5850–5862 *Nucleic Acids Research, 2017, Vol. 45, No. 10* *Published online 22 March 2017*
doi: 10.1093/nar/gkx201

Yeast Sub1 and human PC4 are G-quadruplex binding proteins that suppress genome instability at co-transcriptionally formed G4 DNA

Christopher R. Lopez[1], Shivani Singh[1], Shashank Hambarde[1], Wezley C. Griffin[2], Jun Gao[2], Shubeena Chib[2], Yang Yu[3], Grzegorz Ira[3], Kevin D. Raney[2] and Nayun Kim[1,4,*]

[1]Department of Microbiology and Molecular Genetics, University of Texas Health Science Center at Houston, Houston, TX 77030, USA, [2]Department of Biochemistry and Molecular Biology, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA, [3]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA and [4]The University of Texas Graduate School of Biomedical Sciences, Houston, TX 77030, USA

Received January 02, 2017; Revised February 18, 2017; Editorial Decision March 14, 2017; Accepted March 15, 2017

ABSTRACT

G-quadruplex or G4 DNA is a non-B secondary DNA structure consisting of a stacked array of guanine-quartets that can disrupt critical cellular functions such as replication and transcription. When sequences that can adopt Non-B structures including G4 DNA are located within actively transcribed genes, the reshaping of DNA topology necessary for transcription process stimulates secondary structure-formation thereby amplifying the potential for genome instability. Using a reporter assay designed to study G4-induced recombination in the context of an actively transcribed locus in *Saccharomyces cerevisiae*, we tested whether co-transcriptional activator Sub1, recently identified as a G4-binding factor, contributes to genome maintenance at G4-forming sequences. Our data indicate that, upon Sub1-disruption, genome instability linked to co-transcriptionally formed G4 DNA in Top1-deficient cells is significantly augmented and that its highly conserved DNA binding domain or the human homolog PC4 is sufficient to suppress G4-associated genome instability. We also show that Sub1 interacts specifically with co-transcriptionally formed G4 DNA *in vivo* and that yeast cells become highly sensitivity to G4-stabilizing chemical ligands by the loss of Sub1. Finally, we demonstrate the physical and genetic interaction of Sub1 with the G4-resolving helicase Pif1, suggesting a possible mechanism by which Sub1 suppresses instability at G4 DNA.

INTRODUCTION

Guanine-rich DNA sequences with the minimal sequence requirement of $GGGN_{1-7}GGGN_{1-7}GGGN_{1-7}GGG$ can adopt a four-stranded non-B secondary structure called G-quadruplex DNA or G4 DNA (reviewed in (1,2)). Four guanine bases interact *via* Hoogsteen bonding to form a planar G-quartet; multiple G-quartets stack upon each other to form G4 DNA while the intervening sequence are extruded as loops. G4 DNA-forming sequences or G4 motifs are found in the genomes of a variety of organisms and significantly enriched near transcription start sites in the human genome, suggesting that transcriptional regulation as their putative function (3). Adding strength to this hypothesis, G4 DNA was recently shown to negatively regulate transcription of c-Myc and HIV-1 LTR promoters (4,5).

A significant correlation between G4 motifs and genome instability has been described in multiple systems. In mammalian cells, G4 motifs are found at genomic loci associated with elevated rearrangements and recombination such as immunoglobulin switch regions, ribosomal DNA loci, telomeres and at breakpoints associated with translocations and somatic copy number alterations in human cancers (6,7). In chicken DT40 cells lacking the translesion polymerase Rev1, replication stalling at a G4 forming sequence is thought to lead to epigenetic instability and de-repression of heterochromatic regions (8). In microorganisms, *N. gonorrhoeae* and *B. burgdorferi*, G4 motifs are implicated in regulation of targeted DNA rearrangement responsible for generating antigenic variation at pilE and vlsE loci, respectively (9,10).

In yeast, G4 motifs have been shown to induce a large array of genome instability events including deletion/duplications, gene conversions and gross chromosomal rearrangements (11–14). When a G4-forming se-

*To whom correspondence should be addressed. Tel: +1 713 500 5597; Fax: +1 713 500 5499; Email: nayun.kim@uth.tmc.edu
Present address: Shubeena Chib, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802, USA.

quence from the mouse immunoglobulin switch Mu region (Sμ) was inserted into the yeast genome and transcribed from an inducible promoter, genome instability was elevated by active transcription only when the G-run containing strand was located on the non-transcribed strand (15). The genome instability associated with the co-transcriptionally formed G4 DNA was exacerbated by disruption of yeast topoisomerase I and suppressed by overexpressing *Escherichia coli topA* (encoding topoisomerase I), but not *E. coli* gyrase (16). These data indicate that the function of topoisomerase I in removing negative supercoils is necessary for preventing G4 DNA formation. Overexpression of yeast RNaseH1 does not suppress the increase in G4-mediated recombination, further supporting the notion that preventing the accumulation of negative supercoils and not R-loops by Topoisomerase I is important in suppressing G4 DNA-associated genome instability. Overexpression of the gene encoding human topoisomerase I, which was shown to bind specifically to G4 DNA (17,18), is able to suppress the recombination at a G4 forming sequence in Top1-deficient yeast cells supporting the hypothesis that the increased recombination seen in topoisomerase I deficient cells is specifically due to G4 DNA formation (16).

A number of proteins in yeast and metazoans have been characterized to bind G4 DNA with high affinity and specificity (19). These include RecQ, FANCJ and Pif1 family helicases, which are capable of unwinding G4 DNA structures *in vitro* and evidently play a role in suppressing genome instability associated with G4 DNA formation. In the nematode *C. elegans*, large deletions near runs of guanines are dramatically elevated upon disruption of Dog-1, which is a homolog of mammalian FANCJ DNA helicase (20). In yeast *S. cerevisiae*, disruption of Pif1 leads to frequent stalling of DNA replication and highly elevated gross chromosomal rearrangements (GCRs) proximal to G4 motifs (13,21). Also in yeast, disruption of RecQ family helicase Sgs1 increases ectopic recombination at a G4-motif when the sequence is actively transcribed (15). Although a G4-binding property has been recognized for several other non-helicase proteins such as Top1, Nucleolin, and Ku proteins, the biological significance of such specific binding, particularly in respect to the G4-associated genome instability, is yet to be elucidated.

Sub1 (Suppressor of TFIIB) is a single stranded DNA binding protein and a highly conserved transcription factor that interacts with RNA polymerase II and III complexes (22). Sub1 is involved in multiple facets of transcription including initiation and elongation, and notably in regulation of sporulation genes (23,24). With respect to its role in genome instability, Sub1 cooperates with nucleotide excision repair proteins to enhance tolerance to oxidizing agents and protect from oxidation induced DNA breaks (24,25). Loss of Sub1 results in elevated mutagenesis caused by cytosine deamination, particularly at the 5′ end of genes, suggesting that Sub1 normally binds to the promoter regions of genes and protects them when they are single stranded during transcription (26). Recently, yeast Sub1 was identified as a G4 DNA-binding protein in a proteomic study surveying yeast whole cell extracts for factors binding to a G4 DNA-forming oligonucleotide (27). Subsequent analysis with purified, recombinant proteins showed that yeast Sub1 has significantly higher binding affinity to G4 DNA than a single stranded DNA substrate. The human homolog of Sub1, variously referred to as hSub1, p15 or PC4, is highly expressed in several different types of cancers and binds to the promoters of proto-oncogenes PLK1 and c-Myc to regulate their expression suggesting a role in cancer development (28–30). A recent study showed that PC4 co-localizes with single-strand DNA binding protein RPA at sites of replication fork stalling and aids in ensuring genome stability (31). Like yeast Sub1, purified, recombinant PC4 preferentially binds to G4 DNA structures *in vitro* (27), although whether this biochemical property extends to a relevant biological function is yet to be investigated.

To investigate whether the *in vitro* G4 DNA binding activities of yeast Sub1 and human PC4 have functional relevance, we explored the possible role of these proteins in maintaining genome stability when G4 DNA is formed *in vivo*. Using a recombination reporter system in *S. cerevisiae*, we show here that the deletion of *SUB1* significantly elevates G4-mediated genome instability in Top1-deficient cells but not R-loop mediated-genome instability in RNase H-deficient cells. The expression of PC4 complements the loss of Sub1. Sub1 protein interacts with the G4 DNA helicase Pif1 and is enriched at the co-transcriptionally formed G4 DNA in yeast cells. Overall, our data indicate that Sub1, together with its interacting partners including Pif1, plays an important role in suppressing genome instability associated with unresolved G4 DNA structure.

MATERIALS AND METHODS

Yeast strains and plasmids

Yeast strains used for the mutation and recombination assays were derived from YPH45 (*MATa, ura3-52 ade2-101 trp1Δ1)*. Construction of strains containing the *pTET-lys2-GTOP or –GBTM* constructs were previously described (15). Gene deletions were carried out through one-step allele replacement by amplification of the loxP-flanked marker cassettes (32). Sub1-expression plasmid was constructed by amplifying and cloning of SUB1 ORF along with 490 nt upstream and 250 nt downstream into the yeast *CEN* vector pRS316 (33). Full length and truncated PC4 expression plasmids pMV854 and pMV860 have been previously described (24). The standard pop-in-pop-out allele replacement method was used to introduce pif1-M2 allele using pVS31 that has been described previously (34). The complete null mutation of *PIF1* (*pif1Δ*) allele was introduced as described above except that, prior to the gene deletion, the cells were first treated with 20 μg/ml Ethidium Bromide and '*petite*' derivatives were selected.

Recombination rates

Recombination rates and 95% confidence intervals were determined using the method of the median as described previously (35). Twelve to 24 individual cultures were used to determine each rate and the associated 95% confidence intervals. Recombination rates are considered to be statistically different when the 95% confidence intervals, which are indicated in each graphs as error bars, do not overlap (36).

Western blotting to confirm Sub1 truncation

Whole cell extracts (WCE) were prepared from 10^8 cells using NaOH lysis method from strains containing full length or truncated FLAG-tagged alleles of Sub1. WCE were run on a 12% polyacrylamide gel (BioRad), and transferred using a Semi-Dry Trans-Blot Cell (BioRad). After blocking, the membrane was incubated with anti-FLAG M2 antibody (Sigma F1804) followed by anti-Mouse-HRP antibody (Santa Cruz SC2031). As a loading control, Nsr1 protein was detected using anti-Nsr1 antibody (Clone 31C4; ThermoScientific). The blot was developed using West-Q Pico Dura ECL Solution (GenDepot) and the ChemiDoc MP Imaging system (BioRad).

Circular dichroism analysis

Circular dichroism analysis was performed on a Jasco J-715 spectropolarimeter. DNA substrates were heated to 95°C for 10 min. and slow cooled to room temperature before analysis. Spectra were recorded at 25°C with 10 μM DNA in 10 mM Tris–Cl, pH 7.5 with 100 mM KCl or NaCl unless indicated otherwise.

Hf2-binding assay

Hf2 antibody expression and purification was carried out as previously described with following modifications (37). Hf2 was purified from cell pellets resuspended in lysis buffer (25mM Tris-HCl, 100 mM NaCl, 10% glycerol, 1% NP-40 and 10 mM immidazole) and sonicated using HisPur Ni-NTA resin (Thermo Scientific) following the manufacturer-suggested protocol. Eluted protein was concentrated using Amicon Ultra-4 Centrifugal Filter and stored at –20°C in 50% glycerol. For binding assay, 5′ Cy5 labeled oligoes (Sigma) were resuspended in either 100 mM KCl/10 mM Tris–Cl or 100 mM LiCl/10 mM Tris–Cl, denatured at 95°C for 5 min and finally slowly cooled overnight to allow G-quadruplex formation. Annealed oligoes (0.2 μM) were mixed with 2 μM Hf2 in 100 mM KCl or LiCl, 20 mM HEPES pH 7.5, 0.01% NP40, 5% glycerol, 5 mM $MgCl_2$ and incubated on ice for 15 min before running on 10% non-denaturing TBE-polyacrylamide gel. Gel images were captured using BioRad Chemidoc imager.

Fluorescence equilibrium binding assay

Equilibrium binding analysis was performed with a PerkinElmer Life Sciences Victor^{3}V 1420 multilabel counter with filters set at 485 and 535 nm. Fluorescence polarization measurements were recorded at 25°C in assay buffer (10 mM Tris–Cl pH 7.5, 1 mM EDTA, 100 mM KCl and 0.1 mg/ml BSA). 6-carboxyfluorescein (FAM)-labeled DNA substrates (5 nM final concentration) were titrated with increasing concentrations of recombinant Sub1 or PC4 to measure binding. The FAM-labeled substrate without protein was used to normalize the change in fluorescence polarization observed in the protein titrated samples. Fluorescence data were plotted as change in anisotropy against protein concentration with Kaleidagraph software (Synergy Software, Reading PA) and fit to the quadratic equation to obtain dissociation constants (K_D). Recombinant Sub1 and PC4 were expressed and purified as described previously (27,38)

Chromatin immunoprecipitation

ChIP was performed following a previously described protocol (39). Briefly, chromatin fraction was isolated from yeast cells that were cross-linked with 1% formaldehyde and quenched with 136 mM glycine. Following the sonication step to shear the DNA to ∼750 bp (QSONICA sonicator with a microtip), samples were incubated with anti-FLAG antibody (clone M2 - Sigma) conjugated to Protein G Dynabeads (Life Technologies) overnight at 4°C. After washing, cross-link was reversed by incubation with proteinase K at 42°C for four hours and at 65°C for 12 h. DNA was isolated using MiniElute PCR Purification Kit (Qiagen). qPCR was performed using SensiFAST SYBR no-ROX Mastermix (Bioline) and CFX Connect instrument (Biorad). For each PCR reaction, 10 ng of input or ChIP DNA was used as template. The final concentration of the primer was 0.4 μM each. Cycling conditions were 95°C for 3 min followed by 40 cycles of 95°C for 5 s, 60°C for 10 s, and 72°C for 10 s. Ct values were determined using the CFX Manager software. Ct values for each ChIP samples were first normalized to the ChIP experiment carried out with yeast cells expressing untagged-Sub1 proteins and then divided by the values for the *CAN1* locus to calculate the relative fold enrichment. Primers used in the qPCR analysis are listed in the Supplementary Table S1.

TMPyP4 sensitivity

For each indicated strain, TMPyP4 (EMD Millipore) to 200 μM final concentration was added to six individual cultures that were grown to early log phase ($O.D._{600}$ of 0.4–0.5) in YEPD at 30°C. After 20 h incubation with TMPyP4 or DMSO at 30°C, cells were collected by centrifugation and washed twice with water. 10-fold serial dilutions of washed cells were spotted onto YEPD-agar plates and incubated at 30°C and pictures were taken 2 days after spotting.

Co-immunoprecipitation

Immunoprecipitation was performed as described (40) with some modifications. Approximately 100 OD_{600} units of mid-log-phase cells grown in YEPD were harvested, washed with water and resuspended with 350 μl of lysis buffer (50 mM HEPES–NaOH, pH 7.5, 100 mM NaCl, 1 mM EDTA, 10% glycerol, 0.05% NP-40, 1 mM DTT, 1 mM PMSF, 1× protease inhibitor cocktail (Roche)) and 300 μl of 0.5-mm glass beads. Following mechanical lysis of the cells with Biospec Mini-Beadbeater for 4 × 60 s periods with 5 min cooling at interval, additional 200 μl of lysis buffer was added to each sample. The cell lysate was centrifuged at 12K RPM for 30 min. For each IP, 400 μl of the supernatant was incubated with 20 μl Anti-FLAG M2 affinity gel (Sigma, Cat# A2220) for 4 h at 4°C. The beads were washed with the lysis buffer three times for 5 min each and then eluted by boiling in 1× SDS-PAGE loading buffer followed by immunoblotting analysis using anti-HA-HRP (Sigma; Cat# H6533-1VL) or anti-FLAG-HRP (Sigma; Cat# A8592).

Interaction of purified Sub1 and Pif1

Yeast Pif1 and Sub1 were overexpressed and purified as described previously (27,41). Purified recombinant Sub1 or BSA were covalently cross-linked onto epoxy-activated Dynabeads M-270 (Invitrogen) according to the manufacturer's instructions. Briefly, saturating amounts of protein were used for coating 5 mg of Dynabeads M-270 at 37°C for 24 h in the presence of 1.5 M ammonium sulfate. Co-precipitation experiments were performed by incubating purified Pif1 (50 μg) with Sub1- or BSA-coated dynabeads in a 200 μl binding buffer (25 mM HEPES pH 7.4, and 100 mM NaCl) with rotation at 4°C overnight. Dynabeads were captured using a magnet and washed five times in 1 ml of buffer containing 0.2% Tween 20. Pif1 protein that co-precipitated with the beads was eluted by addition of 30 μl of Laemmli sample buffer and heating at 95°C for 10 min. The sample was resolved by 10% SDS-PAGE gel and the proteins were visualized by Coomassie staining.

RESULTS

Sub1 suppresses G4-mediated recombination in Top1-deficient cells

In the current study, we employed the recombination reporter assay designed to determine how G4-forming sequence leads to elevated genome instability. The details of the reporter assay and the sequence of the G4 motif containing insert are described in previous publications (15,35). Briefly, in this system, a G-run containing sequence from the mouse immunoglobulin heavy chain switch Mu region (Sμ) was inserted into the yeast genome within the context of the *LYS2* gene transcribed from a tetracycline/doxycycline-repressible promoter (*pTET*). The Sμ sequence was inserted in two different orientations placing the G-run-containing strand either on the top/non-transcribed strand (*GTOP*) or on the bottom/transcribed strand (*GBTM*). The formation of G4 DNA is favored when the G-rich strand located on the non-transcribed or top strand is transformed into single strands during transcription freeing the guanine bases to interact with each other through Hoogsteen base-pairing. When G-runs are located in the transcribed or bottom strand, they will be occupied in base pairing with the nascent RNA strand and will not be free to fold into G4 DNA. Thus, any factors involved in the formation or stability of G4 DNA should affect the recombination occurring at the *pTET-lys2-GTOP* construct, with little to no effect on the rate of recombination at the *pTET-lys2-GBTM* construct. We previously demonstrated such orientation-dependent elevation in recombination when these constructs are transcribed actively indicating that co-transcriptionally formed G4 DNA elevates instability at this locus (16,35,42).

In order to determine the role of Sub1 in maintaining genome stability at G4 DNA loci, we disrupted Sub1 in strains containing the *pTET-lys2-GTOP* or *–GBTM* construct and determined the recombination rates at this locus. Under low transcription conditions where *pTET* promoter was maximally repressed by addition of 2μg/ml doxycycline, the rate of recombination for *pTET-lys2-GTOP* was about 2.5-fold higher in *sub1Δ* than in *WT* backgrounds (Supplementary Figure S1). Sub1-disruption, however, did not further elevate the rate of recombination in *WT* background under high transcription conditions (Figure 1A). For the *pTET-lys2-GBTM* construct, Sub1-disruption did not affect the rates of recombination under high or low transcription conditions (Figure 1A and Supplementary Figure S1).

In *top1Δ* backgrounds, accumulation of negative supercoils during highly active transcription promotes the formation of G4 DNA structure and significantly elevates G4-associated genome instability (16,35). We deleted *SUB1* in *top1Δ* background to determine whether Sub1 plays a role in resolving G4 DNA that accumulates at actively transcribed regions in the absence of Top1. Under low transcription conditions, there was a small—1.5- or 2.5-fold—increase in the rates of recombination for the *pTET-lys2-GBTM* or *–GTOP* constructs, respectively (Supplementary Figure S1). When highly transcribed, the rate of recombination for the *pTET-lys2-GTOP* was elevated by ∼7-fold in *top1Δ sub1Δ* strain compared to *top1Δ* (Figure 1B). For the *pTET-lys2-GBTM,* the rates of recombination were indistinguishable between *top1Δ* and *top1Δ sub1Δ* backgrounds. Mutation of the catalytic tyrosine of Top1 (Top1 Y727F) results in a non-functional protein in respect to removal of supercoils and in dominant negative effect in respect to G4-associated genome instability with recombination rate even higher than *top1Δ* (16). This is likely due to this mutant protein still retaining the high affinity G4 binding activity described for the wildtype enzyme (17,18). Loss of Sub1 in the Top1 Y727F-expressing strain resulted in a significant elevation in the rate of recombination at the *pTET-lys2-GTOP* but not at the *–GBTM* constructs (Figure 1C).

We previously reported that defects in RNase H function elevated recombination for both *pTET-lys2-GTOP* and *–GBTM* constructs (15). Accumulation of RNA:DNA hybrids, which ensues when RNase H encoding genes, *RNH1* and *RNH201*, are deleted in *S. cerevisiae* (*rnh1Δ rnh201Δ*), is the likely cause of such elevated genome instability since instability is exacerbated by transcription (15) and suppressed by overexpression of *RNH1* (16). Upon deletion of *SUB1*, we did not observe any significant elevation in recombination rate in *rnh1Δ rnh201Δ* backgrounds for the *pTET-lys2-GTOP* or *–GBTM* construct (Figure 1D). Overexpression of *RNH1*, which reduced the rates of recombination both the *pTET-lys2-GTOP* and *–GBTM* constructs by ∼3-fold in *rnh1Δ rnh201Δ* backgrounds, did not affect the rates of recombination in *top1Δ sub1Δ* backgrounds (Figure 1E). Overall, these data indicate that the role of Sub1 in genome instability is specific to G4 DNA formed due to transcription-induced negative helical torsion and not broadly related to the transcription process or RNA:DNA hybrid accumulation.

Human transcription co-activator PC4 can complement the loss of Sub1 and suppress G4-associated recombination

PC4, a human homolog of Sub1, which shares 34% identity and 45% homology with Sub1, was previously reported to functionally complement Sub1 in the tolerance of oxidative stress (Figure 2A) (24,25). To determine whether PC4 can

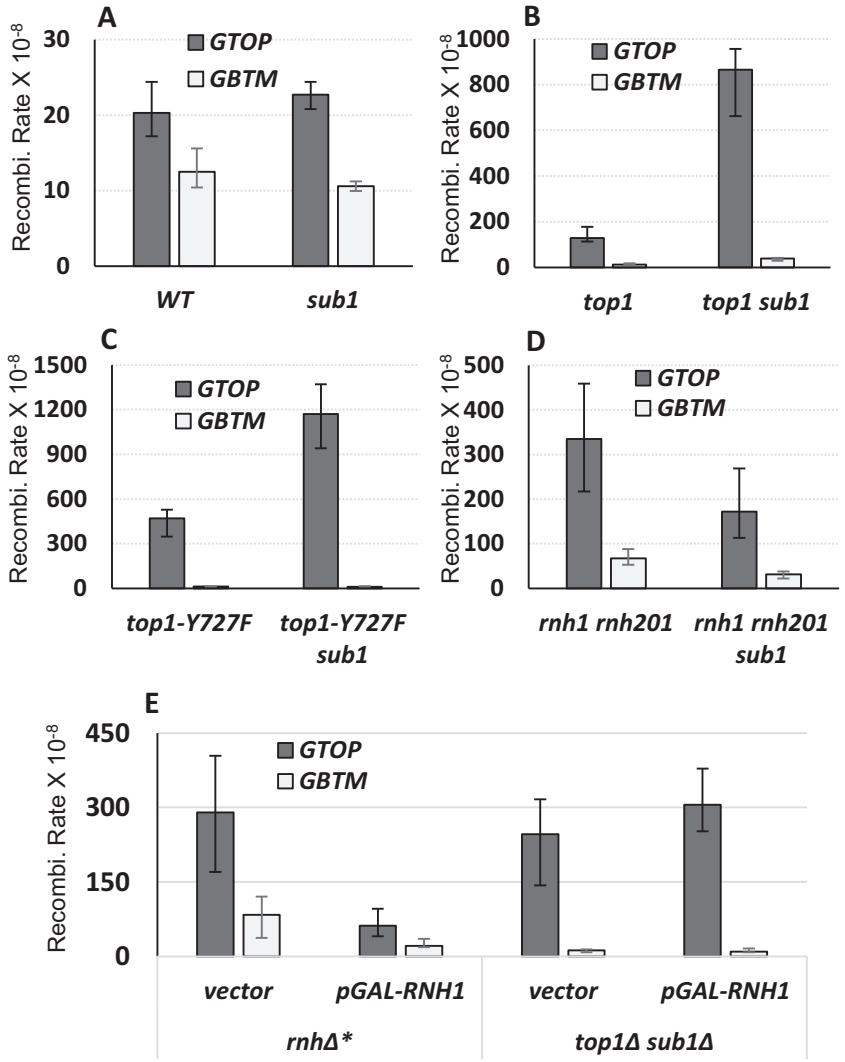

Figure 1. *SUB1* deletion elevates G4-associated recombination in *top1Δ* backgrounds. The recombination assay and the sequence of the Sμ fragment used to construct G4 used to construct *pTET-lys2-GTOP* or *–GBTM* cassettes are described in detail in previous publications (15,35). '*GTOP*' and '*GBTM*' refers to the orientation of the Sμ sequence within the *LYS2* gene relative to the direction of transcription. When in '*GTOP*' orientation, the guanine runs are present on the non-transcribed, top strand; When in '*GBTM*' orientation, the guanine runs are present on the transcribed, bottom strand. All graphs are showing the rates of recombination ($\times 10^{-8}$). Error bars indicate 95% confidence intervals. (**A**) Recombination rates in *WT* backgrounds. (**B**) Recombination rates in *top1Δ* backgrounds. (**C**) Recombination rates in *top1 Y727F* backgrounds. (**D**) Recombination rates in *rnh1Δ rnh201Δ* backgrounds. (**E**) Recombination rates in *rnh1Δ rnh201Δ* (*rnhΔ**) or *top1Δ sub1Δ* backgrounds with expression of the indicated plasmids.

sufficiently complement Sub1 for its function in suppressing G4-associated instability, we expressed full-length Sub1 or PC4 from plasmids in *top1Δ sub1Δ* strains. The ectopic expression of Sub1 or PC4 reduced the rate of recombination for the *pTET-lys2-GTOP* construct by 9.6- and 10.3-fold, respectively, thereby resulting in the rates of recombination that are significantly lower than in *top1Δ* single deletion strains (Figure 2B). In *top1-Y727F sub1Δ* strains, ectopic expression of Sub1 or PC4 resulted in >10-fold decrease in the rate of recombination for the *pTET-lys2-GTOP* construct (Figure 2C). The rate of recombination for the *pTET-lys2-GBTM* construct remained unaffected by the expression of Sub1 or PC4 from plasmids in *top1Δ sub1Δ* or *top1-Y727F sub1Δ* strains. There is a ≥10-fold difference between the rates at the *pTET-lys2-GTOP* and *–GBTM* in *top1Δ sub1Δ* or *top1-Y727F sub1Δ* backgrounds. Upon expression of PC4 or Sub1, this difference between the rates at the *pTET-lys2-GTOP* and *–GBTM* was reduced to <2-fold, which is similar to those in WT background. This result suggests that the level of endogenous Sub1 in *top1Δ* or *top1-Y727F* cells could be a rate-limiting factor in preventing G4 DNA-mediated recombination. This was supported by the data in Figure 2D and E showing that when Sub1 or PC4 were expressed in *top1Δ* or *top1-Y727F* strains where the expression of WT Sub1 is unperturbed, the rate of recombination was further decreased for the *pTET-lys2-GTOP* construct but remained unchanged for the *pTET-lys2-GBTM*.

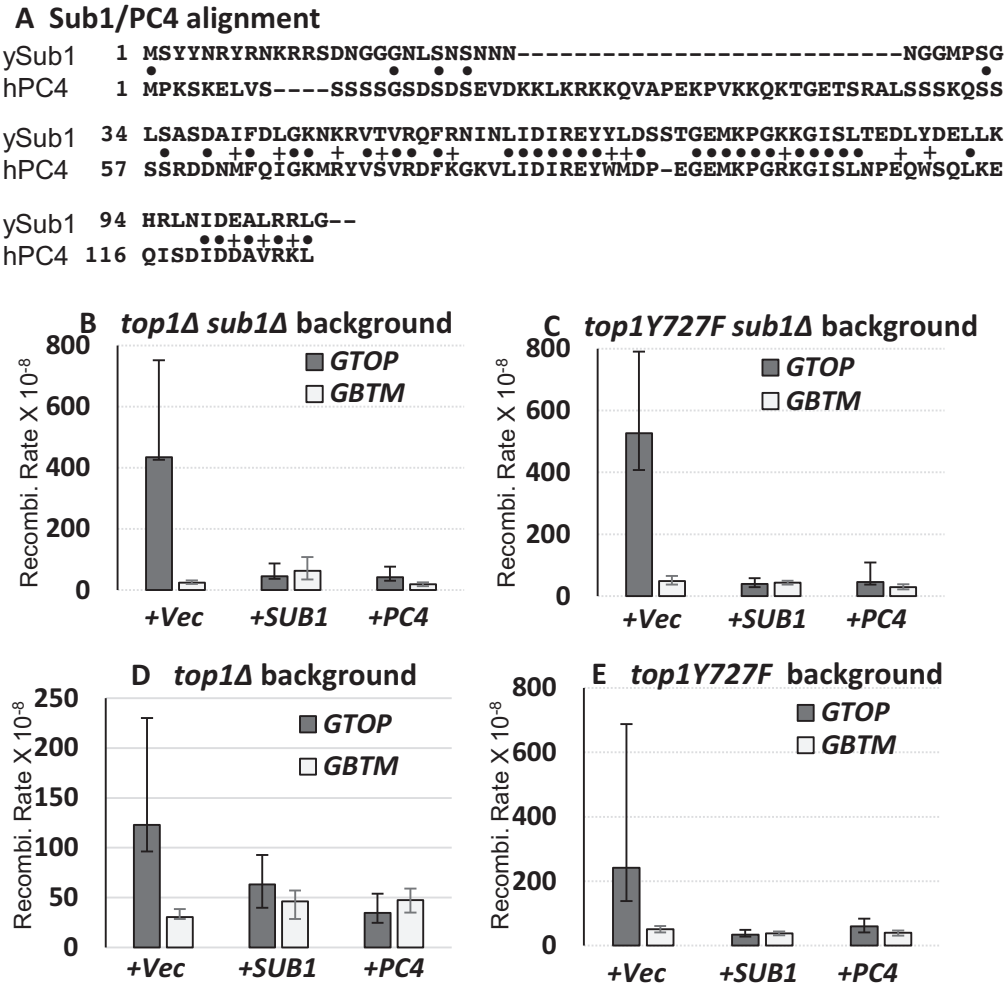

Figure 2. Complementation of *SUB1* deletion in *top1Δ* and *top1Y727F* backgrounds. (**A**) Protein sequence alignment between human PC4 (hPC4) and *S. cerevisiae* Sub1 (ySub1). Identical and conservative residues are indicated by bullets (•) and plus signs (+), respectively. Only the first 116 a.a. of Sub1 (292 a.a.) are shown. (**B**–**E**) All graphs are showing the rates of recombination ($\times 10^{-8}$). Error bars indicate 95% confidence intervals. Prior to fluctuation analysis to determine the rates of recombination, indicated yeast strains were transformed with empty vector (+*Vec*), Sub1-expression plasmid (+*SUB1*), or PC4-expression vector (+*PC4*).

DNA binding domain of Sub1 and PC4 are required for suppression of G4-mediated recombination

In order to identify which regions of Sub1 are required to prevent G4-associated recombination, we constructed a series of C-terminal truncation mutants of Sub1. As shown in Figure 1B, the complete deletion of *SUB1* resulted in an increase in G4-associated recombination in the *top1Δ* strains containing the *pTET-lys2-GTOP* construct. The rate of recombination for the *pTET-lys2-GTOP* was not affected by replacing the full length *SUB1* allele with $sub1_{1–200}$ or $sub1_{1–150}$ truncation alleles that is missing the C-terminal 92 or 142 amino acids, respectively (Figure 3A). This indicates that the coiled-coil domains in the C-terminal end of Sub1 are not required in preventing G4-mediated genome instability. The rate of recombination at the *pTET-lys2-GTOP* construct, however, was elevated to the level indistinguishable from the complete deletion mutation (*sub1*) when the C-terminal 172 or 212 amino acids were truncated ($Sub1_{1-120}$ or $Sub1_{1-80}$, respectively). Although the homology between human PC4 and yeast Sub1 is limited to amino acids a.a. 1–105 of Sub1 (Figure 2A), a conserved domain (PTHR13215) is present in yeast Sub 1 a.a. 35–146 when aligned to members of RNA polymerase II transcriptional co-activator subfamily 2 (Figure 3B) (43). Alternatively, the insufficient functional complementation of $Sub1_{1–120}$ or $Sub1_{1–80}$ could be due to the reduced Sub1 protein levels. Western analysis of the yeast strains expressing the C-terminal truncation alleles used for the determination of recombination rates in Figure 3A showed that the steady-state protein level of $Sub1_{1–120}$ or $Sub1_{1–80}$ were significantly less than that of the full length Sub1, $Sub1_{1–200}$ or $Sub1_{1–150}$ (Figure 3C). This indicates that the region between a.a. 120 and 150, which is highly enriched for charged residues, might be important for protein folding and stability.

To determine whether the ssDNA binding region alone was sufficient for resolving G4 DNA, we transformed *top1Δ sub1Δ* and *top1-Y727F sub1Δ* strains with a plasmid ex-

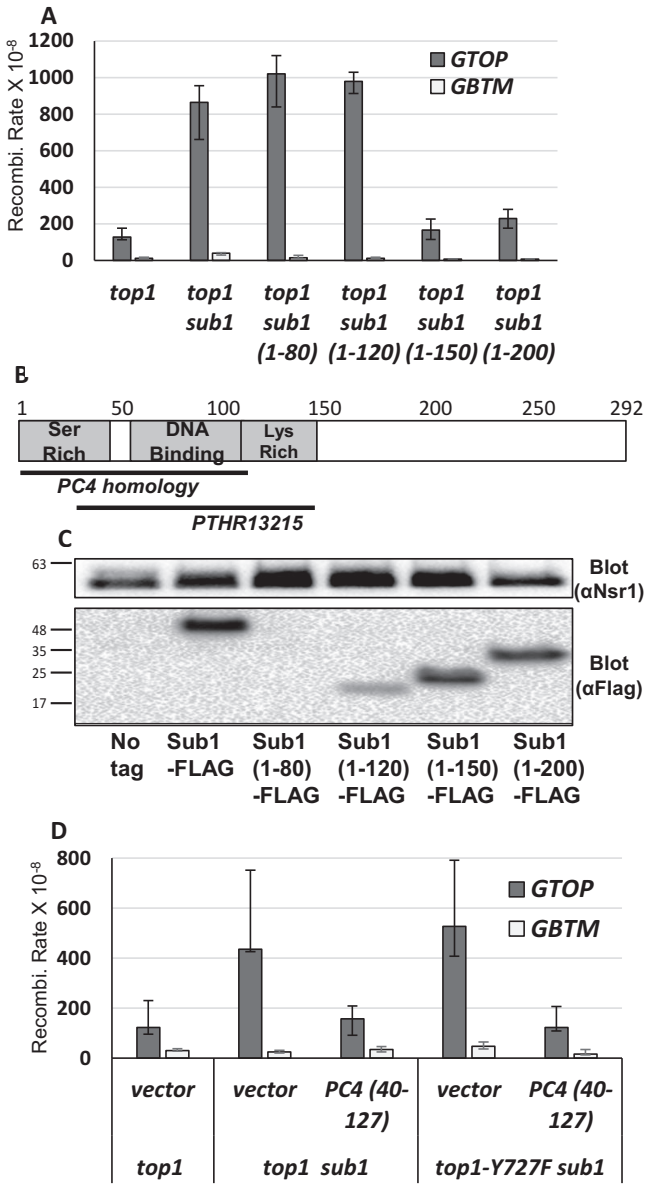

Figure 3. ssDNA binding domain of Sub1 and PC4 required for suppression of G4-associated recombination. (**A**) The graph is showing the rates of recombination ($\times 10^{-8}$). Error bars indicate 95% confidence intervals. Relevant genotypes are indicated below. (**B**) Domains of yeast Sub1. Regions with homology to human PC4 and the conserved PTHR13215 domain are indicated. (**C**) Whole cell extracts from yeast strains expressing the indicated Sub1-truncation constructs with C-terminal 3XFLAG tags were analyzed by western blotting. Nsr1 bands as loading control are shown in the top panel and Sub1 constructs detected by anti-FLAG antibody ware shown in the bottom panel. Bars to the left indicate positions of the molecular weight markers (kDa). (**D**) The graph is showing the rates of recombination ($\times 10^{-8}$). Error bars indicate 95% confidence intervals. Relevant genotypes are indicated below.

pressing PC4 truncation mutant ($PC4_{40–127}$). This mutant protein, which contains only the ssDNA binding domain of PC4 (24), sufficiently complemented the loss of Sub1 and significantly reduced the recombination rates at the *pTET-lys2-GTOP* in both strain backgrounds (Figure 3D). Consistent with its G4-specific role, the recombination rates at the *pTET-lys2-GBTM* construct were not significantly affected by the expression of $PC4_{40–127}$. Compared to the full length PC4, however, the $PC4_{40–127}$ truncation mutant resulted in more modest suppression of recombination; the rate of recombination at the *pTET-lys2-GTOP* was still significantly higher than at the *pTET-lys2-GBTM* (Figures 2C and 3D). The $PC4_{40–127}$ truncation mutant also did not further reduce the recombination rate in *top1Δ* single deletion strain (Supplementary Figure S2).

Sub1 binds the G4-forming switch region sequence *in vitro* and *in vivo*

Sub1 was previously shown to preferentially bind to a 21-mer G4-forming oligonucleotide derived from the c-MYC promoter (5′-GAGGGTGGGTAGGGTGGGTAA-3′) (27,44). In order to determine whether Sub1 also binds to G4 DNA formed by the Ig switch Mu sequence integrated into the recombination reporter constructs used in the genetic assays above (*pTET-lys2-GTOP* and *–GBTM*), we first performed circular dichroism (CD) analysis with a 45-mer oligonucleotide partially representing this sequence (SμG; Supplementary Table S1). The CD spectrum with a shallow minimum at ∼240 nm and a large maximum near 265 nm, which is characteristic of a parallel G-quadruplex structure (45), was observed for the SμG oligo in 100 mM KCl but not in the absence of salt or in the presence of 100 mM NaCl (Figure 4A). G4 DNA formation by the SμG sequence was further confirmed by its binding to the G4-specific recombinant antibody Hf2 (37), demonstrated by the change in mobility when SμG oligo was incubated with the antibody in presence of G4-stabilizing K+ cation (Figure 4B). The C-run containing complimentary sequence (SμC) did not bind Hf2 antibody in presence of K+ or Li+.

To demonstrate the direct recognition of SμG DNA by Sub1, fluorescence equilibrium binding analysis (fluorescence anisotropy assay) was performed at a final concentration of 5 nM for each DNA substrate. Fluorescence anisotropy measured with increasing concentrations of Sub1 showed that Sub1 bound SμG with a $K_D = 8.7 \pm 1.7$ nM which is ∼2.4-fold tighter than that observed for ssDNA (21.2 ± 1.1 nM) (Figure 4C). No appreciable binding of Sub1 was observed for duplex DNA. PC4 bound Sμ G with a $K_D = 6.0 \pm 0.9$ nM, which is similar to the K_D observed for ssDNA (9.0 ± 0.9 nM) (Supplementary Figure S3). PC4 bound a duplex substrate with a $K_D = 38.7 \pm 8.0$ nM which is ∼6 and ∼4-fold higher than Sμ G and ssDNA, respectively.

Having shown that Sub1 binds to the G4 forming Sμ *in vitro*, we examined whether there was Sub1-G4 DNA interaction *in vivo*. First, we constructed a *SUB1* allele with a 3XFLAG tag at the C-terminus and confirmed that the C-terminal tag does not interfere with its function in suppressing G4-associated recombination (Supplementary Figure S4). Chromatin immuno-precipitation (ChIP) was performed with antibody against the FLAG epitope starting with either *WT* or *top1Δ* yeast cells containing either the *pTET-lys2-GTOP* or *–GBTM* construct. Following the pull-down, qPCR analysis was used to determine the enrichment of Sub1-FLAG at each locus normalizing to the enrichment at the *CAN1*, which does not contain any G4

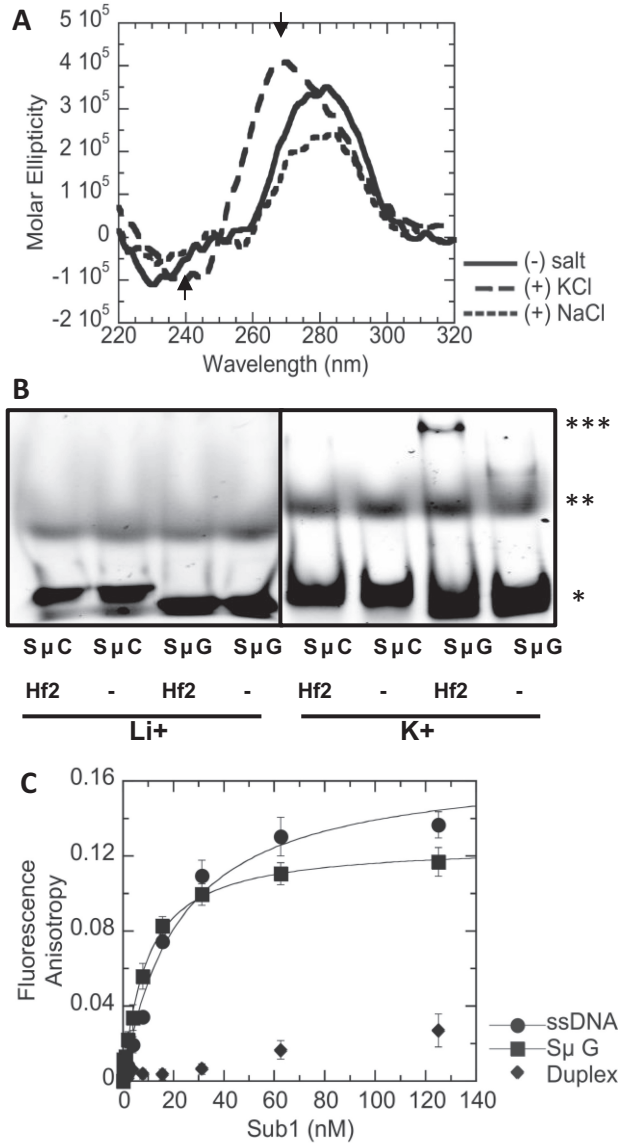

Figure 4. G4-formation by Sμ oligo and Sub1-Sμ binding *in vitro*. (**A**) CD spectra of Sμ oligo solution with no salt, 100 mM KCl or 100 mM NaCl. The maximum and minimum ellipticity at 265 nm and 240 nm for the +KCl sample are indicated with arrow heads. See Supplementary Table S1 for sequence of the oligo. (**B**) Elecrophoretic mobility shift assay. Cy5-conjugated SμC or SμG oligos incubated with Hf2 antibody (Hf2) or buffer alone (–) containing either 100 mM LiCl or 100 mM KCl. *; free oligo, **; secondary structure, ***; Hf2-bound oligo. (**C**) Fluorescence anisotropy of 5 nM ssDNA (●), Sμ G (<), or duplex (♦) with indicated concentrations of Sub1 was measured. Data represent the mean of three separate experiments. Data were fit to the quadratic equation to obtain the K_D values for SμG (8.7 ± 1.7 nM) and ssDNA (21.1 ± 1.71 nM). of See Supplementary Table S1 for sequences of the oligos.

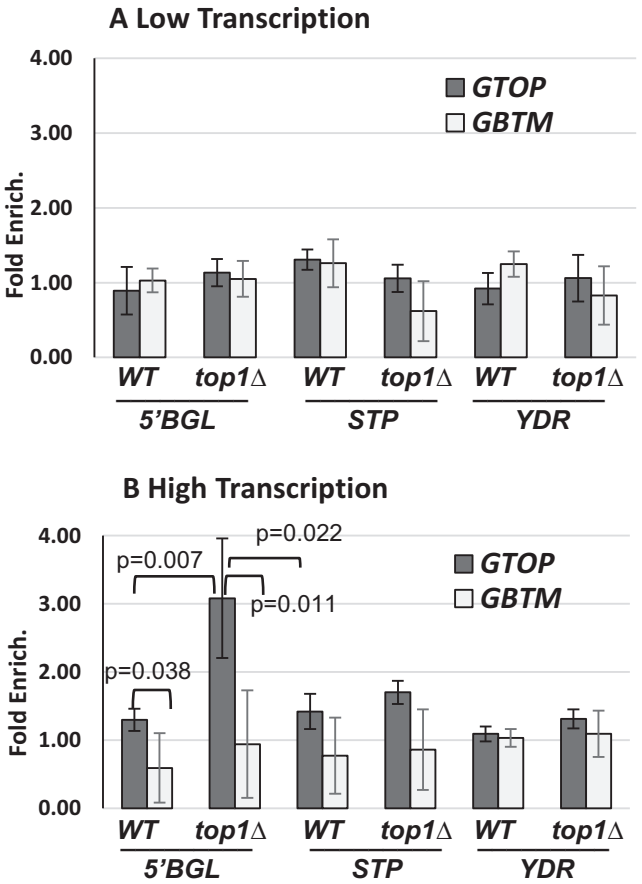

Figure 5. ChIP shows Sub1-association with G4 DNA *in vivo*. Chromatin immunoprecipitation was carried out with anti-FLAG antibody and chromatin fractions prepared from yeast cells with indicated genetic backgrounds. In all strains, 3XFLAG epitope tagged *SUB1* allele was present. 5′BGL and STP primers anneal within *LYS2* ORF, ∼100 and ∼3000 bp from the Sμ sequence insertion site, respectively. YDR primers anneal to YDR554C region of the yeast genome. All primer sequences are in Supplementary Table S1. All values are based on four independent experiments and the standard deviations are indicated by error bars. (**A**) Transcription from *pTET* was repressed by culturing cells in rich media with 2μg/ml doxycycline. (**B**) Cells were cultured in rich media with no doxycycline to ensure *pTET* is fully turned on. *P* values were calculated by Student's *t*-test.

motifs (QGRS Mapper; (46)). When cells were cultured with 2μg/ml doxycycline to repress the transcription from the *pTET* promoter, no significant enrichment of Sub1-Flag was detected at any of the loci tested (Figure 5A). Under high transcription conditions, in *top1Δ* background, Sub1-FLAG was significantly enriched (>3-fold) at the G4 insertion site (5′BGL) when the G4 sequence was in *GTOP* orientation but not in *GBTM* orientation (Figure 5B). In *top1Δ* cells under high transcription conditions, the enrichment of Sub1 was significantly higher at the switch region G4 sequence insertion site (5′BGL) than the 3′ region of the *lys2* sequence 3 kb away (STP). A slight but significant enrichment of Sub1 was detected at the G4 insertion site in *WT* background for *GTOP* but not for *GBTM* orientation under high transcription conditions. Overall, these data indicate that Sub1 is not simply being recruited to highly transcribed regions but specifically associates with the region of G4 DNA accumulated under high transcription conditions particularly in the absence of Top1. Additionally, at the *YDR544C* locus (YDR in Figure 5), where a 2-fold enrichment of Sub1 was reported previously (27), the enrichment of Sub1-FLAG was not significantly higher than the negative control locus, *CAN1*, in both WT and *top1Δ* backgrounds indicating that the co-transcriptionally formed G4

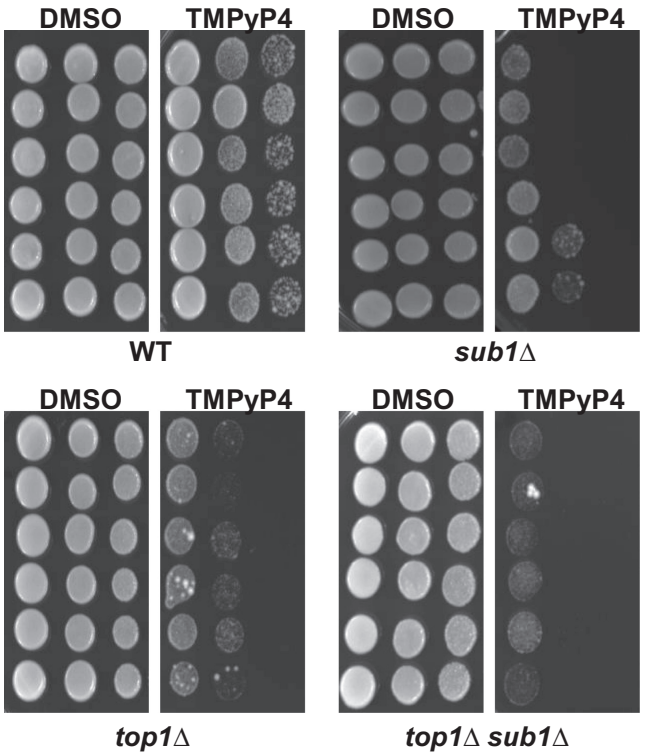

Figure 6. TMPyP4 treatment. For each indicated strains, 6 individual colonies were grown to early log phase and treated with 200 μM TMPyP4 or DMSO for 20 h. Dilutions (10-fold) were spotted on YEP media with 2% glucose and images were taken 2 days after spotting.

DNA at the *pTET-lys2-GTOP* cassette is a high-affinity interaction site for Sub1.

Sub1-disruption results in elevated sensitivity to the G4-binding ligand TMPyP4

Cationic porphyrin TMPyP4 binds G-quadruplex DNA *in vivo* and stabilizes the secondary structure disrupting cell proliferation in mammalian and yeast cells (47–49). According to the elevated recombination rate observed at the *pTET-lys2-GTOP* reporter upon the loss of Sub1 in a *top1Δ* background as well as the *in vitro* and *in vivo* binding assays, Sub1 binds to G4 DNA and consequently suppresses the instability induced by these secondary DNA structures. We tested whether Sub1 can also suppress the cytotoxicity associated with the interaction between TMPyP4 and G4 DNA. Yeast cells were treated with 200 μM TMPyP4 for 20 h. After washing, cells were spotted on rich media with no drug and monitored for recovery from the effect of TMPyP4. After 2 days on rich media, *top1Δ* cells showed greater sensitivity to TMPyP4 than *WT* cells (Figure 6). In both WT and top1Δ backgrounds, deletion of SUB1 led to enhanced TMPyP4 sensitivity indicating that Sub1 does play a protective function to alleviate the toxic effect of TMPyP4.

Sub1 physically and genetically interact with the helicase Pif1

How G4 DNA-binding by Sub1 can suppress G4-associated genome instability is not yet clear. One class of proteins that can undoubtedly promote genome stability at G4-forming sequences are G4 specific helicases such as RecQ and Pif1 family helicases (19). Since Sub1 does not contain any recognizable helicase domain, it could be acting in a mode analogous to *S. Pombe* Pot1, which can disrupt G4 structures at telomeres by binding to these sequences (50). Alternatively, Sub1 could function by recruiting other proteins such as G4-resolving helicase to G4 DNA. To test this hypothesis, we immuno-precipitated Sub1 (with a 3X FLAG tag) from yeast whole cell extracts using beads coated with anti-FLAG antibody. In the co-IP fractions, we detected the G4 DNA helicase Pif1 (tagged with HA), suggesting that Sub1 does form a complex with one such G4 helicase Pif1 (Figure 7A). Pif1-HA co-immunoprecipitated with Sub1-FLAG in both WT and *top1Δ* backgrounds. In a reciprocal co-IP experiment, IP of Pif1-HA with anti-HA coated beads resulted in the co-immunoprecipitation of Sub1-Flag. We confirmed the physical interaction between Sub1 and Pif1 by showing that purified Pif1 protein is much more efficiently captured by the beads conjugated with over-expressed, purified Sub1 protein than by the beads conjugated with BSA (Figure 7B).

To test whether Sub1 and Pif1 work in the same pathway to suppress G4-associated genome instability, we introduced the *pif1-M2* mutation, into *WT*, *top1Δ* and *top1Δ sub1Δ* strains. This mutation is located in the nuclear localization signal of Pif1 protein and specifically disrupts its nuclear function while leaving its mitochondrial function intact (34). The rates of recombination for the *pTET-lys2-GTOP* or *–GBTM* reporter was not elevated by *pif1-M2* mutation (Supplementary Figure S5). Because Pif1-M2 is not completely null for its nuclear function and showed significant residual activity in previous studies (51), we proceeded to introduce a complete null mutation of Pif1 (*pif1Δ*) into *WT*, *top1Δ, sub1Δ* and *top1Δ sub1Δ* strains. Deletion of *PIF1* did not have any effect on the rates of recombination for either of the recombination reporter construct in strains containing WT Top1 (Figure 7C). In the *top1Δ* strains, however, the rate of recombination for the *pTET-lys2-GTOP* was elevated by ∼3 fold upon deletion of PIF1 gene while the rate for the *pTET-lys2-GBTM* was not affected (Figure 7D). When *SUB1* and *PIF1* were both disrupted *(top1Δ sub1Δ pif1Δ)*, the rate of recombination for the *pTET-lys2-GTOP* in triple mutant strain was statistically indistinguishable (i.e. 95% confidence level overlaps) from the rate in *top1Δ sub1Δ* or in *top1Δ pif1Δ* strains indicating an epistatic relationship between these two factors.

DISCUSSION

Recently emerging information regarding the genetic changes in cancer genomes have identified discrete hotspots where genome arrangements, CNVs, and mutations occur more frequently than stochastic distribution (7). Often, repetitive and/or unusual sequences that can form non-canonical DNA structures such as hairpins, H-DNA, cruciform and G4 DNA overlap with these cancer-associated genome instability hotspots. Additionally, non-B DNA forming sequences are frequently linked with genome instability underlying degenerative neurological disorders such as Huntington's disease and Fredriech's ataxia. In partic-

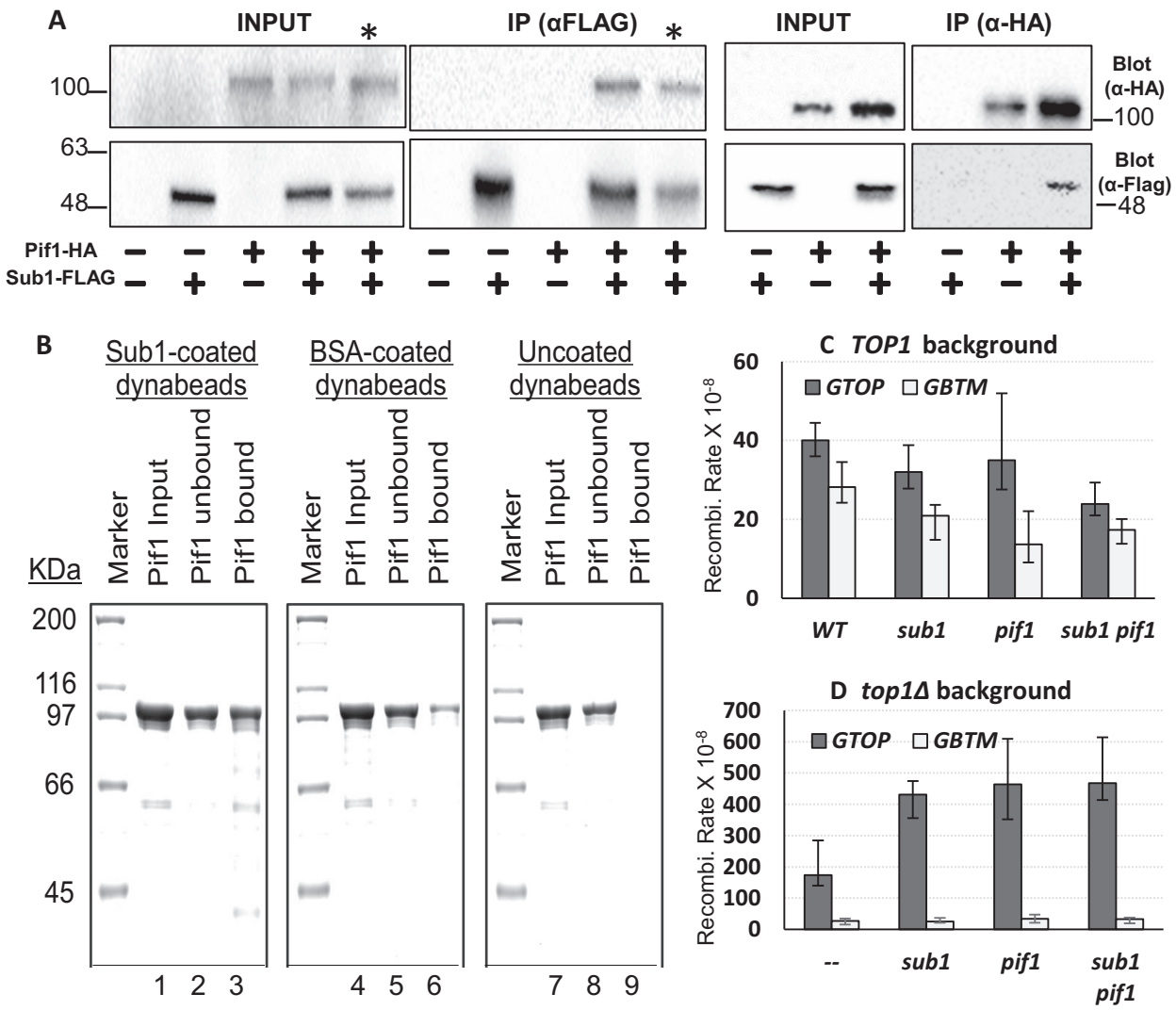

Figure 7. Sub1-Pif1 interaction. (**A**) Co-immunoprecipitation experiment. Cell lysates from yeast strains containing the indicated HA or FLAG epitope tagged protein constructs were used in immunoprecipitation with anti-FLAG or anti-HA antibody as indicated. Pif1-HA or Sub1-FLAG proteins were detected by immunoblotting with the indicated antibodies. * indicates cell lysate and co-IP sample from *top1Δ* yeast cells. All other strains were in *WT* background. Molecular weight markers (in kDa) are shown on the sides of the blots. (**B**) Interaction between purified Sub1 and Pif1. The schematic representation of the capture and elution of Pif1 by Sub1- or BSA-coated dynabeads is described in Supplementary Figure S6. The input, unbound and bound/eluted fractions from the Sub1-coated (lanes 1–3), BSA-coated (lanes 4 -6) and uncoated (lanes 7 -9) dynabeads were resolved in 10% SDS-PAGE gel. After coomassie staining, Pif1 appears as a band around 100 kDa. (**C** and **D**) All graphs are showing the rates of recombination ($\times 10^{-8}$). Error bars indicate 95% confidence intervals.

ular, G4 DNA forming sequences that consist of multiple runs of guanines have been identified at several major break points associated with recurrent translocations found in various types of hematological malignancies (6). G4 DNA is also implicated in the expansion of (GGGGCC)n present at C9orf72 gene responsible for the neurological disorder amyotrophic lateral sclerosis (ALS) (52). Non-B forming sequences located within highly transcribed regions are especially problematic because the dynamic structural transformation required for transcription to efficiently proceed can also promote the secondary structure formation of these sequence motifs. Identifying what underlying factors function to minimize the effect of the co-transcriptionally formed non-B structure on the genome integrity is important.

In this report, we focused on genome instability induced by G4 DNA that is co-transcriptionally formed from an archetypal G4 motif previously identified at the mammalian immunoglobulin heavy chain switch Mu region (Sμ). We previously described that the transient separation of the DNA strands during transcription can promote the structural shift of guanine-rich DNA strand within Sμ sequence into G4 DNA resulting in elevated genome instability. Such G4-associated genome instability is compounded by failure to remove negative supercoils (i.e. Top1-deficiency). Because the transcription-associated separation of DNA strand would render only the non-transcribed or 'top' strand as single strand while the transcribed or 'bottom' strand remains base-paired with the nascent RNA, genome instability associated with the co-transcriptionally formed

G4 DNA displayed a clear orientation-bias. The consequence of Top1-disruption was a specific elevation in recombination when the guanine runs were located on the non-transcribed strand (*pTET-lys2-GTOP*) and not on the transcribed strand (*pTET-lys2-GBTM*). We show here that loss of Sub1 elevates genome instability under the conditions that would be expected of co-transcriptionally formed G4 DNA. That is, the loss of Sub1 elevated recombination in *top1Δ* background when *pTET-lys2-GTOP* was actively transcribed (Figure 1).

The suppression of G4-associated genome instability by Sub1 is highly correlated with its physical association with G4 DNA. Sub1 is enriched at the G4 construct under the same conditions that promote G4 DNA formation – guanines on the top strand of an actively transcribed gene (*pTET-lys2-GTOP*) in Top1-deficient strain (Figure 5). We previously showed that transcription is elevated to the similar levels in *pTET-lys2-GTOP* and *pTET-lys2-GBTM* in this genetic background (35). However, Sub1 was not significantly enriched at the *pTET-lys2-GBTM* indicating that Sub1-association shown by the ChIP analysis here is not solely due to activated transcription. Sub1-enrichment is significantly reduced ∼3.5 kb away from the guanine-runs even though it is within the same transcribed unit, which further stresses the specificity of Sub1 for interaction with G4 DNA. Moreover, recombinant PC4 that contains only the DNA binding domain, but not the C-terminal truncation of Sub1 that eliminates DNA binding domain, complemented the loss of Sub1 to reduce the recombination at the actively transcribed *pTET-lys2-GTOP* (Figure 3).

Sub1 has been known to be a general transcription factor and implicated in regulation of transcription elongation and termination. Disturbance in transcription elongation process can promote annealing of the nascent RNA to template DNA strand leading to R-loop accumulation (53). Recombination elevation due to R-loop accumulation are observed in yeast cells with the disruption of mRNA packaging/export process as in mutations in THO/TREX complex subunits and with inhibition of RNA:DNA hybrid removal as in deletion of RNAse H-encoding genes. In mammalian cells, disruption of splicing factor ASF1/SF2 (54) or RNA:DNA helicase Aquarius or Senataxin (55) also lead to R-loop accumulation and genome instability. In all of these cases, overexpression of RNAse H complemented the defect and reduced the R-loop-associated recombination. In contrast, our data show that R-loop accumulation is not the pathological structure inducing genome instability in Sub1-deficient yeast cells. Disruption of Sub1 did not affect the recombination rate for the highly *pTET-lys2-GBTM* construct where the location of guanine runs in the transcribed strand disfavors G4 DNA formation (Figure 1). The rates of recombination for both the *pTET-lys2—GTOP* and *GBTM* constructs were unaffected by the disruption of Sub1 in the *RNase HΔ* backgrounds. In addition, RNase H overexpression did not reduce the highly elevated recombination for the *pTET-lys2-GTOP* construct in *top1Δ sub1Δ* strains. The suppression of G4-associated genome instability is thereby specific to G4 DNA and related to its G4 DNA binding property.

Another evidence that Sub1 does have a G4-specific function besides its role as a general transcription factor is its physical interaction with Pif1 (Figure 7). Pif1 is a DNA helicase involved in multiple important cellular pathways including lagging strand DNA replication, Break-induced Replication pathway of DSB repair, mitochondrial genome maintenance and telomere length homeostasis (56). Its role in unwinding G4 DNA and thereby suppressing G4-associated genome instability has been demonstrated several times in yeast model systems (11,14,21) and confirmed in our reporter assay (Figure 7B). Based on our data demonstrating the physical interaction and the genetic epistasis between Sub1 and Pif1, we propose that Sub1 suppresses G4-associated genome instability by facilitating the recruitment of Pif1 helicase to co-transcriptionally formed G4 DNA structures.

In summary, we found a novel function of the highly conserved and ubiquitously present protein Sub1. This protein, previously characterized as a general transcription co-activator, specifically localizes to co-transcriptionally formed G4 DNA *in vivo* and contributes in suppressing genome instability induced by these obstructive structures. Further investigation into Sub1-interacting partners in this function will likely yield useful insight into the cellular mechanism keeping these unusual sequences in check. Finally, we identified Sub1 as a substantial factor modulating TMPyP4 sensitivity in yeast cells, which underscores the significance of this protein and its mammalian homologs, especially since a variety of G4 binding ligands are under investigation for their possible efficacy in anti-neoplastic therapy (50).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Authors thank S. Balasubramanian of Cambridge University (Cambridge, UK) for the HF2 expression plasmid, M. R. Volkert of University of Massachusetts Medical School (Worcester, MA) for PC4 expression plasmids and S.H. Yoo of University of Texas Health Science Center at Houston (Houston, TX) for additional reagents.

FUNDING

National Institutes of Health (NIH) [GM080600 to G.I., GM098922 to K.D.R., GM116007 to N.K.]; Welch Foundation [AU1875 to N.K.]. Funding for open access charge: NIH [GM116007 to N.K.].
Conflict of interest statement. None declared.

REFERENCES

1. Bochman,M.L., Paeschke,K. and Zakian,V.A. (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.*, **13**, 770–780.
2. Davis,L. and Maizels,N. (2011) G4 DNA: at risk in the genome. *EMBO J.*, **30**, 3878–3879.
3. Maizels,N. and Gray,L.T. (2013) The G4 genome. *PLoS Genet.*, **9**, e1003468.
4. Gonzalez,V. and Hurley,L.H. (2010) The C-terminus of nucleolin promotes the formation of the c-MYC G-quadruplex and inhibits c-MYC promoter activity. *Biochemistry*, **49**, 9706–9714.

5. Tosoni,E., Frasson,I., Scalabrin,M., Perrone,R., Butovskaya,E., Nadai,M., Palu,G., Fabris,D. and Richter,S.N. (2015) Nucleolin stabilizes G-quadruplex structures folded by the LTR promoter and silences HIV-1 viral transcription. *Nucleic Acids Res.*, **43**, 8884–8897.
6. Katapadi,V.K., Nambiar,M. and Raghavan,S.C. (2012) Potential G-quadruplex formation at breakpoint regions of chromosomal translocations in cancer may explain their fragility. *Genomics*, **100**, 72–80.
7. Bacolla,A., Tainer,J.A., Vasquez,K.M. and Cooper,D.N. (2016) Translocation and deletion breakpoints in cancer genomes are associated with potential non-B DNA-forming sequences. *Nucleic Acids Res.*, **44**, 5673–5688.
8. Sarkies,P., Reams,C., Simpson,L.J. and Sale,J.E. (2010) Epigenetic instability due to defective replication of structured DNA. *Mol. Cell*, **40**, 703–713.
9. Walia,R. and Chaconas,G. (2013) Suggested role for G4 DNA in recombinational switching at the antigenic variation locus of the Lyme disease spirochete. *PLoS One*, **8**, e57792.
10. Cahoon,L.A. and Seifert,H.S. (2009) An alternative DNA structure is necessary for pilin antigenic variation in Neisseria gonorrhoeae. *Science*, **325**, 764–767.
11. Paeschke,K., Bochman,M.L., Garcia,P.D., Cejka,P., Friedman,K.L., Kowalczykowski,S.C. and Zakian,V.A. (2013) Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature*, **497**, 458–462.
12. Piazza,A., Boule,J.B., Lopes,J., Mingo,K., Largy,E., Teulade-Fichou,M.P. and Nicolas,A. (2010) Genetic instability triggered by G-quadruplex interacting Phen-DC compounds in Saccharomyces cerevisiae. *Nucleic Acids Res.*, **38**, 4337–4348.
13. Piazza,A., Serero,A., Boule,J.B., Legoix-Ne,P., Lopes,J. and Nicolas,A. (2012) Stimulation of gross chromosomal rearrangements by the human CEB1 and CEB25 minisatellites in Saccharomyces cerevisiae depends on G-quadruplexes or Cdc13. *PLoS Genet.*, **8**, e1003033.
14. Ribeyre,C., Lopes,J., Boule,J.B., Piazza,A., Guedin,A., Zakian,V.A., Mergny,J.L. and Nicolas,A. (2009) The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences in vivo. *PLoS Genet.*, **5**, e1000475.
15. Kim,N. and Jinks-Robertson,S. (2011) Guanine repeat-containing sequences confer transcription-dependent instability in an orientation-specific manner in yeast. *DNA Repair (Amst.)*, **10**, 953–960.
16. Yadav,P., Owiti,N. and Kim,N. (2016) The role of topoisomerase I in suppressing genome instability associated with a highly transcribed guanine-rich sequence is not restricted to preventing RNA:DNA hybrid accumulation. *Nucleic Acids Res.*, **44**, 718–729.
17. Marchand,C., Pourquier,P., Laco,G.S., Jing,N. and Pommier,Y. (2002) Interaction of human nuclear topoisomerase I with guanosine quartet-forming and guanosine-rich single-stranded DNA and RNA oligonucleotides. *J. Biol. Chem.*, **277**, 8906–8911.
18. Arimondo,P.B., Riou,J.F., Mergny,J.L., Tazi,J., Sun,J.S., Garestier,T. and Helene,C. (2000) Interaction of human DNA topoisomerase I with G-quartet structures. *Nucleic Acids Res.*, **28**, 4832–4838.
19. Fry,M. (2007) Tetraplex DNA and its interacting proteins. *Front. Biosci.*, **12**, 4336–4351.
20. Cheung,I., Schertzer,M., Rose,A. and Lansdorp,P.M. (2002) Disruption of dog-1 in Caenorhabditis elegans triggers deletions upstream of guanine-rich DNA. *Nat. Genet.*, **31**, 405–409.
21. Paeschke,K., Capra,J.A. and Zakian,V.A. (2011) DNA replication through G-quadruplex motifs is promoted by the Saccharomyces cerevisiae Pif1 DNA helicase. *Cell*, **145**, 678–691.
22. Knaus,R., Pollock,R. and Guarente,L. (1996) Yeast SUB1 is a suppressor of TFIIB mutations and has homology to the human co-activator PC4. *EMBO J.*, **15**, 1933–1940.
23. Gupta,R., Sadhale,P.P. and Vijayraghavan,U. (2015) SUB1 Plays a Negative Role during Starvation Induced Sporulation Program in Saccharomyces cerevisiae. *PLoS One*, **10**, e0132350.
24. Wang,J.Y., Sarker,A.H., Cooper,P.K. and Volkert,M.R. (2004) The single-strand DNA binding activity of human PC4 prevents mutagenesis and killing by oxidative DNA damage. *Mol. Cell. Biol.*, **24**, 6084–6093.
25. Yu,L., Ma,H., Ji,X. and Volkert,M.R. (2016) The Sub1 nuclear protein protects DNA from oxidative damage. *Mol. Cell. Biochem.*, **412**, 165–171.
26. Lada,A.G., Kliver,S.F., Dhar,A., Polev,D.E., Masharsky,A.E., Rogozin,I.B. and Pavlov,Y.I. (2015) Disruption of transcriptional coactivator Sub1 leads to genome-wide re-distribution of clustered mutations induced by APOBEC in active yeast genes. *PLoS Genet.*, **11**, e1005217.
27. Gao,J., Zybailov,B.L., Byrd,A.K., Griffin,W.C., Chib,S., Mackintosh,S.G., Tackett,A.J. and Raney,K.D. (2015) Yeast transcription co-activator Sub1 and its human homolog PC4 preferentially bind to G-quadruplex DNA. *Chem. Commun. (Camb.)*, **51**, 7242–7244.
28. Kim,J.M., Kim,K., Schmidt,T., Punj,V., Tucker,H., Rice,J.C., Ulmer,T.S. and An,W. (2015) Cooperation between SMYD3 and PC4 drives a distinct transcriptional program in cancer cells. *Nucleic Acids Res.*, **43**, 8868–8883.
29. Chen,L., Du,C., Wang,L., Yang,C., Zhang,J.R., Li,N., Li,Y., Xie,X.D. and Gao,G.D. (2014) Human positive coactivator 4 (PC4) is involved in the progression and prognosis of astrocytoma. *J. Neurol. Sci.*, **346**, 293–298.
30. Chakravarthi,B.V., Goswami,M.T., Pathi,S.S., Robinson,A.D., Cieslik,M., Chandrashekar,D.S., Agarwal,S., Siddiqui,J., Daignault,S., Carskadon,S.L. *et al.* (2016) MicroRNA-101 regulated transcriptional modulator SUB1 plays a role in prostate cancer. *Oncogene*, **35**, 6330–6340.
31. Mortusewicz,O., Evers,B. and Helleday,T. (2016) PC4 promotes genome stability and DNA repair through binding of ssDNA at DNA damage sites. *Oncogene*, **35**, 761–770.
32. Gueldener,U., Heinisch,J., Koehler,G.J., Voss,D. and Hegemann,J.H. (2002) A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res.*, **30**, e23.
33. Sikorski,R.S. and Hieter,P. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics*, **122**, 19–27.
34. Schulz,V.P. and Zakian,V.A. (1994) The saccharomyces PIF1 DNA helicase inhibits telomere elongation and de novo telomere formation. *Cell*, **76**, 145–155.
35. Yadav,P., Harcy,V., Argueso,J.L., Dominska,M., Jinks-Robertson,S. and Kim,N. (2014) Topoisomerase I plays a critical role in suppressing genome instability at a highly transcribed g-quadruplex-forming sequence. *PLoS Genet.*, **10**, e1004839.
36. Spell,R.M. and Jinks-Robertson,S. (2004) Determination of mitotic recombination rates by fluctuation analysis in Saccharomyces cerevisiae. *Methods Mol. Biol.*, **262**, 3–12.
37. Lam,E.Y., Beraldi,D., Tannahill,D. and Balasubramanian,S. (2013) G-quadruplex structures are stable and detectable in human genomic DNA. *Nat. Commun.*, **4**, 1796.
38. Werten,S., Langen,F.W., van Schaik,R., Timmers,H.T., Meisterernst,M. and van der Vliet,P.C. (1998) High-affinity DNA binding by the C-terminal domain of the transcriptional coactivator PC4 requires simultaneous interaction with two opposing unpaired strands and results in helix destabilization. *J. Mol. Biol.*, **276**, 367–377.
39. Williams,S.K., Truong,D. and Tyler,J.K. (2008) Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 9000–9005.
40. Li,J., Yu,Y., Suo,F., Sun,L.L., Zhao,D. and Du,L.L. (2014) Genome-wide screens for sensitivity to ionizing radiation identify the fission yeast nonhomologous end joining factor Xrc4. *G3 (Bethesda)*, **4**, 1297–1306.
41. Ramanagoudr-Bhojappa,R., Chib,S., Byrd,A.K., Aarattuthodiyil,S., Pandey,M., Patel,S.S. and Raney,K.D. (2013) Yeast Pif1 helicase exhibits a one-base-pair stepping mechanism for unwinding duplex DNA. *J. Biol. Chem.*, **288**, 16185–16195.
42. Lippert,M.J., Kim,N., Cho,J.E., Larson,R.P., Schoenly,N.E., O'Shea,S.H. and Jinks-Robertson,S. (2011) Role for topoisomerase 1 in transcription-associated mutagenesis in yeast. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 698–703.
43. Mi,H., Poudel,S., Muruganujan,A., Casagrande,J.T. and Thomas,P.D. (2016) PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic Acids Res.*, **44**, D336–D342.
44. Mathad,R.I., Hatzakis,E., Dai,J. and Yang,D. (2011) c-MYC promoter G-quadruplex formed at the 5′-end of NHE III1 element:

insights into biological relevance and parallel-stranded G-quadruplex stability. *Nucleic Acids Res.*, **39**, 9023–9033.
45. Paramasivan,S., Rujan,I. and Bolton,P.H. (2007) Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding. *Methods*, **43**, 324–331.
46. Kikin,O., D'Antonio,L. and Bagga,P.S. (2006) QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.*, **34**, W676–W682.
47. Le,D.D., Di Antonio,M., Chan,L.K. and Balasubramanian,S. (2015) G-quadruplex ligands exhibit differential G-tetrad selectivity. *Chem. Commun. (Camb.)*, **51**, 8048–8050.
48. Zhuang,X.Y. and Yao,Y.G. (2013) Mitochondrial dysfunction and nuclear-mitochondrial shuttling of TERT are involved in cell proliferation arrest induced by G-quadruplex ligands. *FEBS Lett.*, **587**, 1656–1662.
49. Andrew,E.J., Merchan,S., Lawless,C., Banks,A.P., Wilkinson,D.J. and Lydall,D. (2013) Pentose phosphate pathway function affects tolerance to the G-quadruplex binder TMPyP4. *PLoS One*, **8**, e66242.
50. Patel,D.J., Phan,A.T. and Kuryavyi,V. (2007) Human telomere, oncogenic promoter and 5′-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.*, **35**, 7429–7455.
51. Stundon,J.L. and Zakian,V.A. (2015) Identification of Saccharomyces cerevisiae genes whose deletion causes synthetic effects in cells with reduced levels of the nuclear Pif1 DNA helicase. *G3 (Bethesda)*, **5**, 2913–2918.
52. Haeusler,A.R., Donnelly,C.J., Periz,G., Simko,E.A., Shaw,P.G., Kim,M.S., Maragakis,N.J., Troncoso,J.C., Pandey,A., Sattler,R. *et al.* (2014) C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature*, **507**, 195–200.
53. Hamperl,S. and Cimprich,K.A. (2014) The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair (Amst.)*, **19**, 84–94.
54. Li,X. and Manley,J.L. (2005) Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell*, **122**, 365–378.
55. Sollier,J., Stork,C.T., Garcia-Rubio,M.L., Paulsen,R.D., Aguilera,A. and Cimprich,K.A. (2014) Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol. Cell*, **56**, 777–785.
56. Boule,J.B. and Zakian,V.A. (2006) Roles of Pif1-like helicases in the maintenance of genomic stability. *Nucleic Acids Res.*, **34**, 4147–4153.