Identification and characterization of novel factors in the DNA damage response

Singh, J.K.

Citation

Version:	Publisher's Version
License:	[Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)
Downloaded from:	https://hdl.handle.net/1887/3485639

Note: To cite this publication please use the final published version (if applicable).

Identification and characterization of novel factors in the DNA damage response

Jenny Kaur Singh

Identification and characterization of novel factors in the DNA damage response

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 9 november 2022
klokke 11:15 uur

door

Jenny Kaur Singh
geboren te Amsterdam
in 1991

1e Promotor   Prof. Dr. H. van Attikum

2e Promotor   Prof. Dr. Ir. S.M. van der Maarel

Leden promotiecommissie   Prof. Dr. M. van Vugt (Universitair Medisch
                              Centrum Groningen)
                              Prof. Dr. J.A.F. Marteijn (Erasmus Medisch
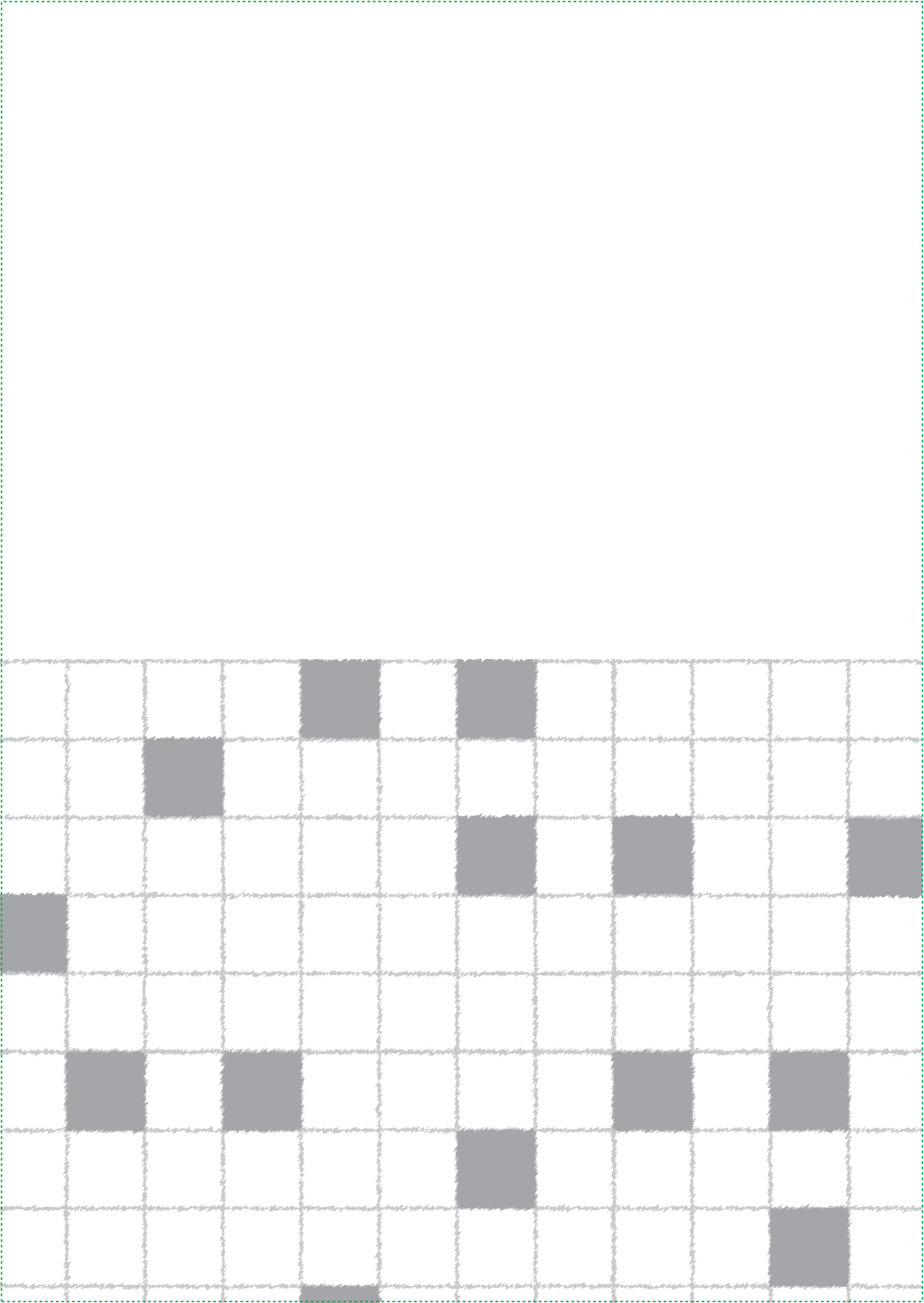   Centrum Rotterdam)
                              Prof. Dr. A.C.O. Vertegaal
                              Dr. F. Mattiroli (Hubrecht Instituut)

Voor Papa, Mama en Simran

CONTENTS

CHAPTER 1

General introduction

The cellular responses to DNA damage

The genetic information in our cells is constantly challenged by several sources that can cause DNA damage. These include exogenous sources such as ionizing radiation (IR), ultraviolet radiation (UV), alkylating compounds, viral infections and chemotherapeutic drugs, as well as endogenous sources such as aberrant DNA replication or reactive oxygen species (ROS) that are generated as a byproduct of cell metabolism [1]. Amongst the various lesions that can occur, DNA double-strand breaks (DSBs) are considered as most dangerous [2]. Furthermore, during DNA replication the genome is particularly vulnerable to lesions that block replication fork progression and prolonged stalling of replication forks can lead to fork collapse, DSBs and genome instability [3]. Cells combat these lesions and provide surveillance to the genome by activation of the DNA damage response (DDR). The DDR is a complex evolutionary conserved cellular network that has the ability to sense, signal and repair DNA lesions in time and space while coordinating gene expression, modulating chromatin structure, regulating cell-cycle progression or in the case of irreparable DNA damage, promote apoptosis [4]. An inefficient or inaccurate DDR may lead to mutations and genome instability, contributing to the development of diseases such as cancer, neurodegenerative disorders and ageing.

1

DSB repair

DNA double-strand break repair pathways

In somatic cells, two main pathways dictate the repair of DSBs either by homologous recombination (HR) or non-homologous end joining (NHEJ). Canonical non-homologous end joining (cNHEJ) is the dominant repair pathway in human cells, which is active throughout the entire cell cycle. This repair process is initiated by the binding of Ku70/Ku80 (Ku) heterodimers to the broken ends, followed by activation of DNA-PKcs and recruitment of APLF via its conserved Ku-binding motif (KBM). Ku and DNA-PKcs recruit and activate several enzymes involved in the processing of DNA ends, including the Artemis nuclease and DNA polymerases μ and λ. Finally, the broken ends are sealed in a mainly error-free manner by the assembly of the XLF-XRCC4-LIG4 complex, which is stimulated by PAXX [5, 6]. HR is the error-free repair pathway, which is active in the S and G2 phase of the cell cycle. It is initiated by extensive end resection of the broken ends by endo- and exo-nucleases such as MRE11, CtIP, DNA2 and EXO, resulting in the formation of 3'overhangs at the DSB ends, which become coated by the ssDNA-binding complex RPA. This is followed by the BRCA1-dependent loading of PALB2, which facilitates the assembly of BRCA2 and the recombinase RAD51, the latter of which catalyzes HR by using the sister chromatid as a template for repair. When cNHEJ or HR are disabled, DSB repair can also occur via alternative non-homologous end joining (aNHEJ) or single-strand annealing (SSA). During aNHEJ, microhomology is used that ranges between 2 base pairs (bp) and 20 bp, which seals the broken ends in an error-prone and in a manner dependent on the XRCC1-Ligase III complex and POLQ [6]. Particularly, the repair of blunt DNA ends is favored by POLQ-mediated end joining [7]. Alternatively, in the case of more extensive end-resection (> 20 bp) microhomology usage may lead to deleterious, RAD52-dependent repair of DSBs via single-strand annealing (SSA) (Figure 1 and [8].
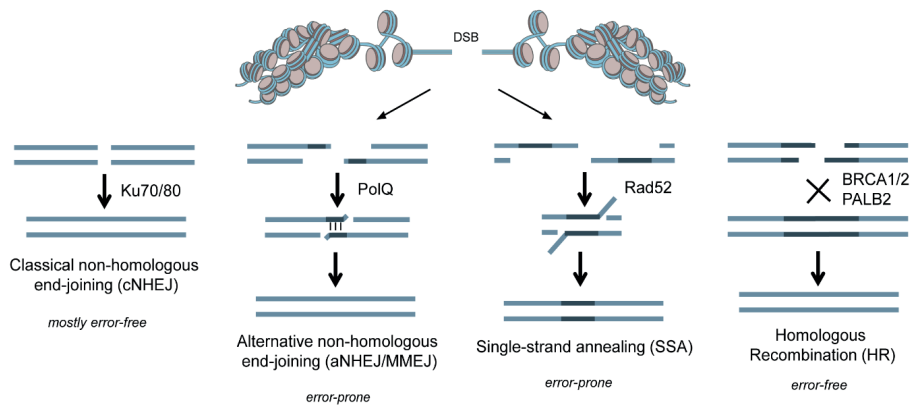
Figure 1. Overview of the different DSB repair pathways. DSBs that arise in chromatin can be repaired via two main pathways homologous recombination (HR) or non-homologous end joining (NHEJ). Classical NHEJ (cNHEJ) relies on the Ku78/Ku80 heterodimer and is mostly error-free. HR on the other hand, is a high fidelity repair pathway dependent on BRCA1/2-PALB2-RAD51, which loads RAD51. Subsequently, RAD51 invades the sister chromatid that serves as a template for repair. When HR and cNHEJ are disabled, DSB repair relies on the more error-prone pathway alternative NHEJ (aNHEJ), in which microhomology is used by PolQ to guide repair. In the case of more extensive resection, DSBs can be repaired via the RAD52-dependent highly error-prone single-strand annealing (SSA) pathway.

DSB repair: pathway choice

DSB pathway choice is tightly regulated by several mechanisms which ensure that HR is strictly regulated in S/G2 phase. During this phase, CDK-specific phosphorylation of CtIP enables its resection activity [9]. Furthermore, during G1 phase the interaction of BRCA1 with PALB2 is inhibited by the suppressive ubiquitylation of the BRCA1 interacting domain in PALB2, impairing HR by limiting BRCA1-PALB2-BRCA2-RAD51 complex formation [10]. In addition, the KU-binding protein CYREN inhibits NHEJ in S/G2 phase by protecting breaks with overhangs [11]. Also, in G1 phase TIRR-mediated recruitment of 53BP1 to DSBs inhibits DNA end-resection through 53BP1's effectors RIF1, MAD2L2, REV7 and the Shieldin complex, thereby impairing HR and to favor NHEJ. [12]. The 53BP1-RIF1-Shieldin complex mediates the recruitment of the CTC1-STN1-TEN1 (CST) complex to DSBs. The CST complex and Polα promote the fill-in of resected DSB ends, which may be protected against DNA end resection by Shieldin [13]. Remarkably, it has also been reported that small tandem duplications (TD) are generated by CST-Polα at DSBs containing 3' overhangs in a 53BP1 and shieldin dependent manner, leading to Ku-dependent cNHEJ [14]. Finally, it was shown that ZMYM3 facilitates the recruitment of ABRA1 and BRCA1 to DSBs, thereby antagonizing the HR-suppressive effects of the

BRCA1-A complex and facilitating the loading of RAD51 and HR [15]. In addition to the above-mentioned mechanisms, DSB repair pathway choice also depends on the complexity of the DSB ends. During cNHEJ, compatible DNA ends are not processed, and incompatible DNA ends are minimally processed in a ligation-competent complex or otherwise protected by Ku70/Ku80 and DNA-PKcs [5]. This type of repair also plays an important role during V(D)J recombination and class-switch recombination (CSR), which are processes occurring during lymphocyte differentiation that ensure the generation of a diverse class of antibodies in order to eliminate pathogens [16]. During V(D)J recombination, RAG1 and RAG2 induce DSBs between the V, D and J segments, and the conserved recombination signal sequences (RSS) of the Ig and T cell receptor (TCR). This is followed by the repair of these DSBs by NHEJ, resulting in the diversification of antigen receptor genes [17]. A key step during B cell maturation is isotype switching of immunoglobulins (Igs) through CSR. This process relies on the formation and repair of DSBs induced by activation-induced (cytidine) deaminase (AID) at conserved motifs within the switch (S) regions, which are upstream of the constant regions of antibody heavy chains [18]. Two S regions are joined by NHEJ upon break formation, resulting in loss of intervening DNA between the S regions, and consequently a change in the class of Igs expressed by a B cell.

An important step in response to DSBs is the activation of the ATM kinase. ATM is recruited to DSBs by the MRE11-RAD50-NBS1 (MRN) complex, where it initiates a signaling cascade by phosphorylating histone variant H2A.X. Phosphorylated H2AX, also referred to as γH2AX, then serves as a platform to load MDC1 [19]. The ATM-dependent phosphorylation of MDC1 leads to the recruitment of the E3 ubiquitin ligase RNF8, which ubiquitylates histone H1 and allows the accrual of another E3 ubiquitin ligase RNF168. RNF168 further decorates the DSB-flanking chromatin with H2AK15ub. The simultaneous association of 53BP1 to the RNF8/RNF168 dependent H2AK15ub, and H4K20me2 via its tudor domain, results in 53BP1 dependent cNHEJ [20]. While high levels of H4K20me2 in G1 cells promote 53BP1 recruitment, this mark is gradually diluted during S-phase and lowers 53BP1 binding at replicated areas of the genome, leading to a switch in HR [21]. On the other hand, the BRCA1-associated RING domain protein 1 (BARD1) recruits BRCA1 to DSBs by engaging H2AK15ub. The binding of BARD1 to H2AK15Ub and unmethylated H4K20 via its BRCT-domain-associated ubiquitin-dependent recruitment motif (BUDR) and Ankyrin repeat domains respectively, promotes BRCA1-BARD1-dependent DSB repair via HR during replication (Figure 2 and [22]).
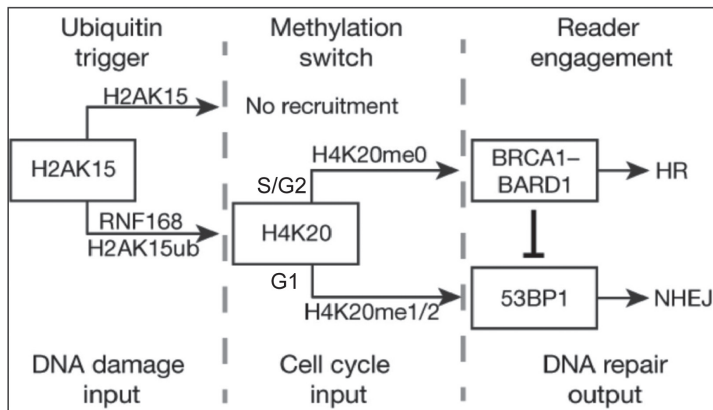
Figure. 2 DSB repair pathway choice by post-translational modifications (PTMs). Combinatorial H2AK15 and H4K20 PTMs govern the choice of the DSB repair pathway. On one hand, the DNA damage mediated recruitment of RNF168 results in the ubiquitylation of H2AK15 and recruitment of BRCA1-BARD1 to promote HR. On the other hand, high levels of H4K20me2 in G1 promote the recruitment of reader protein 53BP1 to mediate NHEJ. The simultaneous dilution of H4K20me2 during S/G2-phase leads to a switch in HR [22].

DSB repair in the context of chromatin

Efficient detection and repair of DSBs is complicated by the packaging of DNA into chromatin. ATP-dependent chromatin remodelers and post-translational modifications (PTMs) are therefore required to change chromatin structure at DSBs to facilitate repair [23, 24]. Over the past years it has become clear that besides the PTMs that are involved in regulating 53BP1 and BRCA1-BARD1 recruitment, several other PTMs and structural chromatin changes driven by ATP-dependent chromatin remodelers affect DSB repair. One key protein involved in circumventing the chromatin barrier during cNHEJ is PARP1. PARP1 is rapidly recruited to DSB sites where it promotes the local expansion of chromatin by the formation of PAR-chains on itself and its target proteins, including histones [25]. This results in the accumulation of DNA repair proteins which bind to PAR-chains directly [26] or to the exposed DNA, as was observed for several chromatin remodelers and cNHEJ factors that directly associate with DNA [27, 28]. Importantly, following DSB repair, the chromatin structure must be restored to its original state.

Chromatin remodeling is important for both HR and NHEJ, as it allows for the recruitment to DNA damage sites of factors involved in both these repair processes [29-31]. During short-range resection, chromatin is extensively remodeled to facilitate the loading of end-resection factors MRE11 and CtIP [32]. This is followed Fun30/

SMARCAD1 chromatin remodeling-driven long-range resection by the EXO1 and DNA2 nuclease. During this process, it is the Fun30/SMARCAD1-depent chromatin remodeling that counteracts the suppressive effect of 53BP1 on end-resection [33-35]. The Chromodomain Helicase DNA-binding (CHD) class of chromatin remodelers have also been implicated in DSB repair. Amongst these are CHD2 and CHD7, which have been shown to promote NHEJ [29, 30, 36]. CHD2 is recruited to sites of DNA damage by binding to PAR moieties associated with PARP1. Subsequently, it promotes DSB-induced chromatin expansion and deposits histone variant 3.3 to facilitate the efficient recruitment and proper functioning of the NHEJ machinery [30]. CHD7, on the other hand, does not bind PAR moieties but is recruited to DSBs via the PARP1-induced relaxation of damaged chromatin. CHD7 stimulates further chromatin relaxation and recruits HDAC1/2 for local chromatin de-acetylation and compaction. This allows for the efficient loading and correct positioning of KU and XRCC4/LIG4, thereby facilitating cNHEJ. Loss of CHD7 reinforces 53BP1 and shifts repair to mutagenic aNHEJ [29]. While the interplay between chromatin modulators and HR and cNHEJ has gained more insight over the last years, it remains unknown how they regulate aNHEJ and SSA. Future endeavors are required to understand the combinatorial effect of chromatin remodelers and histone modifications on DSB repair, particularly during aNHEJ and SSA.

DSB repair and transcription

DNA double strand breaks (DSBs) not only affect chromatin structure, but also the impact on other DNA-dependent processes such as transcription. Given the fact that DSBs occur in both inactive and actively transcribed regions, it is of utmost importance that transcription and DSB repair are coordinated properly in order to prevent collisions between the two processes [37]. Several studies have observed a direct link with transcription repression and DSB repair. In fact, HR and NHEJ factors are implicated in the repression of transcription upon DSBs. One such player is PARP1, which targets the RNAPII complex for PARylation in response to DSBs. This triggers an interaction between RNAPII and the NELF-E subunit of the negative elongation factor (NELF) complex, which leads to NELF-E dependent repression of transcription elongation in the vicinity of DSBs to promote efficient DSB repair [38]. Moreover, PARylated PARP1 is also responsible for the recruitment of KDM5A, which induces demethylation of H3K4me3. In turn, this modification allows for the recruitment of ZMYND8, which binds PAR and H4K16Ac to promote transcriptional repression and HR by facilitating the accrual

of the repressive NuRD complex [39]. PARP1 also functions with TIMELESS to recruit the E3 ubiquitin ligase complex FRRUC to DSBs. On one hand, FRUCC represses transcription by inducing the monoubiquitylation of H2AK119, while on the other hand it promotes the exchange of H2A with H2A.Z [40]. Two members of the PI3K kinase family, ATM and DNA-PK, have been shown to play crucial roles in the suppression of transcription upon DSBs. ATM dependent phosphorylation of transcription elongation factor ENL affects the accumulation and/or retention of Ku70 at DSBs [41]. Moreover, ATM can repress transcription in an RNF8/RNF168 dependent manner. RNF8 triggers the assembly of K11 linked ubiquitin moieties on histones H2A/H2AX in coordination with the Ube2S/C E2-conjugating enzymes to induce transcription silencing at DSBs by enforcing RNAPII pausing and repression of transcription of genes in close vicinity of a break [42-43]. On the other hand, ATM can also repress transcription in response to DSBs in a manner independent of RNF8/RNF168. This involves the ATM dependent phosphorylation of the PBAF subunit BAF180, which is crucial for the mono-ubiquitylation of H2A on K119, thereby promoting transcription silencing and NHEJ repair of DSBs [44]. More recently, the HECT E3 ubiquitin ligase WWP was identified as a critical mediator of DNA-PKcs-dependent transcription silencing at DSBs [24]. DNA-PK complex (consisting of Ku70/Ku80 and DNA-PKcs) and WWP2 are recruited to DSBs in actively transcribed genes. WWP2, in turn, modifies the RNAPII subunit RPB1 by K48 polyubiquitylation via DNA-PK. This triggers the recruitment of the proteasome at DSB sites and targets RNAPII for proteasomal degradation. Finally, this prevents collisions between the NHEJ factors and the transcription machinery, promoting efficient retention of the DNA-PK complex to repair the DSBs by NHEJ [24].

While it is evident that transcription becomes silencing occurs at DSBs, transcription termination is another crucial process for transcription regulation at DSBs. This process involves the combined action of several different enzymes that target nascent RNAs, leading to transcription termination and RNAPII removal from chromatin. The 5'-3' exoribonuclease XRN2 and the RNA/DNA helicase Senataxin (SETX) are examples of these enzymes [45]. SETX interacts with RNAPII and thereby promotes the cleavage of nascent RNA at poly(A) sites in a manner dependent on XRN2. This in turn, allows for the release of the pre-mRNA and the degradation of the remaining chromatin-associated RNA [46].

While transcription silencing and transcription termination at DSBs is better understood, mechanisms that promote transcription restart following the completion of DSB repair are still very poorly understood. Given that chromatin structure changes dramatically

surrounding DSB sites [47, 48], restoring the pre-existing chromatin state is foremost important to restart transcription following DSB repair. For instance, the removal of H2AK119Ub by the deubiquitinase USP16 is crucial for transcription restart after DSB repair [43]. Counterintuitively, WWP2-DNA-PKcs trigger the proteasomal degradation of RNAPII [24]. Further studies are needed to resolve how restart of transcription is achieved following RNAPII pausing and/or degradation and the completion following DSB repair in different chromatin contexts.

Replication stress

DNA replication

The ability of cells to survive and proliferate depends on the faithful duplication of their DNA via DNA replication during each cell cycle. Inaccurate or incomplete DNA replication can result in genomic instability, cell death and cancer. In eukaryotes, replication origins are licensed through the assembly of the pre-replication complex (pre-RC). This process is dependent on the six-subunit origin recognition complex (ORC), which provides a platform for the loading of the minichromosome maintenance complex (MCM) in a Cdc6- and Cdt1-dependent manner in G1. Upon S-phase entry, cyclin-dependent kinases (CDKs), and Dbf4-Cdc7 (DDK) activate a subset of replication origins. This is followed by the formation and activation of the CMG (CDC45-MCMs-GINS) helicase complex, which unwinds the DNA duplex, forming a replication bubble in which DNA synthesis is initiated by the replication machinery [49]. DNA replication can be initiated stochastically at thousands of individual replication origins, which form bidirectional forks. It is of utmost importance that cells balance accuracy and speed, as well as the consumption of resources such as nucleotides and replication factors, to complete DNA replication. Therefore, the firing of replication origins is strictly regulated and divided into early replication origins and late replicating origins. Also, not all licensed origins fire during an unperturbed S-phase, but can be activated following replication stress to ensure the completion of DNA replication [3, 50].

ATR signaling

Replication stress (RS) is defined as the stalling or slowing of replication fork progression or DNA synthesis. Several different sources can induce RS, including hydroxyurea (HU)-induced uncoupling of the MCM helicases from the replicative polymerases, leaving behind nicks or gaps with stretches of single-stranded DNA (ssDNA). DNA lesions caused

by UV light or chemical mutagens, DNA secondary structures, replication-transcription conflicts (TRCs), common fragile sites (CFS) and oncogene-induced replication stress are common DNA polymerase-blocking lesions [3]. ATR is the major kinase responding to RS, which coordinates a wide range of cellular processes including preserving the RPA pool, suppression of new origin firing, preventing fork collapse and promoting fork restart at stalled replication forks before cells enter mitosis (Figure 3 and [51]). RPA-coated ssDNA generated by DNA polymerase-helicase uncoupling during fork stalling triggers the activation of ATR by its binding partner ATRIP which binds RPA directly. This results in the recruitment of the ATR activating protein TOPBP1, which is recruited to ssDNA-dsDNA 5' junctions through its interaction with the MRE11-RAD50-NBS1 (MRN) complex and the 9-1-1 (RAD9-RAD1-HUS1) complex [52]. TOPBP1 also interacts with RHINO, and this interaction is required for the full activation of ATR upon replication stress [53]. On the other hand, ATR can be activated via ETAA1, which is, unlike TOPBP1, recruited to RPA-coated ssDNA via a direct interaction with RPA. This interaction is required for the activation of ATR via its ATR activation domain (AAD) and the subsequent activation of the S-phase checkpoint kinase CHK1 [54].

Fork reversal

Protection of ssDNA by RPA and checkpoint signaling is not sufficient to prevent fork collapse under prolonged replication stress conditions. In fact, RPA is highly dynamic on ssDNA and is displaced in a timely manner by RAD51, thereby promoting replication fork reversal [55]. During fork reversal the three-way fork junction is migrated backward to displace and anneal the nascent DNA strands to form a 'chicken foot' structure (Figure 3). This process is catalyzed by several enzymes including the SNF2 family of chromatin remodelers SMARCAL1, ZRANB3 and HLTF, as well as the RecQ helicases BLM and RECQ1 [54]. SMARCAL1 travels with the replication fork and remodels forks with persistent ssDNA gaps into reversed forks in an ATR dependent manner. ZRANB3 interacts with polyubiquitinated PCNA to promote fork remodeling and HLTF affects fork reversal directly via its HIRAN domain [56]. Fork reversal has the following benefits on replication fork fidelity: 1) backtracking and annealing of nascent DNA strands prevents replication fork collapse, 2) preventing the formation of excessive amounts of ssDNA and allowing more time for the repair of replication impediments which can be repositioned back on the dsDNA template, and 3) generating a Holliday junction with a single-ended DSB that can be repaired via HR [57]. On the other hand, reversed forks

are common substrates for nucleolytic degradation by several nucleases such as MRE11, CtIP, DNA2 or EXO1, which act predominantly in the absence of fork protection factors RAD51, BRCA1 and BRCA2. This ultimately leads to nascent DNA strand shortening, fork collapse and chromosomal aberrations [56].

DNA damage tolerance (DTT) pathways

A pathway acting in parallel to fork reversal is PrimPol dependent repriming. PrimPol is an RNA/DNA primase that is recruited to replication forks via its interaction with RPA coated ssDNA under conditions of fork stalling induced by UV irradiation or HU treatment. Primpol leaves behind ssDNA gaps during repriming on the leading strand, which are enlarged by MRE11 and EXO1. This in turn, results in the loading of RAD51 to promote HR-mediated post-replicative repair (PRR) (Figure 3 and [58, 59]). Importantly, depletion of RAD51 or fork reversal factors such SMARCAL1 or HLTF leads to PrimPol mediated repriming on the leading strand (Figure 3 and [58, 59]). Thus, the balance between fork reversal factors such as RAD51-SMARCAL1 and PrimPol, determines the choice between fork reversal and repriming and thereby controls the fate of the replication fork. When the replication fork obstacle (usually a DNA lesion) is present on the lagging strand, the replication machinery can bypass these DNA lesions by using so-called DNA damage tolerance (DDT) pathways, which comprise translesion synthesis (TLS) and template switching (TS) (Figure 3). During TLS, replication across a DNA lesion is resumed by low fidelity DNA polymerases including Y-family polymerases (POLH, POLI, POLK and REV1), B-family polymerase Pol ζ and A-family polymerases POLQ and POLN. Each polymerase has the ability to recognize a specific type of DNA lesion, which can either be bypassed in a relatively error-free mode, for example by POLN, or via an error-prone mechanism by Pol ζ and REV1 [60]. TS is the more faithful DDT pathway in which the stalled replication fork uses the newly synthesized strand as a template to avoid the lesion via pathways that are less well understood [61]. The choice between the DDT pathways is regulated by PCNA, the replicative sliding clamp that is essential for DNA replication. In response to RS, PCNA is either mono- or polyubiquitylated on Lysine 164 (K164). TLS is mediated by the RAD18 dependent mono-ubiquitylaton of PCNA for the loading of TLS polymerases, whereas TS is regulated by HLTF and SHPRH dependent polyubiquitylation of PCNA [62]. However, the exact mechanism by which the PCNA ubiquitylation status triggers TLS or TS remains unknown.
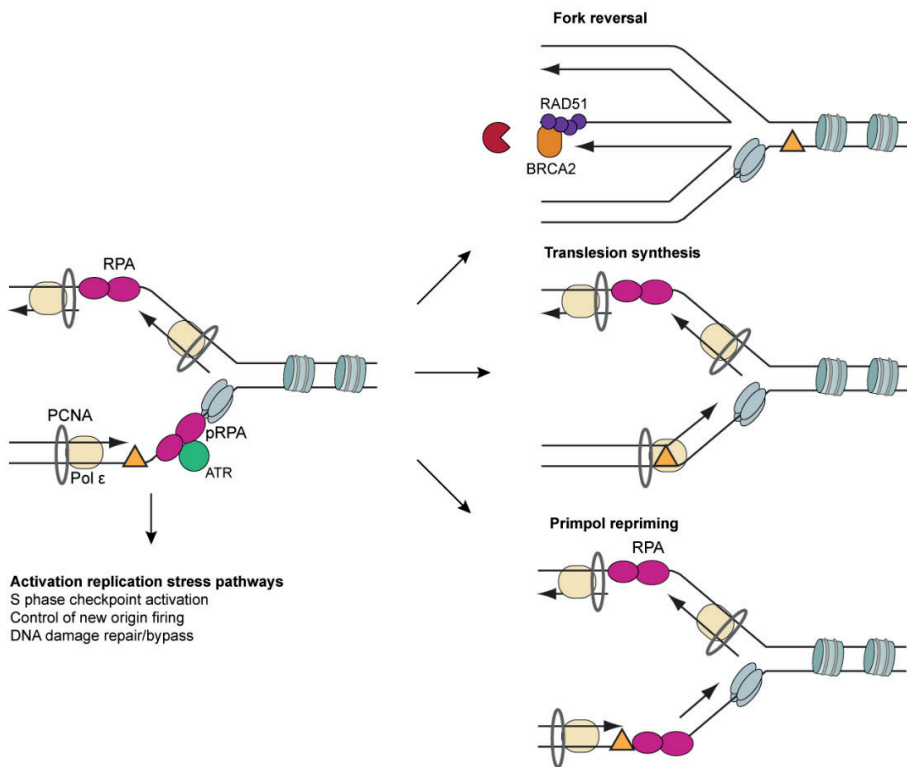
Figure. 3 Fork repair mechanisms. Polymerase-helicase uncoupling leads to the activation of the S-phase checkpoint which prevents new origin firing. Alternatively, fork reversal, translesion synthesis (TS), fork repriming may act to overcome DNA lesions caused by polymerase-helicase uncoupling.

Break induced replication (BIR)

The resumption of DNA replication at stalled forks is mediated by several different mechanisms. One of the most favorable mechanisms of fork restart is the resumption of semiconservative DNA synthesis supported by the CMG helicase. However, when the CMG helicase is disassembled or lost, a collapsed fork is prone to cleavage by the MUS81 nuclease resulting in a one-ended DSB [63]. This is followed by resection and strand invasion of the DNA end using the undamaged opposite strand, which can resume DNA synthesis using a modified replisome on a migrating D-loop during a process called break induced replication (BIR). The molecular mechanism of BIR is, however, poorly understood in mammalian cells. While canonical BIR relies on RAD51, several

studies have reported RAD51-independent BIR in mammalian cells, which relies on RAD52 and may be occurring at common fragile sites with large amounts of RS [54]. Furthermore, BIR occurs during mitotic DNA synthesis (MiDAS). MiDAS completes DNA synthesis at replication origin poor chromosomal regions referred to as 'common fragile sites' during mitosis in a MUS81-dependent manner. BIR is considered an error-prone pathway, which can lead to genomic rearrangements such as the formation of tandem duplications and copy number variations. Moreover, defects in BIR can also lead to impaired restart of reversed replication forks in the absence of BRCA2 [64]. Future work is required to better understand the genomic context in which BIR occurs, and which factors drive this important process.

Transcription-replication conflicts

Co-directional and head-on transcription-replication conflicts

During the S-phase of each cell cycle, eukaryotic cells are the most vulnerable as the replication and transcription machineries co-exist on the same DNA template. It is inevitable that the two machineries could interfere with each other, giving rise to transcription-replication conflicts (TRCs). Depending on the orientation of transcription relative to the direction of the replication forks, TRCs may occur in a head-on (HO) or co-directional (CD) orientation (Figure 4). During HO collisions, the transcription machinery and replication machinery progress in opposite directions. CD collisions, on the other hand, are caused by replisomes that encounter RNA polymerases that progress in the same direction [65]. Studies in bacteria and eukaryotes have shown that HO collisions induce pausing and blockage of the replication fork to a greater extent than CD collisions do. Consequently, HO collisions are considered more toxic.

Several different scenarios have been suggested for how TRCs can occur. For instance, the MCM2-7 helicase unwinds the parental DNA ahead of the replisome. This could lead to a frontal clash of the helicase with the RNA polymerase and may lead to the inactivation of the helicases [66]. Furthermore, transcription and replication require unwinding of the DNA. During transcription, this leads to positive and negative supercoiling upstream and downstream of the RNA polymerase causing torsional stress which is relieved by DNA topoisomerase enzymes. Indeed, depletion of Topoisomerase I in mammalian cells increases negative supercoiling in actively transcribed regions which results in replication fork stalling and DNA breaks [67]. In addition to torsional stress, non-B structures on DNA are also common sources of TRCs. Such non-B DNA structures are formed at a

particular DNA sequence (e.g. repetitive sequences or at common fragile sites) and exist in the form of hairpins, triplex DNA, G-quadruplexes or alternatively as RNA-DNA hybrids (R-loops) [65].
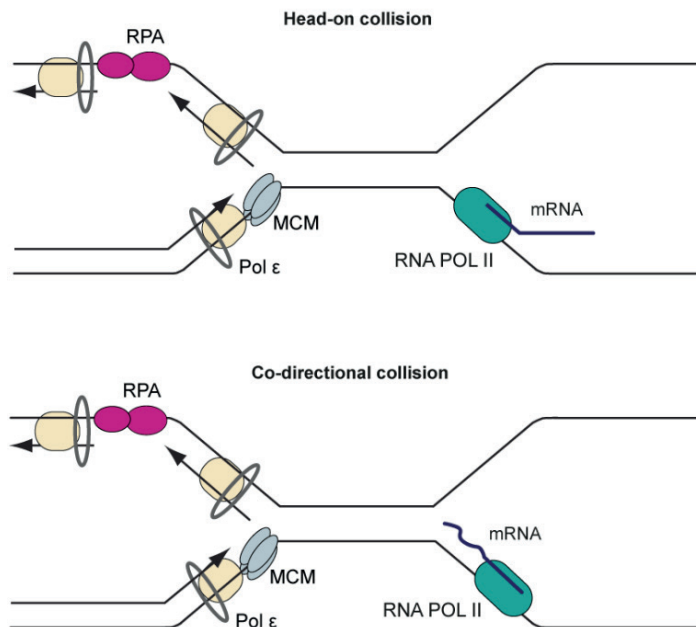

Fig. 4 Transcription-replication conflicts (TRCs). Progression of RNA polymerase (RNAPII) and a replication fork in the opposite direction leads to head-on collisions (HO), which induce pausing and blockage of the replication fork, and may lead to fork collapse and the formation of DNA breaks. Progression of RNAPII and a replication fork in the same direction leads to co-directional (CD) collisions if the fork moves faster than RNAPII. CD collisions can be resolved by the displacement of RNAPII from DNA.

RNA-DNA hybrids

R-loops are important players in a number of physiological processes. These structures are formed due to nascent RNA base pairing with its template DNA behind the RNA polymerase leaving behind the non-templated strand as ssDNA. Their presence in the genome is required during DNA replication at the start of leading strand synthesis, as well as during Okazaki fragment synthesis at the lagging strand. In addition, R-loops can also be a source of genome instability. Consequently, the factors that resolve or prevent their formation such as RNA binding proteins, helicases, nucleases and DNA replication and repair factors, are of utmost importance for genome maintenance [68].

Genome wide mapping studies revealed that R-loops are primarily located at promoters of RNA polymerase II (RNAPII) transcribed genes. Here, they activate transcription by the inhibition of transcriptional repressors such as DNA-methylation enzymes (DNMT), or act as binding sites for transcription factors and chromatin remodelers. Alternatively, R-loops can also repress transcription by the induction of repressive chromatin marks [69]. Furthermore, R-loops are also involved in efficient transcription termination by stalling RNAPII downstream of the poly-adenylation tail. This is followed by the recruitment of the R-loop helicases senataxin (SETX) and DHX9 which resolve R-loops and promote the release of nascent RNA for degradation by the exonuclease XRN2 leading to transcription termination [68].

Unscheduled R-loop formation can be a source of genomic instability as the ssDNA generated during this process can be susceptible to DNA damage. Furthermore, when R-loops are not removed properly they can cause TRCs. Recent studies on an engineered bacterial genome and a mammalian episomal system have shed light on the orientation of TRCs and how they influence the levels of R-loops [70, 71]. Interestingly, both studies revealed that HO collisions enhance R-loops formation due to increased topological stress, whereas CD collisions contribute to resolving R-loops, the latter which occurs via a mechanism that remains largely unknown. R-loops can also cause transcriptional stress by the excessive stalling, arrest or backtracking of RNAPII. It is unclear what distinguishes R-loops that promote transcription from those that cause transcription stress. One possibility is that regulatory R-loops are rapidly degraded in cells and therefore do not persist long enough to block RNAPII [72].

Cells have evolved several mechanisms to prevent or remove unscheduled R-loops in order to maintain genomic stability. One such mechanism involves the prevention of newly transcribed RNA to hybridize with DNA. This mechanism is driven by mRNA binding proteins (RBPs), which promote the processing and export of the RNA, thereby impeding hybrid formation [73]. Indeed, mutants of the THO/TREX protein complex involved in transcription and mRNA export, initially identified in yeast, showed increased R-loop-mediated genomic instability [74]. Another mechanism is the relieve of DNA supercoiling at transcription sites. During this process, specialized topoisomerase enzymes such as TOP1 prevent the formation of R-loops by relieving negative supercoiling behind RNA polymerases [67]. More recently, chromatin factors have emerged as important players in R-loop homeostasis. An interaction between the THO and SIN3A histone deacetylase complex (HDAC) was described to regulate acetylation of histones to prevent

the accumulation of harmful R-loops [75]. Furthermore, depletion of the SWI/SNF complex member BRG1 in mammalian cells increases R-loop dependent DNA breaks and TRCs [76]. Besides factors that suppress R-loop formation, cells are also equipped with enzymes that directly remove R-loops such as RNase H1 and RNase H2 that cleave the RNA strand in the hybrid. In addition, Rnase H2 is also involved in the removal of mis-incorporated ribonucleotides in a process called ribonucleotide excision repair (RER) [77]. While Rnase H1 is mainly involved in the removal of R-loops following replication stress, Rnase H2 is required to remove R-loops in a post-replicative manner [78]. In addition, several helicases are involved in resolving R-loops, either by translocating along the RNA strand, unwinding double-strand regions and displacing proteins, or by acting locally at R-loops [79]. Examples of such helicases are DDX5, which is stimulated by BRCA2, aquarius (AQR), and Senataxin (SETX) [80]. *In vitro* studies of Sen1, the yeast ortholog of SETX, revealed that this protein displays 5' to 3' translocation activity, as well as 5' to 3' unwinding properties along DNA and RNA. Loss of RNase H and SETX activity was shown to cause R-loop-induced replication stress and genomic instability [73]. Interestingly, ectopic expression of Rnase H enzymes supresses replication stress in cells accumulating R-loops, which indicates that R-loops interfere with replication fork progression [70]. Furthermore, R-loops can be processed by nucleotide excision repair (NER) factors XPF and XPG. While it remains unclear how XPF and XPG recognize and process R-loops, one possibility may be that they may recognize and cleave both strands of an R-loop, removing the R-loop and generating a DSB. Another possibility could be that the transcription-coupled NER (TC-NER) machinery is recruited to R-loops during DNA replication and generates a ssDNA gap that is converted into a DSB [81].

R-loops and DSB repair

Several recent studies have shown that R-loops can also form at DSBs, where they can both inhibit and promote repair. For instance, the loss of RNase H enzymes in budding yeast is associated with an increase in RAD52 foci, indicative of persistent breaks [82]. In contrast, a study in fission yeast showed that R-loops are enriched at *I-Ppol* induced DSBs, and that both overexpression and deletion of Rnase H impair HR, indicating the importance of a temporary R-loop near the break site for efficient repair [83]. To this end, an important observation is the fact that R-loops may impact DNA end resection. In budding yeast, persistent R-loops block resection at the break site adjacent to the R-loop [84]. On the other hand, factors that promote resection, including Sae2 in budding

yeast and its ortholog CtIP in mammalian cells, have been shown to promote R-loops resolution, illustrating the complex relationship between R-loops and resection [85].

R-loops also contribute to an alternative type of HR in which a homologous RNA molecule instead of a homologous DNA molecule is used as the template for DSB repair [86]. There is evidence that this type of HR occurs in G0/G1 cell cycle phase of post-mitotic neurons. During this process, R-loops may guide RNA-templated HR by promoting RAD52 recruitment to DSBs [87]. Although the usage of RNA as a template for HR has been observed in both yeast and human cells, the physiologic relevance of R-loops herein needs to be resolved [86]. Interestingly, DNA damage-induced long non-coding RNAs (dilncRNAs) which are transcribed from broken DNA ends have also been implicated in R-loop formation [88]. dilncRNAs pair to the resected DNA ends during S/G2 phase and form R-loops, which are recognized by BRCA1. This is followed by BRCA2-mediated recruitment of RNase H2, which degrades R-loops to ensure efficient HR-mediated repair [89]. Finally, R-loops can also be generated as a consequence of transcription inhibition induced upon DNA damage [37]. Indeed, pausing of RNAPII distal to the DNA damage site has been associated with increased R-loop accumulation [71, 90]. Altogether, these findings highlight the complexity of R-loops and their link to genome stability and pinpoint the importance of R-loop processing factors, genomic location and transcriptional status as determinants on how they affect DSB repair.

DSB repair, replication stress response and cancer therapy

Recent technological advances in the field (e.g. CRISPR/Cas9 based screens), have enabled us to explore synthetic lethal interactions relevant to DNA repair and replication in normal cells and cancer cells [91]. Synthetic lethality describes the cellular condition in which a defect in either one of two genes has little or no effect, but the combination of both gene defects results in death (synthetic lethality) or sickness (synthetic sickness) [92]. The ever-growing focus on exploring synthetic lethality has stemmed from the success of PARP inhibitors (PARPi) in the clinic for the treatment of tumors with mutations that disrupt BRCA1/BRCA2 genes or other HR factors. Mechanistically, PARPi does not only block PARP1 activity for the repair of DNA lesions such as single-stranded DNA breaks or DNA replication intermediates (e.g. unligated okazaki fragments), but also traps PARP1 on to damaged DNA [93]. PARPi-induced trapping of PARP1 on DNA possesses a block to the replisome and thereby causes cytotoxicity. In order to resolve these blocks and resume cell cycle progression, cells require functional HR. Defects in

HR lead to the persistence of DNA damage, thereby causing cell death. However, over the past years, studies in mouse models and patients have shown that PARPi responses are frequently associated with resistance. Intriguingly, PARPi resistance is often correlated with resistance to platinum-based agents such as cisplatin [94]. Several resistance mechanisms have been described which can be classified into four main categories: (1) influence of the cellular availability of the inhibitor, (2) direct impact on the activity and abundance of PAR chains, (3) reactivation of HR and (4) the influence of replication fork protection [12].

Long-term treatment of *Brca1*-deficient breast tumors in mouse murine models with PARPi displayed resistance due to overexpression of transporter genes *Abcb1a* and *Abcb1b* encoding P-glycoprotein efflux pumps. Indeed, coadministration of the P-glycoprotein inhibitor reversed their resistance to PARPi [95]. Upregulation of ABCB1 was also observed in chemotherapy-treated ovarian and breast cancer cells as a result of chromosomal translocations involving these genes [96]. However, whether these translocations are present in BRCA1/2 mutated tumors, and whether P-glycoprotein efflux pumps are upregulated in PARPi-resistant tumors needs to be addressed in the future.

PARylation induced by PARP1 is a reversible and transient modification of proteins by the covalent addition of PAR upon DNA damage [97]. While depletion of PARP1 is synthetic lethal with BRCA1/2 deficiency, the cytotoxicity caused by PARPi treatment is much larger than that following genetic depletion of PARP1 [98, 99]. Therefore, mutations in PARP1 that influence PARP1 trapping induce PARPi resistance, even in HR deficient cells [100]. A mechanism related to PARP inhibition involves the inhibition of PAR glycohydrolase (PARG), which is the enzyme that degrades PAR chains in cells. Genetic screens performed using *Brca2* mutated mouse mammary tumors revealed that loss of PARG represented a major PARPi resistance in these tumors. PARG loss partially restored PARylation in PARPi-treated cells, also diminished PARP1 trapping on DNA, and partially rescued DNA damage signaling caused by PARP1 [101].

The majority of HR-disrupting mutations such as in BRCA1/2 are frameshift mutations or single nucleotide mutations. The occurrence of secondary- or reversion mutations, would enable frame-restoration and potentially restore protein activity and HR. Interestingly, these reversion mutations are often showing a microhomology signature, suggesting that they result from the repair of DSBs via aNHEJ [102]. The identification of mutational signatures has allowed for the identification of several synthetic lethal interactions to

therapeutically exploit BRCA-deficient tumors. There are several examples of synthetic lethal interactions described in recent years. For instance, FANCD2 supports POLQ mediated TMEJ and protects stalled replication forks from degradation. Consequently, loss of FANCD2 leads to cell death in BRCA1/2 deficient cells [103]. More recently, the POLQ inhibitor ART558 was found to effectively target BRCA-defective cancers and even enhance the effect of the PARPi treatment [104]. Furthermore, FEN1 is involved in the 5'flap processing following aNHEJ. It also plays a role in long-patch SSB repair. FEN1 inhibition or downregulation results in increased replication stress due to compromised 5'flap processing and BER in BRCA-deficient cells, leading to cell death [105]. Another feature of PARPi resistance in BRCA1, but not BRCA2 deficient cells, is the loss of resection inhibitors by factors that regulate NHEJ including 53BP1 and its downstream factors in the Shieldin complex [106]. Interestingly, the POLQ inhibitor ART558 could overcome the PARPi resistance observed in 53BP1/Shieldin defective cells, due to DNA nuclease-mediated hyperresection [104]. In addition to 53BP1/Shieldin, CRISPR/Cas9 screens have enabled the identification for other suppressors of PARPi resistance in BRCA1 deficient cells, including DYNLL1 and its transcriptional activator ATMIN [107]. DYNLL1 associates with DNA end resection factors, including MRE11, and these factors to promote DNA end-resection. Indeed, loss of DYNLL1 leads to increased resection in BRCA1 deficient cells and thereby causing PARPi resistance [107].

Finally, the restoration of replication fork protection also serves as a mechanism for PARPi resistance. In addition to their role in HR, BRCA1/2 also protect stalled replication forks [108]. BRCA1/2 loss leads to nascent strand shortening, fork collapse and chromosomal aberrations due to nucleases such as MRE11 and MUS81, which act upon the deprotected fork. Secpifically, EZH2 and PTIP were found to mediate PARPi sensitivity by recruiting MUS81 and MRE11 to stalled forks [109, 110]. In addition, RADX is involved in the stabilization and resolution of stalled forks which prevents MUS81-dependent fork cleavage and inhibits RAD51 mediated fork remodeling [111]. Fork remodeling is another process required for MRE11-dependent degradation of nascent DNA in BRCA1/2 deficient human cells, and depletion of fork reversal factors SMARCAL1, ZRANB3 and HLTF cause PARPi resistance [112].

More recently, another mechanism of BRCA1/2 deficiency and synthetic lethality has been described. This mechanism underlies the accumulation of ssDNA gaps in BRCA1 deficient cells, a phenomenon that was found to correlate with PARPi sensitivity. Interestingly, LIG3 loss promotes the formation of MRE11 mediated post-replicative

ssDNA gaps in BRCA1 deficient and BRCA1/53BP1 double deficient cells. Moreover, LIG3 depletion enhances the efficacy of PARPi treatment in Brca1-deficient mammary tumors in mice, suggesting LIG3 as a potential therapeutic target [113, 114]. CRISPR/Cas9 screens also allowed the identification of the PAR-binding chromatin remodeler ALC1/CHD1L as a key determinant of PARPi toxicity in HR-deficient cells. ALC1 loss results in decreased chromatin accessibility, thereby impairing the loading of base excision repair (BER) factors. This results in the accumulation of replication associated unrepaired ssDNA gaps, increased PARP trapping and a critical dependence on HR repair. Therefore, targeting ALC1 alone or as a PARPi sensitizer could be specifically employed in HR deficient tumors.

AIMS AND OUTLINE

The genetic information in our cells has to be properly protected in order to allow its faithfull duplication prior to it being transferred to daughter cells during cell division. This process is challenged by the continuous induction of DNA damage by exogenous and endogenous sources. To circumvent adverse effects of DNA damage, cells activate a DNA damage response (DDR) that consist of pathways that detect and repair lesions, as well as pathways leading to the activation of checkpoints that delay cell cycle progression to allow for repair of DNA damage. Insight into the regulation of repair and signaling cascades, particularly those that counteract DSBs and RS, may improve our understanding of the DDRs and how their dysfunction causes devastating diseases such as neurodegenerative disorders and cancer [4].

In **Chapter 1** this thesis I provide an overview of DDR mechanisms with a focus on DSB repair and RS. This chapter describes the core factors involved in HR, cNHEJ, aNHEJ and SSA and how these mechanisms are key in repairing DSBs and preventing genome instability. Furthermore, it describes how DSB repair pathway choice is regulated in the context of chromatin. This is followed by a description of how DSB repair is linked with transcription regulation at sites of DSBs. Moreover, this chapter also describes mechanisms that help to preserve DNA replication fork integrity in response to RS, with a particular emphasis on the role transcription-replication conflicts therein. Finally, this chapter describes the link between these pathways and cancer, focus on their exploitation in cancer therapy.

In **Chapter 2**, I highlight the importance of zinc-finger domain-containing proteins in a diverse range of cellular processes, including DSB repair. ZnF-domains exist in 5% of all

human proteins and bind to a large variety of substrates such as DNA, RNA and several PTMs including PAR, ubiquitin and SUMO [28]. Importantly, ZnF-domain containing proteins have been shown to be important players in DSB repair, demonstrating their importance as caretakers of the genome and beyond.

In **Chapter 3**, I describe a new role of ZnF-protein ZNF384 in DSB repair. ZNF384 is a member of the C2H2 family of ZnF proteins which binds DNA ends *in vitro* and is recruited to DSBs *in vivo*. The accumulation of ZNF384 at DSBs requires PARP1-dependent chromatin expansion, which allows ZNF384 to bind to the exposed DNA via its C2H2 motifs. ZNF384 interacts with core NHEJ proteins Ku70/Ku80 via its N-terminus domain, thereby promoting Ku70/Ku80 assembly and the accrual of downstream factors including APLF and XRCC4/LIG4. In conclusion, this chapter highlights the role of ZNF384 as a 'Ku-adaptor' that binds broken DNA and Ku70/Ku80 to facilitate cNHEJ.

In **Chapter 4**, I focus on the molecular basis of the non-specific lethal (NSL) chromatin remodeling complex in R-loop homeostasis, and I specifically describe the role of the NSL complex member KANSL3 in this process [76]. KANSL3 protects cells from replication stress by promoting the restart of stalled replication forks, which may arise due to the formation of R-loop -induced TRCs. Indeed, KANSL3 has a suppressive role on R-loop formation, thereby preventing collisions between the replication fork and transcription machinery. Future work is required to gain insight into how KANSL3 affects R-loop formation and consequently replication fork stability/recovery following stress. Future work is required to gain mechanistic insight into how the dual role of KANSL3 protects genomic integrity.

In **Chapter 5**, I focus on the RNA processing enzyme ERI1. ERI1 is an evolutionary conserved 3'-5' exonuclease that participates in several RNA processing pathways, such as those in ribosomal RNA processing and the degradation of histone mRNAs. Here, I describe a novel function of ERI1 in suppressing R-loop formation and preserving replication fork stability. Together, my findings unveil a novel role of the RNA-binding and RNA-processing protein ERI1 in protecting cells form R-loop induced replication fork instability.

Finally, In **Chapter 6** I discuss how these findings contribute to our understanding of genome surveillance mechanisms, highlighting unresolved issues and making recommendations for follow-up studies.

REFERENCES

1. Borde, V., A.S. Goldman, and M. Lichten, *Direct coupling between meiotic DNA replication and recombination initiation.* Science, 2000. **290**(5492): p. 806-9.

2. Ciccia, A. and S.J. Elledge, *The DNA damage response: making it safe to play with knives.* Mol Cell, 2010. **40**(2): p. 179-204.

3. Zeman, M.K. and K.A. Cimprich, *Causes and consequences of replication stress.* Nat Cell Biol, 2014. **16**(1): p. 2-9.

4. Jackson, S.P. and J. Bartek, *The DNA-damage response in human biology and disease.* Nature, 2009. **461**(7267): p. 1071-8.

5. Stinson, B.M., et al., *A Mechanism to Minimize Errors during Non-homologous End Joining.* Molecular Cell, 2020. **77**(5): p. 1080-1091.e8.

6. Chang, H.H.Y., et al., *Non-homologous DNA end joining and alternative pathways to double-strand break repair.* Nat Rev Mol Cell Biol, 2017. **18**(8): p. 495-506.

7. Schimmel, J., et al., *Mutational signatures of non-homologous and polymerase theta-mediated end-joining in embryonic stem cells.* Embo j, 2017. **36**(24): p. 3634-3649.

8. Bhargava, R., D.O. Onyango, and J.M. Stark, *Regulation of Single-Strand Annealing and its Role in Genome Maintenance.* Trends Genet, 2016. **32**(9): p. 566-575.

9. Huertas, P. and S.P. Jackson, *Human CtIP mediates cell cycle control of DNA end resection and double strand break repair.* J Biol Chem, 2009. **284**(14): p. 9558-65.

10. Orthwein, A., et al., *A mechanism for the suppression of homologous recombination in G1 cells.* Nature, 2015. **528**(7582): p. 422-6.

11. Arnoult, N., et al., *Regulation of DNA repair pathway choice in S and G2 phases by the NHEJ inhibitor CYREN.* Nature, 2017. **549**(7673): p. 548-552.

12. Noordermeer, S.M. and H. van Attikum, *PARP Inhibitor Resistance: A Tug-of-War in BRCA-Mutated Cells.* Trends Cell Biol, 2019. **29**(10): p. 820-834.

13. Mirman, Z., et al., *53BP1–RIF1–shieldin counteracts DSB resection through CST- and Polα-dependent fill-in.* Nature, 2018. **560**(7716): p. 112-116.

14. Schimmel, J., et al., *Small tandem DNA duplications result from CST-guided Pol α-primase action at DNA break termini.* Nat Commun, 2021. **12**(1): p. 4843.

15. Leung, J.W., et al., *ZMYM3 regulates BRCA1 localization at damaged chromatin to promote DNA repair.* Genes Dev, 2017. **31**(3): p. 260-274.

16. Alt, Frederick W., et al., *Mechanisms of Programmed DNA Lesions and Genomic Instability in the Immune System.* Cell, 2013. **152**(3): p. 417-429.

17. Deriano, L. and D.B. Roth, *Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage.* Annu Rev Genet, 2013. **47**: p. 433-55.

18. Dong, J., et al., *Orientation-specific joining of AID-initiated DNA breaks promotes antibody class switching.* Nature, 2015. **525**(7567): p. 134-139.

19. Ceccaldi, R., B. Rondinelli, and A.D. D'Andrea, *Repair Pathway Choices and Consequences at the Double-Strand Break.* Trends Cell Biol, 2016. **26**(1): p. 52-64.

20. Fradet-Turcotte, A., et al., *53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark.* Nature, 2013. **499**(7456): p. 50-4.

1

21. Pellegrino, S., et al., *Replication-Coupled Dilution of H4K20me2 Guides 53BP1 to Pre-replicative Chromatin.* Cell Rep, 2017. **19**(9): p. 1819-1831.

22. Becker, J.R., et al., *BARD1 reads H2A lysine 15 ubiquitination to direct homologous recombination.* Nature, 2021.

23. Dantuma, N.P. and H. van Attikum, *Spatiotemporal regulation of posttranslational modifications in the DNA damage response.* Embo j, 2016. **35**(1): p. 6-23.

24. Caron, P., et al., *WWP2 ubiquitylates RNA polymerase II for DNA-PK-dependent transcription arrest and repair at DNA breaks.* Genes Dev, 2019. **33**(11-12): p. 684-704.

25. Jungmichel, S., et al., *Proteome-wide identification of poly(ADP-Ribosyl)ation targets in different genotoxic stress responses.* Mol Cell, 2013. **52**(2): p. 272-85.

26. Helfricht, A., et al., *Loss of ZBTB24 impairs non-homologous end-joining and class-switch recombination in patients with ICF syndrome.* Journal of Experimental Medicine, 2020.

27. Smith, R., et al., *Poly(ADP-ribose)-dependent chromatin unfolding facilitates the association of DNA-binding proteins with DNA at sites of damage.* Nucleic Acids Res, 2019. **47**(21): p. 11250-11267.

28. Singh, J.K. and H. van Attikum, *DNA double-strand break repair: Putting zinc fingers on the sore spot.* Semin Cell Dev Biol, 2020.

29. Rother, M.B., et al., *CHD7 and 53BP1 regulate distinct pathways for the re-ligation of DNA double-strand breaks.* Nature Communications, 2020. **11**(1): p. 5775.

30. Luijsterburg, M.S., et al., *PARP1 Links CHD2-Mediated Chromatin Expansion and H3.3 Deposition to DNA Repair by Non-homologous End-Joining.* Mol Cell, 2016. **61**(4): p. 547-562.

31. Khurana, S., et al., *A macrohistone variant links dynamic chromatin compaction to BRCA1-dependent genome maintenance.* Cell Rep, 2014. **8**(4): p. 1049-62.

32. Kari, V., et al., *Loss of CHD1 causes DNA repair defects and enhances prostate cancer therapeutic responsiveness.* EMBO reports, 2018. **19**(10): p. e46783.

33. Densham, R.M., et al., *Human BRCA1–BARD1 ubiquitin ligase activity counteracts chromatin barriers to DNA resection.* Nature Structural & Molecular Biology, 2016. **23**(7): p. 647-655.

34. Costelloe, T., et al., *The yeast Fun30 and human SMARCAD1 chromatin remodellers promote DNA end resection.* Nature, 2012. **489**(7417): p. 581-4.

35. Chen, X., et al., *The Fun30 nucleosome remodeller promotes resection of DNA double-strand break ends.* Nature, 2012. **489**(7417): p. 576-80.

36. Rother, M.B. and H. van Attikum, *DNA repair goes hip-hop: SMARCA and CHD chromatin remodellers join the break dance.* Philos Trans R Soc Lond B Biol Sci, 2017. **372**(1731).

37. Caron, P., J. van der Linden, and H. van Attikum, *Bon voyage: A transcriptional journey around DNA breaks.* DNA Repair (Amst), 2019. **82**: p. 102686.

38. Awwad, S.W., et al., *NELF-E is recruited to DNA double-strand break sites to promote transcriptional repression and repair.* EMBO reports, 2017. **18**(5): p. 745-764.

39. Gong, F. and K.M. Miller, *Double duty: ZMYND8 in the DNA damage response and cancer.* Cell Cycle, 2018. **17**(4): p. 414-420.

40. Rona, G., et al., *PARP1-dependent recruitment of the FBXL10-RNF68-RNF2 ubiquitin ligase to sites of DNA damage controls H2A.Z loading.* Elife, 2018. **7**.

41. Ui, A., Y. Nagaura, and A. Yasui, *Transcriptional Elongation Factor ENL Phosphorylated by ATM Recruits Polycomb and Switches Off Transcription for DSB Repair.* Molecular Cell, 2015. **58**(3): p. 468-482.

42. Paul, A. and B. Wang, *RNF8- and Ube2S-Dependent Ubiquitin Lysine 11-Linkage Modification in Response to DNA Damage.* Mol Cell, 2017. **66**(4): p. 458-472.e5.

43. Shanbhag, N.M., et al., *ATM-Dependent Chromatin Changes Silence Transcription In cis to DNA Double-Strand Breaks.* Cell, 2010. **141**(6): p. 970-981.

44. Meisenberg, C., et al., *Repression of Transcription at DNA Breaks Requires Cohesin throughout Interphase and Prevents Genome Instability.* Mol Cell, 2019. **73**(2): p. 212-223.e7.

45. Proudfoot, N.J., *Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut.* Science, 2016. **352**(6291): p. aad9926.

46. Skourti-Stathaki, K., Nicholas J. Proudfoot, and N. Gromak, *Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination.* Molecular Cell, 2011. **42**(6): p. 794-805.

47. Schep, R., et al., *Impact of chromatin context on Cas9-induced DNA double-strand break repair pathway balance.* Molecular Cell, 2021. **81**(10): p. 2216-2230.e10.

48. Clouaire, T., et al., *Comprehensive Mapping of Histone Modifications at DNA Double-Strand Breaks Deciphers Repair Pathway Chromatin Signatures.* Molecular Cell, 2018. **72**(2): p. 250-262.e6.

49. Técher, H. and P. Pasero, *The Replication Stress Response on a Narrow Path Between Genomic Instability and Inflammation.* Frontiers in Cell and Developmental Biology, 2021. **9**(1671).

50. McIntosh, D. and J.J. Blow, *Dormant origins, the licensing checkpoint, and the response to replicative stresses.* Cold Spring Harb Perspect Biol, 2012. **4**(10).

51. Haahr, P., et al., *Activation of the ATR kinase by the RPA-binding protein ETAA1.* Nature Cell Biology, 2016. **18**(11): p. 1196-1207.

52. Zou, L., *DNA Replication Checkpoint: New ATR Activator Identified.* Current Biology, 2017. **27**(1): p. R33-R35.

53. Cotta-Ramusino, C., et al., *A DNA damage response screen identifies RHINO, a 9-1-1 and TopBP1 interacting protein required for ATR signaling.* Science, 2011. **332**(6035): p. 1313-7.

54. Cortez, D., *Replication-Coupled DNA Repair.* Mol Cell, 2019. **74**(5): p. 866-876.

55. Petermann, E., M. Woodcock, and T. Helleday, *Chk1 promotes replication fork progression by controlling replication initiation.* Proc Natl Acad Sci U S A, 2010. **107**(37): p. 16090-5.

56. Quinet, A., D. Lemaçon, and A. Vindigni, *Replication Fork Reversal: Players and Guardians.* Molecular Cell, 2017. **68**(5): p. 830-833.

57. Neelsen, K.J. and M. Lopes, *Replication fork reversal in eukaryotes: from dead end to dynamic response.* Nat Rev Mol Cell Biol, 2015. **16**(4): p. 207-20.

58. Quinet, A., et al., *PRIMPOL-Mediated Adaptive Response Suppresses Replication Fork Reversal in BRCA-Deficient Cells.* Mol Cell, 2020. **77**(3): p. 461-474.e9.

59. Calvo, Patricia A., et al., *Motif WFYY of human PrimPol is crucial to stabilize the incoming 3'-nucleotide during replication fork restart.* Nucleic Acids Research, 2021. **49**(14): p. 8199-8213.

60. Ghosal, G. and J. Chen, *DNA damage tolerance: a double-edged sword guarding the genome.* Translational cancer research, 2013. **2**(3): p. 107-129.

1

61. Lin, J.-R., et al., *SHPRH and HLTF Act in a Damage-Specific Manner to Coordinate Different Forms of Postreplication Repair and Prevent Mutagenesis.* Molecular Cell, 2011. **42**(2): p. 237-249.

62. Chang, D.J. and K.A. Cimprich, *DNA damage tolerance: when it's OK to make mistakes.* Nature chemical biology, 2009. **5**(2): p. 82-90.

63. Scully, R., et al., *Recombination and restart at blocked replication forks.* Current Opinion in Genetics & Development, 2021. **71**: p. 154-162.

64. Kramara, J., B. Osia, and A. Malkova, *Break-Induced Replication: The Where, The Why, and The How.* Trends in Genetics, 2018. **34**(7): p. 518-531.

65. García-Muse, T. and A. Aguilera, *Transcription–replication conflicts: how they occur and how they are resolved.* Nature Reviews Molecular Cell Biology, 2016. **17**(9): p. 553-563.

66. Hamperl, S. and K.A. Cimprich, *Conflict Resolution in the Genome: How Transcription and Replication Make It Work.* Cell, 2016. **167**(6): p. 1455-1467.

67. Tuduri, S., et al., *Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription.* Nat Cell Biol, 2009. **11**(11): p. 1315-24.

68. Crossley, M.P., M. Bocek, and K.A. Cimprich, *R-Loops as Cellular Regulators and Genomic Threats.* Mol Cell, 2019. **73**(3): p. 398-411.

69. Skourti-Stathaki, K., K. Kamieniarz-Gdula, and N.J. Proudfoot, *R-loops induce repressive chromatin marks over mammalian gene terminators.* Nature, 2014. **516**(7531): p. 436-439.

70. Kotsantis, P., et al., *Increased global transcription activity as a mechanism of replication stress in cancer.* Nature Communications, 2016. **7**(1): p. 13087.

71. Hamperl, S., et al., *Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses.* Cell, 2017. **170**(4): p. 774-786.e19.

72. Niehrs, C. and B. Luke, *Regulatory R-loops as facilitators of gene expression and genome stability.* Nat Rev Mol Cell Biol, 2020. **21**(3): p. 167-178.

73. Aguilera, A. and T. García-Muse, *R Loops: From Transcription Byproducts to Threats to Genome Stability.* Molecular Cell, 2012. **46**(2): p. 115-124.

74. Huertas, P. and A. Aguilera, *Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination.* Mol Cell, 2003. **12**(3): p. 711-21.

75. Salas-Armenteros, I., et al., *Human THO-Sin3A interaction reveals new mechanisms to prevent R-loops that cause genome instability.* Embo j, 2017. **36**(23): p. 3532-3547.

76. Bayona-Feliu, A., et al., *The SWI/SNF chromatin remodeling complex helps resolve R-loop-mediated transcription–replication conflicts.* Nature Genetics, 2021.

77. Zimmer, A.D. and D. Koshland, *Differential roles of the RNases H in preventing chromosome instability.* Proc Natl Acad Sci U S A, 2016. **113**(43): p. 12220-12225.

78. Lockhart, A., et al., *RNase H1 and H2 Are Differentially Regulated to Process RNA-DNA Hybrids.* Cell Rep, 2019. **29**(9): p. 2890-2900.e5.

79. Bourgeois, C.F., F. Mortreux, and D. Auboeuf, *The multiple functions of RNA helicases as drivers and regulators of gene expression.* Nature Reviews Molecular Cell Biology, 2016. **17**(7): p. 426-438.

80. Sessa, G., et al., *BRCA2 promotes DNA-RNA hybrid resolution by DDX5 helicase at DNA breaks to facilitate their repair‡.* Embo j, 2021. **40**(7): p. e106018.

81. Sollier, J. and K.A. Cimprich, *Breaking bad: R-loops and genome integrity.* Trends Cell Biol, 2015. **25**(9): p. 514-22.

82. Amon, J.D. and D. Koshland, *RNase H enables efficient repair of R-loop induced DNA damage.* eLife, 2016. **5**: p. e20533.

83. Ohle, C., et al., *Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair.* Cell, 2016. **167**(4): p. 1001-1013.e7.

84. Costantino, L. and D. Koshland, *Genome-wide Map of R-Loop-Induced Damage Reveals How a Subset of R-Loops Contributes to Genomic Instability.* Molecular Cell, 2018. **71**(4): p. 487-497. e3.

85. Makharashvili, N., et al., *Sae2/CtIP prevents R-loop accumulation in eukaryotic cells.* Elife, 2018. **7**.

86. Keskin, H., et al., *Transcript-RNA-templated DNA recombination and repair.* Nature, 2014. **515**(7527): p. 436-439.

87. Welty, S., et al., *RAD52 is required for RNA-templated recombination repair in post-mitotic neurons.* J Biol Chem, 2018. **293**(4): p. 1353-1362.

88. Michelini, F., et al., *Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks.* Nature Cell Biology, 2017. **19**(12): p. 1400-1411.

89. D'Alessandro, G., et al., *BRCA2 controls DNA:RNA hybrid level at DSBs by mediating RNase H2 recruitment.* Nature Communications, 2018. **9**(1): p. 5376.

90. Zhang, X., et al., *Attenuation of RNA polymerase II pausing mitigates BRCA1-associated R-loop accumulation and tumorigenesis.* Nat Commun, 2017. **8**: p. 15908.

91. Setton, J., et al., *Synthetic Lethality in Cancer Therapeutics: The Next Generation.* Cancer Discovery, 2021. **11**(7): p. 1626-1635.

92. Ashworth, A., C.J. Lord, and J.S. Reis-Filho, *Genetic interactions in cancer progression and treatment.* Cell, 2011. **145**(1): p. 30-8.

93. Hanzlikova, H., et al., *The Importance of Poly(ADP-Ribose) Polymerase as a Sensor of Unligated Okazaki Fragments during DNA Replication.* Mol Cell, 2018. **71**(2): p. 319-331.e3.

94. Fong, P.C., et al., *Poly(ADP)-ribose polymerase inhibition: frequent durable responses in BRCA carrier ovarian cancer correlating with platinum-free interval.* J Clin Oncol, 2010. **28**(15): p. 2512-9.

95. Rottenberg, S., et al., *High sensitivity of BRCA1-deficient mammary tumors to the PARP inhibitor AZD2281 alone and in combination with platinum drugs.* Proc Natl Acad Sci U S A, 2008. **105**(44): p. 17079-84.

96. Christie, E.L., et al., *Multiple ABCB1 transcriptional fusions in drug resistant high-grade serous ovarian and breast cancer.* Nature Communications, 2019. **10**(1): p. 1295.

97. Hanzlikova, H., et al., *Overlapping roles for PARP1 and PARP2 in the recruitment of endogenous XRCC1 and PNKP into oxidized chromatin.* Nucleic acids research, 2017. **45**(5): p. 2546-2557.

98. Zandarashvili, L., et al., *Structural basis for allosteric PARP-1 retention on DNA breaks.* Science, 2020. **368**(6486): p. eaax6367.

99. Murai, J., et al., *Trapping of PARP1 and PARP2 by Clinical PARP Inhibitors.* Cancer Research, 2012. **72**(21): p. 5588-5599.

100. Pettitt, S.J., et al., *Genome-wide and high-density CRISPR-Cas9 screens identify point mutations in PARP1 causing PARP inhibitor resistance.* Nature Communications, 2018. **9**(1): p. 1849.

101. Gogola, E., et al., *Selective Loss of PARG Restores PARylation and Counteracts PARP Inhibitor-Mediated Synthetic Lethality.* Cancer Cell, 2018. **33**(6): p. 1078-1093.e12.

102. Koh, G., et al., *Mutational signatures: emerging concepts, caveats and clinical applications.* Nature Reviews Cancer, 2021. **21**(10): p. 619-637.

103. Patel, P.S., A. Algouneh, and R. Hakem, *Exploiting synthetic lethality to target BRCA1/2-deficient tumors: where we stand.* Oncogene, 2021. **40**(17): p. 3001-3014.

104. Zatreanu, D., et al., *Polθ inhibitors elicit BRCA-gene synthetic lethality and target PARP inhibitor resistance.* Nature Communications, 2021. **12**(1): p. 3636.

105. Guo, E., et al., *FEN1 endonuclease as a therapeutic target for human cancers with defects in homologous recombination.* Proceedings of the National Academy of Sciences, 2020. **117**(32): p. 19415-19424.

106. Noordermeer, S.M., et al., *The shieldin complex mediates 53BP1-dependent DNA repair.* Nature, 2018. **560**(7716): p. 117-121.

107. He, Y.J., et al., *DYNLL1 binds to MRE11 to limit DNA end resection in BRCA1-deficient cells.* Nature, 2018. **563**(7732): p. 522-526.

108. Schlacher, K., H. Wu, and M. Jasin, *A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2.* Cancer Cell, 2012. **22**(1): p. 106-16.

109. Rondinelli, B., et al., *EZH2 promotes degradation of stalled replication forks by recruiting MUS81 through histone H3 trimethylation.* Nature Cell Biology, 2017. **19**(11): p. 1371-1378.

110. Ray Chaudhuri, A., et al., *Replication fork stability confers chemoresistance in BRCA-deficient cells.* Nature, 2016. **535**(7612): p. 382-7.

111. Dungrawala, H., et al., *RADX Promotes Genome Stability and Modulates Chemosensitivity by Regulating RAD51 at Replication Forks.* Mol Cell, 2017. **67**(3): p. 374-386.e5.

112. Taglialatela, A., et al., *Restoration of Replication Fork Stability in BRCA1- and BRCA2-Deficient Cells by Inactivation of SNF2-Family Fork Remodelers.* Mol Cell, 2017. **68**(2): p. 414-430.e8.

113. Paes Dias, M., et al., *Loss of nuclear DNA ligase III reverts PARP inhibitor resistance in BRCA1/53BP1 double-deficient cells by exposing ssDNA gaps.* Molecular Cell, 2021.

114. Cong, K., et al., *Replication gaps are a key determinant of PARP inhibitor synthetic lethality with BRCA deficiency.* Molecular Cell, 2021. **81**(15): p. 3128-3144.e7.
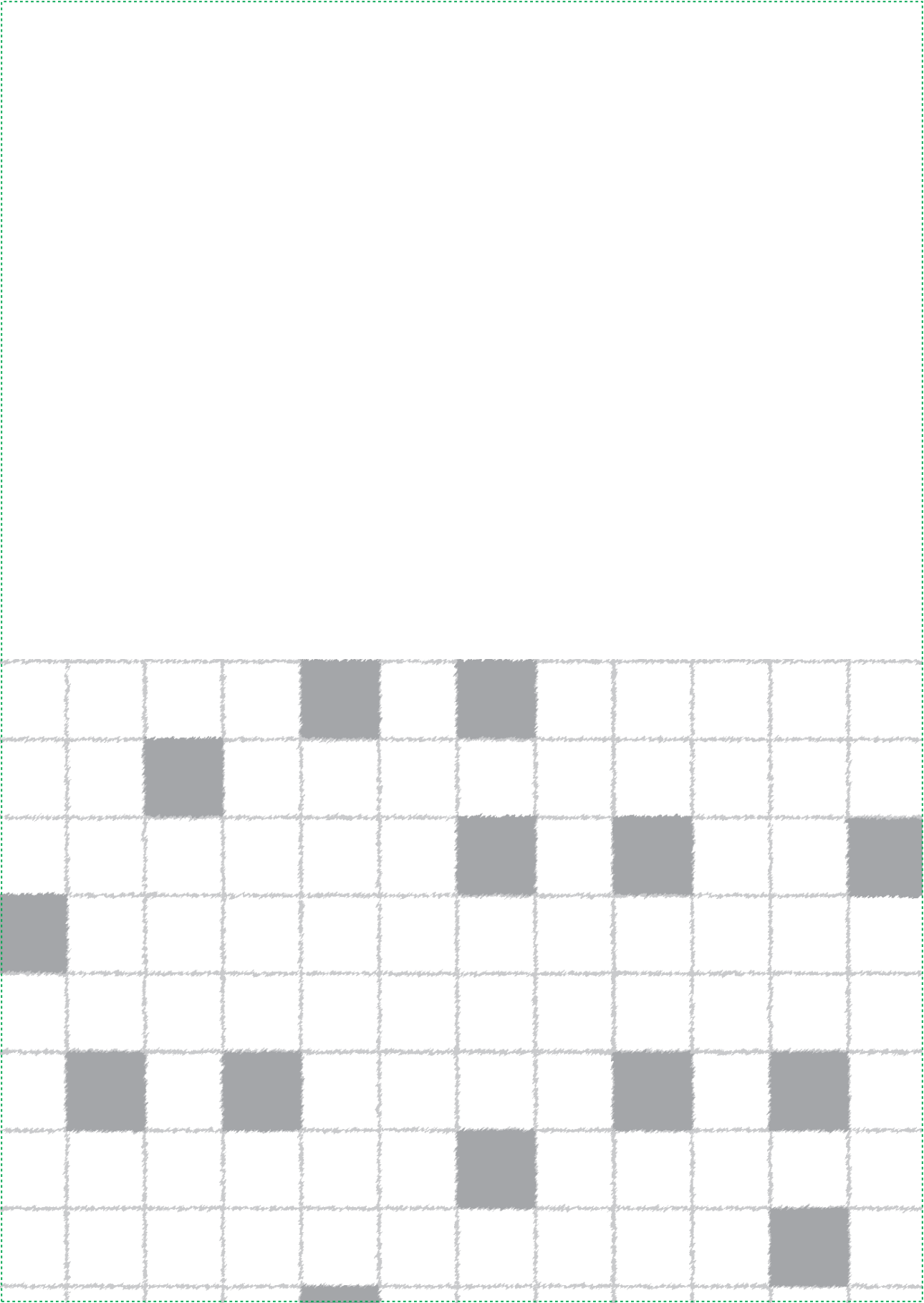
CHAPTER 2

DNA double-strand break repair: putting zinc fingers on the sore spot

Jenny Kaur Singh and Haico van Attikum*

ABSTRACT

Zinc-Finger (ZnF) proteins represent to the most abundant group of proteins in the human genome. At first characterized as DNA binding proteins, it has become increasingly clear that ZnF-proteins have the ability to bind a large variety of substrates such as RNAs, proteins and post-translational modifications, suggesting potential roles in a variety of biological processes. Indeed, several studies have implicated ZnF-proteins for instance in transcription regulation, signal transduction and cell migration. Intriguingly, more recently these proteins have emerged as important protectors of the genome, particularly by orchestrating the repair of highly deleterious DNA double-strand breaks. Here we provide a comprehensive summary of the roles of ZnF domain-containing proteins in DNA double-strand break repair and discuss how their dysfunction impacts genome stability and human disease.

Keywords: Zinc-finger (ZnF) proteins, DNA double-strand break (DSB), DSB repair, non-homologous end joining (NHEJ), homologous recombination (HR), genome stability

INTRODUCTION

Our genome is constantly challenged by endogenous and exogenous DNA damage, causing tens of thousands of DNA lesions on a daily basis [1]. DNA double-strand breaks (DSB) are considered one of the most toxic lesions that can occur in the genome. If left unrepaired or repaired inaccurately, they can lead to mutations and chromosomal translocations, thereby increasing the risk of developing human disorders such as cancer, neurodegeneration or immunodeficiency [2]. To protect the integrity of our genome, cells have evolved specialized molecular machines to detect and repair DSBs, the latter of which involves two main pathways: non-homologous end-joining (NHEJ) and homologous recombination (HR) [3]. The dominant repair pathway is classical non-homologous end joining (cNHEJ). During this repair process, the broken ends are bound by Ku70/Ku80, followed by the assembly of DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and aprataxin and polynucleotide kinase/phosphatase-like factor (APLF). This leads to the recruitment of a ligation machinery consisting of XRCC4, LIG4 and XLF, in a manner stimulated by PAXX, which seals the broken ends. cNHEJ requires no or minimal end-processing, the latter of which may lead to small deletions and insertions at the repair site [4]. HR, on the contrary, is the more faithful repair pathway and is restricted to the S and G2 phases of the cell cycle as it requires the presence of a sister chromatid. During HR, extensive resection of the DSB occurs involving the activities of endo- and exonucleases, including MRE11, CtIP, DNA2 and EXO1, to generate 3'single-stranded (ss)DNA overhangs that become coated by the single-strand DNA binding complex RPA. This triggers the recruitment of BRCA1-PALB2-BRCA2 complexes, whose docking on damaged DNA occurs in a manner dependent on physical interactions between PALB2 and RNF168 on the one hand, and RNF168 and ubiquitylated H2A on the other hand [5]. This allows for the removal of RPA by BRCA2 and loading of the recombinase RAD51, which promotes homology search and strand invasion using the undamaged sister chromatid as a repair template [6]. When cNHEJ and HR are compromised, repair can occur via alternative non-homologous end joining (aNHEJ), which is an error-prone pathway that uses short stretches of microhomology to seal the broken ends. This pathway is dependent on XRCC1-Ligase III complex or the DNA polymerase POLQ [7]. Alternatively, larger stretches of resection can lead to repair by single-strand annealing (SSA), which is dependent on RAD52 and ERCC1 [8].

DNA repair pathway choice is strictly regulated throughout the cell cycle by the RING-finger E3 ubiquitin ligases RNF8 and RNF168, which promote the ubiquitin-dependent

assembly of 53BP1 and the BRCA1-Abraxas-RAP80-MERIT40 (BRCA1-A) complex at DSBs [9]. 53BP1 inhibits DNA end-resection in G1 phase by its various effectors, including RIF1 and the Shieldin complex, to impair HR and favor NHEJ [10], whereas the BRCA1-A complex suppresses HR by sequestering BRCA1 away from the repair site. In addition, BRCA1's interaction with PALB2 is inhibited in G1 due to suppressive ubiquitylation of the BRCA1-interacting domain in PALB2, further impairing HR [11]. Conversely, RIF1 accumulation at DSB sites in S and G2 phase is compromised by BRCA1 and CtIP, allowing end resection and HR to occur [12].

The repair of DSBs is challenged by the fact that genomic DNA is interweaved by histone and non-histone proteins in a high-order structure called chromatin. DNA repair machineries have to overcome this barrier to gain access to damaged DNA and repair the lesions [13]. Over the recent years it became evident that chromatin structure is altered by different mechanisms such as DNA methylation, ATP-dependent chromatin remodeling and post-translational modifications (PTMs), including but not limited to phosphorylation, acetylation, methylation, S-acylation, poly(ADP)ribosylation (PAR), SUMOylation and ubiquitylation [14, 15]. One key enzyme that helps to overcome the chromatin barrier is PARP1. PARP1 is among the first proteins that binds to DNA breaks, where it promotes the rapid local expansion of chromatin by the formation of PAR-chains on itself and several target proteins, including histones [16]. This allows for the accumulation of several DNA repair proteins, either by binding to DNA damage-associated PAR-chains or to the exposed damaged DNA, as was observed for the CHD2 chromatin remodeling enzyme and several cNHEJ factors, respectively [17, 18]. In addition to PARP1, PARP3 also partakes in cNHEJ. This involves its auto-mono-ADP ribosylation (MAR), which is a prerequisite for the interaction with APLF and the subsequent assembly of functional cNHEJ complexes during DSB repair [19].

Several studies reported that a number of transcription factors are also recruited to sites of DNA damage, in a manner often dependent on PARP/PAR [17, 20]. A large class of these transcription factors belongs to the Zinc-Finger (ZnF) domain-containing protein family, which represent one of the most abundant groups of proteins in the eukaryotic cell. It has become evident that ZnF protein function is not limited to transcription regulation, but is also key to many other cellular processes such as for instance signal transduction, cell migration and DSB repair [21]. Here we provide an overview of the current knowledge on the role of ZnF proteins in DSB repair, most notably NHEJ and HR, highlighting their importance as protectors of the genome.

2

FUNCTIONAL CLASSIFICATION OF ZNF DOMAIN-CONTAINING PROTEINS

The ZnF-domain was recognized over more then 30 years ago as a repeated zinc-binding motif consisting of conserved histidine and cysteine residues in the *Xenopus* transcription factor TFIIIA [22]. Over the years, numerous other zinc-binding domains have been identified, which are encoded by nearly 5% of all human genes [23]. ZnF-domains usually consist of multiple unique zinc-finger motifs that are responsible for the contact with different target molecules. The numerous structural folds, diverse binding modes and sequence recognition sites of zinc-finger motifs have led to the evolution of zinc-finger domain families. To date, ZnFs comprise more then 50 unique domains such as the FCS-, PBZ- and PHD-domains (Table 1). Although, zinc-finger motifs within a ZnF-domain are evolved to bind zinc-ions, it has become clear they are also capable of binding other metals in a cellular environment. For instance cobalt, cadmium, nickel and copper can compete with zinc for binding to ZnF-domains. These metal-substitutions cause heteregenous structural and chemical changes of the zinc-finger motifs and thereby change their substrate-recognition properties [24].

While ZnFs are mostly known as DNA-binding domains, it has become clear that ZnF- domains also have the ability to bind RNA, lipids, and methylated DNA, as well as to protein PTMs such as SUMO, ubiquitin, PAR and methyl (Figure 1 and [25]). Interestingly, many ZnF proteins possess multiple different types of ZnF-domains. Consequently, they can exhibit very different binding specificities for target molecules [21]. ZnF-domains therefore occur in several unrelated protein super-families (e.g. transcription factors, nuclear hormone receptors, integrase enzymes, E3-ubiquitin ligases, chromatin remodelers, tumor suppressors, RAS-GTPases, membrane transport proteins and chaperones) varying in domain structure and sequence, and displaying considerable versatility in their binding modes (e.g., some bind to DNA, others to RNA or proteins). This suggests that ZnF-domains act as scaffolds that have evolved a diversity of functions. Indeed, ZnF domain-containing proteins have been shown to function e.g. in transcription, signal transduction and cell migration, as well as in different DSB repair pathways (Figure 1). In this review, we provide an overview of ZnF-proteins whose ZnF-domains were shown to be functionally relevant for NHEJ or HR.
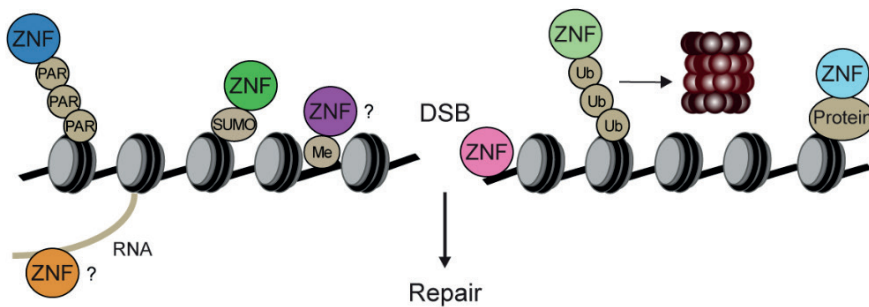
Figure 1. Binding specificities of zinc-finger (ZnF) domains. At DSBs, ZnF-domains may bind to a variety of substrates including DNA, RNA, methylated DNA, proteins and PTMs such as PAR, SUMO and K48-linked ubiquitin chains.

ZNF PROTEINS IMPLICATED IN NHEJ REPAIR OF DNA BREAKS

Several different ZnF domain-containing proteins have been implicated in the repair of DSBs via NHEJ. Below we discuss the function of several of these ZnF domain-containing proteins in this repair process (Table 1).

PBZ-type ZnF protein: APLF

First classified as a classical C2H2 domain-containing protein, structural and biochemical studies revealed that APLF contains shorter inter-cysteine-histidine loops which have binding specificity towards PAR. This led to the classification of APLF as a member of the PBZ-domain family [26]. Interestingly, the two PBZ-domains of APLF bind to auto-MARylated PARP3 at sites of DNA damage. APLF also contains a Ku-binding motif (KBM) that mediates its interaction with Ku and is important for APLF's recruitment to sites of DNA damage. Moreover, APLF interacts with XRCC4 via its N-terminal fork-head associated domain (Figure 2A and [19]). This multitude of interactions mediated by APLF ensures the binding and retention of LIG4 through its cofactors XLF and XRCC4, thereby promoting cNHEJ [27]. Importantly, mutations in the PBZ-domains, as well as in the KBM motif, abolished APLF recruitment, impaired XRCC4 loading and compromised both the efficiency and accuracy of NHEJ, indicating the relevance of its ZnF-domain for DSB repair [28].

C2H2-type ZnF proteins: ZBTB24 and ZNF281

Two classical C2H2-type ZnF proteins, ZBTB24 and ZNF281, have been implicated in NHEJ. Sequence analysis revealed that ZBTB24 contains a BTB- or POK- (POZ and

Krüppel zinc- finger) domain, which is frequently found in transcription factors [29], as well as a Krüppel-type C2H2 zinc-finger domain. Interestingly, mutations in ZBTB24 are causally linked to immunodeficiency, centromeric instability and facial anomalies (ICF2) syndrome [30]. We unveiled that loss of ZBTB24 in both B cells from mice and ICF2 patients causes a cNHEJ defect during immunoglobulin class-switch recombination (CSR) and consequently impaired immunoglobulin production [31]. Domain-mapping revealed that the C2H2-, but not the BTB- domain of ZBTB24 has PAR-binding affinity and mediates its interaction with PARP1-associated PAR-chains, an event that is important for ZBTB24 recruitment to DNA breaks. Moreover, the C2H2-domain of ZBTB24 protects PAR-chains on PARP1 and this is critical for the assembly of XRCC4-LIG4 complexes and NHEJ at DNA breaks (Figure 2B and [31]). Together, these findings show the importantce of a C2H2-type ZnF protein in cNHEJ, providing a molecular basis for the observed CSR defect in ICF2 patients.

The C2H2 domain-containing protein ZNF281 was identified in a screen aimed to measure the localization of transcription factors at laser micro-irradiation induced DNA damage [20]. ZNF281 recruitment to sites of DNA damage was dependent on its C2H2-domain and the activity of PARP1 [32]. Whether its recruitment relies on binding of the C2H2-domain to DNA damage-associetd PAR-chains or the exposed damaged DNA remains to be established. ZNF281 also interacts with XRCC4, thereby promote its loading at DNA breaks. Interestingly, point-mutations in ZNF281's C2H2-motifs abolished the interaction with XRCC4, impaired XRCC4 recruitment to sites of DNA damage and lead to a defect in cNHEJ (Figure 2C and [32]). Interestingly, however, another study reported that ZNF281 binds to the promotor of XRCC4 in manner dependent on its C2H2-domain, thereby controlling XRCC4 expression and cNHEJ (Figure 2C and [33]). Future studies will have to clarify precisely how ZNF281's apparent dual mode-of-action impacts this repair process.

PHD-type ZnF proteins: PHF6 and ACF1

PHD-domains were discovered more than 25 years ago and numerous sequence and functional analyses of PHD containing-proteins have pinpointed a role in the regulation of chromatin structure [34]. Their intimate association with histones and their ability to recruit multi-protein complexes to damaged chromatin has suggested important roles for these proteins in DNA-based processes, including DNA repair [34]. Indeed, two PHD-type ZnF proteins, PHF6 and ACF1, have been implicated in NHEJ.

PHF6 and several other PHD-domain family-members such as PHF3 and PHF12, were among the top hits in a G2-checkpoint recovery RNAi-screen [35]. PHF6 is recruited

to DSBs in a PAR/PARP dependent manner, where it promotes 53BP1 accummulation and repair via NHEJ. Importantly, G2-checkpoint recovery and 53BP1 accumulation were also impaired in cells expressing a mutant form of PHF6 lacking its PHD-domain. Although the exact molecular mechanism underlying PHF6's role in G2-checkpoint recovery remains to be established, the existing data suggest that the PHD-domain plays a pivotal role in promoting 53BP1-dependent NHEJ, thereby preventing persistent unrepaired DSBs that inhibit recovery from a G2- checkpoint arrest [35].

ACF1 is a non-catalytic PHD-domain containing subunit of the ACF1-ISWI chromatin remodeling complex. ACF1 binds histones and is recruited to sites of DNA damage through its PHD-domain [36]. Moreover, ACF1 recruits components of the ACF1-ISWI complex, most notably its catalytic ATPase subunit SNF2H, as well as Ku70/Ku80 to DSBs. The latter is dependent on the interaction between ACF1 and Ku70/Ku80 and SNF2H-driven chromatin remodeling. In addition, ACF1 and SNF2H also promote DSB repair by HR in cooperation with the ACF1-associated proteins CHRAC15 and CHRAC17 in the ISWI-complex, which induce chromatin changes to target ACF1/SNF2H to damaged chromatin. This is followed by the SNF2H-dependent reruitment of HR-factors such as RPA and RAD51 at DSBs [37]. Taken together, these data suggest that chromatin changes induced by SNF2H are important for both HR and NHEJ. The latter is dependent on SNF2H dependent recruitment of ACF1, which in turn leads to the loading of Ku70/Ku80 at DSBs, thereby facilitating repair of these lesions via NHEJ [36].

RING-type ZnF proteins: RNF8 and RNF126

RING-domains are present in E3 ubiquitin ligases to make direct contact with E2 ubiquitin conjugating enzymes to ensure ubiquitin transfer to target proteins [38]. Two RING finger domain-containing proteins have been implicated in NHEJ. RING-finger protein RNF8 is involved in the proteasomal degradation of Ku80 through its Lys48-linked ubiquitination. This event triggers the release of Ku80 from DNA damage sites. While it is evident that the RING-domain is required for the ubiquitination of Ku80, the E2 enzyme involved in this process remains to be identified [39]. More recently, RNF126, in conjunction with the UBE2D3 E2 conjugating enzyme, was found to associate with Ku80 via its RING-domain. Similar to RNF8, RNF126 also ubiquitylates Ku80 to trigger its release from DNA damage sites during NHEJ [40]. Although, RNF8 and RNF126 both promote NHEJ by regulating Ku70/Ku80 release during DSB repair, unclear is whether they cooperate or act redundantly during this process (Figure 2D).
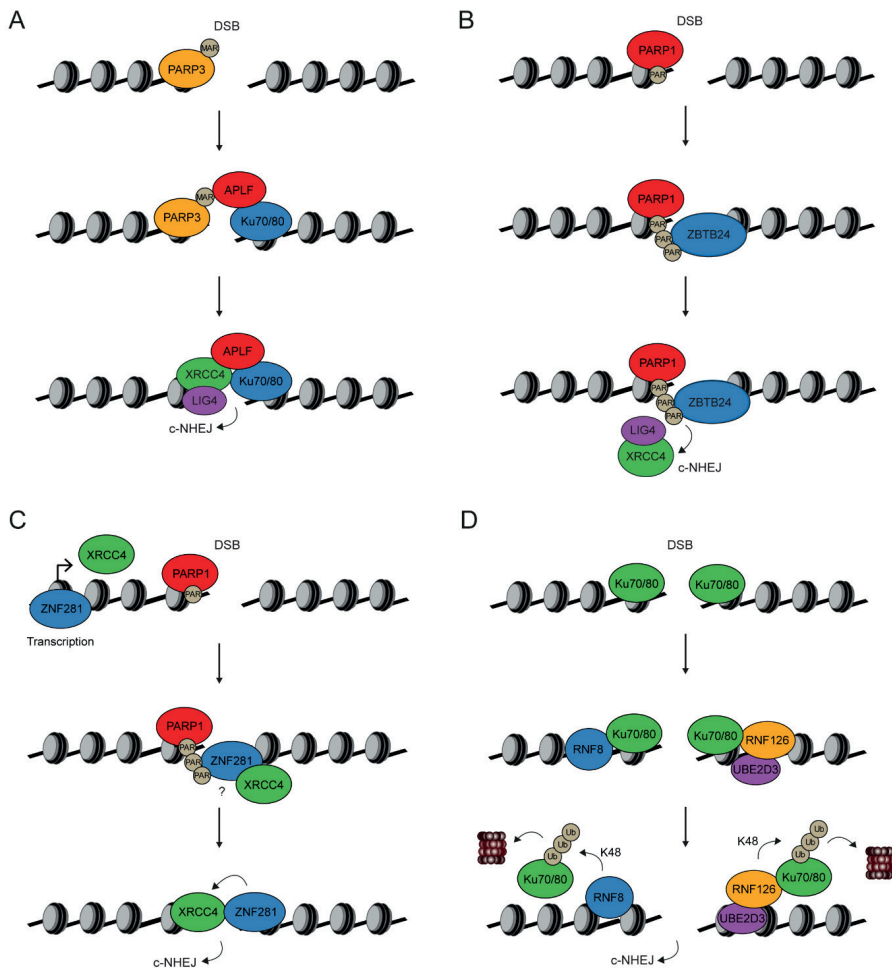
Figure 2 Model for the roles of ZnF-proteins APLF, ZBTB24, ZNF281, RNF8 and RNF126 in NHEJ. (A) APLF is recruited to DSBs by binding to PARP3-associated MAR through its PBZ-domain, and by associating with Ku70/Ku80. It also interacts with XRCC4, thereby promoting recruitment of XRCC4/LIG4 and c-NHEJ. (B) ZBTB24 is recruited to DSBs by binding to PARP1-associated PAR through its C2H2-domain. It promotes PARP1-mediated PAR synthesis, thereby functioning as a scaffold to protect PAR from degradation. This is followed by the PARP1-dependent recruitment of XRCC4/LIG4 and c-NHEJ. (C) ZNF281 is recruited to DSBs via PARP activity and through its C2H2-domain. Whether the latter involves its binding to PAR or DNA is unclear. The C2H2-domain also interacts with XRCC4, thereby facilitating XRCC4 recruitment and c-NHEJ. ZNF281 also binds to the *XRCC4* promotor via its C2H2-domain, thereby controlling XRCC4 expression and c-NHEJ. (D) RNF8 and RNF126 are recruited to DSBs, where they modify Ku70/Ku80 by K48-ubiquitylation, triggering the proteasome-dependent degradation of Ku70/Ku80 during NHEJ. RNF126 does so by interacting with Ku70/Ku80, an event that is also critical for its recruitment, and by operating with the E2 conjugating enzyme UBE2D3.

ZNF PROTEINS IMPLICATED IN HR REPAIR OF DNA BREAKS

In addition to the role of ZnF domain-containing proteins in NHEJ, it became evident that several ZnF proteins are implicated in DSB repair by HR. Below we discuss the function of several of these ZnF domain-containing proteins in this repair process (Table 1).

PHD and CXXC-type ZnF protein: KDM2A

PHD-domain containing proteins have not only been implicated in NHEJ (see above), but also impact HR, particularly the PHD and CXXC domain-containing protein KDM2A interacts with and becomes phosphorylated by the DSB-responsive kinase ATM at threonine 632 located within its PHD-motif. This counteracts KDM2A's binding to damaged chromatin and enhances H3K36me2 levels at DSB sites. Moreover, H3K26me2 serves as a platform to recruit MRE11 via its binding partner NBS1, which binds this histone mark via its BRCT2 domain. Utlimately, this faciliates HR by promoting MRE11-dependent DNA-end resection (Figure 3B and [41]) Interestingly, however, a more recent study reported a role for the CXXC-domain in the recruitment of KDM2A to *bona fide* nuclease-induced DSBs [42]. In addition, proteomics approaches identified 53BP1 as a binding partner of KDM2A. The KDM2A-53BP1 interaction was found to be dependent on the CXXC-domain of KDM2A and was required for the KDM2A-mediated ubiquitination and recruitment of 53BP1. Unclear is how this histone demethylase promotes 53BP1 ubiquitylation and how this modification of 53BP1 affects its accumulation at DNA damage sites (Figure 3B). Nevertheless, KDM2A-depleted cells displayed impaired 53BP1 foci formation, elevated levels of unrepaired DSBs and premature exit from the G2/M checkpoint. Re-expression of ΔCXXC-version of KDM2A failed to rescue these defects and resulted in an increase of micronuclei formation [42]. These findings suggest that the ZnF-domains of KDM2A support DSB repair through distinct modes. While its impact on HR is evident, it remains unclear whether KDM2A also affects NHEJ, which may be expected given its role in regulating 53BP1. Future studies will have to resolve precisely how KDM2A dictates DSB-repair outcome.

C2H2-type ZnF proteins: CTCF

CTCF is commonly known for its role in genome organziation and transcription [43]. However, proteomics approaches identified CTCF as a DNA damage-dependent binding-partner of MRE11 and CtIP. These interactions were found to be dependent on

its 11 C2H2-motifs and facilitate CtIP recruitment to DNA breaks, allowing for DNA end-resection and HR to take place [44]. In line with this, CTCF was also described to interact with, and recruit RAD51 and BRCA2 at sites of DNA damage [45, 46], suggesting a multifaceted role during different stages of the HR process.

ePHD-type ZnF protein: PHF11

PHF11, which is a member of the extended PHD (ePHD) family of ZnF proteins, contains a ZnF-domain that consists of two parts, one of which is the pre-PHD region that binds a single zinc ion, and the other is a PHD-finger motif that binds two additional zinc-ions. PHF11 was identified at uncapped telomeres using the proteomics of isolated chromatin segments (PICh) approach, and its PHD-finger motif was found to interact with RPA, suggesting it may act at resected DNA ends [47]. Indeed, PHF11 mediates the removal of RPA, thereby providing access for EXO1 or DNA2 to partially resected ends generated by MRN, which are otherwise inaccessible for these nucleases [47]. Consequently, loss of PHF11 impaired HR and rendered cells sensitive to DSB-inducing agents. Whether PHF11 is recruited to DNA breaks via its PHD domain-dependent interaction with RPA or via its ePHD-domain remains to be established.

A20-type ZnF protein: A20/TNFAIP3

The DSB-response involves the RING-type ZnF proteins and E3 ubiquitin ligases RNF8 and RNF168. RNF8 interacts with phosphorylated MDC1 and ubiquitylates histone H1 to recruit RNF168 through ubiquitin binding. RNF168 then catalyzes the mono-ubiquitination of H2A and H2AX, which initiates the subsequent formation of K63-linked ubiquitin chains for the assembly of the BRCA1-A complex and 53BP1, thereby promoting NHEJ [9]. A recently discovered zinc-finger protein A20/TNFAIP3 was described to function in the RNF168-53BP1 axis. This protein contains an A20-type ZnF-domain that was identified in a cDNA-based screen for regulatory factors in the tumor necrosis factor (TNF) signaling cascade. Sequence analysis showed that this domain contains multiple repeats of Cys_2/Cys_2 fingers [48]. Structural analysis and functional assays confirmed that A20 binds mono-ubiquitin and K63-linked poly-ubiquitin chains [49]. Indeed, following its transcriptional upregulation by NFkβ after DNA damage, A20 directly binds to RNF168 via its ZnF-domain. This disrupts the binding of RNF168 to ubiquitinated H2A and H1, thereby antagonizing RNF168-dependent ubiquitylation and 53BP1 binding at DNA damage sites (Figure 3C and [50]). Accordingly, loss of A20 lead to increased NHEJ levels,

concomitantly with a decrease in HR. This establishes a new link between NFkB-signaling and the regulation of an A20-type ZnF protein during DSB repair pathway choice.

RanBP2-type ZnF protein: RYBP

RYBP belongs to the non-canonical PcG protein complex and possesses an ubiquitin-binding motif (UBM) within its RanBP2-domain, which may be involved in the recognition and/or amplification of ubiquitin at DSBs. Indeed, RYBP preferentially binds to K63-linked ubiquitin chains via its ZnF domain to suppresses BRCA1 binding. However, upon DNA damage RYBP's ZnF domain becomes polyubiquitinated by RNF8, which leads to its rapid removal from damaged chromatin by the VCP/p97 segregase, allowing BRCA1-recruitment and repair via HR to occur [51]. This implies a dual function of the RanBP2-type motif, which on one hand binds to ubiquitin, and on the other hand is required for the ubiquitination-dependent removal of RYBP1. A similar behavior was described for the ZnF proteins TAB2 and TAB3, which belong to the same family [52], suggesting a widespread role for RanBP2-type ZnF proteins in HR.

UBZ4-type ZnF protein: RAD18

The ability of cells to repair DSBs via HR relies on the recombinase RAD51. Vertebrates contain five different RAD51 paralogs which form two distinct protein complexes. Mutations in any of these paralogs leads to defects in HR and genome instability [53]. One such paralog, RAD51C, is regulated by the UBZ4-type ZnF protein RAD18. Similar to the more classical C2H2 ZnFs, such as those found in the UBZ and UBM domains of Y-family polymerases [54], the UBZ4-type zinc-finger was also shown to bind to ubiquitin [55]. Indeed, RAD18 binds to K63-linked ubiquitin chains generated by RNF8/UBC13 through its UBZ4-domain, whereas it associates with RAD51C via its RING-domain, serving as an adaptor for RAD51C loading on damaged chromatin [56, 57]. Importantly, both UBZ4- and RING-domain mutants failed to rescue the HR-defect observed in RAD18-depleted cells, illustrating the importance of these ZnF domains in regulating RAD18-dependent HR. Although it is evident that RAD18 does so by loading RAD51C at DSBs that undergo HR, precisely how RAD51C impacts this repair process remains unclear.

ADD-type ZnF protein: ATRX

The ADD-domain of ATRX consist of an N-terminal GATA-like ZnF domain and a PHD-finger. However, the PHD-finger of ATRX is different from the classical PHD-finger domains

found in PHF2 and KDM2A, as it consists of an additional N-terminal C2C2 motif. Sequence analysis revealed that the only proteins that share this feature are DNMT3 and DNMT3L. Hence, the domain is called ATRX-DNMT3-DNMT3L (ADD) [58]. ATRX recruitment to sites of damage and its binding to damaged chromatin relies on the imperfect PHD domain [59]. This is followed by H3.3 incorporation in cooperation with its chaperone DAXX to promote extended repair synthesis following RAD51-dependent strand invasion during HR.

MYM-type ZnF protein: ZMYM3

The MYM-domains are only found in six mammalian proteins, one of which was identified in a comparative proteomic analysis as a chromatin-interacting protein [60]. This protein, ZMYM3, was also found to interact with members of the BRCA1-A complex (RAP80, ABRA1 and BRE), and promote BRCA1-A accumulation at sites of DNA damage, particularly by regulating ABRA1 accrual. While, the BRCA1-A complex is known to inhibit HR by restricting end-resection in the S/G2 phase of the cell cycle, ZMYM3 was also found to counteract the RAP80-dependent accumulation of BRCA1-A at DNA damage sites, allowing BRCA1-PALB2-BRCA2-RAD51-mediated HR to occur (Figure 3D and [61]). ZMYM3 binds DNA and chromatin via its N-terminal domain, which is distinct from its MYM domain. Perhaps surprisingly, it is this MYM domain that is required for its recruitment to sites of DNA damage [60]. The exact binding substrate of ZMYM3's MYM domain remains, however, to be determined.

RING-type ZnF proteins: BARD1, RNF138 and FRUCC

BARD1 is a RING-type ZnF protein whose RING-finger is required for its dimerization with BRCA1. This stimulates BRCA1-BARD1 E3 ligase activity and the ubiquitylation of H2A at sites of DNA damage [62]. Ubiquitylated H2A serves as a binding platform for the ATP-dependent chromatin remodeler SMARCAD1, which repositions or evicts nucleosomes, thereby counteracting 53BP1-mediated inhibition of DNA resection [63]. On the other hand, BARD1-BRCA1 also binds to DNA, specifically to the D-loop formed after RAD51-dependent strand invasion, a process that is enhanced by BRCA1-BARD1 through direct interaction with RAD51. Thus, both the RING-finger dependent E3-ligase activity and the DNA-binding capabilities of BRCA1 and BARD1 contribute to efficient HR repair [64].

MRE11-RAD50-NBS1 has been implicated in the removal of Ku70/Ku80 from DSBs, allowing end-resection and HR to occur [65]. One protein that supports HR by the removal of Ku is the E3-ubiquitin ligase RNF138, which contains three different ZnF domains

including a RING, C2HC and C2H2 domain. While the deletion of either of these domains separately did not affect RNF138 recruitment to DSBs, deleting all domains simultaneously completely abolished its recruitment, particularly its binding to ssDNA overhangs at these DNA lesions. Consequently, it remains unclear which combination of domains is responsible for its DNA binding. Nevertheless, it is evident that the RING domain mediates RNF138's interaction with Ku70/Ku80. RNF138, in conjunction with the E2 UBED2, ubiquitylates Ku70/Ku80 in a manner dependent on its RING, C2HC and C2H2 domains to promote Ku70/Ku80 eviction from DSBs. This in turn allows for the recruitment of CtIP/EXO1 and extensive end-resection, promoting DSB repair by HR [66].

Since DSBs occur in both inactive and actively transcribed regions, it is of utmost importance that transcription and DNA repair are coordinated properly. This is to prevent collisions between transcription and DNA repair machineries that may otherwise interfere with DSB repair. Indeed, several studies have observed a direct link between transcription repression and HR [67], and implicated a role for the FRUCC-complex in regulating these processes [68]. FRUCC was identified as a E3 ubiquitin ligase complex consisting of FBXL10 and the RING-domain proteins RNF68-RNF2, which ensure the recruitment of the BMI-RNF2 and MEL18-RNF2 E3 ubiquitin ligase complexes. These complexes are responsible for H2A-K119 ubiquitylation, a mark associated with the repression of transcription. In addition, FRUCC also represses transcription by promoting the exchange of H2A with H2A.Z [68]. Loss of FRUCC results in a defect of transcriptional silencing and impaired the loading of HR proteins at DSBs, thereby affecting this repair process.

MYND-type ZnF protein: ZMYND8

The MYND-domain consists of a ZnF-motif that primarily mediates protein-protein interactions in the context of transcription regulation [69]. One of the MYND-domain containing proteins, ZMYND8, was identified as a factor in a screen for bromo-domain proteins that localize to sites of DNA damage [70]. ZMYND8 contains a triple PHD-BRD-PWWP chromatin-binding module in its N-terminus and a C-terminal MYND domain. The PHD-BRD-PWWP domain is responsible for binding acetylated histones and DNA. Proteomics analysis revealed that ZMYND8 associates with the NuRD chromatin remodeling complex, as well as with the ZnF proteins ZNF532, ZNF592 and ZNF687, all of which are recruited to sites of DNA damage. The MYND domain in ZMYND8 binds to a PPPLΦ motif in the NuRD subunit and GATA-type ZnF protein GATAD2A. This interaction is

important for the rapid, poly(ADP-ribose)-dependent recruitment of GATAD2A/NuRD to sites of DNA damage [71]. In addition, more recently it became evident that the association of ZMYND8-NuRD with damaged chromatin also depends on the removal of H3K4me3, which is a mark associated with active transcription. The H3K4me3 demethylase and PHD containing ZnF protein KDM5A, which is recruited to sites of damage via PAR and its PHD domain, was shown to ensure ZMYND8-NuRD binding by demethylation of H3K4me3 [72]. Taken together, these studies demonstrated that the combined action of several ZnF proteins ensures H4K4me3 demethylation at DSBs, allowing ZMYND8-NuRD to bind the damaged chromatin, repress transcription and promote HR [70].
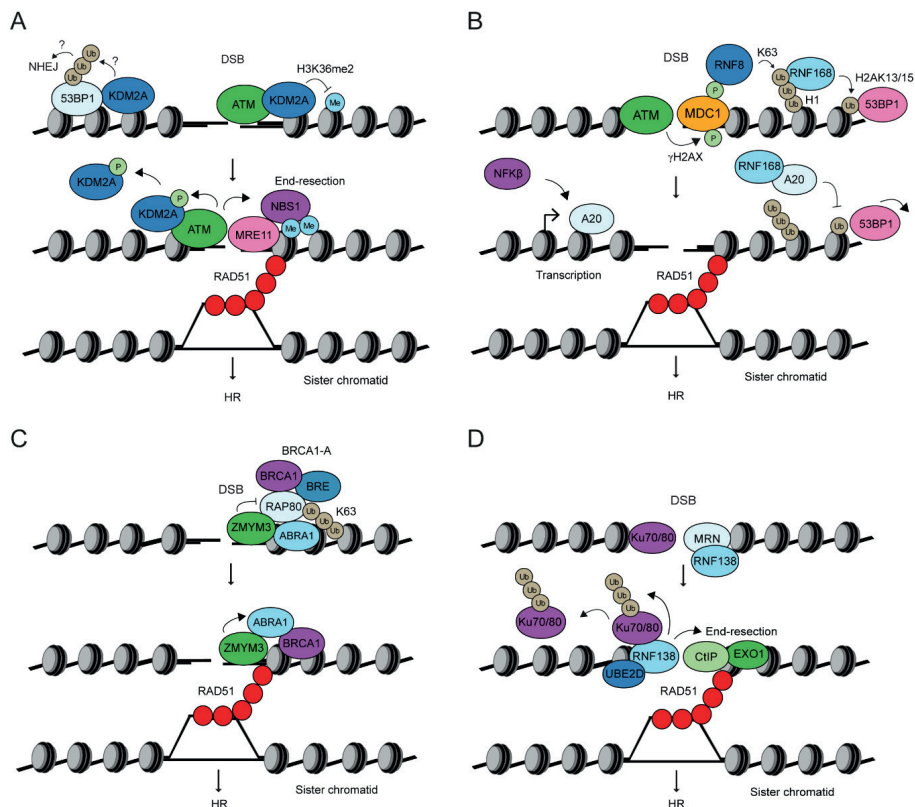
Figure 3 Model for the roles of KDM2A, A20/TNFAIP3, RNF138 and ZMYM3 in HR. (A) ATM-induced phosphorylation of KDM2A's PHD-domain counteracts its chromatin-binding, resulting in increased H3K36me2 levels at DSBs. This enhances the binding of MRE11 via its interaction partner NBS1, which binds to H3K36me2, thereby stimulating end-resection and HR. KDM2A also interacts with 53BP1 through its CXXC-domain promoting the ubiquitination-dependent recruitment of 53BP1 via an unknown mechanism. Whether KDM2A promotes 53BP1-dependent NHEJ is equally unclear. (B) In response to DSBs, A20 is transcriptionally upregulated by NFkβ. A20 binds to RNF168 through its ZnF-domain, abrogating RNF168-

binding to ubiquitinated H1. This impairs the RNF168-dependent ubiquitination of H2AK13/15 and accrual of 53BP1, allowing end-resection and HR to occur. (C) Ku70/80 and MRN complexes bind DSBs independently. In S/G2 phase, MRE11 processes DSB ends to create short overhangs, which are recognized and bound by the ZnF-domains of RNF138. This leads to the displacement of DNA-bound Ku70/80 complex through RNF138-UBE2D-mediated ubiquitylation. Ku removal allows for binding of the CtIP/EXO1 nucleases, which further resect the DSB ends, thereby promoting HR. (D) ZMYM3 is recruited to DSBs through interactions with H2A/H2AX and dsDNA, and via its interaction partners in the BRCA1-A complex: RAP80, BRE and ABRA1. ZMYM3 facilitates the recruitment of ABRA1 and BRCA1 to DSBs, while antagonizing the HR-suppressive effects of BRCA1-A, thereby facilitating RAD51 loading and HR.

CONCLUSIONS AND FUTURE PERSPECTIVES

ZnF domains are present in in at least 5% of all human proteins [23]. Their numerous structural folds and sequence recognition motifs allow them to bind to a plethora of substrates. The fact that ZnF proteins often contain multiple different ZnF domains and that the sequence recognition motifs within a domain exhibit different binding-specificities, also allows them to recognize a combination of substrates. Due to these multifaceted features, ZnF proteins have been implicated in a broad range of molecular processes such as signal transduction, cell migration and transcription regulation. Over the recent years, considerable efforts have also highlighted important roles of ZnF proteins in DSB repair (Table 1). Several studies have demonstrated the role of ZnF domains in facilitating the recruitment and binding of DNA-repair factors to damaged chromatin, by regulating PTMs, chromatin remodeling, protein-protein interactions and/or transcription. Besides the somewhat more well-described role of some larger ZnF domain families in DSB repair (PHD and RING), the functional relevance and mode-of-action of the majority of these proteins in this process remains poorly understood. Not only is it unclear to which substrates and/or combinations of substrates different ZnF domain containing-proteins bind, also the lack of biochemical and cellular complementation studies with mutant proteins lacking functional ZnF domains disallowed their characterization during DSB repair.

Interestingly, several uncharacterized ZnF proteins have been shown to localize at sites of laser induced micro-irradiation [20]. Laser micro-irradiation is a commonly used method to characterize the spatiotemporal dynamics of proteins using imaging of fluorescent proteins at DNA damage sites in fixed or living cells. This method induces a wide range of DNA lesions, including DSBs, single-strand breaks and oxidative DNA lesions, suggesting that ZnF proteins may have the ability to bind different types of DNA lesions. Moreover, ZnF proteins also ranked high in CRISPR/Cas9 screen aimed at identifying protein networks that protect cells against DNA damaging agents, including chemotherapeutics [73]. Collectively, this work suggests a broader role of

ZnF proteins, not only in DNA repair but likely also in tumor resistance mechanisms, than previously anticipated.

Emerging evidence has shown that several ZnF proteins also play key roles in the development of human diseases such as cancer and neurodegeneration [21]. For instance, ZNF281 is overexpressed in colorectal cancer (CRC) and causes cancer metastasis through regulation of epithelial to mesenchymal transition (EMT) [21]. Mutations in ZBTB24 are causally linked to ICF syndrome [30], whereas ZMYM3 is frequently mutated in several cancers, including chronic lymphocytic leukemia (CLL), medulloblastoma, and Ewing sarcoma [23]. It is, however, not completely understood how mutations and changes in the expression of ZnF proteins contribute to disease etiology, warranting the further functional characterization ZnF proteins in human diseases associated with DNA repair alterations. Moreover, mutational signature analyses through next generation sequencing will likely expand our knowledge on ZnF-mutations and their link to human disease, most notably cancer. Such knowledge may not only lead to a better understanding of disease mechanisms, but may also pave the way for the development of drugs that target ZnF proteins in anti-cancer therapies.

ACKNOWLEDGEMENTS

AUTHOR'S CONTRIBUTIONS

J.K.S. and H.v.A. wrote the manuscript.

COMPETING INTERESTS

J.K.S. and H.v.A. have no competing interest.

REFERENCES

1. Ciccia, A. and S.J. Elledge, *The DNA damage response: making it safe to play with knives.* Mol Cell, 2010. **40**(2): p. 179-204.

2. Jackson, S.P. and J. Bartek, *The DNA-damage response in human biology and disease.* Nature, 2009. **461**(7267): p. 1071-8.

3. Scully, R., et al., *DNA double-strand break repair-pathway choice in somatic mammalian cells.* Nature Reviews Molecular Cell Biology, 2019. **20**(11): p. 698-714.

4. Lieber, M.R., *The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway.* Annu Rev Biochem, 2010. **79**: p. 181-211.

5. Luijsterburg, M.S., et al., *A PALB2-interacting domain in RNF168 couples homologous recombination to DNA break-induced chromatin ubiquitylation.* Elife, 2017. **6**.

6. Mattiroli, F., et al., *The nucleosome acidic patch plays a critical role in RNF168-dependent ubiquitination of histone H2A.* Nat Commun, 2014. **5**: p. 3291.

7. Chang, H.H.Y., et al., *Non-homologous DNA end joining and alternative pathways to double-strand break repair.* Nat Rev Mol Cell Biol, 2017. **18**(8): p. 495-506.

8. Bhargava, R., D.O. Onyango, and J.M. Stark, *Regulation of Single-Strand Annealing and its Role in Genome Maintenance.* Trends Genet, 2016. **32**(9): p. 566-575.

9. Ceccaldi, R., B. Rondinelli, and A.D. D'Andrea, *Repair Pathway Choices and Consequences at the Double-Strand Break.* Trends Cell Biol, 2016. **26**(1): p. 52-64.

10. Noordermeer, S.M. and H. van Attikum, *PARP Inhibitor Resistance: A Tug-of-War in BRCA-Mutated Cells.* Trends Cell Biol, 2019. **29**(10): p. 820-834.

11. Orthwein, A., et al., *A mechanism for the suppression of homologous recombination in G1 cells.* Nature, 2015. **528**(7582): p. 422-6.

12. Escribano-Díaz, C., et al., *A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice.* Mol Cell, 2013. **49**(5): p. 872-83.

13. Lukas, J., C. Lukas, and J. Bartek, *More than just a focus: The chromatin response to DNA damage and its role in genome integrity maintenance.* Nat Cell Biol, 2011. **13**(10): p. 1161-9.

14. Narlikar, G.J., R. Sundaramoorthy, and T. Owen-Hughes, *Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes.* Cell, 2013. **154**(3): p. 490-503.

15. Dantuma, N.P. and H. van Attikum, *Spatiotemporal regulation of posttranslational modifications in the DNA damage response.* Embo j, 2016. **35**(1): p. 6-23.

16. Gibbs-Seymour, I., et al., *HPF1/C4orf27 Is a PARP-1-Interacting Protein that Regulates PARP-1 ADP-Ribosylation Activity.* Mol Cell, 2016. **62**(3): p. 432-442.

17. Smith, R., et al., *Poly(ADP-ribose)-dependent chromatin unfolding facilitates the association of DNA-binding proteins with DNA at sites of damage.* Nucleic Acids Res, 2019. **47**(21): p. 11250-11267.

18. Luijsterburg, M.S., et al., *PARP1 Links CHD2-Mediated Chromatin Expansion and H3.3 Deposition to DNA Repair by Non-homologous End-Joining.* Mol Cell, 2016. **61**(4): p. 547-562.

19. Rulten, S.L., et al., *PARP-3 and APLF function together to accelerate nonhomologous end-joining.* Mol Cell, 2011. **41**(1): p. 33-45.

20. Izhar, L., et al., *A Systematic Analysis of Factors Localized to Damaged Chromatin Reveals PARP-Dependent Recruitment of Transcription Factors.* Cell Rep, 2015. **11**(9): p. 1486-500.

21. Cassandri, M., et al., *Zinc-finger proteins in health and disease.* Cell Death Discov, 2017. **3**: p. 17071.

22. Miller, J., A.D. McLachlan, and A. Klug, *Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes.* Embo j, 1985. **4**(6): p. 1609-14.

23. Vilas, C.K., et al., *Caught with One's Zinc Fingers in the Genome Integrity Cookie Jar.* Trends Genet, 2018. **34**(4): p. 313-325.

24. Kluska, K., J. Adamczyk, and A. Krężel, *Metal binding properties, stability and reactivity of zinc fingers.* Coordination Chemistry Reviews, 2018. **367**: p. 18-64.

25. Laity, J.H., B.M. Lee, and P.E. Wright, *Zinc finger proteins: new insights into structural and functional diversity.* Curr Opin Struct Biol, 2001. **11**(1): p. 39-46.

26. Macrae, C.J., et al., *APLF (C2orf13) facilitates nonhomologous end-joining and undergoes ATM-dependent hyperphosphorylation following ionizing radiation.* DNA Repair (Amst), 2008. **7**(2): p. 292-302.

27. Hammel, M., et al., *An Intrinsically Disordered APLF Links Ku, DNA-PKcs, and XRCC4-DNA Ligase IV in an Extended Flexible Non-homologous End Joining Complex.* J Biol Chem, 2016. **291**(53): p. 26987-27006.

28. Grundy, G.J., et al., *APLF promotes the assembly and activity of non-homologous end joining protein complexes.* Embo j, 2013. **32**(1): p. 112-25.

29. Stogios, P.J., et al., *Sequence and structural analysis of BTB domain proteins.* Genome Biol, 2005. **6**(10): p. R82.

30. Thijssen, P.E., et al., *Mutations in CDCA7 and HELLS cause immunodeficiency-centromeric instability-facial anomalies syndrome.* Nat Commun, 2015. **6**: p. 7870.

31. Helfricht, A., et al., *Loss of ZBTB24 impairs non-homologous end-joining and class-switch recombination in patients with ICF syndrome.* Journal of Experimental Medicine, 2020.

32. Nicolai, S., et al., *ZNF281 is recruited on DNA breaks to facilitate DNA repair by non-homologous end joining.* Oncogene, 2020. **39**(4): p. 754-766.

33. Pieraccioli, M., et al., *ZNF281 contributes to the DNA damage response by controlling the expression of XRCC2 and XRCC4.* Oncogene, 2016. **35**(20): p. 2592-2601.

34. Musselman, C.A. and T.G. Kutateladze, *PHD fingers: epigenetic effectors and potential drug targets.* Mol Interv, 2009. **9**(6): p. 314-23.

35. Warmerdam, D.O., et al., *PHF6 promotes non-homologous end joining and G2 checkpoint recovery.* EMBO Rep, 2020. **21**(1): p. e48460.

36. Lan, L., et al., *The ACF1 complex is required for DNA double-strand break repair in human cells.* Mol Cell, 2010. **40**(6): p. 976-87.

37. Rother, M.B. and H. van Attikum, *DNA repair goes hip-hop: SMARCA and CHD chromatin remodellers join the break dance.* Philos Trans R Soc Lond B Biol Sci, 2017. **372**(1731).

38. Gamsjaeger, R., et al., *Sticky fingers: zinc-fingers as protein-recognition motifs.* Trends Biochem Sci, 2007. **32**(2): p. 63-70.

39. Feng, L. and J. Chen, *The E3 ligase RNF8 regulates KU80 removal and NHEJ repair.* Nature Structural & Molecular Biology, 2012. **19**(2): p. 201-206.

40. Ishida, N., et al., *Ubiquitylation of Ku80 by RNF126 Promotes Completion of Nonhomologous End Joining-Mediated DNA Repair.* Mol Cell Biol, 2017. **37**(4).

41. Cao, L.L., et al., *ATM-mediated KDM2A phosphorylation is required for the DNA damage repair.* Oncogene, 2016. **35**(3): p. 301-13.

42. Bueno, M.T.D., et al., *Recruitment of lysine demethylase 2A to DNA double strand breaks and its interaction with 53BP1 ensures genome stability.* Oncotarget, 2018. **9**(22): p. 15915-15930.

43. Phillips, J.E. and V.G. Corces, *CTCF: master weaver of the genome.* Cell, 2009. **137**(7): p. 1194-211.

44. Hwang, S.Y., et al., *CTCF cooperates with CtIP to drive homologous recombination repair of double-strand breaks.* Nucleic Acids Research, 2019. **47**(17): p. 9160-9179.

45. Lang, F., et al., *CTCF prevents genomic instability by promoting homologous recombination-directed DNA double-strand break repair.* Proc Natl Acad Sci U S A, 2017. **114**(41): p. 10912-10917.

46. Hilmi, K., et al., *CTCF facilitates DNA double-strand break repair by enhancing homologous recombination repair.* Sci Adv, 2017. **3**(5): p. e1601898.

47. Gong, Y., et al., *PHF11 promotes DSB resection, ATR signaling, and HR.* Genes Dev, 2017. **31**(1): p. 46-58.

48. Opipari, A.W., Jr., M.S. Boguski, and V.M. Dixit, *The A20 cDNA induced by tumor necrosis factor alpha encodes a novel type of zinc finger protein.* J Biol Chem, 1990. **265**(25): p. 14705-8.

49. Bosanac, I., et al., *Ubiquitin binding to A20 ZnF4 is required for modulation of NF-κB signaling.* Mol Cell, 2010. **40**(4): p. 548-57.

50. Yang, C., et al., *A20/TNFAIP3 Regulates the DNA Damage Response and Mediates Tumor Cell Resistance to DNA-Damaging Therapy.* Cancer Res, 2018. **78**(4): p. 1069-1082.

51. Ali, M.A.M., et al., *RYBP Is a K63-Ubiquitin-Chain-Binding Protein that Inhibits Homologous Recombination Repair.* Cell Rep, 2018. **22**(2): p. 383-395.

52. Kanayama, A., et al., *TAB2 and TAB3 activate the NF-kappaB pathway through binding to polyubiquitin chains.* Mol Cell, 2004. **15**(4): p. 535-48.

53. Masson, J.Y., et al., *Identification and purification of two distinct complexes containing the five RAD51 paralogs.* Genes Dev, 2001. **15**(24): p. 3296-307.

54. Bienko, M., et al., *Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis.* Science, 2005. **310**(5755): p. 1821-4.

55. Notenboom, V., et al., *Functional characterization of Rad18 domains for Rad6, ubiquitin, DNA binding and PCNA modification.* Nucleic Acids Res, 2007. **35**(17): p. 5819-30.

56. Rodrigue, A., et al., *Interplay between human DNA repair proteins at a unique double-strand break in vivo.* Embo j, 2006. **25**(1): p. 222-31.

57. Huang, J., et al., *RAD18 transmits DNA damage signalling to elicit homologous recombination repair.* Nat Cell Biol, 2009. **11**(5): p. 592-603.

58. Argentaro, A., et al., *Structural consequences of disease-causing mutations in the ATRX-DNMT3-DNMT3L (ADD) domain of the chromatin-associated protein ATRX.* Proc Natl Acad Sci U S A, 2007. **104**(29): p. 11939-44.

59. Juhász, S., et al., *ATRX Promotes DNA Repair Synthesis and Sister Chromatid Exchange during Homologous Recombination.* Mol Cell, 2018. **71**(1): p. 11-24.e7.

60. Leung, J.W., et al., *ZMYM3 regulates BRCA1 localization at damaged chromatin to promote DNA repair.* Genes Dev, 2017. **31**(3): p. 260-274.

2

61. Coleman, K.A. and R.A. Greenberg, *The BRCA1-RAP80 complex regulates DNA repair mechanism utilization by restricting end resection.* J Biol Chem, 2011. **286**(15): p. 13669-80.

62. Irminger-Finger, I., M. Ratajska, and M. Pilyugin, *New concepts on BARD1: Regulator of BRCA pathways and beyond.* Int J Biochem Cell Biol, 2016. **72**: p. 1-17.

63. Densham, R.M., et al., *Human BRCA1–BARD1 ubiquitin ligase activity counteracts chromatin barriers to DNA resection.* Nature Structural & Molecular Biology, 2016. **23**(7): p. 647-655.

64. Zhao, W., et al., *BRCA1–BARD1 promotes RAD51-mediated homologous DNA pairing.* Nature, 2017. **550**(7676): p. 360-365.

65. Myler, L.R., et al., *Single-Molecule Imaging Reveals How Mre11-Rad50-Nbs1 Initiates DNA Break Repair.* Mol Cell, 2017. **67**(5): p. 891-898.e4.

66. Schmidt, C.K., et al., *Systematic E2 screening reveals a UBE2D-RNF138-CtIP axis promoting DNA repair.* Nat Cell Biol, 2015. **17**(11): p. 1458-1470.

67. Caron, P., J. van der Linden, and H. van Attikum, *Bon voyage: A transcriptional journey around DNA breaks.* DNA Repair (Amst), 2019. **82**: p. 102686.

68. Rona, G., et al., *PARP1-dependent recruitment of the FBXL10-RNF68-RNF2 ubiquitin ligase to sites of DNA damage controls H2A.Z loading.* Elife, 2018. **7**.

69. Spellmon, N., et al., *Structure and function of SET and MYND domain-containing proteins.* Int J Mol Sci, 2015. **16**(1): p. 1406-28.

70. Gong, F. and K.M. Miller, *Double duty: ZMYND8 in the DNA damage response and cancer.* Cell Cycle, 2018. **17**(4): p. 414-420.

71. Spruijt, C.G., et al., *ZMYND8 Co-localizes with NuRD on Target Genes and Regulates Poly(ADP-Ribose)-Dependent Recruitment of GATAD2A/NuRD to Sites of DNA Damage.* Cell Rep, 2016. **17**(3): p. 783-798.

72. Gong, F., et al., *Histone demethylase KDM5A regulates the ZMYND8-NuRD chromatin remodeler to promote DNA repair.* J Cell Biol, 2017. **216**(7): p. 1959-1974.

73. Balmus, G., et al., *ATM orchestrates the DNA-damage response to counter toxic non-homologous end-joining at broken replication forks.* Nature Communications, 2019. **10**(1): p. 87.

74. Vancurova, M., et al., *PML nuclear bodies are recruited to persistent DNA damage lesions in an RNF168-53BP1 dependent manner and contribute to DNA repair.* DNA Repair (Amst), 2019. **78**: p. 114-127.

75. Wu, S., et al., *A YY1-INO80 complex regulates genomic stability through homologous recombination-based repair.* Nat Struct Mol Biol, 2007. **14**(12): p. 1165-72.

76. Peng, M., et al., *BACH1 is a DNA repair protein supporting BRCA1 damage response.* Oncogene, 2006. **25**(15): p. 2245-2253.

77. Liu, X.-S., et al., *LRF maintains genome integrity by regulating the non-homologous end joining pathway of DNA repair.* Nature Communications, 2015. **6**(1): p. 8325.

78. Ismail, I.H., et al., *The RNF138 E3 ligase displaces Ku to promote DNA end resection and regulate DNA repair pathway choice.* Nat Cell Biol, 2015. **17**(11): p. 1446-57.

79. Khoury-Haddad, H., et al., *PARP1-dependent recruitment of KDM4D histone demethylase to DNA damage sites promotes double-strand break repair.* Proc Natl Acad Sci U S A, 2014. **111**(7): p. E728-37.

80. Xie, H.Y., et al., *Dimerization of MORC2 through its C-terminal coiled-coil domain enhances chromatin dynamics and promotes DNA repair.* Cell Commun Signal, 2019. **17**(1): p. 160.
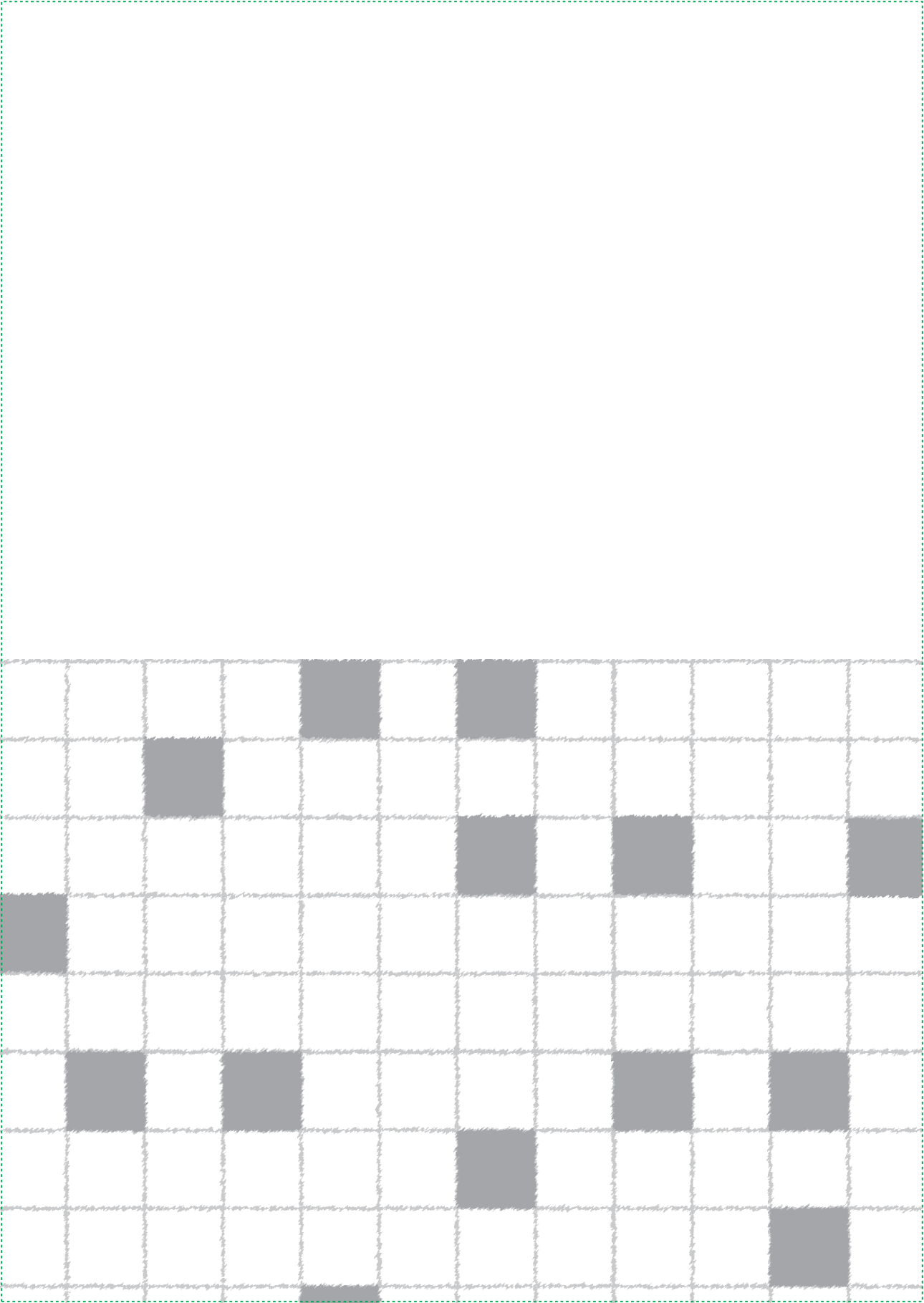
81. Nowsheen, S., et al., *L3MBTL2 orchestrates ubiquitin signalling by dictating the sequential recruitment of RNF8 and RNF168 after DNA damage.* Nat Cell Biol, 2018. **20**(4): p. 455-464.

82. Galanty, Y., et al., *Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks.* Nature, 2009. **462**(7275): p. 935-9.

83. Gong, F., et al., *Screen identifies bromodomain protein ZMYND8 in chromatin recognition of transcription-associated DNA damage that promotes homologous recombination.* Genes Dev, 2015. **29**(2): p. 197-211.

84. Chen, Y.J., et al., *SMYD3 Promotes Homologous Recombination via Regulation of H3K4-mediated Gene Expression.* Sci Rep, 2017. **7**(1): p. 3842.

85. Ray Chaudhuri, A. and A. Nussenzweig, *The multifaceted roles of PARP1 in DNA repair and chromatin remodelling.* Nat Rev Mol Cell Biol, 2017. **18**(10): p. 610-621.

86. Liu, C., et al., *CHFR is important for the first wave of ubiquitination at DNA damage sites.* Nucleic Acids Res, 2013. **41**(3): p. 1698-710.

87. Cao, L.L., et al., *ATM-mediated KDM2A phosphorylation is required for the DNA damage repair.* Oncogene, 2016. **35**(3): p. 301-313.

88. Martín, V., et al., *Sws1 is a conserved regulator of homologous recombination in eukaryotic cells.* Embo j, 2006. **25**(11): p. 2564-74.

89. Ogiwara, H., et al., *Histone acetylation by CBP and p300 at double-strand break sites facilitates SWI/SNF chromatin remodeling and the recruitment of non-homologous end joining factors.* Oncogene, 2011. **30**(18): p. 2135-46.

90. Nakajima, S., et al., *Ubiquitin-specific protease 5 is required for the efficient repair of DNA double-strand breaks.* PLoS One, 2014. **9**(1): p. e84899.

91. Mosbech, A., et al., *The deubiquitylating enzyme USP44 counteracts the DNA double-strand break response mediated by the RNF8 and RNF168 ubiquitin ligases.* J Biol Chem, 2013. **288**(23): p. 16579-87.

Table 1. Overview of ZnF proteins involved in DSB repair

ZnF domains, the number of genes containg a particular domain, the number of genes implicated in DSB repair and their known binding affinities are indicated. Examples of ZnF proteins involved in DSB repair are shown, several of which play poorly understood roles in NHEJ (marked as [1]), HR (marked as [2]) or both NHEJ and HR (marked as [3]). Well-characterized ZnF domain-containing proteins in DSB repair, which are discussed in in this review, are marked with an asterisk (*).

Domain name	Number of genes	Number of genes in DSB repair	Binding specificities	Examples	References
A20	7	1	Ubiquitin	A20/TNFAIP3[2,*]	[50]
ADD	4	1	DNA, modified histones	ATRX[2,*]	[59]
AN-1	8	0	DNA, RNA, Protein, Lipid		
B-Box	75	5	DNA	PML[2]	[74]
BED	6	0	DNA		
BTB/POZ	139	6	DNA, Protein	ZBTB7A[1], YY1[1], BACH1[2]	[75-77]
C2C2	6	1	DNA	TCEA[3]	[20]
C2H2	759	13	DNA, RNA, Protein, Lipid, Methylated DNA	ZBTB24[1,*], ZNF281[1,*], PHF11[2,*], ZNF830[2]	
C2HC	6	3	DNA	RNF138[2,*]	[78]
C2CH	13	0	DNA		
C3H1	59	2	RNA	ZC3H11A[3]	[20]
C4	56	2	DNA	ESR2[3], NR1H4[3]	[20]
C5HC2	24	3	Modified histones	KDM4D[2]	[79]
CXXC	12	2	DNA	KDM2A[1,*]	[42]
CCHC	38	0	DNA, RNA		
CCHHC	7	0	DNA		
CHHC	4	0	RNA		
CW	7	1	Modified histones	MORC2[3]	[80]
DBF	3	0	DNA, Protein		

Domain name	Number of genes	Number of genes in DSB repair	Binding specificities	Examples	References
ePHD	23	1	Modified histones	PHF11[2,*]	[47]
FCS	5	2	RNA	L3MBTL2[1]	[81]
FYVE	32	0	Lipid, methylated DNA		
GATA	15	3	DNA	MTA2[3]	[20]
HIT	6	0	Protein, DNA		
KRAB	362	1	DNA, Ubiquitin	ZNF829[3]	[20]
LIM	71	0	Protein		
MATRIN	8	1	RNA	ZMAT1[3]	[20]
MIZ (SP-RING)	7	3	SUMO	PIAS1 and PIAS4[3]	[82]
MYM	6	1	SUMO	ZMYM3[2,*]	[60]
MYND	21	2	Protein	ZMYND8[2,*], SMYD3[2]	[83, 84]
PARP	2	2	DNA	PARP1[3]	[85]
PBZ	2	2	Poly(ADP)ribose	APLF[1,*], CHFR[3]	[28, 86]
PHD	71	27	Modified histones	PHF6[1,*], ACF1[1,*], KDM2A[2,*]	[35, 36, 87]
RAD18 (UBZ4)	8	3	Ubiquitin	RAD18[2,*]	[57]
RBZ (RANBP2)	23	4	Ubiquitin	RYBP[2,*]	[51]
RING	282	28	Protein, Ubiquitin	RNF8[1,*], RNF126[1,*], BARD1[2,*], RNF138[2,*], FRUCC[2,*]	[39, 40, 63, 68, 78]
SCA7	6	0	Protein		
SWIM	9	1	DNA, Protein	ZSWIM7[2]	[88]
TAZ	2	2	Protein, DNA	CBP/p300[1]	[89]
TFIIB	3	0	DNA		
TFIIS	6	1	DNA, RNA, Protein	TCEA1[3]	[20]
TRAF	23	0	Ubiquitin, Protein		
UBP	14	4	Ubiquitin	USP44[3], USP5[2]	[90, 91]
ZBR	2	0	Protein		
ZZ	18	3	Protein	CBP/p300[1]	[89]

2

CHAPTER 3

Zinc finger protein ZNF384 is an adaptor of KU to DNA during classical non-homologous end-joining

Jenny Kaur Singh, Rebecca Smith, Magdalena B. Rother, Anton J.L. de Groot, Wouter W. Wiegant, Kees Vreeken, Ostiane D'Augustin, Robbert Q. Kim, Haibin Qian, Przemek M. Krawczyk, Román González-Prieto, Alfred C.O. Vertegaal, Meindert Lamers, Sébastien Huet and Haico van Attikum.

ABSTRACT

DNA double-strand breaks (DSBs) are among the most deleterious types of DNA damage as they can lead to mutations and chromosomal rearrangements, which underlie cancer development. Classical non-homologous end-joining (cNHEJ) is the dominant pathway for DSB repair in human cells, involving the DNA-binding proteins XRCC6 (KU70) and XRCC5 (KU80). Other DNA-binding proteins such as Zinc Finger (ZnF) domain-containing proteins have also been implicated in DNA repair, but their role in cNHEJ remained elusive. Here we show that ZNF384, a member of the C2H2 family of ZnF proteins, binds DNA ends *in vitro* and is recruited to DSBs *in vivo*. ZNF384 recruitment requires the poly(ADP-ribosyl) polymerase 1 (PARP1)-dependent expansion of damaged chromatin, followed by binding of its C2H2 motifs to the exposed DNA. Moreover, ZNF384 interacts with KU70/KU80 via its N-terminus, thereby promoting KU70/KU80 assembly and the accrual of downstream cNHEJ factors, including APLF and XRCC4/LIG4, for efficient repair at DSBs. Altogether, our data suggest that ZNF384 acts as a 'Ku-adaptor' that binds damaged DNA and KU70/KU80 to facilitate the build-up of a cNHEJ repairosome, highlighting a role for ZNF384 in DSB repair and genome maintenance.

Keywords: DNA double-strand break (DSB) repair, classical non-homologous end-joining (cNHEJ), DNA binding, ZNF384, KU70/KU80

INTRODUCTION

DNA double-strand breaks (DSBs) represent one of the most toxic lesions that can occur in the human genome. If left unrepaired or repaired incorrectly, they can lead to loss of genetic information, thereby contributing to the development of diseases, including cancer [1]. In order to maintain genomic stability, cells have evolved pathways for the signaling and repair of these DSBs [1]. DSB repair can occur by either homologous recombination (HR) or non-homologous end joining (NHEJ). HR is the more faithful repair pathway, which is active in the S and G2 phase of the cell cycle. It requires end-resection to form large stretches of single-stranded DNA (ssDNA), which in turn become coated by the ssDNA-binding complex RPA and the recombinase RAD51. Collectively, these and several other auxiliary factors contribute to HR by using the sister chromatid as a template for repair [2]. In contrast, the dominant repair pathway in human cells is canonical non-homologous end joining (cNHEJ), which requires minimal DNA-end processing and is initiated by the binding of KU70/KU80 heterodimers to the broken ends, followed by activation of DNA-PKcs and recruitment of APLF via its conserved KU-binding motif (KBM). This facilitates the assembly of the XLF-XRCC4-LIG4 complex, which stimulated by PAXX, ligates the broken ends predominantly in an error-free manner [3]. When cNHEJ is disabled, DSB repair can also occur via alternative non-homologous end joining (aNHEJ), which seals the broken ends in an error-prone fashion by microhomology usage and in a manner dependent on the XRCC1-Ligase III complex or POLQ [3]. Alternatively, in the case of more extensive end-resection microhomology usage may lead to deleterious, RAD52-dependent repair of DSBs via single-strand annealing (SSA) [4].

Efficient detection and repair of DSBs is complicated by the packaging of DNA into chromatin. ATP-dependent chromatin remodeling enzymes and a wide plethora of enzymes that induce post-translational modifications (PTMs) on damaged chromatin, including but not limited to acetylation, methylation and ubiquitylation, are therefore required to change chromatin structure at DSB sites to facilitate repair [5-7]. One of these enzymes is poly(ADP-ribosyl) polymerase 1 (PARP1), which becomes activated upon binding to DNA breaks and promotes chromatin expansion by the formation of poly(ADP-ribose) (PAR) chains on itself and adjacent nuclear proteins, such as histones, as well by promoting the recruitment of ATP-dependent chromatin remodelers in the vicinity of these breaks [8-10]. This increases chromatin accessibility and the recruitment of several DSB repair proteins, including KU70/KU80 and XRCC4, via direct PAR binding or DNA binding [9, 11].

3

Interestingly, a number of transcription factors have also been shown to localize at sites of DNA damage either in a PARP/PAR-dependent manner or via their DNA binding domains [11, 12]. Their role in DSB repair, is however, largely unknown. One such class of transcription factors are Zinc Finger (ZnF) domain-containing proteins. ZnF domains exist in approximately 5% of all human proteins and bind to a large variety of substrates including DNA, RNA, lipids and post-translational modifications (PTMs) [13, 14]. Due to their versatile binding ability, ZnF proteins play roles in different cellular processes such as transcription regulation, signal transduction and cell migration [15]. Interestingly, recent studies have implicated ZnF domain-containing proteins as new players in DSB repair [13]. For instance, ZMYND8 was found to play a role in transcription repression during DSB repair via HR [16], whereas ZNF830 promotes HR by facilitating RBBP8 (CTIP)-dependent DNA end resection [17]. ZNF281, on the other hand, was shown to promote XRCC4-dependent NHEJ of DSBs [18]. Together, these findings suggest a more important role for ZnF domain proteins in DSB repair than previously anticipated, although their mode of action is still poorly understood.

Here we describe an important regulatory role for the C2H2-type ZnF protein ZNF384 in DSB repair by cNHEJ. ZNF384 is recruited to sites of DNA damage in a manner dependent on PARP/PAR-mediated chromatin expansion followed by binding to the exposed DNA via its internal C2H2 domain. Moreover, ZNF384 physically interacts with KU70/KU80 via its N-terminus and both its interaction with DNA and KU70/KU80 are critical for efficient KU70/KU80 loading. This in turn allows for the assembly of a complete repairosome that includes cNHEJ proteins such as APLF and XRCC4/LIG4, thereby facilitating cNHEJ. Collectively, our data show that zinc-finger protein ZNF384 is an adapter of KU to DNA during DSB repair via cNHEJ.

RESULTS

ZNF384 is recruited to DNA damage sites and interacts with NHEJ proteins

ZNF384 was among the candidate ZnF proteins that localize to sites of DNA damage induced by laser micro-irradiation [12]. To validate this finding, we transiently co-expressed GFP-tagged ZNF384 (isoform 2, containing 6 C2H2 motifs) and the DNA damage sensor mCherry-NBS1 in *ZNF384* knockout (KO) U2OS cells (Supplementary Fig. 1A) and measured their recruitment to sites of DNA damage induced by multiphoton irradiation using live cell imaging (Supplementary Fig. 1B). GFP-ZNF384 was recruited

to NBS1-marked DNA damage sites within 1 minute and remained enriched at these sites for at least 3 minutes (Fig. 1A). We also observed the accumulation of endogenous ZNF384 at UV-A laser-induced DNA damage, which was completely abolished following ZNF384 knockdown, confirming the specificity of the ZNF384 antibody (Supplementary Fig. 1C). Since multiphoton and UV-A laser micro-irradiation may induce a variety of lesions other than DSBs, we next examined whether ZNF384 is specifically recruited to DSBs. First, we monitored its accumulation at chromatin regions micro-irradiated by ultra-soft X-rays (USX) [19]. Endogenous ZNF384 accumulated at USX-induced DSBs, co-localizing with γH2AX (Supplementary Fig. 1D-E), and with the core NHEJ proteins KU70 and XRCC4 (Supplementary Fig. 1F-G). Second, we measured the co-localization between ZNF384 and γH2AX at *Asi*SI nuclease-induced DSBs by proximity ligation assay (PLA) (Fig. 1B). We observed a clear co-localization between ZNF384 and γH2AX at these DSBs (Fig. 1C), as well as between TP53BP1 (53BP1) and γH2AX (Supplementary Fig. 1H-I and [20], showing the validity of the approach. Together, these observations demonstrate that ZNF384 is recruited to DSBs.

To gain insight into ZNF384's function at sites of DNA damage, we aimed to identify possible interaction partners of ZNF384. To this end, we generated U2OS Flp-In/T-Rex cells stably expressing inducible GFP-tagged ZNF384 or GFP-NLS, and performed GFP-trap-based pulldowns followed by label-free mass spectrometry (MS) (Fig. 1D). Our analysis revealed that ZNF384 interacts with 24 proteins that were at least 2-fold enriched in GFP-ZNF384 pulldowns when compared to those of GFP-NLS (Fig. 1D and Supplementary Table 1). Interestingly, KU70/KU80 and PARP1 belonged to the top-interactors, all of which regulate DSB repair via cNHEJ [9, 21]. GFP pulldowns coupled to western blot analysis confirmed that GFP-tagged ZNF384 interacts with endogenous KU70, KU80 and PARP1 (Fig. 1E), while reciprocal pulldowns of both GFP-KU80 (Fig. 1F) and GFP-PARP1 (Fig. 1G) revealed interactions with endogenous ZNF384. Moreover, we also confirmed the ZNF384-KU70 interaction endogenously (Fig. 1H), using *ZNF384* knockdown cells to show specificity of the ZNF384 signal in the KU70 immunoprecipitation (Supplementary Fig. 1J). To rule out an indirect interaction between these proteins we purified His-MBP-tagged ZNF384 and addressed its ability to bind KU70/KU80 *in vitro*. In agreement with our *in vivo* pulldown results, we found that recombinant ZNF384 bound recombinant KU70/KU80, demonstrating a direct protein-protein interaction (Fig. 1I). These findings suggest that ZNF384 forms a complex with KU70/KU80 and PARP1, manifesting a possible role for ZNF384 in NHEJ.
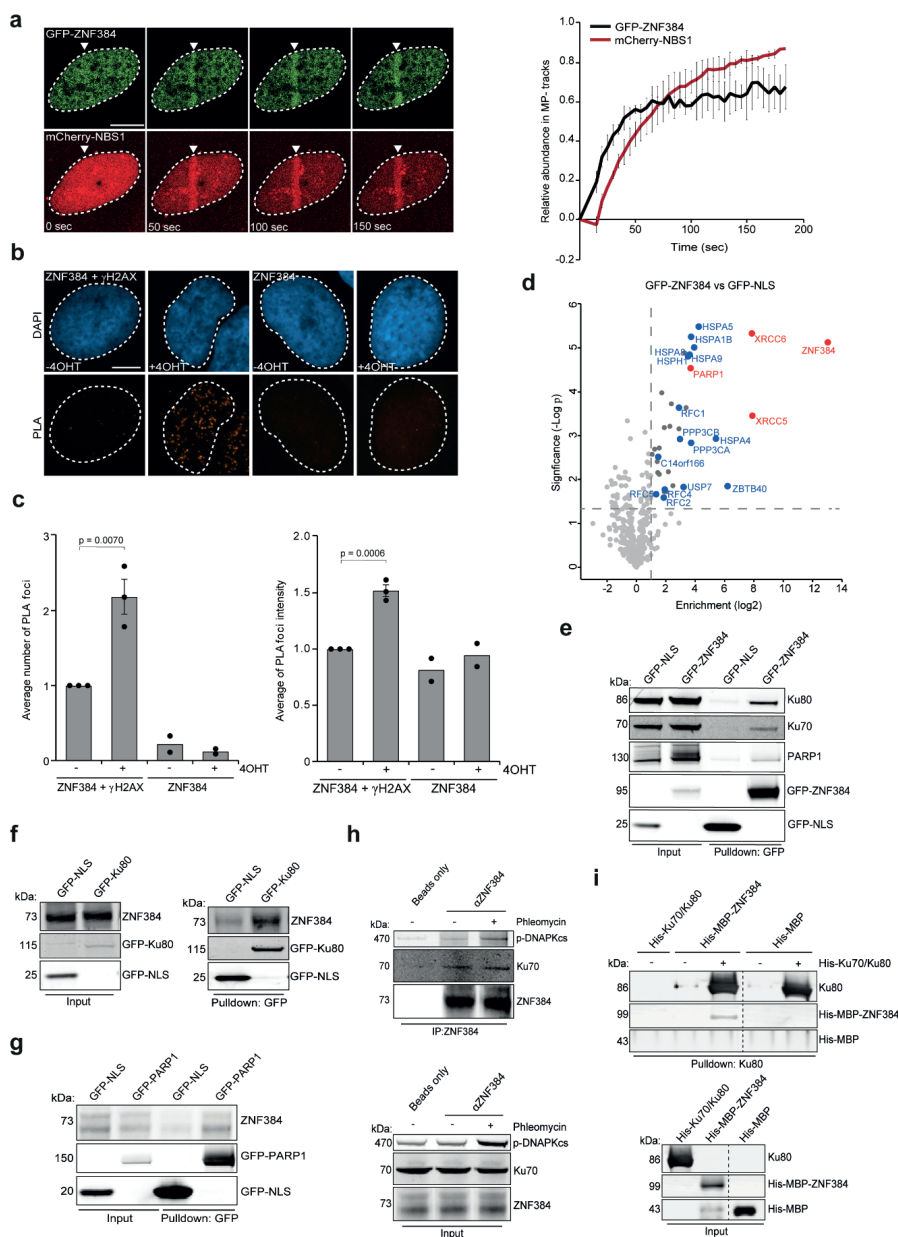
Figure 1. ZNF384 is recruited to DNA damage sites and interacts with NHEJ proteins. (A) Recruitment of GFP-ZNF384 to 800 nm multiphoton tracks in U2OS Flp-In/T-Rex *ZNF384* KO cells. mCherry-NBS1 was used as a damage marker (left panel). White triangles indicate irradiated regions. Quantification of the data is plotted on a time scale as relative abundance in tracks. Peak values were set to 1. The graph represents the mean ±SD of >30 cells acquired in 2-3 independent

experiments (right panel). (B) PLA of ZNF384 and γH2AX in *Asi*SI–ER-U2OS cells treated with 4-OHT for DSB induction. PLA foci were quantified after 5 hours of DSB induction. (C) Quantification of B. The mean ±SEM of PLA foci formation and foci intensity from >200 cells acquired in 2-3 independent experiments are shown. Statistical significance was calculated using the two-tailed Student's *t* test. (D) Volcano plot depicting the statistical differences of the MS analysis on GFP-ZNF384 versus GFP-NLS pulldowns. The enrichment is plotted on the x-axis and the significance (t-test -log2 p-value) is plotted on the y-axis. NHEJ factors are shown in red and several hits are shown in blue (see also Supplementary Table 1). (E) Pulldowns of the indicated GFP fusion proteins in U2OS Flp-In/T-Rex cells. Blots were probed for GFP, KU70, KU80 and PARP1. (F) Pulldowns of the indicated GFP fusion proteins in Hela cells. Blots were probed for GFP and ZNF384. (G) Pulldowns of the indicated GFP fusion proteins in Hela cells. Blots were probed for GFP and ZNF384. (H) Immunoprecipitation (IP) of endogenous ZNF384 from 500µM Phleomycin treated U2OS cells. Control IP contained beads only. Blots were probed for ZNF384, p-DNA-PKcs (S2056) and KU70. (I) In *vitro* KU80 pulldown in the presence or absence of His-KU70/KU80 and His-MBP or His-MBP-ZNF384. Control IP contained beads only. Blots were probed for KU80 and MBP. Scale bar 5 µm. Source data are provided as Source Data file.

ZNF384 recruitment to DNA damage sites requires the activity of PARP1

Because of ZNF384's interaction with KU70/KU80 and PARP1, we first analyzed whether it is recruited to DNA breaks via KU80. We observed that GFP-ZNF384 recruitment to UV-A laser micro-irradiation induced DNA damage, as well as that of the DNA damage marker NBS1 (Supplementary Fig. 2A-B), remained unchanged in cells depleted of KU80 or its interaction partner in the DNA-PK complex, DNA-PKcs kinase (Supplementary Fig. 2A-B) [22]. Moreover, we also found endogenous ZNF384 at γH2AX marked UV-A laser-inflicted DNA damage to remain unaffected in cells depleted for KU70 (Supplementary Fig. 2C-D), KU80 or DNA-PKcs (Supplementary Fig. 2E), which was confirmed in *Xrcc6* KO hamster cells (Supplementary Fig. 2F-G, ruling out effects of incomplete knockdown. In addition, inhibition of DNA-PKcs kinase activity did not exert any effect on the recruitment of GFP-ZNF384 (Supplementary Fig. 2H) and endogenous ZNF384 (Supplementary Fig. 2I-J), suggesting that ZNF384 recruitment is independent of DNA-PK.

Next, we treated cells with the PARP1 and PARP2 inhibitor (PARPi) olaparib (Fig. 2A), and found this treatment to impair the recruitment of GFP-ZNF384 at multiphoton laser micro-irradiation (Fig. 2A), as well as UV-A laser micro-irradiation induced DNA damage (Supplementary Fig. 3A-B). Similarly, knockout of PARP1 alone or in combination with PARP2 completely impaired ZNF384 recruitment, whereas knockout of PARP2 alone, and knockdown of PARP3 had no major effect (Fig. 2B and Supplementary Fig. 3C-F). PARP3 knockdown, however, impaired KU80 recruitment as expected (Supplementary Fig. 3G-H) [23]. Importantly, we previously showed that PARP1 itself is still recruited

to sites of DNA damage in PARPi-treated cells [24], suggesting that the recruitment of ZNF384 does not involve a physical interaction between ZNF384 and PARP1, but rather relies on PARP1's activity. Possibly, this interaction is important for the PARP1-dependent PARylation of ZNF384, which we and others observed in response to DSB induction (Supplementary Fig. 3I), in accordance with a previous report [25]. Collectively, these results suggest that ZNF384 is rapidly recruited to DSB-containing tracks in a manner dependent on the activity of PARP1.

PARP1-dependent chromatin unfolding facilitates DNA binding of ZNF384 at DNA damage sites

We next asked whether the PARP1 activity-dependent recruitment of ZNF384 to DNA breaks could be due to direct binding of ZNF384 to PARP1-generated PAR moieties. To investigate this, we used a previously established fluorescence three-hybrid assay (Supplementary Fig. 4A) [26]). This assay measures the ability of a lacO-anchored putative 'ADP-ribose-binding' protein of interest to interact with PARylated PARP1 that is naturally released from sites of laser-induced DNA damage and is then free to diffuse and bind to the lacO-anchored PAR-binder. Indeed, we observed recruitment of PARylated PARP1 to the lacO-anchored macrodomain of macroH2A1.1, a well-characterized PAR binding protein [11, 27], which was abolished following treatment with PARPi , but remained unaffected by treatment with PARGi (likely to due to the availability of a limited number of lacO-anchored macrodomain molecules) (Supplementary Fig. 4B-D). In contrast, we did not observe an interaction between lacO-anchored ZNF384 and PARylated PARP1 (Supplementary Fig. 4E,F). To corroborate and extend these findings, we used a second independent approach that can discriminate whether ZNF384 is recruited by binding to PAR, or to DNA that becomes exposed upon PAR-dependent chromatin relaxation. In this assay *ZNF384* KO cells were micro-irradiated and, 240 seconds post-irradiation, after the completion of the initial wave of PARP/PAR-dependent chromatin relaxation [9, 26, 28], PARPi was added to acutely block PARP enzymatic activity. Under these conditions PAR-binding proteins are rapidly released from the damaged area, while proteins that bind DNA are maintained [11]. Indeed, we observed that the WWE domain of RNF146, which is a known PAR-binder [29], was rapidly released from the irradiated area (Fig. 2C-D and Supplementary Fig. 5A). While its binding was nearly completely reversed at 450s post-irradiation due to rapid degradation of PAR (Supplementary Fig. 5A), counteracting the removal of PAR

3

chains by adding PARGi to the PARPi-treated cells preserved the accumulation of the WWE domain at sites of DNA damage (Supplementary Fig. 5A-B). In contrast, PARPi treatment did not revert the recruitment of GFP-ZNF384 at the damaged area (Fig. 2E-F), suggesting that ZNF384 does not bind PAR, but rather associates with DNA. To corroborate these findings, we investigated whether DNA binding of ZNF384 depends on PARP/PAR-dependent chromatin relaxation, which we and others have shown to facilitate the association of DNA-binding proteins with DNA at sites of damage [9, 26, 28]. We found that overexpression of the ATP-dependent chromatin remodeler ALC1, which enhances chromatin unfolding without affecting PAR-signaling (Supplementary Fig. 5C and), increased ZNF384 accumulation at sites of DNA damage when compared to that after overexpression of ATPase-dead (E175Q) ALC1 (Fig. 2G). PARPi treatment inhibited ZNF384 recruitment in both WT ALC1 and ATPase-dead (E175Q) ALC1 overexpressing cells (Fig. 2G), consistent with the PARP-dependent recruitment of these proteins [8]. Interestingly, we did not find a significant enrichment of ZNF384 in chromatin-enriched extracts from cells treated with phleomycin (Supplementary Fig. 5D-E), nor a change in ZNF384 turnover at DNA lesions as measured by FRAP (Supplementary Fig. 5F-H), suggesting that the binding of ZNF384 to damaged DNA is not qualitatively different from its binding to undamaged DNA and is mostly triggered by the increased accessibility of DNA consecutive to PAR-driven chromatin relaxation. This behaviour of ZNF384 is comparable to that of the DNA binding domain BZIP of C/EBPa (Supplementary Fig. 5I and [11]) and the chromatin remodeler CHD4 (Supplementary Fig. 5I and [26]), which were both shown to recruit to DNA lesions due to increased accessibility of damaged DNA through PAR-dependent chromatin unfolding. In contrast, the ZNF384 interaction partners and DNA-end binding proteins KU70/KU80 showed a clear slowing of its turnover following DNA damage (Supplementary Fig. 5J). Thus, ZNF384 recruitment is dependent on PAR-dependent chromatin unfolding, allowing ZNF384 to bind to the exposed damaged DNA.
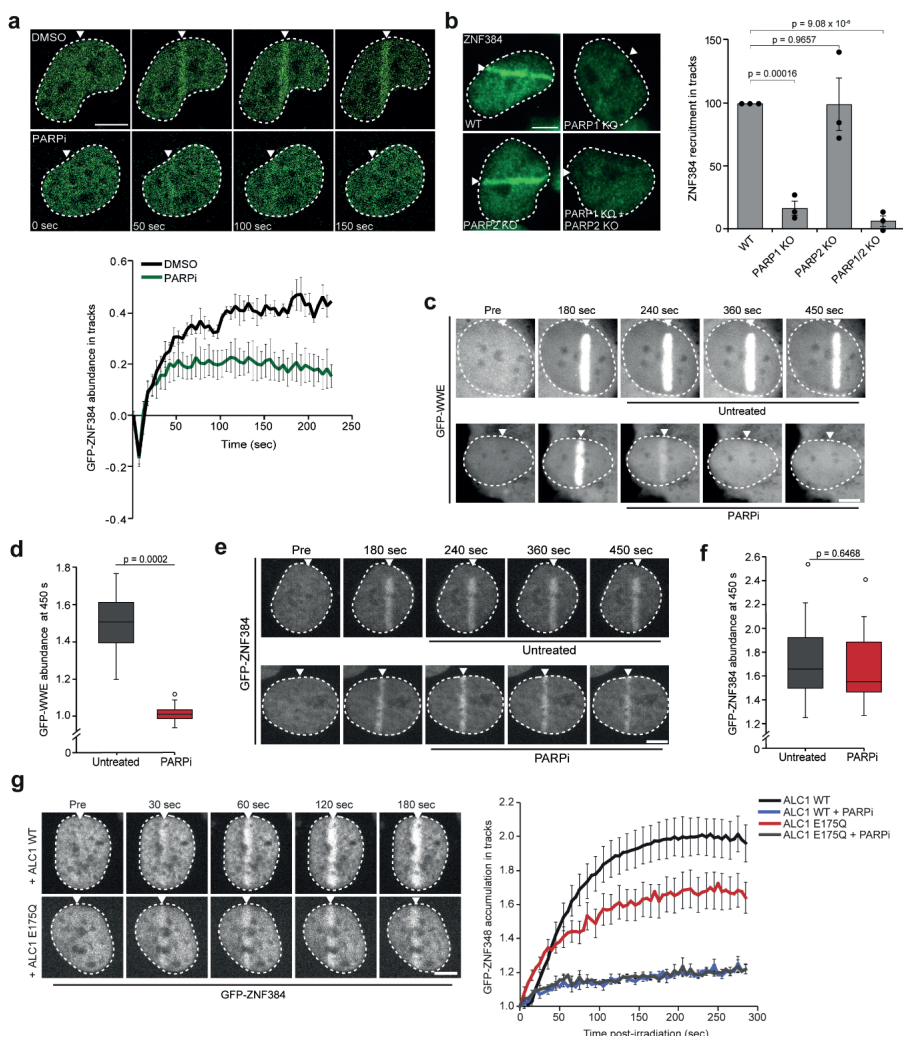
Figure 2. PARP1 activity facilitates DNA binding of ZNF384 at sites of damage. (A) GFP-ZNF384 recruitment to 800 nm multiphoton tracks in stable U2OS Flp-In/T-Rex cells treated with PARP inhibitor (PARPi) for 1 hour before micro-irradiation (top panel). Quantification of the data is presented as the mean ±SEM of >35 cells acquired in 3 independent experiments (bottom panel). (B) ZNF384 recruitment to 365 nm UV-A tracks 10 minutes after DNA damage induction in BrdU-sensitized wildtype (WT) and indicated KO U2OS cells (left panel). The mean ±SEM of >180 cells from 3 independent experiments is shown (right panel). (C) Confocal images showing accumulation of GFP-WWE at sites of 405 nm laser micro-irradiation in Hoechst-sensitized U2OS cells. Cells were left untreated or treated with PARPi 180 seconds after DNA damage induction. (D) Boxplot limits correspond to the 25th and 75th percentiles and the center line in the box indicates the median value of the accumulation of GFP-WWE at 450 sec post irradiation from 23-25 cells from a representative of 3 independent experiments. (E) As in C, except for GFP-ZNF384. (F) As in D, except for GFP-ZNF384 from 21-27 cells. (G) GFP-ZNF384 recruitment to 405 nm laser tracks in U2OS cells overexpressing iRFP-ALC1 wildtype (WT) and iRFP-ALC1 ATPase-dead (E175Q)

treated with PARPi for 1 hour before micro-irradiation (left panel). GFP-ZNF384 recruitment is displayed as intensity integrated over the damaged area (right panel). The mean ±SEM from 13-16 cells from a representative of 3 independent experiments is shown. White triangles indicate irradiated regions. Scale bar 5 μm. Source data are provided as a Source Data file. Boxplot limits correspond to the 25th and 75th percentiles and the center line in the box indicates the median value. The whiskers extend 1.5 times the interquartile range. All P-values were calculated using the unpaired Student's *t* test, assuming unequal variances.

Previous work indeed showed that ZNF384 belongs to one of the few C2H2-type of ZnF proteins reported to have unique DNA binding affinity, particularly for homopolymeric dA·dT DNA consensus elements enriched in the genome [30]. We performed biolayer interferometry (BLI) experiments and *in vitro* DNA pulldown experiments using purified His-MBP-tagged ZNF384 and confirmed that His-MBP-ZNF384, in contrast to His-MBP alone, has a high affinity to bind T-rich single-strand (ss)DNA (Supplementary Fig. 6A-B), as opposed to A-rich ssDNA (Fig. 3A). Extending this finding, we also found ZNF384 to bind double-stranded (ds)DNA with either a 3'- or 5' overhang in both BLI and *in vitro* DNA pulldown assay, albeit with a seemingly lower affinity when compared to ssDNA (Fig. 3A-B). Importantly, ZNF384 bound poorly, if at all, to dsDNA, and did not show RNA binding (Supplementary Fig. 6C). Moreover, it showed reduced colocalization with ethynyl uridine labeled RNA compared to Hoechst labeled DNA as quantified by the Pearson correlation coefficient (Supplementary Fig. 6D). Collectively, these results suggest that ZNF384 recruitment to DNA damage sites is dependent on PARP-induced chromatin relaxation and its affinity to bind damaged DNA *in vivo* and *in vitro*.

ZNF384 is recruited to sites of DNA damage via its C2H2 DNA binding motif

The fact that ZNF384 is recruited to DNA damage and binds DNA, encouraged us to investigate whether the C2H2, N-terminal or C-terminal domain of ZNF384 is implicated in this process (Fig. 3C). To this end, we purified His-MBP-tagged versions of these ZNF384 domains (Supplementary Fig. 6A), and assessed their ability to bind dsDNA with a 3'overhang, a common substrate found at DNA breaks, *in vitro* by BLI and DNA-pulldown assay. As expected, full-length ZNF384 was able to bind this DNA substrate, whereas its N-terminus C-terminus revealed very poor to no binding (Supplementary Fig. 7A-B). In contrast, the C2H2 domain showed a stronger binding to this substrate, which was even comparable to that of full-length ZNF384 in the BLI assay (Supplementary Fig. 7C). Finally, we also observed that the C2H2 domain, similar to full-length ZNF384, has affinity for ssDNA, as well as for dsDNA with a 3'- or 5'-overhang,

albeit that the affinity for the latter two substrates was seemingly lower when compared to that for ssDNA (Supplementary Fig. 7D). Together, these findings show that ZNF384 binds different DNA substrates through its C2H2 domain.

Given the different DNA binding affinities of these ZNF384 domains, we next assessed their relevance for ZNF384 recruitment to DNA damage sites. To this end, we studied the recruitment of GPF-tagged versions of the C2H2, N-terminal or C-terminal domains to UV-A laser micro-irradiation inflicted DNA damage (Fig. 3D) in *ZNF384* KO U2OS. This was to avoid the possible dimerization of any of the GFP-tagged domains with endogenous ZNF384 as observed for GFP-ZNF384 and endogenous ZNF384 (Supplementary Fig. 7E). mCherry-NBS1 was co-expressed in these cells to control for DNA damage induction (Supplementary Fig. 7F). In agreement with our in *vitro* experiments, we found that the C2H2 domain was still recruited, while the N-terminus and C-terminus recruitment was completely abolished (Fig. 3D). Consistently, we also found the colocalization between ZNF384 and DNA to be dependent on its C2H2 motif (Supplementary Fig. 6C).

Next, we sought to test the DNA binding affinity of ZNF384 deletion mutants *in vivo*. We generated a ΔC2H2 mutant lacking the 6 C2H2 motifs, (ΔC2H2), as well as mutants lacking the N-terminus (ΔN-terminus) or C-terminus (ΔC-terminus) (Supplementary Fig. 7G), and estimated the ability of full-length ZNF384 and ΔC2H2 to bind DNA using FRAP. In this assay, strong DNA binding affinity corresponds to a slow FRAP recovery and *vice versa*. While full-length ZNF384 showed a slow recovery after photobleaching of the damaged area, we observed a fast recovery of the ΔC2H2 ZNF384. In contrast, ZNF384 ΔN-terminus- and ΔC-terminus showed a slower FRAP recovery compared to full-length ZNF384, suggesting that both domains slightly destabilize ZNF384's interaction with DNA (Fig. 3E and Supplementary Fig. 7H). We then assessed the recruitment of these GFP-tagged deletion mutants to DNA damage 405 laser micro-irradiation in Hoechst-sensitized *ZNF384* KO U2OS. Core histone H2B fused to photoactivatable PATagRFP (PTR) was co-expressed to define the damaged area. We found that ΔC2H2 dramatically impaired ZNF384 recruitment, while ΔN-terminus and ΔC-terminus had no major effect on recruitment (Fig. 3F). Collectively, these findings suggests that the C2H2 motif is important for DNA binding and recruitment of ZNF384 at DNA damage sites.
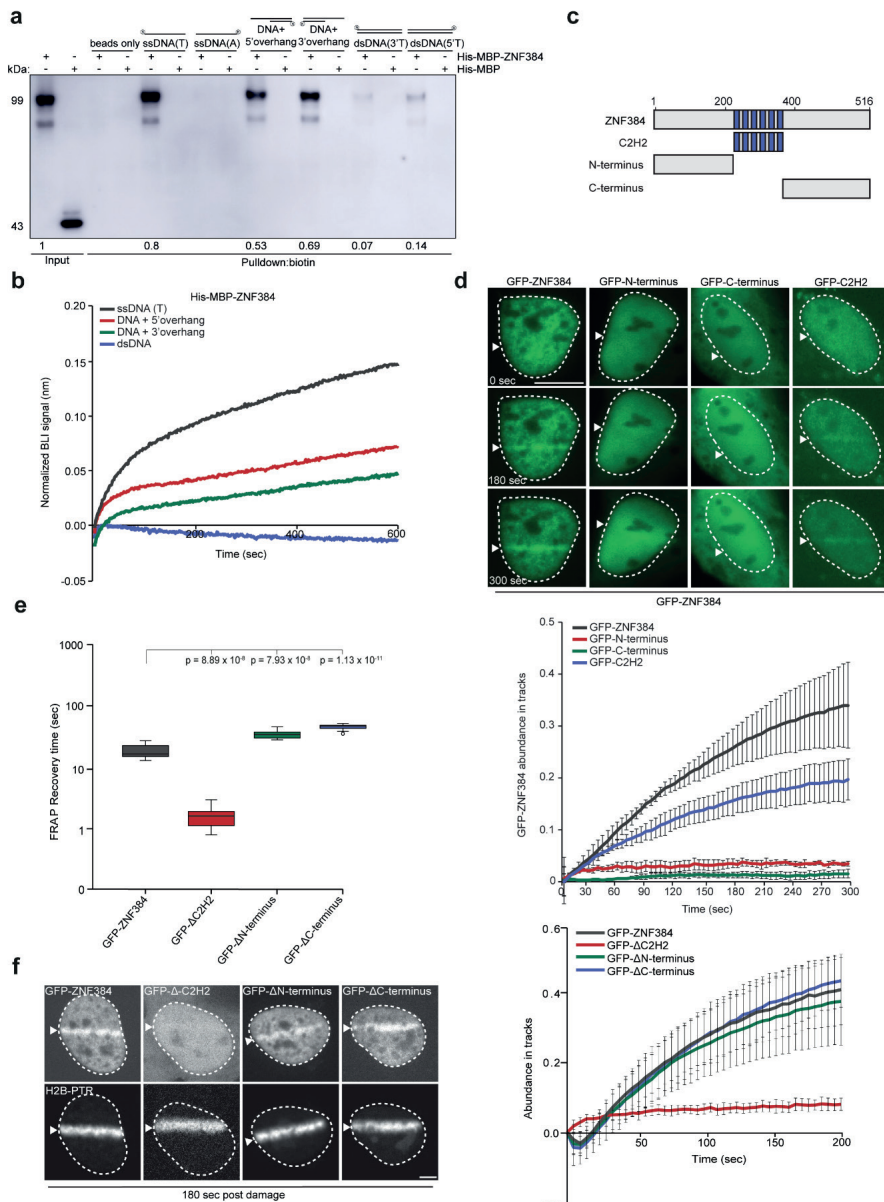
Figure. 3. ZNF384 is recruited to sites of DNA damage via its C2H2 DNA binding motif. (A) DNA pulldowns of the indicated biotinylated DNA substrates in the presence of His-MBP-C2H2 or His-MBP (control). Blots were probed for MBP. His-MBP-ZNF384 pulldown signals were normalized to the Input signal, which was set to 1. The mean from 4 independent experiments is indicated below the blot. His-MBP was not detectable in pulldowns. A representative experiment is

shown. (B) DNA binding of His-MBP-ZNF384 to the indicated biotinylated DNA substrated as measured by BLI. Quantified data is plotted on a time scale and normalized to His-MBP (control). (C) Schematic representation of full-length ZNF384 protein and its domains (C2H2, N-terminus and C-terminus). (D) Live cell imaging of the recruitment of the indicated GFP-ZNF384 proteins to DNA damage tracks generated by 365 nm UV-A tracks in BrdU-sensitized *ZNF384* KO U2OS Flp-In/T-Rex cells. mCherry-NBS1, which was co-expressed with GFP-ZNF384 proteins, served as a DNA damage marker. Representative images are shown. White triangles indicate irradiated regions. Scale bars: 10 μm (upper panel). Quantification of the data is shown as mean ±SEM from 30-40 cells (lower panel). (E) Quantification of FRAP measurements to assess the local dynamics of the indicated GFP-ZNF384 constructs. 12 cells per condition were analyzed. Boxplot limits correspond to the 25th and 75th percentiles and the center line in the box indicates the median value. The whiskers extend 1.5 times the interquartile range. P-values were calculated using an unpaired Student's t test, assuming unequal variances. (F) Live cell imaging of the recruitment of the indicated GFP-ZNF384 proteins to DNA damage tracks generated by 405 nm laser micro-irradiation in Hoechst sensitized ZNF384 KO U2OS Flp-In/T-Rex cells. White triangles and photoactivatable H2B-PTR, which was co-expressed with the indicated GFP-ZNF384 proteins, indicate irradiated regions. Representative images are shown Scale bars: 4 μm (left panel). Data are presented as mean values ± SEM. P-values were calculated using an unpaired Student's t test, assuming unequal variances (right panel). Source data are provided as a Source Data file.

ZNF384 modulates KU70/KU80 dynamics at DNA damage sites

The repair of DSBs by cNHEJ depends on the binding of the KU70/KU80 heterodimer to broken DNA ends, followed by the recruitment of the DNA-PKcs kinase (Mari, Florea et al. 2006). This in turn leads to the recruitment of the XRCC4-LIG4 complex, which ultimately seals the broken ends (Frit, Ropars et al. 2019). Given the interaction between ZNF384 and KU70/KU80, we sought to address if ZNF384 is involved in the loading of KU70/KU80 at DNA. To this end, we first monitored the levels of KU70 in chromatin-enriched extracts from cells depleted of ZNF384 and KU80 (Supplementary Fig. 8A). While *KU80* knockdown reduced KU70 levels on chromatin, depletion of ZNF384 had no impact (Supplementary Fig. 8A). Next, we monitored the impact of ZNF384 on KU70 dynamics at DNA by FRAP. Within two seconds after photobleaching we observed a small increase in fluorescence recovery of GFP-KU70 in *ZNF384*-depleted cells compared to that in control cells (Supplementary Fig. 8B), suggesting that ZNF384 has a modest effect on KU70's DNA binding. To further investigate this finding, we assessed the DNA binding affinities of both ZNF384 and KU70 by comparing their relative residence times as measured by Fluorescence Correlation Spectroscopy (FCS) (Supplementary Fig. 8C). Interestingly, we observed that ZNF384 has a higher residence time compared to KU70 (Supplementary Fig. 8D), suggesting that ZNF384 is more tightly bound to DNA, thereby impacting the association of KU70/KU80 with DNA.

Based on these findings, we next examined whether ZNF384 has a potential stimulatory role on KU70/KU80's binding to dsDNA with a 3'overhang, a common DSB repair substrate,

in vitro. To this end, we performed DNA-pulldown assays using purified ZNF384 and KU70/KU80. Importantly, while KU70/KU80 bound dsDNA with a 3' overhang in the absence of ZNF384 [31] their binding was enhanced in the presence of increasing amounts of ZNF384 (Fig. 4A and Supplementary Fig. 8E). To examine whether ZNF384 also affects the loading of KU70/KU80 at damaged DNA *in vivo*, we depleted ZNF384 in cells expressing endogenously GFP-tagged KU70 (Britton, Coates et al. 2013), and subjected these cells to multiphoton laser micro-irradiation. Importantly, ZNF384 depletion reduced GFP-KU70 accumulation at sites of DNA damage as compared to that in control cells (Fig. 4B-C). Given that ZNF384 is recruited to DNA breaks in a PAR-dependent manner (Fig. 2A) and regulates the loading of KU70/KU80, we asked whether KU70/KU80 is also recruited to sites of DNA damage in a PAR-dependent manner. Indeed, we found PARPi treatment to impair the recruitment of KU70 to DNA break sites (Supplementary Fig. 8F), while the accumulation of the DNA damage marker NBS1 remained unaffected (Supplementary Fig. 8G-H). To better understand the impact of ZNF384 on the dynamics of KU70, we assessed its turnover at DNA lesions by FRAP (Fig. 4D-E). Within 10 seconds after photobleaching, we observed a faster fluorescence recovery of GFP-KU70 in ZNF384-depleted cells compared to that in control cells (Fig. 4E), suggesting that ZNF384 contributes to the association of KU70 with DNA lesions. To be able to extract quantitative characteristics from the FRAP data, we first tried to fit the recovery curves with single-population models (Sprague, Pego et al. 2004). However, none of them could accurately fit the experimental curves (Supplementary Fig. 9A), indicating that KU displays a more complex behavior. We then examined KU dynamics by FCS as this method is able to assess fast protein turnover more accurately. The fit of the correlation curves showed that KU follows a two-population dynamics at DNA damage sites (Supplementary Fig. 9B). We infer that the fast population refers to KU molecules diffusing thought the nucleus and displaying only very transient interactions with chromatin, while the slow population corresponds to KU molecules that bind DNA lesions more stably. To characterize specifically the behavior of this slow population, we restricted the fitting of the FRAP curves to the timepoints after 3 seconds post photobleaching, which could be well adjusted by a reaction-limited model (Supplementary Fig. 9C). Using this model, we were able to estimate binding (k'_{on}) and unbinding (k_{off}) rates of KU at sites of DNA damage by FRAP. The k'_{on} is a pseudo first-order association rate corresponding to the product of the actual binding rate k_{on} and the local density of DNA damage sites dictated by the irradiation conditions, which was similar between the different conditions. To confirm the applicability of this model, we also verified that the k_{off} parameter estimated with this model was independent of the size of the

bleached area (Supplementary Fig. 9D). Based on these fits, we observed a reduction in the k'_{on} in ZNF384-depleted cells, suggesting a decreased KU70 binding rate (Fig. 4F). In contrast, no significant impact of ZNF384 depletion on the estimated k_{off} was observed (Fig. 4G). Together these results imply that ZNF384 facilitates the recruitment and subsequent binding, rather than the retention of KU70/KU80, at sites of DNA damage.
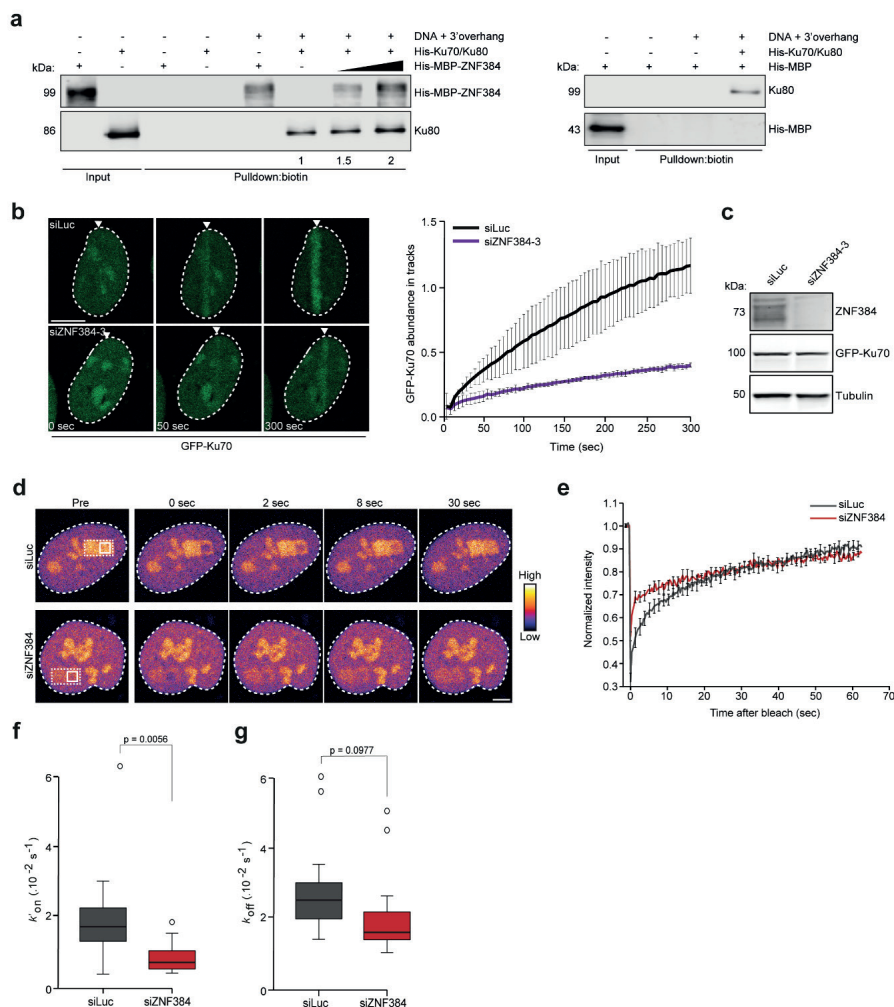

Figure 4. ZNF384 modulates KU70/KU80 dynamics at DNA damage sites. (A) DNA pulldowns of biotinylated DNA with a 3'overhang in the presence of His-KU70/KU80, His-MBP-ZNF384 or His-MBP alone or His-KU70/KU80 in combination with His-MBP-ZNF384 or His-MBP. Blots were probed for MBP and KU80. KU80 pulldown signals were normalized to that in the pulldown lacking His-MBP-ZNF384, which was set to 1. The mean from 4 independent experiments is indicated below the blot. His-MBP was not detectable in pulldowns. (B) GFP-KU70

recruitment to 800 nm multiphoton tracks in RPE1-hTERT cells transfected with the indicated siRNAs (left panel). White triangles indicate irradiated regions. Quantification of the data is presented as the mean ±SD from >60 cells acquired in 2 independent experiments. Scale bar: 5 μm (C) Western blot analysis of ZNF384 expression in cells from B. Tubulin is a loading control. Data shown represent 2 independent experiments. (D) Representative images of RPE1-hTERT cells transfected with the indicated siRNAs in which FRAP measurements were performed to assess the local turnover of GFP-KU70 at the sites of DNA damage. DNA damage was induced in the region indicated with a dashed line. Subsequent FRAP was induced in a subarea within the DNA damage region, as indicated with an unbroken line. Images are pseudocolored according to the look-up table displayed on the right. Scale bar: 4 μm. (E) Normalized FRAP curves from D. (F) Association (k'_{on}) rates of GFP-KU70 measured by FRAP after fitting of the curves from E. (G) Dissociation (k_{off}) rates of GFP-KU70 after fitting of the curves from E. Data from E-G was collected from 15 cells per condition. Boxplot limits correspond to the 25th and 75th percentiles and the center line in the box indicates the median value from a representative of 2 independent replicates. Statistical significance was calculated using the unpaired Student's *t* test, assuming unequal variances. Source data are provided as a Source Data file.

The C2H2 and N-terminal domains of ZNF384 are critical for loading KU70/KU80 at DSBs

Having shown that ZNF384 promotes the loading of KU70/KU80 at DNA breaks and considering that ZNF384 and KU70/KU80 physically interact, we next sought to investigate whether the interaction between these proteins is important for the efficient recruitment of KU70/KU80 to DNA breaks. We therefore set out to map the region in ZNF384 that is required for the interaction with KU70/KU80, making use of our set of ZNF384 deletion mutants (Supplementary Fig. 7G). Using GFP-pulldown assays we found that GFP-tagged ZNF384, ΔC2H2 ZNF384 and ΔC-terminus ZNF384 associated with endogenous KU70 and KU80 with equal efficiency (Fig. 5A), suggesting that the C2H2 motifs and C-terminus are dispensable for the interaction. On the contrary, ΔN-terminus ZNF384 showed an almost complete loss of KU70/KU80 binding, indicating that the N-terminus of ZNF384 mediates its interaction with KU70 and KU80 (Fig. 5A). In agreement with our *in vivo* pulldown results, we found that recombinant ΔN-terminus ZNF384 bound less efficiently to KU70/KU80 when compared to recombinant ZNF384 (Figure 5B).

We next asked if this domain has any functional relevance for KU70/KU80 recruitment. To this end, we used the Flp-In/T-rex system to establish U2OS cells stably expressing inducible siRNA-resistant GFP-tagged versions of either ZNF384, ΔC2H2 ZNF384 or ΔN-terminus ZNF384. U2OS cells stably expressing GFP-NLS served as a control (Fig. 5C). While expression of GFP-ZNF384 fully restored KU80 accumulation at UV-A laser micro-irradiation inflicted damage, expression of ΔN-terminus ZNF384 only partially rescued this KU80 defect (Fig. 5D). In contrast, expression of ΔC2H2

ZNF384 completely failed to rescue the KU80 accumulation defect (Fig. 5D). DNA damage induction was similar in all conditions based on equal PAR levels in laser tracks (Fig. 5E). These results suggest that the binding of ZNF384 at DNA breaks via its C2H2 motif, as well as the interaction between its N-terminus and KU70/KU80 contribute to efficient KU80 recruitment.

Finally, we asked if ZNF384 is responsible for KU70/KU80 complex formation. To this end, we performed GFP-pulldowns on cells expressing endogenously GFP-tagged KU70 (Britton, Coates et al. 2013) that were depleted for ZNF384 and left untreated or exposed to ionizing radiation (IR). While we observed that GFP-KU70 and endogenous KU80 interact, as expected, loss of ZNF384 did not impact this interaction, neither in untreated nor in IR-exposed cells (Supplementary Fig. 10A). This suggests that ZNF384 is not involved in KU70/KU80 complex formation, but rather the loading of this complex directly at sites of DNA damage.
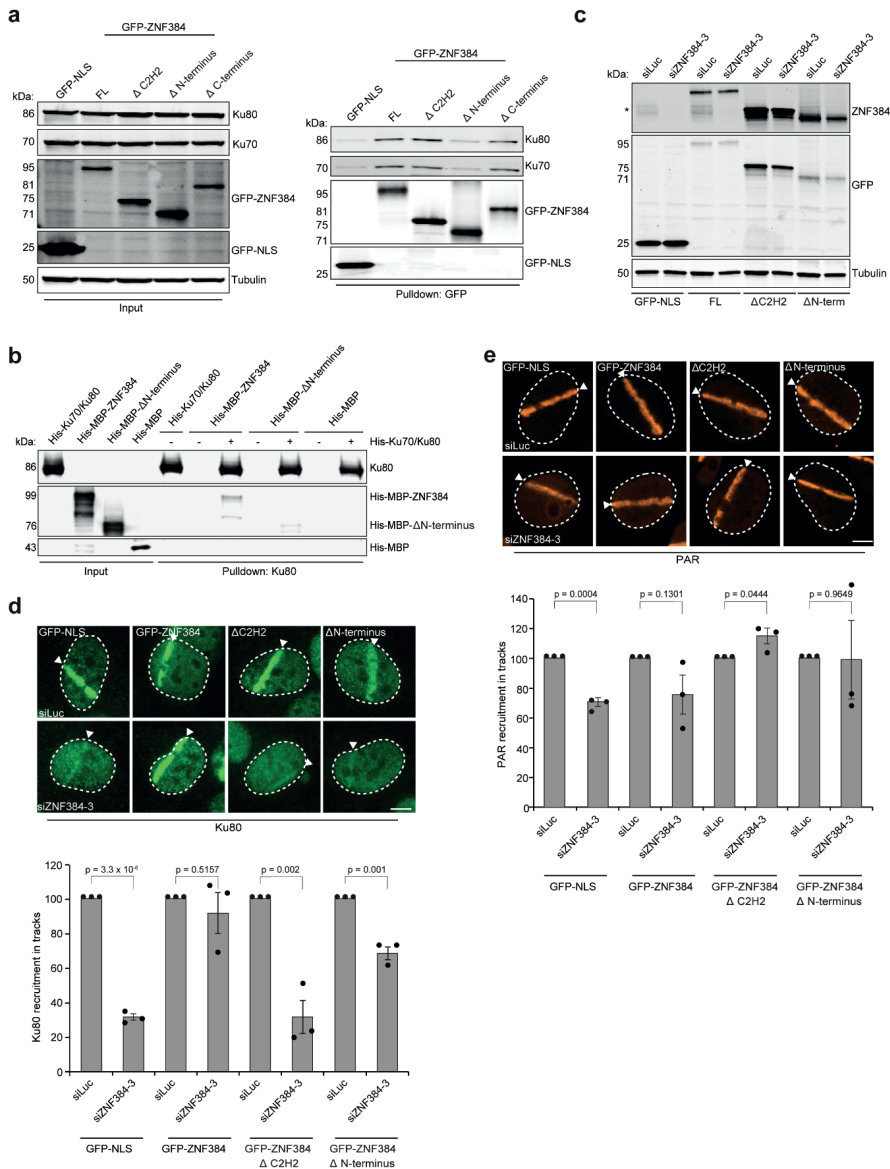
Figure 5. The C2H2 motifs and N-terminus of ZNF384 are required K70/Ku80 loading at DSBs. (A) Pulldowns of the indicated GFP fusion proteins in U2OS Flp-In/T-Rex cells. Blots were probed for Ku70, Ku80, Tubulin and GFP. Data shown represent 3 independent experiments. (B) In *vitro* KU80 pulldown in the presence or absence of His-KU70/KU80 and His-MBP, His-MBP-ZNF384 or His-MBP-ΔN-terminus. Control IP contained beads only. Blots were probed for KU80 and MBP. Data shown represent 2 independent experiments. (C) Western blot analysis of the expression of endogenous ZNF384 and ectopic GFP-ZNF384 full-length and deletion mutants. Tubulin is a loading control. The asterisk (*) indicates endogenous ZNF384. Data shown

represent 3 independent experiments. (D) Accumulation of endogenous KU80 at 365 nm UV-A tracks in BrdU-sensitized U2OS Flp-In/T-Rex cells expressing siRNA-resistant doxycycline (dox)-inducible GFP-ZNF384, GFP-ZNF384 ΔC2H2 and GFP-ZNF384 ΔN-terminus following transfection with the indicated siRNAs. Cells were subjected to laser micro-irradiation and 10 minutes later fixed and immunostained. White triangles indicate irradiated regions (upper panel). Quantification of endogenous Ku80 levels in laser tracks is presented as ±SEM of >150 cells acquired in 3 independent experiments (lower panel). (E) As in D, except for PAR (upper panel). Quantification of endogenous PAR levels in laser tracks from is presented as the mean ±SEM from 3 independent experiments. Data were normalized to siLuc, which was set to 100% (lower panel). Statistical significance was calculated using the unpaired Student's *t* test, assuming unequal variances. Scale bar: 5 μm. Source data are provided as a Source Data file.

ZNF384 promotes KU-dependent loading of APLF and XRCC4/LIG4 at DSBs

Next, we asked if ZNF384 affects the accumulation of factors that act downstream of Ku70/Ku80. We first measured the recruitment of APLF, which physically interacts with Ku80 at DSBs via its conserved Ku-binding motif (KBM) [32]. To this end, YFP-tagged APLF and mCherry-NBS1 were co-expressed in ZNF384-, KU80- and ZNF384/KU80-depleted U2OS cells, and examined for their localization at sites of DNA damage. ZNF384 and KU80 depletion impaired APLF recruitment to sites of DNA damage, while the DNA damage marker NBS1 remained unaffected (Fig. 6A-B and [32]. KU80 depletion also reduced nuclear retention of APLF [33], which we validated in KU80 knockdown cells (Supplementary Fig. 10B-C), as well as in *Ku80−/−* mouse embryonic stem cells (Supplementary Fig. 10D). Interestingly, double knockdown of ZNF384 and KU80 did not further impair APLF accumulation, suggesting that these proteins function epistatically to recruit APLF (Fig. 6A-B).

The fact that ZNF384 promotes the consecutive accumulation of KU70/KU80 and APLF made us wonder if ZNF384 could be involved in Ku-APLF complex formation. To assess this, we performed GFP-pulldowns after transient expression of YFP-APLF in wild-type and ZNF384-depleted U2OS cells. Endogenous KU80 was co-precipitated in both conditions (Supplementary Fig. 10E). Moreover, ZNF384 and Ku70 showed comparable recruitment kinetics (Fig. S10F-G), while APLF was recruited earlier (Supplementary Fig. 10H). This suggests that ZNF384 does not impact Ku-APLF complex stability, but rather promotes the loading of these factors at sites of DNA damage.

The Ku-APLF interaction is critical for recruitment of the XRCC4/LIG4 complex to DNA beaks [3, 32, 34, 35]. Given that ZNF384 promotes both Ku-APLF loading at sites of DNA damage, we next examined if ZNF384 also affects XRCC4 recruitment. To this

end, we measured the levels of endogenous XRCC4 at *bona fide* DSBs induced by tethering of a Lactose repressor (LacR)-tagged FokI nuclease at a stably integrated Lactose operator (LacO) array in U2OS cells [36] (Fig. 6C). While we detected XRCC4 accumulation at FokI-induced DSBs in control cells, as expected, its levels were dramatically reduced in ZNF384-depleted cells (Fig. 6D-E). his indicates that ZNF384 acts at *bona fide* DSBs to facilitate the Ku-APLF-dependent accumulation of XRCC4/LIG4 complexes.

3
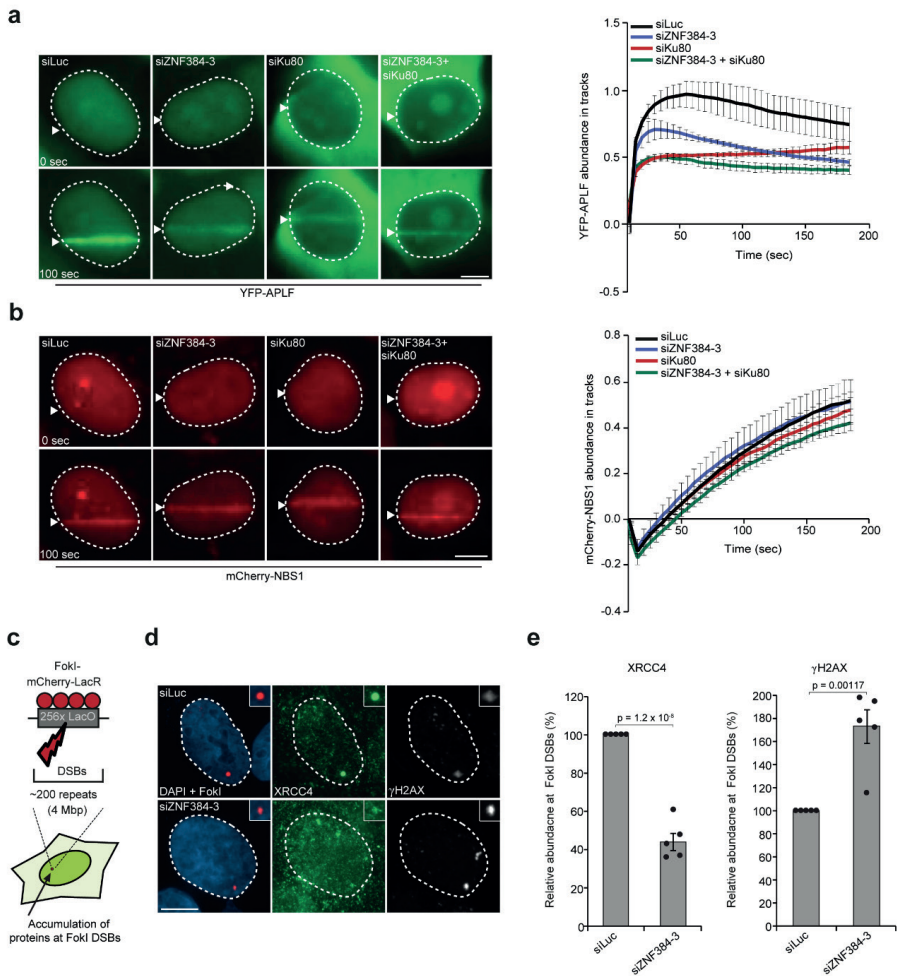
Fig. 6. ZNF384 promotes KU-dependent loading of APLF and XRCC4/LIG4 at DSBs. (A) Live cell imaging of the recruitment of YFP-APLF to 365 nm UV-A tracks in BrdU-sensitized U2OS cells transfected with the indicated siRNAs mCherry-NBS1, which was co-expressed with YFP-APLF, served as a DNA damage marker. Representative images are shown. White triangles indicate irradiated regions (left panel). Quantification of data is presented as the mean values ±SEM from 60

cells acquired in 3 independent experiments (right panel). (B) As in A, except for mCherry-NBS1 (left panel). Quantification of data is presented as the mean values ±SEM from 60 cells acquired in 3 independent experiments (right panel) (C) Schematic of the system in U2OS 2-6-3 cells used to locally induce multiple DSBs upon tethering of the FokI endonuclease. (D) Accumulation of XRCC4 (green) to γH2AX-marked (white) DSBs induced by FokI-mCherry-LacR at a LacO array (red) in cells transfected with the indicated siRNAs. (E) Quantification of XRCC4 and γH2AX in cells from D is presented as the mean ±SEM of >200 cells acquired in 5 independent experiments. Data were normalized to siLuc control which was set to 100%. Statistical significance was calculated using the two-tailed unpaired Student's *t* test, assuming unequal variances. Scale bar: 5 μm. Source data are provided as a Source Data file.

ZNF384 promotes recruitment of cNHEJ proteins independently of DNA-PKcs

Given that ZNF384 promotes the accumulation of XRCC4/LIG4 in a Ku70/Ku80-dependent manner, we wondered how this is linked to DNA-PKcs, which is recruited and activated by DSB-bound Ku to promote the loading of XRCC4/LIG4 [37, 38]. Western-blot analysis detected an induction of phosphorylated (p)-DNA-PKcs (S2056) after IR (Supplementary Fig. 11A). Surprisingly, neither KU80 nor ZNF384 depletion affected p-DNA-PKcs (S2056) activation in U2OS cells (Supplementary Fig. 11A). Moreover, normal levels of p-DNA-PKcs (S2056) activation were observed in *Ku80*[-/-] mouse embryonic stem cells, ruling out that the lack of phenotype in KU80-depleted cells was due to incomplete knockdown (Supplementary Fig. 11B and [39]). To corroborate these findings, we examined the interplay between ZNF384, KU and DNA-PKcs during the loading of XRCC4 at sites of DNA damage (Supplementary Fig. 11C). We found that ZNF384, KU80 and DNA-PKcs depletion impaired XRCC4 accumulation (Supplementary Fig. 11D), which is in line with previous results [37, 40]. Double knockdown of ZNF384 and KU80 did not result in an additive effect on XRCC4 accumulation when compared to that in ZNF384- or KU80-depleted cells (Supplementary Fig. 11D-F). This suggest that these proteins act epistatically to recruit XRCC4, which is in line with their epistatic role in recruiting APLF to DNA breaks (Fig. 6A-B). In contrast, double knockdown of ZNF384 and DNA-PKcs resulted in a larger effect on XRCC4 accumulation when compared to that of ZNF384 or DNA-PKcs depletion alone (Supplementary Fig. 11C-E), suggesting redundant functions for these proteins in XRCC4 recruitment.

Given that ZNF384 and Ku were recruited to sites of DNA damage via PAR-driven processes (Fig. 2A and Supplementary Fig. 8F), and that ZNF384 and DNA-PKcs function redundantly, we next asked whether DNA-PKcs is also recruited to DNA breaks in a PAR dependent manner. Indeed, PARPi treatment impaired the recruitment of

DNA-PKcs (Supplementary Fig. 11G), while the accumulation of the DNA damage marker NBS1 remained unaffected (Supplementary Fig. 11H-I). To further understand the interplay between ZNF384, KU and DNA-PKcs with PARP1 epistasis analysis were performed. We found that ZNF384, DNA-PKcs and PARP1 depletion impaired XRCC4 accumulation, which is in line with our previous results (Supplementary Fig. 12A-B and [9]. However, double knockdown of either ZNF384 or DNA-PKcs with PARP1 did not result in an additive effect on XRCC4 accumulation (Supplementary Fig. 12C-D). Together these results suggest that PARP activity drives two parallel pathways for DSB repair by NHEJ, one of which relies on the ZNF384-mediated ligation of broken ends via KU-APLF-XRCC4, the other on DNA-PKcs-XRCC4.

ZNF384 promotes DSB repair via cNHEJ

The interaction of ZNF384 with cNHEJ factors (Fig. 1D), its PARP/PAR-dependent recruitment (Fig. 2A), as well as its ability to load Ku (Fig. 4B), APLF (Fig. 6A) and XRCC4/LIG4 (Fig. 6E) at DSBs, encouraged us to assess whether ZNF384 supports DSB repair via NHEJ. To this end, we first used the well-established EJ5-GFP reporter assay, which relies on the restoration of GFP expression following repair of I-*Sce*I endonuclease-induced DSBs that flank a puromycin gene that separates a GFP gene from a CMV promoter (Fig. 7A). Flow cytometric analysis of GFP fluorescence revealed that NHEJ was reduced following ZNF384 knockdown, which was comparable to the effect observed upon XRCC4 knockdown (Fig. 7A). Cell cycle profiles remained unaffected in these cells, ruling out effects of cell cycle mis-regulation (Supplementary Fig. 13A-B). Importantly, knockdown of ZNF384 did not affect the steady-state levels of several factors involved in NHEJ (Supplementary Fig. 13C), albeit that the expression of XRCC4 was slightly reduced (Supplementary Fig. 13C). However, a semi-quantitative analysis revealed that this effect was not consistent, neither in U2OS (Supplementary Fig. 13D) nor in HeLa Flp-In cells (Supplementary Fig. 13E), suggesting that indirect effects due to transcriptional misregulation are unlikely. Furthermore, ZNF384 depletion did not have a major impact on the steady state levels of the checkpoint kinases ATM and CHK1 (Supplementary Fig. 13F-H), nor affected the IR-induced phosphorylation of ATM (at S1981) and CHK1(at S345) (Supplementary Fig. 13I-J), the latter of which was used as a readout for ATR activation [41]. This suggest that ZNF384 does not contribute to DSB repair by regulating ATM or ATR activation.

The EJ5-GFP reporter provides a readout for total NHEJ activity, including both cNHEJ and aNHEJ [42]. To address if ZNF384 specifically plays a role in KU70/80-, APLF- and XRCC4/LIG4-dependent cNHEJ, we measured random plasmid integration into the genome via cNHEJ (Fig. 7B and [9, 43]. Indeed, ZNF384 knockdown, similar to that of KU80 depletion, impaired random plasmid integration, indicating an important role for ZNF384 in cNHEJ (Fig. 7B).

Next, we asked whether ZNF384 plays a unique role in cNHEJ or also affects HR. Interestingly, loss of ZNF384 only very moderately impaired HR in the well-established DR-GFP reporter (Supplementary Fig. 14A), and rendered cells only mildly sensitive to treatment with PARPi (Supplementary Fig. 14B-C), in contrast to HR-deficient cells depleted of BRCA1 [44]. Given that ZNF384 is critical for KU70/KU80 loading at DSBs and loss of KU70/KU80 has been linked to increased DNA end resection and HR levels [45, 46], we assessed whether ZNF384 affects DNA end resection. Cells depleted of ZNF384 did not show a significant difference in RPA foci number and foci intensity (Supplementary Fig. 14D-G), which is consistent with the fact that ZNF384 loss did not impact HR (Supplementary Fig. 14A). KU80-depleted cells also did not show changes in DNA end resection (Supplementary Fig. 14D-G), which is in agreement with another report showing that end-resection remained unaffected in *Ku70/Ku80* double KO MEF cells [47]. Corroborating these findings, we observed that the accumulation of the core HR protein RAD51 into DSB-containing foci induced by IR was not affected in ZNF384-depleted S-phase cells (Supplementary Fig. 14H). Together these results suggest that ZNF384 does not play a major role in DSB repair via HR, but rather promotes cNHEJ.

When cNHEJ is compromised, DSB repair mostly occurs via aNHEJ, causing a loss of accurate end-joining and a switch to error-prone repair due to deletion formation and microhomology usage [48]. To test whether impaired cNHEJ in ZNF384-depleted cells impacts the mutational signature at repair junctions, we used a previously published NHEJ reporter in GC92 cells [49]. This reporter consists of a CD4 gene separated from a CMV promoter by an H2Kd-CD8 cassette that is flanked by I-*Sce*I cleavage sites. DSB induction by I-*Sce*I expression leads to re-ligation of the CMV promoter to the CD4 gene after which repair junctions can be amplified and Sanger-sequenced (Supplementary Fig. 15A). Indeed, depletion of the cNHEJ factor KU80 resulted in an increase in the formation of larger deletions and usage of larger stretches of microhomology (Supplementary Fig. 15B-C). Moreover, although ZNF384 depletion only caused a modest effect on the total deletion frequency, among these events the proportion of larger deletions and use of

microhomology during repair was increased and resembled that observed after KU80 depletion (Supplementary Fig. 15B-C), corroborating a role for ZNF384 in cNHEJ.

In further support of these findings, we found that ZNF384 loss impaired clonogenic survival of U2OS cells and VH10-SV40-immortalized fibroblasts following exposure to IR-induced DSBs similar to that observed after XRCC4 depletion (Fig. 7C-D and Supplementary Fig. 15D-E). Interestingly, double depletion of ZNF384 and XRCC4 did not result in an increased sensitivity to IR (Fig. 7C-D), again indicating that these proteins function epistatically to promote cNHEJ-dependent repair DSBs. Finally, we employed the Flp-In/T-Rex system to establish HeLa cells stably expressing inducible siRNA-resistant GFP-tagged ZNF384 or GFP-NLS (Fig. 7E). We found that expression of siRNA-resistant GFP-ZNF384, but not that of GFP-NLS, almost fully rescued the IR-sensitivity observed after ZNF384 loss, the latter of which being comparable to that observed after XRCC4 loss (Fig. 7F). This indicates that the IR sensitivity and underlying cNHEJ defect are not caused by off-targets of the siRNAs against ZNF384. The increase in IR sensitivity likely resulted from an accumulation of unresolved DSBs, as indicated by the increase in γH2AX foci following IR-exposure of ZNF384-depleted G1 cells, which resembled the phenotype of KU80-depleted cells (Supplementary Fig. 15F-H and [50]. Interestingly, KU80 has been implicated in telomere length maintenance (Jaco, Muñoz et al. 2004). To study whether ZNF384 is similarly involved in this process, we performed Fluorescence *in situ* hybridization (FISH) using a PNA probe to label all telomeres and use their intensity as a proxy for telomere length (Fig. 7G) [51]. Strikingly, knockdown of ZNF384, similar to that of KU80 [52], significantly reduced telomere length in metaphase spreads when compared to that in control cells (Fig. 7H-I). Collectively, our results demonstrate that ZNF384 promotes efficient DSB repair via cNHEJ and is involved in telomere length maintenance (Fig. 7J).
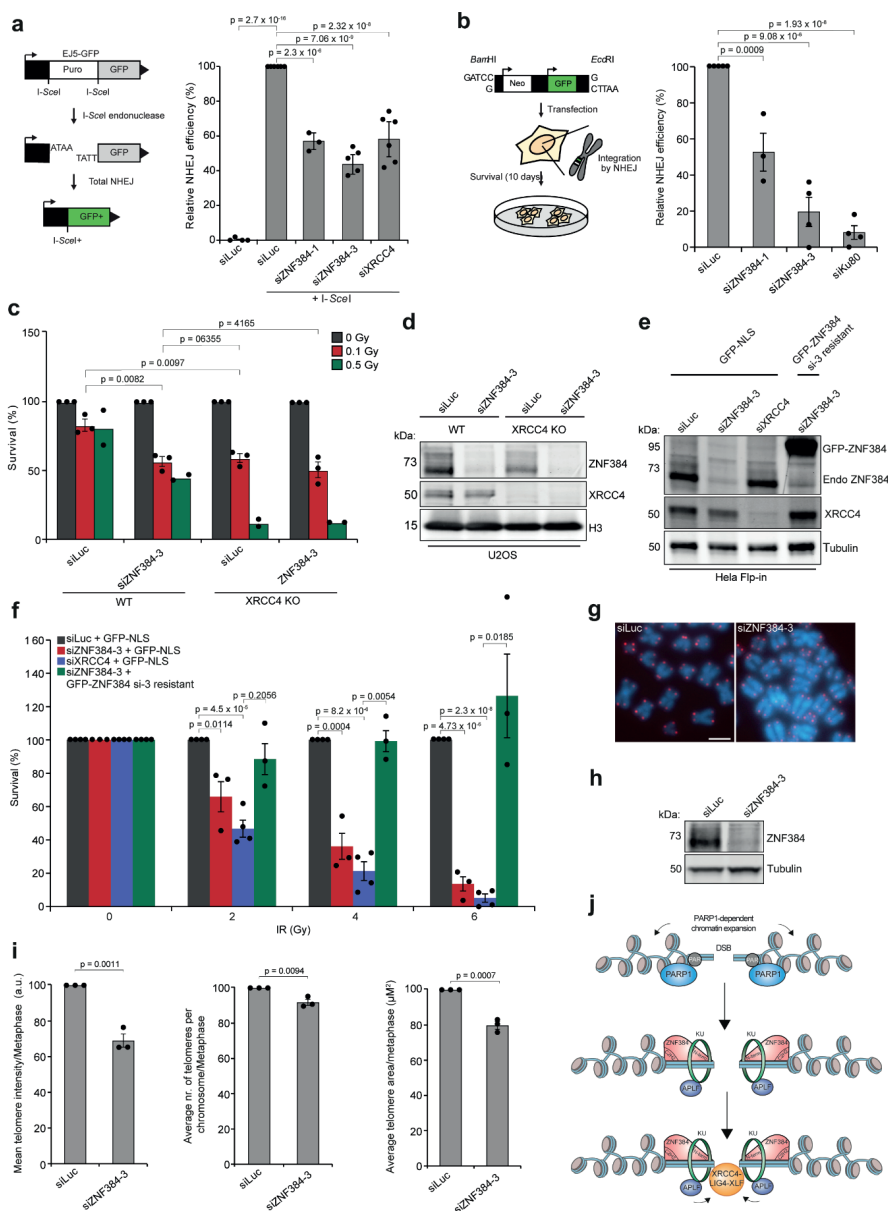
Fig. 7. ZNF384 promotes DSB repair via cNHEJ. (A) Schematic of the EJ5-GFP reporter for NHEJ (left panel). Quantification of the EJ5-GFP-positive U2OS cells transfected with the indicated siRNAs and I-*Sce*I expression vector. I-*Sce*I transfection was corrected by co-transection with mCherry expression vector. The mean ±SEM of 3-5 independent experiments is shown (right panel). Data was normalized to the siLuc control which was set to 100%. (B) Schematic of the random plasmid integration assay (left panel). Quantification of plasmid integration efficiencies in

U2OS cells transfected with the indicated siRNAs (right panel). The mean ±SEM of 3-6 independent experiments is shown. Data were normalized to siLuc control which was set to 100%. (C) Relative survival efficiency in WT and XRCC4 KO U2OS cells transfected with the indicated siRNAs and exposed to the indicated doses of IR. The mean ±SEM of 2-3 independent experiments is shown. Data were normalized to un-irradiated conditions and set to 100%. (D) Western blot analysis of the expression of endogenous ZNF384 from cells in C. Tubulin is a loading control. (E) Expression levels of endogenous ZNF384 and dox-inducible siRNA-resistant GFP-ZNF384 in Hela Flp-In/TRex cells. Tubulin is a loading control. (F) Effect of inducible expression of GFP-NLS and siRNA resistant GFP-tagged ZNF384 on the survival of stable Hela Flp-In/TRex after transfection with indicated siRNAs and IR treatment. The mean ±SEM of 3-4 independent experiments is shown. Data was normalized to siLuc control which was set to 100%. (G) Representative FISH images of metaphases from HCT116 cells transfected with the indicated siRNAs. Scale bar: 5 μm (H) Western blot analysis of the expression of endogenous ZNF384 from cells in G. Tubulin is a loading control. (I) Quantifications per metaphase from H are presented as the mean ±SEM of 75 chromosomes acquired in 3 independent experiments. Data were normalized to siLuc control, which was set to 100%. (J) Model for how ZNF384 works as an adaptor of KU to DNA during DSB repair by cNHEJ (see text for details). Source data are provided as a Source Data file. All P-values were calculated using the two-tailed unpaired Student's t test, assuming unequal variances.

DISCUSSION

In this study, we uncover an important role of the poorly characterized ZnF protein ZNF384 in DSB repair via cNHEJ (Fig. 7J). First, we demonstrate that ZNF384 is recruited to sites of DNA damage and interacts with KU70/KU80 and PARP1. Second, we show that ZNF384 recruitment requires PARP1/PAR-dependent chromatin remodeling, which promotes binding of ZNF384 to the exposed DNA via its C2H2 motifs. Third, ZNF384 stimulates the binding of KU70/KU80 at DNA breaks, on the one hand through physical interaction with this complex, and on the other hand through its affinity for DNA. This way, it promotes the assembly of a functional cNHEJ complex that includes APLF and the XRCC4/LIG4 complex. Finally, ZNF384 promotes NHEJ in EJ5-GFP reporter assays and random plasmid integration assays, and functions epistatically with both Ku and XRCC4 during this repair process. Thus, ZNF384 functions as an 'adaptor station' for the proper assembly of repair proteins at DSBs, thereby promoting efficient repair by cNHEJ (Fig. 7J).

PARP1/PAR-dependent chromatin unfolding allows ZNF384 binding to damaged DNA

Our findings reveal that ZNF384 is recruited to sites of DNA damage following chromatin unfolding driven by the activity of PARP1, but not PARP2 or PARP3. Several other cNHEJ repair proteins, including KU70/KU80 and XRCC4, are also recruited in a manner dependent on the activity of PARP1. This may involve their binding to PARP1-associated PAR chains or to the damaged DNA itself [9, 53]. Using three independent

approaches we confirmed that ZNF384 does not bind to PARP1-associated PAR chains at sites of DNA damage. Instead ZNF384 recruitment occurs in a manner dependent on the PARP1-induced relaxation of the damaged chromatin, making the DNA available for ZNF384 to bind via its C2H2 DNA binding motif. This is consistent with other reports showing that ZNF384 directly binds to DNA, particularly to homopolymeric dA·dT consensus sequences *in vitro* [30, 54]. We confirmed and extended this finding by showing that ZNF384, through its C2H2 domain, also efficiently binds dsDNA substrates containing 5' or 3' overhangs as opposed to dsDNA with blunt ends. DSBs with such protruding ends have been shown to be preferentially repaired by cNHEJ, while blunt ends are mostly subjected to polymerase theta-mediated end-joining [39]. Thus, the preferential binding of ZNF384 to dsDNA with 5' or 3' overhangs is consistent with its role in cNHEJ. However, it is important to note that the DNA substrates used in our study also contain T-rich consensus sequences. These sequences are considered the most abundant simple repetitive motifs in the human genome that are frequently expanded due to DNA replication slippage [55]. The fact that DSBs can occur throughout the genome, including at homopolymeric dA·dT repeats, raises the question whether the C2H2 DNA binding domain of ZNF384 binds to a specific DNA context or binds to lesions in any given sequence context, the latter of which would be in line with a more general role of ZNF384 in cNHEJ. Elucidating the nature of its DNA sequence-specific binding mode will be key in further unraveling how ZNF384 acts at DSBs to promote cNHEJ

ZNF384 serves as a 'KU-adapter' at sites of DNA damage

ZNF384 promotes the accumulation of KU70/KU80, but not their retention at sites of damage. This raises the question as to how the recruitment of these proteins is regulated at the level of DNA binding. FCS analysis revealed that ZNF384 had a significantly higher residence time compared to KU70, suggesting that ZNF384 binds stronger to DNA as compared to KU70. Moreover, ZNF384 stimulates the binding of KU70/KU80 on DNA in vitro, suggesting that ZNF384 is the dominant binding force within the ZNF384-KU complex, in which it serves as a platform that assists in the positioning of KU70/KU80 on DNA. Indeed, we identified two regions in ZNF384 that contribute to KU70/KU80 recruitment at sites of DNA damage: the N-terminus, which mediates the interaction with KU70/KU80, and the internal C2H2 motifs, which ensure DNA binding. The region in KU70/KU80 that is responsible for the KU-ZNF384 interaction, as well as its relevance for cNHEJ, remain to be established. Collectively, these findings

suggest that ZNF384 may act as a 'KU-adapter' that 1) senses DNA damage, 2) binds to DNA upon PARP/PAR-induced chromatin relaxation, and 3) guides KU70/KU80 for efficient loading at DNA breaks. To further support this model, which suggests a co-operative mode of action between ZNF384 and KU70/KU80, future studies may focus on understanding the spatio-temporal dynamics of these proteins at individual DSBs.

ZNF384 and DNA-PKcs act redundantly during cNHEJ protein assembly

How does ZNF384 promote the proper KU-dependent build-up of downstream NHEJ proteins at sites of DNA damage? ZNF384 and KU80 act epistatically to promote APLF recruitment to sites of DNA damage. Moreover, a direct interaction between APLF and KU's conserved KBM region has previously been shown to promote XRCC4 recruitment [32]. Our data reveal that ZNF384 is not implicated in KU-APLF complex formation, instead suggesting that ZNF384 loads XRCC4 at DSBs by promoting the recruitment of KU70/KU80 and thereby also APLF. Although our mass spectrometry analysis for ZNF384 -interacting proteins did not detect APLF, we cannot exclude the possibility that physical interactions between these proteins also contributes to XRCC4/LIG4-dependent cNHEJ.

Previous work suggested that XRCC4/LIG4 assembly also depends on the recruitment and activation of DNA-PKcs by KU-bound DNA ends [56]. However, DNA-PKcs activity remained unaffected in ZNF384- or KU80-depleted human cells, as well as in *Ku80⁻ᐟ⁻* mouse embryonic stem cells. This raises the question whether ZNF384 and KU70/KU80 promote XRCC4/LIG4 accumulation independently of DNA-PKcs. Our epistasis analysis suggests that ZNF384 cooperates with KU70/KU80, but functions independently of DNA-PKcs to promote XRCC4 accumulation at DNA breaks. This is in line with a recent report showing that KU and XRCC4/LIG4 are sufficient for DNA end synapsis independently of DNA-PKcs *in vitro* [57]. Furthermore, it may be possible that APLF dictates the functional redundancy between ZNF384 and DNA-PKcs, as APLF is recruited to sites of DNA damage via KU, PARP3 [34, 35], and ZNF384 to promote the loading of XRCC4. Finally, DNA-PKcs has been reported to have additional roles beyond NHEJ such as during mitosis, in which DNA-PKcs autophosphorylation appears to be largely independent of KU [58, 59]. This suggests the existence of KU and DNA-PKcs independent mechanisms and may explain the redundancy of ZNF384 and DNA-PKcs during cNHEJ. Future work may not only provide more insight into how

3

ZNF384 functions independently of DNA-PKcs during the assembly of a functional cNHEJ repairosome, but would also help to deepen our understanding of how DNA-PKcs function is linked to cNHEJ driven by ZNF384-KU70/KU80.

ZNF384 and other ZnF proteins in cNHEJ

We reveal a criticial role for ZNF384 in stimulating efficient cNHEJ in human cells. However, ZNF384 is not the only ZnF protein involved in cNHEJ. For instance, APLF and ZBTB24 possess distinct ZnF domains (PBZ in APLF and C2H2 in ZBTB24) that are required for the build-up of a functional NHEJ complex by binding to auto-mono(ADP-ribosyl)ated (MAR) PARP3, PARylated PARP1, or DNA, respectively, at DSBs [18, 32, 60]. This suggests that the versatile substrate recognition ability by distinct domains in ZnF proteins may play an important role in the cNHEJ process. To this end, it is interesting to note that several other ZnF proteins are recruited to sites of DNA damage in a PAR-dependent manner [12]. Although it remains to be established whether this involves direct PAR-binding or binding to damaged DNA, these findings suggest that ZnF proteins may play a more important role in DNA repair than previously anticipated. Future mechanistic studies will undoubtedly improve our understanding of their crucial role in diverse biological process, including DNA damage repair, thereby increasing our understanding of genome stability maintenance.

METHODS

Cell lines

U2OS, Hela, VH10-SV40 and SV40 T-transformed GM639 human fibroblasts cells were cultured in 5% CO_2 at 37°C in DMEM (Dulbecco's modified Eagle's medium), and DMEM and DMEM F-12 (Ham) supplemented with 10% fetal calf serum and antibiotics. RPE1-hTERT cells expressing endogenous GFP-Ku70 were a gift from Steve Jackson [61] . U2OS cells with stably integrated EJ5-GFP or DR-GFP reporters were a gift from Jeremy Stark and Maria Jasin [42, 62]. SV40 large T-transformed GM639 human fibroblasts with a stably integrated GC92 reporter were a gift from Bernard Lopez [49]. U2OS cells stably expressing cell cycle marker mKO-Cdt1 were previously generated [9]. U2OS *Asi*SI–ER- cells were a gift from Gaelle Legube [63]. U2OS 2-6-5 cells stably expressing ER-mCherry-LacR-FokI-DD were a gift from Roger Greenberg [36]. HeLa and stable GFP-Ku80 expressing HeLa cells were a gift from Dik van Gent (Erasmus Medical Center, Rotterdam, the Netherlands). 129/Ola-derived IB10 mouse

embryonic stem cells WT and Ku80[-/-] were a kind gift from Marcel Tijsterman [39]. PARP1, PARP2 and PARP1/PARP2 knockout U2OS cells were a kind gift from Nicholas Lakin [64]. XRCC4 knockout U2OS cells were generated by co-transfection of pKLV-U6gRNA-EF(BbsI)-PGKpuro2ABFP (Addgene) containing XRCC4 gRNA (5'-GATGACATGGCAATGGAAA-3') with pSpCas9(BB)-2A-GFP (PX458) containing Cas9 (Addgene). ZNF384 knockout U2OS Flp-In/T-rex cells were generated by co-transfection of pKLV-U6gRNA-EF(BbsI)-PGKpuro2ABFP (Addgene) containing ZNF384 gRNA (5'- CCACCTCTGAGAACAGGAGACTC-3') with pSpCas9(BB)-2A-GFP (PX458) containing Cas9 (Addgene). HeLa Flp-In/T-Rex and U2OS Flp-In/T-Rex cells, which were generated using the Flp-In/T-REx system (Thermo Fisher Scientific), were a gift of Geert Kops (University Medical Center Utrecht, the Netherlands) and Stephen Taylor (University of Manchester, UK). These cells were used to stably express inducible versions of GFP-NLS and GFP-APLF as well as siRNA-resistant GFP-ZNF384^{WT}, GFP-ZNF384$^{\triangle N\text{-terminus}(1\text{-}209)}$, GFP-ZNF384$^{\triangle C2H2(205\text{-}410)}$, GFP-ZNF384$^{\triangle C\text{-terminus}(401\text{-}516)}$ by co-transfection of pCDNA5/FRT/TO-Puro plasmid encoding GFP or GFP-tagged ZNF384 (WT or deletion mutants) (5 µg), together with pOG44 plasmid encoding the Flp recombinase (1 µg). After selection on 1 µg/mL puromycin, single clones were isolated and expanded. Both HeLa Flp-In/T-REx clones and U2OS Flp-In/T-Rex were incubated with 2 µg/mL doxycycline for 24h to induce expression of cDNAs. All cells were authenticated by STR profiling and tested negative in routinely performed mycoplasma tests.

Chemicals

Cells were treated with Phleomycin (InvivoGen) at the indicated concentrations for 1 hour and collected for further analysis. The PARP inhibitor olaparib (Selleck Chemicals) and DNA-PK inhibitor NU-7441 (Selleck Chemicals) were both used at a final concentration of 10 µM, whereas PARGi (PDD00017273, Sigma) inhibitor was used at a concentration of 25 mM. H_2O_2 was used at a concentration of 0.5 mM.

Transfections, siRNAs and plasmids

Cells were transfected with siRNAs using RNAiMAX (Invitrogen) according to the manufacturer's instructions. Cells were transfected twice with siRNAs at 0 and 24 hours at a concentration of 40 nM and analyzed 48 hours after the second transfection unless otherwise indicated. siRNA sequences are listed in Supplementary Table 2. Cells

were transfected with plasmid DNA using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions and analyzed 24-48 hours after transfection. The expression vector for full length human ZNF384 (pCDNA3.1-FLAG-ZNF384-WT, isoform 2 with 6 C2H2 motifs), which was a gift from Myriam Alcalay (Bajpai, Chen et al. 2010), was amplified and cloned into pCDNA5/FRT/TO-Puro as a *HindIII/KpnI* fragment (Supplementary Table 3). Deletion constructs were generated by site-directed mutagenesis PCR (Supplementary Table 3). siZNF384-3-resistant ZNF384 cDNA was generated by introducing the underlined mutations CGACAGCATAATAAGGACAAG by overlap PCR and cloned as *HindIII/KpnI* fragment into pCDNA5/FRT/TO-Puro-ZNF384-WT (Supplementary Table 2). All ZNF384 expression constructs were verified using Sanger sequencing. All other plasmids were described previously: pmCherry-PARP1 [65], pmEGFP-macroH2A1.1 macrodomain [11], GFP-WWE (from RNF146) [8], H2B-PTR, GFP-BZIP [11], GFP-CHD4 and YFP-APLF [66].

Generation of DSBs by ionizing radiation (IR)

IR was delivered to cells by an YXlon X-ray generator machine (200 KV, 4 mA, dose rate 1 Gy/minute) or a Faxitron Cabinet X-ray System Model RX-650 (130 kVp, dose rate 1.85 Gy/minute).

365 nm UV-A laser micro-irradiation

Cells were grown on 18 mm coverslips and sensitized with 10 μM 5'-bromo-2-deoxyuridine (BrdU) for 24 hours as described [9]. For micro-irradiation, the cells were placed in a Chamlide TC-A live-cell imaging chamber that was mounted on the stage of a Leica DM IRBE widefield microscope stand (Leica) integrated with a pulsed nitrogen laser (Micropoint Ablation Laser System; Andor). The pulsed nitrogen laser (16 Hz, 364 nm) was directly coupled to the epifluorescence path of the microscope and focused through a Leica 40x HCX PLAN APO 1.25–0.75 oil-immersion objective. The growth medium was replaced by CO_2-independent Leibovitz's L15 medium supplemented with 10% FCS and cells were kept at 37°C. The laser output power was set to 72-80 to generate strictly localized sub-nuclear DNA damage. Cells were micro-irradiated (two iterations per pixel) within 5 minutes using Andor IQ software (version 3.6). Following micro-irradiation, cells were incubated for the indicated time points at 37°C in Leibovitz's L15 and subsequently fixed with 4% formaldehyde before immunostaining. Images of fixed samples were acquired on a Zeiss AxioImager M2 or D2 widefield fluorescence

microscope equipped with 40x, 63x and 100x PLAN APO (1.4 NA) oil-immersion objectives (Zeiss), an HXP 120 metal-halide lamp used for excitation and the following filters: DAPI (excitation filter: 350/50 nm, dichroic mirror: 400 nm, emission filter: 460/50 nm), GFP/Alexa 488 (excitation filter: 470/40 nm, dichroic mirror: 495 nm, emission filter: 525/50 nm), mCherry (excitation filter: 560/40 nm, dichroic mirror: 585 nm, emission filter: 630/75 nm), Alexa 555 (excitation filter: 545/25 nm, dichroic mirror: 565 nm, emission filter: 605/70 nm), Alexa 647 (excitation filter: 640/30 nm, dichroic mirror: 660 nm, emission filter: 690/50 nm). Images were recorded using ZEN 2012 software (blue edition, version 1.1.0.0) and analyzed in Image J (version 1.48) as described previously (Luijsterburg, de Krijger et al. 2016). Briefly, the average pixel intensity of laser tracks was measured within the locally irradiated area (Idamage), in the nucleoplasm outside the locally irradiated area (Inucleoplasm), and in a region not containing cells in the same field of view (Ibackground). The level of protein accumulation relative to the protein level in the nucleoplasm was calculated as follows: ((Idamage - Ibackground)/(Inucleoplasm - Ibackground) – 1).

405 nm laser micro-irradiation

Laser micro-irradiation for local photoactivation and DNA damage induction at 405 nm was performed using a single-point scanning head (iLas2 from Roper Scientific) coupled to the epifluorescence backboard of a Nikon Ti-E inverted microscope equipped with a spinning-disk scan head CSU-X1 from Yokogawa at a rotation speed of 5000 rpm, a Plan APO 60x/1.4 N.A oil-immersion objective lens and a sCMOS ORCA Flash 4.0 camera. The fluorescence of EGFP and mCherry/activated PATagRFP were excited with lasers at 490 nm and 561 nm, respectively. Bandpass filters adapted to the fluorophore emission spectra were used for fluorescence detection. Images were acquired using Metamorph software (version 7.8.2.0). Cells were sensitized with media containing 0.3 µg/mL Hoechst 33342 for 1 hour at 37°C. Prior to imaging, the medium was replaced with CO_2-independent phenol red-free Leibovitz's L15 medium (Life Technologies) supplemented with 20% FCS. Cells were irradiated with a 16 µm line through the nucleus to simultaneously induce DNA damage and photoactivate PATagRFP. The 405 nm laser power was measured at the beginning of each experiment and set to 125 µW at the sample level to ensure reproducibility. For PAR-3H experiments and ZNF384 recruitment experiments, images were collected every 5 seconds for 10 minutes. For ZNF384 and WWE recruitment with late PARP inhibitor treatment, Z-stacks (1

µm steps) of irradiated nuclei were collected every 30 seconds for 15 minutes. Image collection was paused 3 minutes post damage and olaparib was added to the imaging media to a final concentration of 30 µM. For protein recruitment analysis, a custom-made MATLAB (MathWorks) program R2014b (version 8.4.0.150421) (available upon request). For reviewing purposes the following link can be used: https://github.com/sehuet/Singh-image-processing was used to segment the site of damage (I_d) as determined by the photoactivated H2B area, the total nuclear fluorescence (I_{nd}), and an area of background outside of the cell (I_{bg}). Protein accumulation at sites of damage (A_d) was calculated as:

$$A_d = \frac{I_d - I_{bg}}{I_n - I_{bg}} \ (1)$$

The intensity within the micro-irradiated area was then normalized to the intensity prior to damage induction. Chromatin relaxation was determined by measuring the change in thickness of the photoconverted H2B line [8].

Multiphoton laser micro-irradiation

Cells grown were grown on 18 mm coverslips. For micro-irradiation, cells were placed in a Chamlide CMB magnetic chamber and the growth medium was replaced by CO_2-independent Leibovitz's L15 medium supplemented with 10% FCS and antibiotics. Laser micro-irradiation was performed on a Leica SP5 confocal microscope equipped with an environmental chamber set to 37°C. DNA damage-containing tracks (1.5 µm width) were generated with a Mira mode locked titanium-sapphire (Ti:Sapphire) laser (l = 800 nm, pulse length = 200 fs, repetition rate = 76 MHz, output power = 80 mW) using a UV-transmitting 63x 1.4 NA oil immersion objective (HCX PL APO; Leica). Confocal images were recorded before and after laser irradiation at 5 seconds time interval over a period of 3-5 minutes. Images after multi-photon micro-irradiation of living cells were recorded using LAS-AF software (Leica, light version 1.0.0) and analyzed with Image J (version 1.48) as described previously [9]. The average pixel intensity of laser tracks was measured within the locally irradiated area (Idamage), in the nucleoplasm outside the locally irradiated area (Inucleoplasm) and in a region not containing cells in the same field of view (Ibackground). The level of protein accumulation relative to the protein level in the nucleoplasm was calculated as follows: ((Idamage - Ibackground)/(Inucleoplasm - Ibackground) – 1).

Ultra-soft X-ray irradiation and imaging

The U2OS Flp-In/T-Rex cells stably expressing inducible GFP-tagged ZNF384 were incubated with doxycycline (2 µM) for 2.5 hours before being irradiated using the previously described ultra-soft X-ray system [19]. To obtain locally concentrated DSBs, a custom-designed irradiation mask with parallel apertures (2.5 µm wide) was placed between the bottom of the cell culture dishes and the X-ray source. Cells were irradiated for 10 seconds at 40 mA emission current and 6 KeV acceleration voltage, resulting in approximately 1,000 DSBs per irradiated area. Images of cells and of the irradiation mask were collected five minutes after exposure. For experiments involving immunofluorescence imaging, cells were exposed as described above, without doxycycline preincubation. Five minutes after irradiation, cells were fixed and immunostained. Wide-field 3D images were acquired using a Leica DMi8 microscope (63x/1.4 NA) and deconvolved using Huygens Professional (version 19.10). Confocal images of immunostained cells were captured using a Leica SP8-X SMD (63x/1.4 NA).

Proximity Ligation Assay (PLA)

U2OS AsiSI-ER cells were seeded on 12-mm coverslips and after 24 hours treated with 1 µM 4-Hydroxytamoxifen (4-OHT, Sigma) for 5h to induce DSBs. Subsequently, cells were fixed in 4% paraformaldehyde and permeabilized in 0.5% TritonX-100. Primary antibodies rabbit anti-53BP1 (NOVUS Biologicals NB100-304), mouse anti-γH2AX (Millipore clone JBW301), and rabbit anti-ZNF384 (ATLAS antibodies HPA004051) were used to stain selected proteins. Proximity Ligation Assay was performed with Duolink In Situ PLA Probe Anti-Mouse Plus (Sigma) and Anti-Rabbit Minus (Sigma), and with Duolink In Situ Detection Reagents Orange (Sigma) according to manufacturer's instructions. Finally, secondary antibodies anti-rabbit coupled to Alexa488 (Invitrogen) and anti-mouse coupled to Alexa 647 (Invitrogen) were used to stain selected proteins in immunofluorescence. Number and intensity of PLA foci per cell were analysed by the imaging software ImageJ version 1.48.

Fluorescence Recovery After Photobleaching of ZNF384

FRAP of GFP-tagged ZNF384 constructs was performed on performed on a Nikon Ti-E inverted microscope equipped with a spinning-disk scan head CSU-X1 from Yokogawa at a rotation speed of 5000 rpm, a Plan APO 60x/1.4 N.A oil-immersion objective lens and a sCMOS ORCA Flash 4.0 camera. The fluorescence of EGFP was excited with lasers at 490

nm. Bandpass filters adapted to the fluorophore emission spectra were used for fluorescence detection. Local bleaching within a 4 μm diameter circular area in the cell nucleus was performed using a dedicated single-point scanning head (iLas2 from Roper Scientific) coupled to the epifluorescence backboard of the microscope. Images were collected at 2 images/second. To estimate fluorescence recovery kinetics, the mean fluorescence intensity inside the bleached area was measured by automatic segmentation using a custom-made custom-made MATLAB (MathWorks) program R2014b (version 8.4.0.150421) (available upon request). This routine allowed for background subtraction from the intensity measurements and correction for photobleaching due to imaging by dividing the intensity in the bleached area with the one measured for the whole nucleus. The recovery time was the time required to recover half of the fluorescence signal lost upon photobleaching.

Fluorescence Recovery After Photobleaching of Ku70

Fluorescence Recovery After Photobleaching (FRAP) of Ku70 was performed on a Zeiss LSM880 confocal setup equipped with a Plan APO 63x/1.2 N.A. water immersion objective. Samples were maintained at 37°C using a heating chamber. GFP fluorescence was excited at 488 nm and emission was detected at 500-590 nm. DNA damage was induced in a 6x2 μm area of the cell nucleus with a pulsed infrared laser set at 800 nm (Mai Tai, Spectra-Physics). Regions of interest of sizes ranging between 1 and 4 μm² located inside the previously irradiated area were bleached using a 488 nm laser. Images of the subsequent fluorescence recovery were collected at 4 frames per second using Zen Black (version 14.0.9.201). After background subtraction, the fluorescence recovery kinetics were obtained by dividing the signal within the bleached area to the one measured in the unbleached part of the damaged region.

Fluorescence correlation spectroscopy of Ku70

Fluorescence correlation spectroscopy (FCS) of Ku70 was performed on a Zeiss LSM880 confocal microscope equipped with a C-Apo 40x/1.2 N.A water immersion lens. GFP fluorescence was excited with a 488 nm laser and single emitted photons at wavelengths ranging between 500 and 550 nm were detected and counted on the GaAsP spectral detector. The laser power used for FCS measurements was adjusted to minimize photobleaching. FCS acquisition lasted 30 seconds to reduce the noise in the autocorrelation curves. Samples were maintained at 37°C using a heating chamber. FCS curves were detrended for slow fluctuations using Fluctuation Analyzer 4G [67].

Mathematical models for fitting of the FRAP and FCS data

The diffusion-limited model used to fit the FRAP curves is expressed as,

$$frap(t) = e^{\frac{-T_D}{2t}} \left[I_0\left(\frac{T_D}{2t}\right) + I_1\left(\frac{T_D}{2t}\right) \right] \quad (2)$$

where T_D is the characteristic diffusion time within the bleached area and I_0 and I_1 are modified Bessel functions of the first kind. The parameter T_D varies with the diffusion coefficient of Ku70 but also, in case of transient interactions with DNA, it depends on the K_d of this interaction [68]. The reaction-limited model assumes that Ku70 interacts with the DNA breaks according to the following reaction:

$$Free + BS \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} Bound \quad (3)$$
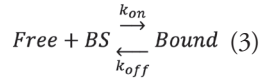
with *Free* and *Bound* referring to the binding state of Ku and *BS* to the break site. The mathematical expression of the reaction-limited model is then as follows:

$$frap(t) = 1 - \frac{k'_{on}}{k'_{on} + k_{off}} e^{-k_{off}t} \quad (4)$$

where k'_{on} is the pseudo first-order association rate corresponding to the product of the association rate k_{on} by the local concentration of break sites [*BS*] and k_{off} is the dissociation rate. The one-population model used to fit the correlation curves is expressed as

$$G(t) = \frac{1}{2^{3/2}N} \left(1 + \frac{t}{\tau}\right)^{-1} \left(1 + \frac{t}{\omega^2\tau}\right)^{-1/2} \quad (5)$$

where N is the number of tagged molecules in the focal volume, τ is the residence time in the focal volume, and ω is the structural parameter of the focal volume, which was fixed to 6. Similarly, the two-population model used to fit the correlation curves is expressed as

$$G(t) = \frac{1}{2^{3/2}N} \left[f_1 \left(1 + \frac{t}{\tau_1}\right)^{-1} \left(1 + \frac{t}{\omega^2\tau_1}\right)^{-1/2} + (1 - f_1)\left(1 + \frac{t}{\tau_2}\right)^{-1} \left(1 + \frac{t}{\omega^2\tau_2}\right)^{-1/2} \right] \quad (6)$$

where τ_1 and τ_2 are the residence times of the two populations in the focal volume and f_1 is the fraction of the molecules belonging to the population displaying a residence time τ_1.

Fluorescence three-hybrid assay

Fluorescence three-hybrid/PAR-3H assays were performed as described [26]. Briefly, GFP-tagged proteins were tethered to a genomically integrated LacO array using a LacI-GFP trap in U2OS-2B2 cells [69] expressing mCherry-PARP1. Cells were sensitized with Hoechst and micro-irradiated with 405 nm light to induce DNA damage. If the GFP-tagged protein of interest is able to bind PAR, PARylated mCherry-PARP1, which

is generated at sites of DNA damage, will enrich at the LacO array after DNA damage induction. The mCherry-PARP1 signal intensity at the LacO array was quantified pre and 30, 60 and 120 seconds post DNA damage induction. The average intensity at the lacO array was normalized to the average intensity of the nucleus and corrected for background signal.

Immunofluorescence analysis

Cells were either directly fixed with 2% formaldehyde in PBS for 20 minutes at room temperature (RT), or pre-extracted with 0.5% Triton-X100 (Serva) in PBS on ice for 2 minutes prior to fixation. Alternatively, cells were fixed, post-extracted with 0.25% Triton-X100 (Serva) in PBS and treated with 100 mM glycine in PBS for 20 minutes to block unreacted aldehyde groups. Cells were then rinsed with PBS and equilibrated in wash buffer (PBS containing 0.5% BSA). Antibody incubation steps and washes were in wash buffer. Primary antibodies were incubated for 1-2 hours at room temperature. Detection was done using goat anti-mouse or goat anti-rabbit Ig coupled to Alexa 488, 555 or 647 (1:1500; Invitrogen Molecular probes) or Cy3-conjugated goat anti-mouse secondary antibody (1:100; Jackson Immuno research). All antibodies are listed in Supplementary Table 4. Samples were incubated with 0.1 µg/mL 4',6-Diamidino-2-Phenylindole Dihydrochloride (DAPI) and mounted in Polymount.

DSB reporter assay

U2OS 2-6-5 cells stably expressing ER-mCherry-LacR-FokI-DD [36] were treated for 5 hours with 1 µM Shield-1 (Clontech Laboratories UK Ltd) and 1 µM 4-hydroxytamoxifen (4-OHT, Sigma-Aldrich) to induce DSBs.

Pulldown and co-immunoprecipitation assays

GFP-pulldowns were performed on U2OS Flp-In/T-Rex cells expressing GFP-NLS, GFP-ZNF384[WT] or the indicated mutants and on HeLa and GFP-Ku80-expresing HeLa cells, while untransfected U2OS cells were used for co-immunoprecipitation assays. Cells were lysed in EBC buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.5% NP-40, 2 mM $MgCl_2$, protease inhibitor cocktail tablets) with 500 units benzonase. Samples were incubated for 1 hour at 4ºC under constant mixing. 50 µL input sample was collected in a separate tube and mixed with 2x Laemmli buffer. The cleared lysates were subjected to GFP pulldown with GFP-Trap

beads (Chromotek) or immunoprecipitation using a specific antibody (or corresponding IgG control) that was conjugated to Protein G-coupled agarose beads (Millipore 16-201). The beads were then washed 6 times with EBC buffer and boiled in 2x Laemmli buffer along with the input samples. Samples were subjected to western blot analysis.

Sample preparation and mass spectrometry

For mass spectrometry, U2OS Flp-In/T-Rex cells expressing GFP-NLS and GFP-ZNF384[WT] were treated with 0.5 mM Phleomycin for 1 hour or left untreated. Cell pellets were lysed in EBC-1 buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.5% NP-40, 2 mM $MgCl_2$, protease inhibitor cocktail tablets) with 500 units benzonase. Samples were incubated for 1 hour at 4ºC under constant mixing followed by high speed centrifugation for 10 minutes at 4ºC. Protein concentration was measured by Qubit in the cleared lysates, equalized and transferred to tubes containing GFP-Trap beads (Chromotek). After 90 minutes of incubation at 4ºC under rotating condition, the beads were washed 4 times with EBC-2 buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, and protease inhibitor cocktail tablets) and 3 times with 50 mM ammonium bicarbonate followed by overnight digestion using 2.5 μg trypsin at 37°C under constant shaking. Digestion was terminated with 1% trifluoroacetic acid and centrifuged for 5 minutes at high speed to precipitate insoluble fractions. Consequently, C18 cartridges were prepared by washing 2 times with acetonitrile followed by 2 times with 0,1% Acetic acid. Peptides were loaded on the cartridge, while bound peptides were washed 2 times with 0.1% acetic acid and eluted with 1 mL 80% acetonitrile / 0.1% acetic acid and lyophilized.

Mass spectrometry was performed essentially as previously described [70]. Samples were analyzed on a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific, Germany) coupled to an EASY-nanoLC 1000 system (Proxeon, Odense, Denmark). Digested peptides were separated using a 15 cm fused silica capillary (ID: 75 μm, OD: 375 μm, Polymicro Technologies,California, US) in-house packed with 1.9 μm C18-AQ beads (Reprospher-DE, Pur, Dr. Maisch, Ammerburch-Entringen, Germany). Peptides were separated by liquid chromatography using a gradient from 2% to 95% acetonitrile with 0.1% formic acid at a flow rate of 200 nl/minute for 65 minutes. The mass spectrometer was operated in positive-ion mode at 2.8 kV with the capillary heated to 250°C, and in a Data-Dependent Acquisition (DDA) mode with a top 7 method. Full scan MS spectra were obtained with a resolution of 70,000, a target value of

3×10^6 and a scan range from 400 to 2,000 m/z. Maximum Injection Time (IT) was set to 50 ms. Higher-Collisional Dissociation (HCD) tandem mass spectra (MS/MS) were recorded with a resolution of 35,000, a maximum IT of 20 ms, a target value of 1×10^5 and a normalized collision energy of 25%. The precursor ion masses selected for MS/MS analysis were subsequently dynamically excluded from MS/MS analysis for 60 seconds. Precursor ions with a charge state of 1 and greater than 6 were excluded from triggering MS/MS events. Three replicates were included per condition with two technical repeats each.

Mass spectrometry data analysis

Raw mass spectrometry files were analysed with MaxQuant software (v1.5.5.1) as described [71] with the following modifications from default settings: the maximum number of mis-cleavages by trypsin/p was set to 3, Label Free Quantification (LFQ) was enabled disabling the Fast LFQ feature. Match-between-runs feature was enabled with a match time window of 0.7 minutes and an alignment time window of 20 minutes. We performed the search against an in silico digested UniProt reference proteome for Homo sapiens (8th June 2020). Analysis output from MaxQuant was further processed in the Perseus computational platform (version 1.5.5.3) [71]. Proteins identified as common contaminants, only identified by site and reverse peptide, were filtered out, and then all the LFQ intensities were log2 transformed. Different biological repeats of each condition were grouped and only protein groups identified in all three biological replicates in at least one condition were included for further analysis. Missing values were imputed using Perseus software by normally distributed values with a 1.8 downshift (log2) and a randomized 0.3 width (log2) considering total matrix values. Volcano plots were generated, and Student's T-tests were performed to compare the different conditions. Spreadsheets from the statistical analysis output from Perseus were further processed in Microsoft Excel for comprehensive visualization and analysis of the data (Supplementary Table 1).

Western blot analysis

Cells were lysed in 2x Laemmli buffer and proteins were separated by Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis (SDS-PAGE) using 4-12% pre-cast polyacrylamide gels (BioRad or Invitrogen) and MOPS running buffer (Invitrogen). Next, proteins were transferred onto nitrocellulose membranes (Millipore). Protein expression was analyzed by immunoblotting with the indicated primary antibodies

(Table 3) and secondary CF680 goat anti-rabbit or CF770 goat anti-mouse Ig antibodies (1:5000, Biotium). Membranes were scanned and analyzed using an Odyssey Infrared Imaging System (Licor; V3.0) and Image Studio Lite (version 5.2). Uncropped blots are provided in the source data file.

Chromatin fractionation

Chromatin fractionation was performed using a previously published protocol [72] with few modifications. Briefly, 100.000-150.000 cells were grown per 6-cm dish for 24 h and then transfected with siRNAs. Next, the cells were treated with 500 µM phleomycin for 1 hour, washed three times with PBS, and incubated in NETN extraction buffer (100mM NaCl, 1mM EDTA, 20mM Tris-Cl pH8, 0,5% NP-40 + Proteasome inhibitors). After 15 minutes of incubation on ice, samples were taken for the chromatin-unbound fraction and mixed with the same amount of 2x Laemmli buffer. Cells were washed with PBS, lysed, and incubated in Laemmli buffer with benzonase for 15 minutes to obtain the chromatin-bound fraction. Samples were heated for 7 minutes at 80°C and subjected to Western blot analysis.

MBP-based protein purifications

For MBP-based purification, cultures of *Escherichia Coli* BL21-CodonPlus (DE3)-RIL cells containing pET-His6-MBP,pET-His6-MBP-ZNF384, pET-His6-MBP-C2H2, pET-His6-MBP-N-terminus and pET-His6-MBP-C-terminus plasmids were grown to an OD_{600} of 0.3 absorbance units. To start induction of protein expression, 0.3 mM IPTG was added to the culture followed by incubation overnight at 20 °C. After centrifugation, cell pellets from were frozen and stored at -80 °C. For protein purification, cell pellets were lysed in 5 ml B-per™ Bacterial Protein Extraction Reagent (Thermo Fisher Scientific) supplemented with protease inhibitors (Sigma) and 15 kU rLysozyme (Merck) until the lysates were clear. Viscosity of the lysate was decreased by the addition of 125 units benzonase or sonication. The lysate was centrifuged for centrifuge for 10 min at 21000 g at 4°C in a table centrifuge 21000 g. For the ZNF384 FL, N-terminus and C- terminus fragments, the supernatant was loaded on a column packed with 0.75 ml Amylose Resin High Flow (NEB) installed in ÄKTA pure protein purification system (Cytiva). The column was washed with buffer A (20 mM Tris pH 7.4, 200 mM NaCl, 1 mM EDTA, 10 mM β-ME) and the proteins were eluted with buffer B (buffer A + 10 mM maltose). For the ZNF384 C2H2 fragment, two purification steps were performed. First, the supernatant was loaded on HiTrap SP

HP Strong Cation Exchange column (Cytiva) and proteins were eluted using a linear gradient from 50 to 1000 mM NaCl in 20 mM Tris pH 7.4, 10 mM β-ME collecting 2 ml fractions. Second, the fractions containing C2H2 were loaded on to the Amylose column, followed by washing and elution with buffer A and B, respectively.

Biotinylated DNA substrates

Biotinylated DNA substrates (Supplementary Table 5) were used at a concentration of 1 pmol/µl. dsDNA substrates were made by annealing complementary oligo's (Table 2 and 4) in reaction buffer (10 mM Tris pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 0.25% Tween-20, 3.5 mM DTT, 5 % glycerol). Annealing was done in a PCR machine heating for 2 minutes to 95 °C, then gradually cooling over a period of 45 minutes to 25 °C.

DNA pulldown assay

DNA binding reactions were done at 4 °C in 40 µl reaction buffer containing 0.1% BSA, 10 pmol biotinylated DNA substrate (Supplementary Table 5), and ~50 fmol of purified MBP or the different MBP-tagged ZNF384 proteins. After 30 minutes, reaction buffer containing 10 µl Dynabeads M-280 Streptavidin suspension and 0.1% BSA was added and samples were incubated for 15 minutes at 4 °C. After this incubation, beads were washed 3 times using 200 µl reaction buffer and loaded on 4-12% polyacrylamide Bis-Tris gel. After electrophoresis, proteins were blotted onto PVDF membranes for one hour at 50 Volt. Membranes were incubated at room temperature for 1 hour with mouse monoclonal anti-MBP antibody (NEB), followed by 1 hour incubation with goat anti-mouse HRP antibody (Bethyl Laboratories), and imaged by AI680 imager (GE) with ECL. For Ku70/Ku80 DNA binding reactions to 3'overhang DNA, His-MBP-ZNF384 or His-MBP, 100 fmol Ku70/Ku80 heterodimer and 10 pmol 3'-overhang biotinylated oligo was incubated in 40 ul reaction buffer (10 mM Tris pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 0.25% Tween-20, 3.5 mM DTT, 5 % glycerol) with 0.1% BSA for 30 minutes at 4 °C. After 30 minutes, reaction buffer containing 10 µl Dynabeads M-280 Streptavidin suspension and 0.1% BSA was added and samples were incubated for 30 minutes at 4 °C. After this incubation, beads were washed three times with 200 µl reaction buffer, Laemmli sample buffer was added and samples were incubated for 10 minutes at 95 °C. For separation by electrophoresis samples were loaded on Bolt 4-12% polyacrylamide Bis-Tris Mini Protein gel (ThermoFisher). After electrophoresis, western blotting was performed and the blots were stained using mouse anti-MBP (NEB) and rabbit anti-

Ku80 (Santa Cruz) as primary antibodies and goat anti-Rabbit[CF680] and goat anti-Mouse[CF770] (Biotium) as secondary antibodies for detection with an Odyssey Infrared Imaging System (Licor; V3.0).

In vitro pulldown assay

Protein reactions and washing steps were all done at room temperature. 15 µl Dynabeads™ M-280 Sheep anti-Mouse IgG (Thermo Scientific) were incubated with 2 µg mouse anti-Ku80 antibody (Santa Cruz) in 40 µl PBS containing 0.1% BSA for 30 minutes. Beads were collected using a magnetic rack. After discarding the supernatant, beads were washed with 200 µl PBS containing 0.1% BSA and incubated for 30 minutes with 6 pmol recombinant His-Ku70/Ku80 (Sino Biological) in 10 µl reaction buffer (10 mM Tris pH 7.5, 150 mM KCl, 5 mM MgCl2, 0.25% Tween-20, 3.5 mM DTT, 5 % glycerol). Beads were collected and supernatant was discarded. Beads were then washed with 200 µl reaction buffer and 10 µl of 0.2% BSA was added to the reaction buffer. After 15 minutes, 10 µl reaction buffer with 1 pmol of purified His-MBP-tagged protein was added and samples were incubated for 30 minutes. Beads were collected and the supernatant was removed. Beads were subsequently washed 4 times with 200 µl reaction buffer. Samples were heated in Laemmli sample buffer for 10 minutes at 95 °C. For electrophoresis, samples were loaded on Bolt 4-12% polyacrylamide Bis-Tris Mini Protein gel (ThermoFisher). After electrophoresis, Western blotting was performed and the blots were stained using mouse anti-MBP (NEB) and rabbit anti-Ku80 (Santa Cruz) as primary antibodies and goat anti-Rabbit[CF680] and goat anti-Mouse[CF770] (Biotium) as secondary antibodies for detection with an Odyssey Infrared Imaging System (Licor; V3.0).

Biolayer interferometry (BLI) measurements

BLI measurements were done on an OctedRed System (Sartorius), shaking the assay plate (1000 rpm) at 298K. All steps were performed in BLI buffer (10 mM Tris PH7.5, 150 mM KCl, 5 mM $MgCl_2$, 0.05% (v/v) Tween-20, 0.1% (w/v) BSA, 1 mM DTT). Biotinylated DNA substrates (Supplementary Table 5) were immobilised on streptavidin sensors pre-equilibrated in BLI buffer, after which a washout in BLI buffer was done. Then, 400 nM of purified MBP or the different MBP-tagged ZNF384 proteins was used to measure the association of the analyte. Resulting data were processed using the ForteBio Data Analysis software (version 7.1.0.38).

Quantitative Fluorescence *In Situ* Hybridization (FISH) of telomeres

Telomere FISH was based on a previously published protocol [51]. Briefly, HCT116 cells were harvested following 2 hours of colcemid (Sigma) incubation. After hypotonic swelling, cells were fixed in methanol/acetic acid, dropped on slides and dried at 37°C overnight. The next day slides were treated with RNAseA (R4642; Sigma), pepsin (P7000; Sigma) at pH 2, followed by formaldehyde fixation, washes in PBS, dehydration in ethanol and air drying. Hybridization mixture containing 70% formamide, 0.3 mg/ml Cy-3-conjugated (C3TA2)3 peptide nucleic acid (PNA) probe in 10 mM Tris (pH 7.5) was added to the slide, after which a coverslip was added followed by DNA denaturation for 1.5 minute at 80°C. After hybridization for 2 hours at room temperature, slides were washed with 70% formamide / 10 mM Tris / pH 7.2, and 0.05 M Tris / 0.15 M NaCl containing 0.05% Tween-20. Slides were stained with DAPI, dehydrated with ethanol, air dried, and mounted in Aquapolymount (Polysciences). Images were acquired on a Zeiss AxioImager M2 widefield fluorescence microscope with 63x PLAN APO (1.4 NA) oil-immersion objectives (Zeiss). Integrated density and area of single telomeres was measured with ImageJ (version 1.48) by using Threshold, Polygon and Analyze Particles functions, subsequently. While obtaining images we noticed variation of telomere signals between metaphases and less so within one metaphase spread. Therefore, after subtraction of background values the average integrated density per telomere of each metaphase was calculated and plotted.

DR-GFP and EJ5-GFP reporter assays

U2OS cells containing either a stably integrated copy of the DR-GFP or EJ5-GFP reporter were used to measure the repair of I-*Sce*I-induced DSBs by HR or NHEJ [42, 62]. Briefly, DR-GFP U2OS cells or EJ5-GFP U2OS cells treated with siRNA for 48 hours were co-transfected with an mCherry expression vector and the I-*Sce*I expression vector pCBASceI [62]. 48 hours later the percentage of GFP-positive cells among the mCherry-positive cells was determined by FACS on a BD LSRII flow cytometer (BD Bioscience) using FACSDiva software version 5.0.3. An example of the gating strategy can be found Fig S15I. Quantifications were performed with FACSDiva™ (BD Biosciences).

Random plasmid integration assay

U2OS cells were seeded (day 1) and transfected with siRNAs the following day (day 2). Later at day 2, the cells were transfected with 2 μg gel-purified *Bam*HI/*Eco*RI-linearized pEGFP-C1 plasmid. The cells were subsequently transfected twice with siRNAs at 24 and 36 hours after the first transfection (day 3 and day 4, respectively). On day 5, cells were collected, counted, seeded and grown in medium without or with 0.5 mg/mL G418. The transfection efficiency was determined on the same day by FACS analysis using GFP fluorescence as a measure. The cells were incubated at 37°C to allow colony formation and medium was refreshed on day 8 and 12. On day 15, the cells were washed with 0.9% NaCl and stained with methylene blue (2.5 g/L in 5% ethanol, Sigma-Aldrich). Colonies of more than 50 cells were scored. Random plasmid integration efficiency was scored as the number of G418-resistant colonies normalized by the plating efficiency, which was determined by the number of colonies formed on plates without G418 and corrected for the transfection efficiency.

Analysis of repair junctions in the GC92 reporter

Sequence analysis of repair junctions in the GC92 reporter was performed as described [49]. Briefly, GC92 fibroblasts were first transfected with siRNAs and 48 hours later with the I-*Sce*I expression vector pCBASce [62]. 48 hours later, genomic DNA was extracted using phenol:chloroform:isoamyl alcohol (25:24:1 v/v, Invitrogen). PCR was performed on the genomic DNA using the CMV1 and CD4int primers (Supplementary Table 3) to amplify repair junctions. PCR products were cloned into pGEM-T easy vector (Promega). Colony PCR was performed using M13 primers (Supplementary Table 3) on individual bacterial colonies to amplify repair junctions, which were subjected to Sanger sequencing using the M13 FW primer (Supplementary Table 3). Sequences were analyzed using a custom Sanger sequence analyzer as described previously [39].

Cell survival assays

VH10-SV40 cells were transfected with siRNAs, trypsinized, seeded at low density and exposed to IR. For HeLa Flp-In/T-Rex, cDNAs were expressed by adding Dox for 24 hours after siRNA transfection. U2OS cells were seeded at low densities and exposed to increasing doses of olaparib. After 7 days, the cells were washed with 0.9% NaCl and stained with methylene blue (2.5 g/L in 5% ethanol, Sigma-Aldrich). Colonies of more than 20 cells were scored.

Cell cycle profiling

Cells were fixed in 70% ethanol, followed by DNA staining with 50 µg/mL propidium iodide in the presence of RNase A (0.1 mg/mL; Sigma). Cell acquisition and quantification was performed on a BD LSRII flow cytometer (BD Bioscience) using FACSDiva software version 5.0.3.

Statistics and reproducibility

Results were confirmed in multiple cell lines or by using complementary approaches. All experiments yielding micrographs, pull-down experiments and western blot analysis were performed independently at least twice, but often three times. The MS experiments were performed in triplicate. Statistical analysis was carried out using the two-tailed Student's t test ($P < 0.05$). Boxplots were generated using R (version 4.0.5) and R Studio (version 1.4.1106).

DATA AVAILABILITY

The mass spectrometry proteomics data generated in this study and shown in Fig. 1D and Table S1 have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (https://www.ebi.ac.uk/pride/). Access can be obtained with the dataset identifier PDX020417 [73]. Additionally, publicly available reference datasets for Homo sapiens (8th June 2020) was used to search against an in sillico digested UniProt reference proteome. Source data are provided with this paper.

CODE AVAILABILITY

The code used in this work is available at https://github.com/sehuet/Singh-image-processing [74].

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTION

J.S. generated partial ZNF384 knockout cells, ZNF384 plasmids and performed laser micro-irradiation, cell cycle profiling, pulldown and co-immunoprecipitation experiments, as well as plasmid integration, EJ5-GFP reporter assay and clonogenic survivals. R.S., O.D. and S.H. performed laser micro-irradiation, fluorescence three-hybrid assay, fluorescence correlation spectroscopy and FRAP. W.W.W. performed DR-GFP, immunofluorescence, FISH, and clonogenic survivals. K.V. generated all ZNF384 deletion constructs. M.B.R. performed GC92 reporter assays, pulldown experiments and generated ZNF384 plasmids. A.J.L.G prepared mass-spectrometry samples, purified recombinant proteins and performed *in vitro* DNA pulldown experiments. H.Q. and P.M.K. performed Ultra-soft X-ray irradiation experiments. R.Q.K. and M.L. performed BLI experiments. R.P. and A.C.O.V. performed mass-spectrometry experiments and data-analysis. H.v.A. conceived and supervised the project. J.S and H.v.A. wrote the manuscript with input and edits from all authors.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

1. Jackson, S.P. and J. Bartek, *The DNA-damage response in human biology and disease.* Nature, 2009. **461**(7267): p. 1071-8.

2. Symington, L.S. and J. Gautier, *Double-strand break end resection and repair pathway choice.* Annu Rev Genet, 2011. **45**: p. 247-71.

3. Chang, H.H.Y., et al., *Non-homologous DNA end joining and alternative pathways to double-strand break repair.* Nat Rev Mol Cell Biol, 2017. **18**(8): p. 495-506.

4. Bhargava, R., D.O. Onyango, and J.M. Stark, *Regulation of Single-Strand Annealing and its Role in Genome Maintenance.* Trends Genet, 2016. **32**(9): p. 566-575.

5. Khurana, S., et al., *A macrohistone variant links dynamic chromatin compaction to BRCA1-dependent genome maintenance.* Cell Rep, 2014. **8**(4): p. 1049-62.

6. Dantuma, N.P. and H. van Attikum, *Spatiotemporal regulation of posttranslational modifications in the DNA damage response.* Embo j, 2016. **35**(1): p. 6-23.

7. Caron, P., et al., *WWP2 ubiquitylates RNA polymerase II for DNA-PK-dependent transcription arrest and repair at DNA breaks.* Genes Dev, 2019. **33**(11-12): p. 684-704.

8. Sellou, H., et al., *The poly(ADP-ribose)-dependent chromatin remodeler Alc1 induces local chromatin relaxation upon DNA damage.* Mol Biol Cell, 2016. **27**(24): p. 3791-3799.

9. Luijsterburg, M.S., et al., *PARP1 Links CHD2-Mediated Chromatin Expansion and H3.3 Deposition to DNA Repair by Non-homologous End-Joining.* Mol Cell, 2016. **61**(4): p. 547-562.

10. Gibbs-Seymour, I., et al., *HPF1/C4orf27 Is a PARP-1-Interacting Protein that Regulates PARP-1 ADP-Ribosylation Activity.* Mol Cell, 2016. **62**(3): p. 432-442.

11. Smith, R., et al., *Poly(ADP-ribose)-dependent chromatin unfolding facilitates the association of DNA-binding proteins with DNA at sites of damage.* Nucleic Acids Res, 2019. **47**(21): p. 11250-11267.

12. Izhar, L., et al., *A Systematic Analysis of Factors Localized to Damaged Chromatin Reveals PARP-Dependent Recruitment of Transcription Factors.* Cell Rep, 2015. **11**(9): p. 1486-500.

13. Vilas, C.K., et al., *Caught with One's Zinc Fingers in the Genome Integrity Cookie Jar.* Trends Genet, 2018. **34**(4): p. 313-325.

14. Singh, J.K. and H. van Attikum, *DNA double-strand break repair: Putting zinc fingers on the sore spot.* Semin Cell Dev Biol, 2020.

15. Cassandri, M., et al., *Zinc-finger proteins in health and disease.* Cell Death Discov, 2017. **3**: p. 17071.

16. Gong, F., et al., *Screen identifies bromodomain protein ZMYND8 in chromatin recognition of transcription-associated DNA damage that promotes homologous recombination.* Genes Dev, 2015. **29**(2): p. 197-211.

17. Chen, G., et al., *ZNF830 mediates cancer chemoresistance through promoting homologous-recombination repair.* Nucleic Acids Res, 2018. **46**(3): p. 1266-1279.

18. Nicolai, S., et al., *ZNF281 is recruited on DNA breaks to facilitate DNA repair by non-homologous end joining.* Oncogene, 2020. **39**(4): p. 754-766.

19. Kochan, J.A., et al., *Ultra-soft X-ray system for imaging the early cellular responses to X-ray induced DNA damage.* Nucleic Acids Research, 2019. **47**(17): p. e100-e100.

20. Galbiati, A., C. Beauséjour, and F. d'Adda di Fagagna, *A novel single-cell method provides direct evidence of persistent DNA damage in senescent cells and aged mammalian tissues.* Aging Cell, 2017. **16**(2): p. 422-427.

21. Fell, V.L. and C. Schild-Poulter, *The Ku heterodimer: function in DNA repair and beyond.* Mutat Res Rev Mutat Res, 2015. **763**: p. 15-29.

22. Davis, A.J., B.P. Chen, and D.J. Chen, *DNA-PK: a dynamic enzyme in a versatile DSB repair pathway.* DNA Repair (Amst), 2014. **17**: p. 21-9.

23. Beck, C., et al., *PARP3 affects the relative contribution of homologous recombination and nonhomologous end-joining pathways.* Nucleic Acids Res, 2014. **42**(9): p. 5616-32.

24. Gogola, E., et al., *Selective Loss of PARG Restores PARylation and Counteracts PARP Inhibitor-Mediated Synthetic Lethality.* Cancer Cell, 2018. **33**(6): p. 1078-1093.e12.

25. Jungmichel, S., et al., *Proteome-wide identification of poly(ADP-Ribosyl)ation targets in different genotoxic stress responses.* Mol Cell, 2013. **52**(2): p. 272-85.

26. Smith, R., et al., *CHD3 and CHD4 recruitment and chromatin remodeling activity at DNA breaks is promoted by early poly(ADP-ribose)-dependent chromatin relaxation.* Nucleic Acids Res, 2018. **46**(12): p. 6087-6098.

27. Kustatscher, G., et al., *Splicing regulates NAD metabolite binding to histone macroH2A.* Nat Struct Mol Biol, 2005. **12**(7): p. 624-5.

28. Smeenk, G., et al., *Poly(ADP-ribosyl)ation links the chromatin remodeler SMARCA5/SNF2H to RNF168-dependent DNA damage signaling.* J Cell Sci, 2013. **126**(Pt 4): p. 889-903.

29. Wang, Z., et al., *Recognition of the iso-ADP-ribose moiety in poly(ADP-ribose) by WWE domains suggests a general mechanism for poly(ADP-ribosyl)ation-dependent ubiquitination.* Genes Dev, 2012. **26**(3): p. 235-40.

30. Torrungruang, K., et al., *DNA binding and gene activation properties of the Nmp4 nuclear matrix transcription factors.* J Biol Chem, 2002. **277**(18): p. 16153-9.

31. Krasner, D.S., et al., *Interplay between Ku and Replication Protein A in the Restriction of Exo1-mediated DNA Break End Resection.* J Biol Chem, 2015. **290**(30): p. 18806-16.

32. Grundy, G.J., et al., *APLF promotes the assembly and activity of non-homologous end joining protein complexes.* Embo j, 2013. **32**(1): p. 112-25.

33. Shirodkar, P., et al., *Identification and functional characterization of a Ku-binding motif in aprataxin polynucleotide kinase/phosphatase-like factor (APLF).* J Biol Chem, 2013. **288**(27): p. 19604-13.

34. Rulten, S.L., et al., *PARP-3 and APLF function together to accelerate nonhomologous end-joining.* Mol Cell, 2011. **41**(1): p. 33-45.

35. Nemoz, C., et al., *XLF and APLF bind Ku80 at two remote sites to ensure DNA repair by non-homologous end joining.* Nat Struct Mol Biol, 2018. **25**(10): p. 971-980.

36. Tang, J., et al., *Acetylation limits 53BP1 association with damaged chromatin to promote homologous recombination.* Nat Struct Mol Biol, 2013. **20**(3): p. 317-25.

37. Lu, H., et al., *DNA-PKcs promotes chromatin decondensation to facilitate initiation of the DNA damage response.* Nucleic Acids Res, 2019. **47**(18): p. 9467-9479.

38. Blackford, A.N. and S.P. Jackson, *ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response.* Mol Cell, 2017. **66**(6): p. 801-817.

3

39. Schimmel, J., et al., *Mutational signatures of non-homologous and polymerase theta-mediated end-joining in embryonic stem cells.* Embo j, 2017. **36**(24): p. 3634-3649.

40. Drouet, J., et al., *DNA-dependent protein kinase and XRCC4-DNA ligase IV mobilization in the cell in response to DNA double strand breaks.* J Biol Chem, 2005. **280**(8): p. 7060-9.

41. Brown, E.J. and D. Baltimore, *Essential and dispensable roles of ATR in cell cycle arrest and genome maintenance.* Genes Dev, 2003. **17**(5): p. 615-28.

42. Bennardo, N., et al., *Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair.* PLoS Genet, 2008. **4**(6): p. e1000110.

43. Galanty, Y., et al., *Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks.* Nature, 2009. **462**(7275): p. 935-9.

44. McCabe, N., et al., *Deficiency in the repair of DNA damage by homologous recombination and sensitivity to poly(ADP-ribose) polymerase inhibition.* Cancer Res, 2006. **66**(16): p. 8109-15.

45. Ronato, D.A., et al., *Limiting the DNA Double-Strand Break Resectosome for Genome Protection.* Trends Biochem Sci, 2020. **45**(9): p. 779-793.

46. Lee, K.J., et al., *Phosphorylation of Ku dictates DNA double-strand break (DSB) repair pathway choice in S phase.* Nucleic Acids Res, 2016. **44**(4): p. 1732-45.

47. Shao, Z., et al., *Persistently bound Ku at DNA ends attenuates DNA end resection and homologous recombination.* DNA Repair, 2012. **11**(3): p. 310-316.

48. Kabotyanski, E.B., et al., *Double-strand break repair in Ku86- and XRCC4-deficient cells.* Nucleic Acids Res, 1998. **26**(23): p. 5333-42.

49. Taty-Taty, G.C., et al., *Control of alternative end joining by the chromatin remodeler p400 ATPase.* Nucleic Acids Res, 2016. **44**(4): p. 1657-68.

50. Rother, M.B., et al., *CHD7 and 53BP1 regulate distinct pathways for the re-ligation of DNA double-strand breaks.* Nature Communications, 2020. **11**(1): p. 5775.

51. Lansdorp, P.M., et al., *Heterogeneity in Telomere Length of Human Chromosomes.* Human Molecular Genetics, 1996. **5**(5): p. 685-691.

52. Jaco, I., P. Muñoz, and M.A. Blasco, *Role of human Ku86 in telomere length maintenance and telomere capping.* Cancer Res, 2004. **64**(20): p. 7271-8.

53. Caron, M.C., et al., *Poly(ADP-ribose) polymerase-1 antagonizes DNA resection at double-strand breaks.* Nat Commun, 2019. **10**(1): p. 2954.

54. Nakamoto, T., et al., *CIZ, a zinc finger protein that interacts with p130(cas) and activates the expression of matrix metalloproteinases.* Mol Cell Biol, 2000. **20**(5): p. 1649-58.

55. Dechering, K.J., et al., *Distinct frequency-distributions of homopolymeric DNA tracts in different genomes.* Nucleic Acids Research, 1998. **26**(17): p. 4056-4062.

56. Costantini, S., et al., *Interaction of the Ku heterodimer with the DNA ligase IV/Xrcc4 complex and its regulation by DNA-PK.* DNA Repair (Amst), 2007. **6**(6): p. 712-22.

57. Zhao, B., et al., *The essential elements for the noncovalent association of two DNA ends during NHEJ synapsis.* Nat Commun, 2019. **10**(1): p. 3588.

58. Jette, N. and S.P. Lees-Miller, *The DNA-dependent protein kinase: A multifunctional protein kinase with roles in DNA double strand break repair and mitosis.* Progress in Biophysics and Molecular Biology, 2015. **117**(2): p. 194-205.

59. Douglas, P., et al., *Polo-like kinase 1 (PLK1) and protein phosphatase 6 (PP6) regulate DNA-dependent protein kinase catalytic subunit (DNA-PKcs) phosphorylation in mitosis.* Biosci Rep, 2014. **34**(3).

60. Helfricht, A., et al., *Loss of ZBTB24 impairs non-homologous end-joining and class-switch recombination in patients with ICF syndrome.* Journal of Experimental Medicine, 2020.

61. Britton, S., J. Coates, and S.P. Jackson, *A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair.* J Cell Biol, 2013. **202**(3): p. 579-95.

62. Pierce, A.J., et al., *XRCC3 promotes homology-directed repair of DNA damage in mammalian cells.* Genes Dev, 1999. **13**(20): p. 2633-8.

63. Iacovoni, J.S., et al., *High-resolution profiling of gammaH2AX around DNA double strand breaks in the mammalian genome.* The EMBO journal, 2010. **29**(8): p. 1446-1457.

64. Ronson, G.E., et al., *PARP1 and PARP2 stabilise replication forks at base excision repair intermediates through Fbh1-dependent Rad51 regulation.* Nature Communications, 2018. **9**(1): p. 746.

65. Gottschalk, A.J., et al., *Poly(ADP-ribosyl)ation directs recruitment and activation of an ATP-dependent chromatin remodeler.* Proc Natl Acad Sci U S A, 2009. **106**(33): p. 13770-4.

66. Mehrotra, P.V., et al., *DNA repair factor APLF is a histone chaperone.* Mol Cell, 2011. **41**(1): p. 46-55.

67. Wachsmuth, M., et al., *High-throughput fluorescence correlation spectroscopy enables analysis of proteome dynamics in living cells.* Nat Biotechnol, 2015. **33**(4): p. 384-9.

68. Sprague, B.L., et al., *Analysis of binding reactions by fluorescence recovery after photobleaching.* Biophys J, 2004. **86**(6): p. 3473-95.

69. Czarna, A., et al., *Structures of Drosophila cryptochrome and mouse cryptochrome1 provide insight into circadian function.* Cell, 2013. **153**(6): p. 1394-405.

70. Kumar, R., et al., *The STUbL RNF4 regulates protein group SUMOylation by targeting the SUMO conjugation machinery.* Nat Commun, 2017. **8**(1): p. 1809.

71. Tyanova, S., T. Temu, and J. Cox, *The MaxQuant computational platform for mass spectrometry-based shotgun proteomics.* Nat Protoc, 2016. **11**(12): p. 2301-2319.

72. Feng, W., et al., *TRAIP regulates replication fork recovery and progression via PCNA.* Cell Discov, 2016. **2**: p. 16016.

73. Perez-Riverol, Y., et al., *The PRIDE database and related tools and resources in 2019: improving support for quantification data.* Nucleic Acids Res, 2019. **47**(D1): p. D442-d450.

74. Singh J.K., R.S., M.B. Rother, A.J.L. de Groot, W.W. Wiegant, K. Vreeken, O. D'Augustin, R.Q. Kim, P.M. Krawcyk, R. González-Prieto, A.C.O. Vertegaal, M. Lamers, S. Huet and H. van Attikum,, *Zinc finger protein ZNF384 is an adaptor of KU to DNA during classical non-homologous end-joining, Singh-image-processing.* 2021.
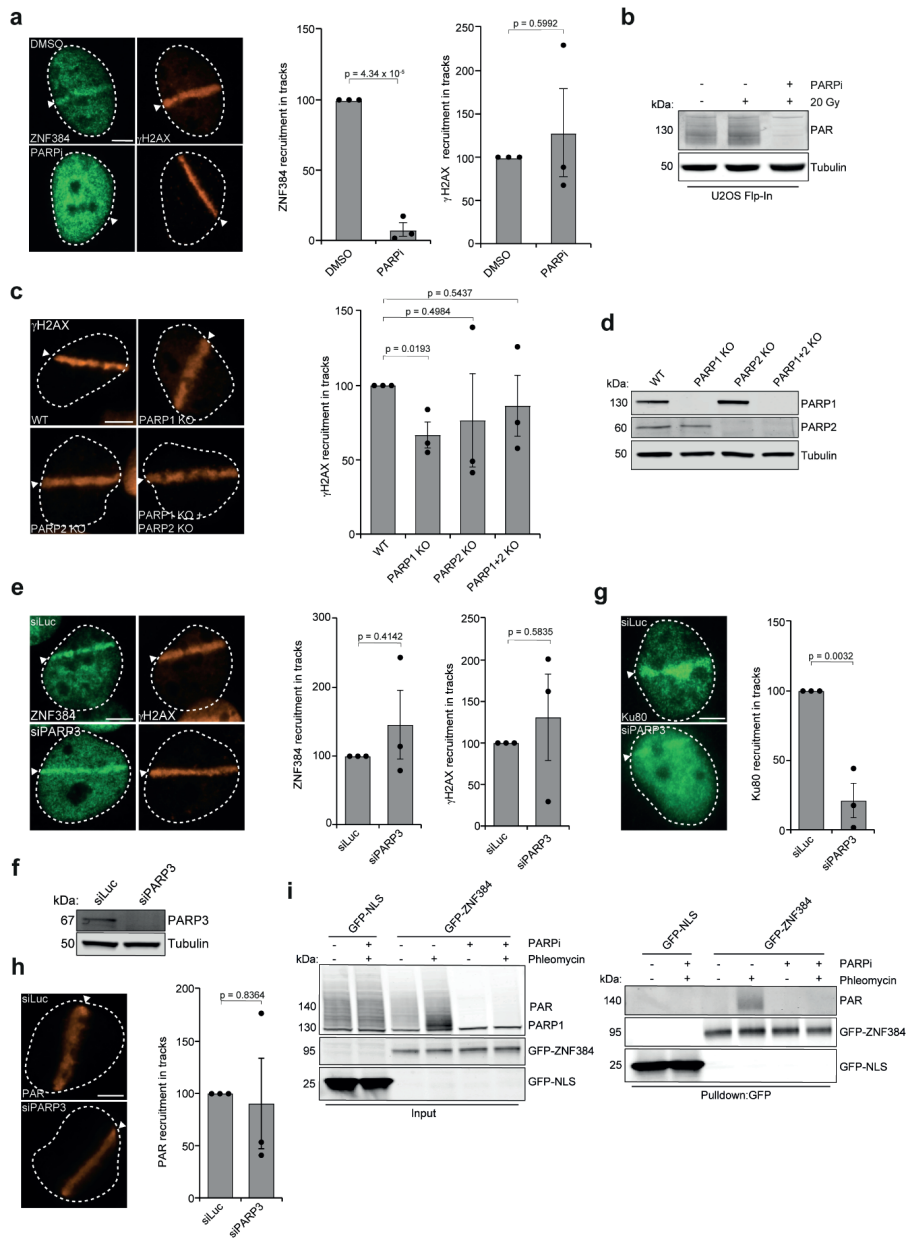
SUPPLEMENTARY FIGURES, LEGENDS AND TABLES
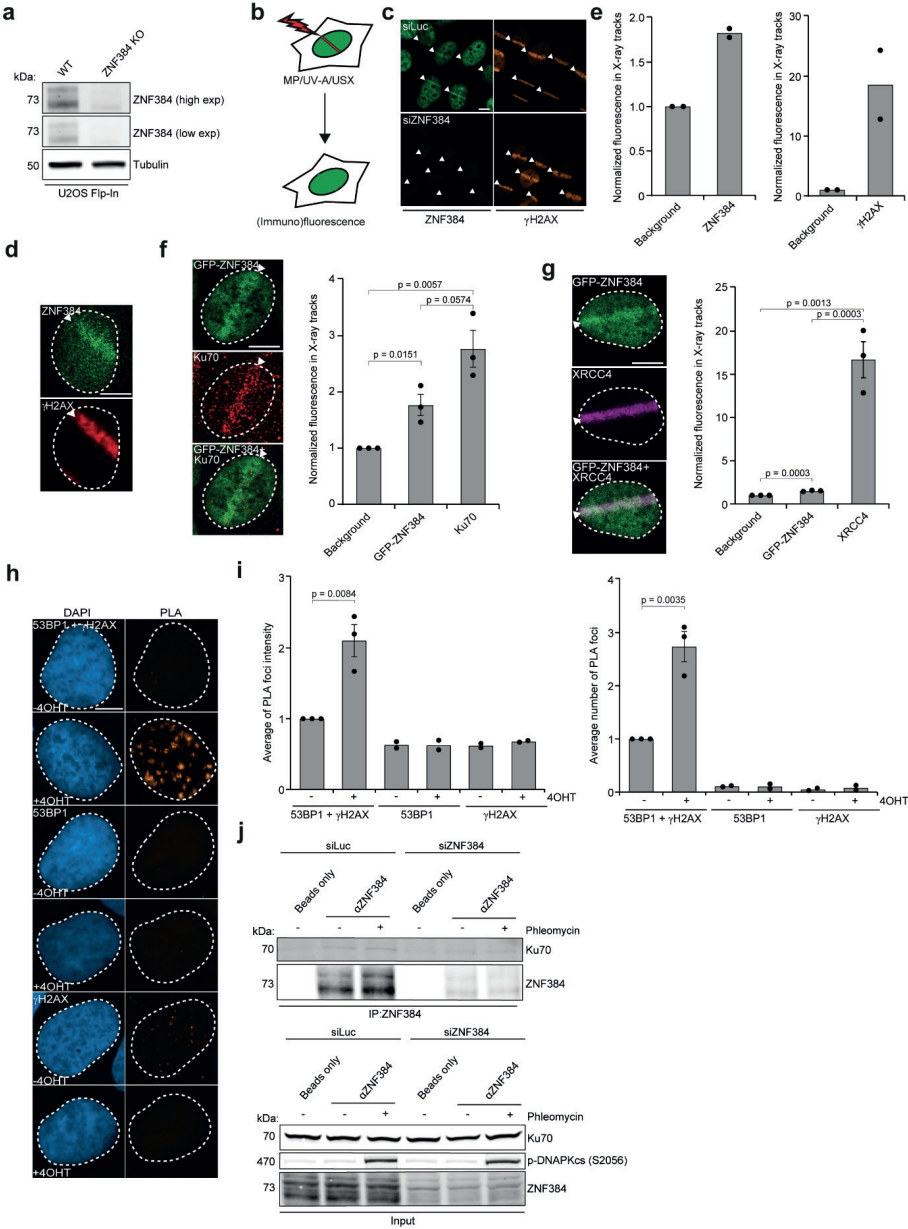
3

▶ **Supplementary Figure. 1 – related to Figure. 1. ZNF384 is recruited to DNA damage sites.**
(A) Western blot analysis of ZNF384 expression in cells from Figure 1A. Tubulin is a loading control. Data shown represent 2 independent experiments. (B) Schematic representation of the laser micro-irradiation approach to study protein accumulation at sites of DNA damage. An 800 nm multiphoton (MP) laser, 365 nm UV-A laser or Ultra-soft X-ray (USX) irradiation was used. (C) Representative images of two independent experiments of the accumulation of ZNF384 at γH2AX-marked 365 nm UV-A tracks in cells transfected with the indicated siRNAs. Cells were fixed and immunostained 10 minutes after laser micro-irradiation. White triangles indicate irradiated regions. (D) Accumulation of endogenous ZNF384 at γH2AX-marked DNA damage inflicted by USX irradiation in U2OS Flp-In/T-Rex cells expressing doxycycline (dox)-inducible GFP-ZNF384. Cells were fixed and immunostained 5 minutes after USX irradiation. White triangles indicate irradiated regions. (E) Quantification of ZNF384 and γH2AX accumulation in cells from D. The mean of at least 40 cells acquired in 2 independent experiments is shown. Data were normalized to the nuclear background outside the irradiated area, which was set to 1. (F) Accumulation of GFP-ZNF384 at sites of KU70-marked DNA damage inflicted by USX irradiation in U2OS Flp-In/T-Rex cells expressing doxycycline (dox)-inducible GFP-ZNF384. Cells were fixed and immunostained 5 minutes after laser micro-irradiation. White triangles indicate irradiated regions. (left panel). Quantification of ZNF384 and KU70 is presented as the ±SEM of >45 cells acquired in 3 independent experiments. Data were normalized to the nuclear background outside the irradiated area, which was set to 1 (right panel). P-values were calculated using the two-tailed unpaired Student's *t* test. (G) As in F, but for GFP-ZNF384 and XRCC4. Quantification of ZNF384 and XRCC4 is presented as the ±SEM of 60 cells acquired in 3 independent experiments. P-values were calculated using the two-tailed unpaired Student's *t* test, assuming unequal variances. (H) PLA of 53BP1 and ZNF384 and γH2AX in *Asi*SI–ER-U2OS cells treated with 4-OHT for DSB induction. PLA foci were scored after 5 hours of DSB induction. (I) Quantification of H. PLA foci formation and foci intensity from >200 cells acquired in 2-3 independent experiments are shown. Statistical significance was calculated with the two-tailed Student's *t* test, assuming unequal variances. (J) Immunoprecipitation (IP) of endogenous ZNF384 from 500μM Phleomycin treated U2OS cells transfected with the indicated siRNAs. Control IP contained beads only. Blots were probed for ZNF384 and KU70. Data shown represent 2 independent experiments. Scale bar 5 μm, except for C in which scale bar is 10 μm. Source data are provided as Source Data file.

3
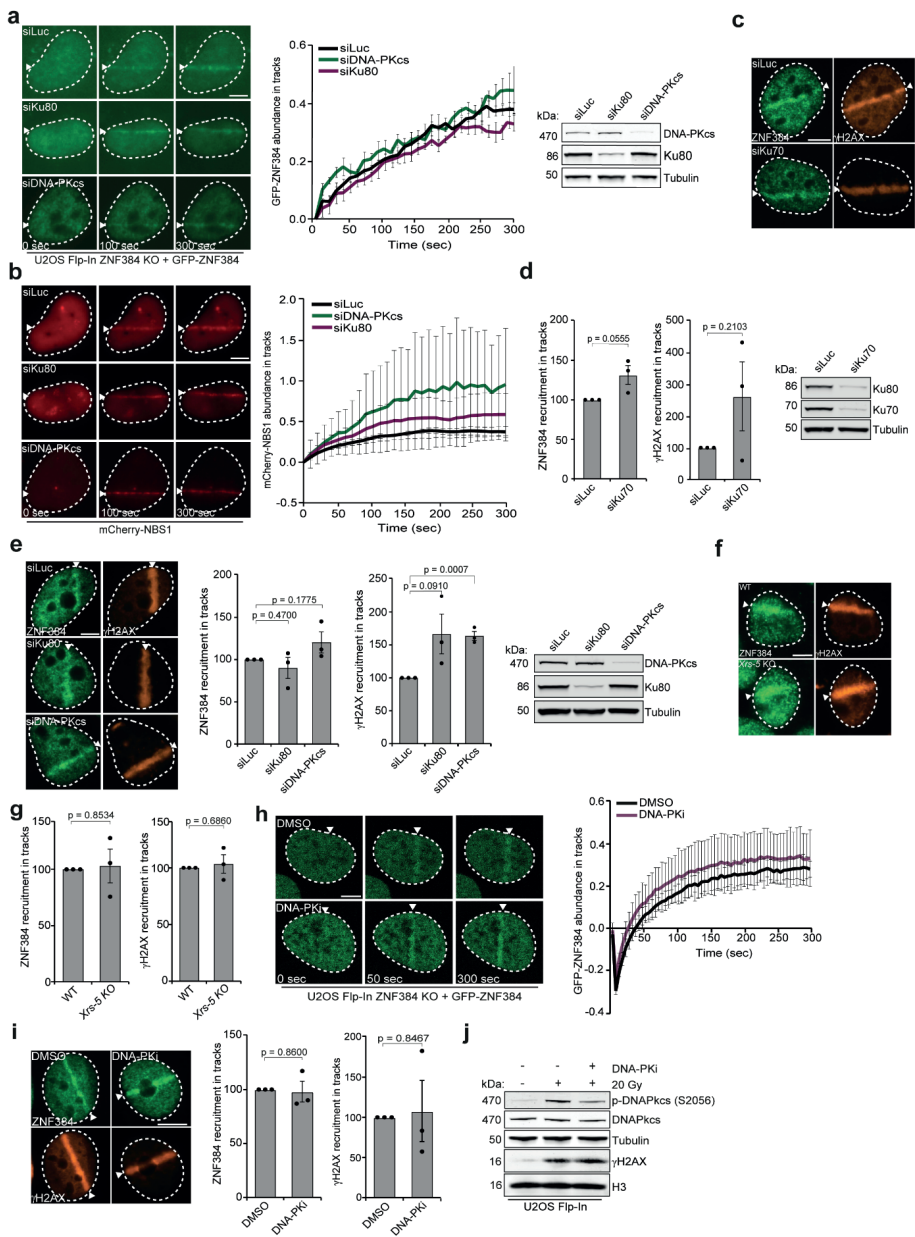
▶ **Supplementary Figure. 2 – related to Figure. 1. ZNF384 is recruited to DNA damage sites independent of KU70/KU80 and DNA-PKcs.** (A) Recruitment of GFP-ZNF384 to 365 nm UV-A tracks in BrdU-sensitized U2OS Flp-In/T-Rex ZNF384 KO cells. White triangles indicate irradiated regions (left panel). Quantification of the data is plotted on a time scale as relative abundance in tracks. The graph represents the mean ±SEM of >40 cells acquired in 3 independent experiments (middle panel). Western blot analysis of KU80 and DNA-Pkcs expression. Tubulin is a loading control. Data shown represent 2 independent experiments (right panel). (B) As in A, except for mCherry-NBS1. (C) Accumulation of endogenous ZNF384 at 365 nm UV-A tracks in BrdU-sensitized U2OS cells following transfection with siLuc and siKU70. Cells were fixed and immunostained 10 minutes after laser micro-irradiation. White triangles indicate irradiated regions. γH2AX was used as a DNA damage marker. (D) Quantification of endogenous ZNF384 and γH2AX levels in laser tracks in cells from D is presented as the mean ±SEM of >180 cells acquired in 3 independent experiments (left panel). Western blot analysis of KU70 and K80 expression in cells. Tubulin is a loading control. Data shown represent 2 independent experiments (right panel). (E) Accumulation of endogenous ZNF384 at 365 nm UV-A tracks in BrdU-sensitized U2OS cells following transfection with siLuc, siKU80 and siDNA-PKcs. Cells were fixed and immunostained 10 minutes after laser micro-irradiation. White triangles indicate irradiated regions. γH2AX was used as a DNA damage marker (left panel). Quantification of endogenous ZNF384 and γH2AX levels in laser tracks is presented as the mean ±SEM of >180 cells acquired in 3 independent experiments (middle panel). Western blot analysis of KU80 and DNA-Pkcs expression. Tubulin is a loading control. Data shown represent 2 independent experiments (right panel). (F) Accumulation of endogenous ZNF384 at 365 nm UV-A tracks in BrdU-sensitized wildtype (WT) and xrs-5 KO hamster cells. Cells were fixed and immunostained 10 minutes after laser micro-irradiation. White triangles indicate irradiated regions. γH2AX was used as a DNA damage marker. (G) Quantification of F. Endogenous ZNF384 and γH2AX levels in laser tracks are presented as the mean ±SEM of >150 cells acquired in 3 independent experiments. (H) Kinetics of GFP-ZNF384 recruitment to 800 nm laser tracks in U2OS Flp-In/T-Rex ZNF384 KO cells treated for 1 hour with DNA-PKcs inhibitor prior to laser micro-irradiation. White triangles indicate irradiated regions (left panel). Quantification of GFP-ZNF384 in laser tracks is presented as the mean ±SD from 2 independent experiments (right panel). (I) Accumulation of endogenous ZNF384 at 365 nm UV-A tracks in BrdU-sensitized U2OS cells treated for 1 hour with DNA-PKcs inhibitor prior to laser micro-irradiation. Cells were fixed and immunostained 10 minutes after laser micro-irradiation. White triangles indicate irradiated regions. γH2AX was used as a DNA damage marker (left panel). Quantification of endogenous ZNF384 and γH2AX levels in laser tracks is presented as the mean ±SEM of >180 cells acquired in 3 independent experiments (right panel). Statistical significance was calculated with the two-tailed Student's t test, assuming unequal variances. Scale bar 5 μm. (J) Western blot analysis of p-DNA-PKcs (S2056) levels in IR- and DNA-PKcs inhibitor-treated cells used in I. Cells were treated with DNA-PKi 1 hour before IR exposure. Blots were probed for p-DNA-PKcs (S2056), DNA-PKcs, γH2AX. Tubulin and H3 are loading controls. Source data are provided as Source Data file.

3

▶ **Supplementary Figure. 3 – related to Figure. 1. ZNF384 is recruited to DNA damage sites via PARP1/PAR.** (A) Accumulation of endogenous ZNF384 at 365 nm UV-A tracks in BrdU-sensitized U2OS cells treated for 1 hour with PARPi prior to laser micro-irradiation. Cells were fixed and immunostained 10 minutes after laser micro-irradiation. White triangles indicate irradiated regions. γH2AX was used as a damage marker (left panel). Quantification of endogenous ZNF384 and γH2AX levels in laser tracks is presented as the mean ±SEM of >200 cells acquired in 3 independent experiments (right panel). (B) Western blot analysis of PAR levels in IR treated and PARPi treated cells used in Figure 2A and Figure S3A. Cells were treated with PARPi 1hr before the IR treatment. Tubulin is a loading control. Data shown represent 2 independent experiments. (C) γH2AX accumulation at 365 nm UV-A tracks 10 minutes after DNA damage induction in BrdU-sensitized wildtype, PARP1 KO, PARP2 KO and PARP1/PARP2 KO U2OS cells. White triangles indicate irradiated regions (left panel). The mean ±SEM of >180 cells from 3 independent experiments is shown (right panel). (D) Western blot analysis of PARP1 and PARP2 expression in cells from Figure 2B and Figure S3C. Tubulin is a loading control. Data shown represent 2 independent experiments. (E) Accumulation of endogenous ZNF384 at 365 nm UV-A tracks in BrdU-sensitized U2OS cells transfected with the indicated siRNAs. Cells were fixed and immunostained 10 minutes after laser micro-irradiation. White triangles indicate irradiated regions. γH2AX was used as a damage marker (left panel). Quantification of endogenous ZNF384 and γH2AX levels in laser tracks is presented as the mean ±SEM of > 200 cells acquired in 3 independent experiments (right panel). (F) Western blot analysis of PARP3 expression in cells from Figure S3E. Tubulin is a loading control. Data shown represent 2 independent experiments. (G) Accumulation of endogenous KU80 at 365 nm UV-A tracks in BrdU-sensitized U2OS cells transfected with the indicated siRNAs. Cells were fixed and immunostained 10 minutes after laser micro-irradiation. White triangles indicate irradiated regions (left panel). Quantification of endogenous KU80 levels in laser tracks is presented as the mean ±SEM of >120 cells acquired in 3 independent experiments (right panel). (H) As in G, except for PAR. (I) Pull-downs of GFP-ZNF384 and GFP-NLS under denaturing conditions in untreated and phleomycin (Phleo)-treated U2OS Flp-In/T-Rex cells expressing doxycycline (dox)-inducible GFP-ZNF384 or GFP-NLS. Cells were also treated with PARPi (Olaparib) for 1 hour prior to the phleomycin treatment. Blots were probed for PAR, PARP1 and GFP. Data shown represent 2 independent experiments. Statistical significance was calculated with the two-tailed Student's *t* test, assuming unequal variances. Scale bar 5 μm. Source data are provided as Source Data file.
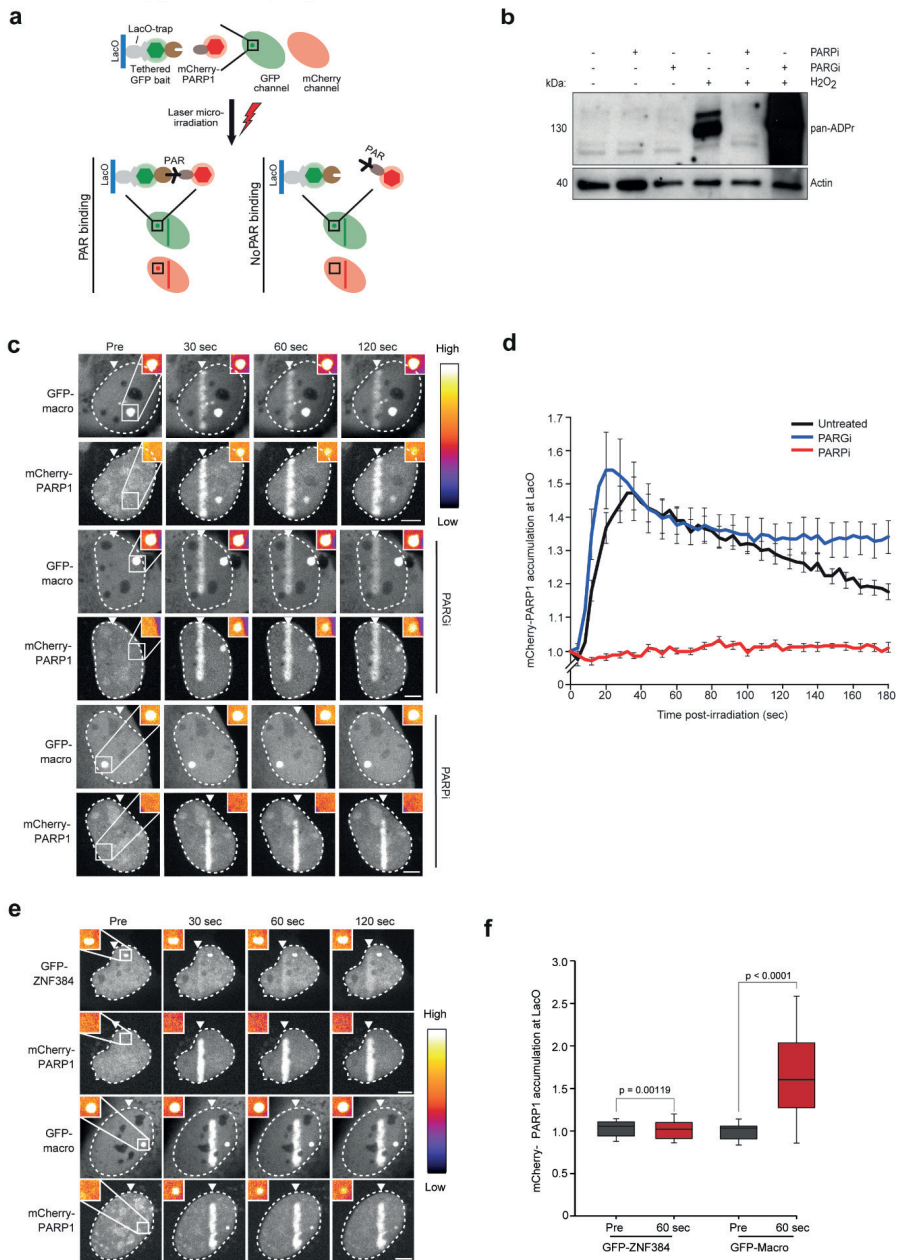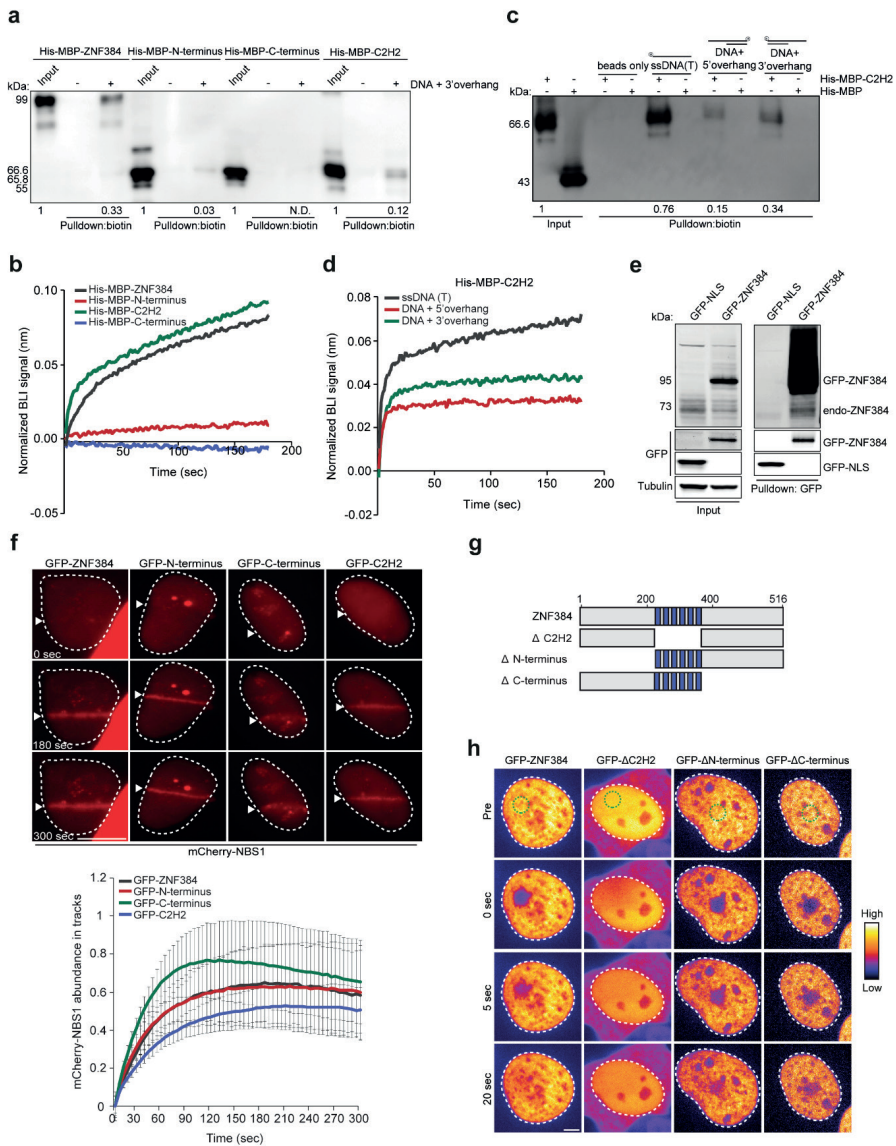
▶ **Supplementary Figure. 4 – related to Figure. 2. ZNF384 is recruited to DNA damage sites via PARP1/PAR dependent chromatin unfolding.** (A) Schematic of the fluorescence three-hybrid assay (F3H). Prior to laser irradiation, a GFP-tagged bait protein which is tethered to a genomically integrated LacO array through a LacI-GFP binding protein (LacO-trap) does not show enrichment of mCherry-PARP1 at the LacO array. After laser micro-irradiation, mCherry-PARP1 diffuses away from the site of damage in a PARylated state. If the tethered GFP-tagged bait protein can interact with PARylated PARP1 (bottom left), mCherry-PARP1 will enrich at the LacO array. If the tethered protein of interest does not interact with the tethered bait protein (bottom right), there is no enrichment of mCherry-PARP1 at the LacO array. (B) Western blot analysis of PAR levels in H_2O_2 and PARPi or PARGi treated cells used in Figure S4D. Cells were treated with 2 mM H_2O_2 for 10 min to induce PARylation. Blot was probed with panADPr. Actin is a loading control. Blots are representative of 3 independent replicates. (C) Fluorescence images of *in situ* PAR binding three-hybrid assays. LacI-GFP-macroH2A1.1 (GFP-macro) was tethered to a LacO array in Hoechst-sensitized cells co-expressing mCherry-PARP1 treated for 1 hour with PARPi or PARGi prior to 405 nm laser micro-irradiation. White triangles indicate irradiated regions. Inset shows the magnified LacO array. Insets are pseudocolored according to the look-up table displayed on the right. Scale bar 5 μm. (D) The mean intensity of mCherry signal at the LacO array from 19-20 cells was quantified pre damage and 60 seconds post damage. Data shows mean intensity of the mCherry signal at the LacO array ±SEM. Cells were treated for 1 hour with PARPi or PARGi prior to micro-irradiation. (E) Fluorescence images of *in situ* PAR binding three-hybrid assays. GFP-ZNF384 or GFP-macro was tethered to a LacO array in Hoechst-sensitized cells co-expressing mCherry-PARP1 prior to 405 nm laser micro-irradiation. White triangles indicate irradiated regions. Inset shows the magnified LacO array. Insets are pseudocolored according to the look-up table displayed on the right. (F) The mean intensity of mCherry signal at the LacO array from 12-15 cells was quantified pre damage and 60 seconds post damage. The boxplot limits correspond to the 25th and 75th percentiles and the center line in the box indicates the median value. The whiskers extend 1.5 times the interquartile range. P-values were calculated using the two-tailed Student's *t* test, assuming unequal variances. Scale bar 5 μm. Source data are provided as Source Data file.

3
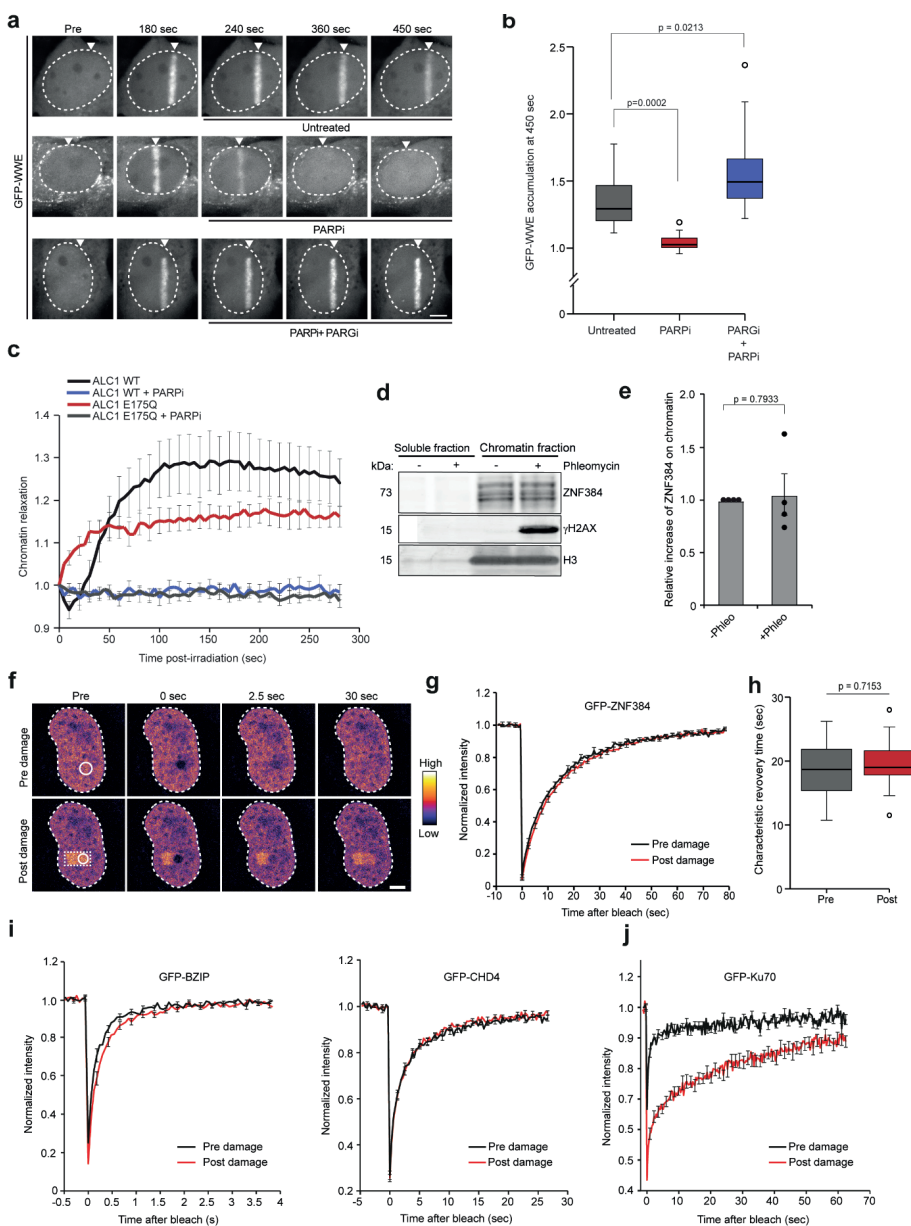
▶ **Supplementary Figure. 5 – related to Figure. 2. ZNF384 is recruited to DNA damage sites via PARP1/PAR dependent chromatin unfolding.** (A) Confocal images showing accumulation of GFP-WWE at sites of 405 nm laser micro-irradiation in Hoechst-sensitized U2OS cells. Cells were left untreated or treated with PARPi or PARPi and PARGi 240 seconds after DNA damage induction. White triangles indicate irradiated regions. Scale bar 5 μm. (B) Boxplot showing the accumulation of GFP-WWE at 450 seconds post irradiation from cells in A. Boxplots show the first, median and third quartiles from a representative of 3 independent replicates. The whiskers extend 1.5 times the interquartile range. P-values were calculated using the two-tailed unpaired Student's t test, assuming unequal variances. (C) Chromatin relaxation was measured for each of the overexpression conditions from Figure 2G. The mean ±SEM from 13-16 cells from a representative of 3 independent experiments is shown. (D) Western blot analysis of the indicated proteins in soluble and chromatin fractions from untreated and phleomycin treated U2OS cells. H3 is the loading control. γH2AX is a DNA damage marker. A representative experiment is shown. (E) Quantification of ZNF384 levels on chromatin in cells from D. H3 is a loading control, which was used for normalization of ZNF384 levels on chromatin. The mean ±SEM from 4 independent experiments is shown. P-values were calculated using the Student's t test, assuming unequal variances. (F) Representative images of Hoechst-sensitized U2OS Flp-In/T-Rex *ZNF384* KO cells in which FRAP measurements were performed to assess the local turnover of GFP-ZNF384 at DNA damage sites. DNA damage was induced in the region indicated with a dashed line using 405 laser micro-irradiation. Subsequent FRAP was induced in a subarea within the damaged region, as indicated with an unbroken line. Images are pseudocolored according to the look-up table displayed on the right. Scale bar 4 μm. (G) Normalized FRAP curves from F. Data show mean ±SEM from 14 cells per condition from a representative of 3 independent replicates (H) Boxplot showing the characteristic recovery times estimated from the FRAP curves shown in F. The boxplot limits correspond to the 25th and 75th percentiles and the center line in the box indicates the median value. The whiskers extend 1.5 times the interquartile range. P-values were calculated using the two-tailed paired Student's t test. (I) Normalized FRAP curves of the GFP-tagged DNA binding domain BZIP from the transcription factor C-EBPa and DNA damage induction in U2OS cells (left panel). Normalized FRAP curves of GFP-tagged CHD4 and DNA damage induction in U2OS cells (right panel). (J) Normalized FRAP curves of GFP-KU70 and DNA damage induction in RPE1-hTERT cells. The mean ±SD is shown from 14 cells per condition from a representative of 2 independent replicates. Source data are provided as Source Data file.
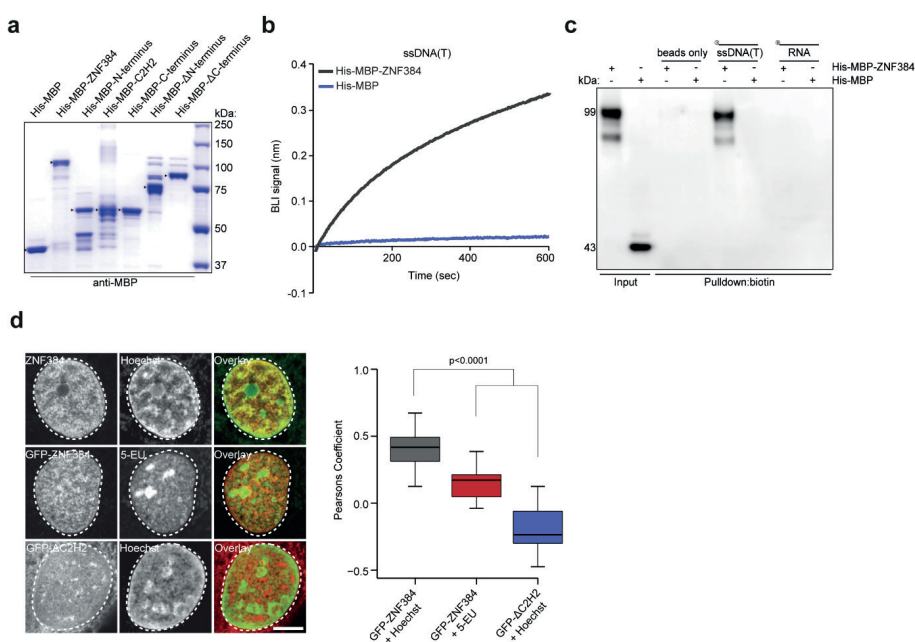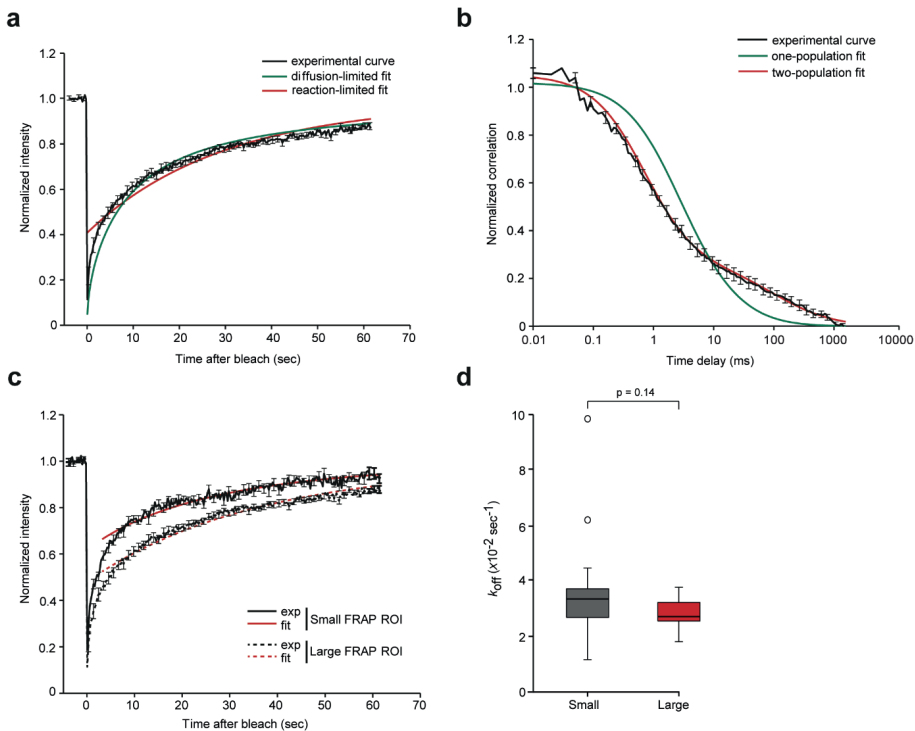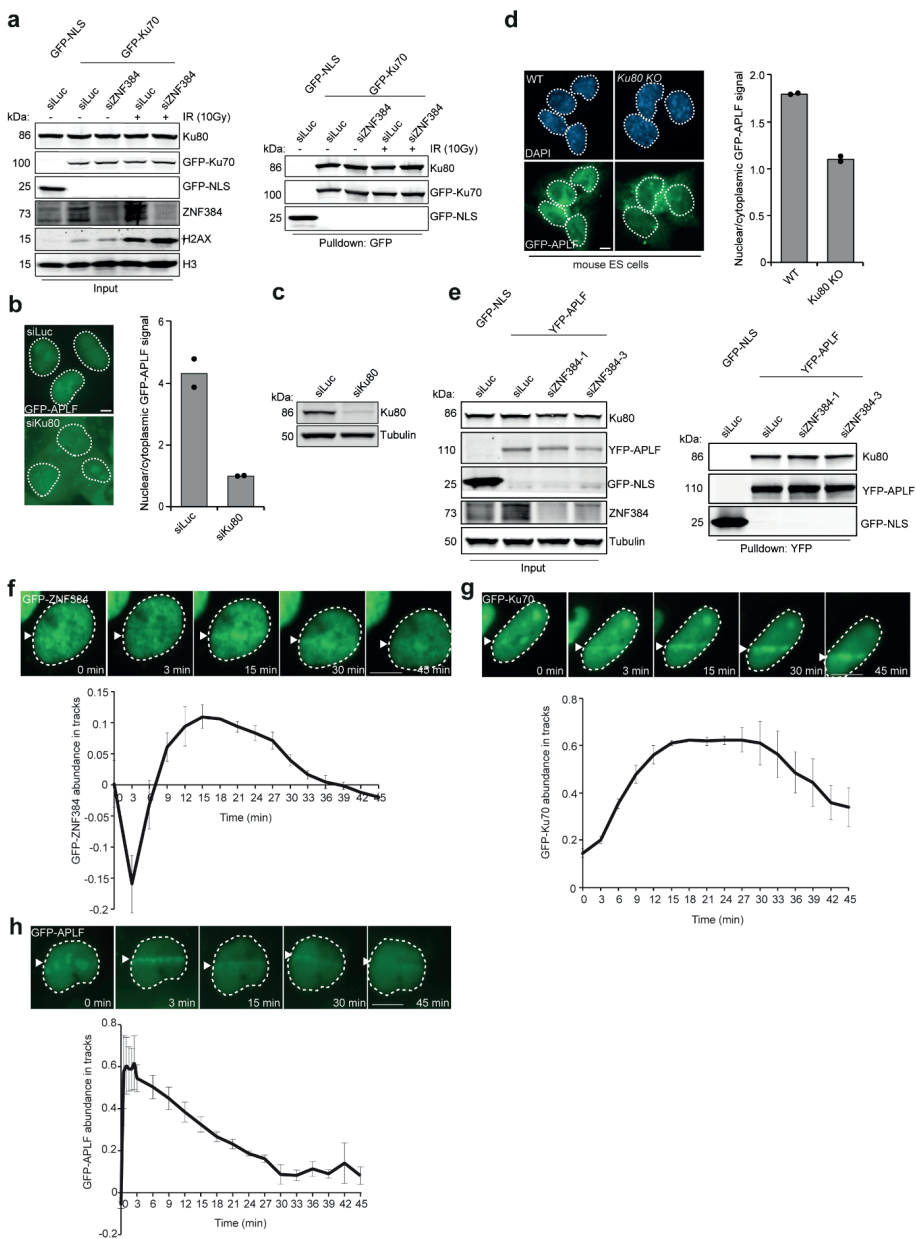
Supplementary Figure. 6 – related to Figure. 3. ZNF384 is recruited to sites of DNA damage in manner dependent on its ability to bind DNA. (A) SDS-PAGE analysis of the indicated purified His-MBP-tagged proteins. Data shown represent 2 independent experiments. (B) DNA binding by His-MBP-ZNF384 and His-MBP to biotinylated T-rich ssDNA (ssDNA(T)) by BLI. (C) Pulldowns of the indicated biotinylated DNA or RNA substrates in the presence of His-MBP-ZNF384 or His-MBP. Blots were probed for MBP. Data shown represent 2 independent experiments. (D) Representative images of U2OS Flp-In/T-Rex *ZNF384* KO cells expressing either GFP-tagged full-length ZNF384 or ΔC2H2. Cells were co-stained with Hoechst or 5-Ethynyl Uridine (5-EU) to assess their colocalization with DNA or RNA (left panel). The levels of colocalization between ZNF384 or ΔC2H2 and DNA or RNA were assessed using the Pearson coefficient (right panel). The mean ±SD is shown from >23 cells per condition. The boxplot limits correspond to the 25th and 75th percentiles and the center line in the box indicates the median value. The whiskers extend 1.5 times the interquartile range. For GFP-ZNF384 + Hoechst vs GFP-ZNF384 + 5-EU, p= 7.1156×10^{-7}. For GFP-ZNF384 + Hoechst vs GFP-ΔC2H2 + Hoechst, p= 4.027×10^{-12}. P-values were calculated using the two-tailed unpaired Student's *t* test, assuming unequal variances. Source data are provided as Source Data file.

a

His-MBP-ZNF384 His-MBP-N-terminus His-MBP-C-terminus His-MBP-C2H2

DNA + 3'overhang

kDa:
99
66.6
65.8
55

1   0.33   1   0.03   1   N.D.   1   0.12

Pulldown:biotin Pulldown:biotin Pulldown:biotin Pulldown:biotin

c

beads only ssDNA(T) DNA+ 5'overhang DNA+ 3'overhang

His-MBP-C2H2
His-MBP

kDa:
66.6
43

1   0.76   0.15   0.34

Input   Pulldown:biotin

b

Normalized BLI signal (nm)

His-MBP-ZNF384
His-MBP-N-terminus
His-MBP-C2H2
His-MBP-C-terminus

Time (sec)

d

His-MBP-C2H2

Normalized BLI signal (nm)

ssDNA (T)
DNA + 5'overhang
DNA + 3'overhang

Time (sec)

e

kDa:
GFP-NLS GFP-ZNF384 GFP-NLS GFP-ZNF384

95
73

GFP-ZNF384
endo-ZNF384

GFP

GFP-ZNF384
GFP-NLS

Tubulin

Input   Pulldown: GFP

f

GFP-ZNF384 GFP-N-terminus GFP-C-terminus GFP-C2H2

0 sec

180 sec

300 sec

mCherry-NBS1

mCherry-NBS1 abundance in tracks

GFP-ZNF384
GFP-N-terminus
GFP-C-terminus
GFP-C2H2

Time (sec)

g

1   200   400   516

ZNF384
Δ C2H2
Δ N-terminus
Δ C-terminus

h

GFP-ZNF384 GFP-ΔC2H2 GFP-ΔN-terminus GFP-ΔC-terminus

Pre

0 sec

5 sec

20 sec

High
Low

▶ **Supplementary Figure. 7 – related to Figure. 3. ZNF384 is recruited to sites of DNA damage via its C2H2 domain.** (A) DNA pulldowns of biotinylated DNA with a 3'overhang in the presence of the indicated ZNF384 proteins. Blots were probed for MBP. For each protein, input signals were used for normalization of the pulldown signals. The mean of 2 independent experiments is indicated below the blot. A representative experiment is shown. N.D. = not detectable. (B) DNA binding by the indicated ZNF384 proteins to DNA with a 3'overhang as measured by BLI. Quantification of the data is plotted on a time scale and normalized to His-MBP alone. (C) DNA pulldowns of biotinylated T-rich ssDNA (ssDNA(T)), biotinylated DNA with a 5'overhang or biotinylated DNA with a 3'overhang in the presence of His-MBP-C2H2 or His-MBP (control). Blots were probed for MBP. His-MBP-C2H2 pulldown signals were normalized to that in the input, which was set to 1. The mean from 3 independent experiments is indicated below the blot. His-MBP was not detectable in pulldowns. A representative experiment is shown. (D) DNA binding of His-MBP-C2H2 to biotinylated T-rich ssDNA (ssDNA(T)), biotinylated DNA with a 5'overhang or biotinylated DNA with a 3'overhang as measured by BLI. Quantification of the data is plotted on a time scale and normalized to His-MBP (control). (E) Pulldowns of the indicated GFP-fusion proteins in U2OS Flp-In/T-Rex cells. Blots were probed for GFP, ZNF384 and Tubulin. Data shown represent 2 independent experiments. (F) Live cell imaging of mCherry-NBS1 to 365 nm UV-A tracks in BrdU- sensitized *ZNF384* KO U2OS Flp-In/T-Rex cells. mCherry-NBS1, which was co-expressed with GFP-ZNF384 proteins, served as a DNA damage marker for Figure 3D. Representative images are shown. White triangles indicate irradiated regions. Scale bars 10 µm (upper panel). Quantification of the data is shown as mean ±SEM from 30-40 cells (lower panel). (G) Schematic representation of full-length ZNF384 and three deletion mutants (ΔC2H2, ΔN-terminus and ΔC-terminus). (H) Representative images of cells from Figure 3E in which the FRAP measurements are performed to assess the local dynamics of the indicated GFP-ZNF384 constructs. FRAP area is indicated with a dashed line. Images are pseudocolored according to the look-up table displayed on the bottom. Scale bar: 4 µm. Source data are provided as Source Data file.

a

b

c

d

e

f

g

h
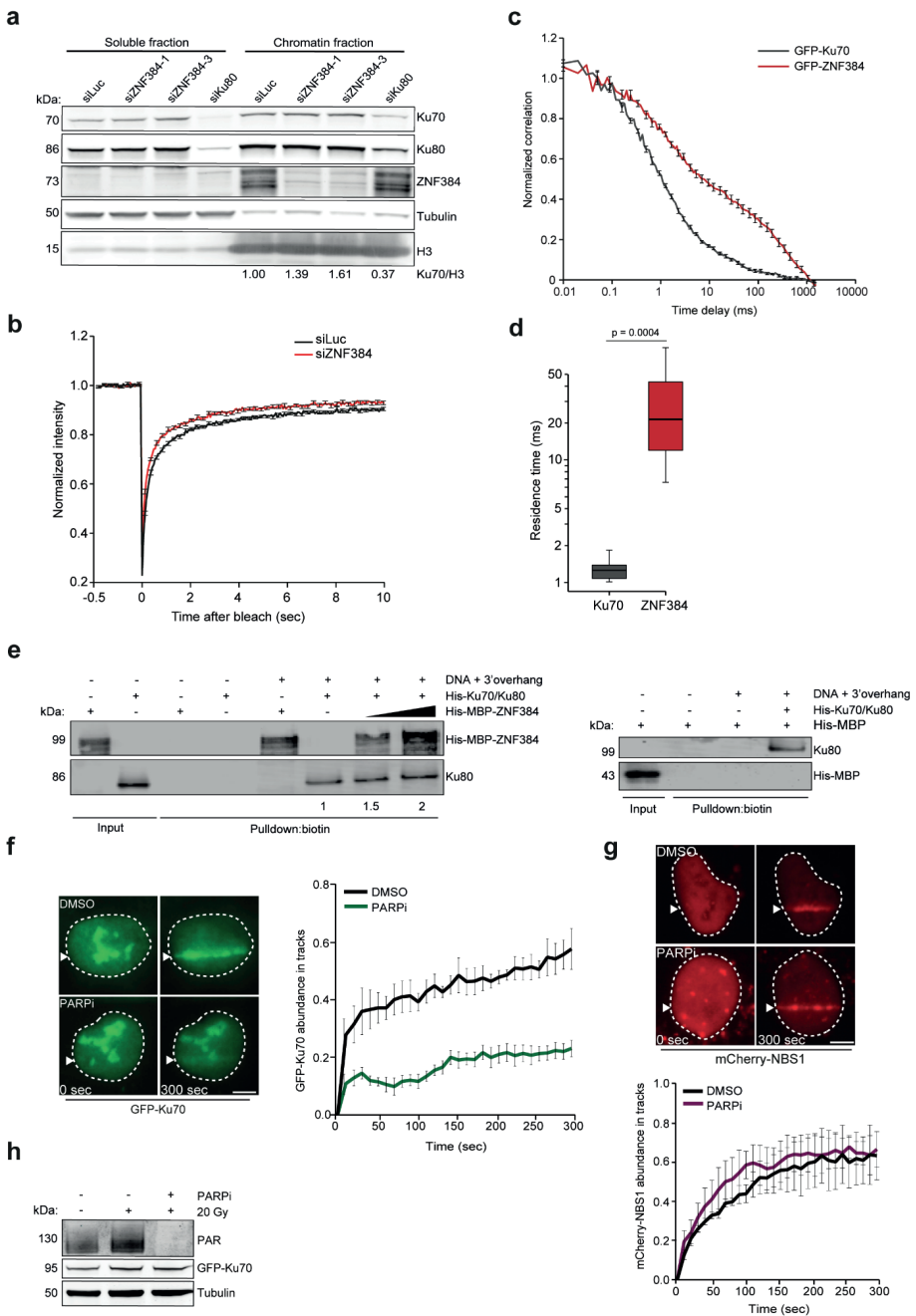
▶ **Supplementary Figure. 8 – related to Figure. 4. ZNF384 modulates the loading of KU70/ KU80 at DNA damage sites.** (A) Western blot analysis of the indicated proteins in soluble and chromatin fractions from U2OS cells. Blots were probed for the indicated proteins. H3 and Tubulin are loading controls. A representative experiment is shown. KU70 levels in chromatin were first normalized to H3 levels. KU70/H3 ratios were then normalized to that for siLuc, which was set to 1. Quantifications are shown as the mean from 2 independent experiments. (B) Normalized FRAP curves of GFP-KU70 in RPE-hTERT cells transfected with the indicated siRNAs in the absence of DNA damage. The mean ±SD is shown from 20-21 cells per condition. (C) Normalized FCS autocorrelation curves of GFP-KU70 in RPE1-hTERT cells and GFP-ZNF384 in U2OS cells in the absence of DNA damage. The mean ±SEM is shown from 10-16 cells analyzed per condition. (D) Boxplot showing the residence times of GFP-KU70 and GFP-ZNF384 in RPE1-hTERT and U2OS cells, respectively, as measured by FCS. 10-15 cells were analyzed per condition. The boxplot limits correspond to the 25th and 75th percentiles and the center line in the box indicates the median value. The whiskers extend 1.5 times the interquartile range. P-values were calculated using the two-tailed unpaired Student's *t* test, assuming unequal variances. (E) DNA pulldowns of biotinylated DNA with a 3'overhang in the presence of His-KU70/KU80, His-MBP-ZNF384 or His-MBP (control) alone or His-KU70/KU80 in combination with His-MBP-ZNF384 or His-MBP (control). Blots were probed for MBP and KU80. KU80 pulldown signals were normalized to that in the pulldown lacking His-MBP-ZNF384, which was set to 1. The mean from 4 independent experiments is indicated below the blot. His-MBP was not detectable in pulldowns. (F) GFP-KU70 recruitment to 365 nm UV-A tracks in cells treated with PARPi for 1 hour before laser micro-irradiation. White triangles indicate irradiated regions (left panel). Quantification of the data is presented as the mean ±SEM of >100 cells acquired in 3 independent experiments (right panel). (G) As in F, except for mCherry-NBS1. (H) Western blot analysis of PAR levels, which likely reflect PARP1 autoPARylation, in phleomycin- and PARPi-treated cells from F and G. Cells were treated with PARPi for 1 hour before IR treatment. Blots were probed for PAR, GFP and Tubulin. Tubulin is a loading control. Data shown represent 2 independent experiments. Statistical significance was calculated using the two-tailed unpaired Student's *t* test, assuming unequal variances. Scale bar 5 μm. Source data are provided as Source Data file.
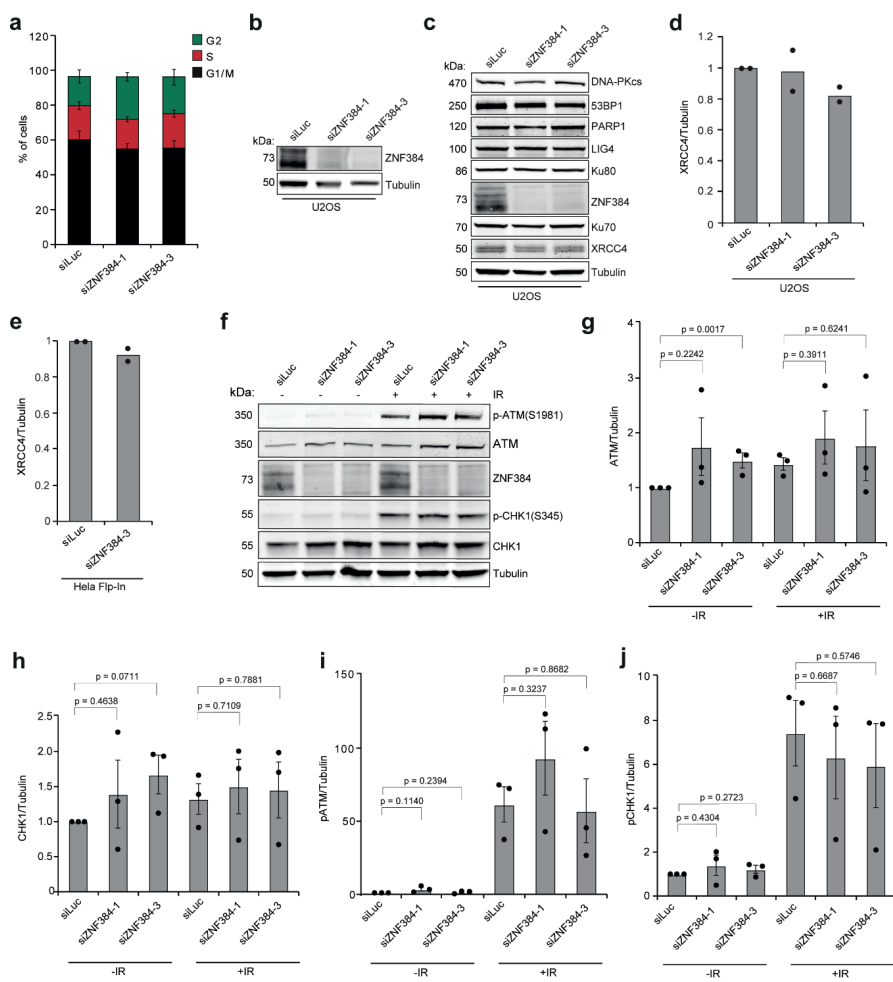
a

b

c

d

3

Supplementary Figure. 9 – related to Figure. 4. ZNF384 facilitates the recruitment rather than the retention of KU70/KU80 at sites of DNA damage. (A) FRAP curves of GFP-KU70 at the DNA lesions fitted with two alternative models: i) a diffusion-limited model which assumes that KU interacts only very transiently with DNA at the sites of damage, and ii) a reaction-limited model which is valid if the fluorescence recovery is limited by the long lifetime of the interaction between KU and the DNA lesions [1]. The mean ±SEM is shown from 15 cells from a representative of 3 independent experiments is shown. (B) FCS curves of GFP-KU70 at the DNA lesions fitted with models assuming either one or two population dynamics. Based on the 2-population fit, the fast population displays a characteristic diffusion time of 590±70 µseconds within the confocal volume. Given that the molecular weight of the GFP-tagged KU70/KU80 complex (180 kDa) is ~6 times larger than GFP only, and that GFP shows a diffusion time of approximately 300 µseconds when diffusing through the nucleus [2], the Stockes-Einstein equation predicts that the diffusion time on an unbound KU70/KU80 complex should be equal to $300*6^{1/3}$=550 µseconds. Therefore, it seems reasonable to assume that the fast population of KU seen by FCS corresponds to freely diffusing KU complexes. The mean ±SEM is shown from 10 cells from a representative of 3 independent experiments. (C) FRAP curves of GFP-KU70 at the DNA lesions obtained for two different bleaching area and fitted with a reaction-limited model for the timepoints at more than 3 seconds post photobleaching. The mean ±SEM is shown from 15 cells from a representative of 3 independent experiments is shown. (D) k_{off} parameters estimated from the reaction-limited model for the two different sizes of bleached area from C. The mean ±SEM is shown from 15 cells from a representative of 3 independent experiments is shown. The boxplot limits correspond to the 25th and 75th percentiles and the center line in the box indicates the median value. The whiskers extend 1.5 times the interquartile range. P-values were calculated using the two-tailed unpaired Student's *t* test, assuming unequal variances. Source data are provided as Source Data file.

▶ **Supplementary Figure. 10 – related to Figure. 5 and 6. ZNF384 is dispensable for cNHEJ complex stability.** (A) Pull-downs of the indicated GFP fusion proteins in IR treated RPE-hTERT cells transfected with indicated siRNAs. Blots were probed for GFP, KU80, ZNF384, γH2AX and H3. Data shown represent 2 independent experiments. (B) Nuclear localization of GFP-APLF in U2OS Flp-In/T-Rex cells expressing doxycycline (dox)-inducible GFP-APLF transfected with the indicated siRNAs (left panel). Quantification of the data is presented as the mean of >200 cells acquired in 2 independent experiments (right panel). (C) Western blot analysis of KU80 in cells from B. Tubulin is a loading control. Data shown represent 2 independent experiments. (D) Nuclear localization of GFP-APLF in wildtype and *Ku80*[-/-] mouse embryonic stem (ES) cells (left panel). Quantification of the data is presented as the mean of >60 cells acquired in 2 independent experiments (right panel). (E) Pull-downs of the indicated GFP fusion proteins transfected with indicated siRNAs in U2OS cells. Blots were probed for GFP, KU80, ZNF384 and tubulin. Data shown represent 2 independent experiments. (F) Kinetics of GFP-ZNF384 recruitment to 365 nm UV-A tracks in BrdU-sensitized U2OS Flp-In/T-Rex *ZNF384* KO cells. White triangles indicate irradiated regions (left panel). Quantification of GFP-ZNF384 is presented as the mean ±SD from >120 cells acquired in 2 independent experiments (right panel). (G) Kinetics of GFP-KU70 recruitment to 365 nm UV-A tracks in BrdU-sensitized RPE1-hTERT cells. White triangles indicate irradiated regions. (left panel). Quantification of GFP-KU70 is presented as the mean ±SD from >120 cells acquired in 2 independent experiments (right panel). (H) Kinetics of GFP-APLF recruitment to 365 nm UV-A tracks in BrdU-sensitized U2OS Flp-In/T-Rex GFP-APLF cells. White triangles indicate irradiated regions (left panel). Quantification of GFP-APLF is presented as the mean ±SD from <120 cells acquired in 2 independent experiments (right panel). Statistical significance was calculated with the two-tailed unpaired Student's *t* test, assuming unequal variances. Scale bar 5 μm. Source data are provided as Source Data file.
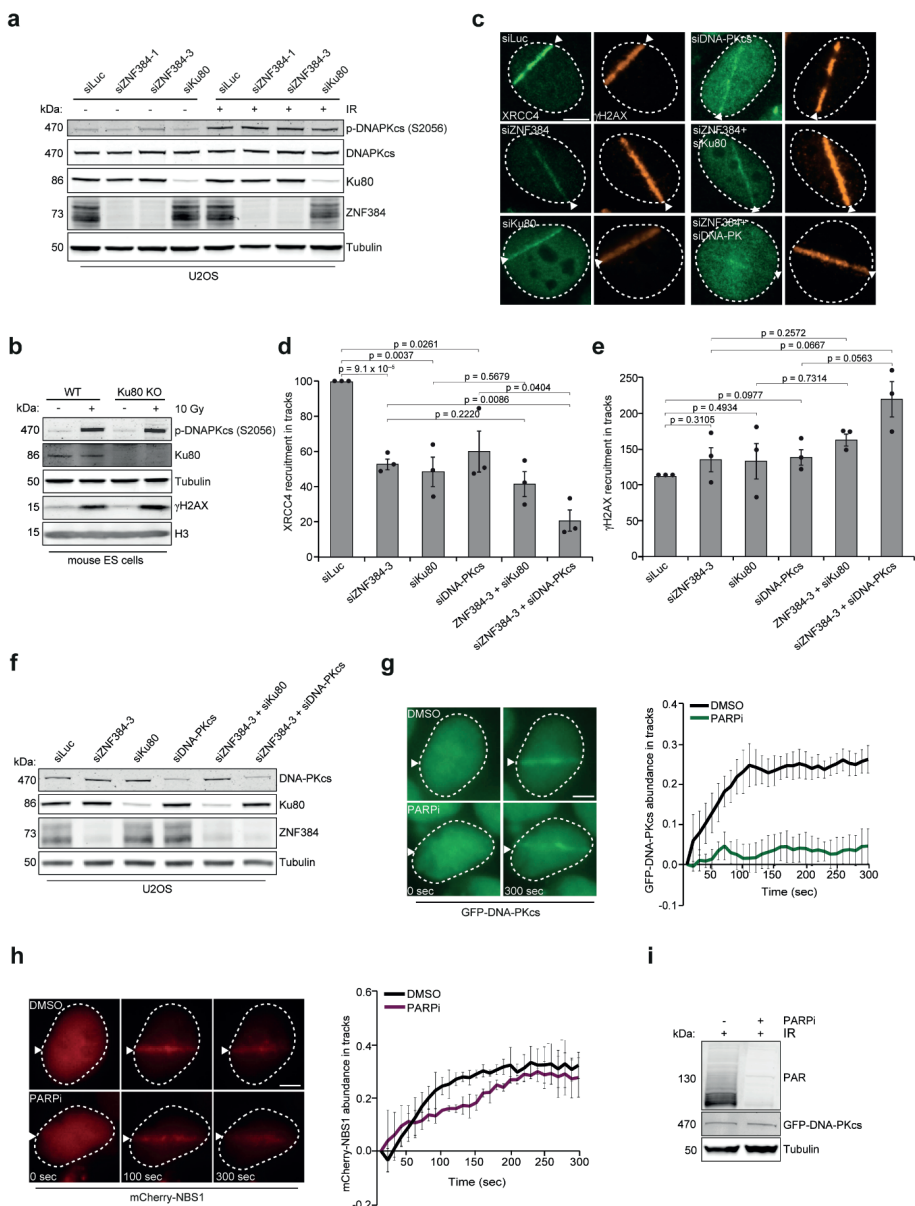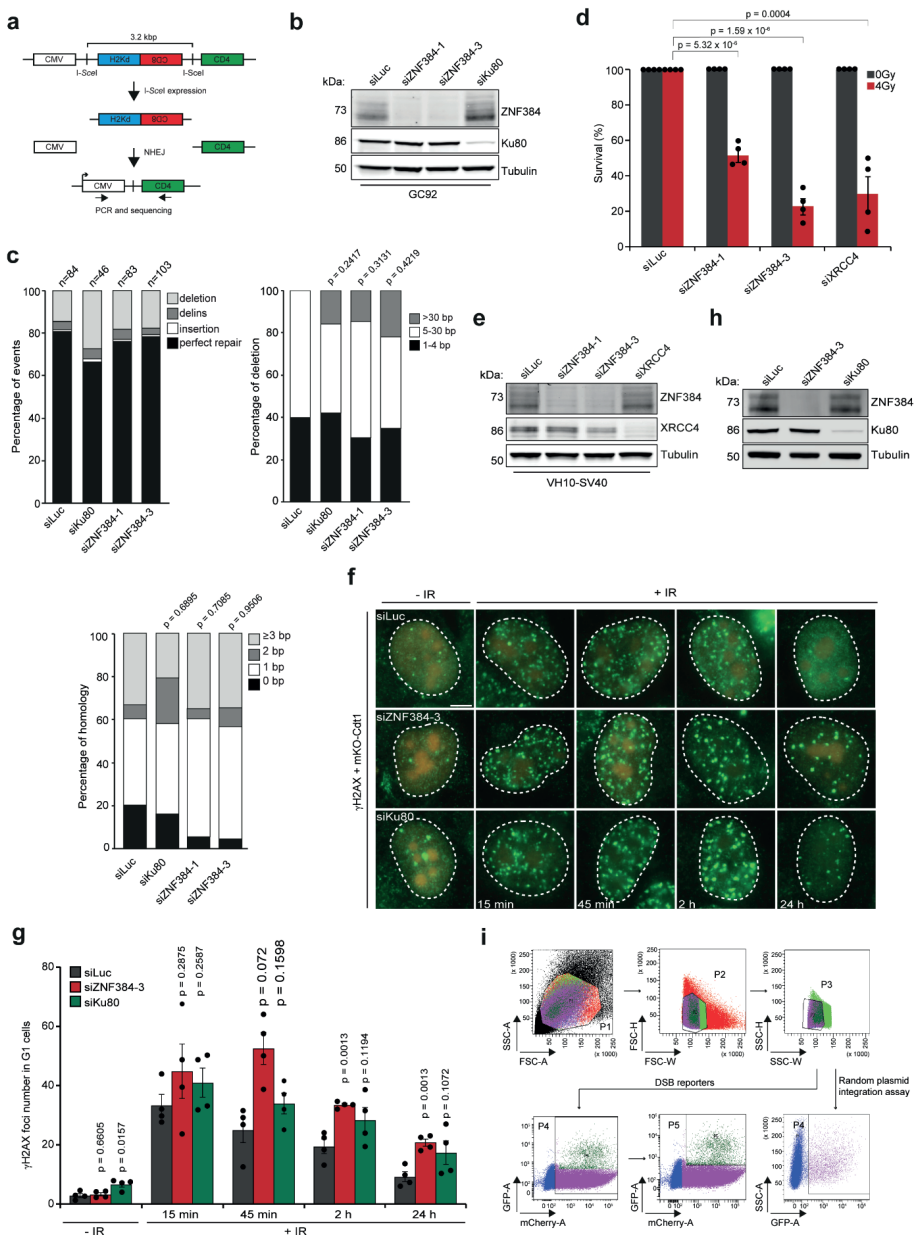
▶ **Supplementary Fig. 11 – related to Fig. 6. ZNF384 promotes recruitment of cNHEJ proteins independently of DNA-PKcs.** (A) Western blot analysis of DNA-PK activation 1 hour after 10 Gy of irradiation radiation (IR) in U2OS cells transfected with the indicated siRNAs. Blots were probed for p-DNA-PKcs (S2056), DNA-PKcs, KU80 and ZNF384. Tubulin is a loading control. Data shown represent 3 independent experiments. (B) Western blot analysis of DNA-PK activation 1 hour after 10 Gy of irradiation radiation (IR) in wildtype and *Ku80*[-/-] mouse embryonic stem cells. Blots were probed for p-DNA-PKcs (S2056), KU80 and γH2AX. H3 and tubulin are loading controls. Data shown represent 2 independent experiments. (C) Endogenous XRCC4 recruitment to 365 nm UV-A tracks 10 minutes after damage induction in U2OS cells transfected with the indicated siRNAs. γH2AX is a damage marker. White triangles indicate irradiated regions. (D) XRCC4 quantification of C. Data shows the mean ±SEM of >150 cells acquired in 3 independent experiments (E) γH2AX quantification of C. The mean ±SEM from 3 independent experiments is shown. Data were normalized to siLuc, which was set to 100%. (F) Western blot analysis of DNA-PKcs, KU80, ZNF384, γH2AX and Tubulin in cells from C. Tubulin is a loading control. Data shown represent 2 independent experiments. (G) GFP-DNA-PKcs recruitment to 365 nm UV-A tracks in BrdU-sensitized HCT116 cells treated with PARPi for 1 hour before laser micro-irradiation. White triangles indicate irradiated regions (left panel). Quantification of the data is presented as the mean ±SEM of >100 cells acquired in 3 independent experiments (right panel). (H) As in G, except for mCherry-NBS1. (I) Western blot analysis of PAR levels in IR- and PARPi -treated cells from G. Cells were treated with PARPi for 1 hour before IR treatment. Blots were probed for PAR, GFP and Tubulin. Tubulin is a loading control. Data shown represent 2 independent experiments. Statistical significance was calculated with the two-tailed unpaired Student's *t* test, assuming unequal variances. Scale bar 5 μm. Source data are provided as Source Data file.

3
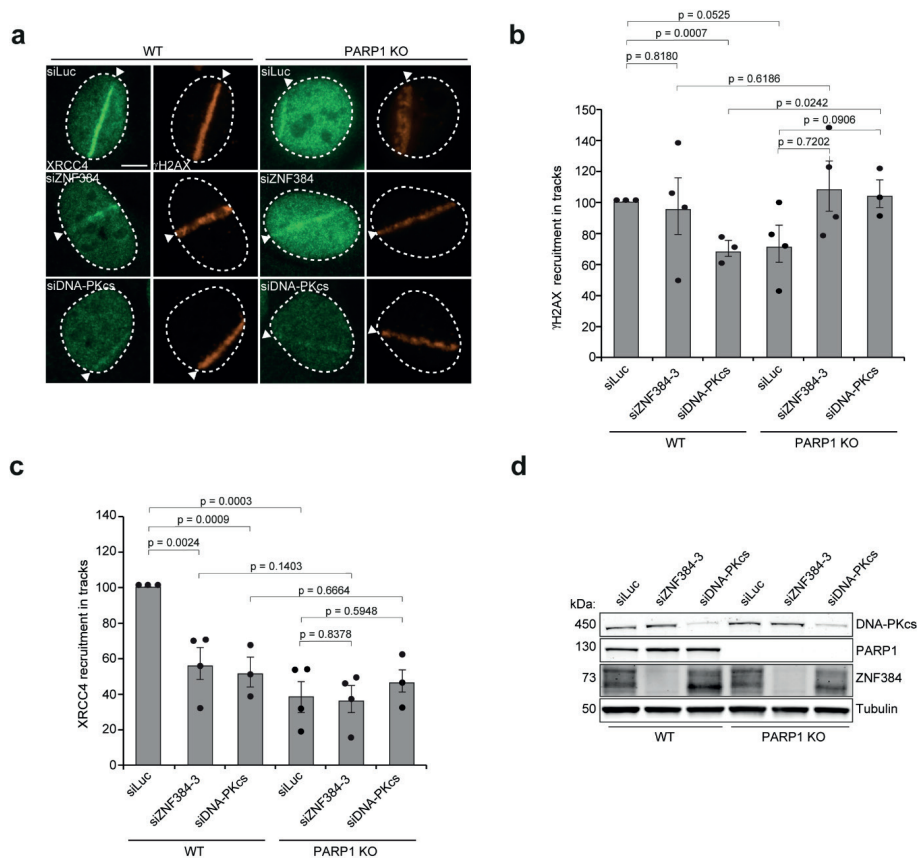
Supplementary Figure. 12 – related to Figure. 6. ZNF384 and PARP1, as well as DNA-Pkcs and PARP1, function epistatically in cNHEJ. (A) Endogenous XRCC4 recruitment to 365 nm UV-A tracks in BrdU-sensitized U2OS cells transfected with the indicated siRNAs. Cells were fixed and immunostained 10 minutes after laser micro-irradiation. White triangles indicate irradiated regions. γH2AX is a damage marker. (B) γH2AX quantification of A. Data shows the mean ±SEM of >240 cells acquired in 4 independent experiments (C) XRCC4 quantification of A. The mean ±SEM from 3 independent experiments is shown. Data were normalized to siLuc, which was set to 100%. (D) Western blot analysis of ZNF384, DNA-PKcs and PARP1 in cells from A. Tubulin is a loading control. Data shown represent 2 independent experiments. Statistical significance was calculated with the two-tailed unpaired Student's *t* test, assuming unequal variances. Scale bar 5 µm. Source data are provided as Source Data file.

a

% of cells — siLuc, siZNF384-1, siZNF384-3 (G2, S, G1/M)

b

ZNF384, Tubulin — U2OS

c

kDa: 470 DNA-PKcs, 250 53BP1, 120 PARP1, 100 LIG4, 86 Ku80, 73 ZNF384, 70 Ku70, 50 XRCC4, 50 Tubulin — siLuc, siZNF384-1, siZNF384-3 — U2OS

d

XRCC4/Tubulin — siLuc, siZNF384-1, siZNF384-3 — U2OS

e

XRCC4/Tubulin — siLuc, siZNF384-3 — Hela Flp-In

f

kDa: 350 p-ATM(S1981), 350 ATM, 73 ZNF384, 55 p-CHK1(S345), 55 CHK1, 50 Tubulin — IR: -, -, -, +, +, + — siLuc, siZNF384-1, siZNF384-3

g

ATM/Tubulin
- p = 0.0017
- p = 0.2242
- p = 0.6241
- p = 0.3911

-IR, +IR

h

CHK1/Tubulin
- p = 0.0711
- p = 0.4638
- p = 0.7881
- p = 0.7109

-IR, +IR

i

pATM/Tubulin
- p = 0.2394
- p = 0.1140
- p = 0.8682
- p = 0.3237

-IR, +IR

j

pCHK1/Tubulin
- p = 0.5746
- p = 0.6687
- p = 0.2723
- p = 0.4304

-IR, +IR

3

135

▶ **Supplementary Figure. 13 – related to Figure. 7. ZNF384 promotes efficient NHEJ independently of transcription.** (A) Cell cycle profile of the indicated siRNA-transfected U2OS cells. The fraction of G1-, S- and G2-phase cells was determined by propidium iodide staining and FACS analysis. >30000 cells per condition were analyzed. Data show mean ±SEM from 3-4 independent experiments. (B) Western blot analysis of ZNF384 expression in U2OS cells from A. Tubulin is a loading control. Data shown represent 2 independent experiments. (C) Western blot analysis of the expression of NHEJ factors in siRNA transfected ZNF384 knockdown U2OS cells. Blots were probed for ZNF384, 53BP1, DNA-PKcs, KU80, KU70, PARP1, XRCC4 and LIG4. Tubulin is a loading control. (D) Quantification of XRCC4 expression in cells from C. Tubulin is a loading control, which was used for normalization of XRCC4 expression. The mean ±SD from 2 independent experiments is shown. Data were normalized to WT + siLuc, which was set to 1. (E) Quantification of XRCC4 expression in cells from Figure 7E. Tubulin is a loading control, which was used for normalization of XRCC4 expression. The mean ±SD from 2 independent experiments is shown. Data were normalized to WT + siLuc, which was set to 1. (F) Western blot analysis of IR-induced ATM phosphorylation (p-ATM) at Ser1981 and CHK1 phosphorylation (p-CHK1) at Ser345. Cells were collected 1 hour after exposure to 10 Gy of IR. Blots were also probed for ATM, ZNF384, CHK1 and Tubulin. Tubulin is a loading control. A representative blot from 3 independent experiments is shown. (G) Quantification of ATM expression in cells from F. Tubulin was used for normalization of ATM expression. The mean ±SEM from 3 independent experiments is shown. Data were normalized to siLuc - IR, which was set to 1. (H) As in G, except for CHK1. (I) As in G, except for pATM. (J) As in G, except for pCHK1. Statistical significance was calculated with the two-tailed unpaired Student's *t* test, assuming unequal variances. Source data are provided as Source Data file.
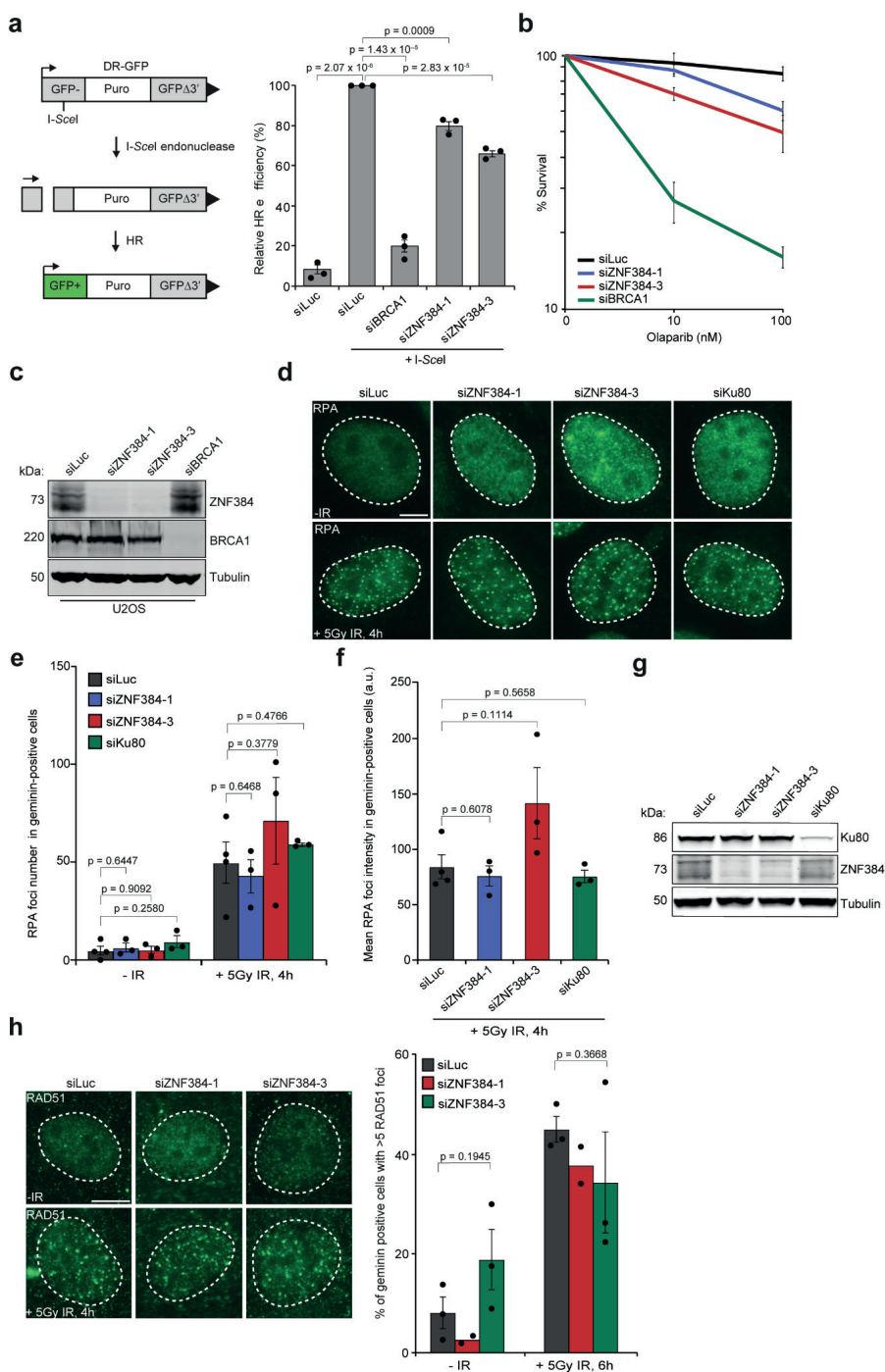
▶ **Supplementary Figure. 14 – related to Figure. 7. ZNF384 is dispensable for HR.** (A) Schematic of the DR-GFP reporter for HR (left panel). Quantification of HR efficiencies in 3 independent DR-GFP- U2OS transfected cells with the indicated siRNAs. I-*Sce*I transfection was corrected by co-transfection with mCherry expression vector. The mean ±SEM from 3 independent experiments is shown (right panel). (B) Clonogenic survival of U2OS cells transfected with the indicated siRNAs following treatment with the PARPi. The mean ±SEM from 3 independent experiments is shown. (C) Western blot analysis of ZNF384 and BRCA1 expression in U2OS cells. Tubulin is a loading control. Data shown represent 2 independent experiments. (D) RPA foci formation in U2OS cells transfected with the indicated siRNAs. Cells were exposed to 5 Gy of IR and foci were scored after 4 hours. (E) Quantification of RPA foci formation in cells from D. Foci were quantified in immuno-stained, geminin-positive S/G2 phase cells. The mean ±SEM from >200 cells from 3-4 independent experiments is shown. (F) As in E, except that RPA foci intensity was measured. (G) Western blot analysis of ZNF384 and KU80 expression in cells from D. Tubulin is a loading control. (H) RAD51 foci formation in U2OS cells transfected with the indicated siRNAs. Cells were exposed to 5 Gy of IR and foci were quantified after 6 hours (left panel) The mean ±SEM from >300 cells from 2-3 independent experiments is shown. Foci were quantified in immuno-stained, geminin-positive S/G2 phase cells (right panel). Statistical significance was calculated with the two-tailed unapired Student's *t* test, assuming unequal variances. Scale bar 5 μm. Source data are provided as Source Data file.

▶ **Supplementary Fig. 15 – related to Fig. 7. ZNF384 promotes DSB repair via cNHEJ.** (A) Schematic of the GC92 reporter for NHEJ. (B) Western blot analysis of ZNF384 and KU80 expression in GC92 fibroblasts. Tubulin is a loading control. Data shown represent 2 independent experiments. (C) Mutational signatures (top left panel, deletion sized (top right panel) and microhomology usage (in case of deletion formation) (bottom panel) at repair junctions in GC92 reporter for NHEJ. GC92 cells were transfected with the indicated siRNAs and I-*Sce*I expression vector. Repair junctions were amplified by PCR and Sanger-sequenced. (D) Clonogenic survival of VH10-SV40 cells transfected with the indicated siRNAs and exposed to 4 Gy of IR. The mean ±SEM from 4 independent experiments is shown. P-values were calculated using the two-tailed unpaired Student's *t* test, assuming unequal variances. (E) Western blot analysis of ZNF384 and XRCC4 expression in cells used in D. Tubulin is a loading control. Data shown represent 2 independent experiments. (F) γH2AX foci formation (green) in stable Monomeric Kusabira-Orange (mKO)-Cdt1 expressing U2OS cells transfected with the indicated siRNAs. Cells were exposed to 2 Gy of IR and foci were scored after 15 minutes, 45 minutes, 2 hours and 24 hours. (G) Quantification of γH2AX foci formation in Cdt1-positive G1 phase cells from F. The mean ±SEM from 2-4 independent experiments. Statistical significance was calculated with the two-tailed unpaired Student's t-test, assuming unequal variances. (H) Western blot analysis of ZNF384 and KU80 expression in cells from F. Tubulin is a loading control. Data shown represent 2 independent experiments. (I) Gating strategy for flow cytometric analysis of cells for DR-GFP and EJ5- GFP reporter assays, as well as random plasmid integration assays. Scale bar 10 μm. Source data are provided as a Source Data file.

Supplementary Table 1.

Table S1. Statistical analysis comparing LFQ intensities of proteins identified by LC-MS/MS in GFP-pulldowns from U2OS Flp-In/T-Rex cells expressing either GFP-ZNF384 or GFP-NLS. -Log p-value and average fold-enrichment (log2) are indicated. Additional information from the analysis of identified proteins is also included.

3

141

3

3

sp	Q96H79	ZCCH Zinc finger CCCH-ty ZC3HAV1L	+	1.59	-0.69	4	4	4	16.7	16.7	16.7	32,962	0	15.37	3.86E+08	19	Q96H79	(Q96H79	26.90	26.55	28.24	25.99	25.62	25.80
sp	Q9H4A6	KT3K K dosamino-3-kinas(FN3KRP	+	1.85	-0.71	6	6	6	23.9	23.9	23.9	34,412	0	32.44	2.83E+09	43	Q9HA-64, Q9HA-64	29.59	29.40	29.40	28.54	28.67	29.07
tr	I3QR08	I3QR0(Ribosomal protein L RPL19		0.32	-0.72	4	4	4	22.8	22.8	22.8	23,134	0	29.32	4.48E+08	15	J3QR09, J3QR09	26.76	26.35	26.33	24.16	27.33	25.77
sp|P16527|C-SRP Cysteine and glycin(C-SRP2		2.04	-0.73	5	5	5	35.8	35.8	35.8	20,954	0	33.58	2.78E+09	49	Q16527|F Q16527|I	29.70	29.29	29.32	28.60	28.65	28.87
sp|P10599|THIO_I Thioredoxin TXN	+	2.17	-0.74	5	5	5	42.9	42.9	42.9	11,737	0	28.86	3.03E+09	47	P10599|F P10599|F	29.46	29.73	29.60	28.64	29.05	28.86
sp	P62906	RL10A 60S ribosomal prote RPL10A		0.41	-0.75	3	3	3	13.4	13.4	13.4	24,831	0	5.087	2.17E+08	6	P62906 P62906	25.48	25.54	25.79	23.50	26.18	24.96
sp	P84103	RS-RS S-arginine/arginine- h SRSF3	+	1.37	-0.77	4	4	3	30.6	30.6	30.6	14,203	0	21.2	1.59E+09	31	P84103-2 P84103-2	29.08	28.57	28.21	27.72	27.87	27.97
sp|P50454|ERP Serpin H1 SERPINH1		0.36	-0.78	3	3	3	11.7	11.7	11.7	46.44	0	16.11	2.04E+08	11	P50454|E P50454|E	25.68	25.50	25.12	25.41	25.67	22.88
sp|P04899|GNAI2 Guanine nucleotide l GNAI2|GNAS |GNA0 1|SN		0.54	-0.79	2	2	2	7.3	7.3	7.3	40.45	0	2.967	2.32E+08	10	P04899|F P04899|F	25.51	25.68	25.56	25.41	25.60	23.59
tr	E9PPU1	E9PP 40S ribosomal prote RPS3		0.57	-0.80	9	9	9	63.3	63.3	63.3	17,407	0	21.97	1.34E+09	40	E9PPU1 E9PPU1	28.35	28.13	28.14	26.40	28.54	27.29
tr	A0A590UJK4	A S-arine/arginine-rich s SRSF6	+	1.56	-0.80	5	5	3	20.6	20.6	11.5	26,306	0	8.505	5.58E+08	19	A0A590, A0A590L	27.44	27.17	26.95	26.76	26.21	28.19
tr	E7EX73	E7EX Eukaryotic translatio EIF4B1		0.64	-0.82	2	2	2	1.7	1.7	1.7	156.64	0	8.418	8.18460X0	4	E7EX73; E7EX73;	24.24	24.39	24.42	24.12	22.36	24.09
sp	P46781	RS9_F 40S ribosomal prote RPS9		0.61	-0.83	10	10	10	43.6	43.6	43.6	22,591	0	16.33	1.2E+09	36	P46781;B P46781;B	28.28	27.90	28.11	26.42	28.43	28.96
sp	P05109	S10A6 Protein S100-A6,Pro S100A6		0.56	-0.83	3	3	3	28.9	28.9	28.9	10,634	0	9.963	9.86E+08	20	P05109 P05109	27.61	27.34	28.66	26.40	27.96	26.95
tr	A0A494C0J7	A(Protein glutamine-ga TGM3		1.07	-0.85	6	6	6	8.6	8.6	8.6	78,656	0	11.45	4.84E+08	15	A0A494C A0A494C	26.36	26.96	27.47	25.77	26.42	28.09
sp	P62277	RS13_ 40S ribosomal prote RPS13		1.04	-0.85	6	6	6	35.6	35.6	35.6	17,222	0	25.31	1.2E+09	26	P62277,I P62277,I	28.47	28.22	27.96	26.92	28.06	27.12
sp|P38373|RL13_I 60S ribosomal prote RPL13		0.26	-0.85	7	7	7	29.9	29.9	29.9	24,261	0	45.13	6.39E+08	25	P38373|F P38373|F	27.61	26.60	27.34	24.17	28.51	26.53
sp	P04040	CATA_ Catalase CAT		0.93	-0.86	6	6	6	16	16	16	59,755	0	42.37	6.33E+08	18	P04040 P04040	26.82	27.91	27.13	26.05	26.99	26.24
tr	J3KT73	J3KT73_ 60S ribosomal prote RPL36		1.03	-0.87	3	3	3	39.1	39.1	39.1	7,5649	0	39.6	1.12E+09	23	J3KT73;J J3KT73;I	28.57	28.22	27.91	26.71	27.47	27.90
sp	P53007	TXTP_ Tricarboxylate transp SLC25A1		0.66	-0.88	2	2	2	7.7	7.7	7.7	34,012	0	11.8	1.28E+08	6	P53007 P53007	25.43	25.41	23.72	23.53	24.16	24.22
sp	P42357	2HUTI Histidine ammonia-l HAL		0.84	-0.89	2	2	2	3.4	3.4	3.4	64,667	0	2.736	1.03E+08	6	P42357-2 P42357-2	24.42	24.57	25.50	23.28	24.46	24.08
sp	Q9HCY8	S10A Protein S100-A14 S100A14		0.81	-0.91	2	2	2	25	25	25	11,662	0	12.49	1.08E+08	9	Q9HCY8 Q9HCY8	24.83	24.64	24.96	24.37	24.46	22.89
sp	Q9UBV8	PEF_ Peflin PEF1		2.04	-0.92	4	4	4	12	12	12	30,381	0	8.4	2.55E+08	13	Q9UBV8 Q9UBV8	26.26	25.91	26.20	25.22	24.92	25.46
sp	Q92896	2	AN3_ Acidic leucine-rich n ANP32B		0.68	-0.93	5	4	4	28.7	22.1	22.1	22,276	0	10.01	6.79E+08	16	Q92896-2 Q92896-2	27.85	27.96	28.13	26.74	26.20	28.25
sp	P61313	RL15_I 60S ribosomal prote RPL15		0.32	-0.93	6	6	6	32.8	32.8	32.8	24,146	0	30.22	4.7E+08	16	P61313;A P61313;A	26.60	26.17	26.62	23.39	27.45	25.90
tr	H7C468	H7C46(Cathepsin D,Cathep CTSD		1.08	-0.93	4	4	4	12.2	12.2	12.2	35,785	0	21.92	3.93E+08	13	H7C468;/ H7C468;/	26.18	26.63	27.08	25.69	26.33	25.29
tr|Q5JR96|Q5JR9(40S ribosomal prote RPS8		0.56	-0.95	4	4	4	27.7	27.7	27.7	21,879	0	54.49	1.2E+09	27	Q5JR96;| Q5JR96;|	28.11	28.19	28.15	25.90	28.52	27.17
sp|P35998|PRS7_ 26S protease regulat P SMC2		0.39	-0.96	3	3	3	7.6	7.6	7.6	48,633	0	4.264	2.63E+08	5	P35998|F P35998|F	26.11	25.78	25.99	26.16	25.86	22.97
sp	P17096	HMGA High mobility group (HMGA1	+	1.32	-0.96	2	2	1	22.4	22.4	15	11,676	0	16.07	1.96E+08	8	P17096;F P17096;F	26.28	25.54	25.17	24.81	24.52	24.77
tr	H0YEN5	H0YE 40S ribosomal prote RPS2		0.69	-0.99	4	4	4	22.6	22.6	22.6	21,154	0	29.05	1.13E+09	23	H0YEN5 H0YEN5	28.21	28.15	27.62	26.10	28.29	26.81
sp	Q96PE8	SPB1 Serpin B12 SERPINB12	+	1.85	-1.03	6	6	6	17.8	17.8	17.8	46,276	0	35.11	9.79E+08	29	Q96PE8; Q96PE8;	27.87	28.33	27.99	26.63	27.26	27.21
tr	J3KRG2	J3KRG Gasdermin-A GSDMA		0.71	-1.04	2	2	2	13.9	13.9	13.9	17,791	0	4.491	7104100X	3	J3KRG2, J3KRG2,	23.77	23.79	25.43	22.53	23.72	23.61
tr	M0R210	M0R21(40S ribosomal prote RPS16		0.90	-1.04	5	5	5	40.3	40.3	40.3	14,419	0	16.68	9.64E+08	33	M0R210,I M0R210,I	28.28	27.53	27.95	28.09	27.80	26.74
tr	H0YM70	H0YM(Peroxisomal acyl-co P SME2		0.41	-1.06	5	5	5	23.2	23.2	23.2	26,011	0	25.41	7.45E+08	8	H0YM70, H0YM70,	27.58	26.97	27.41	27.30	24.09	27.36
sp	Q6FIN16-2	H5(Hydroxysteroid dehy HSDL2		0.57	-1.07	2	2	2	10.1	10.1	10.1	37.32	0	18.41	2.28E+08	13	Q6FIN16- Q6FIN16-	28.40	25.56	25.21	25.40	25.25	23.10
sp	Q59FF8	H3BS2 Putative heat shock HSP90AB2P		0.41	-1.08	6	1	1	16.3	3.9	3.9	44,348	0	3.401	4.9E+08	13	Q59FF8 Q59FF8	27.05	26.36	26.55	26.68	26.68	23.37
sp	P46109	CRKL_ Crk-like protein CRKL	+	2.03	-1.14	7	7	7	33	33	33	33,777	0	60.5	1.35E+09	31	P46109 P46109	28.69	28.64	28.46	27.09	27.40	27.88
tr	X6R700	X6R70(Chromatin target of P CHTOP		1.11	-1.17	2	2	2	11.7	11.7	11.7	23,661	0	18.42	1.54E+08	12	X6R700;(X6R700;(	25.47	24.58	24.65	23.36	23.29	24.54
tr	J3KTL2	J3KTL2_ S-arine/arginine-rich s SRSF1	+	2.06	-1.19	7	7	7	28.1	28.1	28.1	28,329	0	20.69	8.86E+08	32	J3KTL2C J3KTL2C	28.36	27.87	27.52	26.65	26.70	26.83
sp	P47929	LEG7_ Galectin-7 LGALS7		1.28	-1.23	4	4	4	35.3	35.3	35.3	15,075	0	12.66	6.76E+08	17	P47929 P47929	26.78	27.71	26.26	26.21	26.61	26.22
sp	Q01650	LAT1_I Large neutral amino (SLC7A5		1.10	-1.27	2	2	2	6.3	6.3	6.3	55.01	0	16.79	2.41E+08	17	Q01650 Q01650	26.07	25.96	36.28	25.39	23.76	25.37
tr	Q2L6C6	Q2L6C(Cornecdermain CDSN		0.89	-1.28	3	3	3	9.6	9.6	9.6	51,539	0	45.11	8.93E+08	16	Q2L6C6;I Q2L6C6;I	27.80	28.61	26.77	26.06	27.12	26.35
tr	H3BV65	H3BV6 BolA-like protein 2 BOLA2B	BOLA2		0.89	-1.29	2	2	2	46.3	46.3	46.3	6,3612	0	15.71	1.39E+08	5	H3BV65; H3BV65;	25.07	25.61	24.99	22.63	24.49	24.68
sp|P53996|TCP4_ Activated RNA poly SUB1		1.64	-1.29	7	7	7	46.5	46.5	46.5	14,395	0	28.62	3.66E+09	34	P53996|C P53996|	30.31	30.36	29.34	28.86	28.43	28.62
sp	P14923	PLAK_ Junction plakoglobin JUP	+	1.56	-1.31	23	23	23	33.8	33.8	33.8	81,744	0	227.4	1.31E+10	179	P14923;C P14923	31.56	32.39	31.25	30.13	30.72	30.44
sp	Q06554-2	DSC Desmocollin-1 DSC1		1.23	-1.31	10	10	10	16.2	16.2	16.2	93,634	0	136.1	4.06E+09	69	Q06554-2 Q06554-2	30.05	30.65	29.28	28.43	29.16	28.65
sp|P05089|ARGI1 Arginase-1 ARG1		1.66	-1.32	13	13	13	57.1	57.1	57.1	34,735	0	105.6	2.5E+09	53	P05089|F P05089|F	28.97	29.64	29.52	27.66	28.54	28.19
sp	O75940	2PDC Programmed cell dea PDCD6	+	1.36	-1.35	2	2	2	10.1	10.1	10.1	21,664	0	4.348	9542000X	6	O75940-2 O75940-2	25.05	24.79	24.77	23.44	22.76	24.33
tr	H3BV11	H3BV1 Nuclear pore comple NUP93		0.62	-1.37	3	3	3	33.8	33.8	33.8	15.23	0	6.954	1.87E+08	4	H3BV11; H3BV11;	25.26	24.99	25.07	25.72	22.79	22.69
tr	H0YN28	H0YN: Acidic leucine-rich n ANP32A		0.83	-1.39	6	6	5	24.9	24.9	24.9	19,997	0	25.62	1.12E+09	17	H0YN28, H0YN28,	25.82	28.67	26.61	27.10	26.35	26.69
sp|P15924|DESP Desmoplakin DSP		1.46	-1.42	102	102	102	32.6	32.6	32.6	331.77	0	323.3	3.12E+10	553	P15924|F P15924|F	32.97	33.74	32.27	31.26	31.78	31.67
sp	J3KU69	ITA3_I Integrin alpha-3,Integ ITGA3	+	1.90	-1.42	2	2	2	1.9	1.9	1.9	146,61	0	4.294	1.3E+06	5	P26084;P P26084;P	25.07	24.49	25.30	23.16	24.27	23.65
tr	P8WBP5	P8W6F Homeglobin subunit H88	HBD-2		2.28	-1.45	4	4	4	64.4	64.4	64.4	9.67	0	11.53	3.57E+08	16	P8W6P5, P8W6P5,	26.98	26.69	26.22	25.04	25.40	25.31
sp	Q02413	D-SG1_ Desmoglein-1 DSG1		1.33	-1.45	18	18	18	24.9	24.9	24.9	113.75	0	313.4	1.02E+10	124	Q02413;C Q02413	31.20	32.15	30.68	29.37	30.26	30.02
sp|C0149|FAB P1 Fatty acid-binding pr FABP5		1.17	-1.46	5	5	5	25.9	25.9	25.9	15,164	0	12.35	2.72E+09	33	C0149|H C0149|I	28.97	28.93	28.09	28.09	27.99	27.92
sp	P04264	K2C6B Keratin, type II cytos KRT6B		0.95	-1.50	52	4	3	62.4	8.7	5.7	60,066	0	25.12	4.03E+09	10	P04259 P04259	30.51	30.67	28.89	27.89	28.26	29.44
tr	A0A2R8Y7D1	A Dynamin-like 120 kD QPAI1		0.97	-1.51	3	3	3	8.8	8.8	8.8	72,406	0	8.039	1.46E+08	6	A0A2R8\ A0A2R8\	25.83	25.89	24.59	24.54	24.37	22.46
sp	Q14574-2	DSC Desmocollin-3 DSC3		1.62	-1.58	5	5	5	10.1	10.1	10.1	99,453	0	105.4	8.33E+08	14	Q14574-2 Q14574-2	27.83	28.64	27.21	26.03	26.22	28.39
sp	Q6UWP8	S BS Suprabasin SBSN		1.38	-1.59	9	9	9	34.9	34.9	34.9	60.54	0	24.52	1.58E+08	56	Q6UWP(Q6UWP(	28.22	29.51	28.22	26.95	27.25	27.26
sp	O43927	CYTB_ Cystatin-B C6T8		1.19	-1.64	2	2	2	24.5	24.5	24.5	11,139	0	8.507	1.68E+08	7	P04080;A P04080;A	25.04	25.17	27.04	24.62	24.14	24.17
sp	O13835-2	PKP Plakophilin-1 PKP1		1.88	-1.66	13	13	13	18.5	18.5	18.5	80,496	0	64.49	1.34E+09	40	O13835-2 O13835-2	29.20	29.18	27.97	26.62	27.01	27.65
tr	F8VYN9	F8VYN A DP-ribosylation fac ARL1	+	1.48	-1.69	2	2	2	12.9	12.9	12.9	21,778	0	5.286	1.54E+08	9	F8VYN9; F8VYN9;	25.53	25.66	25.26	23.56	24.84	23.08
sp	O6MVV4	PCF Protein POF1B POF1B		1.98	-1.71	6	6	6	12.7	12.7	12.7	68,064	0	28.77	6.61E+08	26	Q6MVV4 Q6MVV4	27.38	28.00	27.56	25.41	26.82	25.92
sp	O9NZT1	CALL1 Calmodulin-like prote CALML5		0.51	-1.78	6	6	6	47.3	47.3	47.3	15,892	0	290.4	1.52E+09	23	Q9NZT1 Q9NZT1	29.04	30.01	28.11	26.05	23.73	25.85
tr|F8VWU4|F8VW Ras-related protein R RAB5C|RAB5B|RAB5A		0.64	-1.79	2	2	2	24.2	24.2	24.2	9,9152	0	6.238	3.43E+08	13	F8VWU4 F8VWU4	26.44	26.36	26.73	22.22	25.95	26.01
tr	A0A0B4J121	A(S-arine/arginine-rich s SRSF7		0.86	-1.80	3	3	2	21.9	21.9	21.9	15,763	0	20.31	2.66E+08	9	A0A0B4, A0A0B4,	26.49	26.94	26.54	25.70	26.21	23.11
tr	H3BPC4	H3BP SUMO-conjugating (UBE2I		0.90	-1.85	2	2	2	32.9	32.9	32.9	7,8691	0	4.248	1.66E+08	3	H3BPC4 H3BPC4	25.50	25.41	25.28	22.82	22.37	25.45
tr	M0QZC5	M0QZC 40S ribosomal prote RPS11		1.82	-2.00	3	3	3	22	22	22	13,997	0	5.095	2.73E+08	18	M0QZC5, M0QZC5,	27.54	26.26	26.01	23.99	25.10	23.47
tr	J3QL05	J3QL05_ S-arine/arginine-rich s SRSF2	SRSF8		0.74	-2.26	4	4	4	32.3	32.3	32.3	15,156	0	21.73	8.48E+08	20	J3QL05;J J3QL05;,	28.30	27.65	27.16	26.92	22.79	28.04
tr	A0A0C4DP9	A Protein SETP potein SETS6	ETS1P		1.00	-2.95	4	4	4	19.5	19.5	19.5	31,124	0	13.91	6.51E+08	18	A0A0C4(A0A0C4(	27.69	27.93	27.02	27.17	22.57	24.07

Supplementary Table 2. List of siRNAs

Target	Sequence (5'-3')
BRCA1	AGAUAGUUCUACCAGUAAA
ZNF384-1	GCACAUCCGUAUACACUCA
ZNF384-3	GGCAACACAACAAAGAUAA
DNA-PKcs	CUUUAUGGUGGCCAUGGAG
Ku80	CAAGGAUGAGAUUGCUUUAGU
Luciferase	CGUACGCGGAAUACUUCGA
PARP1	CCAUCGAUGUCAACUAUGA
XRCC4	AUAUGUUGGUGAACUGAGA
PARP3	SMARTpool: ON-TARGETplus (Dharmacon)
Ku70	SMARTpool: siGENOME (Dharmacon)

Supplementary Table 3. List of primers

Name	Sequence (5'-3')
HindIII FW	TAAAAAGCTTATATGGAAGAATCTCACTTCAATTCTAAC
KpnI RV	TAATAATGGTACCCTAAGAGCTGGCCAGGTGC
ZNF384-3_siRNAres_HindII Rv	ACTTGAAGGGCTTGTCCTTATTATGCTGTCGTCT
ZNF384-3_siRNAres_KpnI FW	AGACGACAGCATAATAAGGACAAGCCCTTCAAGT
CMV1 FW	TGGCCCGCCTGGCATTATGCC
CD4int RV	GCTGCCCCAGAATCTTCCTCT
M13 FW	GTAAAACGACGGCCAGT
M13 RV	CAGGAAACAGCTATGAC
ZNF384 △N-terminus(1-209) Fw	CGAGCTCAAGCTTATATGGACCATCAGAAAGAC
ZNF384 △N-terminus(1-209) Rv	GCCGTCTTTCTGATGGTCCATATAAGCTTGAGC
ZNF384 △C2H2(205-410) Fw	ATGCTGGAATCAGGGCTGGCAGTGGCCCAG
ZNF384 △C2H2(205-410) Rv	GGCCTGGGCCACTGCCAGCCCTGATTCGTC
ZNF384 △C-terminus(401-516) Fw	CCGCCTGATCTTCAGCCGTTTAAACCCGCT
ZNF384 △C-terminus(401-516) Rv	AGCGGGTTTAAACGGCTGAAGATCAGGCGG
ZNF384 N-terminus Fw	ATTAAAAAGCTTCTATGGAAGAATCTCACTTC AATTCTAACCC

Name	Sequence (5'-3')
ZNF384 N-terminus Rv	TAATAATGAATTCTTAGAATTCTTAATCATCATCCTCAGGGGAGAGGAC
ZNF384 C-terminus Fw	ATTAAAAAGCTTCTCAACAGGTGCAGGCAGCAG
ZNF384 C-terminus Rv	TAATAATGAATTCCTAAGAGCTGGCCAGGTGC
ZNF384 C2H2 Fw	ATTAAAAAGCTTCTCCCGAGATGAATGACCCTTATGTCC
ZNF384 C2H2 Rv	TAATAATGAATTCTTATGCCGCTGCTGCTGCTG
pET-His6-MBP-ZNF384 Fw	TACTTCCAATCCAATGCAATGGAAGAATCTCACTTCAATTCTAAC
pET-His6-MBP-ZNF384 Rv	TTATCCACTTCCAATGTTATTACTAAGAGCTGGCCAGGTG
pET-His6-MBP-N-term Fw	ATTAAAAATATTATGGAAGAATCTCACTTCAATTCTAACCC
pET-His6-MBP-N-term Rv	TAATAATGGATVVTTAATCATCATCCTCAGGGGAGAGGAC
pET-His6-MBP-C-term Fw	ATTAAAAATATTCAACAGGTGCAGGCAGCAG
pET-His6-MBP-C-term Rv	TAATAATGGATCCCTAAGAGCTGGCCAGGTGC
pET-His6-MBP-C2H2 Fw	ATTAAAAATATTCCCGAGATGAATGACCCTTATGTCC
pET-His6-MBP-C2H2 Rv	TAATAATGGATCCTTATGCCGCTGCTGCTGCTG
ssDNA T-strand / 3'overhang FW / dsDNA$^{3'-T}$ FW	/5Biosg/CGCTATCGTCTACGTCATGATCGCAGAAAAAGAAAAAAAAAGAA
ssDNA A-strand / dsDNA$^{5'T}$ FW	TTCTTTTTTTTTTCTTTTTCTGCGATCATGACGTAGACGATAGCG
forward 5'overhang / dsDNA$^{5'T}$ RV	CGCTATCGTCTACGTCATGATCGC
5'overhang RV	/5Biosg/CGCTATCGTCTACGTCATGATCGC
3'overhang RV	/5Biosg/GCGATCATGACGTAGACGATAGCGTTCTTTTTTTTCTTTTTCT
dsDNA$^{3'-T}$ RV	AGAAAAAGAAAAAAAAAGAACGCTATCGTCTACGTCATGATCGC
RNA U-strand	/5Biosg/GCGAUCAUGACGUAGACGAUAGCGUUCUUUUUUUUUCUUUUUCU

Supplementary Table 4. List of primary antibodies

Protein	Host	Company	IF	WB
BRCA1	Mouse	Santa Cruz (sc-6954)		1:500
DNA-PKcs	Mouse	Abcam (clone 18-2)		1:750
p-DNA-PKcs (S2056)	Rabbit	Abcam (ab18192)		1:750
GFP	Mouse	Roche (11814460001)		1:2500
Ku80	Rabbit	Santa Cruz (H-300, sc-9034)	1:500	1:1000
PAR	Rabbit	Enzo Lifesciences	1:1000	1:2000
Ku70	Mouse	Santa Cruz (sc-17789)		1:1000
RAD51	Mouse	GeneTex (clone 14B4)	1:200	
α-Tubulin	Mouse	Sigma (cloneDM1A, T6199)		1:5000
XRCC4	Rabbit	Gift from D. van Gent	1:500	
XRCC4	Mouse	SAB (40455)		1:1000
ZNF384	Rabbit	Abcam (ab176689)		1:1000
γH2AX	Mouse	Millipore (clone JBW301, 05-636)	1:2000	1:1000
PARP1	Rabbit	Cell signaling (9542S)		1:1000
GFP	Mouse	Sigma (11814460001)		1:1000
H3	Rabbit	Abcam (ab1791)		1:10000
LIG4	Rabbit	Abcam (ab193353)		1:1000
53BP1	Rabbit	Novus Biologicals (NB100-304)		1:2000
γH2AX	Mouse	Upstate (clone JBW301, #05-636)	1:100	
ZNF384	Rabbit	ATLAS antibodies	1:1000	
PARP3	Rabbit	Gift from F. Dentzer		1:10000
Geminin	Rabbit	Proteintech (10802-1AP)	1:400	
ATM	Rabbit	Cell signaling (clone D2E2)		1:2000
pATM (S1981)	Mouse	Cell signaling (4526)		1:2000
CHK1	Mouse	Santa Cruz (clone G-4, sc8408)		1 :1000
pCHK1 (S345)	Rabbit	Cell signaling (clone 133D3)		1 :1000
RPA32	Mouse	Abcam (ab1275)	1 :1000	
PARP2	Mouse	ENZO lifesciences		1 :1000
Pan-ADPr	Rabbit	Sigma (MABE1016)		1 :1500
Actin	Rabbit	Sigma (A5060)		1 :1000
MBP	Mouse	NEB (E8032S)		1 :10000

Supplementary Table 5. List of biotin-labeled DNA substrates

Name	Sequence
ssDNA T-strand	/5Biosg/-GCGATCATGACGTAGACGATAGCGTTCTTTTTTTTTCTTTTTCT
ssDNA A-strand	TTCTTTTTTTTTTCTTTTTCTGCGATCATGACGTAGACGATAGCG-/5Biosg
5'overhang	TTCTTTTTTTTTTCTTTTTCTGCGATCATGACGTAGACGATAGCGCGCTAGTACTGCATCTGCTATCGC-/5Biosg/
3'overhang	/5Biosg/-GCGATCATGACGTAGACGATAGCGTTCTTTTTTTTTCTTTTTCTCGCTAGTACTGCATCTGCTATCGC
dsDNA$^{3'\text{-}T}$	/5Biosg/-GCGATCATGACGTAGACGATAGCGTTCTTTTTTTTTCTTTTTCTCGCTAGTACTGCATCTGCTATCGCAAGAAAAAAAAAGAAAAAGA
dsDNA$^{5'T}$	TTCTTTTTTTTTTCTTTTTCTGCGATCATGACGTAGACGATAGCGAAGAAAAAAAAAGAAAAAGACGCTAGTACTGCATCTGCTATCGC-/5Biosg/

Supplementary References

1. Sprague, B.L., et al., *Analysis of binding reactions by fluorescence recovery after photobleaching.* Biophys J, 2004. **86**(6): p. 3473-95.

2. Smith, R., et al., *Poly(ADP-ribose)-dependent chromatin unfolding facilitates the association of DNA-binding proteins with DNA at sites of damage.* Nucleic Acids Res, 2019. **47**(21): p. 11250-11267.

CHAPTER 4

KANSL3 suppresses R-loop formation and replication fork instability

Jenny Kaur Singh, Jolanthe Lingeman, Wouter W. Wiegant, Anton J.L. de Groot, Román González-Prieto, Alfred C.O. Vertegaal, and Haico van Attikum

ABSTRACT

KANSL3 is a member of the NSL chromatin remodeling complex. This complex is mainly involved in the transcriptional regulation of housekeeping genes. Whether this complex plays a role in processes other than transcription is largely unclear. Here we show that KANSL3 is a novel factor involved in the replication stress response. We found that KANSL3-depleted cells display an increased sensitivity to hydroxyurea. Furthermore, we observed that after hydroxyurea treatment, KANSL3-depleted cells show increased pRPA levels, reduced fork restart as measured by DNA fiber analysis and increased R-loop formation. Interestingly, increased levels of pRPA and R-loops were also observed in untreated cells. These results suggest that KANSL3 has a suppressive role on R-loop formation, thereby preventing collisions between the replication fork and the transcription machinery that can lead to replication fork arrest/collapse. How KANSL3's role in transcription regulation affects R-loop formation and consequently replication fork stability/recovery following stress is subject of future studies.

Keywords: Replication stress (RS), R-loops, KANSL3, NSL complex

INTRODUCTION

DNA replication is a tightly regulated process that guarantees the accurate duplication of the genome once per cell cycle. DNA replication is, however, constantly challenged by several sources, which can cause a phenomenon known as replication stress (RS) [1]. RS is defined as the slowing or stalling of replication fork progression and/or DNA synthesis, and can be caused by for instance limiting nucleotides, ribonucleotide incorporation, repetitive DNA elements, transcription-replication conflicts, DNA secondary structures, fragile sites, and oncogene-induced stress [2]. Fortunately, cells have evolved sophisticated S phase and mitotic checkpoint pathways that help DNA replication to complete in the face of RS. Defects in these checkpoint pathways can lead to the stalling or collapse of replication forks, which in turn can cause mutations, chromosome rearrangements or the missegragation of chromosomes, thereby contributing to the development of diseases such as cancer [1].

ATR is the major kinase involved in dealing with RS and is recruited to RPA-coated single-stranded DNA (ssDNA) generated by uncoupling of the replicative MCM helicase from DNA polymerases during replication fork stalling [3]. This is followed by the ATR activation via two parallel pathways mediated by either ETAA1 or TOPBP1. ETAA1 stimulates ATR signaling by binding RPA on ssDNA, whereas TOPBP1 does so via its interaction with the RAD17 and 9-1-1 complexes at single-strand/double-strand DNA junctions [4, 5]. Once activated, ATR phosphorylates a wide variety of substrates, including CHK1[6]. ATR-CHK1 then promote the restart of stalled replication forks, which may e.g. have arisen by uncoupling of the replicative MCM helicase from DNA polymerases, while enabling the slowdown or arrest of cell cycle progression until replication is completed [2], and controlling the firing of new origins to preserve the RPA pool [3]. Another protective mechanism that ensures fork integrity is replication fork reversal. During this process, the nascent DNA strands on each sister chromatid are reannealed to form a fourth regressed arm, which becomes actively converted into a Holliday junction (HJ)-like structure [7]. Several key players are involved in fork reversal such as the SWI/SNF chromatin remodelers SMARCAL1 and ZRANB3, as well as the helicase HLTF [8]. Reversed forks are substrates for degradation by the nuclease MRE11 and must be properly protected by BRCA1, BRCA2 and Fanconi anemia (FA) proteins FANCA, FANCD2 and FANCJ to prevent fork collapse and ensure fork restart [9]. Alternatively, replication forks that stall at DNA lesions can restart downstream of the lesion by PrimPol-dependent re-priming on the leading strand, leaving unreplicated

ssDNA gaps to be filled post-replicatively by translesion synthesis (TLS) polymerases such as POL ŋ and REV1, or by template switching (TS), during which an undamaged homologous template on the sister chromatid is used to bypass DNA lesions [10].

Given the fact that DNA replication and transcription both utilize DNA, it is unavoidable that the two processes will interfere with each other, giving rise to transcription-replication conflicts (TRCs) [2]. TRCs can occur in a head-on (HO) or co-directional orientation (CD), depending on the orientation of the genes relative to the replication fork directionality. In mammalian cells the majority of replication forks progress through genes in a CD orientation to avoid the more deleterious HO collisions [11]. TRCs that occur in an HO orientation are frequently associated with the formation and persistence of R-loops. These are structures in which RNA is annealed to genomic DNA to create an RNA:DNA hybrid surrounded by a loop of non-templated single-stranded DNA which is protected by RPA [12, 13]. While R-Loops have several regulatory and topological roles in genome regulation, their persistence and association with TRCs can cause replication-stress associated genome instability [14]. However, although the network of R-loop regulators is expanding, a full understanding of the context in which R-loops arise and cause TRC-associated genome instability is lacking [15].

Chromatin factors have emerged as important players in R-loop formation and the associated genome instability [16]. These factors include the member of the MSL (male specific lethal) histone acetyltransferase (HAT) complex MOF, which has been reported to suppress replication fork stalling and R-loop formation [17]. Besides the MSL chromatin remodeling complex, MOF is also shared within the non-specific lethal (NSL) complex which consists of four core members KANSL1, KANSL2, KANSL3 and PHF20 [18]. The NSL complex was found to associate with the promotors of transcriptionally active genes, particularly housekeeping genes. [18], and regulates the expression of several genes essential for cell proliferation [19]. Also, more recently the loss of NSL complex members was associated with abnormalities in nuclear morphology and genome instability [20]. However, despite several studies linking the MSL and NSL chromatin remodeling complexes to transcription regulation and RS, the underlying molecular mechanism remains unclear.

Here we describe an important regulatory role of the NSL-complex member KANSL3 in the RS response. We found that KANSL3, unlike MOF, is not involved in the regulation of H4K16Ac in chromatin. KANSL3 protects cells from HU-induced RS and is implicated in the restart of stalled replication forks. We speculate that KANSL3 may suppress R-loop induced TRCs that lead to replication fork stalling and/or collapse.

RESULTS

The NSL complex member KANSL3 is dispensable for H4K16Ac levels

MOF forms the catalytic core of the NSL and MSL complex. [18]. As part of the MSL complex, MOF is the key lysine acetyltransferase (KAT) responsible for histone H4 lysine (K) 16 acetylation (H4K16Ac)-dependent decompaction of chromatin structure in mammalian cells (Fig. 1A and [20]. Since MOF is also part of the NSL complex, we asked whether this complex could also be involved in H4K16Ac. To examine this, we monitored the levels of H4K16Ac in chromatin-enriched extracts from cells depleted of the NSL complex members KANSL3 or MOF (Fig. 1A). As expected, depletion of MOF decreased the H4K16Ac levels in chromatin, while KANSL3 knockdown showed no effect (Fig. 1B). During the course of this work, several other studies also observed no change in H4K16Ac levels in KANSL3-depleted cells, suggesting distinct MSL- and NSL-complex characteristics [19, 20]. In fact, MOF acetylates lamin A/C as part of the NSL complex in order to maintain nuclear architecture [20]. Moreover, MOF induces H4K5Ac and H4K8Ac at transcription start sites (TSS) of genes involved in cell proliferation as part of the NSL-complex [19]. Together, our data in agreement with other studies suggests that the MSL and NSL complexes function in an H4K16Ac-dependent and -independent manner.
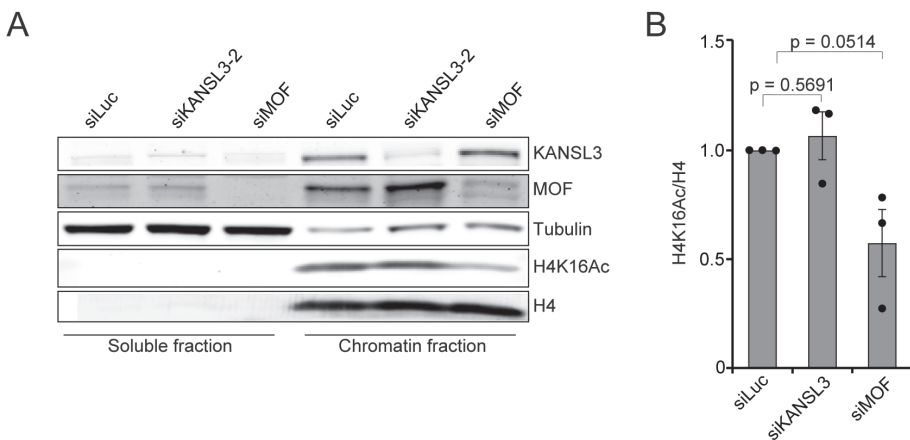
Figure 1. KANSL3 is dispensable for H4K16Ac. (A) Western blot analysis of the indicated proteins in soluble and chromatin fractions from U2OS cells. H4 is the loading control. A representative experiment is shown (B) Quantification of H4K16Ac levels on chromatin in cells from D. H4 is a loading control, which was used for normalization of H4K16Ac levels on chromatin. The mean ±SEM from 3 independent experiments is shown.

KANSL3 interacts with members of the NLS and MSL complex

To further unravel how the NSL complex may operate in a manner distinct from MSL, we asked whether KANSL3 has additional interaction partners besides its members in the NSL complex. To this end, we generated U2OS Flp-In/T-Rex cells stably expressing inducible GFP-tagged KANSL3 or GFP-NLS and performed GFP-trap based pulldowns followed by label-free mass spectrometry (MS). Our analysis revealed that KANSL3 interacts with 30 proteins that were at least 2-fold enriched in GFP-KANSL3 pulldowns compared to those of GFP-NLS (Fig. S1A and Table S1). Although, most interactors were known members of the NSL and MSL complexes (e.g. KANSL1, KANSL2 and KAT8), we also found additional interaction partners including NAT10, SUPT6H, SUPT16H, SCAF1 and PARP2 (Table S1). Since KANSL3's interaction partner MOF has previously been implicated in the response to RS [17], we performed this interaction analysis also in response to hydroxyurea (HU), which induces RS by dNTP depletion (Fig. S1B and Table S1). Interestingly, we found that KANSL3 interacts with PARP2, a protein that stabilizes RAD51 at stalled replication forks [21], with subunits of the RNA polymerase II complex (e.g. POLR2A, POLR2B and POLR2E), and with factors involved in RNA splicing (e.g. SRSF1 and SCAF1). Together, these results suggest that KANSL3 may interact with proteins involved in transcription and the RS response, manifesting a possible role of the NSL complex in these processes, although these interactions need to be confirmed.

KANSL3 promotes cell survival upon replication stress

To assess whether KANSL3, similar to MOF, is implicated in the RS response [17], we performed clonogenic cell survival assays in the presence and absence of HU. We found that KANSL3 loss rendered cells sensitive to HU, an effect that was similar to that observed after depletion of RPA-interacting protein ETAA1 (Fig. 2A-B). Cell cycle profiles remained unaffected in KANSL3-depleted cells, ruling out effects of cell cycle mis-regulation (Fig. S2A). Finally, we employed the Flp-In/T-Rex system to establish HCT116 cells stably expressing inducible and siRNA-resistant GFP-tagged KANSL3 (Fig. 2C). Confirming our previous data (Fig. 2A), we found that KANSL3 knockdown rendered cells sensitive to HU. Importantly, expression of siRNA-resistant GFP-KANSL3 almost fully rescued the HU sensitivity observed after KANSL3 knockdown (Fig. 2B), ruling out off-target effects of the siRNA against KANSL3.
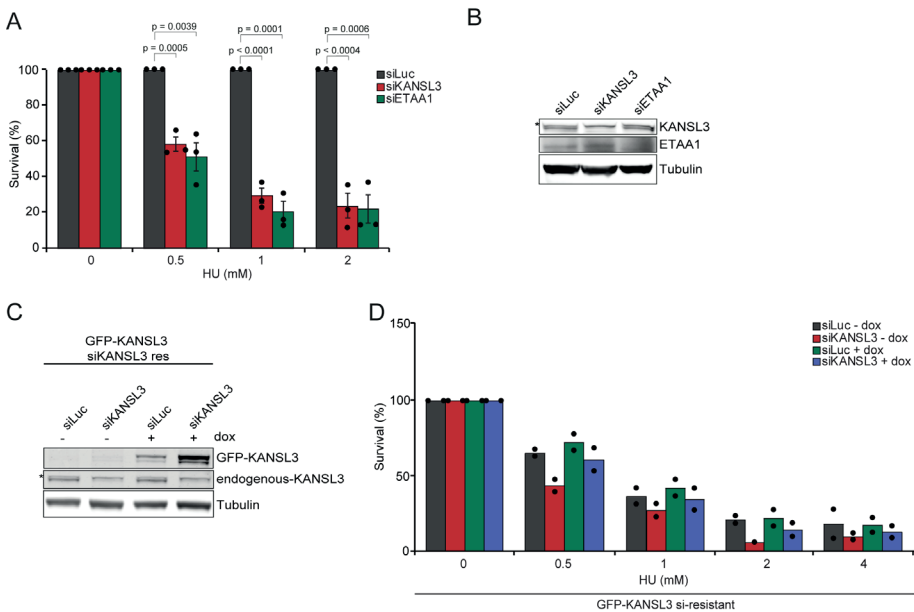
Figure 2. KANSL3-depleted cells are HU-sensitive. (A) Relative survival efficiency in HCT116 cells transfected with the indicated siRNAs and treated with 0.5, 1 and 2 mM HU. The mean ±SEM of 3 independent experiments is shown. Data were normalized to un-irradiated conditions and set to 100%. Statistical significance was calculated using Student's *t* test. (B) Western blot analysis of the expression of endogenous KANSL3 and ETAA1 from cells in A. Tubulin is a loading control. (C) Expression levels of endogenous KANSL3 and dox-inducible siRNA-resistant GFP-KANSL3 in HCT116 Flp-In/TRex cells. Tubulin is a loading control. (D) Effect of inducible expression of GFP-NLS and siRNA resistant GFP-tagged KANSL3 on the survival of stable HCT116 Flp-In/TRex after transfection with indicated siRNAs and treatment with indicated doses of ionizing irradiation (IR). The mean of 2 independent experiments is shown. Data was normalized to siLuc control which was set to 100%. Statistical significance was calculated using the Student's *t* test.

KANSL3 does not regulate replication fork speed and replication fork protection

To better understand how the absence of KANSL3 sensitizes cells to RS, we performed DNA fiber assays to monitor DNA replication in unperturbed conditions. We monitored progression rates of individual replication forks in control and KANSL3-depleted cells. For this, we sequentially labeled cells with CldU (green) and IdU (red), which was followed by tract length analysis. In line with previous results, PARPi treated cells showed an increase in fork speed [22]. However, KANSL3 depletion did not significantly impact tract lengths when compared to that in control cells (Fig. 3A). These data suggest that KANSL3 does not affect DNA replication in unperturbed conditions.

Next, we asked whether KANSL3 plays a role in stabilizing DNA replication forks under stressed conditions. In order to study this, siRNA-transfected U2OS cells were sequentially labeled with CldU and IdU, and on-going replication forks were then stalled by HU treatment. The relative shortening of the IdU tract length upon HU treatment served as a measure for fork degradation. As expected, BRCA2 knockdown cells displayed a significant reduction in IdU tract length compared to that in control cells [23]. In contrast, KANSL3-depleted cells displayed no significant reduction in IdU tract length (Fig. 3B). This suggests that KANSL3 is not involved in protecting stalled replication forks from degradation.
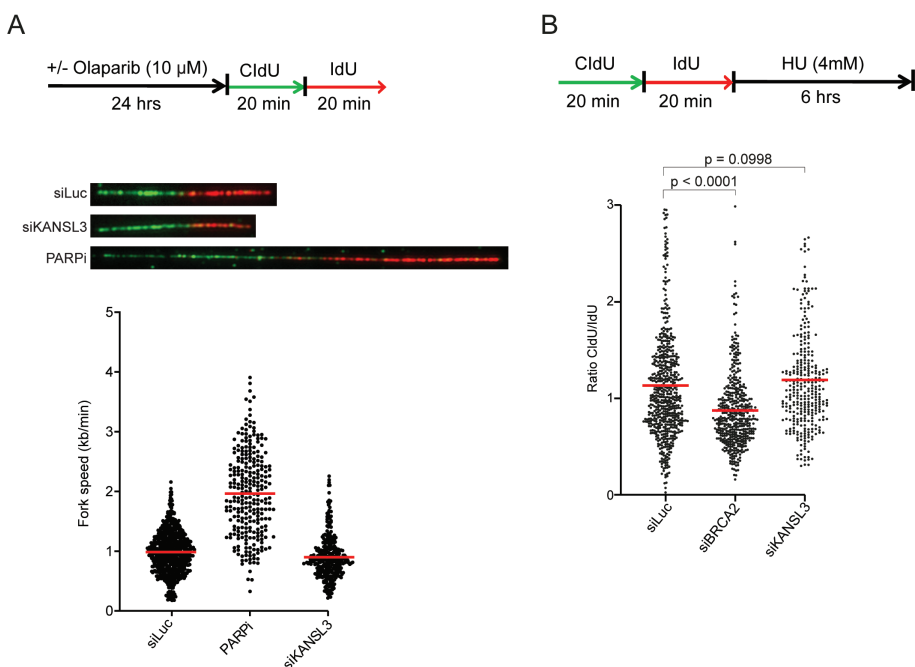
Figure 3. KANSL3 is dispensable for replication fork speed and replication fork protection.
(A) DNA fiber speed analysis in U2OS cells transfected with the indicated siRNAs or treated with 10 μM Olaparib (PARPi) for 24 hrs. Cells were labelled for 20 minutes with CldU (green), followed by labelling with IdU (red) for 20 minutes. Fork speed was calculated as labelled IdU length by pulse time. Red bars indicate the mean fork speed. Representative fibers of different conditions are shown. (B) Fork degradation analysis of siRNA transfected U2OS cells labelled as in A and treated with 4 mM HU for 6 hrs. Plotted values represent ratios of CldU versus IdU. Red bars indicate the mean fork speed. *P*-values were derived from Kruskal–Wallis ANOVA Dunn's post test.

KANSL3 promotes replication fork restart following replication fork stalling

Since the loss of MOF has been attributed to S-phase progression defects, as well as impaired fork restart upon RS [17], we aimed at further substantiating our findings by unraveling whether KANSL3 is also implicated in this process. To address this, we performed an ethynyl-2′-deoxyuridine (EdU) incorporation assay in which MOF or KANSL3-depleted cells were left untreated or treated with HU to stall replication forks, followed a release in medium containing EdU to allow and monitor replication restart (Fig. 4A). We found that control cells fully recovered from the HU treatment during the course of two hours, whereas KANSL3-depleted cells displayed reduced EdU incorporation at all timepoints. Strikingly, MOF depletion did not exert any effect on EdU incorporation (Fig. 4B). To our surprise, we also found KANSL3 depletion to reduce EdU staining in unperturbed conditions (Fig. 4B). Given that KANSL3-depleted cells did not display any change in fork speed, it may be possible that the reduced EdU incorporation observed in these cells, is due to a decrease in origin firing [24]. On another note, we also found that cell cycle profiles displayed a mild increase in the fraction of KANSL3-depleted cells in the G2/M phase (Fig. S2A). This may be indicative of an increase in the levels of under-replicated DNA in KANSL3-depleted cells, which could eventually trigger a G2/M cell cycle arrest.

Next, we sought to investigate whether the decrease in EdU incorporation observed in KANSL3-depleted cells (Fig. 4B) is due to defects in the restart of stalled replication forks. To this end, we performed DNA fiber assays in which siRNA-transfected U2OS cells were labeled with CldU, treated with HU, amd released into IdU. We found that HU-treated KANSL3-depleted cells, similar to RAD51-depleted cells, exhibited a decrease in the restart of stalled forks (Fig. 4C-D), the latter of which is in line with previous results [25].

To further substantiate these findings, we sought to address if KANSL3 is involved in ATR signaling following HU-induced helicase-polymerase uncoupling. ATR is the major kinase, which preserves the nuclear RPA pool by restraining the firing of new origins in order to prevent replication for instability. Inhibition of ATR in HU-treated cells leads to excessive double-strand break (DSB) formation and RPA hyperphosphorylation [3]. Consistent with a decrease in replication fork restart likely caused by replication fork instability, KANSL3-depleted cells showed increased levels of RPA2 phosphorylation at Serine 4 and 8 (Fig. 4E-F), as well as a mild increase in γH2AX levels (Fig. 4G). Importantly, RPA levels remained unchanged in HU-treated KANSL3-depleted cells (Fig. S2B-C), suggesting that KANSL3 loss results in the increase of stalled and collapsed replication forks upon HU-induced RS. Taken together, these results suggest that KANSL3 is important to maintain fork integrity under RS.
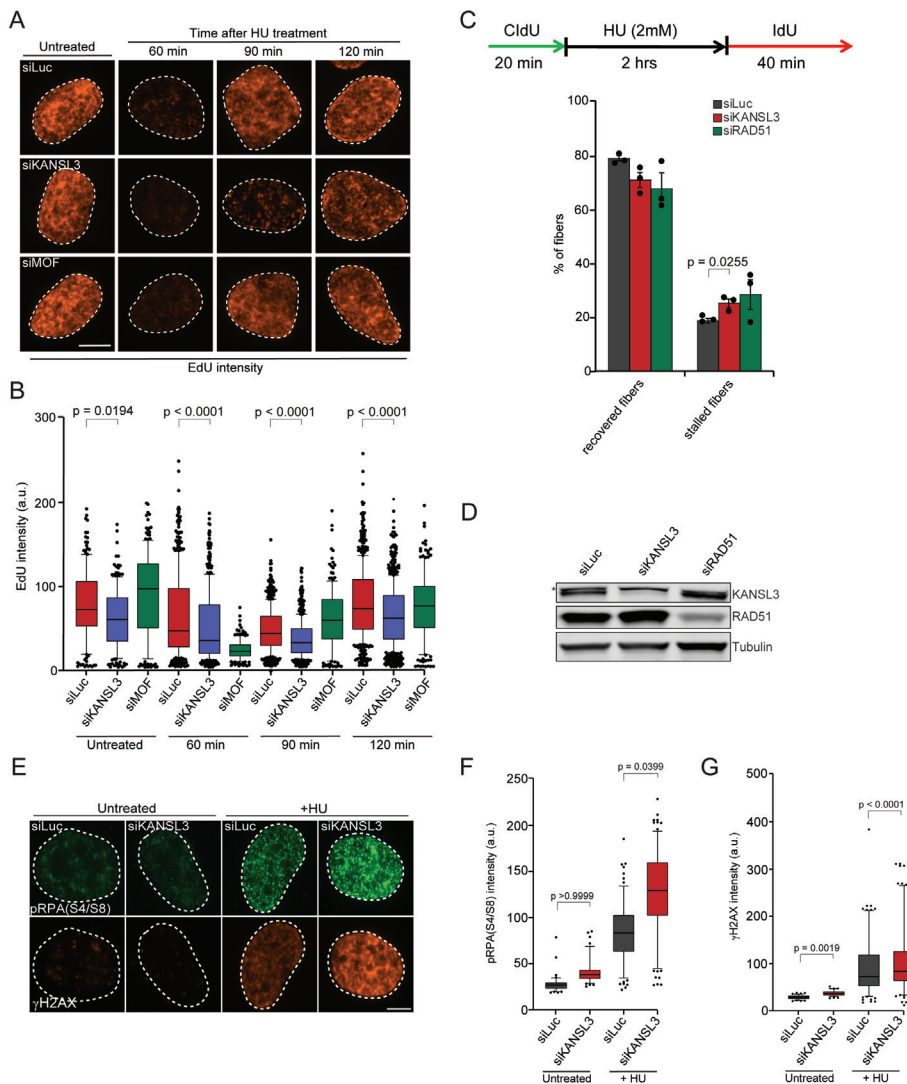
Figure 4. KANSL3 promotes recovery from HU-stalled replication forks. (A) Representative results of immunofluorescence microscopy of the BrdU incorporation assay. (B) The BrdU intensities of individual cells in the results in A are depicted in a boxplot. *P*-values were derived from Kruskal–Wallis ANOVA Dunn's post test. (C) Fork recovery analysis in U2OS cells transfected with the indicated siRNAs, followed by CldU labelling, 4 mM HU treatment for 6 hours and IdU labelling for 40 minutes. Fork progression (IdU-positive tracks) in U2OS cells transfected with the indicated siRNAs. Statistical significance was calculated using the Student's *t* test. (D) Western blot analysis of the expression of endogenous KANSL3 and RAD51 from cells in C. Tubulin is a loading control. (E) Phosphorylated RPA at Ser4/8 foci formation in U2OS cells transfected with the indicated siRNAs. Cells were exposed to 2 mM of HU and foci intensity was measured after 4 hours. (F) Quantification of RPA (S4/8) foci intensity in cells from E. Mean foci intensity was quantified in immuno-stained p-RPA (S4/8) cells. The mean of 3 independent experiments is shown. (F) As in E, except that γH2AX foci intensity was measured. Statistical significance was calculated using the Student's *t* test.

KANSL3-depleted cells suffer from aberrant RNA-DNA hybrid formation

Given the role of the NSL complex in regulating transcription [26], we hypothesized that transcription-replication conflicts (TRCs) might be responsible for the observed RS in KANSL3-depleted cells. TRCs are often associated with increased levels of RNA-DNA hybrid structures (R-loops). We therefore tested the impact of KANSL3 on R-loop formation using the S9.6 antibody for immunofluorescence, which mostly recognizes R-loops, but also has low affinity for dsRNA structures [27]. In order to validate our S9.6 immunofluorescence staining, we depleted a well-known R-loop processing factor AQR and found, as expected, that this led to an increase in the R-loop levels (Fig. 5A-B and [28]. Next, we measured the accumulation of R-loops in KANSL3- or MOF-depleted cells in both undamaged and HU-treated cells. While loss of KANSL3 severely increased the accumulation of R-loops under both conditions, MOF depletion did not, the latter of which being in contrast with a previous study, (Fig. 5A-B and [17]. Together, these results suggest that KANSL3, but not MOF, is involved in suppressing R-loop formation in unperturbed and RS conditions. We speculate that KANSL3 prevents the formation of TRCs by suppressing R-loop formation and the subsequent stalling/collapse of replication forks (Fig. 5C).
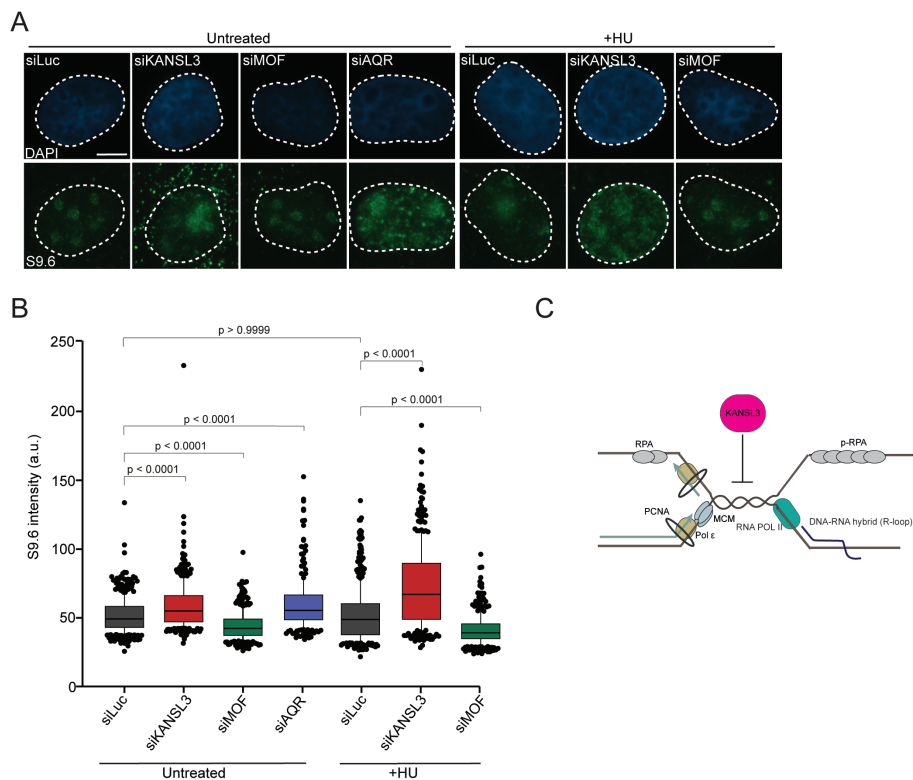
Figure 5. KANSL3 suppresses the formation of R-loops. (A) S9.6 nuclear intensity was measured in U2OS cells transfected with the indicated siRNAs. Cells were exposed to 2 mM of HU and foci intensity was measured after 4 hours. The S9.6 signal intensity per nucleus was determined by subtracting the S9.6 staining with that from nucleoli in each nucleus. (B) Quantification of S9.6 intensity in cells from A. Mean intensity was quantified in immuno-stained S9.6 cells. *P*-values were derived from Kruskal–Wallis ANOVA Dunn's post test. (C) Model for how KANSL3 works during RS and R-loops formation. KANSL3 suppresses the formation of aberrant R-loops, thereby preventing replication fork arrest/collapse as a result collisions between ongoing replication forks and the transcription machinery.

DISCUSSION

In this study, we uncover an important role of the NSL chromatin remodeling complex member KANSL3 in the RS response. First, we demonstrate that KANSL3 is not implicated in H4K16Ac formation, although it associates with both MSL and NSL complex members, including the HAT MOF, which is responsible for H4K16Ac. Second, we show that KANSL3 protects cells from HU-induced RS. Third, KANSL3 is not involved in replication fork speed or fork protection, but instead supports the restart

of HU-induced stalled replication forks. Finally, KANLS3 suppresses the formation of R-loops. Thus, KANSL3 may ensure fork stability by preventing TRC collisions that are causally linked to R-loop formation and induce replication fork collapse (Fig. 5C).

KANSL3 is dispensable for H4K16Ac and interacts with members of the NSL and MSL complex

The NSL and MSL complex are both MOF containing complexes with distinct functions. On one hand MOF catalyzes H4K16Ac as part of the MSL complex, while on the other hand it promotes cell survival and regulates the transcription of housekeeping genes as part of the NSL complex [19]. However, more recently, disruption of the MSL complex has been linked to H4K16Ac-independent chromosomal instability [29]. In agreement with previous studies we found the NSL complex member KANSL3 to be dispensable for H4K16Ac [20]. Another study reported that KANSL3 may function separately from other NSL and MSL complex members in maintaining the expression of several essential genes and cell survival [19]. Our mass spectrometry experiments revealed that KANSL3 interacts with known members of both the NSL and MSL complex. However, more interestingly, KANSL3 also interacts with subunits of the RNA polymerase II complex (e.g. POLR2A, POLR2B and POLR2E), factors involved in RNA splicing (e.g. SRSF1 and SCAF1) and replisome associated proteins (e.g. SMARCA5) [30] and Table S1). Moreover, KANSL3 was also reported to interact with PCNA upon camptothecin-induced RS [31]. This may suggest that KANSL3 maintains fork integrity: 1) indirectly via H4K16Ac independent transcription regulation, 2) via the suppression of R-loop-associated TRCs due to its interaction with the RNA polymerase II complex and splicing factors, or 3) its interaction with replisome associated proteins. However, further work will be needed to resolve the functional relevance of interplay between KANSL3 and its interaction partners in response to RS.

KANSL3 protects cells from HU-induced replication stress and regulates recovery of stalled replication forks

Our findings show that KANSL3-depleted cells are sensitive to HU-induced RS. We could exclude that the observed sensitivity was due to changes in replication fork speed and/or fork protection. Because HU treatment leads to the formation of stalled replication forks, we asked whether KANSL3 may be involved in their restart. Using EdU incorporation and DNA fiber assays, we confirmed that KANSL3 is implicated in the

restart of HU-stalled replication forks. Interestingly, EdU incorporation was reduced in KANSL3-depleted cells both in unperturbed and RS conditions. This suggests KANSL3 is also involved in normal DNA replication. Given the modest increase in the fraction of G2/M phase cells following KANSL3 depletion, it may be possible that KANSL-3-depeletd cells suffer from an increase in under-replicated DNA, presumably due to persistent R-loops, which may in tun result in a G2/M arrest, the latter which is in agreement with a previous study [19]. Alternatively, the decrease in EdU incorporation could also be indicative of decreased origin firing. However, further work will be needed to resolve whether KANSL3-depleted cells suffer from increased under-replicated DNA or have a decrease in the number of origins firing.

In response to HU-induced helicase-polymerase uncoupling, ATR kinase is recruited and activated via its direct interaction with ssDNA at stalled replication forks [32]. This is followed by the ATR-dependent phosphorylation of RPA (at Serine 33), which prevents fork breakage by the suppression of new origin firing, thereby preserving the nuclear pool of RPA [3]. Consequently, inhibition of ATR leads to excessive amounts of ssDNA and limits the available amount RPA molecules by a process referred to as ´RPA exhaustion´. This phenomenon in turn leads to replication fork catastrophies and the formation of DSBs [4, 32]. RPA levels, however, remained unaffected in KANSL3-depleted cells, ruling out RPA exhaustion. Instead, we found increased p-RPA (S4/S8) and γH2AX levels upon RS in KANSL3-depleted cells. This suggests that KANSL3 loss results in the conversion of stalled replication forks into collapsed forks, which may ultimately be converted into DSBs. Given the role of KANSL3 in transcription, the emerging question is whether these events depend on its role in transcription and the formation of R-loop dependent TRCs [33].

KANSL3, but not MOF depletion results in increased R-Loop formation in both HU-treated and untreated conditions

Our data reveal that KANSL3 depletion leads to increased R-loop levels in both unperturbed and RS conditions. However, MOF depletion resulted in reduced R-loops, which is contrary to a previous report [17]. This discrepancy may be explained due to different siRNAs used against MOF in the previous study compared to ours. We used the S9.6 antibody for immunofluorescence (IF) to measure R-loop levels. As expected, we found AQR depletion to increase the formation of R-loops, which validates our S9.6 IF to some extent [34]. However, because S9.6 also recognizes double-strand RNA [28,

35], it is of utmost importance to further validate our S9.6 IF experiments using RnaseIII treatment, treatment with recombinant wildtype human RNaseH1 and its catalytic inactive version, and overexpression of RNaseH1 to further validate the specificity of our S9.6 IF [35].

The unscheduled formation of R-loops has been linked to TRCs and RS [36]. This raises the question whether the aberrant R-loops observed in KANSL3-depleted cells, are a potential mechanism driving RS. Future endeavors including RNaseH1 overexpression rescue experiments in KANSL3-depleted cells will give insight into whether the increase in fork stalling measured by p-RPA (S4/8) and DNA breaks measured by γH2AX levels, as well as the HU-sensitivity observed in KANSL3-depleted cells is caused by R-loops. Furthermore, R-loops may pose an obstacle for replication fork progression and thereby cause defects in replication fork recovery upon HU-induced RS observed in KANSL3-depleted cells [37]. To investigate this possibility, proximity ligation assay (PLA) between PCNA, as a marker for replication, and RNAPII may evaluate the occurrence of TRCs in absence of KANSL3 [16].

The emerging questions concerning KANSL3's role in R-loops suppression is to decipher the genomic location in which R-loops are enriched. Are they specifically localized at genes regulated by the NSL complex or also at other genes? In depth analysis of R-loop distribution by DRIP-sequencing in both control and KANSL3-depleted cells will be required in the future to address these questions. Depending on the orientation of transcription relatively to the direction of the replication forks, TRCs can occur co-directionally or head-on [36]. Co-directional (CD) collisions can be resolved by displacement of RNA polymerase II from DNA, whereas head-on (HO) collisions induce pausing and blocking of the replication fork and may lead to fork collapse and DSBs [36]. Given that KANSL3-depleted cells display an increase in markers of RS-induced fork collapse (p-RPA (S4/S8) and γH2AX) , it may bind regions in the genome in which HO TRC occur. Future DRIP-seq (R-loops) and ChIP-seq (p-RPA (S4/S8) and γH2AX) experiments will provide insight into the context in which KANSL3 suppresses R-loop formation and may cause HO collisions [36]. Strikingly, our data revealed that KANSL3 depletion does not only lead to increased R-loops in the nucleus, but also in the cytosol. Cytosolic R-loops are associated with the activation of the innate immune response via the cGAS-STING pathway which regulates the expression of antiviral genes [38]. Future studies will be needed to decipher whether cytosolic R-loops in KANSL3-depleted cells activate the innate immune response and may pave the way for KANSL3's role in anti-tumor immunity [39].

Cooperation of KANSL3 and other NSL complex members

We reveal a critical role for KANSL3 in preventing the formation of R-loops associated RS. However, KANSL3 functions within the NSL-complex and is responsible for the stability and subsequent catalytic activity of the complex via OGT-mediated O-GlcNAcylation of KANSL3 [40]. KANSL1 is the major scaffolding protein within the complex and binds to MOF, PHF20 and MCRS1. Its interaction with MOF determines the catalytic activity of the NSL complex towards H4K16Ac *in vitro* [41]. Importantly, loss of both KANSL2 and KANSL3 is associated with nuclear abnormalities and chromothripsis [20]. Moreover, our KANSL3 interactome data reveals all known NSL-complex members in both unperturbed and RS conditions. Although it remains to be established whether other NSL-complex members are directly involved suppressing R-loop induced RS, these findings suggest that the majority of NSL proteins remain in a complex upon RS. On the one hand, the NSL complex may associate with the replisome via protein-protein interactions with replisome or replisome-associated proteins. On the other hand, given the role of the NSL-complex in transcription regulation and its interaction with members of the RNAPII complex and RNA splicing factors, NSL proteins may regulate the response to RS by controlling R-loop formation. Future mechanistic studies on the NLS protein repertoire will undoubtedly improve our understanding of their crucial role in diverse biological processes, including RS, thereby increasing our understanding of genome stability maintenance.

MATERIAL AND METHODS

Cell lines

U2OS and HCT116 cells were cultured in 5% CO_2 at 37°C in DMEM (Dulbecco's modified Eagle's medium) supplemented with 10% fetal calf serum and antibiotics. HCT116 Flp-In/T-Rex and U2OS Flp-In/T-Rex cells, which were generated using the Flp-In/T-Rex system (Thermo Fisher Scientific), were a gift of Bradley Wouters (Princess Margaret Cancer Centre, Canada) and Geert Kops (University Medical Center Utrecht, the Netherlands). These cells were used to stably express inducible versions of GFP-NLS, as well as siRNA-resistant GFP-KANSL3[WT] by co-transfection of pCDNA5/FRT/TO-Puro plasmid encoding GFP or GFP-tagged ZNF384 (WT or deletion mutants) (5 μg), together with pOG44 plasmid encoding the Flp recombinase (1 μg). After selection on 1 μg/mL puromycin, single clones were isolated and expanded. Both HCT116 Flp-In/T-Rex clones and U2OS Flp-In/T-Rex were incubated with 2 μg/mL doxycycline for 24h to induce expression of cDNAs. All cells were authenticated by STR profiling and tested negative in routinely performed mycoplasma tests.

Chemicals

Cells were treated with Hydroxyurea (Sigma) at a final concentration of 2 mM or 4 mM for the indicated timepoints and collected for further analysis. The PARP inhibitor olaparib (Selleck Chemicals) was used at a final concentration of 10 μM.

Transfections, siRNAs and plasmids

Cells were transfected with siRNAs using RNAiMAX (Invitrogen) according to the manufacturer's instructions. Cells were transfected once with siRNAs 24 hours after seeding at a concentration of 20 nM and analyzed 48 hours after transfection unless otherwise indicated. siRNA sequences are listed in Table 1. Cells were transfected with plasmid DNA using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions and analyzed 24-48 hours after transfection. The expression vector for full length human KANSL3 (64775: from Addgene; originally from Joan Conaway and Ronald Conaway), was amplified and cloned into pCDNA5/FRT/TO-Puro as a *HindIII/KpnI* fragment (Table 2). siKANSL3-resistant KANSL3 cDNA was generated by introducing the underlined mutations CGACGATAACCTTAGGATCAG by overlap PCR and cloned as *HindIII/KpnI* fragment into pCDNA5/FRT/TO-Puro-KANSL3-WT (Table 2). All KANSL3 expression constructs were verified using Sanger sequencing.

Immunofluorescence analysis

Cells were either directly fixed with 2% formaldehyde in PBS for 20 minutes at room temperature (RT), or pre-extracted with 0.5% Triton-X100 (Serva) in PBS on ice for 2-5 minutes prior to fixation. Alternatively, cells were fixed, post-extracted with 0.25% Triton-X100 (Serva) in PBS and treated with 100 mM glycine in PBS for 20 minutes to block unreacted aldehyde groups. Cells were then rinsed with PBS and equilibrated in wash buffer (PBS containing 0.5% BSA). Antibody incubation steps and washes were in wash buffer. Primary antibodies were incubated for 1-2 hours at room temperature. Detection was done using goat anti-mouse or goat anti-rabbit Ig coupled to Alexa 488, 555 or 647 (1:1500; Invitrogen Molecular probes). All antibodies are listed in Table 3. Samples were incubated with 0.1 μg/mL 4',6-Diamidino-2-Phenylindole Dihydrochloride (DAPI) and mounted in Polymount.

Sample preparation and mass spectrometry

For mass spectrometry, U2OS Flp-In/T-Rex cells expressing GFP-NLS and GFP-KANSL3[WT] were treated with 2 mM Hydroxyurea for 4 hours or left untreated. Cell pellets were lysed in EBC-1 buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.5% NP-40, 2 mM $MgCl_2$, protease inhibitor cocktail tablets) with 500 units benzonase. Samples were incubated for 1 hour at 4ºC under constant mixing followed by high-speed centrifugation for 10 minutes at 4ºC. Protein concentration was measured by Qubit in the cleared lysates, equalized and transferred to tubes containing GFP-Trap beads (Chromotek). After 90 minutes of incubation at 4ºC under rotating condition, the beads were washed 4 times with EBC-2 buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, and protease inhibitor cocktail tablets) and 3 times with 50 mM ammonium bicarbonate followed by overnight digestion using 2.5 μg trypsin at 37°C under constant shaking. Digestion was terminated with 1% trifluoroacetic acid and centrifuged for 5 minutes at high speed to precipitate insoluble fractions. Consequently, C18 cartridges were prepared by washing 2 times with acetonitrile followed by 2 times with 0,1% Acetic acid. Peptides were loaded on the cartridge, while bound peptides were washed 2 times with 0.1% acetic acid and eluted with 1 mL 80% acetonitrile / 0.1% acetic acid and lyophilized.

Mass spectrometry was performed essentially as previously described [42]. Samples were nalysed on a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific, Germany) coupled to an EASY-nanoLC 1000 system (Proxeon, Odense, Denmark). Digested peptides were separated using a 15 cm fused silica capillary (ID: 75 μm, OD: 375 μm, Polymicro Technologies,California, US) in-house packed with 1.9 μm C18-AQ beads (Reprospher-DE, Pur, Dr. Maisch, Ammerburch-Entringen, Germany). Peptides were separated by liquid chromatography using a gradient from 2% to 95% acetonitrile with 0.1% formic acid at a flow rate of 200 nl/minute for 65 minutes. The mass spectrometer was operated in positive-ion mode at 2.8 kV with the capillary heated to 250°C, and in a Data-Dependent Acquisition (DDA) mode with a top 7 method. Full scan MS spectra were obtained with a resolution of 70,000, a target value of $3x10^6$ and a scan range from 400 to 2,000 m/z. Maximum Injection Time (IT) was set to 50 ms. Higher-Collisional Dissociation (HCD) tandem mass spectra (MS/MS) were recorded with a resolution of 35,000, a maximum IT of 20 ms, a target value of $1x10^5$ and a normalized collision energy of 25%. The precursor ion masses selected for MS/MS analysis were subsequently dynamically excluded from MS/MS analysis for 60 seconds. Precursor ions with a charge state of 1 and greater than 6 were excluded from triggering MS/MS events. Three replicates were included per condition with two technical repeats each.

Mass spectrometry data analysis

Raw mass spectrometry files were analysed with MaxQuant software (v1.5.3.30) as described [43] with the following modifications from default settings: the maximum number of mis-cleavages by trypsin/p was set to 3, Label Free Quantification (LFQ) was enabled disabling the Fast LFQ feature. Match-between-runs feature was enabled with a match time window of 0.7 minutes and an alignment time window of 20 minutes. We performed the search against an in silico digested UniProt reference proteome for Homo sapiens (8th June 2020). Analysis output from MaxQuant was further processed in the Perseus (v 1.5.5.3) computational platform [43]. Proteins identified as common contaminants, only identified by site and reverse peptide, were filtered out, and then all the LFQ intensities were log2 transformed. Different biological repeats of each condition were grouped and only protein groups identified in all three biological replicates in at least one condition were included for further analysis. Missing values were imputed using Perseus software by normally distributed values with a 1.8 downshift (log2) and a randomized 0.3 width (log2) considering total matrix values. Volcano plots were generated, and Student's T-tests were performed to compare the different conditions. Spreadsheets from the statistical analysis output from Perseus were further processed in Microsoft Excel for comprehensive visualization and analysis of the data (Table S1).

Western blot analysis

Cells were lysed in 2x Laemmli buffer and proteins were separated by Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis (SDS-PAGE) using 4-12% pre-cast polyacrylamide gels (BioRad or Invitrogen) and MOPS running buffer (Invitrogen). Next, proteins were transferred onto nitrocellulose membranes (Millipore). Protein expression was analyzed by immunoblotting with the indicated primary antibodies (Table 3) and secondary CF680 goat anti-rabbit or CF770 goat anti-mouse Ig antibodies (1:5000, Biotium). Membranes were scanned and analyzed using a Licor Odyssey scanner (LI-COR Biosciences).

Cell survival assays

HCT116 cells were transfected with siRNAs, trypsinized, seeded at low density and exposed to HU for 24 hours. For HCT116 Flp-In/T-Rex, cDNAs were expressed by adding Dox for 24 hours after siRNA transfection. After 12 days, the cells were washed with 0.9% NaCl and stained with methylene blue (2.5 g/L in 5% ethanol, Sigma-Aldrich). Colonies of more than 20 cells were scored.

Cell cycle profiling

Cells were fixed in 70% ethanol, followed by DNA staining with 50 µg/mL propidium iodide in the presence of Rnase A (0.1 mg/mL; Sigma). Cell acquisition and quantification was performed on a BD LSRII flow cytometer (BD Bioscience) using FACSDiva software version 5.0.3.

Chromatin fractionation

Chromatin fractionation was based on a previously published protocol [44] and used with modifications. Briefly, 100.000-150.000 cells were grown per 6-cm dish for 24 h and then transfected with siRNAs. Next, the cells were treated with 500 µM phleomycin for 1 h, washed three times with PBS, and incubated in NETN extraction buffer (100mM NaCl, 1mM EDTA, 20mM Tris-Cl pH8, 0,5% NP-40 + Proteasome inhibitors). After 15 minutes of incubation on ice, samples were taken for the chromatin-unbound fraction and mixed with the same amount of 2x Laemmli buffer. Cells were washed with PBS, lysed, and incubated in Laemmli buffer with benzonase for 15 min to obtain the chromatin-bound fraction. Samples were heated for 7 min at 80°C and subjected to Western blot analysis.

DNA fiber assays

DNA fibers were prepared as described previously [45] with minor modifications. Briefly, siRNA transfected U2OS cells were pulse labelled with 10 µM 5-chloro-2-deoxyuridine (CldU) for 20 min, followed by 250 µM 5-iodide-2-deoxyuridine (IdU) chase for 20 min. Next, cells were washed twice with PBS, trypsinized and resuspended in PBS containing 10% FBS. 5 µl Cell suspension was spotted on the top of a Superfrost glass slide (Fisher Scientific) and dried for 5 min. Subsequently, 7 µl lysis buffer (200 mM Tris-HCl [pH 7.4]; 50 mM EDTA; 0.5% SDS) was added in the middle of the cell suspension, mixed by stirring and incubated for 3 min. Slides were tilted slightly at an angle to induce slow running of a stream of DNA down the slide, air dried, and fixed in methanol:acetic acid (3:1) for 10 min. DNA fibers were denatured with 2.5 M HCl for 80 min, washed three times with PBS and blocked twice with TPB (1% BSA in PBS with 0.1% Tween) for 15 min. Labelled DNA fibers were then stained with rat anti-BrdU antibody (Clone Bu1/75; for CldU detection; Abcam) and mouse anti-BrdU antibody (clone B44; for IdU detection; Abcam) in TPB for 1 hr at room temperature. After incubation, slides were washed thrice with PBS and three times with TPB, followed by fixation in 2% formaldehyde for 10 min.

The microscope slides were then incubated with Alexa Fluor 488 donkey anti-rat antibody (1:1000) and Alexa Fluor 555 goat anti-mouse antibody (1:1000) in TPB for 1.5 hr at room temperature in the dark. Next, the slides were washed three times with PBS, air-dried and mounted with aqua polymount (Polysciences). Fluorescent images were acquired, and track lengths were measured manually using ImageJ software. For fork degradation, cells were incubated with 20 µM CldU for 20 min, followed by 25 µM IdU chase for 20 min, and finally 4 mM hydroxyurea treatment for 5 hr, before trypsinization and slide preparation as described above. For fork recovery, cells were labelled with 20 µM CldU for 20 min, followed by 4 mM hydroxyurea treatment for 5 hr, and finally 250 µM IdU labelling.

EdU incorporation assay

U2OS cells were transfected with siRNAs and 48 hours later left untreated or treated with 2 mM HU for 2 hours. After washing, the cells were further incubated in medium containing 5-ethynyl-2′-deoxyuridine (EdU) for 15 min, 60 min, 90 min and 2 hours and pre-extracted with 0.5% ice-cold Triton-X-100 (Serva) in PBS and fixed with 2% formaldehyde in PBS for 20 minutes at room temperature (RT). Next, cells were permeabilized with 0.25% Triton X-100 in PBS for 5 min at room temperature, followed by incubation with click-iT reaction buffer (10 µM biotin azide (life technologies), 10 mM sodium-L-ascorbate, and 2 mM $CuSO_4$) for 30 min in at RT. Cells were then rinsed with PBS and equilibrated in wash buffer (PBS containing 0.5% BSA. Samples were incubated with 0.1 µg/mL 4',6-Diamidino-2-Phenylindole Dihydrochloride (DAPI) and mounted in Polymount.

AUTHOR CONTRIBUTIONS

J.S. performed western blot analysis, immunofluorescence experiments, clonogenic survival assays. J.L. and W.W.W. performed DNA fiber assay. A.J.L.G. prepared mass-spectrometry samples. R.P. and A.C.O.V. performed mass-spectrometry experiments and data analysis. H.v.A. conceived and supervised the project.

ACKNOWLEDGEMENTS

REFERENCES

1. Gaillard, H., T. García-Muse, and A. Aguilera, *Replication stress and cancer.* Nature Reviews Cancer, 2015. **15**(5): p. 276-289.

2. Zeman, M.K. and K.A. Cimprich, *Causes and consequences of replication stress.* Nat Cell Biol, 2014. **16**(1): p. 2-9.

3. Toledo, L.I., et al., *ATR prohibits replication catastrophe by preventing global exhaustion of RPA.* Cell, 2013. **155**(5): p. 1088-103.

4. Haahr, P., et al., *Activation of the ATR kinase by the RPA-binding protein ETAA1.* Nature Cell Biology, 2016. **18**(11): p. 1196-1207.

5. Bass, T.E., et al., *ETAA1 acts at stalled replication forks to maintain genome integrity.* Nat Cell Biol, 2016. **18**(11): p. 1185-1195.

6. Zhao, H. and H. Piwnica-Worms, *ATR-mediated checkpoint pathways regulate phosphorylation and activation of human Chk1.* Molecular and cellular biology, 2001. **21**(13): p. 4129-4139.

7. Neelsen, K.J. and M. Lopes, *Replication fork reversal in eukaryotes: from dead end to dynamic response.* Nature Reviews Molecular Cell Biology, 2015. **16**(4): p. 207-220.

8. Liao, H., et al., *Mechanisms for stalled replication fork stabilization: new targets for synthetic lethality strategies in cancer treatments.* EMBO Rep, 2018. **19**(9).

9. Schlacher, K., H. Wu, and M. Jasin, *A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2.* Cancer Cell, 2012. **22**(1): p. 106-16.

10. Quinet, A., et al., *To skip or not to skip: choosing repriming to tolerate DNA damage.* Molecular Cell, 2021. **81**(4): p. 649-658.

11. Lin, Y.-L. and P. Pasero, *Transcription-Replication Conflicts: Orientation Matters.* Cell, 2017. **170**(4): p. 603-604.

12. Nguyen, H.D., et al., *Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1.* Mol Cell, 2017. **65**(5): p. 832-847.e4.

13. Gaillard, H. and A. Aguilera, *Transcription as a Threat to Genome Integrity.* Annu Rev Biochem, 2016. **85**: p. 291-317.

14. Crossley, M.P., M. Bocek, and K.A. Cimprich, *R-Loops as Cellular Regulators and Genomic Threats.* Mol Cell, 2019. **73**(3): p. 398-411.

15. Rinaldi, C., et al., *Sensing R-Loop-Associated DNA Damage to Safeguard Genome Stability.* Frontiers in Cell and Developmental Biology, 2021. **8**(1657).

16. Bayona-Feliu, A., et al., *The SWI/SNF chromatin remodeling complex helps resolve R-loop-mediated transcription–replication conflicts.* Nature Genetics, 2021.

17. Singh, D.K., et al., *MOF Suppresses Replication Stress and Contributes to Resolution of Stalled Replication Forks.* Mol Cell Biol, 2018. **38**(6).

18. Sheikh, B.N., S. Guhathakurta, and A. Akhtar, *The non-specific lethal (NSL) complex at the crossroads of transcriptional control and cellular homeostasis.* EMBO Rep, 2019. **20**(7): p. e47630.

19. Radzisheuskaya, A., et al., *Complex-dependent histone acetyltransferase activity of KAT8 determines its role in transcription and cellular homeostasis.* Mol Cell, 2021. **81**(8): p. 1749-1765.e8.

20. Karoutas, A., et al., *The NSL complex maintains nuclear architecture stability via lamin A/C acetylation.* Nature Cell Biology, 2019. **21**(10): p. 1248-1260.

21. Ronson, G.E., et al., *PARP1 and PARP2 stabilise replication forks at base excision repair intermediates through Fbh1-dependent Rad51 regulation.* Nature Communications, 2018. **9**(1): p. 746.

22. Maya-Mendoza, A., et al., *High speed of fork progression induces DNA replication stress and genomic instability.* Nature, 2018. **559**(7713): p. 279-284.

23. Mijic, S., et al., *Replication fork reversal triggers fork degradation in BRCA2-defective cells.* Nature Communications, 2017. **8**(1): p. 859.

24. Iyer, D.R. and N. Rhind, *Replication fork slowing and stalling are distinct, checkpoint-independent consequences of replicating damaged DNA.* PLoS Genet, 2017. **13**(8): p. e1006958.

25. Petermann, E., et al., *Hydroxyurea-Stalled Replication Forks Become Progressively Inactivated and Require Two Different RAD51-Mediated Pathways for Restart and Repair.* Molecular Cell, 2010. **37**(4): p. 492-502.

26. Lam, K.C., et al., *The NSL complex-mediated nucleosome landscape is required to maintain transcription fidelity and suppression of transcription noise.* Genes Dev, 2019. **33**(7-8): p. 452-465.

27. Crossley, M.P., et al., *Catalytically inactive, purified RNase H1: A specific and sensitive probe for RNA-DNA hybrid imaging.* J Cell Biol, 2021. **220**(9).

28. Sollier, J., et al., *Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability.* Molecular Cell, 2014. **56**(6): p. 777-785.

29. Monserrat, J., et al., *Disruption of the MSL complex inhibits tumour maintenance by exacerbating chromosomal instability.* Nature Cell Biology, 2021. **23**(4): p. 401-412.

30. Ribeyre, C., et al., *Nascent DNA Proteomics Reveals a Chromatin Remodeler Required for Topoisomerase I Loading at Replication Forks.* Cell Reports, 2016. **15**(2): p. 300-309.

31. Srivastava, M., et al., *Replisome Dynamics and Their Functional Relevance upon DNA Damage through the PCNA Interactome.* Cell Reports, 2018. **25**(13): p. 3869-3883.e4.

32. Liao, H., et al., *Mechanisms for stalled replication fork stabilization: new targets for synthetic lethality strategies in cancer treatments.* EMBO reports, 2018. **19**(9): p. e46263.

33. Gaub, A., et al., *Evolutionary conserved NSL complex/BRD4 axis controls transcription activation via histone acetylation.* Nature Communications, 2020. **11**(1): p. 2243.

34. Sakasai, R., et al., *Aquarius is required for proper CtIP expression and homologous recombination repair.* Scientific Reports, 2017. **7**(1): p. 13808.

35. Crossley, M.P., et al., *Catalytically inactive, purified RNase H1: A specific and sensitive probe for RNA–DNA hybrid imaging.* Journal of Cell Biology, 2021. **220**(9).

36. García-Muse, T. and A. Aguilera, *Transcription–replication conflicts: how they occur and how they are resolved.* Nature Reviews Molecular Cell Biology, 2016. **17**(9): p. 553-563.

37. Gómez-González, B. and A. Aguilera, *Transcription-mediated replication hindrance: a major driver of genome instability.* Genes Dev, 2019. **33**(15-16): p. 1008-1026.

38. Mankan, A.K., et al., *Cytosolic RNA:DNA hybrids activate the cGAS-STING axis.* Embo j, 2014. **33**(24): p. 2937-46.

39. Li, T. and Z.J. Chen, *The cGAS–cGAMP–STING pathway connects DNA damage to inflammation, senescence, and cancer.* Journal of Experimental Medicine, 2018. **215**(5): p. 1287-1299.

40. Wu, D., et al., *O-Linked N-acetylglucosamine transferase 1 regulates global histone H4 acetylation via stabilization of the nonspecific lethal protein NSL3.* J Biol Chem, 2017. **292**(24): p. 10014-10025.

41. Li, X., et al., *Two mammalian MOF complexes regulate transcription activation by distinct mechanisms.* Mol Cell, 2009. **36**(2): p. 290-301.

42. Kumar, R., et al., *The STUbL RNF4 regulates protein group SUMOylation by targeting the SUMO conjugation machinery.* Nat Commun, 2017. **8**(1): p. 1809.

43. Tyanova, S., T. Temu, and J. Cox, *The MaxQuant computational platform for mass spectrometry-based shotgun proteomics.* Nat Protoc, 2016. **11**(12): p. 2301-2319.

44. Feng, W., et al., *TRAIP regulates replication fork recovery and progression via PCNA.* Cell Discov, 2016. **2**: p. 16016.

45. Nieminuszczy, J., R.A. Schwab, and W. Niedzwiedz, *The DNA fibre technique - tracking helicases at work.* Methods, 2016. **108**: p. 92-8.

Table 1. List of siRNAs

Target	Sequence (5'-3')
MOF	GUGAUCCAGUCUCGAGUGAUU
KANSL3	UGAUGACAAUCUCAGAAUA
RAD51	GAGCUUGACAAACUACUUC
BRCA2	GAAGAAUGCAGGUUUAAUA
ETAA1	GAGCAAAACAAGAGGAAUUUU
Luciferase	CGUACGCGGAAUACUUCGA
AQR	SMARTpool (siGENOME)

Table 2. List of primers

Name	Sequence (5'-3')
HindIII FW	TAAAAAGCTTATATGGAAGAATCTCACTTCAATTCTAAC
KpnI RV	TAATAATGGTACCCTAAGAGCTGGCCAGGTGC
KANSL3_siRNAres_HindII Rv	TAATAATGGTACCTCAGGGTGCTGGAGGC
KANSL3_siRNAres_KpnI FW	TAAAAAGCTTATATGGCCCACCGGGGTGG
M13 FW	GTAAAACGACGGCCAGT
M13 RV	CAGGAAACAGCTATGAC

Table 3. List of primary antibodies

Protein	Host	Company	IF	WB
GFP	Mouse	Roche (11814460001)		1:1000
α-Tubulin	Mouse	Sigma (cloneDM1A, T6199)		1:5000
KANSL3	Rabbit	ATLAS (HPA030358)		1:2000
RAD51	Rabbit	BioAcademia		1:15000
H4K16Ac	Rabbit	Active Motif (39930)		1:2000
MOF	Rabbit	Bethyl Laboratories		1:1000
H4	Mouse	Abcam (ab31810)		1:2000
p-RPA (S4/S8)	Rabbit	Bethyl Laboratories (A300-245A)	1:1000	1:1000
ETAA1	Rabbit	Kind gift of Niels Mailand		1:400
RPA	Mouse	Abcam (ab2175)	1:1000	1:1000

SUPPLEMENTARY FIGURES AND TABLES
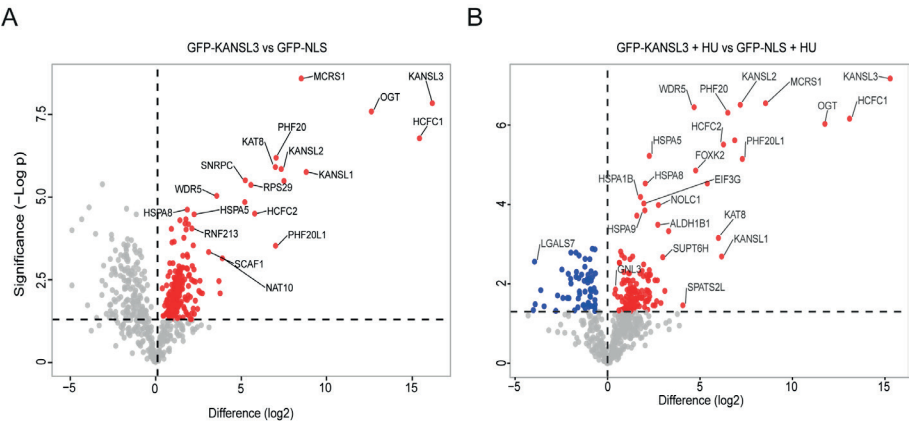
Supplementary Figure 1. – related to Figure 1. KANSL3 interacts with members of the MSL and NSL complex. (A) Volcano plot depicting the statistical differences of the MS analysis on GFP-KANSL3 versus GFP-NLS pulldowns. The enrichment is plotted on the x-axis and the significance (t-test -log2 p-value) is plotted on the y-axis (see also Supplementary Table 1). (B) As in A, except that cells were treated with 2 mM HU for 4 hours before pulldowns.
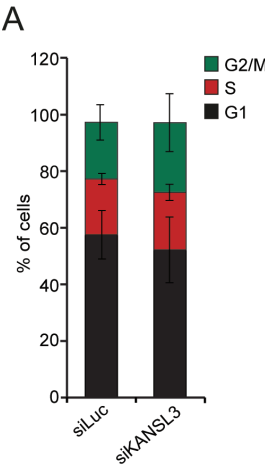
Supplementary Figure 2. – related to Figure 2. Cell cycle profiles remain unaffected in KANSL3-depleted cells. (A) Cell cycle profile of the indicated siRNA-transfected knockdown U2OS cells. The fraction of G1-, S- and G2-phase cells was determined by propidium iodide staining and FACS analysis.

4
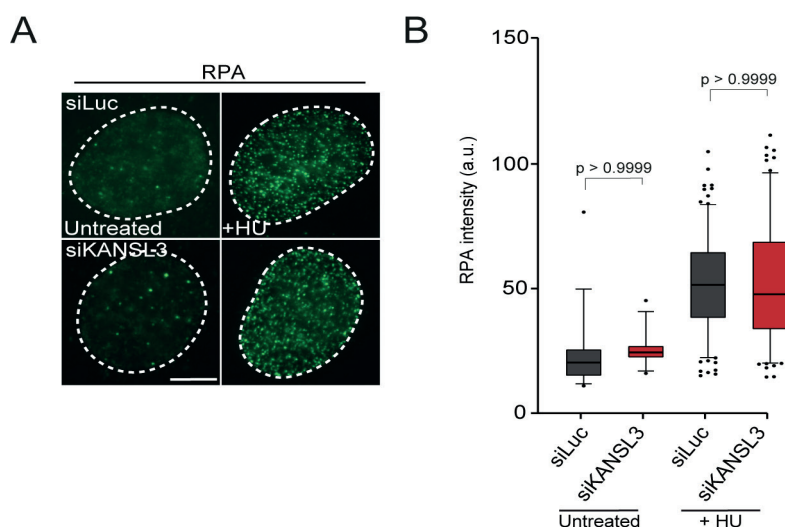
Supplementary Figure 3. – related to Figure 3. RPA foci intensity remains unaffected in KANSL3-depleted cells. (A) Cell cycle profile of the indicated siRNA-transfected knockdown U2OS cells. The fraction of G1-, S- and G2-phase cells was determined by propidium iodide staining and FACS analysis. (B) RPA foci formation in U2OS cells transfected with the indicated siRNAs. Cells were exposed to 2 mM of HU and foci intensity was measured after 4 hours. (F) Quantification of RPA foci intensity in cells from A. Mean foci intensity was quantified in immunostained RPA cells.

Supplementary Table 1

Supplementary Table 1. Statistical analysis comparing LFQ intensities of proteins identified by LC-MS/MS in GFP-pulldowns from U2OS Flp-InT-Rex cells expressing either GFP-KANSL3 or GFP-NLS in the presence and absence of HU. -Log p-value and average fold-enrichment (log2) are indicated. Additional information from the analysis of identified proteins is also included.

Gene			KANSL3 vs GFP		KANSL3 vs GFP (HU)											Miscellaneous										Values												

(Dense multi-column data table)

4

4

CHAPTER 5

Characterization of the exonuclease ERI1 in replication fork stability and R-loop formation

Jenny Kaur Singh, Wouter W. Wiegant, Magdalena B. Rother, Jolanthe Lingeman, Kees Vreeken, Anton J.L. de Groot, Román González-Prieto, Alfred C.O. Vertegaal, and Haico van Attikum.

ABSTRACT

The DNA in our cells is particularly vulnerable in S-phase, as many types of endogenous and exogenous sources can cause lesions that block replication fork progression and impact cell division. It is therefore essential for cells to activate replication stress pathways that stabilize and rescue stalled replication forks, allowing the completion of faithful DNA replication. Here, we describe the RNA processing factor ERI1 to play a role in the replication stress response. ERI1 depletion causes an increase in the levels of DNA-RNA hybrids (R-loops), which may enhance replication fork stalling and collapse. Indeed, ERI1 knockdown cells display increased phospho-RPA (S4/S8) levels after hydroxyurea (HU), which is indicative of replication fork instability. Accordingly, we found that ERI1 loss renders cells hypersensitive to HU-induced replication fork damage. Finally, we found that ERI1 is recruited to (stalled) replication forks, suggesting that ERI1 plays a direct role in the replication stress response. Future research will investigate how ERI1 functions at stalled replication forks and whether its function at these structures relies on its role in RNA-processing.

Keywords: Replication stress (RS), R-loops, ERI1, RNA processing

INTRODUCTION

DNA replication is a tightly regulated process in which the replisome rapidly copies billions of DNA bases during each cell division. This process is, however, constantly challenged by replication stress (RS) [1], which is defined as the slowing or stalling of replication fork progression and/or DNA synthesis, and can be caused by a multitude of impediments including limiting nucleotides, ribonucleotide incorporation, repetitive DNA elements, transcription-replication conflicts, DNA secondary structures, fragile sites, and oncogene-induced stress [2]. Fortunately, cells have evolved sophisticated S-phase and mitotic checkpoint pathways that help DNA replication to complete upon RS. Defects in these checkpoint pathways can lead to the stalling or collapse of replication forks, which in turn can cause mutations, chromosome rearrangements or the missegregation of chromosomes, thereby contributing to the development of diseases such as cancer [1].

ATR is the major kinase involved in dealing with RS. It is is recruited to RPA-coated single-stranded DNA (ssDNA) generated by uncoupling of the replicative MCM helicase from DNA polymerases at stalled replication forks, or for instance at replication-dependent DSBs induced by the topoisomerase I inhibitor camptothecin (CPT) [3, 4]. This is followed by its activation via two parallel pathways mediated by either ETAA1 or TOPBP1. ETAA1 stimulates ATR signaling by binding RPA on ssDNA, whereas TOPBP1 does so via its interaction with the RAD17 and 9-1-1 complexes at single-strand/double-strand DNA junctions [5, 6]. Once activated, ATR phosphorylates a wide variety of substrates, including CHK1 [7]. ATR-CHK1 then promote the restart of stalled replication forks, while enabling the slowdown or arrest of cell cycle progression until replication is completed [2], and controlling the firing of new origins to preserve the RPA pool [3]. Another failsafe mechanism that ensures fork integrity is replication fork reversal. During this process, annealing of newly synthesized and parental DNA occurs to form a fourth regressed arm, which becomes actively converted into a Holliday junction (HJ)-like structure [8]. Several key players are involved in fork reversal such as the SWI/SNF chromatin remodelers SMARCAL1 and ZRANB3, the helicase HLTF, as well as PARP1. PARP1 stabilizes reversed forks by inhibiting the activity of ATP-dependent DNA helicase Q1 (RECQ1), which can otherwise restart the fork [9]. Reversed forks are substrates for degradation by the nuclease MRE11 and must be properly protected by BRCA1, BRCA2 and Fanconi anemia (FA) proteins FANCA, FANCD2 and FANCJ to prevent fork collapse and ensure fork restart [10]. Alternatively, replication forks

that stall at DNA lesions can restart downstream of the lesion by PrimPol-dependent re-priming on the leading strand, leaving unreplicated ssDNA gaps to be filled post-replicatively by translesion synthesis (TLS) polymerases such as POL ŋ and REV1, or by template switching (TS), during which an undamaged homologous template on the sister chromatid is used to bypass DNA lesions [11].

Since DNA replication and transcription both utilize DNA, it is unavoidable that the two processes will interfere with each other, giving rise to transcription-replication conflicts (TRCs) [2]. TRCs can occur in a head-on (HO) or co-directional orientation (CD), depending on the orientation of the genes relative to the replication fork directionality. In mammalian cells the majority of replication forks progress through genes in a CD orientation to avoid the more deleterious HO collisions [12]. However, TRCs that occur in an HO orientation are frequently associated with the formation and persistence of R-loops. These are structures in which RNA is annealed to genomic DNA to create an RNA:DNA hybrid surrounded by a loop of non-templated single-stranded DNA which is protected by RPA [13, 14]. While R-Loops have several regulatory and topological roles in genome regulation, their persistence and association with TRCs can cause replication-stress associated genome instability [15].

RNA binding proteins have emerged as important players in preventing R-loops and genome stability [16]. RnaseH1 and RNaseH2 are evolutionary conserved enzymes that specifically degrade the RNA in R-loops. In addition, RNaseH2 removes single ribonucleotides mis-incorporated during DNA replication [17]. Interestingly, studies have implicated several other RNA processing factors in R-loop associated genome instability. For instance, the THO/TREX complex functions in transcription elongation and mRNP biogenesis to prevent the accumulation of R-loops [18], whereas SRSF1, a protein involved in RNA metabolism, is recruited to nascent transcripts by RNA polymerase II and prevents the formation of mutagenic R-loops [19]. However, despite several studies linking RNA binding proteins to R-loops and genome instability, the function of many RNA binding proteins in R-loop processing remains uncharacterized.

Here, we describe an important regulatory role of the exonuclease ERI1 in preventing R-loop formation following replication stress. ERI1 was found to be enriched at (stalled) replication forks and protects cells against replication stress. Our results suggest that ERI1 may resolve R-loop induced TRCs that lead to replication fork stalling and/or collapse.

RESULTS

ERI1 was among the candidate genes found whose loss causes sensitivity to Hydroxyurea (HU) treatment, which induced RS by dNTP depletion (Fig. 1A and [20]. To validate this finding, we performed clonogenic cell survival assays in the presence and absence of HU. We found that ERI1 loss rendered cells sensitive to HU, an effect that was similar to that observed after depletion of RPA-interacting protein ETAA1 (Fig. 1B and Fig. S1A [5]. To better understand how the absence of ERI1 sensitizes cells to replication stress (RS), we examined whether ERI1 is involved in ATR signaling following HU-induced helicase-polymerase uncoupling. ATR is the major kinase, which preserves the nuclear RPA pool by restraining the firing of new origins in order to prevent replication for instability. Inhibition of ATR in HU-treated cells leads to excessive double-strand break (DSB) formation and RPA hyperphosphorylation [3]. We found that ERI1-depleted cells showed increased levels of RPA2 phosphorylation at Serine 4 and 8 (p-RPA (S4/8) and Fig. 1C-D), as well as a mild increase in γH2AX levels (Fig. 1E). Interestingly, we also observed an increase in p-RPA (S4/8) levels upon camptothecin (CPT) treatment, which induces replication-dependent DSBs (Fig. S1B). Importantly, RPA levels remained unchanged in HU-treated ERI1-depleted cells (Fig. S1C-D). Finally, we employed the Flp-In/T-Rex system to establish U2OS cells stably expressing inducible and siRNA-resistant GFP-tagged ERI1 (Fig. 1F). Confirming our previous data (Fig. 1D), we found that ERI1 knockdown resulted in an increase of p-RPA (S4/8). Importantly, expression of siRNA-resistant GFP-ERI1 almost fully rescued the increase in p-RPA (S4/S8) observed after ERI1 knockdown (Fig. 1G-H), ruling out off-target effects of the siRNA against ERI1. Taken together, these results suggested that ERI1 protects cells against HU-induced RS likely by promoting replication fork stability.

5
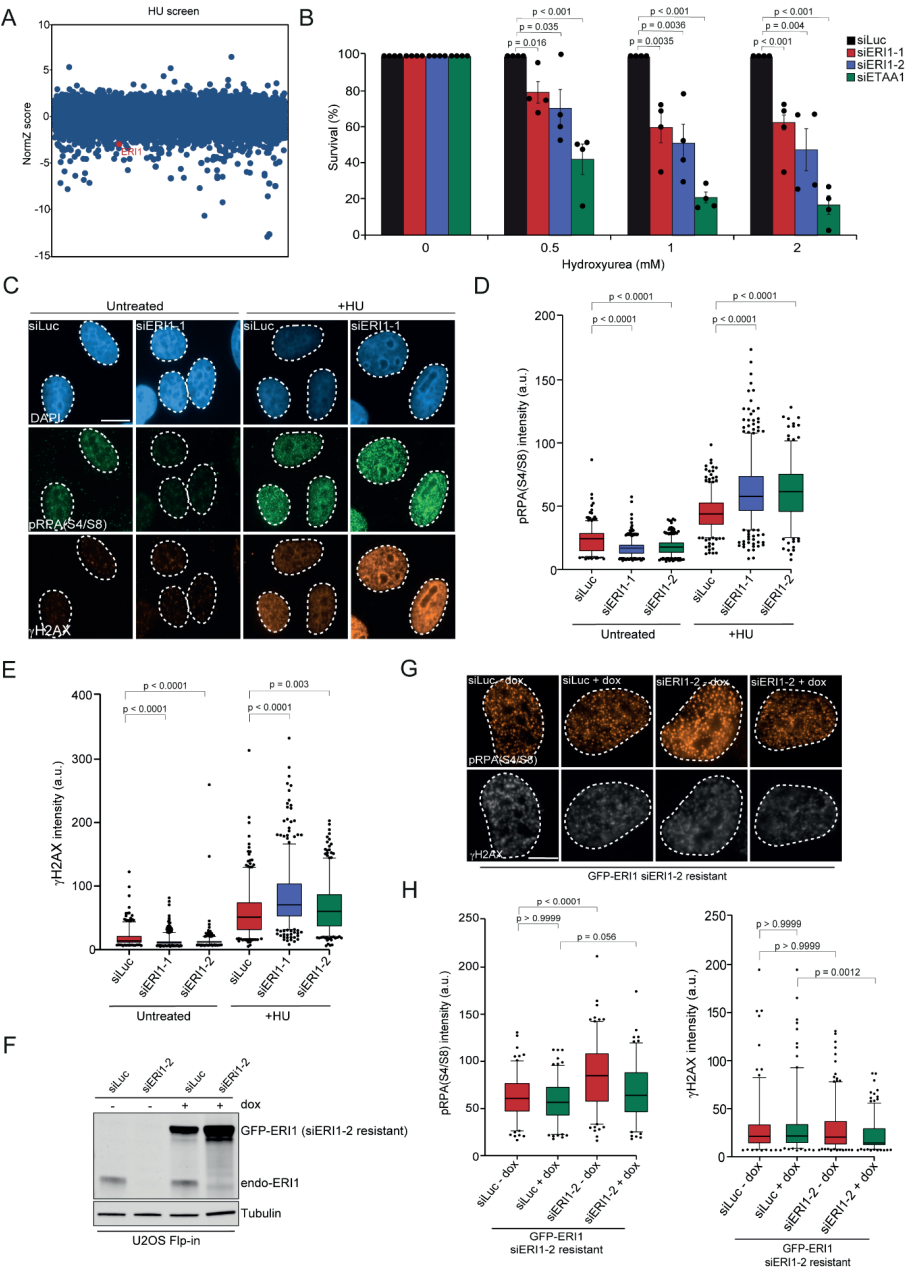
▶ **Figure 1. ERI1 promotes cell survival and ATR signaling upon replication stress.** (A) CRISPR dropout screen results for RPE cells exposed to 2 mM HU. HU was added for 16 h every 3 days for a period of 12-14 days. (B) Relative survival efficiency in HCT116 cells transfected with the indicated siRNAs and treated with 0.5, 1 and 2 mM HU. The mean ±SEM of 3 independent experiments is shown. Data were normalized to un-irradiated conditions and set to 100%. Statistical significance was calculated using Student's *t* test. (C) Phosphorylated RPA at Ser4/8 and γH2AX foci formation in U2OS cells transfected with the indicated siRNAs. Cells were exposed to 2 mM of HU and foci intensity was measured after 4 hours. (D) Quantification of RPA (S4/8) foci intensity in cells from C. Mean foci intensity was quantified in immuno-stained p-RPA (S4/8) cells. The mean of 3 independent experiments is shown. (E) As in D, except that γH2AX foci intensity was measured. (F) Expression levels of endogenous ERI1 and dox-inducible siRNA-resistant GFP-ERI1 in U2OS Flp-In/TRex cells. Tubulin is a loading control. (G) Effect of inducible expression of siRNA resistant GFP-tagged ERI1 on phosphorylated RPA at Ser4/8 and γH2AX foci formation in U2OS cells transfected with the indicated siRNAs. Cells were exposed to 2 mM of HU and foci intensity was measured after 4 hours. (H) Quantification of RPA (S4/8) and γH2AX foci intensity in cells from G. Mean foci intensity was quantified in immuno-stained p-RPA (S4/8) and γH2AX cells. The mean of 3 independent experiments is shown. *P*-values were derived from Kruskal–Wallis ANOVA Dunn's post test.

ERI1 interacts with PARP1

To gain insight into ERI1's potential role in maintaining fork stability, we aimed to identify possible interaction partners of ERI1. To this end, we generated U2OS Flp-In/T-Rex cells stably expressing inducible GFP-tagged ERI1 or GFP-NLS, and performed GFP-trap-based pulldowns followed by label-free mass spectrometry (MS) (Fig. 2A). Our analysis revealed that ERI1 interacts with 173 proteins that were at least 2-fold enriched in GFP-ERI1 pulldowns when compared to those of GFP-NLS (Fig. 2A and Table S1). Although the majority of the interactors were ribosomal proteins, we also found additional interaction partners including RNA binding proteins (RNPS1 and PNN), RNA helicases (DHX15), chromatin remodelers (SMARCA5 and CHD4), histone proteins (HP1BP3, H1F0 and H2AFY) and RNA splicing factors (THRAP3). Interestingly, we found that ERI1 interacts with PARP1, a protein that is implicated in replication fork reversal [21], and TOP2A, a DNA topoisomerase involved in DNA replication [22]. Since ERI1 maintains fork integrity upon RS, we also performed the interaction analysis of GFP-ERI1 in the presence of HU and compared it to that of untreated cells (Fig. 2C and Table S1). In this analysis we found that ERI1 interacts with DHX9, a protein that recruits BRCA1 to RNA to promote DNA end-resection and homologous recombination (HR) repair [23]. Moreover, GNL3, a protein involved in the RS response, was also found to interact with ERI1 under these conditions [24]. Since PARP1 has several roles in the stabilization of stalled replication forks including fork reversal and the regulation of replication fork speed [25], we sought to validate this interaction. GFP pulldowns coupled to western blot analysis confirmed that GFP-tagged ERI1 interacts with endogenous PARP1 (Fig. 2B), manifesting a possible role of this interaction in RS.

5
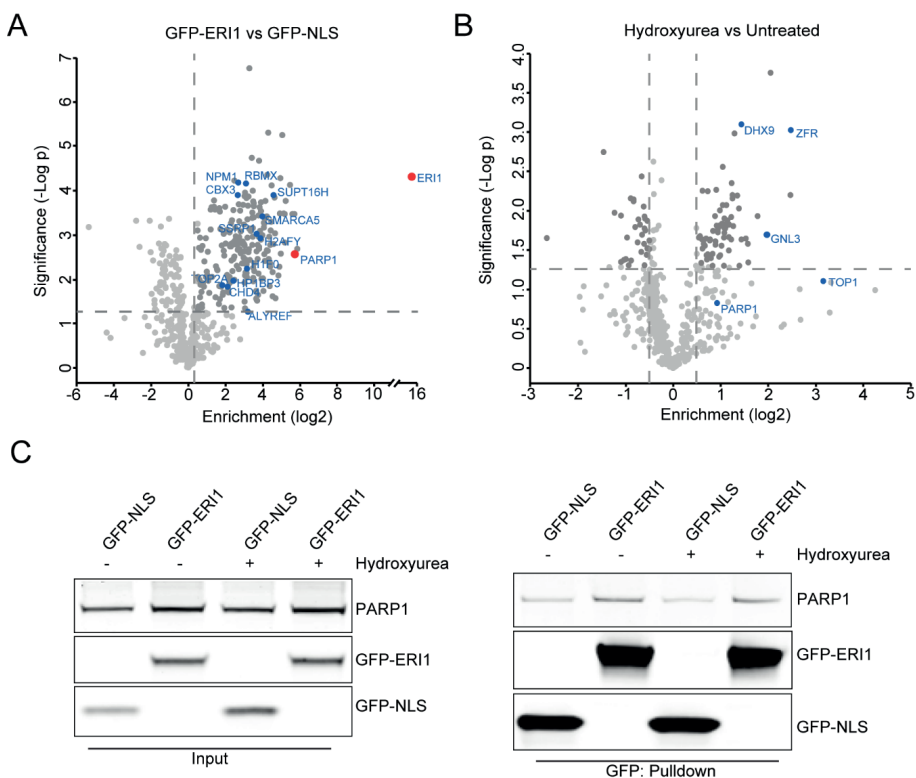
Figure 2. ERI1 interacts with PARP1. (A) Volcano plot depicting the statistical differences of the MS analysis on GFP-KANSL3 versus GFP-NLS pulldowns. The enrichment is plotted on the x-axis and the significance (t-test -log2 p-value) is plotted on the y-axis (see also Supplementary Table 1). (B) As in A, except that cells were treated with 2 mM HU for 4 hours before pulldowns. (C) Pull-down of the indicated GFP fusion proteins in the presence and absence of HU in U2OS Flp-In/T-Rex cells. Blots were probed for GFP and PARP1.

ERI1 is recruited to sites of DNA damage in a manner dependent on PARP activity

Next, we addressed whether ERI1 acts at sites of DNA damage induced by multi-photon laser micro-irradiation (Fig. 3A). Under these conditions, ectopically expressed GFP-tagged ERI1 co-expressed with the DNA damage sensor mCherry-NBS1 in U2OS cells, was rapidly recruited (within 30 s) to DNA damage tracks (Fig. 3B-C). Because of ERI1's interaction with PARP1, we analyzed whether it is recruited to sites of DNA damage via the activity of PARP. To this end, we treated cells with the PARP1 and PARP2 inhibitor (PARPi) Olaparib. This treatment almost completely abrogated the recruitment of GFP-

ERI1 (Fig. 2D-E). Thus, ERI1 is rapidly, but transiently recruited to sites of DNA damage in a manner dependent on the activity of PARP.
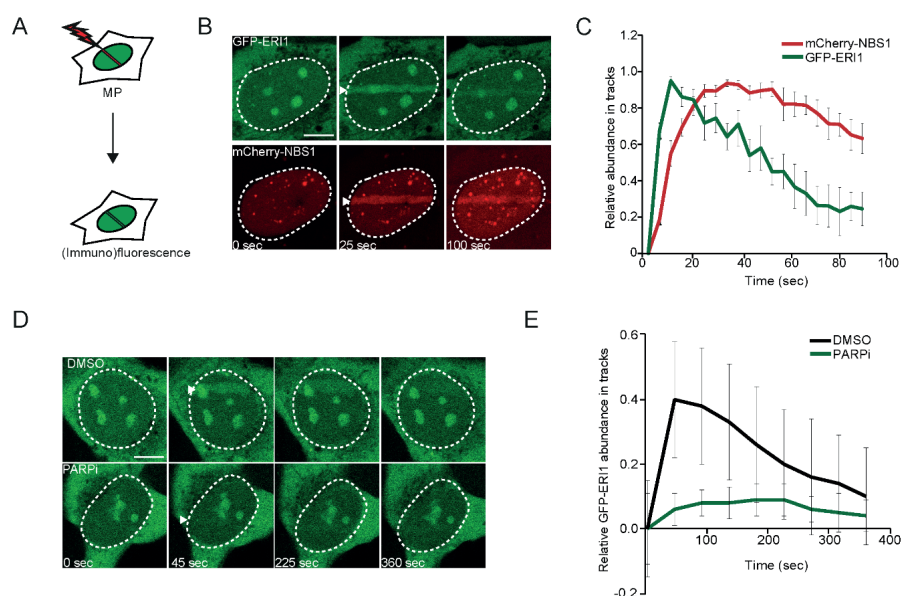
Figure 3. ERI1 is recruited to DNA damage sites via PAR/PARP. (A) Schematic representation of the laser micro-irradiation approach to study protein accumulation at sites of DNA damage. An 800 nm multiphoton (MP) laser was used. (B) Recruitment of GFP-ERI1 to 800 nm multiphoton tracks in U2OS Flp-In/T-Rex cells. mCherry-NBS1 was used as a DNA damage marker. White triangles indicate irradiated regions. (C) Quantification of B. The data are plotted on a timescale as relative abundance in tracks. Peak values were set to 1. The graph represents the mean ± SD of 2 independent experiments. (D) GFP-ERI1 recruitment to 800 nm multiphoton tracks in stable U2OS Flp-In/T-Rex cells treated with PARP inhibitor (PARPi) for 1 h before micro-irradiation. (E) Quantification of D. The data is presented as the mean ± SD.

ERI1 is enriched at (stalled) replication forks

Since multiphoton laser micro-irradiation may induce a variety of lesions other than damaged replication forks, we next investigated the presence of ERI1 on damaged forks [26]. We carried out the isolation of proteins on nascent DNA (iPOND) in unperturbed and replication stress conditions (Fig. 4A). Consistent with previous reports we observed an increase in replication stress proteins, including MRE11 and RPA phosphorylated at Serine 33 (Fig. 4B and [27]. Strikingly, we also found ERI1 enriched at unperturbed forks and increased at HU-induced stalled replication forks (Fig. 4B). This suggests that ERI1 associates with stalled replication forks to maintain their stability.
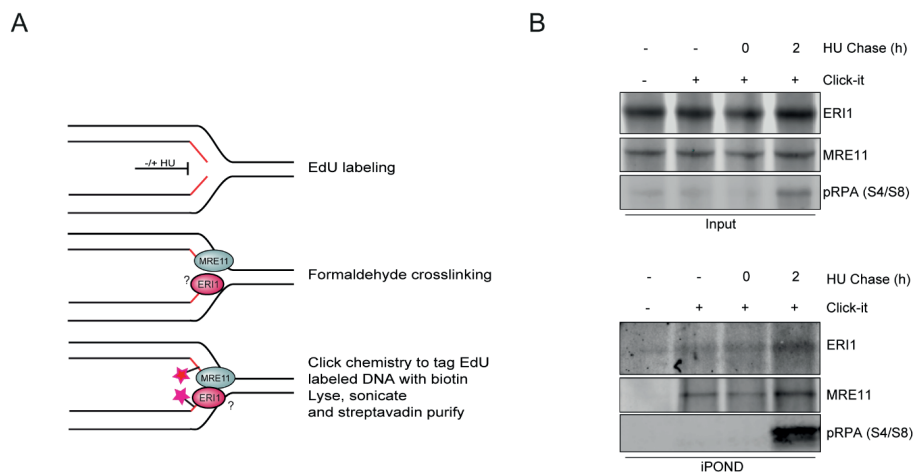
Figure 4. ERI1 is recruited to (stalled) replication forks. (A) Schematic representation of isolation of proteins on nascent DNA (iPOND) procedure. B) iPOND analysis of ERI1 in HEK293T cells after replication stress induced by 2mM HU for the indicated time points. Blots were probed for ERI1, MRE11 and p-RPA (S4/S8).

ERI1-depleted cells suffer from aberrant RNA-DNA hybrid formation

Given the role of ERI1 in several RNA-processing pathways [28], we hypothesized that ERI1 may be involved in the regulation of RNA-DNA hybrid structures (R-loops). R-loops are often associated with transcription-replication conflicts (TRCs), and might be responsible for the observed RS in ERI1-depleted cells. We therefore tested the impact of ERI1 on R-loop formation using the S9.6 antibody for immunofluorescence, which mostly recognizes R-loops, but also has low affinity for dsRNA structures [29]. In order to validate our S9.6 immunofluorescence staining, we depleted a well-known R-loop processing factor AQR and found, as expected, that this led to an increase in R-loop levels (Fig. S2A-B and [30]. In addition, we employed the Flp-In/T-Rex system to establish U2OS cells stably expressing inducible GFP-tagged RNaseH1 (Fig. S2C). Expression of GFP-RNaseH1 almost fully abolished R-loop formation, confirming the specificity of our S9.6 staining (Fig. S2D-E). Next, we measured the accumulation of R-loops in ERI1-depleted cells in both undamaged and CPT-treated cells. We found that loss of ERI1 increased the accumulation of R-loops most clearly following CPT treatment (Fig. 5A-C). Together, these results suggest that ERI1 suppresses R-loop formation, most prominently under RS conditions.
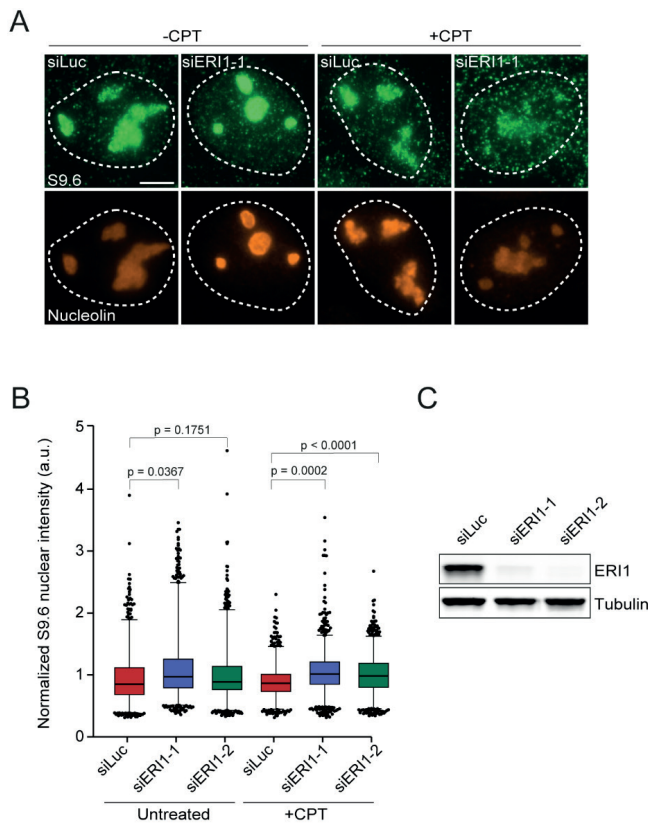
Figure 5. ERI1 suppresses the formation of R-loops. (A) S9.6 nuclear intensity was measured in U2OS cells transfected with the indicated siRNAs. Cells were exposed to 1 μM of CPT and foci intensity was measured after 2 hours. The S9.6 signal intensity per nucleus was determined by subtracting the S9.6 staining with that from nucleoli in each nucleus. (B) Quantification of S9.6 intensity in cells from A. Mean intensity was quantified in immuno-stained S9.6 cells. *P*-values were derived from Kruskal–Wallis ANOVA Dunn's post test. (C) Western blot analysis of ERI1 expression in cells from Figure 5A. Tubulin is a loading control.

The exonuclease domain of ERI1 is critical to suppress replication stress

Since ERI1 is a 3'exonuclease that possesses SAP and exonuclease domains required for binding and processing of RNA, respectively (Fig. 6A and [31, 32]. We next asked whether the decrease in fork stability as measured by phosphorylated RPA at Serine 4/8 relates to ERI1's role in RNA processing (Fig. 1D). Re-expression of siRNA-resistant GFP-tagged wildtype (Fig. 1G), but not exonuclease-dead (D130G and E132G) ERI1 largely suppressed p-RPA (S4/8) intensity in these cells (Fig. 6B-D and [33]. These data

suggest that replication fork stability is strictly dependent on the catalytic activity of ERI1. We speculate that ERI1 prevents the formation of TRCs by the resolution of R-loops and the subsequent stalling/collapse of replication forks (Fig. 6E).
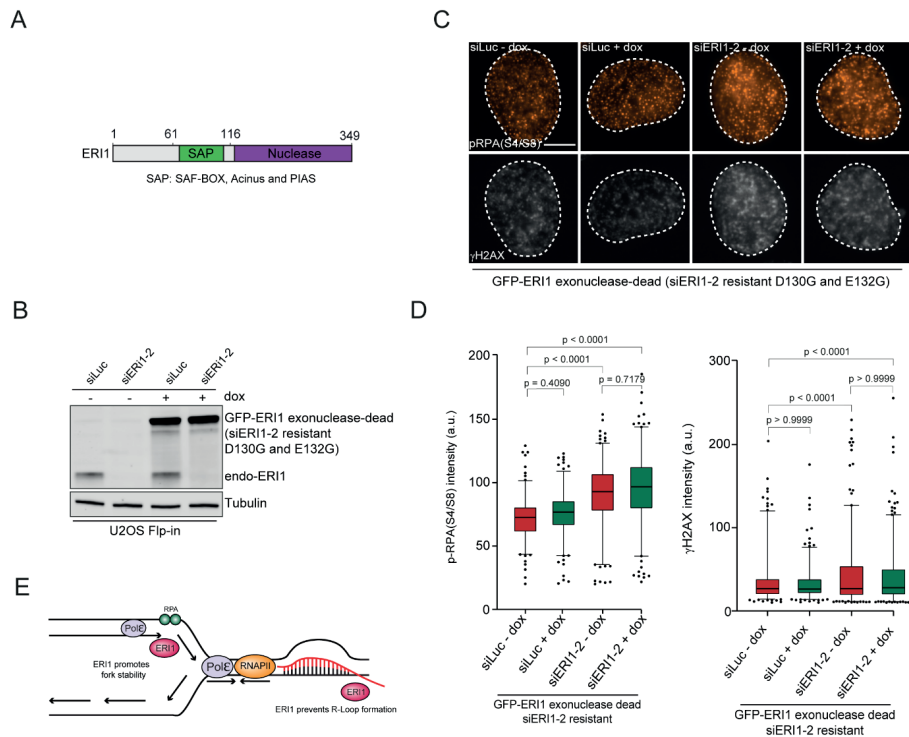
Figure 6. The catalytic activity of ERI1 is critical to suppress replication stress. (A) Schematic representation of full-length ERI1 protein with its indicated domains (SAP and nuclease domain). (B) Expression levels of endogenous ERI1 and dox-inducible siRNA-resistant GFP-ERI1 (exonuclease dead D130G and E132G) in U2OS Flp-In/TRex cells. Tubulin is a loading control. (G) Effect of inducible expression of siRNA resistant GFP-tagged ERI1 exonuclease dead on phosphorylated RPA at Ser4/8 and γH2AX foci formation in U2OS cells transfected with the indicated siRNAs. Cells were exposed to 2 mM of HU and foci intensity was measured after 4 hours. (H) Quantification of RPA (S4/8) and γH2AX foci intensity in cells from G. Mean foci intensity was quantified in immuno-stained p-RPA (S4/8) and γH2AX cells. The mean of 3 independent experiments is shown. *P*-values were derived from Kruskal–Wallis ANOVA Dunn's post test.

DISCUSSION

In this study, we uncover an important role of the exonuclease ERI1 in the RS response. First, we demonstrate that ERI1 protects cells from HU-induced RS and maintains fork stability under these conditions. Second, we show that ERI1 interacts with PARP1 and

is recruited to sites of DNA damage in a manner dependent on PARP activity. Third, iPOND analysis revealed that ERI1 associates with stalled replication forks. Finally, ERI1 suppresses the formation of R-loops and maintains fork stability through its activity in RNA-processing. Thus, ERI1 may ensure fork stability by preventing TRC collisions that are causally linked to R-loop formation and induce replication fork collapse (Fig. 6F).

ERI1 protects cells from HU-induced replication stress and maintains fork stability

In agreement with previous results we confirmed that ERI1-depleted cells are sensitive to HU-induced RS [20]. In response to HU-induced helicase-polymerase uncoupling, ATR kinase is recruited to and activated via its direct interaction with ssDNA at stalled replication forks [34]. This is followed by the ATR-dependent phosphorylation of RPA (at Serine 33), which prevents fork breakage by the suppression of new origin firing, thereby preserving the nuclear pool of RPA [3]. Consequently, inhibition of ATR leads to excessive amounts of ssDNA and limits the available amount RPA molecules by a process referred to as ´RPA exhaustion´. This phenomenon in turn leads to replication fork catastrophies and the formation of DSBs [5, 34]. RPA levels, however, remained unaffected in ERI1-depleted cells, ruling out RPA exhaustion. Instead, we found increased p-RPA (S4/S8) and γH2AX levels upon RS in ERI1-depleted cells. This suggests that ERI1 loss results in the conversion of stalled replication forks into collapsed forks, which may ultimately be converted into DSBs. To investigate this further, DNA fiber analysis may provide insights into whether ERI1 loss affects fork stability/recovery following replication stress.

ERI1 interacts with PARP1 and is enriched at (stalled) replication forks

Our mass spectrometry experiments revealed that the majority of ERI1 interactors are ribosomal proteins. However, we also found several additional interactors such as RNA binding proteins (RNPS1 and PNN), RNA helicases (DHX15), chromatin remodelers (SMARCA5 and CHD4), histone proteins (HP1BP3, H1F0 and H2AFY), RNA splicing factors (THRAP3) and the DNA repair protein PARP1 (Table S1). Moreover, we validated PARP1 as an ERI1 interactor by GFP-pulldowns coupled to western blot analysis, although this needs to be confirmed endogenously and reciprocally. Given that PARP1 is implicated in normal DNA replication and the stabilization of stalled replication forks, it may be possible that ERI1's role in these processes involves its interaction with

PARP1 [35]. Indeed, our findings reveal that ERI1 is recruited to sites of DNA damage via the activity of PARP. However, whether this involves the binding of ERI1 to PAR chains or to PARP1 itself needs to be further investigated [28]. Moreover, our iPOND analysis revealed that ERI1 is enriched at unperturbed forks and increased at HU-induced stalled replication forks. Similarly, PARP1 is activated and binds to HU-induced stalled replication forks [36]. Thus, it seems logical to next assess by iPOND whether ERI1's recruitment to stalled forks is dependent on PAR/PARP1. However, since ERI1 has affinity to bind RNA via its SAP domain, its recruitment to stalled forks may also rely on its interaction with RNA or its exonuclease activity (e.g. to process RNA in R-loops). Further work, including domain mapping analysis, should identify the PARP1 and RNA binding regions in ERI1 to resolve the functional relevance of interplay between ERI1 and PARP1 or RNA in response to replication fork stalling.

ERI1 depletion results in increased R-Loop formation

Our data reveal that ERI1 depletion leads to increased R-loop levels mostly following RS. We used the S9.6 antibody for immunofluorescence (IF) to measure R-loop levels. As expected, we found AQR depletion to increase the formation of R-loops, and overexpression of RnaseH1 to decrease the formation of R-loops, which validates our S9.6 IF [37]. However, because the S9.6 antibody also recognizes double-strand RNA [30, 38], it is of utmost importance to further validate our S9.6 IF experiments using RnaseIII treatment, as well as treatment with recombinant wildtype and catalytic inactive version of human RNaseH1 to further validate the specificity of our S9.6 IF [38]. Since ERI1 is involved in the processing of multiple RNA species [28], and our data reveal that the exonuclease domain of ERI1 is critical to suppress RS as measured by p-RPA (S4/8), it may be interesting to investigate whether the exonuclease domain of ERI1 has affinity to process the RNA in R-loops. Future *in vitro* R-loop formation assays using purified ERI1 wildtype and exonuclease dead will determine whether ERI1 processes R-loops and if this is dependent on its exonuclease domain [39].

The unscheduled formation of R-loops has been linked to TRCs and RS [40]. This raises the question whether the aberrant R-loops observed in ERI1-depleted cells, are a potential mechanism driving RS. Future endeavors including RNaseH1 overexpression rescue experiments in ERI1-depleted cells will give insight into whether the increase in fork stalling measured by p-RPA (S4/8) and DNA breaks measured by γH2AX levels, as well as the HU-sensitivity observed in ERI1-depleted cells is caused by R-loops.

Furthermore, R-loops may pose an obstacle for replication fork progression and thereby cause defects in replication fork instability upon HU-induced RS in ERI1-depleted cells [41]. To investigate this possibility, proximity ligation assay (PLA) between PCNA, as a marker for replication, and RNAPII may evaluate the occurrence of TRCs in absence of ERI1 [42].

A proper balance of histones and DNA during DNA replication is crucial, as dysregulation of the amount of histone levels is associated with chromosomal loss and genome instability [43]. ERI1 is required for the degradation of replication-dependent histone mRNAs [44, 45]. This raises the question whether this function of ERI1 is linked to R-loops and is associated replication stress. It would be important to decipher the genomic location in which R-loops are enriched in ERI1 depleted cells, and whether this correlates with changes in histone incorporation levels at these loci during replication. In depth analysis of R-loop distribution by DRIP-sequencing in both control and ERI1-depleted cells, as well as ChIP-sequencing, including e.g. chromatin occupancy after replication-seq (ChOR-seq) [46], to determine changes in histone levels at these sites will be required in the future to address these questions.

MATERIAL AND METHODS

Cell lines

U2OS and HCT116 cells were cultured in 5% CO_2 at 37°C in DMEM (Dulbecco's modified Eagle's medium) supplemented with 10% fetal calf serum and antibiotics. U2OS Flp-In/T-Rex cells, which were generated using the Flp-In/T-REx system (Thermo Fisher Scientific), were a gift of Geert Kops (University Medical Center Utrecht, the Netherlands) and Stephen Taylor (University of Manchester, UK). These cells were used to stably express inducible versions of GFP-NLS, as well as GFP-ERI1 by co-transfection of pCDNA5/FRT/TO-Puro plasmid encoding GFP or GFP-tagged ERI1 (5 µg), together with pOG44 plasmid encoding the Flp recombinase (1 µg). After selection on 1 µg/mL puromycin, single clones were isolated and expanded. U2OS Flp-In/T-Rex were incubated with 2 µg/mL doxycycline for 24h to induce expression of cDNAs. All cells were authenticated by STR profiling and tested negative in routinely performed mycoplasma tests.

5

Chemicals

Cells were treated with 2mM Hydroxyurea (Sigma) at the indicated timepoints and collected for further analysis. The PARP inhibitor olaparib (Selleck Chemicals) was used at a final concentration of 10 μM. Camptothecin (Sigma-Aldrich) was used at a final concentration of 1 μM.

Transfections, siRNAs and plasmids

Cells were transfected with siRNAs using RNAiMAX (Invitrogen) according to the manufacturer's instructions. Cells were transfected twice with siRNAs at 0 and 24 h at a concentration of 40 nM and were analyzed 48 h after the second transfection unless otherwise indicated. siRNA sequences are listed in Table 1. Cells were transfected with plasmid DNA using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions and analyzed 24-48 hours after transfection. The expression vector for full length human ERI1 (ORFeome collection), was amplified and cloned into pCDNA5/FRT/TO-Puro as a *BspEI/Acc65I* fragment (Table 2). Mutant constructs were generated by site-directed mutagenesis PCR (Table 2). siERI1-resistant cDNA was generated by introducing the underlined mutations GT<u>T</u>AT<u>C</u>GA<u>T</u>TGGATGAA<u>G</u>TT<u>A</u> by overlap PCR and cloned as *BspEI/Acc65I* fragment into pCDNA5/FRT/TO-Puro-ERI1-WT (Table 2). All ERI1 expression constructs were verified using Sanger sequencing.

Multiphoton laser micro-irradiation

Cells grown were grown on 18 mm coverslips. For micro-irradiation, cells were placed in a Chamlide CMB magnetic chamber and the growth medium was replaced by CO_2-independent Leibovitz's L15 medium supplemented with 10% FCS and antibiotics. Laser micro-irradiation was performed on a Leica SP5 confocal microscope equipped with an environmental chamber set to 37°C. DNA damage-containing tracks (1.5 μm width) were generated with a Mira mode locked titanium-sapphire (Ti:Sapphire) laser (l = 800 nm, pulse length = 200 fs, repetition rate = 76 MHz, output power = 80 mW) using a UV-transmitting 63x 1.4 NA oil immersion objective (HCX PL APO; Leica). Confocal images were recorded before and after laser irradiation at 5 seconds time interval over a period of 3-5 minutes. Images after multi-photon micro-irradiation of living cells were recorded using LAS-AF software (Leica) and analyzed with Image J as described previously [47]. The average pixel intensity of laser tracks was measured within the locally irradiated area (Idamage), in the nucleoplasm outside the locally irradiated area (Inucleoplasm)

and in a region not containing cells in the same field of view (Ibackground). The level of protein accumulation relative to the protein level in the nucleoplasm was calculated as follows: ((Idamage – Ibackground)/(Inucleoplasm – Ibackground) – 1).

Immunofluorescence analysis

Cells were either directly fixed with 2% formaldehyde in PBS for 20 minutes at room temperature (RT), or pre-extracted with 0.5% Triton-X100 (Serva) in PBS on ice for 2-5 minutes prior to fixation. Alternatively, cells were fixed, post-extracted with 0.25% Triton-X100 (Serva) in PBS and treated with 100 mM glycine in PBS for 20 minutes to block unreacted aldehyde groups. Cells were then rinsed with PBS and equilibrated in wash buffer (PBS containing 0.5% BSA). Antibody incubation steps and washes were in wash buffer. Primary antibodies were incubated for 1-2 hours at room temperature. Detection was done using goat anti-mouse or goat anti-rabbit Ig coupled to Alexa 488, 555 or 647 (1:1500; Invitrogen Molecular probes). All antibodies are listed in Table 3. Samples were incubated with 0.1 µg/mL 4',6-Diamidino-2-Phenylindole Dihydrochloride (DAPI) and mounted in Polymount.

Sample preparation and mass spectrometry

For mass spectrometry, U2OS Flp-In/T-Rex cells expressing GFP-NLS and GFP-ERI1^{WT} were treated with 2 mM Hydroxyurea for 4 hours or left untreated. Cell pellets were lysed in EBC-1 buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.5% NP-40, 2 mM MgCl$_2$, protease inhibitor cocktail tablets) with 500 units benzonase. Samples were incubated for 1 hour at 4ºC under constant mixing followed by high speed centrifugation for 10 minutes at 4ºC. Protein concentration was measured by Qubit in the cleared lysates, equalized and transferred to tubes containing GFP-Trap beads (Chromotek). After 90 minutes of incubation at 4ºC under rotating condition, the beads were washed 4 times with EBC-2 buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, and protease inhibitor cocktail tablets) and 3 times with 50 mM ammonium bicarbonate followed by overnight digestion using 2.5 µg trypsin at 37°C under constant shaking. Digestion was terminated with 1% trifluoroacetic acid and centrifuged for 5 minutes at high speed to precipitate insoluble fractions. Consequently, C18 cartridges were prepared by washing 2 times with acetonitrile followed by 2 times with 0,1% Acetic acid. Peptides were loaded on the cartridge, while bound peptides were washed 2 times with 0.1% acetic acid and eluted with 1 mL 80% acetonitrile / 0.1% acetic acid and lyophilized.

Mass spectrometry was performed essentially as previously described [48]. Samples were nalysed on a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific, Germany) coupled to an EASY-nanoLC 1000 system (Proxeon, Odense, Denmark). Digested peptides were separated using a 15 cm fused silica capillary (ID: 75 µm, OD: 375 µm, Polymicro Technologies,California, US) in-house packed with 1.9 µm C18-AQ beads (Reprospher-DE, Pur, Dr. Maisch, Ammerburch-Entringen, Germany). Peptides were separated by liquid chromatography using a gradient from 2% to 95% acetonitrile with 0.1% formic acid at a flow rate of 200 nl/minute for 65 minutes. The mass spectrometer was operated in positive-ion mode at 2.8 kV with the capillary heated to 250°C, and in a Data-Dependent Acquisition (DDA) mode with a top 7 method. Full scan MS spectra were obtained with a resolution of 70,000, a target value of $3x10^6$ and a scan range from 400 to 2,000 m/z. Maximum Injection Time (IT) was set to 50 ms. Higher-Collisional Dissociation (HCD) tandem mass spectra (MS/MS) were recorded with a resolution of 35,000, a maximum IT of 20 ms, a target value of $1x10^5$ and a normalized collision energy of 25%. The precursor ion masses selected for MS/MS analysis were subsequently dynamically excluded from MS/MS analysis for 60 seconds. Precursor ions with a charge state of 1 and greater than 6 were excluded from triggering MS/MS events. Three replicates were included per condition with two technical repeats each.

Mass Spectrometry data analysis

Raw mass spectrometry files were analysed with MaxQuant software (v1.5.3.30) as described [49] with the following modifications from default settings: the maximum number of mis-cleavages by trypsin/p was set to 3, Label Free Quantification (LFQ) was enabled disabling the Fast LFQ feature. Match-between-runs feature was enabled with a match time window of 0.7 minutes and an alignment time window of 20 minutes. We performed the search against an in silico digested UniProt reference proteome for Homo sapiens (8th June 2020). Analysis output from MaxQuant was further processed in the Perseus (v 1.5.5.3) computational platform [49]. Proteins identified as common contaminants, only identified by site and reverse peptide, were filtered out, and then all the LFQ intensities were log2 transformed. Different biological repeats of each condition were grouped and only protein groups identified in all three biological replicates in at least one condition were included for further analysis. Missing values were imputed using Perseus software by normally distributed values with a 1.8 downshift (log2) and a randomized 0.3 width (log2) considering total matrix values. Volcano plots were

generated, and Student's T-tests were performed to compare the different conditions. Spreadsheets from the statistical analysis output from Perseus were further processed in Microsoft Excel for comprehensive visualization and analysis of the data (Table S1).

Western blot analysis

Cells were lysed in 2x Laemmli buffer and proteins were separated by Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis (SDS-PAGE) using 4-12% pre-cast polyacrylamide gels (BioRad or Invitrogen) and MOPS running buffer (Invitrogen). Next, proteins were transferred onto nitrocellulose membranes (Millipore). Protein expression was analyzed by immunoblotting with the indicated primary antibodies (Table 3) and secondary CF680 goat anti-rabbit or CF770 goat anti-mouse Ig antibodies (1:5000, Biotium). Membranes were scanned and analyzed using a Licor Odyssey scanner (LI-COR Biosciences).

Isolation of proteins on nascent DNA (iPOND)

iPOND was performed as described by [27] with minor modifications. Briefly, 80-100*10^6 HEK293T cells were pulse labelled with 10 μM 5-ethynyl-2'-deoxyuridine (EdU; Invitrogen) for 20 min. Next, cells were treated with 2 mM HU for 5 hr to induce stalled forks. After treatment, cells were cross-linked with 1% formaldehyde without methanol in PBS for 15 min at room temperature, quenched with 0.125 M glycine and collected by scraping. Fixed cells were spun down for 5 min at 4 °C and 900 g, washed two times with PBS. Next, cells were permeabilized with 0.25% Triton X-100 in PBS for 30 min at room temperature, followed by incubation with click-iT reaction buffer (10 μM biotin azide (life technologies), 10 mM sodium-L-ascorbate, and 2 mM CuSO$_4$) or control buffer for 3 hr at room temperature. The cells were then washed with 0.5% BSA/PBS, and resuspended in lysis buffer (0.5% SDS, 50 mM Tris-HCl pH 8.0, protease inhibitors (Roche)). Samples were then sonicated using a Branson 250 SONIFIER with microtip for 1 s pulse for 10 cycles, repeated 3 times, with 90 s pause between repetition. For input samples, an aliquot of the lysate was taken. The rest of the samples were incubated with pre-washed streptavidin coated beads (Life Technologies) for 1 hr at room temperature. Beads were washed twice with high salt buffer (1% Triton X-100, 20 mM Tris [pH 8.0], 2 mM EDTA, 500 mM NaCl) and twice with lithium chloride salt wash buffer (100 mM Tris [pH 8.0], 500 mM LiCl and 1% Igepal [NP-40]), followed by elution of the captured proteins by boiling the beads in 2X Laemmli for 25 min at 95°C. Proteins were separated on a 4-12% Criterion™ XT Bis-Tris gel (Bio-Rad) and detected by western blot.

Cell survival assays

HCT116 cells were transfected with siRNAs, trypsinized, seeded at low density and exposed to HU for 24 hours. For HCT116 Flp-In/T-Rex, cDNAs were expressed by adding Dox for 24 hours after siRNA transfection. After 12 days, the cells were washed with 0.9% NaCl and stained with methylene blue (2.5 g/L in 5% ethanol, Sigma-Aldrich). Colonies of more than 20 cells were scored.

AUTHOR CONTRIBUTION

ACKNOWLEDGEMENTS

REFERENCES

1. Gaillard, H., T. García-Muse, and A. Aguilera, *Replication stress and cancer.* Nature Reviews Cancer, 2015. **15**(5): p. 276-289.

2. Zeman, M.K. and K.A. Cimprich, *Causes and consequences of replication stress.* Nat Cell Biol, 2014. **16**(1): p. 2-9.

3. Toledo, L.I., et al., *ATR prohibits replication catastrophe by preventing global exhaustion of RPA.* Cell, 2013. **155**(5): p. 1088-103.

4. O'Connell, B.C., et al., *A genome-wide camptothecin sensitivity screen identifies a mammalian MMS22L-NFKBIL2 complex required for genomic stability.* Mol Cell, 2010. **40**(4): p. 645-57.

5. Haahr, P., et al., *Activation of the ATR kinase by the RPA-binding protein ETAA1.* Nature Cell Biology, 2016. **18**(11): p. 1196-1207.

6. Bass, T.E., et al., *ETAA1 acts at stalled replication forks to maintain genome integrity.* Nat Cell Biol, 2016. **18**(11): p. 1185-1195.

7. Zhao, H. and H. Piwnica-Worms, *ATR-mediated checkpoint pathways regulate phosphorylation and activation of human Chk1.* Molecular and cellular biology, 2001. **21**(13): p. 4129-4139.

8. Neelsen, K.J. and M. Lopes, *Replication fork reversal in eukaryotes: from dead end to dynamic response.* Nature Reviews Molecular Cell Biology, 2015. **16**(4): p. 207-220.

9. Berti, M., et al., *Human RECQ1 promotes restart of replication forks reversed by DNA topoisomerase I inhibition.* Nat Struct Mol Biol, 2013. **20**(3): p. 347-54.

10. Schlacher, K., H. Wu, and M. Jasin, *A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2.* Cancer Cell, 2012. **22**(1): p. 106-16.

11. Quinet, A., et al., *To skip or not to skip: choosing repriming to tolerate DNA damage.* Molecular Cell, 2021. **81**(4): p. 649-658.

12. Lin, Y.-L. and P. Pasero, *Transcription-Replication Conflicts: Orientation Matters.* Cell, 2017. **170**(4): p. 603-604.

13. Nguyen, H.D., et al., *Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1.* Mol Cell, 2017. **65**(5): p. 832-847.e4.

14. Gaillard, H. and A. Aguilera, *Transcription as a Threat to Genome Integrity.* Annu Rev Biochem, 2016. **85**: p. 291-317.

15. Crossley, M.P., M. Bocek, and K.A. Cimprich, *R-Loops as Cellular Regulators and Genomic Threats.* Mol Cell, 2019. **73**(3): p. 398-411.

16. Costantino, L. and D. Koshland, *The Yin and Yang of R-loop biology.* Current Opinion in Cell Biology, 2015. **34**: p. 39-45.

17. Hiller, B., et al., *Mammalian RNase H2 removes ribonucleotides from DNA to maintain genome integrity.* J Exp Med, 2012. **209**(8): p. 1419-26.

18. Gómez-González, B., et al., *Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles.* The EMBO journal, 2011. **30**(15): p. 3106-3119.

19. Li, X. and J.L. Manley, *Inactivation of the SR Protein Splicing Factor ASF/SF2 Results in Genomic Instability.* Cell, 2005. **122**(3): p. 365-378.

20. Olivieri, M., et al., *A genetic map of the response to DNA damage in human cells.* 2019: p. 845446.

21. Genois, M.-M., et al., *CARM1 regulates replication fork speed and stress response by stimulating PARP1.* Molecular Cell, 2021. **81**(4): p. 784-800.e8.

5

22. Nitiss, J.L., *DNA topoisomerase II and its growing repertoire of biological functions.* Nature Reviews Cancer, 2009. **9**(5): p. 327-337.

23. Chakraborty, P. and K. Hiom, *DHX9-dependent recruitment of BRCA1 to RNA promotes DNA end resection in homologous recombination.* Nature Communications, 2021. **12**(1): p. 4126.

24. Lin, T., et al., *Nucleostemin and GNL3L exercise distinct functions in genome protection and ribosome synthesis, respectively.* J Cell Sci, 2014. **127**(Pt 10): p. 2302-12.

25. Ray Chaudhuri, A. and A. Nussenzweig, *The multifaceted roles of PARP1 in DNA repair and chromatin remodelling.* Nature Reviews Molecular Cell Biology, 2017. **18**(10): p. 610-621.

26. Reynolds, P., et al., *Spatiotemporal dynamics of DNA repair proteins following laser microbeam induced DNA damage – When is a DSB not a DSB?* Mutation Research/Genetic Toxicology and Environmental Mutagenesis, 2013. **756**(1): p. 14-20.

27. Sirbu, B.M., et al., *Analysis of protein dynamics at active, stalled, and collapsed replication forks.* Genes Dev, 2011. **25**(12): p. 1320-7.

28. Thomas, M.F., N.D. L'Etoile, and K.M. Ansel, *Eri1: a conserved enzyme at the crossroads of multiple RNA-processing pathways.* Trends in genetics : TIG, 2014. **30**(7): p. 298-307.

29. Crossley, M.P., et al., *Catalytically inactive, purified RNase H1: A specific and sensitive probe for RNA-DNA hybrid imaging.* J Cell Biol, 2021. **220**(9).

30. Sollier, J., et al., *Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability.* Molecular Cell, 2014. **56**(6): p. 777-785.

31. Yang, X.C., et al., *Characterization of 3'hExo, a 3' exonuclease specifically interacting with the 3' end of histone mRNA.* J Biol Chem, 2006. **281**(41): p. 30447-54.

32. Ansel, K.M., et al., *Mouse Eri1 interacts with the ribosome and catalyzes 5.8S rRNA processing.* Nat Struct Mol Biol, 2008. **15**(5): p. 523-30.

33. Cheng, Y. and D.J. Patel, *Crystallographic structure of the nuclease domain of 3'hExo, a DEDDh family member, bound to rAMP.* J Mol Biol, 2004. **343**(2): p. 305-12.

34. Liao, H., et al., *Mechanisms for stalled replication fork stabilization: new targets for synthetic lethality strategies in cancer treatments.* EMBO reports, 2018. **19**(9): p. e46263.

35. Ray Chaudhuri, A. and A. Nussenzweig, *The multifaceted roles of PARP1 in DNA repair and chromatin remodelling.* Nature reviews. Molecular cell biology, 2017. **18**(10): p. 610-621.

36. Bryant, H.E., et al., *PARP is activated at stalled forks to mediate Mre11-dependent replication restart and recombination.* The EMBO Journal, 2009. **28**(17): p. 2601-2615.

37. Sakasai, R., et al., *Aquarius is required for proper CtIP expression and homologous recombination repair.* Scientific Reports, 2017. **7**(1): p. 13808.

38. Crossley, M.P., et al., *Catalytically inactive, purified RNase H1: A specific and sensitive probe for RNA–DNA hybrid imaging.* Journal of Cell Biology, 2021. **220**(9).

39. Kasahara, M., et al., *RecA protein-dependent R-loop formation in vitro.* Genes & development, 2000. **14**(3): p. 360-365.

40. García-Muse, T. and A. Aguilera, *Transcription–replication conflicts: how they occur and how they are resolved.* Nature Reviews Molecular Cell Biology, 2016. **17**(9): p. 553-563.

41. Gómez-González, B. and A. Aguilera, *Transcription-mediated replication hindrance: a major driver of genome instability.* Genes Dev, 2019. **33**(15-16): p. 1008-1026.

42. Bayona-Feliu, A., et al., *The SWI/SNF chromatin remodeling complex helps resolve R-loop-mediated transcription–replication conflicts.* Nature Genetics, 2021.

43. Kaygun, H. and W.F. Marzluff, *Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1.* Nat Struct Mol Biol, 2005. **12**(9): p. 794-800.

44. Hoefig, K.P., et al., *Eri1 degrades the stem-loop of oligouridylated histone mRNAs to induce replication-dependent decay.* Nat Struct Mol Biol, 2013. **20**(1): p. 73-81.

45. Hoefig, K.P. and V. Heissmeyer, *Degradation of oligouridylated histone mRNAs: see UUUUU and goodbye.* WIREs RNA, 2014. **5**(4): p. 577-589.

46. Reverón-Gómez, N., et al., *Accurate Recycling of Parental Histones Reproduces the Histone Modification Landscape during DNA Replication.* Mol Cell, 2018. **72**(2): p. 239-249.e5.

47. Luijsterburg, M.S., et al., *PARP1 Links CHD2-Mediated Chromatin Expansion and H3.3 Deposition to DNA Repair by Non-homologous End-Joining.* Mol Cell, 2016. **61**(4): p. 547-62.

48. Kumar, R., et al., *The STUbL RNF4 regulates protein group SUMOylation by targeting the SUMO conjugation machinery.* Nat Commun, 2017. **8**(1): p. 1809.

49. Tyanova, S., T. Temu, and J. Cox, *The MaxQuant computational platform for mass spectrometry-based shotgun proteomics.* Nat Protoc, 2016. **11**(12): p. 2301-2319.

5

Table 1. List of siRNAs

Target	Sequence (5'-3')
ERI1-1	ACGAAUGGCUGUAUUAAUA
ERI1-2	GUAAUUGACUGGAUGAAAU
XRCC4	AUAUGUUGGUGAACUGAGA
ETAA1	GAGCAAAACAAGAGGAAUUUU
Luciferase	CGUACGCGGAAUACUUCGA
AQR	SMARTpool (siGENOME)

Table 2. List of primers

Name	Sequence (5'-3')
HindIII FW	TAAAAAGCTTATATGGAAGAATCTCACTTCAATTCTAAC
KpnI RV	TAATAATGGTACCCTAAGAGCTGGCCAGGTGC
ERI1_siRNAres_HindII Rv	ACTCTCGGCATGGACGAGCTG
ERI1_siRNAres_KpnI FW	GTTAAACCCGCTGATCAGCCTCGA
M13 FW	GTAAAACGACGGCCAGT
M13 RV	CAGGAAACAGCTATGAC
ER1_Catalytic_dead_Fw	GTTATTATGACTACATTTGTATTATTGGCTTTGGAGCCACTTGTGAAGAAGG
ERI1_Catalytic_dead_Rv	CCTTCTTCACAAGTGGCTCCAAAGCCAATAATACAAATGTAGTCATAATAAC

Table 3. List of primary antibodies

Protein	Host	Company	IF	WB
GFP	Mouse	Roche (11814460001)		1:1000
α-Tubulin	Mouse	Sigma (cloneDM1A, T6199)		1:5000
MRE11	Rabbit	Kind gift of Roland Kanaar		1:3000
ERI1	Rabbit	Cell signaling (405SS)		1:2000
p-RPA (S4/S8)	Rabbit	Bethyl Laboratories (A300-245A)	1:1000	1:1000
ETAA1	Rabbit	Kind gift of Niels Mailand		1:400
RPA	Mouse	Abcam (ab2175)	1:1000	1:1000

Supplementary Figures and Tables - Heading

A

B

C

D

5
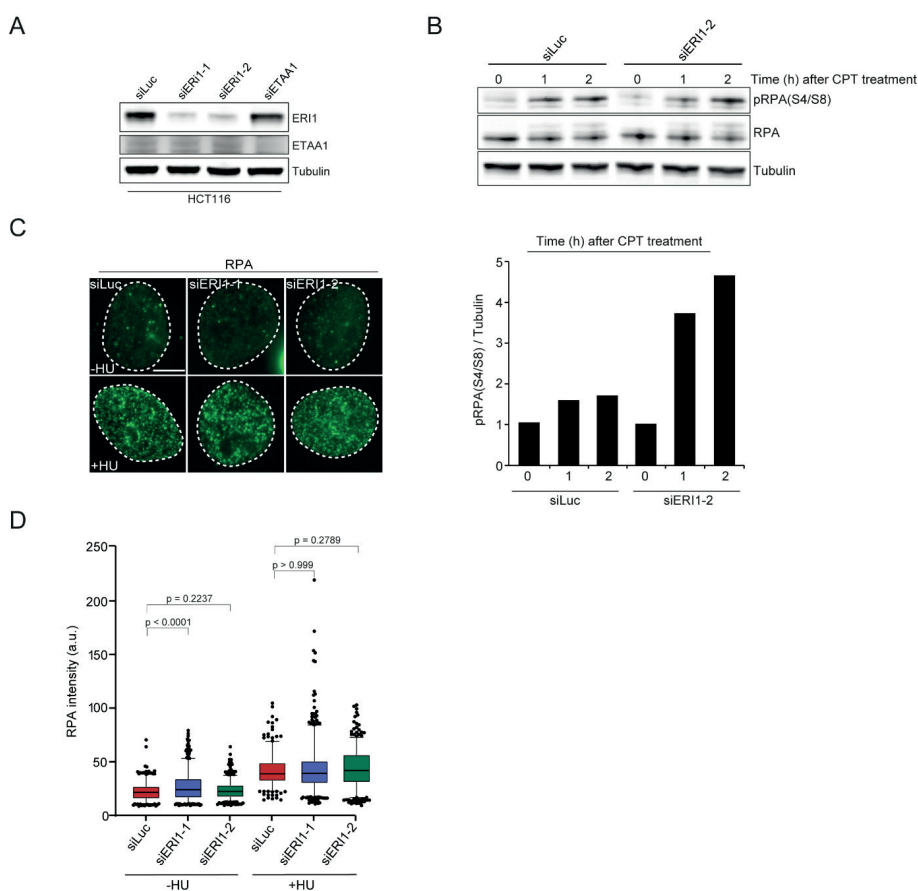
Supplementary Figure. 1 – related to Figure. 1. ERI1 protects cells against HU-induced RS.
(A) Western blot analysis of ERI1 expression in cells from Figure 1B. Blots are probed for ERI1, ETAA1 and Tubulin. Tubulin is a loading control. (B) Western blot analysis of phosphorylated RPA32 (S4/S8) and RPA32 expression in cells transfected with the indicated siRNAs. Cells were examined after the indicated timepoint upon 1µM CPT treatment. Tubulin is a loading control (upper panel). Quantification of phosphorylated RPA32 (S4/S8) expression Tubulin is a loading control, which was used for normalization of phosphorylated RPA32 (S4/S8) expression. Data were normalized to siLuc UT, which was set to 1 (lower panel). (C) RPA foci formation in U2OS cells transfected with the indicated siRNAs. Cells were exposed to 2 mM of HU and foci intensity was measured after 4 hours. (D) Quantification of RPA intensity in cells from B. Mean foci intensity was quantified in immuno-stained RPA cells. The mean of 3 independent experiments is shown.
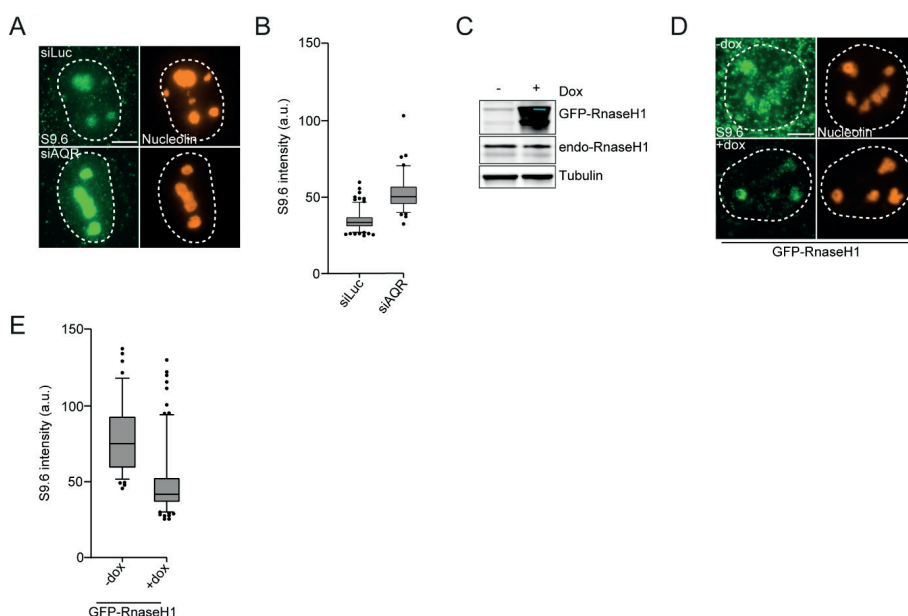
Supplementary Figure. 2 – related to Figure. 5. Validation of S9.6 staining. (A) S9.6 nuclear intensity was measured in U2OS cells transfected with the indicated siRNAs. The S9.6 signal intensity per nucleus was determined by subtracting the S9.6 staining with that from nucleoli in each nucleus. (B) Quantification of S9.6 intensity in cells from A. Mean intensity was quantified in immuno-stained S9.6 cells. *P*-values were derived from Kruskal–Wallis ANOVA Dunn's post test. (C) Expression levels of endogenous RnaseH1 and dox-inducible GFP-RnaseH1 in U2OS Flp-In/TRex cells. Tubulin is a loading control. (D) Effect of inducible expression of GFP-tagged RnaseH1 on S9.6 nuclear intensity U2OS Flp-In/TRex cells. The S9.6 signal intensity per nucleus was determined by subtracting the S9.6 staining with that from nucleoli in each nucleus. (E) Quantification of S9.6 intensity in cells from D. Mean intensity was quantified in immuno-stained S9.6 cells. *P*-values were derived from Kruskal–Wallis ANOVA Dunn's post test.

Supplementary Table 1

Supplementary Table 1. Statistical analysis comparing LFQ intensities of proteins identified by LC-MS/MS in GFP-pulldowns from U2OS Flp-In/T-Rex cells expressing either GFP-ERI1 or GFP-NLS in the presence and absence of HU. -Log p-value and average fold-enrichment (log2) are indicated. Additional information from the analysis of identified proteins is also included.

GENE			Eri1 Vs GFP			Eri1-HU Vs Eri1			Miscellaneus												Values								
Fasta headers	Protein names	Gene names	Significant (p<0.05)	-Log p-value	Difference (log2)	Significant (p<0.05)	-Log p-value	Difference (log2)	Peptides	Razor + unique peptides	Unique peptides	Sequence coverage [%]	Unique + razor sequence coverage [%]	Unique sequence coverage [%]	Mol. weight [kDa]	Q-value	Score	Intensity	MS/MS Count	Protein IDs	Majority protein IDs	ERI1	ERI1-2	ERI1-3	ERI1-HU-1	ERI1-HU-2	ERI1-HU-3	GFP-1	GFP-2	GFP-3

(Remaining data rows of the supplementary table are not legibly reproducible at this resolution.)

5

5

CHAPTER 6

Perspectives

PERSPECTIVES

All cellular organisms contain genomic DNA which provides the instructions for their correct development and functioning. Damage to this DNA may interfere with critical cellular processes such as transcription and replication and has the potential to drive mutagenesis. In turn, this may underlie inherited disorders and accelerate progression of diseases such as cancer and neurodegenerative disorders [1]. The protection of cells and organisms against these devastating effects of DNA damage relies on the DNA damage response (DDR), which comprises a complex network of signaling and repair pathways that coordinate the sensing, signaling and repair of DNA lesions while accommodating suitable adjustments in for instance chromatin structure and cell cycle progression [1, 2]. Not only does the DDR dictate the appropriate repair pathway for several types of DNA damage, including DNA double-strand breaks (DSB), it also modulates replication fork surveillance mechanisms in response to DNA replication stress (RS). While many core proteins have been studied in detail, the full repertoire of factors involved in these pathways remains unknown. Clearly, extending our knowledge on regulators of the DDR will contribute to our understanding of the development, and possibly the treatment, of the numerous disorders that are associated with defects in the DDR [3]. The research described in this thesis has successfully identified and characterized novel factors in DSB repair and the RS response, yet raises additional questions that provide a base for further research and discussion.

The contribution of ZnF proteins to DNA double-strand break repair

DNA-templated reactions such as transcription, DNA replication and DNA damage repair, rely on proteins that bind to DNA to locally regulate these processes. However, the accessibility of these proteins at DNA damage sites is complicated by the packaging of DNA into chromatin. ATP-dependent chromatin remodelers and post-translational modifications (PTMs) are therefore required to change chromatin structure and allow for DNA damage repair to occur [4, 5]. The spatiotemporal recruitment of DNA repair factors is not only regulated by PTM recognition and protein-protein interactions, but also DNA binding at DNA damage sites. One class of proteins that can associate with PTMs, proteins and/or DNA are zinc finger domain-containing proteins (ZnF proteins). Over the last couple of years, ZnF proteins have gained much attention because of their roles in DSB repair by non-homologous end joining (NHEJ) and homologous recombination (HR) (described in chapter 2).

ZnF proteins are categorized in different subfamilies based on their numerous structural folds, diverse binding modes and sequence recognition sites present in their ZnF domains. To date, more than 50 unique domains have been identified, several of which have been shown to be important for DSB repair due to their binding to for instance ubiquitin, poly(ADP-ribose) or DNA [6]. In addition, several ZnF domain families have been implicated in processes other than DSB repair. For instance, FYVE domain-containing proteins are mostly localized at endosomes and golgi compartments where they bind to phospho-inositides via their FYVE domain to regulate specific functions of proteins in these compartments [7]. C2H2-KRAB domain-containing proteins, on the other hand, are involved in the repression of foreign DNA transcription, primarily that of endogenous retroviruses, and bind to the regulatory regions of retrotransposons [8]. However, KRAB domain-containing ZnF proteins are also involved in transcription repression by binding in a sequence-specific manner to cis-regulatory elements and recruiting the KAP-1 corepressor to these loci [9]. This illustrates that ZnF proteins, even within the same domain family, exhibit diverse binding specificities and functions.

With regard to functions in DNA damage repair, it is interesting to note that several ZnF proteins ranked high in CRISPR/Cas9 screens aimed at identifying protein networks that protect cells against DNA damaging agents, including several chemotherapeutics [10]. In addition, several Znf proteins were found to localize at sites of laser micro-irradiation induced DNA damage often in a manner dependent on PAR/PARP [11]. Although these screens highlight the potential importance of ZnF proteins in maintaining genome stability, the role of many of these ZnF proteins in DNA damage repair has remained uncharacterized thus far. Thus, further work is required to dissect their functional impact on the DSB response, but also in preventing human diseases such as cancer and neurodegenerative disorders [12]. The latter is of importance since mutations in ZnF proteins have been linked to broadly varying clinical consequences [12]. For instance, truncation and missense mutations in ZNF750 cause squamous cell carcinoma (SCC). Importantly, these mutations are located in the C2H2 domain, suggesting a critical role of this domain in mediating the tumor suppressor activity of ZNF750 [13]. Similarly, point mutations in the ZnF domain of tumor suppressor MDM2 have been linked to different tumours, including leukemia, hepatocellular carcinoma and osteosarcoma [14], whereas, mutations in E4F1, a ZnF protein that cooperates with PARP1 to promote DSB repair, have been associated with breast cancer [15]. This illustrates that although we have just begun to unravel the exact biological functions of ZnF proteins, their broad

6

implications in preventing human disease have already become apparent. It will be of interest to unravel to which extent their role therein relates to a function in the DDR, most notably DSB repair.

ZNF384, a novel NHEJ regulator

In this thesis, we uncovered a crucial role for the C2H2 type ZnF protein ZNF384 in ensuring the proper regulation of Ku70/Ku80, which are essential for DSB repair via NHEJ (described in chapter 3). ZNF384 regulates the functioning of Ku70/Ku80 by directly interacting with this complex. Our mass spectrometry analysis, however, revealed that ZNF384 not only interacts with Ku70/Ku80 but also with several other proteins including heat shock proteins (HSPA5, HSPA1B, HSPA8, HSPA9 and HSPH1) and replication factor C proteins (RFC1, RFC2, RFC4 and RFC4). Heat shock proteins are molecular chaperones that participate in important cellular processes such as protein folding. Their involvement in anti-cancer drug resistance have made them important targets in DNA damage repair [16]. In fact, several heat shock proteins (HSP) are implicated in DSB repair via HR and NHEJ [17]. For instance, HSPH1 also interacts with Ku70/Ku80, and promotes NHEJ in colorectal cancer cells [18]. This may suggest that ZNF384 protects genome stability in an even broader manner, for instance by the regulation of HSPs. The validation of these interactors by GFP pulldowns and laser micro-irradiation recruitment studies of GFP-tagged HSPs in the presence and absence of ZNF384 would be the first step in identifying the functional interplay between ZNF384 and HSPs in DNA damage repair. RFC complexes, on the other hand, are implicated in the unloading of PCNA during DNA replication [19]. The unloading of PCNA is tightly regulated to prevent premature termination of DNA replication. Since ZNF384 interacts with multiple components of the RFC complex, this may suggest that ZNF384 has a potential role in DNA replication.

In line with previous work, our DNA binding analysis revealed that ZNF384 has preference to bind to homopolymeric dA·dT DNA consensus elements in ssDNA (T-stretch), and does not associate with blunt dsDNA [20, 21]. Importantly, however, we also observed binding of ZNF384 to dsDNA with a 3' or 5' overhang. Consistent with its role in cNHEJ, dsDNA with 3' or 5' overhangs are better substrates for this type of NHEJ repair when compared to blunt dsDNA, which are preferentially repaired by polymerase theta-mediated end-joining [22]. It is important to note that all the DNA substrates used in this study contain the homopolymeric dA·dT DNA consensus

sequence. This raises the question whether ZNF384 binds to the ssDNA (T-stretch) and/or dsDNA moiety of these substrates with protruding ends. The emergence of a more detailed picture using structural biology analysis including Cryo-electron microscopy (cryo-EM) of recombinant ZNF384 bound to these substrates would provide insight into ZNF384's DNA binding mode [23]. Complementary to this, it would be interesting to confirm our *in vitro* results in cells by inducing DSBs with protruding or and blunt ends using the CRISPR/Cas9 nickase wildtype, and measure ZNF384 binding by chromatin-immunoprecipitation (ChIP) [24, 25]. To further extend on this, cell survival assays of ZNF384 depleted cells upon DSB induction with protruding *versus* blunt ends would provide insight into whether ZNF384 indeed promotes the repair of protruding ends. Taken together, future studies will definitively expand our knowledge on ZNF384's DSB end binding mode and NHEJ repair in vivo. Our current data reveal that the N-terminus of ZNF384 is responsible for a direct interaction with Ku70/Ku80 and promotes the loading of this complex at DSBs. However, the domains in Ku70/Ku80 required for their interaction with ZNF384 remain unknown. The identification of these domains will allow better insight into the mechanism by which ZNF384 mediates the contacts of Ku70/Ku80 with DSBs. Since the crystal structure of the Ku heterodimer has already been determined [26], structural analysis with ZNF384 bound to Ku70/Ku80 in the presence of protruding DNA substrates used in this study would definitely provide answers on how ZNF384 assists in Ku making DSB ends meet. However, it is important to note that the C2H2 motifs in ZNF384 have preference to bind T-rich DNA, whereas Ku70/Ku80 bind DSB ends in a sequence independent manner. Whether ZNF384 can also bind other sequences remains yet to be identified. Importantly, its substrate recognition properties are determined by the cysteines and histidine amino acids on each motif which coordinate zinc ions to mediate contact with DNA [20]. Therefore, amino acids substitutions in these residues may modulate ZNF384's DNA binding specificity to regions other than T-rich DNA. The generation of such mutants may significantly aid in determining whether ZNF384's DNA binding sequence specificity is required for the loading of Ku70/Ku80 at DSBs to promote cNHEJ.

Interplay between ZNF384 and DNA-PKcs during cNHEJ

To our surprise we found that ZNF384 promotes the loading of APLF and XRCC4/LIG4 in a manner dependent of PARP1 and Ku, but independently of DNA-PKcs. This is in contrast with previous work suggesting that XRCC4/LIG4 assembly depends on the

recruitment and activation of DNA-PKcs by Ku-bound DNA-ends [27]. We found that DNA-PKcs activity remained intact in Ku70/Ku80 knockdown human cells and *Ku80* KO mouse embryonic stem cells. This may suggest that the presence of DNA, regardless of Ku, is sufficient for DNA-PKcs activation [28]. This is in line with a recent study showing that Ku and XRCC4/LIG4, XLF and PAXX are sufficient, while DNA-PKcs is dispensable for bringing broken DNA ends together *in vitro* in a process called DNA-end synapsis [29]. Extending our epistasis analysis to decipher the interplay between ZNF384 and PAXX or XLF, as well as the interplay between Ku and DNA-PKcs, would be an important next step to gain insight into the functioning of the NHEJ apparatus. Interestingly, DNA-bound Ku is essential for the initial recruitment of PAXX at DSBs [30]. This may suggest that ZNF384 and PAXX cooperate during NHEJ, in which ZNF384 ensures the initial loading of Ku70/Ku80 and PAXX promotes Ku70/Ku80 retention at DSBs. Furthermore, another study reported that destabilization of the APLF or XLF interaction with Ku80 impairs NHEJ due to defects in loading of XRCC4 and Ku [31]. Whether these observations are linked to defects in DNA-Pkcs recruitment and/ or activation remains unclear. To this end, it would be interesting to measure NHEJ of protruding DSB ends using an *in vitro* NHEJ reconstitution system with ZNF384 and combinations of other NHEJ components, including Ku70/Ku80, DNA-PKcs, APLF, PAXX, XLF or XRCC4/LIG4, all of which have been purified previously [32, 33]. In addition, atomic force microscopy (AFM) has been proven to serve as a powerful tool to study protein-DNA interactions and end-tethering by NHEJ proteins, and thus may be a valuable add-on to the *in vitro* studies [34]. Finally, deciphering the spatiotemporal recruitment of the NHEJ apparatus in cells would be an indispensable extension to this research. Since our data reveal that there is redundancy between ZNF384 and DNA-PKcs, it may be fact that they act mutually exclusive at DSBs. Genome-wide recruitment studies of these and other NHEJ components by ChIP-sequencing at AsiSI- and CRISPR/ Cas9 nickase-induced DSBs may determine which NHEJ factors act in concert at these distinct DNA breaks [35, 36].

ZNF384: an oncogene with therapeutic potential

ZNF384 has been implicated in various cancers by acting as an oncogene. For instance, overexpression of ZNF384 results in hepatocellular carcinoma tumor growth via the upregulation of cyclin D1 expression [37]. Furthermore, it has been reported that overexpression of ZNF384 can promote metastasis in melanoma cells [37]. More

interestingly, ZNF384 is found to be rearranged in 48% of mixed phenotype acute leukemia (MPAL) patients [38]. These chromosomal rearrangements involve the fusion of ZNF384 with transcription regulators and epigenetic modifiers such as TCF3, EP300, TAF15 and CREBBP. However, the underlying disease-causing mechanism related to the expression of these fusion proteins remains unclear. Given the role of ZNF384 in NHEJ, one possibility would be to express these ZNF384 fusion proteins in our U2OS Flp-In/ Trex ZNF384 KO cells and examine their effect on NHEJ. Another hallmark of acute leukemia are defects in class switch recombination (CSR) and V(D)J recombination [39], both of which are processes that rely on cNHEJ [40]. A step towards understanding the disease mechanism of MPAL would be to investigate the effect of ZNF384 and these ZNF384-fusion proteins on CSR and V(D)J recombination. Finally, C2H2 ZnF proteins have become attractive therapeutic targets. A recent study screened the human C2H2 proteome for degradation by thalidomide analogs, which induces the ubiquitylation and degradation of C2H2 containing transcription factors (e.g. IKZF1) by targeting this domain to the substrate receptor of the $CRL4_{CRBN}$ E3 ligase [41]. These analogs are clinically approved therapies for the treatment of multiple myeloma and other hematologic malignancies. Whilst ZNF384 function in health and disease mechanism becomes better understood, developing small molecule inhibitors to target ZNF384 may be worthwhile in the future.

ZNF384 and other ZnF proteins in DSB repair

Besides ZNF384, several other ZnF proteins have been implicated in NHEJ. For instance, APLF and ZBTB24 possess distinct ZnF domains (PBZ in APLF and C2H2 in ZBTB24) that are required for the build-up of a functional NHEJ complex by binding to auto-mono(ADP-ribosyl)ated (MAR) PARP3, PARylated PARP1, respectively, at DSBs [42, 43]. Another C2H2 ZnF protein, ZNF281, is recruited to DSBs via PARP activity and through its C2H2-domain. Whether the latter involves its binding to PAR or DNA is unclear. Its C2H2-domain also interacts with XRCC4, thereby facilitating XRCC4 recruitment and cNHEJ [44]. In addition, ZNF281 binds to the *XRCC4* promotor via its C2H2-domain, thereby controlling XRCC4 expression and cNHEJ [45]. It is therefore tempting to speculate that ZNF384 may cooperate with other ZnF proteins to promote efficient NHEJ repair. In fact, our data suggest that ZNF384 cooperates with ZnF protein PARP1, to promote the efficient loading of XRCC4. Possibly, this involves the PARP1-dependent PARylation of ZNF384, which is important for XRCC4 recruitment [42].

Moreover, previous studies reported that other ZnF proteins are also recruited to sites of DNA damage in a manner often dependent on PARP/PAR, indicating their importance in NHEJ or other DNA repair pathways [11, 46]. This raises the question as to why different ZnF proteins are required for efficient NHEJ repair. We speculate that their versatile substrate recognition features may help to load and accommodate the different cNHEJ repairosome components at DSBs, thereby promoting efficient DSB repair.

Several ZnF proteins have been implicated in DNA repair pathways other than NHEJ and HR [6]. For instance, the C2H2 domain-containing protein Aprataxin, encoded by the *APTX* gene that is frequently mutated in the neurological disorder ataxia oculomotor apraxia, catalyzes the removal of adenylate residues from 5' termini at nicks and gaps in duplex DNA to promote efficient ligation via the DNA repair proteins XRCC1 and XRCC4 [47]. Moreover, ZNF451 aids in the resolution of trapped TOP2 cleavage complexes (TOP2cc) by stimulating SUMOylation of TOP2cc via the activity of TDP2, a protein involved in the resolution of TOP2-dependent protein-DNA crosslinks [48]. Interestingly, several ZnF proteins, including ZSCAN4, ZNF827 and TZAP, have also been implicated in telomere maintenance [49]. While ZSCAN4 is required for telomere extension during the early stages of embryogenesis [50], ZNF827 binds telomeres and recruits the NuRD chromatin remodeling complex to facilitate HR and telomere elongation mediated by the alternative lengthening of telomeres (ALT) pathway [51]. Furthermore, TZAP contains 11 C2H2 motifs, three of which bind specifically to canonical telomeric repeats TTAGGG. By binding to these repeats, TZAP promotes the rapid deletion of telomeres to prevent them from becoming extremely elongated, thereby preventing genome instability [52]. Taken together, these studies highlight the significant contribution of ZnF proteins in different genome integrity pathways. However, with regards to DBS repair, their roles in pathways other than HR or NHEJ (for instance aNHEJ and SSA), as well as their roles in DSB repair pathway choice remains to be elucidated.

R-loops and replication fork stability: the NSL-complex comes into focus

Transcription is a DNA-templated process that requires proper regulation. If not regulated properly or incorrectly, the RNAPII transcription machinery may stall and form obstacles to replication fork progression, leading to transcription-replication conflicts (TRCs) that mostly occur in a head-on orientation [53]. TRCs frequently involve the generation and persistence of harmful R-loops, which can be formed by RNAPII stalling and are a

source of genome instability [54]. In the recent years numerous studies have reported on the interplay between transcription and R-loops and its impact on genome instability. However, although through these studies the network of R-loop regulators is expanding, a full understanding of the context in which R-loops arise and cause TRC-associated genome instability is lacking.

Chromatin factors have emerged as important players in R-loop formation [55]. These factors include member of the MSL (male specific lethal) histone acetyltransferase (HAT) complex MOF, which has been reported to suppress replication fork stalling and R-loop formation [56]. In contrast, however, we found that MOF regulates H4K16Ac levels, but has no significant effect on R-loops or S-phase progression. MOF is shared within the non-specific lethal (NSL) complex which consists of four core members KANSL1, KANSL2, KANSL3 and PHF20 [57]. Interestingly, we found that KANSL3, in contrast to MOF, is dispensable for H4K16Ac, agreeing with a previous report [58, 59]. Furthermore, KANSL3 protects cells from HU-induced RS and is implicated in the restart of stalled replication forks. While previous studies have shed light on KANSL3's role in transcription [58, 60], our data suggest that KANSL3 interacts with proteins involved transcription and RNA splicing. Since defects in these processes are causally linked to R-loops, it is therefore more likely that KANSL3 promotes fork stability via its role in the suppression of aberrant R-loops [53]. To this end, it will be interesting to decipher the genomic localization at which these R-loops are enriched in KANSL3-depleted cells, and whether they occur at genes regulated by the NSL complex or also at other genes. In depth analysis of the R-loop distribution by DRIP-sequencing, as well as RNA sequencing to study altered mRNA expression, in KANSL3 depleted cells will therefore be required [58]. Components of the MSL and NSL complex, however, are essential in mammalian cells [61], and siRNA knockdown experiments are not ideal for these type of large scale experiments. Therefore, the endogenous tagging of the NSL complex using CRISPR/Cas9 with the auxin inducible degron in OsTIR1 expressing cells would be a more feasible approach to rapidly and transiently deplete this complex [62]. Given that KANSL3 depleted cells suffer from RS-induced fork collapse, we speculate that this is due to transcription-replication conflicts (TRCs) occurring in the head-on direction (HO), in which the transcription machinery and replication machinery progress in opposite directions [63]. HO collisions induce pausing and blockage of the replication fork and are therefore considered more toxic. Our data suggest that KANSL3 affects replication fork stability/recovery following stress and we speculate that this is due to R-loop induced head-on TRCs.

6

Mutations or deregulation of MSL and NSL proteins are associated with severe human diseases including cancer. For instance, KANSL2 is reportedly upregulated in glioblastoma and drives the stem cell-like features of glioblastoma cells [64]. KANSL1 haploinsufficiency, on the other hand, impairs the activity of MOF to acetylate histone H4 on lysine 16 [65], a defect that has been associated with a wide range of malignancies including skin carcinogenesis [66]. It is also interesting to note that MOF is more stable when associated with the NSL complex, but that cancer cells that suffer from decreased expression of MOF (e.g. ovarian cancer) still display profound levels of NSL activity, while H4K16Ac levels are reduced [58]. However, a recent study reported that H4K16Ac is not required for cell proliferation and differentiation, which are common hallmarks of cancer, questioning how loss of MOF affects cancer development [58]. Our data suggest that MOF does not influence R-loop levels or fork stability, which could otherwise also affect carcinogenesis. Therefore, we infer that low MOF-expressing cancer cells suffer from defects in other MOF-dependent processes such as fatty acid oxidation and mitochondrial respiration, or apoptosis [67]. Moreover, as the catalytic subunit of NSL, MOF also catalyzes H4K5 and H4K8 acetylation, which are important marks for the initiation of transcription of housekeeping genes and for cell survival [58]. Whether loss of this function impacts cancer development is largely unclear. Our data suggest that KANSL3-depleted cells display increased levels of R-loops and fork instability as measured by pRPA (S4/S8), and suffer from reduced DNA synthesis. It is tempting to speculate that these phenotypes reflect a loss of MOF/NSL's acetylation activity towards H4K5 and H4K8, rather than MOF/MSL-dependent H4K16Ac. To this end, an epistasis analysis between members of the NSL and MSL complexes would be a starting point to better understand their interplay in these processes. Ultimately, our data reveal a dual role for the NSL-complex in which on one hand MOF promotes the formation of H4K16Ac which is dispensable for R-loop homeostasis and fork stability. On the other hand, KANSL3 suppresses R-loop formation, and may thereby prevent collisions between the replication fork and the transcription machinery that can lead to replication fork arrest/collapse. Although, the mechanism leading to this dual function within the NSL complex remains to be determined, it would be interesting to decipher whether R-loops and TRCs occur at genomic regions distinct from those enriched for H4K16Ac.

Deciphering the role of exonuclease ERI1 in R-loop suppression and fork stability

The 3'-to-5' exonucleolytic degradation and processing of a variety of RNAs is an essential feature of RNA metabolism in all cells [68]. This involves the activity of RNA processing enzymes during ribosomal RNA maturation, polyadenylation, small RNA production, and histone mRNA regulation. On the other hand, RNA can also exist in RNA-DNA hybrids or R-loops. Several RNA binding proteins are implicated in the processing of R-loops such as RNaseH1 and RNaseH2 [69]. We uncovered a novel function of the DEDDh exonuclease family member ERI1 as a suppressor of R-loop formation and replication fork instability. However, how ERI1 mechanistically impacts these processes remains to be resolved. Our mass spectrometry analysis revealed that ERI1 has various interaction partners including mRNA binding proteins (RNPS1 and PNN), RNA helicases (DHX15), chromatin remodelers (SMARCA5 and CHD4), histone proteins (HP1BP3, H1F0 and H2AFY), RNA splicing factors (THRAP3) and the DNA repair protein PARP1 (Table S1 in Chapter 5). Although, the interaction for the majority of these factors needs to be confirmed, the data may already indicate that ERI1 maintains fork stability in a broader manner then our current study indicates. Interestingly, we also successfully validated the interaction between ERI1 and PARP1 interaction by GFP pulldowns, although it needs to be confirmed reciprocally and endogenously. We found that ERI1 recruits rapidly but transiently at laser micro-irradiation induced DNA damage, in a manner dependent on the activity of PARP. This could be a consequence of the PARylation of ERI1 or the binding of ERI1 to PAR chains directly, which should be further investigated by for instance GFP-ERI1 pulldowns under denaturing conditions followed by western blot analysis and *in vitro* PAR-binding assays using recombinant ERI1 respectively [42]. ERI1 is recruited to (stalled) replication forks as measured by isolation of proteins on nascent DNA (iPOND), and the next logical step would be to investigate whether this recruitment is dependent on PAR/PARP1. PARP1 associates with (stalled) DNA replication forks and PARP1 auto-PARylation triggers fork reversal, a mechanism that ensures fork stabilization upon RS [70, 71]. In addition, it functions as a sensor of unligated Okazaki fragments during unchallenged DNA replication, and facilitates their repair [72]. Investigating the potential functions of ERI1 in these processes and the spatiotemporal modulation of the ERI1-PARP1 interaction could shed light on its contribution to replication fork stability.

6

ERI1 contains a conserved SAF-box, Acinus and PIAS (SAP) domain, which binds to double-stranded DNA and RNA. However, it is unclear whether the SAP domain primarily anchors ERI1 to RNA or additionally recruits proteins to RNA-dependent complexes that modify substrates such as the histone mRNA 3' stem-loop [73]. Alternatively, the SAP domain may prevent the degradation of RNA [74]. In order to establish whether this domain is responsible for ERI1's recruitment at stalled replication forks, investigating the recruitment of SAP domain mutants in iPOND experiments will be required in the future [75]. Furthermore, ERI1 contains an exonuclease domain which is involved in the processing of RNA. Mutations in this domain result in fork instability upon HU-induced RS. Since ERI1 is involved in the regulation of R-loops, which are often associated with transcription-replication conflicts (TRCs), we speculate that R-loops might be responsible for the observed RS in ERI1-depleted cells. To investigate the role of ERI1 in RNA processing and fork stability, several future experiments are required for instance: 1) *in vitro* R-loop formation assays using purified ERI1 wildtype and exonuclease dead to determine whether ERI1 processes R-loops and if this is dependent on its exonuclease domain [76]. 2) RNaseH1 overexpression experiments to study whether fork stability can be restored in ERI- depleted cells and 3) proximity ligation assay (PLA) between PCNA, as a marker for replication, and RNAPII to evaluate the occurrence of TRCs in absence of ERI1 [55].

The fidelity of DNA polymerases is challenged by the 10-200 fold higher abundance of ribonucleotide triphosphates (rNTPs) versus DNA deoxyribonucleotide triphosphates (dNTPs) in mammalian cells [77]. Therefore, rNTPs are frequently mis-incorporated into genomic DNA during normal DNA replication, leading to replication stress, mutagenesis and DNA breaks [78]. RNaseH2 is the main enzyme that removes these rNTPs via a process called ribonucleotide excision repair (RER) [79]. Since ERI1 has a wide range of potential RNA substrates (e.g. ribosomal RNA, histone mRNAs and small regulatory RNAs), it may be possible that ERI1 is responsible for the removal of miss incorporated ribonucleotides during DNA replication, and thereby maintains fork stability. Genome-wide mapping of ribonucleotides by for instance Ribose-seq in both control and ERI1-depleted cells may provide insight into whether ERI1 is involved in RER [78]. On the other hand, ERI1 is required for the degradation of histone mRNAs at the end of DNA replication [80, 81]. In fact, ERI1 binds to the 3' and 5' stem-loop arms of the histone mRNAs to ensure their degradation at the end of each S-phase [82]. The balance of DNA and histone synthesis is important since increased levels of histones

can cause chromosome loss and genome instability [83]. We speculate that the genome instability observed in ERI1-depleted cells is a consequence of TRCs causes by R-loops. To this end, deciphering the genomic locations that are enriched for R-loops in these cells (using DRIP-sequencing), and uncovering whether changes in R-loop levels are affected by changes in histone incorporation (using chromatin occupancy after replication-seq (ChOR-seq) would be worthwhile future endeavors [84]. Morover, ERI1 activity is restricted by SLBP, a key protein recruited to the histone 3' UTR, where it stabilizes and promotes the translation of histone mRNAs. At the end of S-phase, SLBP is degraded by the proteasome, permitting rapid degradation of histone mRNAs by ERI1 [85]. Artificial tethering of SLBP at a genomic site enriched for R-loops could serve as an alternative approach to investigate whether inhibition of histone mRNA degradation is linked to an increase in R-loops [86].

Replication dependent histone supply is required for S-phase entry and DNA replication [87, 88]. This raises the question as to how impaired histone mRNA degradation as previously observed in ERI1 deficient cells, results in the formation of R-loops and replication stress [80]. It is tempting to speculate that disruption of histone mRNA degradation could disturb nucleosome assembly which relies on local recycling of parental histones along with deposition of newly synthesized histones in S-phase. Defects in this process are linked to developmental disorders and ageing [89]. Moreover, impaired nucleosome assembly is associated with delays in PCNA unloading and R-loop formation [87, 90]. Further work will improve our understanding on whether ERI1 suppresses R-loops and replication stress by directly acting at R-loops, or via its role in histone mRNA processing. Ultimately, this will extend our mechanistic insights into how histone biosynthesis is linked to R-loops and fork stability, and how defects in these processes can cause disease.
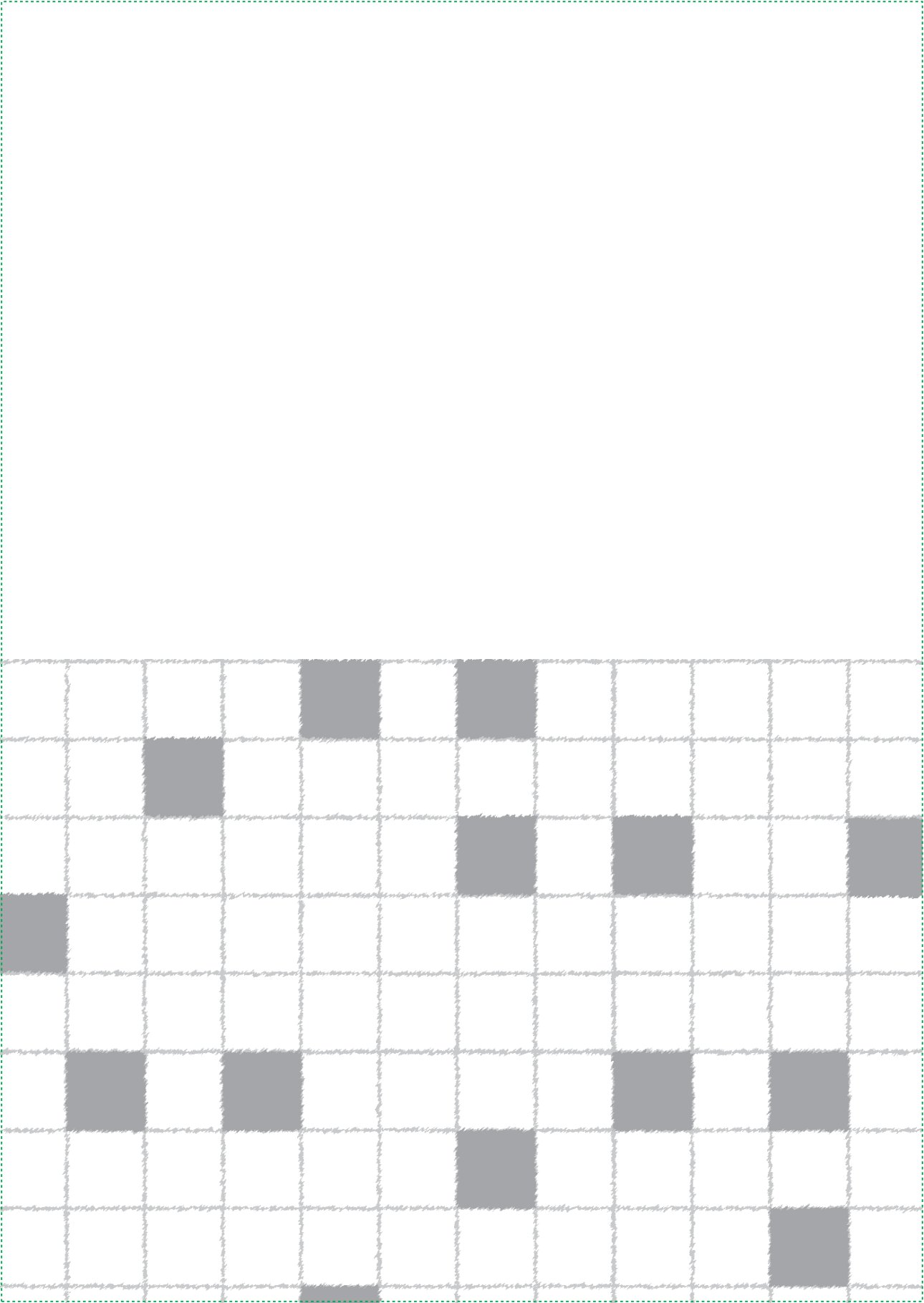
6

REFERENCES

1. Jackson, S.P. and J. Bartek, *The DNA-damage response in human biology and disease.* Nature, 2009. **461**(7267): p. 1071-8.

2. Lukas, J., C. Lukas, and J. Bartek, *More than just a focus: The chromatin response to DNA damage and its role in genome integrity maintenance.* Nat Cell Biol, 2011. **13**(10): p. 1161-9.

3. O'Driscoll, M., *Diseases associated with defective responses to DNA damage.* Cold Spring Harbor perspectives in biology, 2012. **4**(12): p. a012773.

4. Dantuma, N.P. and H. van Attikum, *Spatiotemporal regulation of posttranslational modifications in the DNA damage response.* Embo j, 2016. **35**(1): p. 6-23.

5. Caron, P., et al., *WWP2 ubiquitylates RNA polymerase II for DNA-PK-dependent transcription arrest and repair at DNA breaks.* Genes Dev, 2019. **33**(11-12): p. 684-704.

6. Singh, J.K. and H. van Attikum, *DNA double-strand break repair: Putting zinc fingers on the sore spot.* Semin Cell Dev Biol, 2020.

7. Ridley, S.H., et al., *FENS-1 and DFCP1 are FYVE domain-containing proteins with distinct functions in the endosomal and Golgi compartments.* Journal of Cell Science, 2001. **114**(22): p. 3991-4000.

8. Richardson, S.R., et al., *The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes.* Microbiology spectrum, 2015. **3**(2): p. MDNA3-2014.

9. Schultz, D.C., et al., *SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins.* Genes Dev, 2002. **16**(8): p. 919-32.

10. Balmus, G., et al., *ATM orchestrates the DNA-damage response to counter toxic non-homologous end-joining at broken replication forks.* Nature Communications, 2019. **10**(1): p. 87.

11. Izhar, L., et al., *A Systematic Analysis of Factors Localized to Damaged Chromatin Reveals PARP-Dependent Recruitment of Transcription Factors.* Cell Rep, 2015. **11**(9): p. 1486-500.

12. Cassandri, M., et al., *Zinc-finger proteins in health and disease.* Cell Death Discov, 2017. **3**: p. 17071.

13. Hazawa, M., et al., *ZNF750 is a lineage-specific tumour suppressor in squamous cell carcinoma.* Oncogene, 2017. **36**(16): p. 2243-2254.

14. Schlott, T., et al., *Point mutations and nucleotide insertions in the MDM2 zinc finger structure of human tumours.* J Pathol, 1997. **182**(1): p. 54-61.

15. Moison, C., et al., *Zinc finger protein E4F1 cooperates with PARP-1 and BRG1 to promote DNA double-strand break repair.* Proceedings of the National Academy of Sciences, 2021. **118**(11): p. e2019408118.

16. Sottile, M.L. and S.B. Nadin, *Heat shock proteins and DNA repair mechanisms: an updated overview.* Cell stress & chaperones, 2018. **23**(3): p. 303-315.

17. Dubrez, L., et al., *Heat-shock proteins: chaperoning DNA repair.* Oncogene, 2020. **39**(3): p. 516-529.

18. Causse, S.Z., et al., *HSP110 translocates to the nucleus upon genotoxic chemotherapy and promotes DNA repair in colorectal cancer cells.* Oncogene, 2019. **38**(15): p. 2767-2777.

19. Kang, M.-S., et al., *Regulation of PCNA cycling on replicating DNA by RFC and RFC-like complexes.* Nature Communications, 2019. **10**(1): p. 2420.

20. Torrungruang, K., et al., *DNA binding and gene activation properties of the Nmp4 nuclear matrix transcription factors.* J Biol Chem, 2002. **277**(18): p. 16153-9.

21. Nakamoto, T., et al., *CIZ, a zinc finger protein that interacts with p130(cas) and activates the expression of matrix metalloproteinases.* Mol Cell Biol, 2000. **20**(5): p. 1649-58.

22. Schimmel, J., et al., *Mutational signatures of non-homologous and polymerase theta-mediated end-joining in embryonic stem cells.* Embo j, 2017. **36**(24): p. 3634-3649.

23. Lamers, M.H., et al., *The crystal structure of DNA mismatch repair protein MutS binding to a G·T mismatch.* Nature, 2000. **407**(6805): p. 711-717.

24. Schimmel, J., et al., *Small tandem DNA duplications result from CST-guided Pol -primase action at DNA break termini.* Nat Commun, 2021. **12**(1): p. 4843.

25. Arnoult, N., et al., *Regulation of DNA repair pathway choice in S and G2 phases by the NHEJ inhibitor CYREN.* Nature, 2017. **549**(7673): p. 548-552.

26. Walker, J.R., R.A. Corpina, and J. Goldberg, *Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair.* Nature, 2001. **412**(6847): p. 607-14.

27. Costantini, S., et al., *Interaction of the Ku heterodimer with the DNA ligase IV/Xrcc4 complex and its regulation by DNA-PK.* DNA Repair (Amst), 2007. **6**(6): p. 712-22.

28. Hammarsten, O. and G. Chu, *DNA-dependent protein kinase: DNA binding and activation in the absence of Ku.* Proceedings of the National Academy of Sciences, 1998. **95**(2): p. 525-530.

29. Zhao, B., et al., *The essential elements for the noncovalent association of two DNA ends during NHEJ synapsis.* Nat Commun, 2019. **10**(1): p. 3588.

30. Liu, X., et al., *PAXX promotes KU accumulation at DNA breaks and is essential for end-joining in XLF-deficient mice.* Nature Communications, 2017. **8**(1): p. 13816.

31. Nemoz, C., et al., *XLF and APLF bind Ku80 at two remote sites to ensure DNA repair by non-homologous end joining.* Nature Structural & Molecular Biology, 2018. **25**(10): p. 971-980.

32. Stinson, B.M., et al., *A Mechanism to Minimize Errors during Non-homologous End Joining.* Molecular Cell, 2020. **77**(5): p. 1080-1091.e8.

33. Chang, H.H.Y., et al., *Different DNA End Configurations Dictate Which NHEJ Components Are Most Important for Joining Efficiency.* J Biol Chem, 2016. **291**(47): p. 24377-24389.

34. Yaneva, M., T. Kowalewski, and M.R. Lieber, *Interaction of DNA-dependent protein kinase with DNA and with Ku: biochemical and atomic-force microscopy studies.* Embo j, 1997. **16**(16): p. 5098-112.

35. Caron, M.C., et al., *Poly(ADP-ribose) polymerase-1 antagonizes DNA resection at double-strand breaks.* Nat Commun, 2019. **10**(1): p. 2954.

36. Aymard, F., et al., *Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks.* Nature structural & molecular biology, 2014. **21**(4): p. 366-374.

37. He, L., et al., *Overexpression of zinc finger protein 384 (ZNF 384), a poor prognostic predictor, promotes cell growth by upregulating the expression of Cyclin D1 in Hepatocellular carcinoma.* Cell Death & Disease, 2019. **10**(6): p. 444.

38. Alexander, T.B., et al., *The genetic basis and cell of origin of mixed phenotype acute leukaemia.* Nature, 2018. **562**(7727): p. 373-379.

39. Aplan, P.D., *Causes of oncogenic chromosomal translocation.* Trends Genet, 2006. **22**(1): p. 46-55.

6

40. Alt, Frederick W., et al., *Mechanisms of Programmed DNA Lesions and Genomic Instability in the Immune System.* Cell, 2013. **152**(3): p. 417-429.

41. Sievers, Q.L., et al., *Defining the human C2H2 zinc finger degrome targeted by thalidomide analogs through CRBN.* Science, 2018. **362**(6414): p. eaat0572.

42. Helfricht, A., et al., *Loss of ZBTB24 impairs non-homologous end-joining and class-switch recombination in patients with ICF syndrome.* Journal of Experimental Medicine, 2020.

43. Grundy, G.J., et al., *APLF promotes the assembly and activity of non-homologous end joining protein complexes.* Embo j, 2013. **32**(1): p. 112-25.

44. Nicolai, S., et al., *ZNF281 is recruited on DNA breaks to facilitate DNA repair by non-homologous end joining.* Oncogene, 2020. **39**(4): p. 754-766.

45. Pieraccioli, M., et al., *ZNF281 contributes to the DNA damage response by controlling the expression of XRCC2 and XRCC4.* Oncogene, 2016. **35**(20): p. 2592-2601.

46. Smith, R., et al., *Poly(ADP-ribose)-dependent chromatin unfolding facilitates the association of DNA-binding proteins with DNA at sites of damage.* Nucleic Acids Res, 2019. **47**(21): p. 11250-11267.

47. Ahel, I., et al., *The neurodegenerative disease protein aprataxin resolves abortive DNA ligation intermediates.* Nature, 2006. **443**(7112): p. 713-716.

48. Schellenberg, M.J., et al., *ZATT (ZNF451)-mediated resolution of topoisomerase 2 DNA-protein cross-links.* Science, 2017. **357**(6358): p. 1412-1416.

49. Vilas, C.K., et al., *Caught with One's Zinc Fingers in the Genome Integrity Cookie Jar.* Trends Genet, 2018. **34**(4): p. 313-325.

50. Dan, J., et al., *Zscan4 Inhibits Maintenance DNA Methylation to Facilitate Telomere Elongation in Mouse Embryonic Stem Cells.* Cell Rep, 2017. **20**(8): p. 1936-1949.

51. Conomos, D., R.R. Reddel, and H.A. Pickett, *NuRD-ZNF827 recruitment to telomeres creates a molecular scaffold for homologous recombination.* Nat Struct Mol Biol, 2014. **21**(9): p. 760-70.

52. Li, J.S., et al., *TZAP: A telomere-associated protein involved in telomere length control.* Science, 2017. **355**(6325): p. 638-641.

53. Aguilera, A. and T. García-Muse, *R Loops: From Transcription Byproducts to Threats to Genome Stability.* Molecular Cell, 2012. **46**(2): p. 115-124.

54. Gómez-González, B. and A. Aguilera, *Transcription-mediated replication hindrance: a major driver of genome instability.* Genes & development, 2019. **33**(15-16): p. 1008-1026.

55. Bayona-Feliu, A., et al., *The SWI/SNF chromatin remodeling complex helps resolve R-loop-mediated transcription–replication conflicts.* Nature Genetics, 2021.

56. Singh, D.K., et al., *MOF Suppresses Replication Stress and Contributes to Resolution of Stalled Replication Forks.* Mol Cell Biol, 2018. **38**(6).

57. Sheikh, B.N., S. Guhathakurta, and A. Akhtar, *The non-specific lethal (NSL) complex at the crossroads of transcriptional control and cellular homeostasis.* EMBO Rep, 2019. **20**(7): p. e47630.

58. Radzisheuskaya, A., et al., *Complex-dependent histone acetyltransferase activity of KAT8 determines its role in transcription and cellular homeostasis.* Mol Cell, 2021. **81**(8): p. 1749-1765.e8.

59. Karoutas, A., et al., *The NSL complex maintains nuclear architecture stability via lamin A/C acetylation.* Nature Cell Biology, 2019. **21**(10): p. 1248-1260.

60. Lam, K.C., et al., *The NSL complex-mediated nucleosome landscape is required to maintain transcription fidelity and suppression of transcription noise.* Genes Dev, 2019. **33**(7-8): p. 452-465.

61. Hart, T., et al., *High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities.* Cell, 2015. **163**(6): p. 1515-1526.

62. Natsume, T., et al., *Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors.* Cell Rep, 2016. **15**(1): p. 210-218.

63. García-Muse, T. and A. Aguilera, *Transcription–replication conflicts: how they occur and how they are resolved.* Nature Reviews Molecular Cell Biology, 2016. **17**(9): p. 553-563.

64. Ferreyra Solari, N.E., et al., *The NSL Chromatin-Modifying Complex Subunit KANSL2 Regulates Cancer Stem-like Properties in Glioblastoma That Contribute to Tumorigenesis.* Cancer Res, 2016. **76**(18): p. 5383-94.

65. Li, X., et al., *Two Mammalian MOF Complexes Regulate Transcription Activation by Distinct Mechanisms.* Molecular Cell, 2009. **36**(2): p. 290-301.

66. Fraga, M.F., et al., *Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer.* Nature Genetics, 2005. **37**(4): p. 391-400.

67. Su, J., et al., *The Functional Analysis of Histone Acetyltransferase MOF in Tumorigenesis.* International journal of molecular sciences, 2016. **17**(1): p. 99.

68. Ibrahim, H., J. Wilusz, and C.J. Wilusz, *RNA recognition by 3'-to-5' exonucleases: the substrate perspective.* Biochimica et biophysica acta, 2008. **1779**(4): p. 256-265.

69. Hiller, B., et al., *Mammalian RNase H2 removes ribonucleotides from DNA to maintain genome integrity.* J Exp Med, 2012. **209**(8): p. 1419-26.

70. Zellweger, R., et al., *Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells.* The Journal of cell biology, 2015. **208**(5): p. 563-579.

71. Sugimura, K., et al., *PARP-1 ensures regulation of replication fork progression by homologous recombination on damaged DNA.* J Cell Biol, 2008. **183**(7): p. 1203-12.

72. Hanzlikova, H., et al., *The Importance of Poly(ADP-Ribose) Polymerase as a Sensor of Unligated Okazaki Fragments during DNA Replication.* Molecular Cell, 2018. **71**(2): p. 319-331.e3.

73. Tan, D., et al., *Structure of Histone mRNA Stem-Loop, Human Stem-Loop Binding Protein, and 3′hExo Ternary Complex.* Science, 2013. **339**(6117): p. 318-321.

74. Thomas, M.F., N.D. L'Etoile, and K.M. Ansel, *Eri1: a conserved enzyme at the crossroads of multiple RNA-processing pathways.* Trends in genetics : TIG, 2014. **30**(7): p. 298-307.

75. Yang, X.-c., et al., *Characterization of 3'hExo, a 3' Exonuclease Specifically Interacting with the 3' End of Histone mRNA*.* Journal of Biological Chemistry, 2006. **281**(41): p. 30447-30454.

76. Kasahara, M., et al., *RecA protein-dependent R-loop formation in vitro.* Genes & development, 2000. **14**(3): p. 360-365.

77. Traut, T.W., *Physiological concentrations of purines and pyrimidines.* Molecular and Cellular Biochemistry, 1994. **140**(1): p. 1-22.

78. Zhou, Z.-X., et al., *Ribonucleotide incorporation into DNA during DNA replication and its consequences.* Critical reviews in biochemistry and molecular biology, 2021. **56**(1): p. 109-124.

79. Kellner, V. and B. Luke, *Molecular and physiological consequences of faulty eukaryotic ribonucleotide excision repair.* The EMBO Journal, 2020. **39**(3): p. e102309.

6

80. Hoefig, K.P., et al., *Eri1 degrades the stem-loop of oligouridylated histone mRNAs to induce replication-dependent decay.* Nat Struct Mol Biol, 2013. **20**(1): p. 73-81.

81. Hoefig, K.P. and V. Heissmeyer, *Degradation of oligouridylated histone mRNAs: see UUUUU and goodbye.* WIREs RNA, 2014. **5**(4): p. 577-589.

82. Mullen, T.E. and W.F. Marzluff, *Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'.* Genes Dev, 2008. **22**(1): p. 50-65.

83. Meeks-Wagner, D. and L.H. Hartwell, *Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission.* Cell, 1986. **44**(1): p. 43-52.

84. Reverón-Gómez, N., et al., *Accurate Recycling of Parental Histones Reproduces the Histone Modification Landscape during DNA Replication.* Mol Cell, 2018. **72**(2): p. 239-249.e5.

85. Whitfield, M.L., et al., *Stem-loop binding protein, the protein that binds the 3' end of histone mRNA, is cell cycle regulated by both translational and posttranslational mechanisms.* Mol Cell Biol, 2000. **20**(12): p. 4188-98.

86. Prendergast, L., et al., *Resolution of R-loops by INO80 promotes DNA replication and maintains cancer cell proliferation and viability.* Nature Communications, 2020. **11**(1): p. 4534.

87. Mejlvang, J., et al., *New histone supply regulates replication fork speed and PCNA unloading.* The Journal of cell biology, 2014. **204**(1): p. 29-43.

88. Armstrong, C. and S.L. Spencer, *Replication-dependent histone biosynthesis is coupled to cell-cycle commitment.* Proceedings of the National Academy of Sciences, 2021. **118**(31): p. e2100178118.

89. Alabert, C. and A. Groth, *Chromatin replication and epigenome maintenance.* Nat Rev Mol Cell Biol, 2012. **13**(3): p. 153-67.

90. Kim, S., et al., *ATAD5 restricts R-loop formation through PCNA unloading and RNA helicase maintenance at the replication fork.* Nucleic Acids Res, 2020. **48**(13): p. 7218-7238.

6

CHAPTER 7

Nederlandse Samenvatting

Curriculum Vitae

Publications

Acknowledgements

NEDERLANDSE SAMENVATTING

De cellen in ons lichaam worden continu blootgesteld aan allerlei interne en externe bedreigingen die kunnen leiden tot beschadigingen aan ons DNA. Onze cellen zijn daarom sterk afhankelijk van een aantal essentiële beschermingsmechanismen die DNA-schade opsporen en repareren, en gezamenlijk de DNA-schaderespons worden genoemd. Het niet goed functioneren van de DNA-schaderespons kan leiden tot mutaties en instabiliteit van het genoom, hetgeen weer kan bijdragen tot de ontwikkeling van ziekten zoals kanker, neurodegeneratieve stoornissen en veroudering. Van de verschillende DNA-beschadigingen die kunnen optreden behoren DNA dubbelstrengsbreuken tot de meest genotoxische. Daarnaast zijn er verschillende DNA-beschadigingen die tijdens de verdubbeling van ons DNA, ook wel DNA-replicatie genoemd, de voortgang van de replicatievork kunnen blokkeren. Langdurige blokkade van replicatievorken kan leiden tot instabiliteit en het uiteenvallen van deze vorken, hetgeen de vorming van DNA dubbelstrengs breuken tot gevolg kan hebben. Inzicht in de mechanismen die betrokken zijn bij het repareren van DNA-breuken en het voorkomen van replicatie stress in de vorm van instabiliteit of het uiteenvallen van replicatievorken, kan leiden tot een beter begrip van de DNA-schaderespons en de talrijke aandoeningen die worden geassocieerd met defecten in deze respons. Alhoewel, studies aan eiwitten die betrokken zijn bij de DNA-schaderespons in detail zijn bestudeerd en tot vernieuwende inzichten hebben geleid, blijft het volledige repertoire van eiwitten dat bij deze respons betrokken is onbekend. Het onderzoek beschreven in dit proefschrift is gericht op de identificatie en het functioneel karakteriseren van nieuwe factoren betrokken bij de reparatie van DNA dubbelstrengsbreuken en het voorkomen van replicatie stress. De achtergrond van deze studie wordt uitgebreid beschreven in hoofdstuk 1.

In de afgelopen jaren hebben de matig gekarakteriseerde Zinc Finger (ZnF) eiwitten veel aandacht gekregen, omdat zij mogelijk een rol spelen in de DNA-schaderespons en het ontstaan van kanker. ZnF domeinen komen voor in tenminste 5% van alle humane eiwitten en deze zijn ingedeeld in verschillende ZnF domeinfamilies, waaronder de C2H2 domeinfamilie. ZnF eiwitten met een C2H2 domein staan vooral bekend om hun DNA-bindende vermogen. Verder onderzoek heeft aangetoond dat ZnF domeinen ook in staat zijn om RNA, lipiden en gemethyleerd DNA te binden. Daarnaast kunnen ZnF domeinen ook aan bepaalde chemische modificaties van eiwitten, ofwel post-translationale modificaties (PTMs) binden. Hierbij valt te denken aan modificaties zoals SUMO, ubiquitine, PAR en methyl. ZnF eiwitten kunnen ook meerdere typen ZnF-domeinen

bevatten, waardoor ze een veelzijdige specificiteit voor binding aan verschillende PTMs vertonen. ZnF domeinen komen voor in verschillende niet- eiwitsuperfamilies, zoals bijvoorbeeld in transcriptiefactoren, nucleaire hormoonreceptoren, integrase enzymen, E3 ubiquitine ligases, chromatine remodelers, tumor suppressors, RAS GTPases, membraantransport eiwitten en chaperones. Door de veelzijdige bindings-specificiteit van de ZnF domeinen in deze eiwitten, zijn zij in staat om verschillende processen in onze cellen te reguleren, waaronder het herstel van DNA dubbelstrengsbreuken. In hoofdstuk 2 beschrijf ik de functies van een aantal ZnF eiwitten in de twee belangrijkste DNA dubbelstrengsbreukherstel routes, namelijk niet-homologe eind-verbinding (NHEJ) en homologe recombinatie (HR). De focus ligt voornamelijk op de functionele relevantie van de ZnF domeinen in deze processen en hoe het disfunctioneren van ZnF eiwitten kan leiden tot de ontwikkeling van kanker. Ondanks dat we de functie van een aantal ZnF eiwitten inmiddels beter begrijpen, blijft het merendeel van deze eiwitten tot dusverre onvoldoende gekarakteriseerd. Verschillende studies hebben aangetoond dat ZnF eiwitten lokaliseren op DNA-schade die door bestraling van cellen met een laser is gemaakt. Deze methode induceert een breed spectrum aan DNA-schades, waaronder DNA dubbelstrengsbreuken, enkelstrengsbreuken en oxidatieve DNA-schade, hetgeen suggereert dat ZnF eiwitten aan verschillende DNA-beschadigingen kunnen binden. Bovendien scoren ZnF eiwitten hoog in CRISPR/Cas9 screens gericht op het identificeren van eiwitten die cellen beschermen tegen DNA-beschadigende stoffen, waaronder chemotherapeutica. Alhoewel de huidige kennis een rol voor ZnF eiwitten in DNA-reparatie en de ontwikkeling van kanker impliceert, zal verder onderzoek moeten uitwijzen hoe deze eiwitten bij deze processen betrokken zijn.

Ons DNA is met histon-eiwitten verpakt in chromatine, hetgeen ervoor zorgt dat het in de celkern past. Echter, dit maakt het efficiënt detecteren en repareren van DNA dubbelstrengs breuken uitdagend. Gelukkig zijn eiwitten in onze cellen in staat om de chromatine structuur rondom DNA dubbelstrengs breuken te ontvouwen, zodat het DNA toegankelijk wordt voor DNA-reparatie eiwitten. Een van deze enzymen die betrokken is bij dit proces is poly(ADP-ribosyl)-polymerase 1 (PARP1). Wanneer PARP1 bindt aan DNA-breuken wordt het geactiveerd en bevordert het lokaal de ontvouwing van chromatine door de vorming van poly(ADP-ribose)-ketens (PAR) op zichzelf en op histon-eiwitten. Dit proces leidt tevens het rekruteren van ATP-afhankelijke chromatine-remodelers, die het chromatine vervolgens nog verder ontvouwen. Een eerdere studie heeft aangetoond dat verschillende ZnF eiwitten, waaronder ZNF384, op DNA-schade

7

lokaliseren en dat PARP1 hiervoor noodzakelijk is. In hoofdstuk 3 beschrijf ik een belangrijke rol voor ZNF384 in het herstel van DNA dubbelstrengsbreuken via niet-homologe eind-verbinding (NHEJ), hetgeen in humane cellen het dominante mechanisme is voor het herstel van deze vorm van DNA-schade. Lokalisatie van ZNF384 op DNA dubbelstrengsbreuken vereist de PARP1-afhankelijke ontvouwing van beschadigd chromatine. Deze ontvouwing maak het beschadigd DNA toegankelijk, zodat ZNF384 er via het C2H2 ZnF domein aan kan binden. Door middel van massaspectrometrie, heb ik aangetoond dat ZNF384 interacteert met de twee belangrijke NHEJ-eiwitten, Ku70/Ku80. Deze interactie wordt bewerkstelligd door het N-terminale domein van ZNF384. Het C2H2-domein dat aan DNA bindt en het N-terminale domein dat met Ku70/Ku80 interacteert zijn allebei belangrijk voor de lokalisatie van Ku70/Ku80 op DNA dubbelstrengs breuken, als ook voor de lokalisatie van eiwitten die weer door Ku70/Ku80 worden aangetrokken, zoals APLF en XRCC4/LIG4. Deze resultaten suggereren dat ZNF384 fungeert als een 'Ku-adaptor' die DNA dubbelstrengs breuken en Ku70/Ku80 bindt om de opbouw van de cNHEJ reparatie machine te faciliteren. Hiermee heb ik een nieuwe en belangrijke rol voor ZNF384 in het herstel van DNA dubbelstrengs breuken en behoud van genoomstabiliteit aangetoond.

Het vermogen van cellen om te overleven en zich te vermenigvuldigen hangt af van het kopiëren van DNA via DNA-replicatie tijdens elke celcyclus. Echter, tijdens dit proces is het DNA kwetsbaar voor obstakels die de voortgang van de replicatievork kunnen blokkeren, en daarbij DNA-replicatie stress kunnen veroorzaken. De DNA-schaderespons speelt niet alleen een belangrijke rol in het herstellen van DNA dubbelstrengsbreuken, maar zorgt er ook voor dat DNA-replicatie stress voorkomen of verholpen wordt. Een veel voorkomend obstakel dat de voortgang van replicatievorken kan blokkeren is transcriptie, het proces waarbij genen worden afgelezen en er een kopie in de vorm van RNA wordt gemaakt. Er kunnen botsingen ontstaan tussen de eiwitmachines die transcriptie en DNA-replicatie reguleren. Die botsingen kunnen leiden tot transcriptie-replicatieconflicten (TRC's). TCRs komen meestal in een head-on oriëntatie voor, waarbij frontale botsingen tussen de replicatie- en transcriptie machines optreden. Daarnaast zijn R-loops een veel voorkomende bron van TRCs. R-loops zijn zeer stabiele DNA-RNA hybride structuren die als bijproduct van transcriptie gevormd worden. Door hun impact op de vorming van TRCs zijn R-loops een bron van genoominstabiliteit gebleken. In de afgelopen jaren is er veel onderzoek gedaan naar het samenspel van transcriptie, R-loops en de geassocieerde genoominstabiliteit. Dit onderzoek heeft geleid tot de ontdekking van verschillende

eiwitten die betrokken zijn bij het ontstaan van R-loops. Een volledig begrip van de context waarin R-loops ontstaan en TRC-geassocieerde genoominstabiliteit veroorzaken ontbreekt echter. In hoofdstuk 4 beschrijf ik de functie van het non-specific lethal (NSL) chromatine remodeling complex in het reguleren van R-loops. Mijn bevindingen suggereren dat KANSL3, een unieke subunit van het NSL complex, cellen beschermt tegen replicatie stress door de herstart van vastgelopen replicatievorken te bevorderen, die zijn ontstaan nadat R-loops TRCs hebben veroorzaakt. KANSL3 onderdrukt de vorming van R-loops en voorkomt zo botsingen tussen de replicatie- en transcriptiemachines. Toekomstig onderzoek is nodig om inzicht te krijgen in hoe exact KANSL3 de vorming van R-loops en daarmee de stabiliteit van de replicatievork beïnvloedt.

Naast chromatine remodeling complexen en transcriptie factoren spelen verschillende eiwitten die RNA binden een belangrijke rol in het afbreken van R-loops om replicatie stress te voorkomen. In hoofdstuk 5 beschrijf ik een belangrijke rol voor het RNA afbrekende enzym ERI1 in dit proces. ERI1 is een evolutionair geconserveerde 3'-5' exonuclease die een rol speelt in verschillende RNA-processing routes, inclusief de routes die betrokken zijn bij het afbreken van ribosomaal RNA en de degradatie van histon mRNAs. Ik heb gevonden dat verlies van ERI1 leidt tot een toename van R-loops en de phosphorylatie van RPA (S4/S8), hetgeen indicatief is voor een verhoging van het aantal gestalde/uiteengevallen vorken. Ook heb ik aangetoond dat een inactiverende mutatie in het 3'-5' exonuclease domein van ERI1 leidt tot een verhoogd phosphorylatie-niveau vanRPA (S4/S8). Dit geeft aan dat de exonuclease activiteit van ERI1 belangrijk is bij het voorkomen van gestalde/ uiteengevallen vorken. Tenslotte beschrijf ik dat ERI1 bindt aan gestalde replicatievorken, hetgeen suggereert dat ERI1 een directe rol speelt in het behouden van vorkstabiliteit. Toekomstig onderzoek zal moeten uitwijzen hoe ERI1 een rol speelt in het beschermen van de integriteit van replicatievorken en of dit afhangt van ERI1's rol in het afbreken van R-loops. In het laatste hoofdstuk van dit proefschrift, namelijk hoofdstuk 6, beschrijf ik hoe mijn bevindingen bijdragen aan ons begrip over genoomstabiliteit, richt ik mij op onopgeloste kwesties, en maak ik aanbevelingen voor vervolgstudies.

7

CURRICULUM VITAE

Jenny Kaur Singh werd geboren op 14 september 1991 te Amsterdam. In 2008 behaalde zij haar HAVO diploma aan het Calandlyceum in Amsterdam. Vervolgens startte zij met de opleiding Biologie en Medisch laboratorium onderzoek aan Hogeschool INholland in Amsterdam. Vanwege haar interesse in moleculaire biologie en genetica, en haar leergierigheid om deze vakgebieden in meer detail te begrijpen besloot zij haar opleiding in 2012 te vervolgen met een duale masteropleiding. Deze opleiding bestond uit een samenwerking van de masteropleiding 'Biomolecular Science' aan de Vrije Universiteit in Amsterdam en 'Natural Sciences – Interdisciplinary' at the State University of New York at Buffalo in de Verenigde Staten. In 2014 rondde zij deze opleidingen succesvol af.

In het kader van haar bachelor en master opleiding deed Jenny verschillende stages met een verscheidenheid aan model-organismen om brede ervaring op te doen in moleculaire biologie, genetica en biochemie. In 2011 startte zij een stage onder leiding van Prof. Dr. Niek de Vries, Dr. Paul Klarenbeek en Dr. Marieke Doorenspleet in het Academisch Medisch Centrum (AMC) in Amsterdam, waarin zij werkte aan het karakteriseren van de T-cel en B-cel receptor repertoire in Reumatoïde Artritis patiënten samples en gezonde samples. Na voltooiing van deze stage is Jenny naar London, Verenigd Koninkrijk verhuisd met een Erasmus beurs voor haar tweede bachelor-stage onder leiding van Prof. Dr. Rob de Bruin. Hier heeft zij de rol van het Rpd3 histon de-acetylase complex (HDAC) gekarakteriseerd in gist. Na deze twee stages mocht Jenny haar bachelor diploma in ontvangst nemen. Voor haar masteropleiding heeft zij haar eerste stage volbracht in het lab van Dr. Dirk Bald aan de Vrije Universiteit in Amsterdam, waar zij onderzoek deed naar het testen van medicijnen die het energie metabolisme in de gezonde vorm van Tuberculose kunnen verstoren, om deze vervolgens te kunnen gebruiken tegen Tuberculose. Na het afronden van deze stage heeft Jenny drie maanden een literatuur scriptie geschreven aan de 'Luxembourg Center for Systems Biomedicine' onder leiding van Dr. Jochen Schneider over de rol van integrine eiwitten in macrofagen. Voor de laatste stage van haar master bezocht zij het Roswell Park Cancer Institute in Buffalo, Verenigde Staten. Aldaar werkte zij in het lab van Dr. Xinjiang Wang aan de regulatie van het tumor suppressie eiwit p53 in respons tot DNA schade.

In 2014 begon zij haar promotieonderzoek in het lab van Prof. Dr. Haico van Attikum. Hier heeft zij verschillende nieuwe factoren gevonden die een rol spelen in DNA herstel en DNA replicatie stress. De resultaten van dit onderzoek zijn beschreven in dit proefschrift.

In 2021 vervolgde Jenny haar onderzoek naar het DNA reparatie eiwit BRCA2 in het lab van Dr. Aura Carreira aan het Institut Curie in Parijs. Eind dit jaar zal zij beginnen als Postdoctoraal onderzoeker in het lab van Dr. Annabel Quinet aan het CEA in Parijs. Hier zal zij onderzoeken hoe het PRIMPOL eiwit functioneert tijdens DNA-replicatie stress in de context van chromatine.

PUBLICATIONS

2021   **Zinc finger protein ZNF384 is an adaptor of Ku to DNA during classical non-homologous end-joining**
Jenny Kaur Singh, Rebecca Smith, Magdalena B. Rother, Anton J.L. de Groot, Wouter W. Wiegant, Kees Vreeken, Ostiane D'Augustin, Robbert Q. Kim, Haibin Qian, Przemek M. Krawczyk, Román González-Prieto, Alfred C.O. Vertegaal, Meindert Lamers, Sébastien Huet and Haico van Attikum Zinc finger Protein ZNF384 Is an Adaptor of Ku to DNA during Classical Non-homologous End-Joining.
Nature Communications. 12 (12), 1–21

2020   **DNA double-strand break repair: Putting zinc fingers on the sore spot**
Jenny Kaur Singh and Haico van Attikum DNA double-strand break repair: putting zinc fingers on the sore spot.
Seminars in Cell and Developmental Biology. **113**, 65–74

7

ACKNOWLEDGEMENTS

The PhD trajectory is by no doubt the most challenging and overwhelming, but at the same time a very creative and fruitful time of your life. Looking back at this non-linear trajectory, I can now say that these years taught me self-growth, self-awareness and most importantly made me into an independent researcher. This would have never been possible without the help of my family, friends and peers.

I would like to start by thanking my promotor/mentor Haico. Thank you for the scientific and career-related advices, and showing me how to successfully prepare my work for publication.

Secondly, I would like to thank Silvère for giving constructive feedback during the important and sometimes nerve wrecking PhD-evaluation moments.

The 'van Attikum' group faced constant change in lab members over the past years, I would like to thank each and every one of them for their constant support. Leonie, my PhD-buddy, thank you for being by my side as a friend from day one and later as an 'office-roomie'. Wouter, I feel super lucky to have worked with you. Thank you for teaching me everything you know and appreciating my taste in the radio channel. I would like to thank the early members Albert, Madelon, Angela, Peter, Anton, Kees, Martijn, Thomas and Suming for being there when I started my PhD. I would like to thank Chantal and Amandine for the great times. The current members of the lab, Rick my friend, thank you for all the fun times and good conversations. Bert, I will never forget your positive attitude and your creative scientific input which has been of tremendous help. Sylvie, I feel so grateful to have worked with you. Seeing you becoming a group-leader has been very inspirational, and I simply very much look up to you. A successful PhD would have never been possible without a good work/life balance. I would like to thank Fenna, Nandi, Sophie, Jolanthe, Veronica, Marta, Venda, Anne, Romy and Nuria for keeping me in check. Veronica and Romy, I am so blessed to have you as my best friends. Marta, your determination to drag me to the gym kept me sane. During my PhD I had the opportunity to supervise a very talented Master (and now PhD) student Marnix. It was great having you around and I am very happy you stayed in the department for your PhD.

Jiggling between the lab and the office has been great. But not without my different office mates Suming, Bharath, Pierre, Robin, Juliette, Leonie, Anne and Venda. Besides getting a couch and growing vegetables, I have learned so much from our scientific discussions.

The perks of being part of a huge department is the opportunity to make many friends outside the lab and research-line. Maja, Darina, Haoyu and Mara thank you for the occasional chats we had on the third floor. Being in the lab-day out committee with Remko, Chiara, Yvonne and Saeed turned into an awesome travel-group and I can't wait to plan our next trip. Writing this thesis in combination with a Postdoc has been extremely challenging. I could have not done this without my colleagues/friends at Institut Curie Isaac, Anna, Rady, Yasmina, Charlotte, Kamila and Shrena, and my friends from Paris Umberto, Marta, Audrey, Theresa, Frederik and Camille.

My peer and paranymph from day 1, Remko, I could not thank you enough for all the mental support and your patience while I practiced all my presentations in front of you to such an extent that you started hearing 'foci counting' in your dreams. My other paranymph Magda, you are an amazing scientist and I could not have done this without you. I am very grateful for all your encouraging words and kindness over all these years.

I feel blessed to still have my primary-, high-school and university friends in my life for their constant support, kindness and life lessons. I would like to thank all these amazing people Sharonnita, Vita, Sinony, Geniva, Guus, Naphur, Kadene and Quirine.

Finally and most importantly, I want to thank my family for their unconditional love and support. They moved away from India at a young age to build a new life in the Netherlands as a first generation immigrant family, with the thought of giving their children a better life. Mum, Thank you for always lifting me up when things were hard and having the patience with me. Dad, I admire you hard work and I think I got this from you. My brother Simran, I know you always told me that you would retire before I would get my PhD, but in the end I proved you wrong. Thank you for your (bad) jokes, they kept me smiling throughout the PhD.

7