PSIP1/LEDGF reduces R-loops at transcription sites to maintain genome integrity

Sundarraj Jayakumar[1 2], Manthan Patel[1], Hadicha Aziz[1], Greg Brooke[3], Hemanth Tummala[1] & Madapura M Pradeepa[1*]

1 - Blizard Institute; Faculty of Medicine and Dentistry, Queen Mary University of London, London, UK
2 - Bhabha Atomic Research Centre, Mumbai, India
3 - School of Life Sciences, University of Essex, Colchester, United Kingdom;

[*] Corresponding author Madapura M Pradeepa: p.m.madapura@qmul.ac.uk

Abstract

Accumulation of R-loops, which are transcriptional by-products, poses a persistent threat to genome integrity and are implicated in accelerated ageing and cancer. PSIP1/LEDGF is a chromatin protein, shown to have histone chaperone activity and interact with transcriptional elongation complex. Here, we demonstrated that PSIP1 interacts with R-loops and with the factors involved in R-loop resolution, including PARP1. Reduced PSIP1 levels in human and mouse cells increase R-loop and DNA damage levels at the PSIP1 binding sites. This increased R-loops and DNA damage was specific at the RNA polymerase II transcription sites, causing local transcriptional arrest and reduced replication rate. In addition, we demonstrate that targeting PSIP1 increases the sensitivity of cancer cells to PARP1 inhibitors and transcription-coupled DNA-damaging agents. Overall, our findings reveal the critical role of PSIP1 in maintaining genome stability by reducing the R-loop burden and DNA damage at the transcriptionally active regions of the genome.

Introduction

During transcription, the newly synthesised nascent RNA can bind to the template DNA forming an RNA-DNA hybrid. This hybrid structure combined with the displaced single-stranded DNA is termed R-loops (Petermann et al., 2022). Certain physiological R-loops are generated through a programmed process with the help of specific factors. They play an essential regulatory role in transcription initiation, termination, DNA replication, DNA repair and many other cellular processes (García-Muse and Aguilera, 2019). On the other hand, R-loops are also formed co-transcriptionally, in an un-scheduled fashion, that are referred to as pathological R-loops. The accumulation of these unscheduled R-loops leads to RNA polymerase-II (RNAP2) pausing and transcriptional arrest (Chen et al., 2018). Unscheduled R-loops also cause collision between stalled transcription and replication forks, leading to DNA damage and increased genomic instability (Hamperl et al., 2017). Hence cells are equipped with different mechanisms to prevent and resolve these un-scheduled R-loops, and their dysregulation can potentially cause devastating diseases (Perego et al., 2019; Salvi and Mekhail, 2015). RNA processing and export factors, nucleases (RNAseH1 and RNAseH2) and helicases (SETX, DDX1, DDX19 etc.) reduce RNA-DNA hybrid levels (García-Muse and Aguilera, 2019).

PC4 and SF2 interacting protein (PSIP), also known as LEDGF, is a multifunctional chromatin protein that binds to methylated histone H3 lysine 36 (H3K36me) via the PWWP domain, and PSIP1 binding is enriched at the sites of RNAP2 transcription (Van Nuland et al., 2013; Pradeepa et al., 2012). *PSIP1* encodes two protein isoforms – a shorter PSIP/p52 and a longer PSIP/p75 - generated by alternative splicing (Fig. 1A). PSIP1 aids in transcription elongation by functioning as a histone chaperone in the absence of the FACT complex (LeRoy et al., 2019), and through association with the TFIIS N-terminal domain (Cermakova et al., 2021). PSIP1 binds to RNA processing proteins and modulates alternative splicing (Pradeepa et al., 2012). PSIP1 promotes cancer cell proliferation Deepak(Bhargavan et al., 2012; French et al., 2016; Singh deepak, 2017) and chemotherapy resistance (Basu et al., 2016; French et al., 2016; Ríos-Colón et al., 2017).

Here we discovered that PSIP1 interact with R-loops and factors involved in R-loop homeostasis. PSIP1 reduces R-loop levels at the site of RNAP2 transcription to maintain genome stability. Furthermore, we show that PSIP promotes the repair of DNA damage induced by R-loops by homologous recombination (HR).

Results

PSIP1 interacts with R-loops, and PSIP1 depletion elevates the R-loop level

Proteomics data from immunoprecipitation (IP) of endogenous PSIP1/LEDGF (Fig. 1A) identified several proteins involved in transcription, RNA processing and DNA repair (Pradeepa et al., 2012, 2014; Sundarraj et al., 2021). We found that 46 proteins enriched in the PSIP1 IP overlap with the R-loop interactome described previously (Cristini et al., 2018) (Fig. 1B). These overlapping proteins include many splicing factors, DNA-RNA helicases, and DNA repair factors. We, therefore, sought to investigate the role of PSIP1 in R-loop homeostasis. Firstly, by performing IP with the S9.6 monoclonal antibody that detects RNA-DNA hybrids (R-loops) (Boguslawski et al., 1986), we confirmed the interaction of PSIP1/p75 with the R-loop complex (Fig. 1C). This interaction of PSIP1/p75 with R-loops is sensitive to RNaseH1, showing the specificity of the PSIP1/p75 interaction with R-loop complex. Since PSIP1 interacts with RNAP2 transcriptional elongation machinery (Cermakova et al., 2021) and is associated with transcriptionally active chromatin (Pradeepa et al., 2012, 2014), we hypothesised that PSIP1 could prevent R-loop accumulation or promote the resolution of unscheduled R-loops at the sites of RNAP2 transcription. To test this, we quantified bulk R-loop levels using immunoblotting of genomic DNA using the S9.6 antibody. We detected a stark increase in R-loop levels in PSIP1 knockout mouse embryonic fibroblasts (*Psip1*$^{-/-}$ MEFs) compared to the wild type (Fig. S1A and Fig. 1D). This was rescued upon PSIP1/p75 re-expression, implying the direct role of PSIP1 in reducing R-loop level (Fig. 1D). Elevated R-loop signal is sensitive to RNaseH confirming the specificity of the R-loop detection (Fig. 1D).

We generated PSIP1 knockdown (PSIP1-KD) using two independent short-hairpin (sh) RNAs' that target both isoforms of PSIP1 in normal prostate epithelial cells (RWPE-1) and HEK293T cell lines (Fig. 1E and Fig. S1 and 2). PSIP1-KD cells showed a nearly two-fold increase in R-loop intensity compared to the control (Fig. 1F and S1B). Furthermore, immunofluorescence analysis followed by a high content microscopy analysis revealed a significantly higher level of R-loops in the PSIP1-KD nuclei compared to the control (Fig. 1G and S1C–S1G).

Next, we aimed to map genome-wide occupancy of PSIP1 and R-loops in control and PSIP1-KD RWPE-1 cells by performing Cleavage Under Targets and Tagmentation (CUT&Tag) using the S9.6 antibody (Wang et al., 2021). PSIP1/p75 occupancy was significantly reduced in PSIP1-KD cells, confirming the specificity of the PSIP1/p75 CUT&Tag data (Fig. 1H). Consistent with the slot blot and immunofluorescence data, we

found a nearly five-fold increase in R-loop peaks (n=2741) in PSIP1-KD RWPE-1 cells compared to the control (n=545) (Fig. 1I). Notably, the regions that gained R-loops upon PSIP1 depletion are occupied by PSIP1/p75 in control but not in PSIP1-KD (Fig. 1I and 1J).

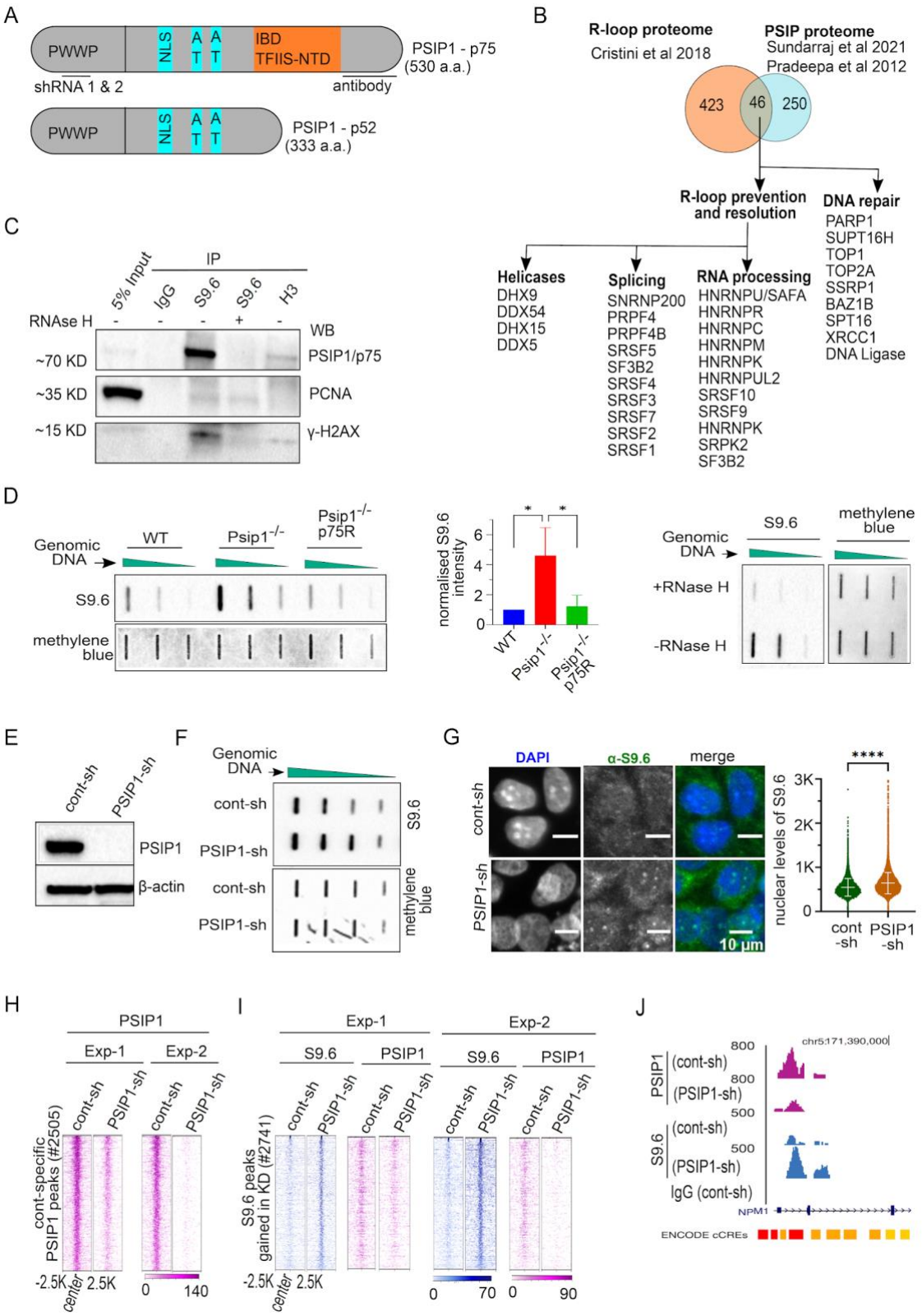

Figure 1. PSIP1 depletion leads to the accumulation of R-loops.

A) Schematic showing PSIP domains, shRNA targeting region and antibody epitope used in the study. B) The list of proteins that overlap between the R-loop and the PSIP1 interactors is shown, along with their functions related to R-loop resolution and DNA repair. C) Western blotting of S9.6, H3 and IgG immunoprecipitated (IP) from RWPE-1 cell nuclear extract with PSIP1/p75, γ-H2AX and PCNA (negative control) antibodies. Lysate treated with RNase-H before the pulldown was used as input control. D) Slot blot for R-loop using S9.6 antibody(S9.6) for genomic DNA extracted from wildtype (WT), *Psip1* knock out (*Psip1*$^{-/-}$) and Psip1$^{-/-}$ after rescue with P75 (*Psip1*$^{-/-}$ p75R) MEFs. The same membrane was treated with methylene blue for the loading control. The relative normalised band intensity was quantified (middle). Genomic DNA from *Psip1*$^{-/-}$ MEFs treated with RNaseH served as a negative control(right). E) Western blotting with PSIP1/p75 and B-actin for lysate from PSIP1-shRNA and non-targeting control shRNA (control-sh). F) Like D, but for PSIP1 knockdown (PSIP1-sh) and control knockdown (control-sh) RWPE-1 cells. G) S9.6 immunofluorescence in RWPE-1 cells (left), and quantification of the S9.6 ab signal from 5000 nuclei (right). ($p<0.0001$, unpaired t-test). H) Heatmaps showing E. coli spike-in normalised CUT&Tag reads for PSIP1/p75 in control and PSIP1 knockdown RWPE-1 cells. I) Heatmap showing R-loop and PSIP1/p75 levels in PSIP1-KD and control. J) Representative genome browser track (hg38) showing the CUT&Tag signal (read counts) for PSIP1 and R-loop in control and PSIP1-KD RWPE-1 cells.

PSIP1 depletion leads to increased DNA damage at R-loop sites.

Accumulation of unscheduled R-loops is known to cause DNA replication stress and activate the DNA damage response pathway (Hamperl et al., 2017). We observed a significant increase in bulk γ-H2AX level and a number of foci in *Psip1*$^{-/-}$ compared to wild-type MEFs, indicating the role of PSIP1 in reducing DNA damage (Fig. 2A and B). This increase in DNA damage in the *Psip1*$^{-/-}$ could be rescued by expressing Psip1/p75 cDNA (Fig. 2A and B), confirming the direct role of PSIP1 in reducing R-loops and accumulation of DNA damage.

Consistently, we also found an elevated level of γ-H2AX levels in PSIP1-KD RWPE-1 (Fig. 2C and D) and HEK293T cells (Fig. S2A and B) compared to the control, confirming the role of PSIP1 in reducing DNA damage. CUT&Tag with γ-H2AX antibodies revealed a marked increase in γ-H2AX peaks upon depletion of PSIP1 (9137 peaks in PSIP1-KD compared to 2505 in control) (Fig. 2E). Notably, accumulated γ-H2AX peaks in PSIP1-KD are at the PSIP1 binding sites (Fig. 2E), confirming that γ-H2AX accumulation is due to the direct consequence of PSIP1 depletion. CUT&Tag data is consistent with the increased bulk γ-H2AX level and a number of γ-H2AX foci in the PSIP1-KD (Fig. 2A-D). γ-H2AX peaks gained in PSIP1-KD also show an increased R-loop signal (Fig. 2F and G), suggesting that the R-loops that accumulate upon PSIP1 depletion are responsible for increased DNA damage in PSIP KD. RNase H1 overexpression and inhibition of RNAP2 transcription by triptolide rescued the elevated level of γ-H2AX. Similarly, a high level of 53BP1, a marker of non-homologous end joining (NHEJ) repair, was also rescued (Fig. 2H and S2C). These data demonstrate that the

elevated DNA damage in the absence of PSIP1 is due to R-loop accumulation and is RNAP2 transcription-dependent.

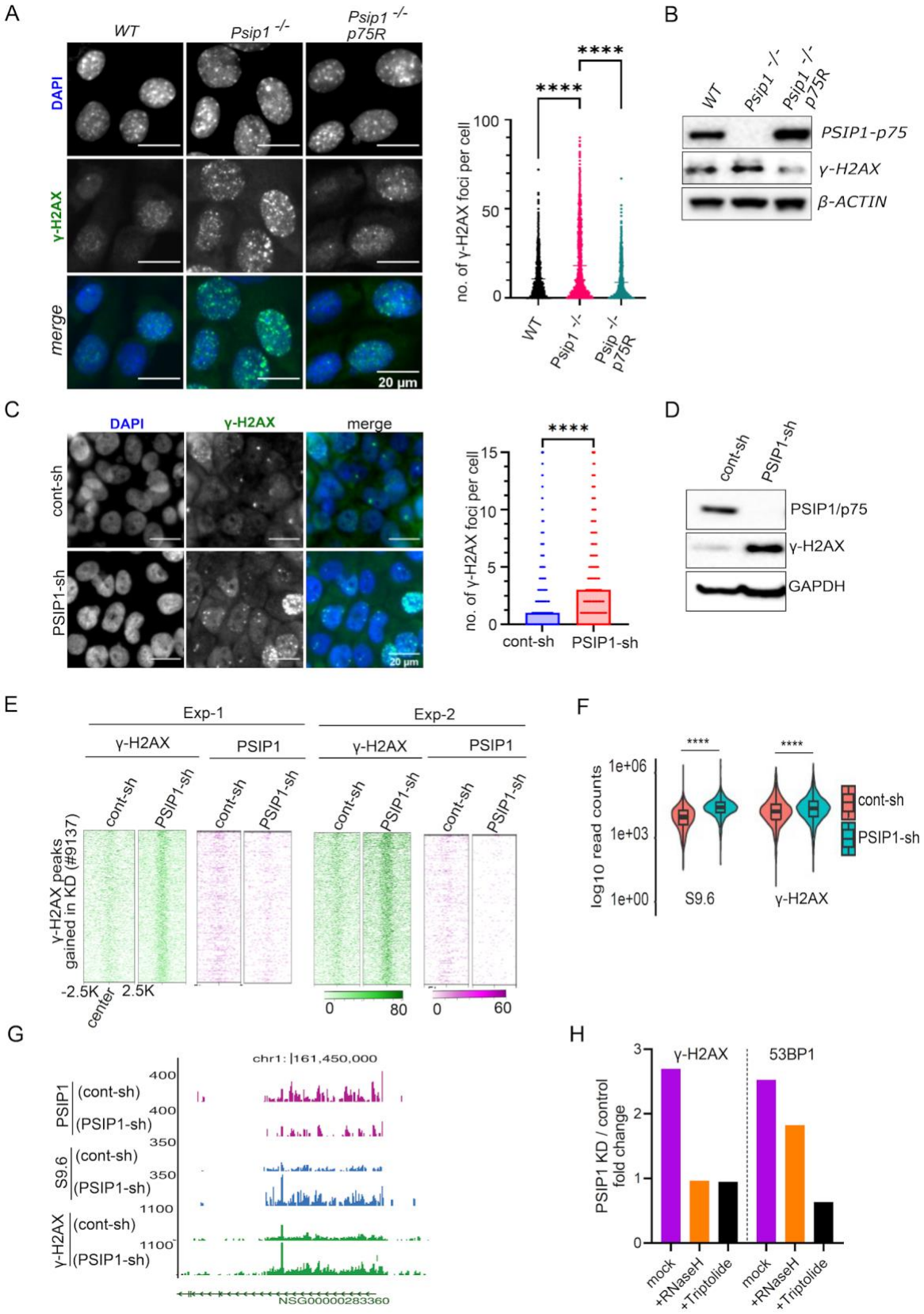

Figure 2. PSIP1 depletion leads to R-loop-mediated DNA damage.

A) Representative immunofluorescence images of γ-H2AX foci (green) in wild-type, *Psip1*$^{-/-}$ and *Psip1*$^{-/-}$ p75R MEFs. The number of γ-H2AX foci per cell is plotted. **** $p < 0.0001$ in one-way ANOVA followed by Tukey's multiple comparison test. B) Western blot for γ-H2AX levels in PSIP1 knock out (*Psip1*$^{-/-}$ MEFs) and upon the rescue with Psip1/p75 (*Psip1*$^{-/-}$ p75R). C) Like A but for PSIP1-KD and control-KD RWPE-1 cells, (n=5000) **** $P<0.0001$ in unpaired t-test. D) Western blot showing the levels of γ-H2AX in RWPE-1 cells. E) Heatmap showing E. coli normalised CUT&Tag reads for γ-H2AX and PSIP1/p75 for γ-H2AX peaks gained in PSIP1-KD F) Violin plot showing E. coli normalised CUT&Tag (log10 CPM) for S9.6 and γ-H2AX antibodies at S9.6 peaks gained in KD. *** $p<0.001$ in Dunn test with Bonferroni correction. G) Genome-browser tracks showing CUT&Tag data for PSIP/p75, R-loops (S9.6 ab) and γ-H2AX in control and PSIP-KD RWPE1 cells. H) Bar plot showing the quantification of western blot signals for γ-H2AX and 53BP1 in the PSIP1-KD cells after expressing RNaseH or treating with triptolide (blots are in Fig S3C).

Elevated R-loops are enriched at strong CpG promoters

R-loop peaks gained, as a result of PSIP1-KD, were enriched around promoter or transcription start sites (TSS) and gene bodies compared to intergenic regions (Fig. 3A). PSIP1/p75 binding is enriched around gene promoters, PSIP1 depletion led to an increase in R-loop levels and DNA damage (γ-H2AX level) at PSIP1 binding sites (Fig. 3B). Strong CpG promoters with more guanine (G) than cytosine (C) in the same strand (high GC skew) are prone to R-loop formation (Ginno et al., 2012). R-loops detected in PSIP1-KD have higher GC skew (Fig. 3C and S3), suggesting that R-loops are enriched at promoters or genes regulated by strong CpG islands and have high transcriptional activity.

R-loops accumulation is transcription-dependent, and elevated R-loops lead to local transcriptional arrest.

R-loops can induce transcriptional termination or arrest (Skourti-Stathaki et al., 2019; Zatreanu et al., 2019a). To investigate whether R-loop accumulation upon PSIP1 depletion leads to local transcriptional arrest at the sites of R-loop accumulation, we performed transient transcriptome sequencing (TT-seq) by labelling newly synthesised transcripts with 4SU (Schwalb et al., 2016). TT-seq data revealed a stark reduction in the nascent transcript levels at the site of R-loop peaks detected explicitly in the PSIP1-KD (Fig. 3D). This suggests the role of PSIP1 in reducing R-loops at the site of transcription, which can otherwise lead to transcriptional arrest.

Triptolide-mediated transcription inhibition reversed the elevated R-loop levels in PSIP1-KD (Fig. 3E). This confirms the role of PSIP1 in reducing the R-loop burden during ongoing transcription. The R-loop-induced transcriptional arrest can cause collisions with the DNA replication forks resulting in a slower replication rate and DNA damage. Reduced incorporation of 5-ethynyl-2'-deoxyuridine (EdU) into DNA in PSIP1-KD implies a slow replication rate (Fig. 3F). These results reveal that PSIP1 reduces transcription induced R-loops, which can otherwise lead to local transcriptional arrest and collision with replication

fork. Minimal changes in the expression of 180 genes known to be involved in R-loop homeostasis (Fig. S3A) suggest that elevated R-loop level is not due to the indirect effect of the altered transcriptional programme.

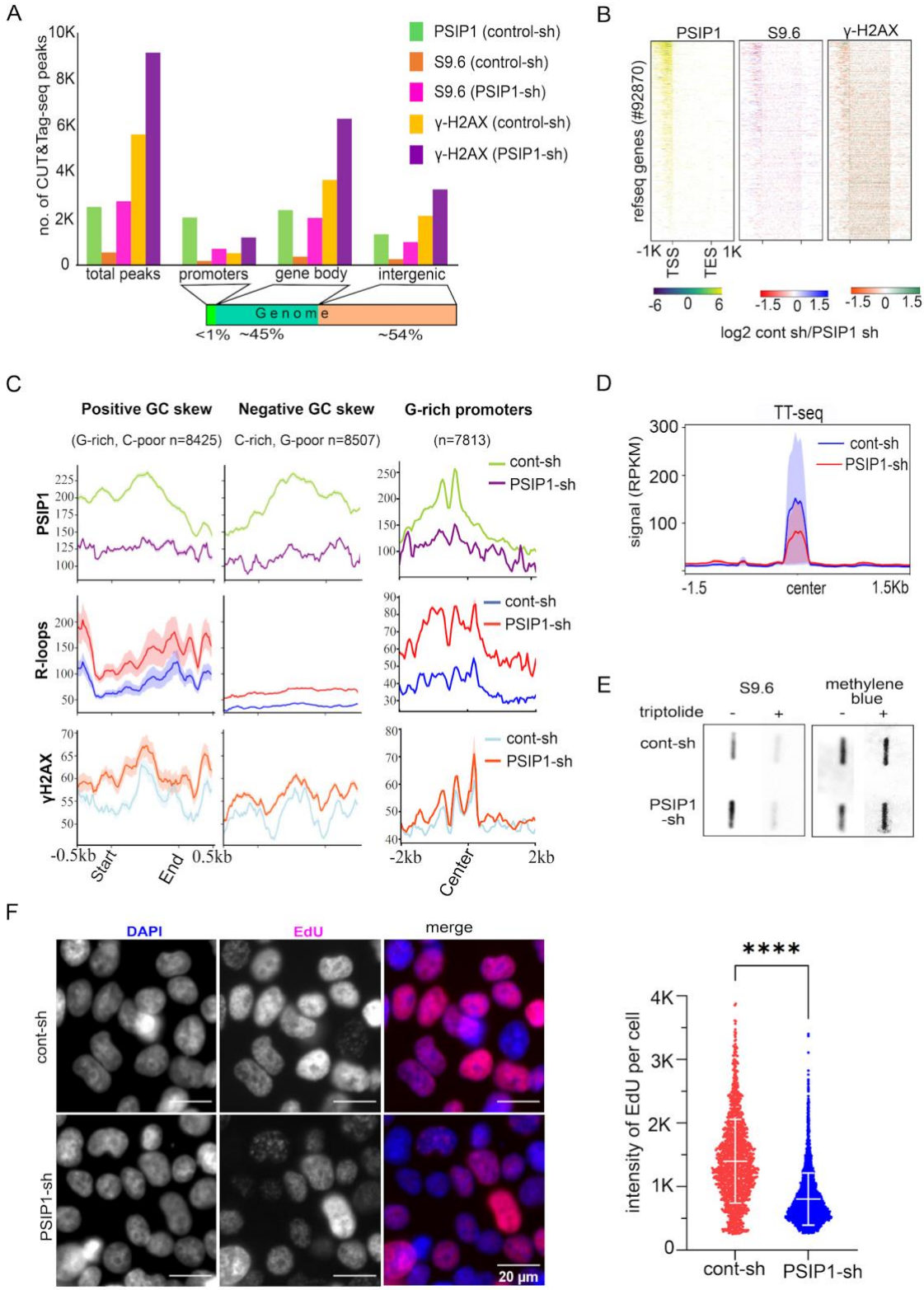

Figure 3. R-loop accumulation in PSIP KD is transcription-dependent.

A) Distribution of PSIP1, R-loop (S9.6 ab), and γ-H2AX CUT&Tag peaks from RWPE-1 cells around the gene transcription start sites (promoters), gene bodies and intergenic regions. B) Heatmaps showing E. coli normalised CUT&Tag reads for PSIP1/p75, R-loops (S9.6 ab) and γ-H2AX in RWPE-1 cells across the NCBI reference genes. The log2 fold change in read counts between control and PSIP1-KD. C) CUT&Tag data showing the enrichment of PSIP1/p75 (top), R-loops (middle) and γ-H2AX (bottom) across positive GC skew, negative GC skew regions along with G rich promoters, in control (cont-sh) and PSIP1 knockdown (PSIP1-sh) RWPE1 cells. D) TT-seq data around the centre of the S9.6 peaks gained in PSIP KD. E) Slot blot using S9.6 antibody in control and PSIP-KD RWPE cells after treatment with transcriptional inhibitor (triptolide), methylene blue staining of the same DNA served as a loading control. F) Microscopic image showing the incorporation of EdU in PSIP1-KD and control cells. Fluorescence intensity per cell for 5000 cells (right) **** P<0.0001 in unpaired t-test.

PSIP1 promotes the repair of transcription and R-loop-induced DNA damage

PSIP1/p75 binds to regions that show elevated R-loops and γ-H2AX levels in PSIP1-KD (Fig. 4A and Fig. 2D-F). Although our data show that increased DNA damage upon PSIP1 depletion is due to elevated R-loops. We cannot rule out the possibility of reduced DNA repair due to PSIP1 depletion, as PSIP1 is implicated in HR, and we show it interacts with many DNA repair factors (Fig. 1B). The DNA damage caused by R-loops is repaired by transcription-coupled double-strand break repair (Cohen et al., 2022). To investigate the role of PSIP1 in the repair of DNA damage at the site of transcription, we treated the cells with illudin-S and etoposide, which are known to cause transcription-coupled DNA damage. PSIP1 depletion led to increased sensitivity to illudin-S mediated DNA lesions that are specifically repaired by the transcription-coupled nucleotide excision repair (TC-NER) pathway but not global nonhomologous end joining (NHEJ) and base excision repair (BER) (Jaspers et al., 2002). Similarly, PSIP1-depletion increased sensitivity to etoposide, which is known to bind topoisomerase II and block transcription and replication (Montecucco et al., 2015) (Fig. 4B). Phleomycin treatment shows similar sensitivity, which induces DNA double strands breaks but not with aphidicolin, that causes cell cycle arrest (Krokan et al., 1981; Reiter et al., 1972) (Fig. S4A and Fig. 4B). This specific sensitivity to clastogens, that interfere with the transcription process, reveals the role of PSIP1 in promoting DNA repair at the site of transcription. Further, western blotting for DNA damage markers such as γ-H2AX and DNA-PK revealed a reduction in DNA repair and a sustained DNA damage response in PSIP1-KD compared to the control (Fig. 4C). Enrichment of PSIP1/p75 in γ-H2AX IP revealed that PSIP1 is recruited at DNA damage sites (Fig. 4D). PSIP1 recruitment to γ-H2AX sites is further increased when treated with camptothecin (CPT), which induces R-loop formation (Fig. 4D) by inhibiting topoisomerase 1 at active promoters (Liu et al., 2000). However, PSIP1 recruitment to γ-H2AX

chromatin was reduced upon overexpression of RNaseH1 (Fig. 4D). Co-localization of PSIP1 with γ-H2AX further confirmed the recruitment of PSIP1 at the DNA damage sites (Fig. S4B). We found an increased level of DNA-PK in PSIP1-KD cells (Fig. 4C), which is involved in the NHEJ pathway. Consistently, 53BP1 foci were significantly higher in PSIP1-KD and upon CPT treatment compared to the control (Fig. 4E). Similar increase in 53BP1 foci was also observed after phleomycin treatment in PSIP1 knockdown cells (Fig. S4C). Whereas the number of RAD51 foci, a marker of the HR pathway, was significantly reduced in PSIP1-KD and upon CPT treatment (Fig. 4F), consistent with the role of PSIP1 in promoting the HR repair pathway (Aymard et al., 2014; Daugaard et al., 2012). These results suggest that PSIP1 facilitates the repair of R-loop mediated DNA damage through the HR pathway at the transcription sites.

Reduced PSIP1 levels sensitise cancer cells to clastogens and PARP1 inhibitor.

PARP1 plays a crucial role in repairing transcription-coupled single- and double-strand breaks (Zhao et al., 2022) and has been shown to play an essential role in R-loop prevention and resolution (Cristini et al., 2018; Lin et al., 2022; Ye et al., 2021). PARP1 was one of the significant interactors with PSIP1/p75 and was also enriched in the R-loop proteome (Cristini et al., 2018; Sundarraj et al., 2021) (Fig. 1B). We confirmed the interaction of PARP1 and PSIP1 with the R-loops by immunoblotting S9.6 ab IP (Fig. 5A). Reciprocal α-PSIP1-IP performed in the presence of universal nuclease (Fig. 5B), together with co-IP of HA-tagged PSIP1/p75 and PSIP1p52 with GFP-tagged PARP1 confirmed the interaction between both PSIP1 isoforms with PARP1, along with γ-H2AX (Fig. 5C).

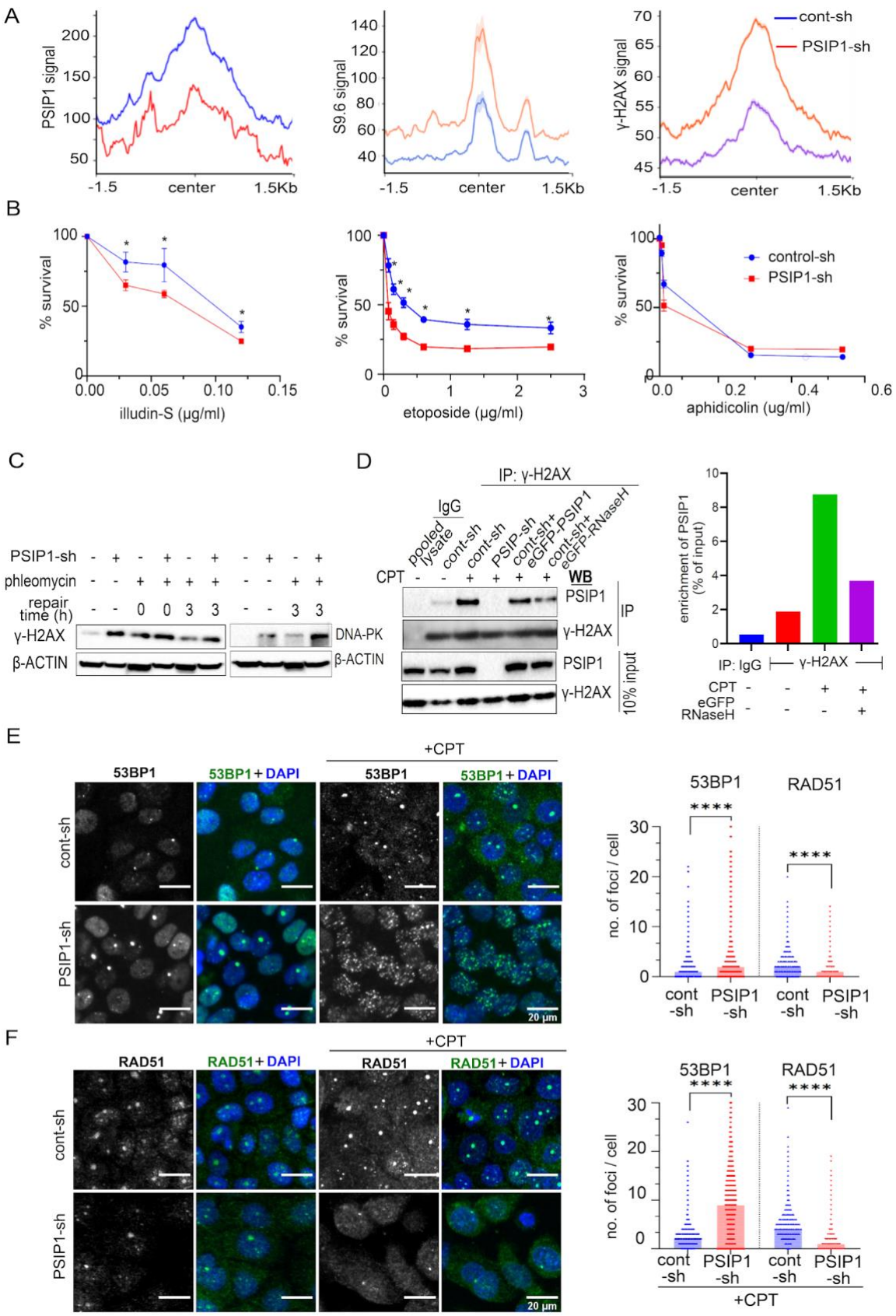

Figure 4. PSIP1 role in the repair of R-loop induced DNA damage.
A) Average profiles for PSIP1/p75, R-loop (S9.6 ab), and γ-H2AX CUT&Tag reads (E.coli normalised) at the γ-H2AX peaks gained in PSIP1-KD RWPE-1 cells. B) Average % survival

of PSIP-KD and control-KD RWPE-1 cells treated with different concentrations of illudin-S, etoposide and aphidicolin drugs. * $p<0.05$ in multiple unpaired t-tests. C) Western blot for γ-H2AX and DNA-PK (right) in control KD and PSIP1-KD cells in cells treated with phleomycin followed by repair for the indicated time. B-ACTIN served as a loading control. D) Western blots for PSIP1/p75 and γ-H2AX, for γ-H2AX IPed HEK293T extracts in control shRNA, PSIP1 shRNA, in the presence of camptothecin (CPT) (+) and DMSO (–). γ-H2AX IP was also performed upon transfection with GFP-control and GFP-RNaseH1 plasmids. IgG is a negative control, and input extracts are blotted (below). The band intensity was quantified, and the PSIP1 enrichment was plotted (right). E) Representative immunofluorescence images from Incell 2200 high-content microscopy showing 53BP1 foci in control and CPT-treated RWPE-1 cells. The average number of foci per cell is shown for 5000 cells per group, *** $p<0.001$ in one-way ANOVA followed by Tukey's multiple comparison test. F) Like (E), but for RAD51 antibody.

Comparison of the PSIP1 transcripts level across different cancer types in the TCGA database, revealed lower PSIP1 levels compared to the respective control tissues (Fig. 5D). PSIP1 levels in high-grade prostate cancer are similar to those found in normal prostate. Cells with elevated levels of R-loops are sensitive to PARP1 inhibitor-induced cell death (Chabanon et al., 2021). Hence, we hypothesised that cancers with low PSIP1 levels are sensitive to PARP inhibitors or agents that cause transcription-coupled DNA damage, such as illudin-S, etoposide or camptothecin, could be an effective synthetic lethal approach for treating cancers. Androgen-sensitive (LNCaP) and castration-resistant (PC3) prostate cancer cell lines showed lower levels of PSIP1 compared to normal prostate epithelial cells (RWPE-1) (Fig. 5E). Importantly, these prostate cancer cells showed significantly higher sensitivity to PARP1 inhibitor – olaparib compared to the normal cells (Fig. 5F). Further depletion of PSIP1 levels led to hypersensitivity of prostate cancer cells to illudin-S (Fig. 5G–I). PSIP1 depletion also sensitised LNCaP cells to olaparib (Fig. 5G and 5I) or a combination of olaparib with etoposide (Fig. 5J). Altogether, these results indicate the possibility of targeting cancers that have a low level of PSIP1 with PARP1 inhibitors or DNA-damaging clastogens that interfere with the transcription-coupled DNA repair mechanism.

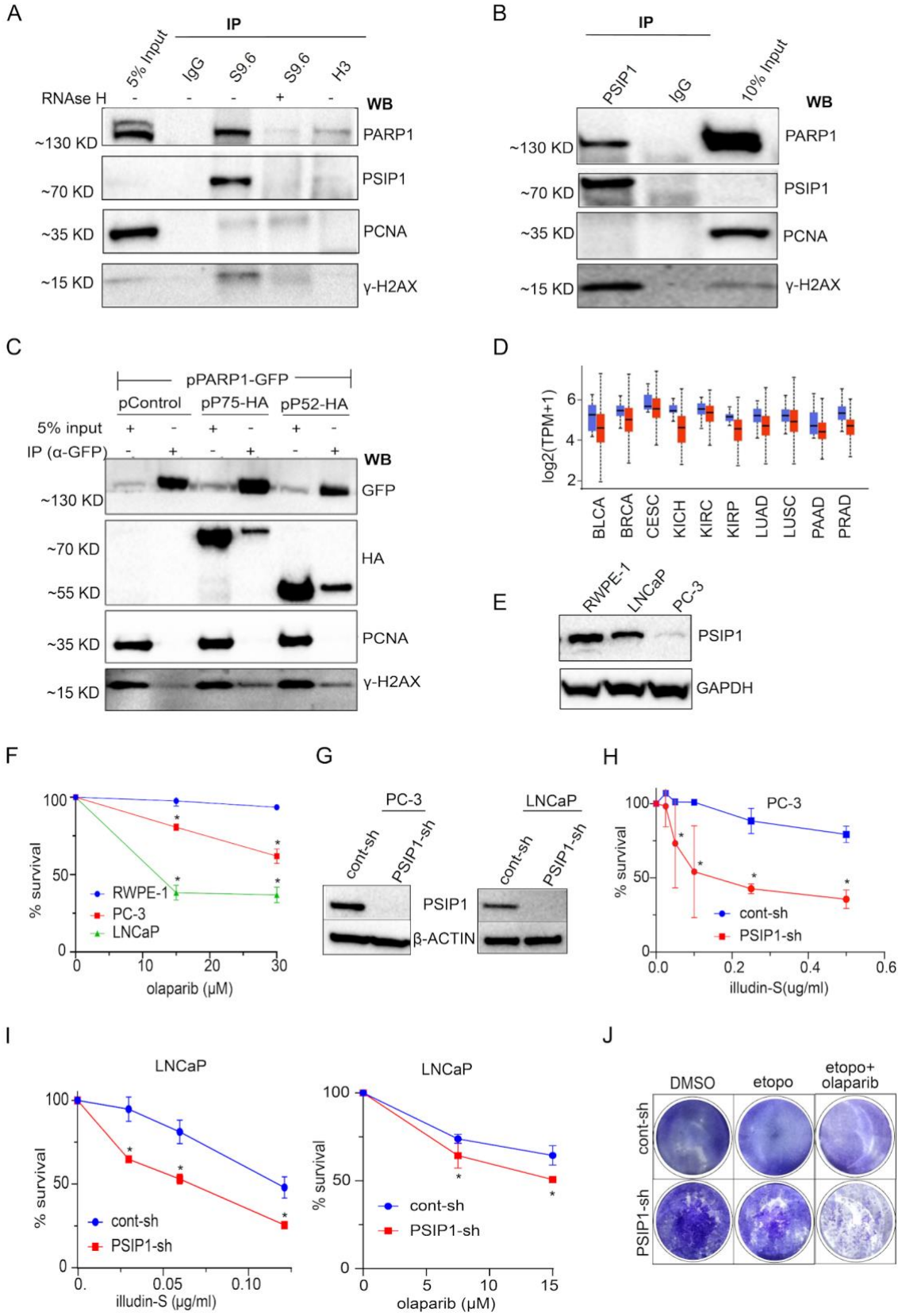

Figure 5. PSIP inhibition sensitises cancer cells to clastogenic. A) Western blotting for S9.6ab-IP with PSIP1/p75, PARP1, γ-H2AX and PCNA antibodies. IgG, histone H3 and RNaseH treatment for the extract was used as negative controls, and 5% of the nuclear extract

of HEK293T cells was used as input. B) Western blotting for PSIP1-IP with PSIP1/p75, PARP1, γ-H2AX and PCNA antibodies, IgG was used as a negative control. 10% of the nuclear extract of HEK293T cells is used as input. C) Western blotting for eGFP-PARP1 and HA-PSIP1 co-IPs performed using GFP-trap beads. αGFP ab was to detect eGFP-PARP1 and αHA to detect HA-PSIP1/p75 and HA-p52, along with γ-H2AX and PCNA antibodies. D) TGCA expression data showing levels of PSIP1 transcripts in different cancers. E) Western blot showing the levels of PSIP1 in prostate normal and cancer cells and their sensitivity to olaparib as determined by CCK-8 assay (F). Mean survival with standard deviation is plotted (n=3) * $p<0.05$, two-way ANOVA test G) Western blot confirming the depletion of PSIP1 by shRNA in PC-3 and LNCaP cells H) Cell survival assay showing % survival of control and PSIP1 knockdown PC3 cells treated with indicated doses of illudin-S determined using the CCK-8 assay (mean± SD; n=3). I) Like (H) but in LNCaP cells treated with illudin-S (left) and olaparib (right). * $p<0.05$, two-way ANOVA. J) Clonogenic survival assay for LNCaP cells treated with etoposide and a combination of olaparib and etoposide.

Discussion

Accumulation of unscheduled, co-transcriptional R-loops affects transcription elongation, compromises genome integrity, and is implicated in diseases. Here we demonstrate that PSIP1/p75 binds to R-loops and proteins associated with R-loops, including PARP1. PSIP1 reduces R-loops formed during transcription to reduce DNA damage and genomic instability (Fig. 6). R-loops and γ-H2AX accumulation are sensitive to inhibition of transcription and overexpression of RNaseH1, which confirms the elevated DNA damage in PSIP1-KD cells due to unresolved R-loops that appear at the site of transcription. Furthermore, PSIP1/p75, R-loops are detected at promoters and gene bodies, confirming the role of PSIP in resolving transcription-induced R-loops. Transcriptional arrest at R-loop sites and reduced replication rate suggest that induced DNA damage is due to R-loop induced collision between RNAP2 transcription with replication fork (Fig. 6).

PSIP1 functions as a histone chaperone in the absence of the FACT complex in differentiated cells to facilitate transcriptional elongation through the nucleosomes (LeRoy et al., 2019). Interestingly, FACT complex subunits (SSRP1 and SPT16) are also known to aid R-loop resolution (Herrera-Moyano et al., 2014). Consistently, subunits of the FACT complex are also detected in both R-loop and PSIP1/p75 interactome (Cristini et al., 2018; Sundarraj et al., 2021). The intrinsic transcript-cleavage activity of RNAP2 requires TFIIS to suppress R-loop formation by reducing RNAP2 pausing and backtracking (Zatreanu et al., 2019b). Intriguingly, PSIP1/p75 also has a TFIIS N-terminal domain that interacts with the transcription elongation complex (Cermakova et al., 2021). All these data suggest the shared function of PSIP1 and FACT in facilitating RNAP2 transcription elongation by resolving R-loops and reducing genomic instability.

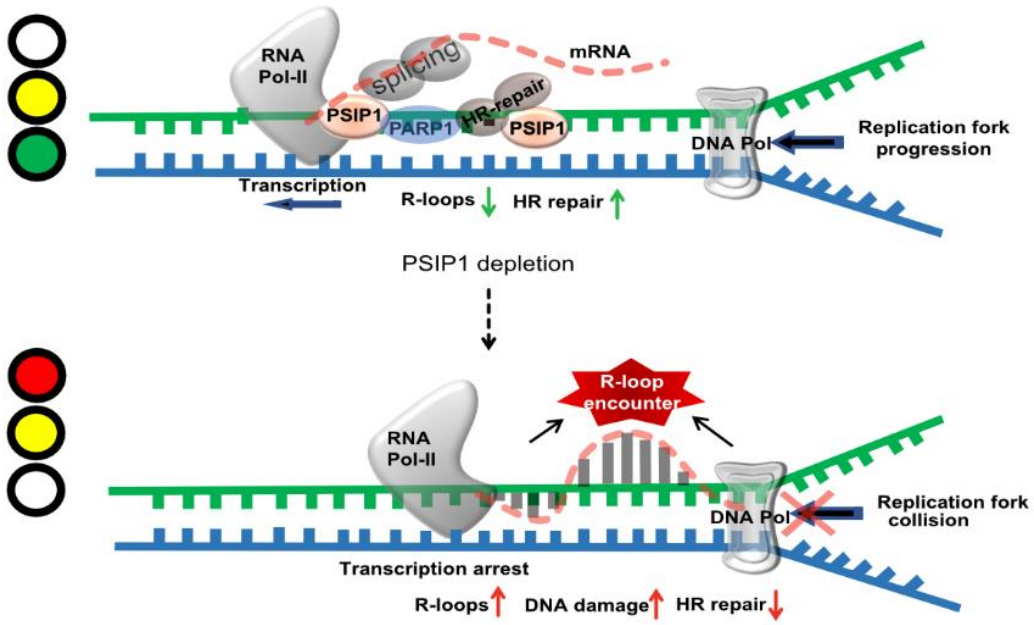

Fig 6. Working model showing the role of PSIP1 in reducing R-loop burden at transcription sites to minimise the replication fork collision with the transcription

Unscheduled R-loops can induce transcriptional repression via various mechanisms (Niehrs and Luke, 2020). DNA damage-induced transcriptional arrest is essential to prevent replication and transcription collision due to aberrant transcripts, which could induce R-loops (Svejstrup, 2010). Reduced nascent transcription at R-loop sites suggests RNAP2 pausing at unresolved R-loops to prevent replication transcription conflict (Zatreanu et al., 2019b). Although PSIP1 binding does not show GC skew, R-loops are enriched at promoters with higher G levels compared to C (GC skew). Although most PSIP1 binding and R-loops are detected around promoters and gene bodies, PSIP1-KD-specific R-loops found in the intergenic regions also show enrichment for γ-H2AX. Interestingly these intergenic regions with accumulated R-loops and γ-H2AX are enriched with histone modifications such as H3K4me1 and H3K27me3, markers of poised regulatory elements (enhancers) (Creyghton et al., 2010).

Several proteins involved in transcription, RNA processing and DNA repairs, such as the PARP1, TFIIS, BRCA1, THO complex, SRSF1, DDX23, SETX, DHX9, TOP1 and others, are known to reduce R-loop burden and genome instability (Alonso and Noordermeer, 2021; Chen et al., 2018; Katyal et al., 2014; Ye et al., 2021; Zatreanu et al., 2019b; Zhao et al., 2022). Interestingly, the splicing factor SRSF1, one of the significant interacting partners of PSIP1, also reduces the genomic instability caused by R-loops (Li and Manley, 2005). The interaction of PSIP1/p75 with CtIP and various DNA repair factors also suggests its role in promoting HR and improving the efficiency of DNA repair during transcription (Daugaard et al., 2012; Sundarraj et al., 2021; Yasuhara et al., 2018). PSIP1 promotes HR in transcribing regions

(Aymard et al., 2014; Daugaard et al., 2012). PARP1 inhibitors are successfully used for cancers deficient in HR due to *BRCA* mutations (Farmer et al., 2005). Importantly, our data show that depletion of PSIP1 results in higher sensitivity of prostate cancer cells to clastogens that cause transcription-coupled DNA damage and the PARP1 inhibitor.

PSIP1 is a chromatin protein that functions as a tethering factor, which recruits proteins involved in RNA processing, DNA repair and transcription elongation. However, since we detect accumulation of both R-loops and DNA damage in the absence of PSIP1, some of the R-loops accumulated in the PSIP1-KD could be due to DNA damage, as double-stranded breaks are known to generate R-loops (Ngo et al., 2021). Although we showed the effect of PSIP1-KD in sensitising cancer cells to clastogens, including PARPi, further studies are needed to validate their efficacy in vivo.

Here we show that PSIP1 interacts with the R-loop complex along with PARP1 and other DNA repair and RNA processing factors. PSIP1 reduces the accumulation of R-loops and genomic instability at the site of transcription. PSIP1 does this by a) reducing the R-loop levels by interacting or recruiting R-loop processing proteins and b) promoting the HR repair pathway to repair damage caused by transcription-induced R-loops. Our data also revealed an increased sensitivity of cancer cells with low levels of PSIP1 to drugs that induce transcription-associated DNA damage and PARP1 inhibitor. With further in vivo validations, insights from this work could be used to modulate R-loops as cancer therapy.

Materials and methods:

Materials:

Material	Company	Identifier
Dimethylsulfoxide	Sigma Life Sciences	D2650
Formaldehyde Solution	Sigma Life Sciences	F8775
Glycerol	Fisher Scientific UK	2002895
HEPES	Sigma Life Sciences	H3375
IGPAL CA-630	Sigma Life Sciences	I8896
Potassium Chloride	Sigma Aldrich	P9333
Skim Milk Powder	Oxoid Ltd	LP0031
Sodium Chloride	Fisher Scientific UK	2315383
Sodium Hydroxide	Fisher Scientific UK	2151855
Sodium Deoxy Cholate	Sigma-Aldrich	D6750

Sodium Dodecy Sulfate	Sigma-Aldrich	436143
Triton-TMX-100	Sigma Life Sciences	T8787
Tween®20	Sigma-Aldrich	P1379
Trizma®base	Sigma Life Sciences	T1503
LB Broth (Miller)	Sigma Life Sciences	L3522
Digitonin high purity	EMD Millipore Corp, USA	300410
DMEM	Gibco	41966-029
RPMI	Gibco	21875-034
Keratinocutes-SFM	Gibco	10724-011
Penicillin-Streptomycin	Gibco	15140122
NUPAGE-LDS sample buffer	Invitrogen	NP0007
Signal Fire (TM) ECL reagent A	Cell signalling Technology	46935P3
Signal Fire (TM) ECL reagent B	Cell signalling Technology	74709P3
Click-iT™ Plus EdU Alexa Fluor™ 647 Assay Kit	Invitrogen	C10634
α-LEDGF/P75 PSIP1	Bethyl	A300-848A
α-RNA-DNA hybrid clone S9.6	EMD Millipore	MABE 1095
α-mouse IgG-HRP	Santa Cruz Biotechnology	SC2031
GFP-Trap®-MA	Chromotek	Gtma-2
Monarch-PCR&DNA cleanup kit	New England Biolabs Limited	T1030L
QubitTMdsDNA HS Assay kit	Invitrogen	Q32854
Delbecco's Phosphate buffered Saline	Sigma-Aldrich	59321C
α-β-actin antibody	Abcam	Ab8226
α-PARP1 antibody	Abcam	Ab191217
α-53BP1 antibody	Abcam	Ab175933
α-γH2AX antibody	Abcam	Ab25350

α-rabbit IgG (Alexa Fluor®594)	Abcam	Ab150080
α-rabbit IgG (Alexa Fluor®488)	Abcam	Ab150077
α-mouse IgG (Alexa Fluor®488)	Abcam	Ab150077
GFP-RNaseH1 plasmid	Martin Regins (Edinburgh)	Addgene ID 108699

Methods:

Cell culture

RWPE-1, PC-3 and LNCaP cells were procured from American Type Culture Collection. *Psip1*$^{-/-}$ and its corresponding WT MEFs were a kind gift from Prof. Alan Engelman (Dana-Farber Cancer Institute, USA) (Shun et al., 2007). The *Psip1*$^{-/-}$ P75 rescue cells were generated in the lab by expressing lentiviral with P75 cDNA (Pradeepa et al., 2017). RWPE-1 cells were cultured in a K-SFM medium. LNCaP cells were cultured in an RPMI medium. PC-3 and MEF cells were cultured in a DMEM medium. The medium was supplemented with 1X penicillin, streptomycin solution, and 10 % FBS. Cells are grown as an adherent monolayer culture in the appropriate medium, in the incubator at 37°C and 5% CO_2.

Generation of PSIP1 knock-down cell lines

Lentiviral shRNA vectors targeting PSIP1 (PSIP1 ShRNA1 TRCN0000298567, PSIP1 shRNA2 TRCN0000286345) and non-targeting control was procured from Sigma Life Sciences. Exponentially growing cells were transduced with lentivirus; 24 h after the transduction media was changed, cells were selected for stable integration using puromycin antibiotic selection at the following concentration RWPE-1- 3ug/ml; PC-3 and LNCaP: 2 ug/ml.

R-loop slot blot

R-loops were estimated by slot blot as described in (Cristini et al., 2018). Cells were lysed in lysis buffer with proteinase K and incubated at 55°C overnight. Isopropanol was added to the lysate, and the DNA was pelleted using centrifugation. The DNA pellet was washed with 70% ethanol, and the pellet was dissolved in TE buffer. The DNA was quantified using a nanodrop spectrophotometer by taking the absorption at 260 nm. Based on the quantification, different quantities of DNA were blotted onto the N+ nitrocellulose membrane using a slot blot apparatus. The membrane was baked at 80°C for 2h, incubated in blocking buffer (5% skimmed

milk in PBS) for 1 hour, followed by incubation with S9.6 antibody overnight at 4ºC in a nutator. Blots were washed three times with PBST (5 min each), incubated with secondary antibody for 2h at room temperature, followed by three washes in PBST 3 times (5 min each) and developed with the ECL reagents A and B (Cell signalling Technology cat no: 46935P3 and 74709P3). The same blots were stained with methylene blue to quantify total DNA.

Immunoprecipitation

The cells were grown to ~90 % confluency, treated with hypotonic solution after washing with PBS, scrapped from the dish using a cell scrapper, and collected in the centrifuge tubes. The nuclear pellet was isolated, and the nuclear extract was prepared using NP-40 lysis buffer. The nuclear extract was subjected to pre-clearing using the protein-A Dyna beads. At this point, a part of the lysate was aliquoted as the input control and the remaining lysate was incubated with the specified primary antibody overnight at 4ºC. Protein-A Dynabeads were added to the lysate and pulled down using the magnet. The supernatant was discarded, and the beads were washed three times (10 min each) with NP-40 lysis buffer. The proteins bound to the beads were extracted using a Bolt Nupage 4X sample loading buffer.

Western blotting

The sample lysates were loaded onto the True-PAGE precast gels and were subjected to electrophoresis at 95 V for 90 min. Proteins were transferred to the PVDF membrane using a Trans-blot turbo transfer system (Bio-Rad). Membranes were blocked with every blot blocking buffer (Bio-Rad), incubated with primary antibody overnight at 4ºC, followed by three washes with 0.05% PBST, incubated with HRP coupled secondary antibodies for 2h at room temperature, followed by three times with 0.05% PBST. Membranes were developed with ECL reagents.

Clonogenic assay

A thousand cells for each cell type were plated in triplicates for each treatment per well of a 96-well plate. The cells were allowed to grow and develop colonies for 14 days after the treatment of the indicated drugs. The colonies were washed and fixed with methanol, stained with crystal violet, and counted manually to calculate the survival fraction using the following formula. Survival fraction = plating efficiency of treatment/plating efficiency of control X 100.

Immunofluorescence

The exponentially growing cells were treated with indicated drugs for a specified time, cells fixing the cells using 4% formaldehyde for 20 min, washed with PBS, permeabilised using 0.3% Triton-X-100, blocked using 3% BSA for 1 h, and incubated with primary antibody at 4ºC for overnight. After washing with PBST three times, the cells were incubated with

secondary antibodies conjugated to different fluorophores and were imaged in the laser-based InCell high-content imaging system (IN Cell Analyzer 6000, GE Healthcare). The images were analysed using dedicated software such as INCarta, Developer and ImageJ.

RNAseH and triptolide treatment

Control and PSIP KD HEK293T cells were transfected with pEGFP-RNASEH1 (Addgene ID 108699, a kind gift from Martin Reigns (MRC-HGU, Edinburgh) (Bubeck et al., 2011). For immunoprecipitation and slot blot, the control and PSIP KD HEK293T cells were treated with 50 nM triptolide, and cells were harvested 48 h after the treatment.

CUT&Tag

CUT&Tag was performed according to (Kaya-Okur et al., 2019) protocol with modifications to tissue processing as described below. Experiments were performed in biological duplicates from each cell type. ~100,000 cells were pelleted by centrifugation for 3 min at 600 x g at room temperature and resuspended in 500µl of ice-cold NE1 buffer (20 mM HEPES-KOH pH 7.9, 10 mM KCl, 0.5 mM spermidine, 1% Triton X-100, and 20 % glycerol and cOmplete EDTA free protease inhibitor tablet) and was let to sit for 10 min on ice. Nuclei were pelleted by centrifugation for 4 min at 1300 x g at 4°C, resuspended in 500µl of wash buffer and held on ice until beads were ready. The required amount of BioMag Plus Concanavalin-A-conjugated magnetic beads (ConA beads, Polysciences, Inc) were transferred into the binding buffer (20 mM HEPES-KOH pH 7.9, 10 mM KCl, 1mM $CaCl_2$ and 1mM $MnCl_2$), washed once in the same buffer, each time placing them on a magnetic rack to allow the beads to separate from the buffer and resuspended in binding buffer. 10 µl of beads was added to each tube containing cells and rotated on an end-to-end rotator for 10 minutes. After a pulse spin to remove liquid from the cap, tubes were placed on a magnet stand to clear, the liquid was withdrawn, and 800 µl of antibody buffer containing 1 µg of primary antibodies was added and incubated at 4 ºC overnight in a nutator. Secondary antibodies (guinea pig α-rabbit antibody, Antibodies online, ABIN101961 or α-mouse antibodies) were diluted 1:100 in dig-wash buffer (5% digitonin in wash buffer), and 100 µL was added in per sample while gently vortexing to allow the solution to dislodge the beads from the sides and incubated for 60 min on a nutator. Unbound antibodies were washed in 1 ml of dig-wash buffer three times. 100 µl of (1:250 diluted) protein-A-Tn5 loaded with adapters in dig-300 buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 0.5 mM spermidine with Roche cOmplete EDTA free protease inhibitor) was added to the samples, placed on nutator for 1 hour and washed three times in 1 ml of dig-300 buffer to remove unbound pA-Tn5. 300 µL tagmentation buffer (Dig-300 buffer + 5 mM $MgCl_2$) was added while gentle vortexing, and samples were incubated at 37ºC for 1 hr. Tagmentation was stopped

by adding 10 µL 0.5M EDTA, 3 µL 10% SDS and 2.5 µL 20 mg/mL Proteinase K to each sample. Samples were mixed by full-speed vortexing for ~2 seconds and incubated for 1 hr at 55ºC to digest proteins. DNA was purified by phenol: chloroform extraction using phase-lock tubes (Quanta Bio) followed by ethanol precipitation. Libraries were prepared using NEBNext HiFi 2x PCR Master mix (M0541S) with a 72ºC gap-filling step followed by 13 cycles of PCR with 10-second combined annealing and extension for the enrichment of short DNA fragments. Libraries were sequenced in Novaseq 6000 (Novogene) with 150bp paired-end reads.

Transient transcriptome sequencing (TT-seq)

TT-seq was performed according to the protocol described in (Vlaming et al., 2022) with modifications including CORALL total RNAseq library preparation kit (Lexogen Cat. no. 147.24) was used for sequencing library preparation.

CCK-8 –assay

Seven thousand cells were plated on a 96-well plate for overnight attachment. The next day, cells were added with different drug concentrations and cultured for 48h in the incubator. Fresh media containing 50ug/ml of WST-8 dye was replaced and incubated for 1h. Then the absorbance was recorded at 450 nm. The absorbance value was used to calculate the percentage of viable cells.

EdU incorporation assay

EdU incorporation was performed using Click-iT™ Plus EdU Alexa Fluor™ 647 Assay Kit (Invitrogen). EdU (5-ethynyl-2'-deoxyuridine, 20 mM) was added to the culture dishes with cells and grown for 4 hours. Then, cells were fixed using 3.8% formaldehyde for 15 min at room temperature and washed twice with 3% BSA in PBS. Cells were permeabilised using 0.5% Triton-X-100 (20 min) and stained using the Click-iT cocktail following the vendor's protocol. Cells were imaged under IN Cell Analyzer 2200 (GE Healthcare Life Sciences), and the fluorescent intensity was quantified (from at least 5000 cells) using IN Carta Image Analysis Software (Molecular Devices).

Data Analysis:

CUT&Tag

The pair-end sequencing reads generated for the CUT&Tag were trimmed for the sequencing adapters with a trimmomatic tool, followed by mapping onto the human genome hg38 using *bowtie2*. The reads were aligned using following parameters: *--very-sensitive-local --no-unal --no-mixed --no-discordant --phred33 -I 10 -X 700*. The data was processed using SAMtools to generate bam files and sort them. All the replicates were pooled using *SAMtools merge*. Merged bam files were used to generate bed files, bedgraph and bigwig.

Peak calling

Peaks on the mapped reads were called using the *SEACR* tool using IgG as a background with *"norm and relaxed"* options. For each replicate per sample, the peaks were called using both individual IgG replicates as background. The peaks obtained were pooled, only the peaks consistent in both replicates were retained, and coordinates were merged. These peaks were used for further analyses.

Peaks were processed using the *bedtools intersect* option to filter the peaks unique to either WT (PSIP parks) or KD (YH2AX and S9.6 peaks). Peak distribution across genomic landmarks (TSS, gene body or intergenic) was done on the coordinates obtained from the UCSC genome browser using *bedtools intersect*.

Bigwig generation and plotting

The signal files for merged samples were generated using the deepTools bamCoverage option with the following parameters: *--binSize 20 –normalise Using CPM --scaleFactor --smoothLength 60 --extended 150 –centerReads.* The signal was normalised between WT and KD for each target by comparing the reads mapping to the *E. coli* genome. The bigwigs generated were used for viewing on the genome browser or plotting as a heatmap or average summary plot on the genomic landmarks and peaks. Matrices generated through *computeMatrix* with *reference-point or scale-region option* were used as input for heatmap *(plotHeatmap) or* average summary *(plotProfile), or* violin *(R package ggplot2)* plots. The genome browser views were captured on the UCSC genome browser.

RNAseq and TTseq

Pair-end reads for the RNA-seq and TT-seq were aligned against the human genome (hg38) via STAR aligner following the Bluebee-CORALL pipeline of mapping. The replicates were merged for each sample using SAMtools merge, followed by filtering out the multi-mapped reads by defining MAPQ as 255. The bigwigs were generated using the deepTools bamCoverage tool with the option to normalise using RPKM. These bigwigs were used for comparing the signal in the genome browser or plotting heatmaps, average summary plots or violin plots on the genes and peaks datasets. Differential gene expression for the RNAseq was performed through the DESeq2 package. The fragments count matrix was generated using the Subreads feature counts option. This count matrix was subjected to DESeq2, and the result obtained from the analysis was used to plot as a Volcano plot.

GC skew analysis

GC skew for the S9.6, PSIP1 and yH2AX CUT&Tag peaks was calculated for the peaks using the ratio of (G-C)/(G+C). The values were used to plot the frequency distribution. For the

background frequency of GC-skew, randomly shuffled genomic coordinates were used to compare the GC-skew. Further, signal distribution of the PSIP1, S9.6 and yH2AX was compared on the GC-skewed regions of our genome (Ginno et al., 2012). The coordinates were lifted from hg18 to hg38 and used for plotting. Likewise, the signal was also plotted for the G-rich promoters (Ginno et al., 2012).

Acknowledgements

We thank QMUL epigenetics hub members and Madapura lab members for discussing and reading the manuscript. We thank Francisco Martin-Zamora and Chema Martin (QMUL) lab for sharing purified pATn5, Martin Regins (MRC-HGU, Edinburgh) for RNAaseH1 plasmid, Benjamin JE Martin (Harvard Medical School) for TT-seq protocol. We thank Blizard core facilities and Luke Gammon and Gary Warnes for their help in image analysis training and FACS analysis. This research utilised Queen Mary's Apocrita HPC facility, supported by QMUL Research-IT.

Funding: Marie Skłodowska-Curie grant (896079-JS), Medical Research Council UKRI/MRC grant and Barts charity grant (MR/T000783/1; MGU0475-MMP).

Author contributions:

Conceptualisation: MMP, JS, HT

Methodology: MMP, JS, HT, MP, GB

Investigation: MMP, JS, MP,

Visualisation: MMP, JS, MP,

Funding acquisition: MMP, JS,

Project administration: MMP

Supervision: MMP, HT

Writing – original draft: MMP, JS,

Writing – review & editing: MMP, JS, HT, MP, GB

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (Sundarraj et al., 2022) and are accessible through GEO Series accession number GSE.

References

Alonso, M.S.M., and Noordermeer, S.M. (2021). Untangling the crosstalk between BRCA1 and R-loops during DNA repair. Nucleic Acids Res. *49*, 4848–4863. https://doi.org/10.1093/NAR/GKAB178.

Aymard, F., Bugler, B., Schmidt, C.K., Guillou, E., Caron, P., Briois, S., Iacovoni, J.S., Daburon, V., Miller, K.M., Jackson, S.P., et al. (2014). Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. Nat. Struct. Mol. Biol. *21*, 366–374. https://doi.org/10.1038/nsmb.2796.

Basu, A., Cajigas-Du Ross, C.K., Rios-Colon, L., Mediavilla-Varela, M., Daniels-Wells, T.R., Leoh, L.S., Rojas, H., Banerjee, H., Martinez, S.R., Acevedo-Martinez, S., et al. (2016). LEDGF/p75 overexpression attenuates oxidative stress-induced necrosis and upregulates the oxidoreductase ERP57/PDIA3/GRP58 in prostate cancer. PLoS One *11*, 1–25. https://doi.org/10.1371/journal.pone.0146549.

Bhargavan, B., Fatma, N., Chhunchha, B., Singh, V., Kubo, E., and Singh, D.P. (2012). LEDGF gene silencing impairs the tumorigenicity of prostate cancer DU145 cells by abating the expression of Hsp27 and activation of the Akt/ERK signaling pathway. Cell Death Dis. *3*, e316. https://doi.org/10.1038/cddis.2012.57.

Boguslawski, S.J., Smith, D.E., Michalak, M.A., Mickelson, K.E., Yehle, C.O., Patterson, W.L., and Carrico, R.J. (1986). Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids. J. Immunol. Methods *89*, 123–130. https://doi.org/10.1016/0022-1759(86)90040-2.

Bubeck, D., Reijns, M.A.M., Graham, S.C., Astell, K.R., Jones, E.Y., and Jackson, A.P. (2011). PCNA directs type 2 RNase H activity on DNA replication and repair substrates. Nucleic Acids Res. *39*, 3652–3666. https://doi.org/10.1093/NAR/GKQ980.

Cermakova, K., Demeulemeester, J., Lux, V., Nedomova, M., Goldman, S.R., Smith, E.A., Srb, P., Hexnerova, R., Fabry, M., Madlikova, M., et al. (2021). A ubiquitous disordered protein interaction module orchestrates transcription elongation. Science *374*, 1113–1121. https://doi.org/10.1126/SCIENCE.ABE2913.

Chabanon, R.M., Morel, D., Eychenne, T., Colmet-Daage, L., Bajrami, I., Dorvault, N., Garrido, M., Meisenberg, C., Lamb, A., Ngo, C., et al. (2021). PBRM1 Deficiency Confers Synthetic Lethality to DNA Repair Inhibitors in Cancer. Cancer Res. *81*, 2888–2902. https://doi.org/10.1158/0008-5472.CAN-21-0628.

Chen, L., Chen, J.Y., Huang, Y.J., Gu, Y., Qiu, J., Qian, H., Shao, C., Zhang, X., Hu, J., Li, H., et al. (2018). The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. Mol. Cell *69*, 412-425.e6. https://doi.org/10.1016/J.MOLCEL.2017.12.029.

Cohen, S., Guenolé, A., Lazar, I., Marnef, A., Clouaire, T., Vernekar, D. V., Puget, N., Rocher, V., Arnould, C., Aguirrebengoa, M., et al. (2022). A POLD3/BLM dependent pathway handles DSBs in transcribed chromatin upon excessive RNA:DNA hybrid accumulation. Nat. Commun. *13*, 2012. https://doi.org/10.1038/s41467-022-29629-2.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. U. S. A. *107*, 21931–21936. https://doi.org/10.1073/PNAS.1016071107/-/DCSUPPLEMENTAL/PNAS.201016071SI.PDF.

Cristini, A., Groh, M., Kristiansen, M.S., and Gromak, N. (2018). RNA/DNA Hybrid Interactome Identifies DXH9 as a Molecular Player in Transcriptional Termination and R-Loop-Associated DNA Damage. Cell Rep. *23*, 1891–1905. https://doi.org/10.1016/j.celrep.2018.04.025.

Daugaard, M., Baude, A., Fugger, K., Povlsen, L.K., Beck, H., Sørensen, C.S., Petersen, N.H.T., Sorensen, P.H.B., Lukas, C., Bartek, J., et al. (2012). LEDGF (p75) promotes DNA-end resection and homologous recombination. Nat. Struct. Mol. Biol. *19*, 803–810. https://doi.org/10.1038/nsmb.2314.

Farmer, H., McCabe, H., Lord, C.J., Tutt, A.H.J., Johnson, D.A., Richardson, T.B., Santarosa, M., Dillon, K.J., Hickson, I., Knights, C., et al. (2005). Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature *434*, 917–921. https://doi.org/10.1038/NATURE03445.

French, J.D., Johnatty, S.E., Lu, Y., Beesley, J., Gao, B., Kalimutho, M., Henderson, M.J., Russell, A.J., Kar, S., Chen, X., et al. (2016). Germline polymorphisms in an enhancer of PSIP1 are associated with progression-free survival in epithelial ovarian cancer. Oncotarget *7*. https://doi.org/10.18632/oncotarget.7047.

García-Muse, T., and Aguilera, A. (2019). R Loops: From Physiological to Pathological Roles. Cell *179*, 604–618. https://doi.org/10.1016/J.CELL.2019.08.055.

Ginno, P.A., Lott, P.L., Christensen, H.C., Korf, I., and Chédin, F. (2012). R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. Mol. Cell *45*. https://doi.org/10.1016/j.molcel.2012.01.017.

Hamperl, S., Bocek, M.J., Saldivar, J.C., Swigut, T., and Cimprich, K.A. (2017). Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. Cell *170*, 774-786.e19. https://doi.org/10.1016/J.CELL.2017.07.043.

Herrera-Moyano, E., Mergui, X., García-Rubio, M.L., Barroso, S., and Aguilera, A. (2014). The yeast and human FACT chromatin-reorganizing complexes solve R-loop-mediated transcription-replication conflicts. Genes Dev. *28*, 735–748. https://doi.org/10.1101/GAD.234070.113.

Katyal, S., Lee, Y., Nitiss, K.C., Downing, S.M., Li, Y., Shimada, M., Zhao, J., Russell, H.R., Petrini, J.H.J., Nitiss, J.L., et al. (2014). Aberrant topoisomerase-1 DNA lesions are pathogenic in neurodegenerative genome instability syndromes. Nat. Neurosci. *17*, 813–821. https://doi.org/10.1038/NN.3715.

Kaya-Okur, H.S., Wu, S.J., Codomo, C.A., Pledger, E.S., Bryson, T.D., Henikoff, J.G., Ahmad, K., and Henikoff, S. (2019). CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat. Commun. 2019 101 *10*, 1–10. https://doi.org/10.1038/s41467-

019-09982-5.

Krokan, H., Wist, E., and Krokan, R.H. (1981). Aphidicolin inhibits DNA synthesis by DNA polymerase alpha and isolated nuclei by a similar mechanism. Nucleic Acids Res. *9*, 4709–4719. https://doi.org/10.1093/NAR/9.18.4709.

LeRoy, G., Oksuz, O., Descostes, N., Aoi, Y., Ganai, R.A., Kara, H.O., Yu, J.R., Lee, C.H., Stafford, J., Shilatifard, A., et al. (2019). LEDGF and HDGF2 relieve the nucleosome-induced barrier to transcription. Sci. Adv. *5 : eaay30*. https://doi.org/10.1101/500926.

Li, X., and Manley, J.L. (2005). Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell *122*, 365–378. https://doi.org/10.1016/j.cell.2005.06.008.

Lin, W.-L., Chen, J.-K., Wen, X., He, W., Zarceno, G.A., Chen, Y., Chen, S., Paull, T.T., and Liu, H. (2022). DDX18 prevents R-loop-induced DNA damage and genome instability via PARP-1. Cell Rep. *40*, 111089. https://doi.org/10.1016/J.CELREP.2022.111089.

Liu, L.F., Desai, S.D., Li, T.K., Mao, Y., Sun, M., and Sim, S.P. (2000). Mechanism of action of camptothecin. Ann. N. Y. Acad. Sci. *922*, 1–10. https://doi.org/10.1111/J.1749-6632.2000.TB07020.X.

Montecucco, A., Zanetta, F., and Biamonti, G. (2015). Molecular mechanisms of etoposide. EXCLI J. *14*, 95–108. https://doi.org/10.17179/EXCLI2015-561.

Ngo, G.H.P., Grimstead, J.W., and Baird, D.M. (2021). UPF1 promotes the formation of R loops to stimulate DNA double-strand break repair. Nat. Commun. 2021 121 *12*, 1–15. https://doi.org/10.1038/s41467-021-24201-w.

Niehrs, C., and Luke, B. (2020). Regulatory R-loops as facilitators of gene expression and genome stability. Nat. Rev. Mol. Cell Biol. *21*, 167–178. https://doi.org/10.1038/s41580-019-0206-3.

Van Nuland, R., Van Schaik, F.M.A., Simonis, M., Van Heesch, S., Cuppen, E., Boelens, R., Timmers, H.T.M., and Van Ingen, H. (2013). Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain. Epigenetics and Chromatin *6*, 1–12. https://doi.org/10.1186/1756-8935-6-12/FIGURES/7.

Perego, M.G.L., Taiana, M., Bresolin, N., Comi, G.P., and Corti, S. (2019). R-Loops in Motor Neuron Diseases. Mol. Neurobiol. *56*, 2579–2589. https://doi.org/10.1007/S12035-018-1246-Y.

Petermann, E., Lan, L., and Zou, L. (2022). Sources, resolution and physiological relevance of R-loops and RNA–DNA hybrids. Nat. Rev. Mol. Cell Biol. https://doi.org/10.1038/s41580-022-00474-x.

Pradeepa, M.M., Sutherland, H.G., Ule, J., Grimes, G.R., and Bickmore, W.A. (2012). Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. PLoS Genet. *8*, e1002717. .

Pradeepa, M.M., Grimes, G.R., Taylor, G.C. a, Sutherland, H.G., and Bickmore, W. a (2014). Psip1/Ledgf p75 restrains Hox gene expression by recruiting both trithorax and polycomb group proteins. Nucleic Acids Res. *42*, 9021–9032. https://doi.org/10.1093/nar/gku647.

Pradeepa, M.M., McKenna, F., Taylor, G.C.A., Bengani, H., Grimes, G.R., Wood, A.J., Bhatia, S., Bickmore, W.A., Lenhard, B., Wells, C., et al. (2017). Psip1/p52 regulates posterior Hoxa genes through activation of lncRNA Hottip. PLOS Genet. *13*, e1006677. https://doi.org/10.1371/journal.pgen.1006677.

Reiter, H., Milewskiy, M., and Kelley, P. (1972). Mode of action of phleomycin on Bacillus subtilis. J. Bacteriol. *111*, 586–592. https://doi.org/10.1128/JB.111.2.586-592.1972.

Ríos-Colón, L., Cajigas-Du Ross, C.K., Basu, A., Elix, C., Alicea-Polanco, I., Sanchez, T.W., Radhakrishnan, V., Chen, C.-S., and Casiano, C.A. (2017). Targeting the stress oncoprotein LEDGF/p75 to sensitize chemoresistant prostate cancer cells to taxanes. Oncotarget *8*, 24915–24931. https://doi.org/10.18632/oncotarget.15323.

Salvi, J.S., and Mekhail, K. (2015). R-loops highlight the nucleus in ALS. Nucleus *6*, 23–29. https://doi.org/10.1080/19491034.2015.1004952.

Schwalb, B., Michel, M., Zacher, B., Hauf, K.F., Demel, C., Tresch, A., Gagneur, J., and Cramer, P. (2016). TT-seq maps the human transient transcriptome. Science (80-.). *352*, 1225–1228. https://doi.org/10.1126/science.aad9841.

Shun, M.C., Raghavendra, N.K., Vandegraaff, N., Daigle, J.E., Hughes, S., Kellam, P., Cherepanov, P., and Engelman, A. (2007). LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. Genes Dev. *21*, 1767. https://doi.org/10.1101/GAD.1565107.

Singh deepak (2017). PSIP1/p75 promotes tumorigenicity in breast cancer cells by promoting the transcription of cell cycle genes. Carcinogenesis *38*, 966–975. https://doi.org/https://doi.org/10.1093/carcin/bgx062.

Skourti-Stathaki, K., Torlai Triglia, E., Warburton, M., Voigt, P., Bird, A., and Pombo, A. (2019). R-Loops Enhance Polycomb Repression at a Subset of Developmental Regulator Genes. Mol. Cell *73*, 930-945.e4. https://doi.org/10.1016/j.molcel.2018.12.016.

Sundarraj, J., Taylor, G.C.A., von Kriegsheim, A., and Pradeepa, M.M. (2021). H3K36me3 and PSIP1/LEDGF associate with several DNA repair proteins, suggesting their role in efficient DNA repair at actively transcribing loci. Wellcome Open Res. *2*, 83. https://doi.org/10.12688/wellcomeopenres.11589.3.

Svejstrup, J.Q. (2010). The interface between transcription and mechanisms maintaining genome integrity. Trends Biochem. Sci. *35*, 333–338. https://doi.org/10.1016/j.tibs.2010.02.001.

Vlaming, H., Mimoso, C.A., Field, A.R., Martin, B.J.E., and Adelman, K. (2022). Screening thousands of transcribed coding and non-coding regions reveals sequence determinants of RNA polymerase II elongation potential. Nat. Struct. Mol. Biol. *29*, 613–620. https://doi.org/10.1038/S41594-022-00785-9.

Wang, K., Wang, H., Li, C., Yin, Z., Xiao, R., Li, Q., Xiang, Y., Wang, W., Huang, J., Chen, L., et al. (2021). Genomic profiling of native R loops with a DNA-RNA hybrid recognition sensor. Sci. Adv. *7*. https://doi.org/10.1126/SCIADV.ABE3516/SUPPL_FILE/ABE3516_TABLES_S1_TO_S3.XLSX.

Yasuhara, T., Kato, R., Hagiwara, Y., Shiotani, B., Yamauchi, M., Nakada, S., Shibata, A., and Miyagawa, K. (2018). Human Rad52 Promotes XPG-Mediated R-loop Processing to Initiate Transcription-Associated Homologous Recombination Repair. Cell *175*, 558-570.e11. https://doi.org/10.1016/J.CELL.2018.08.056.

Ye, B.J., Kang, H.J., Lee-Kwon, W., Kwon, H.M., and Choi, S.Y. (2021). PARP1-mediated PARylation of TonEBP prevents R-loop-associated DNA damage. DNA Repair (Amst). *104*. https://doi.org/10.1016/J.DNAREP.2021.103132.

Zatreanu, D., Han, Z., Mitter, R., Tumini, E., Williams, H., Gregersen, L., Dirac-Svejstrup, A.B., Roma, S., Stewart, A., Aguilera, A., et al. (2019a). Elongation Factor TFIIS Prevents Transcription Stress and R-Loop Accumulation to Maintain Genome Stability. Mol. Cell *76*, 57-69.e9. https://doi.org/10.1016/j.molcel.2019.07.037.

Zatreanu, D., Han, Z., Mitter, R., Tumini, E., Williams, H., Gregersen, L., Dirac-Svejstrup, A.B., Roma, S., Stewart, A., Aguilera, A., et al. (2019b). Elongation Factor TFIIS Prevents Transcription Stress and R-Loop Accumulation to Maintain Genome Stability. Mol. Cell *76*, 57-69.e9. https://doi.org/10.1016/J.MOLCEL.2019.07.037.

Zhao, J., Tian, S., Guo, Q., Bao, K., Yu, G., Wang, X., Shen, X., Zhang, J., Chen, J., Yang, Y., et al. (2022). A PARylation-phosphorylation cascade promotes TOPBP1 loading and RPA-RAD51 exchange in homologous recombination. Mol. Cell *82*, 2571-2587.e9. https://doi.org/10.1016/J.MOLCEL.2022.04.031.