Structures of Nucleic Acids. *J Nucleic Acids* **2017**: 1–11.

Mergny, J.-L., Phan, A.-T., and Lacroix, L. (1998) Following G-quartet formation by UV-spectroscopy. *FEBS J* **435**: 74–78.

Mergny, J.L., Li, J., Lacroix, L., Amrane, S., and Chaires, J.B. (2005) Thermal difference spectra: A specific signature for nucleic acid structures. *Nucleic Acids Res* **33**: 1–6.

Moll, K., Kaneko, A., Scherf, A., and Wahlgren, M. (2013) *Methods in malaria reserach_6th edition*. .

Paeschke, K., Simonsson, T., Postberg, J., Rhodes, D., and Lipps, H.J. (2005) Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nat Struct Mol Biol* **12**: 847–854.

Pagano, B., Margarucci, L., Zizza, P., Amato, J., Iaccarino, N., Cassiano, C., *et al.* (2015) Identification of novel interactors of human telomeric G-quadruplex DNA. *Chem Commun* **51**: 2964–2967.

Qureshi, M.H., Ray, S., Sewell, A.L., Basu, S., and Balci, H. (2012) Replication protein A unfolds G-quadruplex structures with varying degrees of efficiency. *J Phys Chem B* **116**: 5588–5594.

Rhodes, D., and Lipps, H.J. (2015) Survey and summary G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res* **43**: 8627–8637.

Salanti, A., Staalsoe, T., Lavstsen, T., Jensen, A.T.R., Sowa, M.P.K., Arnot, D.E., *et al.* (2003) Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering Plasmodium falciparum involved in pregnancy-associated malaria. *Mol Microbiol* **49**: 179–191.

Saul, A., Myler, P., Elliott, T., and Kidson, C. (1982) Purification of mature schizonts of Plasmodium falciparum on colloidal silica gradients. *Bull World Health Organ* **60**: 755–759.

Scherf, A. (1996) Plasmodium telomeres and telomere proximal gene expression. *Semin Cell Dev Biol* **7**: 49–57.

Scherf, A., Lopez-Rubio, J.J., and Riviere, L. (2008) Antigenic Variation in Plasmodium falciparum. *Annu Rev Microbiol* **62**: 445–470.

Silanes, I.L. De, D'Alcontres, M.S., and Blasco, M.A. (2010) TERRA transcripts are bound by a complex array of RNA-binding proteins. *Nat Commun* **1**: 1–10.

Sissi, C., Gatto, B., and Palumbo, M. (2011) The evolving world of protein-G-quadruplex recognition: A medicinal chemist's perspective. *Biochimie* **93**: 1219–1230.

Smargiasso, N., Gabelica, V., Damblon, C., Rosu, F., Pauw, E. De, Teulade-Fichou, M.-P.P., *et al.* (2009) Putative DNA G-quadruplex formation within the promoters of Plasmodium falciparum var genes. *BMC Genomics* **10**: 1–12.

Stanton, A., Harris, L.M., Graham, G., and Merrick, C.J. (2016) Recombination events among virulence genes in malaria parasites are associated with G-quadruplex-forming DNA motifs. *BMC Genomics* **17**: 1–16.

Takahama, K., Kino, K., Arai, S., Kurokawa, R., and Oyoshi, T. (2011) Identification of Ewing's sarcoma protein as a G-quadruplex DNA- and RNA-binding protein. *FEBS J* **278**: 988–998.

Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013) Defining the RGG/RG Motif. *Mol Cell* **50**: 613–623.

Tonkin, C.J., Carret, C.K., Duraisingh, M.T., Voss, T.S., Ralph, S.A., Hommel, M., *et al.* (2009) Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in Plasmodium falciparum. *PLoS Biol* **7**: 0771–0788.

Vernick, K.D., and McCutchan, T.F. (1988) Sequence and structure of a Plasmodium falciparum telomere. *Mol Biochem Parasitol* **28**: 85–94.

Villar-Guerra, R. del, Trent, J.O., and Chaires, J.B. (2018) G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. *Angew Chemie - Int Ed* **57**: 7171–7175.

Wang, F., Tang, M.L., Zeng, Z.X., Wu, R.Y., Xue, Y., Hao, Y.H., *et al.* (2012) Telomere- and telomerase-interacting protein that unfolds telomere G-quadruplex and promotes telomere extension in mammalian cells. *Proc Natl Acad Sci U S A* **109**: 20413–20418.

Weighardt, F., Biamonti, G., and Riva, S. (1996) The roles of heterogeneous nuclear ribonucleoproteins (hnRNP) in RNA metabolism. *BioEssays* **18**: 747–756.

Windgassen, M., and Krebber, H. (2003) Identification of Gbp2 as a novel poly(A)+ RNA-binding protein involved in the cytoplasmic delivery of messenger RNAs in yeast. *EMBO Rep* **4**: 278–283.

Wong, I., and Lohman, T.M. (1993) A double-filter method for nitrocellulose-filter binding: Application to protein-nucleic acid interactions. *Proc Natl Acad Sci U S A* **90**: 5428–5432.

Zhou, Z., Zhu, J., Zhang, L., Du, Y., Dong, S., and Wang, E. (2013) G-quadruplex-based fluorescent assay of S1 nuclease activity and K +. *Anal Chem* **85**: 2431–2435.

Figures

Figure 1: Biophysical characterization of the AP2_G4.

a) Schematic representation of localization of the selected G4 motif (AP2_G4) (blue box). The AP2_G4 (GGGATTTGGGAGGGGGGG) is present 18 bp upstream of the transcription start site (TSS) of AP2 transcription factor (TF) (Pf3D7_0934400) on the non-coding strand. b) Thermal differential spectra of AP2_G4 and control mut_G4. The TDS of AP2_G4 displays the signature characteristics of G4 while mut_G4 does not show any peaks for G4 formation. c) Circular dichroism (CD) spectra of AP2_G4. The obtained CD spectra of AP2_G4 indicates the formation of parallel G4. The AP2_G4 (GGGATTTGGGAGGGGGGG) and mut_G4 (GAGATTTGAGAGGGAGGG) oligonucleotides was used in these experiments.

Figure 2: *In vitro* binding studies for PfGBP2 and G4 interaction

a) Quantification of PfGBP2 binding to AP2_G4 and mut_G4 by double filter-binding assay. The fraction of DNA bound to PfGBP2 is plotted against the PfGBP2 concentration. The graph was plotted for two replicates using linear regression (GraphPad Prism, San Diego, CA, USA).

Figure 3: Genome-wide distribution of binding sites of PfGBP2

a) Immunofluorescence assay of HA-tagged PfGBP2. The PfGBP2 is expressed in both cytoplasm and nucleus, throughout the intraerythrocytic cycle of the parasites. DIC, differential interference contrast. HA tagged PfGBP2 were detected by α-HA antibody 3F10. Nuclei is stained with

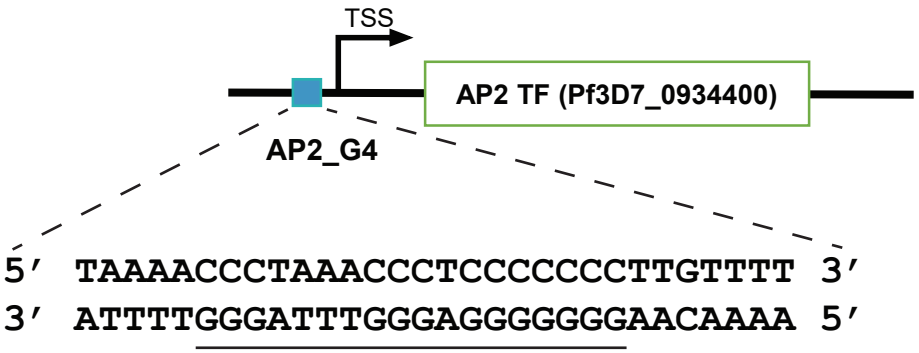
Hoescht. The scale bar = 1 µm. b) Integrative genomic view of ChIP-Seq peaks with -log(Q)>50 of PfGBP2 in all the chromosomes (1-14). The peaks (black lines) of PfGBP2 is distributed throughout all the chromosomes with enriched peaks near the chromosomal ends. The height of peak corresponds to the enrichment of the PfGBP2 peaks over IP input. c) The binding motif analysis of PfGBP2 ChIP-Seq dataset [log(Q)>50] by MEME Suite. The most significant binding motif for PfGBP2 corresponds to the Plasmodium telomere consensus motif [GGGTT(T/C)A]. The e-value describes the statistical significance of the motif while the sites represents the number of sites that contributed to the construction of the motif.

Figure 4: Effect of loss of PfGBP2 protein on parasite proliferation, telomere length and *var* gene expression

a) Growth phenotype assay of △PfGBP2 (△PfGBP2) and control (PfGBP2) parasite lines were performed over two consecutive cycles in the presence of 1 µM G4-stabilizing ligand- Pyridostatin (PDS). Upon PDS treatment, △PfGBP2 parasite line showed growth defect as compared to other control lines. Means and standard error are displayed for three independent replicates. b) TRF Southern blot analysis of △PfGBP2 parasite line (△PfGBP2) and control (PfGBP2) lines for 30 and 60 days to examine the effect of loss of PfGBP2 on the telomere length of the parasites. The telomere length of the △PfGBP2 parasite line and control parasites are observed to be comparable. c) qRT-PCR analysis of expression of var genes in the absence of PfGBP2 protein (△PfGBP2) for 30 days. The bar graphs display the relative expression of var gene expression, which is normalized using the

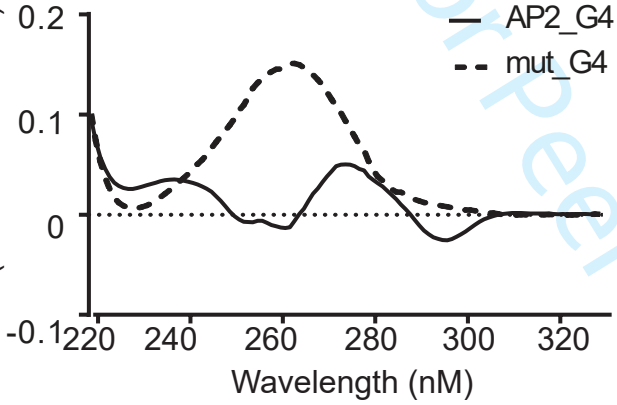
housekeeping gene- Serine-tRNA-ligase. The identical var gene is expressed in both △PfGBP2 and control parasite lines while slight derepression is observed for the majority of the var genes. Data are represented as mean ± SD (n =2).

a



AP_G4:	GGGATTTGGGAGGGGGGG
mut_G4:	GAGATTTGAGAGGGAGGG

b



c

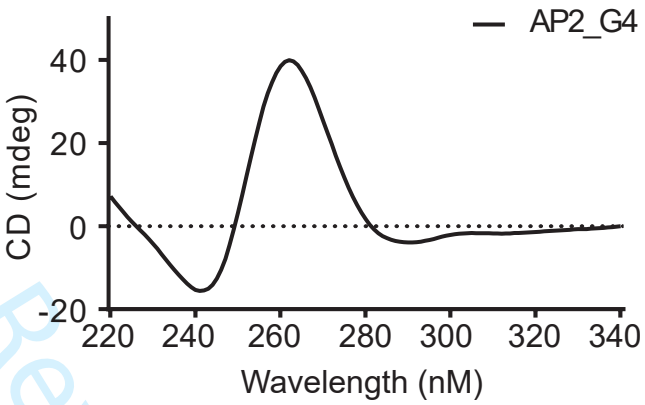
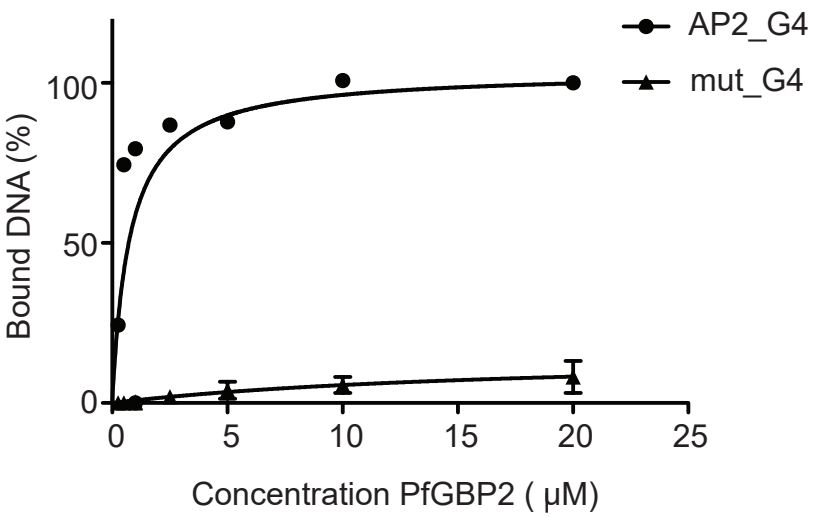
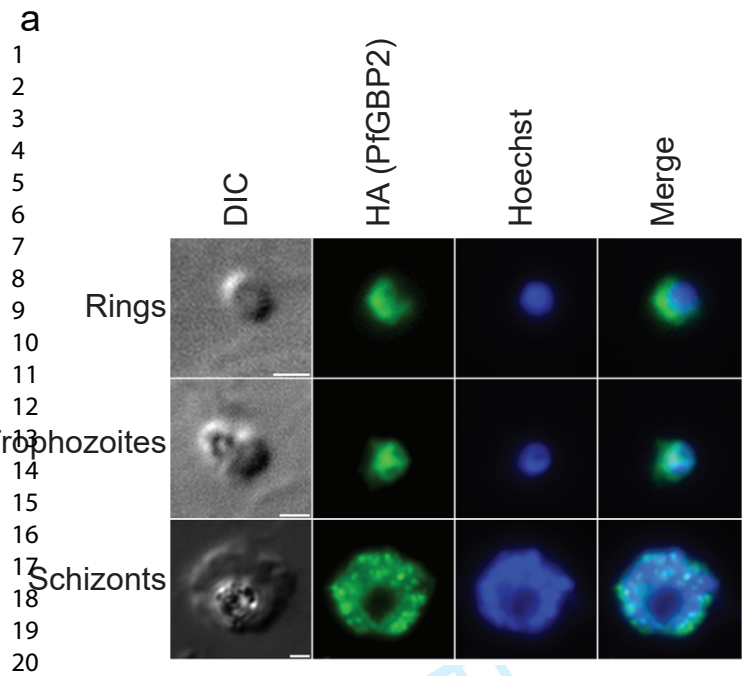
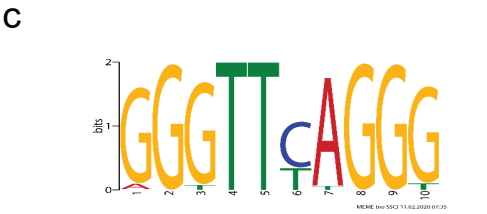


Figure 2

a

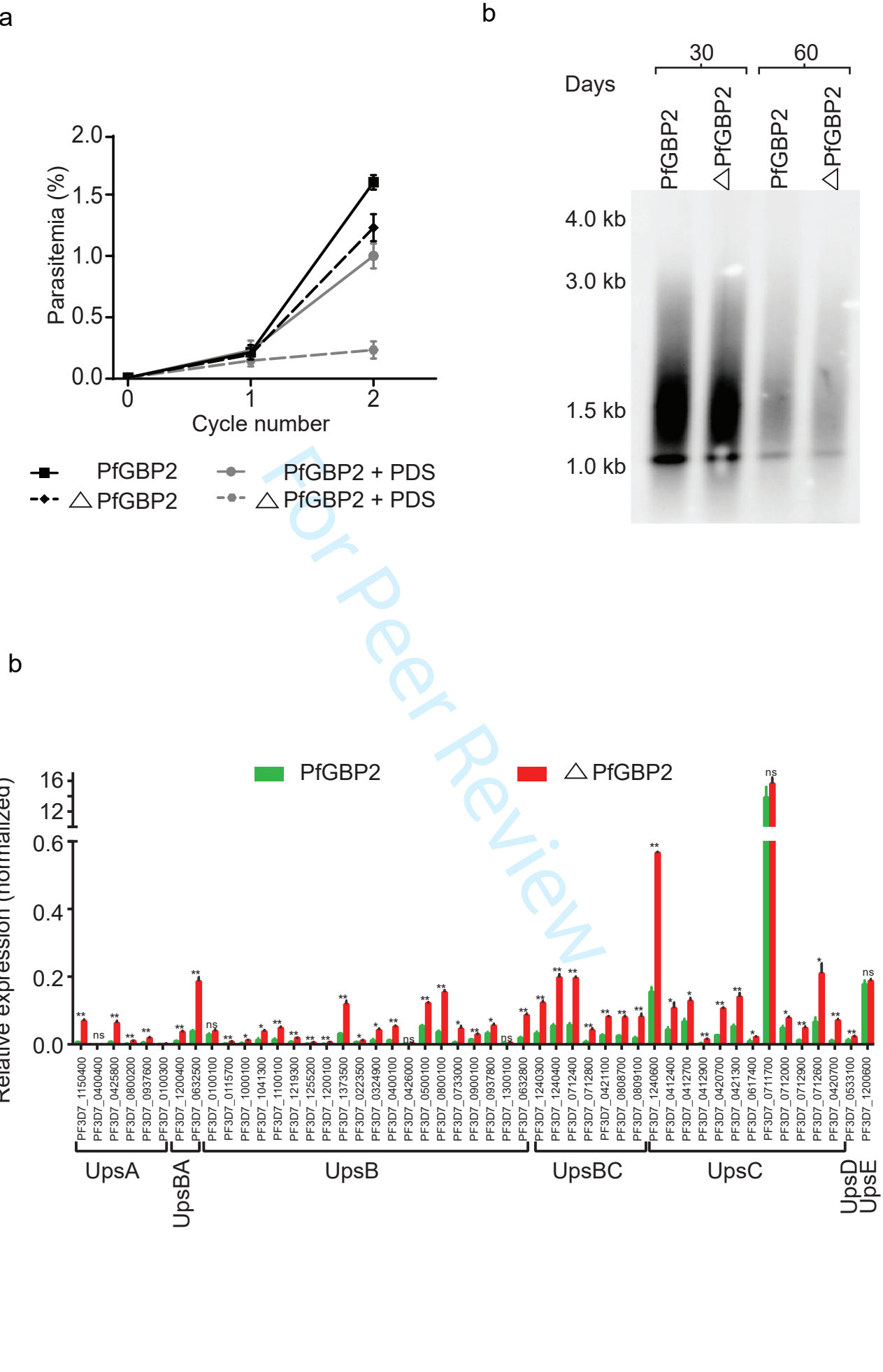






e-value- 4.7e-350; site-915

Figure 4



Supplementary figures

Figure S1: Thermal melting profile of the putative G4 motif (AP2_G4)

The thermal melting profile of AP2_G4 and mut_G4 at 295 nM. The AP2_G4 shows an inverted transition curve, characteristic of G4 structure while mut_G4 does not form any G4.

Figure S2: Identification of G4-BP using DNA pull-down based approach.

a) Outline of DNA pull-down approach. The nuclear lysate of *P. falciparum* is first incubated with the biotinylated mut_G4 to pre-clear the lysate. This pre-cleared lysate is then incubated with the biotinylated AP2_G4. The protein bound to these biotinylated AP2-G4 and mut_G4 were eluted and analyzed by LC-MS/MS. Both the biotinylated oligonucleotides (AP2/mut_G4) are heated at 95°C for 5 min followed by overnight cooling at room temperature to fold them into G4 or secondary structure. b) Pie chart of the gene ontology (GO) analysis of potential G4-BP that have nucleic acid binding function

Figure S3: Bioinformatic analysis of domain organization of PfGBP2 and its homologs.

a) Schematic domain architecture of PfGBP2 protein along with the multiple sequence alignment with its homologs. The analysis shows that PfGBP2 contains highly conserved RNA-recognition motif domains (light grey blocks labelled as RRM1 and RRM2) and RGG motif (red bar). Aligned protein

sequence includes *Plasmodium falciparum* PfGBP2 (Q8IJX3), *Chlamydomonas reinhardtii* CrGBP1p (Q39568) , *Cryptosporidium parvum* CpGBP (Q86PS0), *Saccharomyces cerevisiae* ScGBP2 (P25555), and *Homo sapiens* HnRNP-A2 (P22626). b) Multiple sequence alignment of PfGBP2 and its homologs in *Laverania* species including *P. falciparum* 3D7 (Pf3D7_1006800*), P. billcollinsi* G01 (PBILCG01_1003800), *P. blacklocki* G01 (PBLACG01_1004900), *P. praefalciparum* G01 (PPRFG01_1008000), *P. reichenowi* CDC (PRCDC_1006200), *P. adleri* G01 (PADL01_1005300) and *P. gaboni* G01 (PGABG01_1004800). The analysis shows that the PfGBP2 is highly conserved among *Laverania* species except *P. billcollinsi* G01. Multiple sequence alignment is obtained from the MUSCLE algorithm at www.phylogeny.fr. (*) indicates identical amino acid while (;) indicates conserved amino acid residues.

Figure S4: *In vitro* interaction studies for PfGBP2 and G4.

a) The coomassie stained SDS-PAGE and α-His probed immunoblot of recombinant PfGBP2 protein. The N-terminal His-tagged PfGBP2 protein was produced in *E. coli* and purified by His-Pur Cobalt Spin column, followed by size exclusion chromatography. The expected ~30.53 kDa PfGBP2 protein is detected with α-His antibody. The non-specific band detected by α-His belongs to *E. coli* , which is confirmed by MS analysis (data not shown). NEB pre-stained protein ladder was used as protein size marker. b) Double-filter binding assay for PfGBP2 and G4 interaction. The radiolabelled probes (AP2_G4 or mut_G4) are incubated with increasing concentration of PfGBP2 (0-25 μM). The top membrane (nitrocellulose) shows PfGBP2 bound to the

probes (AP2_G4 or mut_G4) while the bottom membrane (nylon+) shows unbound probes. The nitrocellulose membrane shows that PFGBP2 binds preferentially to AP2_G4 than mut_G4.c) EMSA blot for PfGBP2 and G4 interaction. The biotinylated AP2_G4 or mut_G4 was incubated with the increasing concentration of the PfGBP2 and subjected to run on 10 % native PAGE and visualized by chemiluminescence. The biotinylated AP2_G4 (100 nM) without PfGBP2 displays two different forms: folded G4 formation (upper band) and the linearized form (lower band) (lane 1) while an additional G4-PfGBP2 complex band appears with the increase in the concentration of PfGBP2 protein (0 - 818 nM) (lane 2-6). In contrast, PfGBP2 did not show any binding to mut_G4 (100nM) even at the highest concentration of PfGBP2 protein (lane 7-8).

Figure S5. Confirmation of iKO-PfGBP2 transgenic line.

a) Schematic illustration of the strategy to generate iKO-PfGBP2 parasite line. In this approach, we have replaced the endogenous copy of PfGBP2 with a modified PfGBP2 gene cassette containing two loxp sites and a C-terminal hemagglutinin (HA) tag in P. falciparum p230p line. A donor plasmid containing recodonized PfGBP2 gene (white-shaded box) fused to a 3HA tag (green box) and flanked by homologous regions 1 and 2 (grey colored box). Two loxp regions are represented by red color arrowhead. This cassette is introduced into the WT locus of PfGBP2 gene in the P. falciparum p230p strain by Cas9, which recognizes the gRNA (red flash symbol) present in the third exon. The cas9 and gRNA is cloned into the pDC-Cas9 U6-hdfr/yfcu vector. The primers (blue colored arrows) are used for verification of the intergration of the cassette into the gene locus. b) The PCR analysis of the

genomic DNA of iKO-PfGBP2 parasite line and the control parental Pf3D7 p230p line (3D7-WT) with primer pairs 1+2 (box1) and 3+4 (box 2). The analysis confirms the successful modification at endogenous locus of PfGBP2 exclusively in the iKO-PfGBP2 parasite line with the expected amplicons of 0.81 kb (box1) and 0.94 kb (box 2). The NEB 1kb ladder is used as DNA size marker (lane1). c) Western blot analysis of the endogenous expression of PfGBP2-HA protein in iKO-PfGBP2 parasite line with α-HA antibody. The immunoblot shows that the HAtagged PfGBP2 exclusively expressed in iKO-PfGBP2 parasite line as compared to the parental Pf 3D7 p230p line. The anti-H3 core is used as a loading control.

Figure S6. Optimization of ChIP-Seq assay

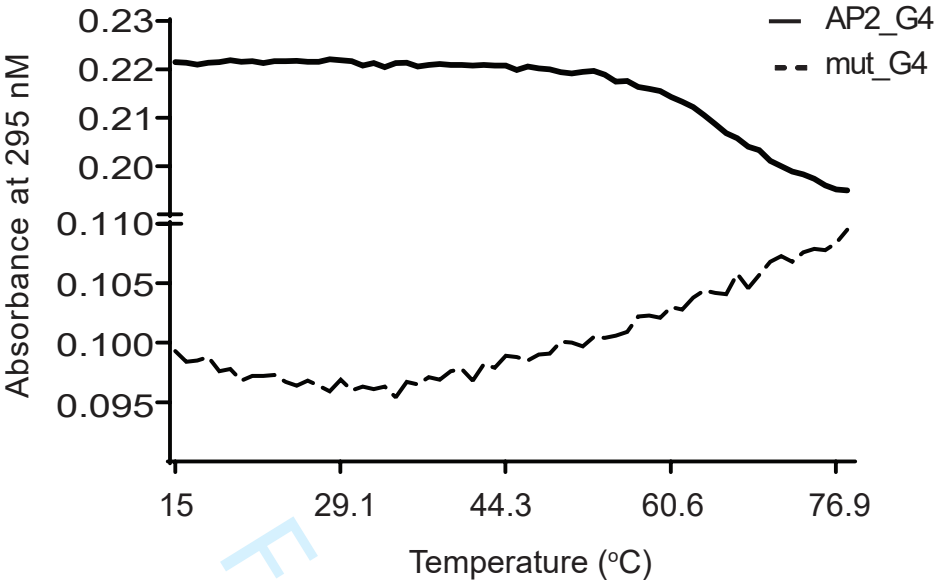
a) Calculation of percentage of reads that map within called peaks obtained in two experiments. In experiment 1 (exp 1) the PfGBP2 samples underwent a pre-clear step while in experiment 2 (exp 2) an IgG control was prepared independently from the IP sample. The bar graph shows that the more reads maps within the called peaks in exp 1 as compared to exp 2, suggesting the effective of pre-clear step to reduce the noise generated by the non-specific binding of antibody. b) Intersection of ChIP-Seq peaks obtained in two experiments (exp 1 and exp 2). The data shows that 89% of the exp 2 peaks overlaps with the exp 1 peaks, confirming the reproducibility of the ChIP-Seq replicates.

Figure S7. Conditional deletion of PfGBP2 in iKO-PfGBP2 transgenic line

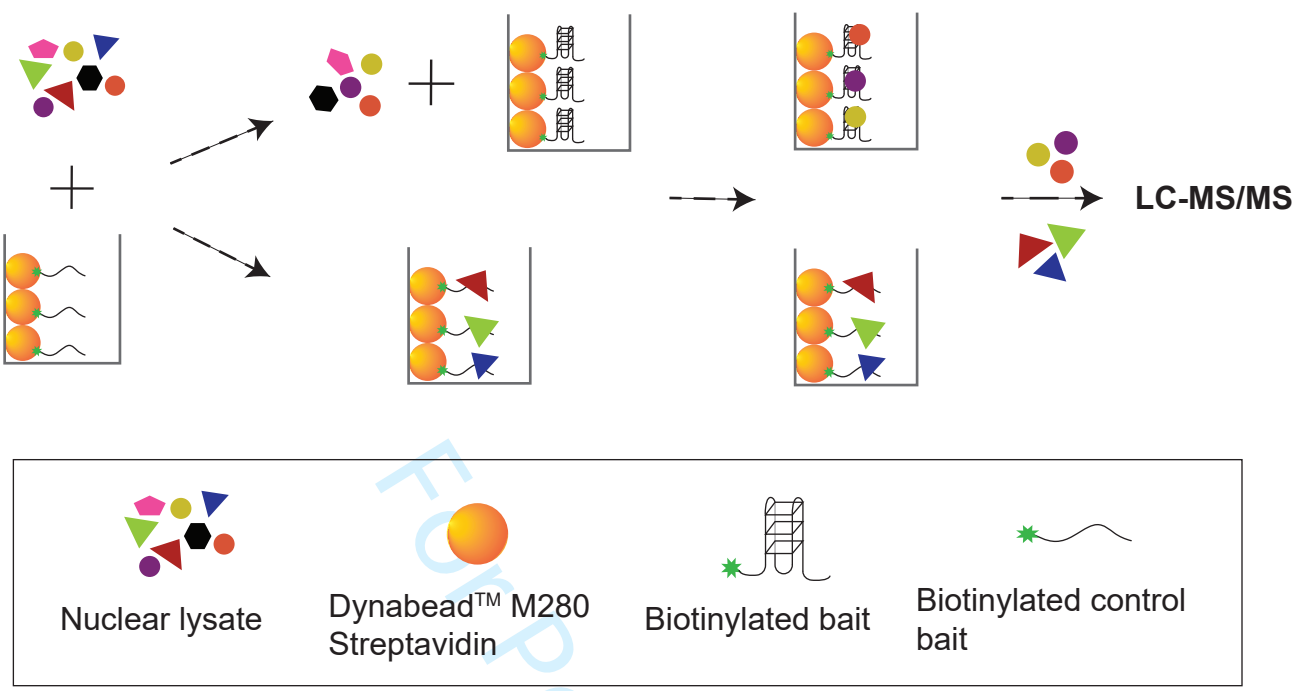
a) Schematic illustration of the strategy to induce the deletion of PfGBP2 in iKO-PfGBP2 transgenic parasite line. The gene is excised upon the rapamycin induction, which provokes the dimerisation of dicre-recombinase subunits (blue-orange shape). The dimerised DiCre-recombinase recognizes the loxp sites to promoter recombination that results into the excision of the respective gene. b) The PCR analysis of iKO-PfGBP2 parasite line treated with or without rapamycin using primer pair (1 and 4). The data shows the truncation of gene encoding PfGBP2-HA in rapamycin treated parasites while the gene remains intact in control parasites. c) Western blot analysis of lysates obtained from iKO-PfGBP2 parasite line treated with or without rapamycin to detect the expression of HA-tagged PfGBP2 using α-HA antibody. The depletion of the PfGBP2 protein (~32.5 kDa) is detected in rapamycin (rap +) treated parasites (i.e. ρPfGBP2 parasite line) as compared to the control (rap -) parasites. α- H3 (~15.4 kDa) was used as loading control. d) Growth phenotype assay of parental Pf 3D7 p230p line and iKO-PfGBP2 parasite line was performed over two consecutive cycles with or without rapamycin treatment. The loss of PfGBP2 protein showed slight effect on the parasite growth as compared to the control and parental parasites. Means and standard error are displayed for three independent replicates.

Figure S1

a



a



b

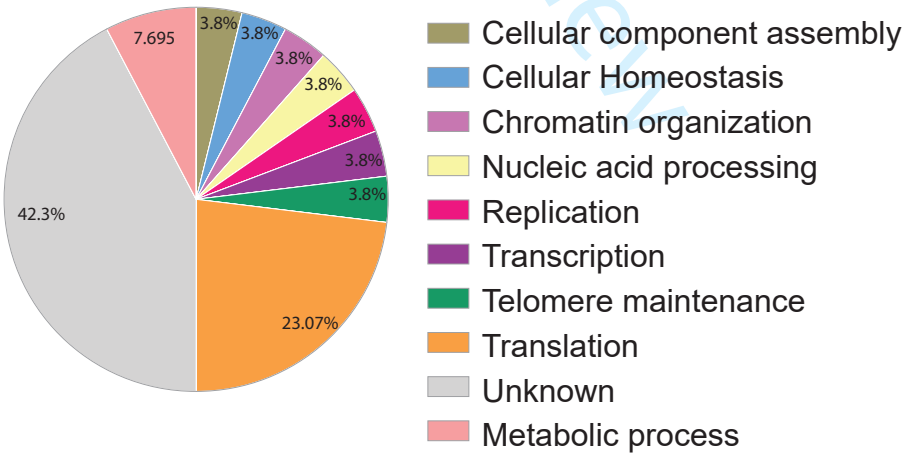
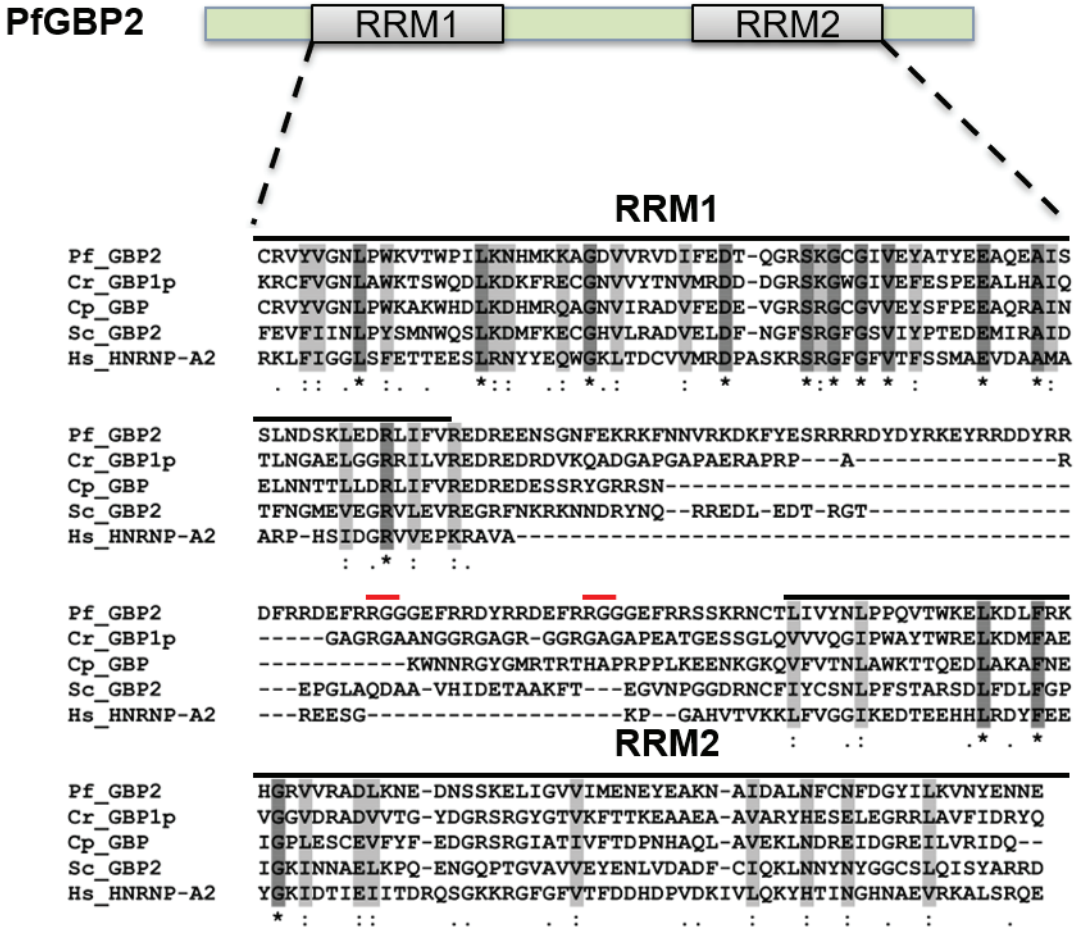
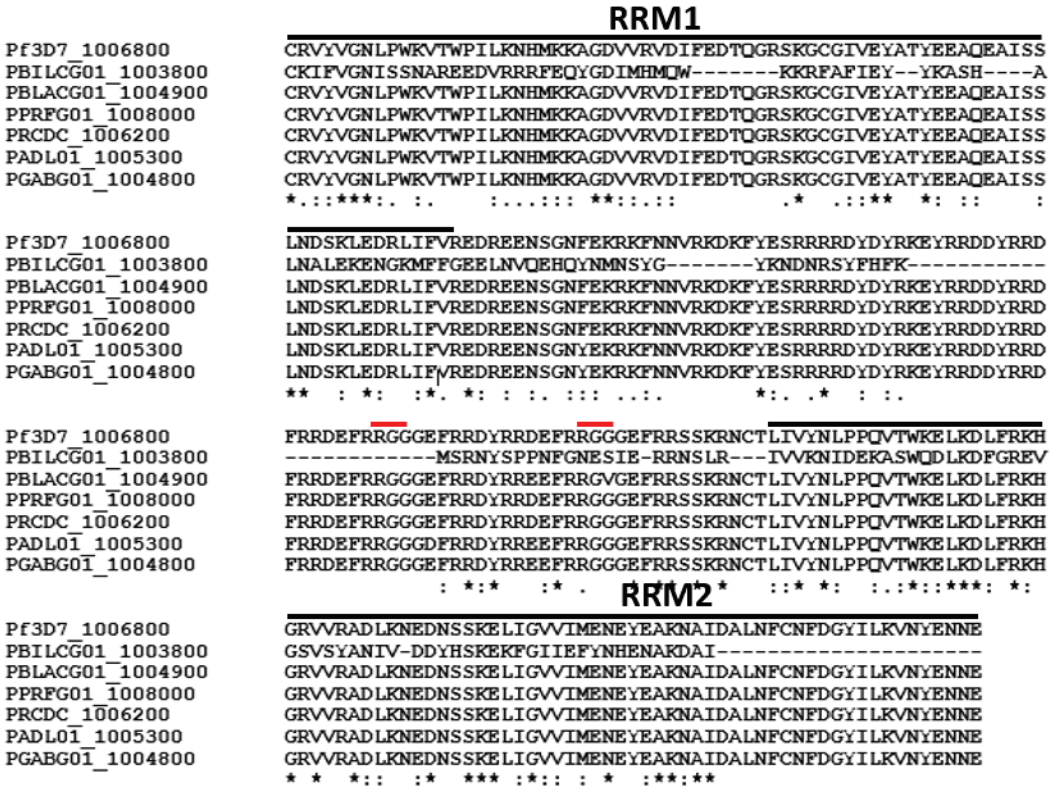


Figure S3

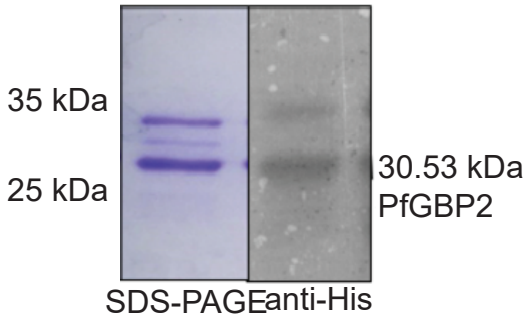
a


PfGBP2
RRM1
RRM2
RRM1
Pf_GBP2 CRVYVGNLPWKVTWPILKNHMKKAGDVVRVDIFEDT-QGRSKGCGIVEYATYEEAQEAIS
Cr_GBP1p KRCFVGNLAWKTSWQDLKDKFRECGNVVYTNVMRDD-DGRSKGWGIVEFESPEEALHAIQ
Cp_GBP CRVYVGNLPWKAKWHDLKDHMRQAGNVIRADVFEDE-VGRSRGCGVVEYSFPEEAQRAIN
Sc_GBP2 FEVFIINLPYSMNWQSLKDMFKECGHVLRADVELDF-NGFSRGFGSVIYPTEDEMIRAID
Hs_HNRNP-A2 RKLFIGGLSFETTEESLRNYYEQWGKLTDCVVMRDPASKRSRGFGFVTFSSMAEVDAAMA
SLNDSKLEDRLIFVREDREENSGNFEKRKFNNVRKDKFYESRRRRDYDYRKEYRRDDYRR
TLNGAELGGRRILVREDREDRDVKQADGAPGAPAERAPRP---A---------------R
ELNNTTLLDRLIFVREDREDESSRYGRRSN----------
TFNGMEVEGRVLEVREGRFNKRKNNDRYNQ--RREDL-EDT-RGT----------
ARP-HSIDGRVVEPKRAVA----------
DFRRDEFRRGGGEFRRDYRRDEFRRGGGEFRRSSKRNCTLIVYNLPPQVTWKELKDLFRK
-----GAGRGAANGGRGAGR-GGRGAGAPEATGESSGLQVVVQGIPWAYTWRELKDMFAE
-----------KWNNRGYGMRTRTHAPRPPLKEENKGKQVFVTNLAWKTTQEDLAKAFNE
---EPGLAQDAA-VHIDETAAKFT---EGVNPGGDRNCFIYCSNLPFSTARSDLFDLFGP
---REESG----------KP--GAHVTVKKLFVGGIKEDTEEHHLRDYFEE
RRM2
HGRVVRADLKNE-DNSSKELIGVVIMENEYEAKN-AIDALNFCNFDGYILKVNYENNE
VGGVDRADVVTG-YDGRSRGYGTVKFTTKEAAEA-AVARYHESELEGRRLAVFIDRYQ
IGPLESCEVFYF-EDGRSRGIATIVFTDPNHAQL-AVEKLNDREIDGREILVRIDQ--
IGKINNAELKPQ-ENGQPTGVAVVEYENLVDADF-CIQKLNNYNYGGCSLQISYARRD
YGKIDTIEIITDRQSGKKRGFGFVTFDDHDPVDKIVLQKYHTINGHNAEVRKALSRQE

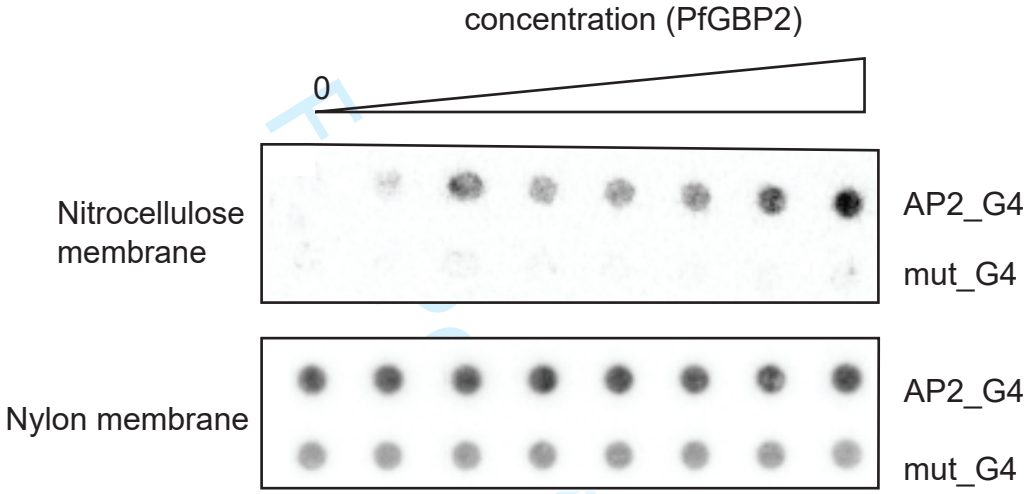
b


RRM1
Pf3D7_1006800 CRVYVGNLPWKVTWPILKNHMKKAGDVVRVDIFEDTQGRSKGCGIVEYATYEEAQEAISS
PBILCG01_1003800 CKIFVGNISSNAREEDVRRRFEQYGDIMHMQW-------KKRFAFIEY--YKASH----A
PBLACG01_1004900 CRVYVGNLPWKVTWPILKNHMKKAGDVVRVDIFEDTQGRSKGCGIVEYATYEEAQEAISS
PPRFG01_1008000 CRVYVGNLPWKVTWPILKNHMKKAGDVVRVDIFEDTQGRSKGCGIVEYATYEEAQEAISS
PRCDC_1006200 CRVYVGNLPWKVTWPILKNHMKKAGDVVRVDIFEDTQGRSKGCGIVEYATYEEAQEAISS
PADL01_1005300 CRVYVGNLPWKVTWPILKNHMKKAGDVVRVDIFEDTQGRSKGCGIVEYATYEEAQEAISS
PGABG01_1004800 CRVYVGNLPWKVTWPILKNHMKKAGDVVRVDIFEDTQGRSKGCGIVEYATYEEAQEAISS
LNDSKLEDRLIFVREDREENSGNFEKRKFNNVRKDKFYESRRRRDYDYRKEYRRDDYRRD
LNALEKENGKMFFGEELNVQEHQYNMNSYG-------YKNDNRSYFHFK----------
FRRDEFRRGGGEFRRDYRRDEFRRGGGEFRRSSKRNCTLIVYNLPPQVTWKELKDLFRKH
----------MSRNYSPPNFGNESIE-RRNSLR---IVVKNIDEKASWQDLKDFGREV
FRRDEFRRGGGEFRRDYRREEFRRGVGEFRRSSKRNCTLIVYNLPPQVTWKELKDLFRKH
RRM2
GRVVRADLKNEDNSSKELIGVVIMENEYEAKNAIDALNFCNFDGYILKVNYENNE
GSVSYANIV-DDYHSKEKFGIIEFYNHENAKDAI----------
GRVVRADLKNEDNSSKELIGVVIMENEYEAKNAIDALNFCNFDGYILKVNYENNE

a


35 kDa
25 kDa
30.53 kDa
PfGBP2
SDS-PAGE
anti-His

b


concentration (PfGBP2)
0
Nitrocellulose membrane
AP2_G4
mut_G4
Nylon membrane
AP2_G4
mut_G4

c

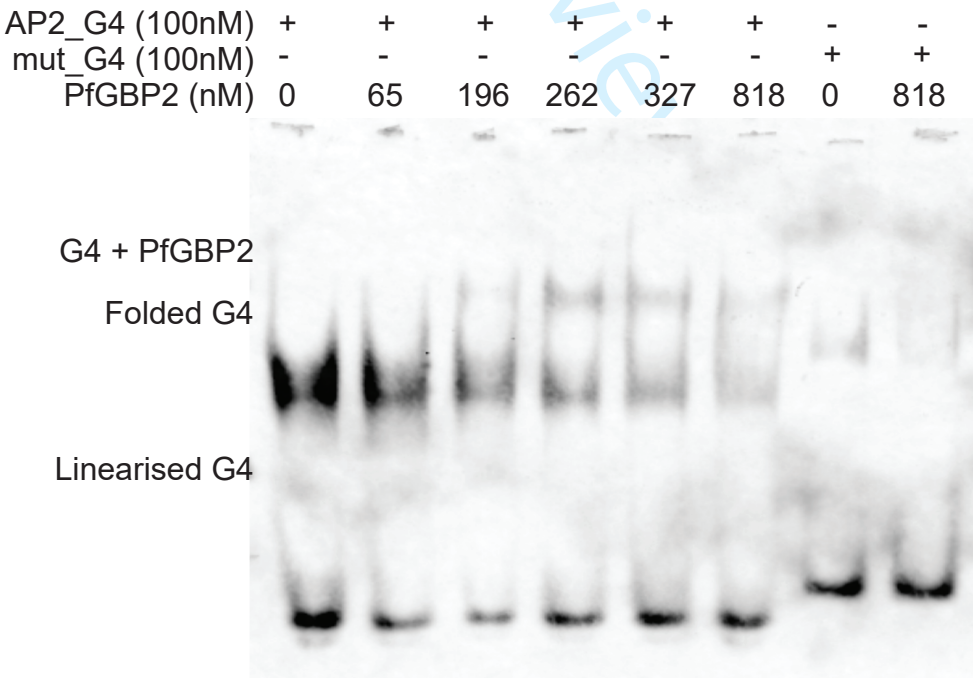
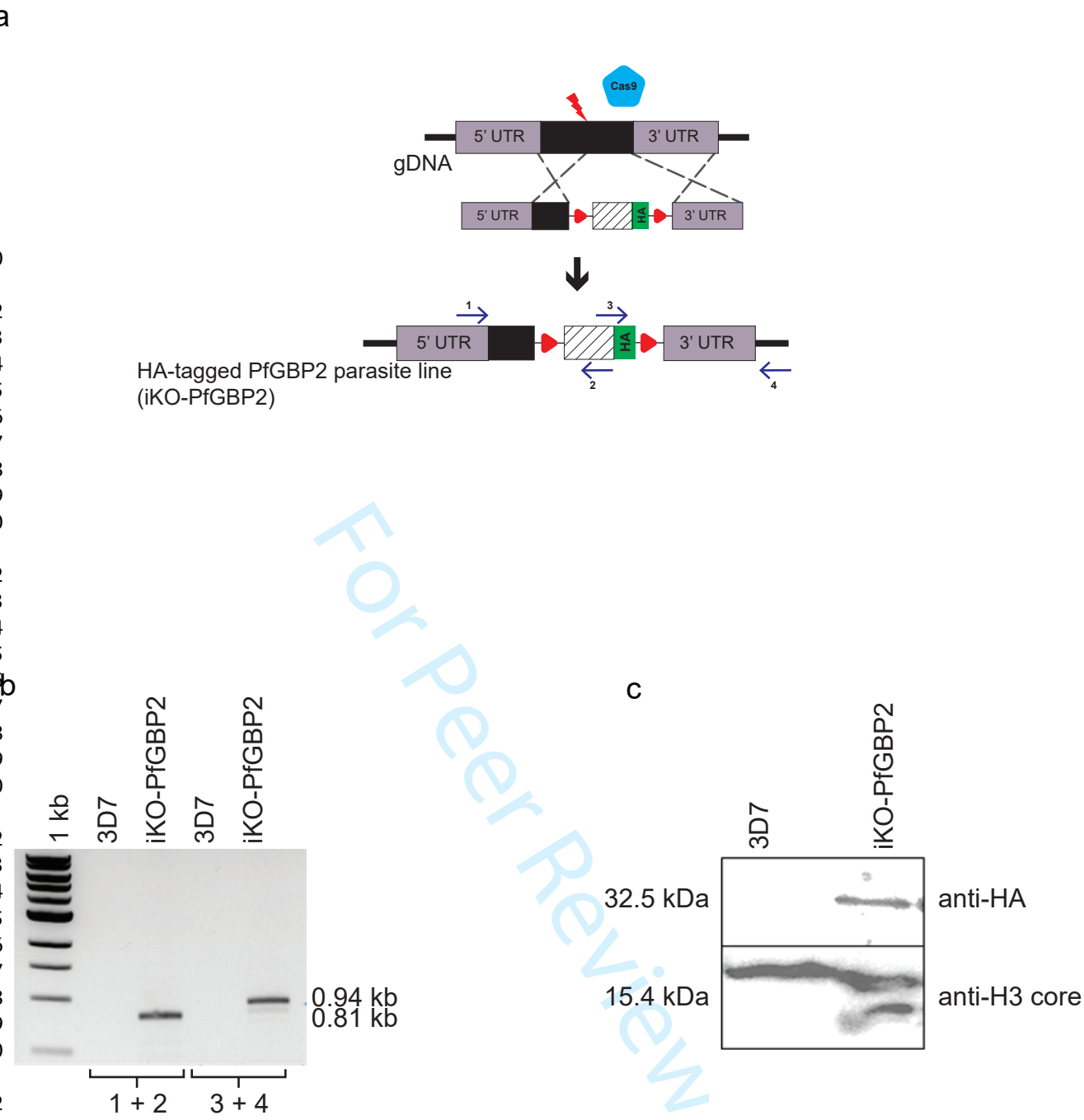

AP2_G4 (100nM) + + + + + + - -
mut_G4 (100nM) - - - - - - + +
PfGBP2 (nM) 0 65 196 262 327 818 0 818
G4 + PfGBP2
Folded G4
Linearised G4

Figure S5



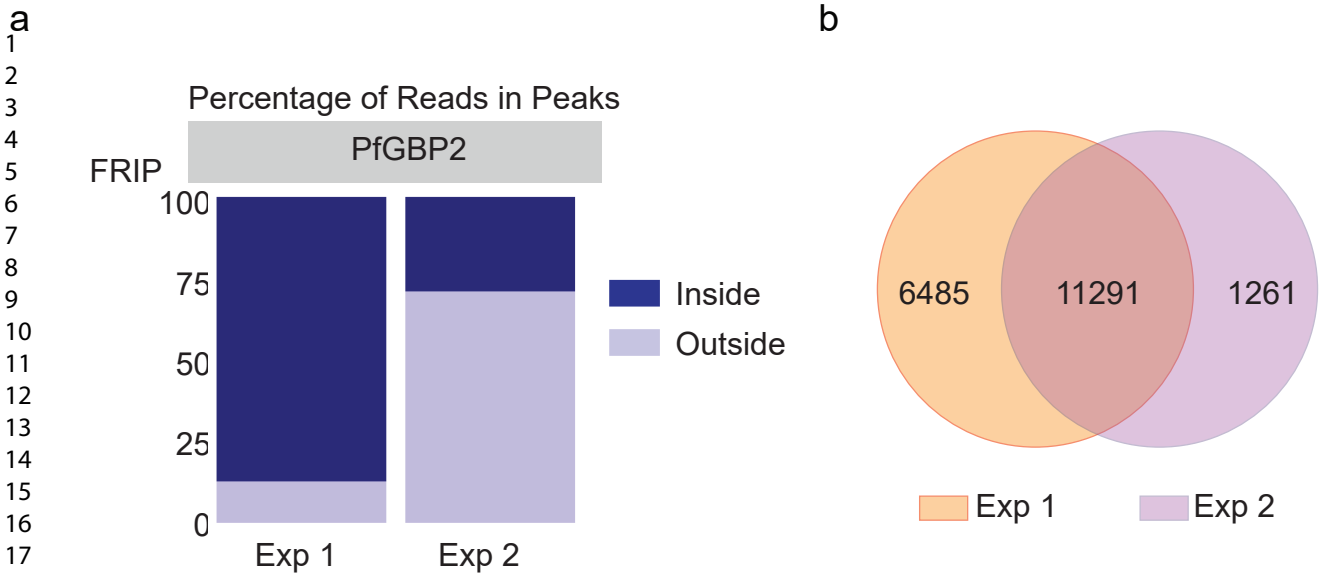


Figure S7