Cellular and Biochemical Characterisation of ERCC6L2

Lucy Armstrong

Lincoln College

Sir William Dunn School of Pathology

University of Oxford

Thesis submitted for the degree of DPhil in Pathology
Hilary term 2021

Abstract

Background and aims: ERCC6L2 is a member of the Snf2 ATPase superfamily that has been linked to the DNA damage response (DDR). Past work from our group demonstrated that *ERCC6L2*^{KO} cells are sensitive to the double-strand break (DSB) inducers etoposide and phleomycin. However, the precise nature of the role that ERCC6L2 plays in the DDR and how it functions to preserve genome stability, remain unclear. Several mutations have been identified in *ERCC6L2* which result in a distinct subtype of inherited bone marrow failure (IBMF). The aim of my project has been to improve understanding of the physiological function of ERCC6L2 using a combination of cellular and biochemical approaches. **Results:** *ERCC6L2*^{KO} cells accumulate higher levels of DNA damage, compared to wildtype cells, under treatment with both etoposide and phleomycin. Cell cycle analysis revealed that this causes a higher proportion of *ERCC6L2*^{KO} cells to accumulate in G2 phase. Assessment of the ability of *ERCC6L2*^{KO} cells to recover from acute phleomycin treatment showed that cells lacking ERCC6L2 have significantly higher levels of DNA damage after a short release but return to levels comparable to untreated cells after a longer recovery period. Following release from phleomycin treatment, *ERCC6L2*^{KO} cells progressively accumulate higher levels of single-stranded DNA (ssDNA) than wildtype cells. We hypothesised that this is due to a DNA end processing defect and that ERCC6L2 functions as an anti-resection factor. Loss or inactivation of anti-resection factors has been shown to suppress PARP inhibitor sensitivity in *BRCA1*-deficient cells. However, I found that loss of ERCC6L2 does not mediate this effect. Furthermore, I found that *ERCC6L2*^{KO} cells are mildly sensitive to the replication inhibitor hydroxyurea (HU). Under HU-induced replication stress, *ERCC6L2*^{KO} cells show significantly lower levels of ssDNA than wildtype cells and have increased chromatin bound levels of the mismatch repair factors MSH2/6. Finally, I show that a fragment of ERCC6L2 containing the uncharacterised C-terminal VIGSSK domain, which localises to sites of DNA damage, can bind DNA. A BioID screen with the VIGSSK fragment identified several exciting interaction candidates which may indicate that ERCC6L2 is important for maintaining genetic stability at specific regions of the genome. **Conclusions:** Loss of ERCC6L2 leads to an increase in DNA end resection following treatment with DSB inducing drugs and therefore likely functions in an anti-resection capacity and in effect, as a non-homologous end joining promoting factor. ERCC6L2 may be involved in DNA replication though it is unclear whether this role is distinct from its function in DSB repair. Collectively, these findings offer useful insights into the physiological function of ERCC6L2 in a cellular and biochemical context and provide a strong foundation for future studies. It is hoped that these results will help to elucidate the pathological basis of *ERCC6L2*-assoicated IBMF.

Acknowledgements

I would like to firstly thank my supervisor Dragana Ahel for providing me the opportunity to work in her lab and for her continued support and guidance throughout the course of this project. I'd also like to thank the EPA trust for funding me throughout my DPhil.

The Dunn School and more specifically OMPI1 has been a truly inspiring place to complete my DPhil. I cannot express how grateful I am for the kindness, patience and encouragement shown by all those who have ever taken the time to help and advise me. In Dragana's lab and beyond, I have been so fortunate to work alongside some of the most hard-working, bright and caring people - the type of people who have made me want to be a better person. So, Chris, Marek, Adam, Lucrezia, Emily, Xiaomeng, Sheng, Io Nam, Kang and Valentina, I have learnt so much from all of you and you have helped me immensely throughout this process. I really couldn't have done it without you.

Writing this thesis has coincided with a particularly challenging time in my life. While it has made an already difficult task more difficult, it has made me really appreciate those in my life who are truly there for me. I have been overwhelmed by the unwavering support and kindness of certain friends over the past few months who I really owe this thesis to. Chris, Adam, Lucrezia, Emily, Marek and Hannah, thank you for being there for me in the hardest moments.

Finally, Mum, Marc and Hugh, you have always been there, and it means everything. Thank you all so much.

Contents

Table of Figures

List of abbreviations

3'PG	3' phosphoglycolated
53BP1	p53 binding protein 1
AID	Activation induced cytosine deaminase
ALT	Alternative lengthening of telomeres
Alt-EJ	Alternative-end joining
AML	Acute myeloid leukaemia
ATM	Ataxia telangiectasia mutated
ATP	Adenosine triphosphate
ATR	Ataxia telangiectasia and Rad3-related
ATRIP	ATR interacting protein
BACs	Bacterial artificial chromosome
BASC	BRCA1-associated surveillance complex
BER	Base excision repair
BIR	Break induced replication
BLM	Bloom's syndrome helicase
BRCA1	Breast cancer type 1 susceptibility protein
BRCA2	Breast cancer type 2 susceptibility protein
BRCT	BRCA1 C-terminus
BrdU	Bromodeoxyuridine
BTR	BLM-TOPOIIIA-RMI1/2
CATD	CENP-A targeting domain
CCAN	Constitutive centromere associated network
CDK	Cyclin dependent kinase
CENP	Centromere protein
CFSs	Common fragile sites
CHK1	Checkpoint kinase 1
CHK2	Checkpoint kinase 2
c-NHEJ	Classical-non homologous end joining
CMG	Cdc45-MCM2-7-GINS
CS	Cockayne Syndrome
CSB	Cockayne Syndrome protein B
CSR	Class switch recombination
CST	CTC1-STN1-TEN1
CtIP	CtBP-interacting protein
D-loop	Displacement loop
DDR	DNA damage response
DDT	DNA damage tolerance
DNA-PKcs	DNA-protein kinase catalytic subunit
DSBs	Double-strand breaks
dsDNA	Double-stranded DNA
ERCC6L2	Excision repair cross complementation group 6 like 2
ETAA1	Ewing's tumour associated antigen
FA	Fanconi anaemia
FBH1	F-box helicase 1
GG-	Global genome
HC	Heterochromatin
HDR	Homology directed repair

HJ	Holliday junction
HJURP	Holliday junction recognition protein
HLTF	Helicase-like transcription factor
HOR	Higher order repeat
HP1	Heterochromatin protein 1
HR	Homologous recombination
HU	Hydroxyurea
IBMF	Inherited bone marrow failure
ICL	Inter-strand crosslink
Ig	Immunoglobulin
IR	Ionising radiation
MCM2-7	Minichromsome maintenance proteins 2-7
MDC1	Mediator of damage checkpoint 1
MDS	Myelodysplastic syndrome
MiDAS	Mitotic DNA synthesis
MMC	Mitomycin C
MMR	Mismatch repair
MRN	MRE11-RAD50-NBS1
NER	Nucleotide excision repair
NONO	non-POU domain-containing octamer-binding protein
PALB2	Partner and localiser of BRCA2
PARP1	Poly-ADP ribose polymerase 1
PAXX	Paralog of XRCC4 and XLF
PCNA	Proliferating cell nuclear antigen
PI3K	Phosphoinositide 3-kinase
PICH	PLK1-interacting checkpoint helicase
PTIP	PAXX transcription activation domain interacting protein
RIF1	RAP interacting factor 1
RMI1/2	RecQ mediated genome instability 1/2
RNF168	Ring finger protein 168
RNF8	Ring finger protein 8
RPA	Replication protein A
RTEL1	Regulator of telomere elongation helicase 1
RVF	Reversed fork
SDSA	Synthesis dependent strand annealing
seDSBs	Single-ended DSBs
SFPQ	Splicing factor proline- and glutamine-rich
SMC2-4	Structural maintenance of chromosomes 2-4
Snf2	Sucrose non-fermentable 2
SSA	Single-strand annealing
SSBs	Single-strand breaks
ssDNA	Single-stranded DNA
TC-	Transcription coupled
TLS	Translesion synthesis
TOPBP1	Topoisomerase II binding protein 1
TOPO I	Topoisomerase I
TOPO II	Topoisomerase II
TOPO IIIA	Topoisomerase IIIA
TPT	Topotecan

TRF2	Telomere repeat binding factor 2
TS	Template switching
UFBs	Ultrafine bridges
UV	Ultraviolet
VDJ	Variability diversity joining
WRN	Werner's syndrome helicase
XLF	XRCC4 like factor
XP	Xeroderma pigmentosum
XRCC1	X-ray repair cross complementing protein1
XRCC4	X-ray repair cross complementing protein 4
γH2AX	Phosphorylation of s139 on histone H2AX
A	Amp
a.u.	Arbitrary units
AUC	Area under curve
BrdU	Bromodeoxyuridine
BSA	Bovine serum albumin
CVs	Column volumes
DMSO	Dimethyl sulfoxide
dNTPs	Deoxyribonucleotide triphosphates
DTT	Dithiothreitol
EMSA	Electrophoretic mobility shift assay
FBS	Foetal bovine serum
GFP	Green fluorescent protein
GST	Glutathione sepharose transferase
IF	Immunofluorescence
IP	Immunoprecipitation
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	Kilodalton
LB	Luria broth
LDS	Lithium dodecyl sulphate
MBP	Maltose binding protein
MW	Molecular weight
Ni-NTA	Nickel- nitrilotriacetic acid
NLS	Nuclear localisation sequence
^{o}C	Degrees Celsius
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PIs	Protease inhibitors
PNK	Polynucleotide kinase
psi	Pound per square inch
rpm	Revolutions per minute
RT	Room temperature
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEM	Standard error of the mean
sgRNA	Single-guide RNA
siRNA	Short interfering RNA
V	Voltage
x g	times gravity
YFP	Yellow fluorescent protein

1. Introduction

1.1. The DNA damage response

The faithful transmission and preservation of genetic material is of utmost importance to the survival of every organism. Yet, genome integrity is constantly under threat from endogenous and environmental insults. DNA lesions arising from these insults can lead to deleterious mutations and chromosomal aberrations (1). They can also affect the fidelity and completion of important nuclear processes such as DNA replication and transcription. In order to tolerate and combat DNA damage, cells have evolved an impressive array of specialised pathways, collectively termed the DNA damage response (DDR). The DDR is a complex network of sensors, mediators and repair factors that work in a highly coordinated fashion to resolve DNA lesions in a timely manner and elicit the appropriate cellular response.

On a cellular level, failure to resolve DNA damage can result in senescence or even cell death. Mutations in DNA repair genes clinically manifest in a diverse range of human diseases which are often associated with complex development and neurological abnormalities, premature ageing and cancer predisposition (2–4). Improving our understanding of defects in the DDR has led to the development of novel chemotherapeutic strategies. For example, tumours harbouring specific DDR mutations render them unable to repair certain types of DNA allowing for selective killing of cancer cells (5). Thus, investigating the role of DDR factors and the repair pathways in which they function will help improve understanding of how this complex and interconnected system operates to maintain genome stability, and how it can be manipulated in the treatment of human diseases such as cancer.

1.2. Origins and repair of DNA damage

DNA lesions come in a wide variety of forms such as DNA base modifications, single and double-strand breaks as well as intra- and inter-strand crosslinks. They can also arise from numerous sources, some which are external or environmental, and those that are naturally occurring. Consequently, cells are armed with a range of DNA repair pathways tailored to resolving specific kinds of DNA damage.

Modification of DNA bases

The intrinsically unstable nature of DNA makes it susceptible to chemical attack from endogenous molecules. For example, the presence of by-products from normal cellular metabolism, like reactive oxygen species and S-adenosyl methionine, can result in base oxidation and alkylation, respectively (2,6). DNA can also undergo spontaneous hydrolysis leading to base deamination or depurination (7). These minor base modifications do not significantly affect the structural conformation of the DNA helix and are typically resolved by the base excision repair (BER) pathway (8). Briefly, this involves damage recognition by a lesion-specific DNA glycosylase which excises the damaged base (8). The abasic site is restored through gap-filling by DNA polymerase β and ligation by DNA ligase III and XRCC1.

More extensive base modifications can be produced by external genotoxic stress. For example, exposure to ultraviolet (UV) light catalyses the formation of cyclobutane pyrimidine dimers and 6-4 photoproducts (2). Additionally, mutagenic chemicals such as aflatoxins and cigarette smoke can also promote formation of bulky base adducts. The presence of such modifications can disrupt the helical structure of DNA which triggers repair by the nucleotide excision repair (NER) pathway (9). NER operates via two sub-pathways, global genome (GG) or transcription-coupled (TC). GG-NER

involves continual scanning of the genome for perturbations in the helix structure, whereas TC-NER is indirectly triggered when such lesions occur in actively transcribed regions and impede progression of RNA polymerase II (9). Ultimately, both pathways result in the removal of a short patch of nucleotides, encompassing the lesion, following dual incision by structure-specific endonucleases at positions flanking the lesion (9). Defects in the NER pathway give rise to a clinically diverse set of human diseases including xeroderma pigmentosum (XP), trichothiodystrophy (TTD) and Cockayne syndrome (CS) which are broadly associated with developmental delay, neurological abnormalities and, in the case of XP, cancer predisposition (10).

DNA breaks

DNA breaks can occur in either one or both strands of the DNA duplex. Single-strand breaks (SSBs) and double-strand breaks (DSBs) can form pathologically after exposure to ionising radiation (IR) (2). They can also be generated as intermediates during certain DNA transactions such as topoisomerase mediated relaxation and decatenation. Topoisomerase I (TOPO I) and topoisomerase II (TOPO II) catalyse nicks in one or both strands of the DNA, respectively, during transcription and DNA replication to relieve torsional stress (11). During this process, the topoisomerase becomes covalently attached to the DNA end creating a cleavage complex. Normally, these breaks are swiftly ligated by the respective topoisomerase. However, failure to complete the ligation step leaves TOPO I or TOPO II – linked breaks which can become exposed if encountered by DNA replication or transcription machineries (12,13).

DSBs are primarily repaired by one of two canonical repair pathways: classical non-homologous end joining (c-NHEJ) or homologous recombination (HR). c-NHEJ essentially involves re-joining the broken DNA ends, therefore there is a risk of

mutation at the break site (14). In contrast to c-NHEJ, HR is a fully restorative repair pathway as it uses the sister chromatid as a template for repair (15). Cell cycle stage is therefore a crucial determinant of DSB repair pathway choice since HR can only be used for repair following DNA replication. HR is therefore thought to only be operative in the S and G2 phases of the cell cycle. Single-strand annealing (SSA) and alternative end joining (alt-EJ) are two additional, non-canonical methods of DSB repair. However, these are highly mutagenic pathways and likely only serve as backup repair mechanisms in situations where c-NHEJ or HR cannot be used (16–18).

Repair of SSBs involves the multifunctional enzyme poly-ADP ribose polymerase 1 (PARP1) (19). PARP1 binds to DNA and modifies an array of substrates, including itself, by synthesising poly-ADP ribose chains from cellular NAD^+ and attaching them to target molecules (19). PARP1 is recruited rapidly to SSB sites where it tethers the DNA ends and undergoes auto-modification. This leads to recruitment of downstream repair factors including XRCC1 and DNA ligase III, which together restore the break (20). PARP inhibitors are commonly used in the treatment of *BRCA1/2*-deficient cancers which are unable to repair DSBs by HR (21). Originally, it was thought that the mechanism underlying PARP inhibitors is linked to the persistence of unrepaired SSBs and the inevitability that these lesions will be processed into DSBs that cannot be repaired by the cancer cells. However, there is evidence to suggest that many of the clinically used PARP inhibitors physically trap PARP1 on DNA which, similar to trapped topoisomerase complexes, can interfere with replication fork progression and rely on HR mechanisms for repair (21,22).

Intra- and inter-strand crosslinks

Many chemotherapeutic agents work by generating DNA damage to exploit DDR defects found in certain tumours. For example, cisplatin and mitomycin C (MMC), are used to treat a wide variety of cancers, and induce intra- and inter- strand crosslinks (23). Inter-strand crosslinks (ICLs) are a highly deleterious form of DNA damage which occur when two separate DNA strands become covalently linked and prevent strand separation. These lesions can be particularly toxic, since their presence can impede progression of transcription and replication machineries (24).

ICLs that occur in S phase are resolved through a complex process involving multiple modes of repair. Central to this are the Fanconi anaemia (FA) proteins. The FA proteins are required for sensing ICLs and for ubiquitin mediated recruitment of downstream repair factors (24). Replication-associated ICLs require HR factors to stabilise replication forks to facilitate endonucleolytic ‘unhooking’ of the ICL and for re-establishing DNA replication. Mutations in the FA proteins lead to the heterogeneous, genetic disease Fanconi anaemia, which is associated with chromosome fragility, bone marrow failure and cancer predisposition (25).

Despite the range of DNA lesions and the diverse range of sources from which they can arise, cells have become well equipped to effectively combat such assaults. In certain instances, repair pathways must cooperate to resolve more complex lesions and many proteins are involved in the repair of more than one type of DNA damage. Regulation of the DDR is also strongly influenced by cell cycle stage since this can determine which pathways are available for repair.

1.3. DNA damage signalling

To combat DNA damage effectively, the DDR must closely coordinate with pathways that control cell cycle progression and cell fate. This ensures that DNA damage is either resolved in a timely manner or, if it proves to be too severe, the appropriate cellular response such as senescence or apoptosis is induced.

The cell cycle is comprised of a series of stages that cells must navigate before dividing into two daughter cells. During G1 phase, cells can either prepare for DNA replication or enter the quiescent G0 phase in which they stop dividing on a permanent or indefinite basis (26). Proliferating cells transition into S phase in which they undergo DNA replication. After S phase, cells move into G2 during which they can resolve DNA replication associated issues before committing to mitosis and cell division (27). Cell cycle checkpoints are in place at each transition point which must be satisfied before progressing to next stage (26). Since DNA damage can arise throughout the cell cycle, these checkpoints are linked to DNA damage signalling to ensure DNA lesions do not persist into later stages of the cell cycle. This is especially important for S phase and mitosis, since unrepaired damage can have catastrophic consequences on genome stability by hindering completion of DNA replication and causing errors during chromosome segregation (27).

DNA damage signalling is orchestrated by three apical protein kinases belonging to the phosphoinositide 3-kinase (PI3K) related kinase family (28). Of these three, ataxia-telangiectasia mutated (ATM) and ataxia-telangiectasia and Rad3-related (ATR) kinases are the two major checkpoint proteins which are critical for integrating damage signalling with cell cycle progression, as well as chromatin remodelling and moderation of other nuclear events.

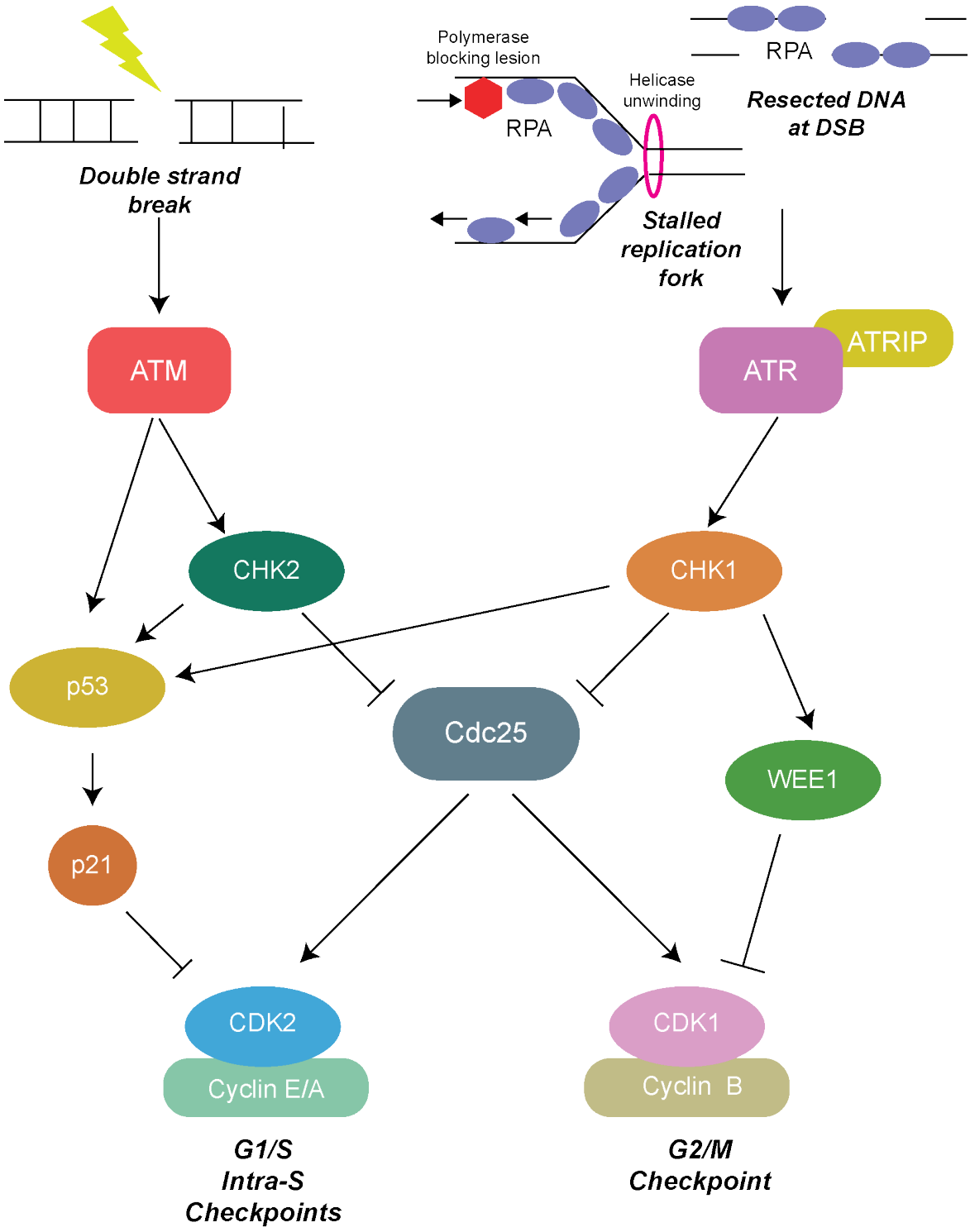


Figure 1-1: DNA damage signal transduction by ATM and ATR kinases

ATM and ATR kinases are activated in response to specific types of DNA damage and activate their downstream effector checkpoint kinases. In turn, these proteins are responsible for signal transduction via various axes to elicit cell cycle arrest.

ATM kinase

ATM kinase is responsible for coordinating DNA damage signalling in response to DSB induction (29). While mice lacking ATM are viable, autosomal recessive mutations in *ATM* result in the neurodegenerative and immune disorder ataxia-telangiectasia, which is unsurprisingly associated with a failure to sense and effectively respond to DSBs (30).

After initial damage sensing, ATM is rapidly recruited to sites of DSBs via its interaction with NBS1, a component of the MRE11-NBS1-RAD50 (MRN) complex which is one of the earliest responders to DSBs (31). Following its targeting to DSBs, ATM is catalytically activated, undergoes auto-modification and phosphorylates an extensive range of downstream substrates.

Histone variant H2AX is phosphorylated at serine 139 (γH2AX), primarily by ATM, in response to DSBs (31). γH2AX is recognised by mediator of damage checkpoint 1 (MDC1), another ATM substrate, that when phosphorylated triggers a ubiquitin signalling cascade which functions to suppress DNA end resection in G1 phase and promote repair through c-NHEJ (32–34). In addition to promoting DNA repair directly, ATM signalling also functions to regulate cell cycle progression by activating cell cycle checkpoints to ensure damage does not persist into later stages of the cell cycle or to daughter cells (Figure 1-1).

If a cell sustains damage in G1 phase it is imperative that it is resolved before the cell commits to entering S phase. In response to damage incurred in G1 phase, ATM activates its effector checkpoint kinase CHK2 by phosphorylating it on threonine 68 (35). CHK2 then modifies the tumour suppressor and transcription factor p53 (36). ATM also phosphorylates p53 directly and activates deubiquitylase enzymes to prevent p53

degradation (37,38). These events result in p53 stabilisation, which promotes transcription of various genes involved in the DDR and cell cycle regulation. Notably, it leads to increased levels of p21, a cyclin-dependent kinase (CDK) inhibitor which restricts S phase entry (39). However, if the DNA damage is too severe, p53 can also promote expression of apoptotic genes which result in programmed cell death (40). Checkpoint activation during S phase and at the G2/M transition can also occur via the ATM-CHK2 axis. In this instance, CHK2 phosphorylates the phosphatase Cdc25A, which prevents it from removing the inhibitory phosphate from CDK2 and CDK1 thus halting cell cycle progression (36,41).

ATR kinase

During S phase, it is inevitable that a cell will face obstacles that challenge the progression of the replication fork and threaten genome stability. Such barriers can come in many forms, including nicks and lesions on the template DNA, secondary DNA structures and collisions with the transcriptional machinery (42,43). This can cause replication fork slowing or stalling and ultimately, in some instances, fork collapse where replisome components dissociate from the DNA (44). Such events are broadly described as replication stress and will be discussed in more depth in Section 1.5.

Checkpoint activation in response to replication stress is governed by ATR kinase which acts to stabilise stalled replication forks and prevent fork breakage (45). When a replication fork stalls, the DNA polymerase becomes uncoupled from the replicative helicase, which continues to unwind the DNA ahead of the lesion in a process known as replication uncoupling (Figure 1-2) (46). This exposes single-stranded DNA (ssDNA) which is rapidly engaged and protected by the ssDNA binding protein replication protein A (RPA). RPA coated ssDNA binds ATR interacting protein (ATRIP)

which recruits ATR to the stalled fork and triggers ATR autophosphorylation and activation (47).

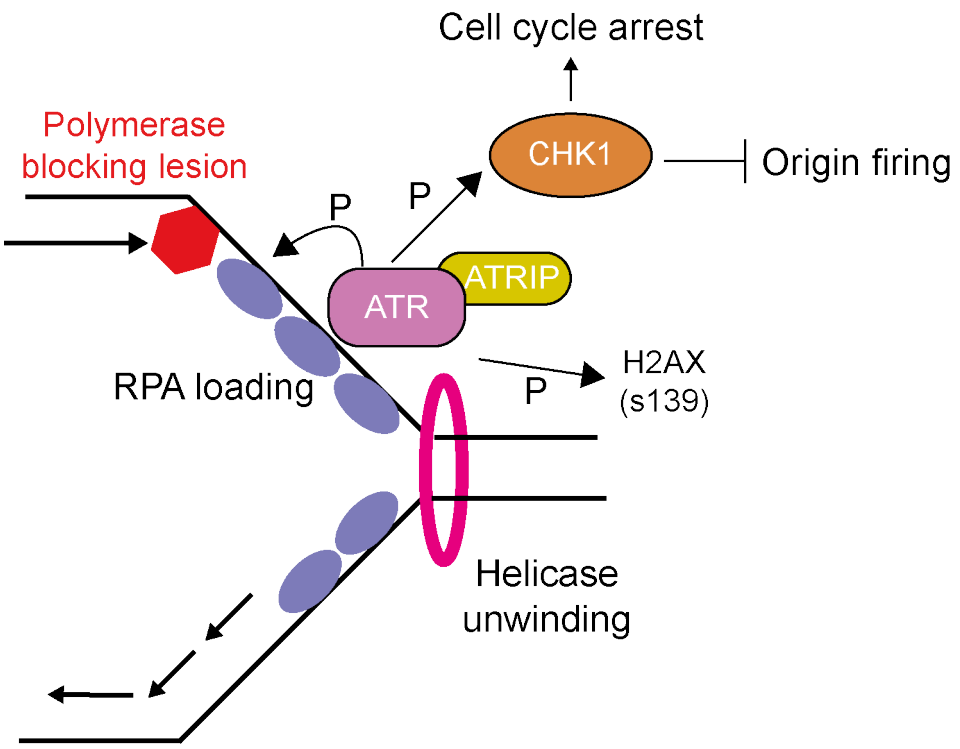


Figure 1-2: Replication uncoupling

Following replication fork stalling at a polymerase blocking lesion, the helicase continues to unwind DNA ahead of the lesion, leading to exposure of ssDNA and RPA loading. This results in ATR activation and its phosphorylation of its downstream targets, which promotes fork stabilisation and recovery.

ATR activity is further stimulated by topoisomerase II binding protein 1 (TOPBP1) which relies on Rad17-RFC and Rad9-Rad1-Hus1 for recruitment (46). More recently, another factor Ewing's tumour associated antigen 1 (ETAA1), which is recruited directly to RPA-ssDNA, was shown to be required for full ATR activity (48). Once fully activated, ATR phosphorylates numerous substrates including RPA, H2AX and its effector kinase CHK1 on serine 317 and 345 (49). Like CHK2, CDC25A is a major CHK1 target whose phosphorylation leads to cell cycle arrest and suppression of global origin firing (Figure 1-1 and 1-2) (46,50).

It should be noted that the cellular role of ATR kinase is not limited to the response to replication stress. RPA-ssDNA is also generated in the context of other types of DNA damage. Thus, ATR kinase also functions in the response to UV-induced damage and resected DSBs (51). Hypomorphic *ATR* mutations have been linked to Seckel syndrome which is a congenital disease associated with microcephaly and dwarfism (52). Unlike ATM, loss of ATR is incompatible with life indicating that it is important for DNA replication in unperturbed conditions (53).

DNA-PKcs

In contrast to ATM and ATR, the final DDR PI3K-like kinase, DNA-protein kinase catalytic subunit (DNA-PKcs), does not play a major role in checkpoint activation and has a significantly lower number of identified substrates (28). Nevertheless, DNA-PKcs plays a crucial role in the DSB repair pathway c-NHEJ. Following DSB formation, exposed DNA ends are rapidly bound by the KU heterodimer consisting of the KU70 and KU80 subunits. DNA-PKcs is then recruited to DSBs through its interaction with KU80 resulting in the formation of the DNA-PK holoenzyme (54). This activates DNA-PKcs leading to a critical autophosphorylation event that facilitates synapsis of the DNA ends (55). DNA-PKcs activation also recruits and stimulates the endonuclease Artemis which processes DNA ends during ‘resection-dependent’ c-NHEJ (56,57). Despite its primary function in DSB repair, there is evidence to suggest that DNA-PKcs may serve as an important backup to S phase checkpoint activation. Despite ATR inhibition, CHK1 phosphorylation is still observed in cells that have been subjected to replication stress, which was shown to be dependent on DNA-PKcs activity (58).

Although the PI3K-like kinases are activated by different substrates, the circumstances under which they function are not exclusive. Instead, they work cooperatively, which is

reflected in the extensive crosstalk between them, to ultimately ensure that genome stability is maintained throughout the cell cycle.

1.4. Double-strand break repair

Regardless of their origin, whether pathological or through deliberate induction during antibody diversification, proper resolution of DSBs is fundamental to cell survival. Repair of DSBs is predominantly mediated by either one of two major pathways: c-NHEJ or HR. As aforementioned, cell cycle stage is key in deciding which of these processes can be used for repair. Many of the pro-c-NHEJ and pro-HR factors, which form the molecular basis of pathway choice, directly antagonise each other at DSB sites. Disruption to the regulation of either c-NHEJ or HR has disastrous consequences for cells, since it may result in cell death or promote the use of mutagenic modes of repair (59–62). If an inappropriate repair pathway is used, it could lead to aberrant events such as chromosomal translocations or telomere fusions (63). It is emerging that many cancers rely on these backup pathways for DSB repair and ultimately survival (64,65). Research is ongoing into how these pathways can be exploited for therapeutic potential. It is therefore crucial to understand not only how these pathways function individually but also the interplay and fine regulatory balance that exists between them.

DSB repair pathway choice

Many factors that regulate DSB repair pathway choice act either to promote or suppress DNA end resection; a critical event that dictates which pathways can be used for repair (66,67). For example, c-NHEJ requires little to no resection whereas homology directed repair (HDR) needs at least 20 bp (68). Mediators controlling DNA

end resection are regulated in a cell cycle dependent manner to ensure that the most appropriate pathway is used.

In G1, ATM mediated phosphorylation of MDC1 recruits the E2 ubiquitin conjugating enzyme UBC13 and E3 ubiquitin ligase ring finger protein 8 (RNF8) (33). The ubiquitylation mediated by these proteins promotes recruitment of another E3 ligase, RNF168, which ubiquitylates histone H2A prompting localisation of several key DSB response factors (34). Among them, is p53 binding protein 1 (53BP1), a critical regulator of DNA end resection. 53BP1 further associates with chromatin through its tandem Tudor domains, which bind methylated lysine residues on histone H4 (69). 53BP1 also undergoes phosphorylation by ATM in its N-terminus, which facilitates interactions with its effector proteins RAP1 interacting factor 1 (RIF1) and PAXX transcription activation domain interacting protein (PTIP) (70,71). Together these proteins inhibit DNA end resection by preventing the tumour suppressor and E3 ligase BRCA1 from localising to DSB sites (Figure 1-3A) (72–74).

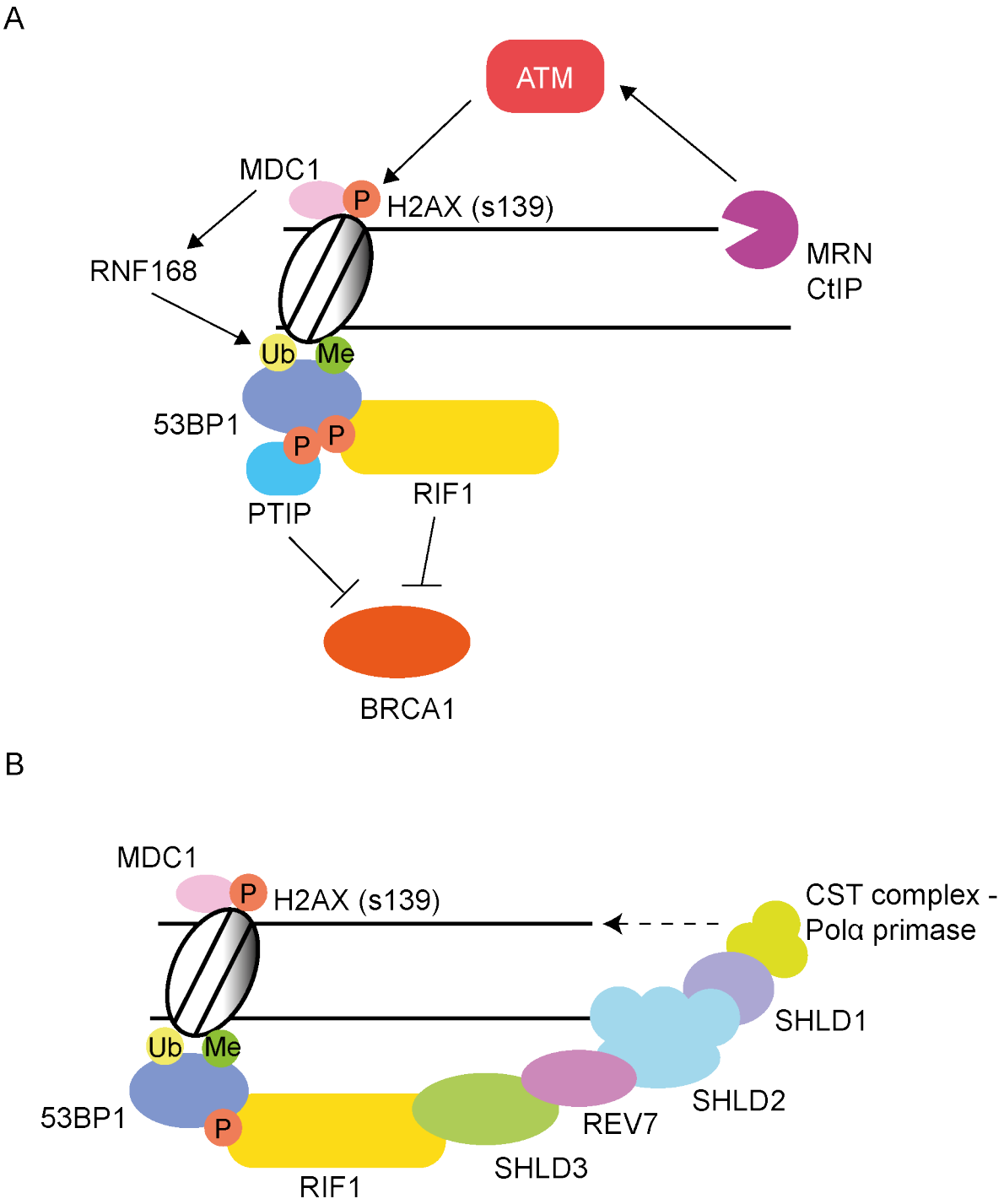


Figure 1-3: Anti-resection functions of 53BP1

A. 53BP1 is recruited to DSBs following ATM signalling events. Phosphorylation of 53BP1 promotes recruitment of its effector proteins PTIP and RIF1, which prevent BRCA1 chromatin accumulation. B. 53BP1-RIF1 recruits the Shieldin complex (SHLD1-3 and REV7), which physically blocks DNA end resection and recruits the CST complex, which promotes gap-filling at resected DNA ends.

53BP1 also antagonises DNA end resection through its recently discovered effector complex Shieldin (Figure 1-3B) (75,76). The Shieldin complex comprises four proteins: Shieldin 1-3 (SHLD1-3) and REV7. These proteins localise to DSBs in a 53BP1 and RIF1 dependent manner and prevent resection by physically blocking DNA ends. More recently it has been shown that these proteins further associate with another complex called the CTC1-STN1-TEN1 (CST) complex. The CST complex associates with polymerase α which promotes fill-in DNA synthesis at resected DNA, both at sites of DSBs and telomeres (72,77).

Once a cell enters S phase, CDK activity promotes phosphorylation of the pro-resection factor CtBP-interacting protein (CtIP). CtIP is an endonuclease which localises to DSBs in an inactive form via an interaction with the MRN complex (78). Phosphorylation of CtIP allows it to bind BRCA1, which leads to eviction of 53BP1 and RIF1 from chromatin thus promoting DNA end resection and use of HR (79).

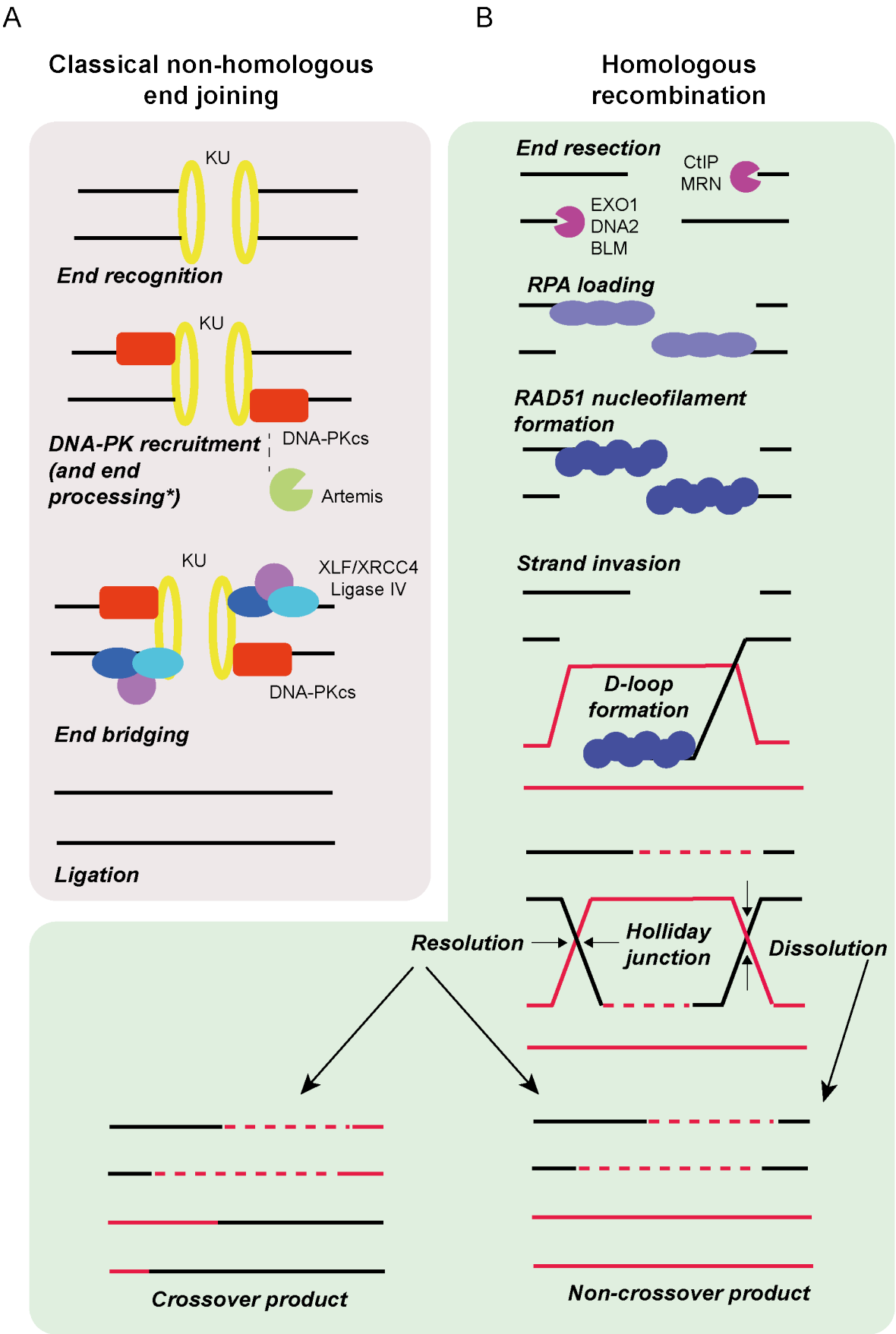
The initiation of DNA end resection represents the major decision point in DSB repair pathway choice. While its initiation is tightly controlled in a cell cycle dependent manner, DNA resection can be reversed and is a highly dynamic process.

Classical non-homologous end joining

The principle of c-NHEJ is relatively simple, in that it ultimately involves the re-joining of broken DNA ends (Figure 1-4A). However, DSBs can present in various forms with different end configurations and thus require different degrees of processing and context-specific repair factors.

In most events of c-NHEJ, the first step involves recognition of the DNA ends by the KU heterodimer which can recognise various kinds of broken DNA termini including blunt ends, incompatible overhangs and 3' phosphoglycolated (3'PG) ends (80). Fundamental to all modes of c-NHEJ is the XRCC4-DNA ligase IV complex which is recruited to DSB sites by KU through an interaction with the BRCA1 C-terminus (BRCT) domains in ligase IV (81,82). XRCC4 associates with the structurally similar protein XRCC4 like factor (XLF), which together form helical filaments encompassing the DNA to promote bridging of DNA ends and enhance ligase IV activity (83,84). The recently characterised paralog of XRCC4 and XLF (PAXX) which directly interacts with KU, also stabilises DNA bridging and promotes ligation (85).

These proteins form the core c-NHEJ machinery but for more complex substrates, the actions of additional factors may be required. For example, the endonuclease Artemis, which is recruited and activated by DNA-PKcs can process 5' and 3' overhangs, hairpins and 3'PG ends to generate substrates suitable for ligation (56). In some instances, polymerase mediated gap-filling is required to complete repair. This is principally carried out by highly mutagenic Pol X family members polymerase λ and polymerase μ which do not rely on a template for nucleotide incorporation (86,87).


A
Classical non-homologous end joining
KU
End recognition
KU
DNA-PK recruitment (and end processing*)
DNA-PKcs
Artemis
KU
XLF/XRCC4
Ligase IV
DNA-PKcs
End bridging
Ligation
B
Homologous recombination
End resection
CtIP
MRN
EXO1
DNA2
BLM
RPA loading
RAD51 nucleofilament formation
Strand invasion
D-loop formation
Resolution
Holliday junction
Dissolution
Crossover product
Non-crossover product

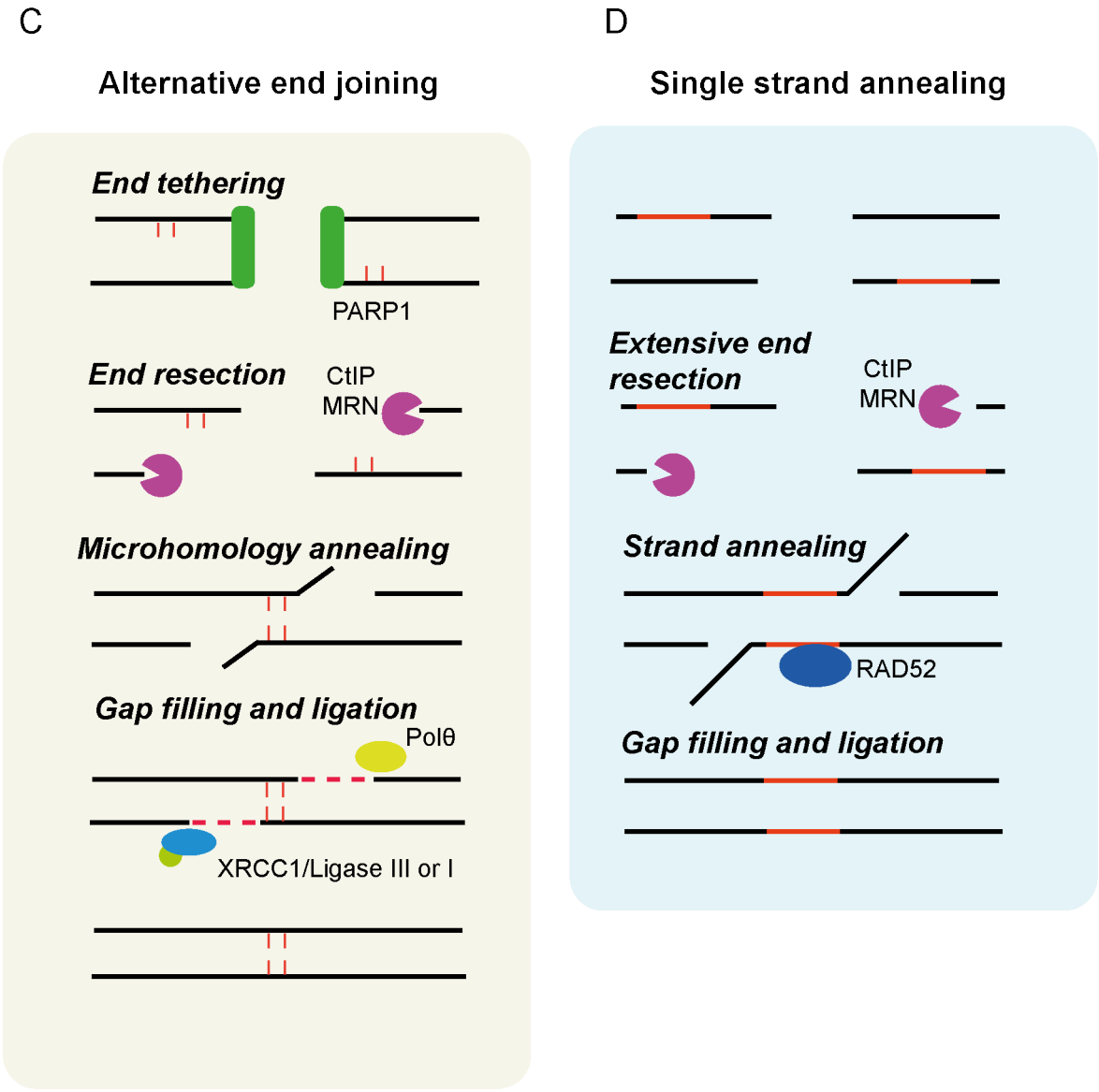


Figure 1-4: Overview of canonical and non-canonical DSB repair pathways

A. Classical non-homologous end joining B. Homologous recombination C. Alternative end joining D. Single-strand annealing (* end processing is only required at certain types of DSBs).

VDJ and Class Switch Recombination

In addition to the repair of pathological DSBs, c-NHEJ factors are vital for the repair of programmed DSBs during antibody diversification. Class switch recombination (CSR) and variability diversity joining (VDJ) recombination are two mechanistically distinct processes that generate genetic diversity in the constant and variable regions of immunoglobulin loci, respectively. VDJ recombination occurs in developing B and T

cells whereas CSR takes place in naïve B cells following contact with an antigen (88). VDJ recombination involves rearrangement of the VDJ genes on the heavy chain and V and J genes on the light chain. CSR is responsible for changing the antibody isotype in the constant region of the heavy chain (88). Both CSR and VDJ recombination occur via a deletional recombination event which is mediated through end-joining (Figure 1-5) (89,90). In both instances, two DSBs are induced at specific DNA sequences, and it is often the case that the recombining elements are located far apart from one another. Thus, the chromatin architecture must be altered to facilitate alignment of these regions to promote end synapsis and deletion of the intervening sequence (91,92). Loss of core c-NHEJ proteins as well as the anti-resection factor 53BP1 and its effector proteins are associated with deficiencies in CSR and VDJ recombination (76,93,94). In the absence of these proteins, the reliance on the mutagenic alt-EJ pathway increases. This involves alignment between microhomologies and can cause chromosomal translocations (95).

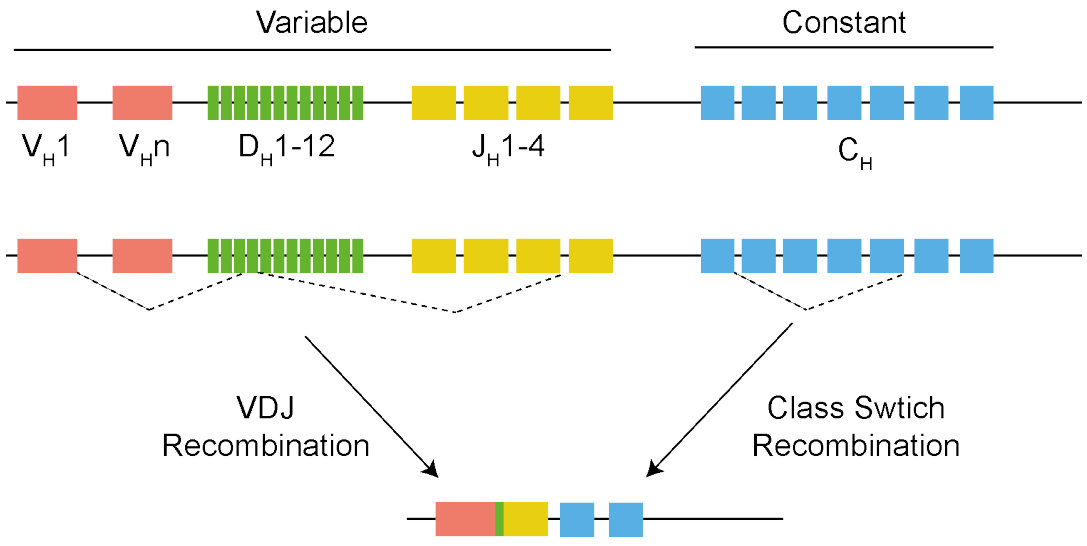


Figure 1-5: VDJ and class switch recombination

Simplistic overview depicting deletional recombination between genetic elements in the variable and constant regions of the immunoglobulin heavy chain locus.

Toxic c-NHEJ

c-NHEJ can be used throughout the cell cycle for DSB repair. However, DSBs that arise during S phase due to replication fork collision with DNA lesions such as TOPO I-linked SSBs or trapped PARP1, are not suitable substrates for c-NHEJ since they only have one DNA end available for repair. These DSBs are termed single-ended DSBs (seDSBs) and strictly rely on HR for repair (68). Consequently, HR-deficient cells are highly sensitive to TOPO I and PARP inhibitors and it has been shown that lack of proficient HR leads to the accumulation of toxic chromosomal fusions mediated by c-NHEJ (5,96). In *BRCA1*-deficient cancers specifically, sensitivity to PARP inhibition can be rescued by depletion of 53BP1, RIF1 and components of the Shieldin and CST complexes (70,75,77,96). Notably, this effect cannot be achieved by depleting core c-NHEJ factors such as ligase IV highlighting the strong influence that DNA resection has in determining pathway choice (96).

Homologous recombination

During DNA replication, sister chromatids become available as repair templates for HDR pathways. Although this means that HR is available for DSB repair in both S and G2 phase, it is now thought that it is primarily relied upon for the repair of seDSBs (97).

Upon entering S phase, DNA resection is licensed by CDK2 activity and BRCA1-mediated displacement of 53BP1, triggering CtIP and MRE11 to produce 3’ overhangs via their 3’-5’ exonuclease activity. More extensive resection is carried out in the 5’-3’ direction by the exonucleases EXO1 and DNA2 with the aid of several DNA helicases (Figure 1-4B) (98). RPA then coats the exposed ssDNA but is subsequently displaced by the recombinase RAD51. BRCA1 is required for recruitment of partner and localiser of BRCA2 (PALB2), which in turn recruits and cooperates with BRCA2 in RAD51

loading (99,100). The RAD51 filament then invades the homologous sister chromatid generating an intermediate known as a D-loop (Fig 1-4B). DNA synthesis commences from the 3' end of the invading strand until it reaches the other DSB end (101). Extended DNA can be annealed directly to the DSB end in a process known as synthesis dependent strand annealing (SDSA) which results in no crossovers (Figure 1-4B). Anti-recombinogenic factors can act to limit these events at different stages. For example, the helicase regulator of telomere elongation helicase 1 (RTEL1) can disengage the D-loop structure (102). F-box helicase 1 (FBH1) and RECQ5B are also helicases but these disrupt formation of the RAD51 nucleofilament (103,104).

Instead of end capture, the other DNA end can also invade the template to form Holliday junctions (HJ). These four-way DNA structures must be processed correctly to ensure faithful chromosome segregation during mitosis (105). Dissolution of HJs is performed by the BTR complex, comprised of the helicase Bloom's syndrome protein (BLM), topoisomerase IIIA (TOPO IIIA), and RECQ mediated genome instability proteins 1 and 2 (RMI1/2) (106,107). This method of coordinated unwinding and decatenation of HJs results solely in non-crossover products. If this method of HJ dissolution fails, structure-specific endonucleases such as SLX1-SLX4, MUS81-EME1 and GEN1 can serve as important backups to ensure timely HJ resolution (108). However, this can generate crossover and non-crossover products, thus the action of these proteins is strictly regulated, in a cell cycle dependent manner, since crossovers in mitotic cells could drive chromosomal instability (103).

Alternative end joining and single-strand annealing

Deficiencies in c-NHEJ and HR can increase cellular reliance on highly mutagenic, backup pathways, such as alt-EJ, for repair. In most cases, alt-EJ uses regions of

microhomology at DSB sites (63). Under normal circumstances, alt-EJ is suppressed by c-NHEJ factors such as KU and XRCC4-ligase IV (109). In the absence of KU binding, it has been proposed that PARP1 engages and tethers DNA ends while promoting recruitment of the MRN complex and CtIP (Figure 1-4C) (16,110). Like in HR, this initiates DNA end resection which means use of alt-EJ is also confined to S and G2 phases of the cell cycle.

Exposure of ssDNA results in RPA binding which normally suppresses alt-EJ and promotes HR (111). There is also evidence to suggest that HR factors act to suppress alt-EJ, revealing that it shares an antagonistic relationship with both major DSB repair pathways (112). Recently, it has been demonstrated that the helicase domain of the gap-filling polymerase Polθ, which is required for alt-EJ, can actively remove RPA from ssDNA (113). Polθ therefore promotes annealing between exposed regions of microhomology and further stabilises the repair substrate through gap-filling. Finally, ligase III is the main ligase responsible for sealing DNA ends during alt-EJ although in some scenarios, likely in the absence of microhomologies, ligase I can also perform this function (114).

In addition to alt-EJ, SSA is another non-canonical DSB repair pathway that can be used to repair breaks that have undergone extensive end resection (Figure 1-4D). RAD52 is the key mediator of SSA and is required for synapsis of the long regions of homology (18). SSA is highly mutagenic since it results in large deletions between the recombining elements. Interestingly, in the absence of both 53BP1 and BRCA1, RAD52 is required for cell viability and SSA has been shown to contribute significantly to the repair of DSBs in these cells (17). Amplification of *RAD52* is observed in some human cancer cells and its loss or inactivation can suppress tumorigenesis (115,116).

Thus, non-canonical DSB repair can drive genome instability and it is important therefore, to understand the contexts in which these backup pathways function.

DSB repair is a highly regulated process and cells have evolved several pathways to ensure swift resolution of these highly toxic lesions. Defects in any one of these processes can have serious consequences by increasing the reliance on other pathways that in certain circumstances, may not be appropriate for repair. Interestingly, mounting evidence suggests that the functions of HR factors extend beyond DSB repair, particularly during DNA replication, in which they play key roles in replication fork protection, even in unperturbed conditions.

1.5. Replication stress

Replicating the human genome is an incredibly complex and demanding task. It is essential that DNA is duplicated faithfully and that any challenges to this process are effectively overcome. DNA replication commences at thousands of replication origins following cell cycle dependent licensing which triggers DNA unwinding by the replicative CDC45-MCM2-7-GINS (CMG) helicase (117,118). DNA replication then proceeds via continuous and discontinuous DNA synthesis on the leading and lagging strand, respectively.

Replication stress is a loosely defined term used to describe events in which replication fork progression is hindered causing fork slowing or stalling and ultimately results in reduced rates of DNA synthesis (42). Genome instability is strongly linked to high levels of replication stress since failure to complete DNA replication, otherwise known as under-replication, is a known source of chromosome segregation errors (39). Under-replication results in unresolved replication intermediates such as ssDNA gaps, which can persist into mitosis and result in chromosome breakage if not properly resolved.

For example, under-replication can lead to elevated levels of DNA structures known as ultrafine anaphase bridges (UFBs) and inherited DNA damage in the form of 53BP1 nuclear bodies (119–121). It can also increase the prevalence of mitotic DNA synthesis (MiDAS) (122).

While DDR related diseases are clinically diverse, a common feature among many though not all, is their predisposition to cancer. Accumulation of various types of unresolved DNA damage can generate high levels of replication stress and promote tumorigenesis (44). As previously discussed, obstacles that impede DNA replication take many forms and can arise from numerous sources. To combat this, cells rely on a range of mechanisms to protect and stabilise replication forks to ensure DNA replication is completed faithfully and genome stability is maintained.

DNA damage tolerance and replication fork rescue

If a replication fork stalls at a DNA lesion, there are several mechanisms of rescuing it from collapse and restarting DNA replication. It is critical that replication is resumed as soon as possible to avoid extensive fork stalling, thus cells possess DNA damage tolerance (DDT) pathways that allow DNA replication to continue without removing the DNA lesion.

For example, translesion synthesis (TLS) involves switching the replicative polymerases used during normal DNA replication for structurally distinct, lower fidelity ones (123,124). These polymerases can replicate over DNA lesions due to their flexible active site but are inherently mutagenic, and therefore have a high rate of misincorporation. Template switching (TS) is another DDT pathway which promotes error-free bypass of DNA lesions by recombination based mechanisms (125). This facilitates continuation of DNA replication by transient switching from the lesion-

containing DNA template to the undamaged sister chromatid. While TS plays an important role in maintaining genome stability, it can also result in genomic rearrangements within repetitive DNA regions (126). In addition to the DDT pathways, DNA can also be re-primed ahead of lesions, but similarly to TS this inevitably leads to the formation of ssDNA gaps that must be filled post-replication (127).

Fork reversal is another mechanism by which replication forks can tolerate DNA lesions (128). This process is mediated primarily by DNA translocases, like SMARCAL1, HLTF and ZRANB3, which belong to the sucrose non-fermentable 2 (Snf2) ATPase superfamily (Figure 1-6). These enzymes catalyse annealing of nascent DNA strands to produce the four-way, reversed fork (RVF) structure (129,130). The recombinase RAD51 is also essential for fork reversal, in a BRCA2 independent manner, but it is still unclear how exactly it contributes to the process (131). Controlled processing of the regressed arm is carried out by nucleases such as DNA2 which generate sufficient ssDNA for RPA binding and subsequent BRCA1 and BRCA2 mediated loading of RAD51 (132). In this instance, RAD51 contributes to fork restart through HDR mechanisms. Failure to protect and stabilise RVFs can result in MRE11 mediated degradation of the regressed arm and lead to genome instability (133,134). *BRCA1/2*-deficient cells exhibit sensitivity to replication stress which can be partially rescued by depleting translocases such as ZRANB3 and SMARCAL1 (135). This highlights that HR factors are critical in both the response to replication stress and DSB repair. Restoration of RVFs can also be catalysed by RECQ1 helicase which is inhibited by PARP1 until the damage is resolved (136).

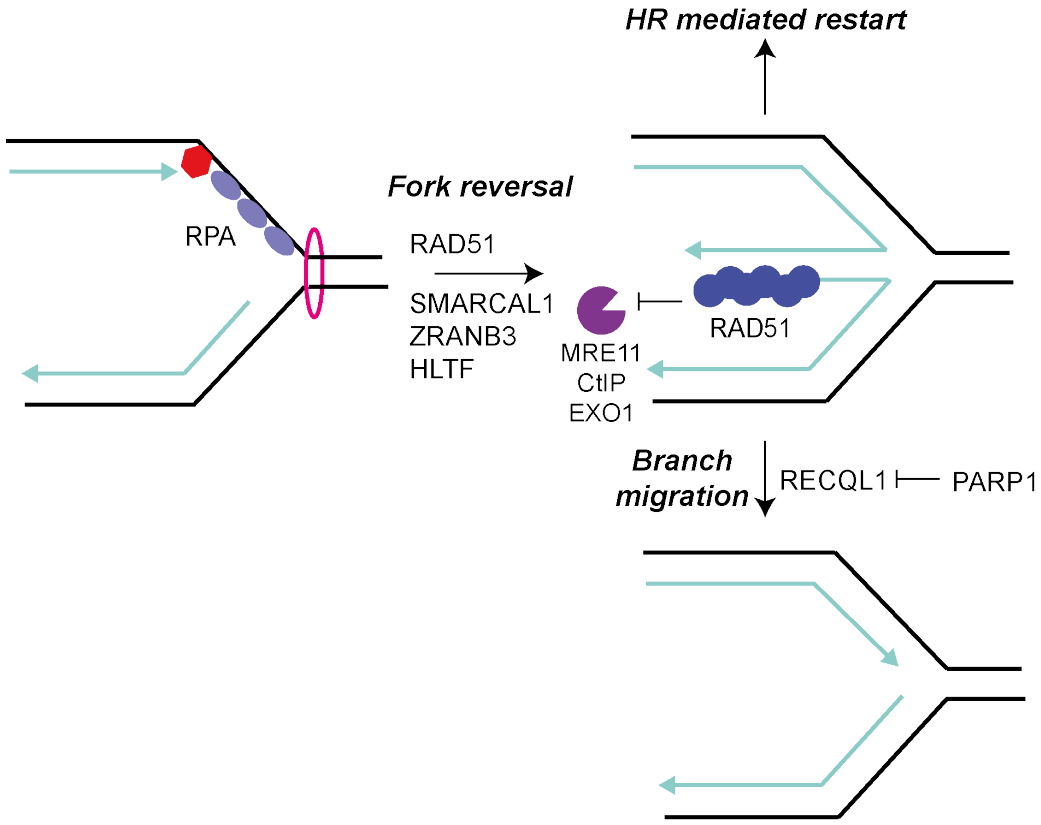


Figure 1-6: Replication fork reversal

Schematic depicting factors involved in catalysing replication fork reversal and restoration.

Fork collapse and S phase associated DSBs

Replication fork collapse can occur after prolonged periods of fork stalling or replication run-off at SSBs (133). Fork collapse is not strictly defined but it is generally associated with disengagement of the replisome from the fork and formation of seDSBs (137). seDSB formation can occur through active cleavage of the fork structure by nucleases, such as MUS81-EME2, or through passive breakage (138,139).

As previously mentioned, HR is used to repair seDSBs during S phase since use of c-NHEJ could result in chromosomal translocations. This mode of HR, called break induced replication (BIR), relies on conservative DNA synthesis (140,141). Consequently, this pathway is used to restart replication forks that have collapsed due to prolonged stalling or upon collision with pre-existing lesions such as SSBs or DNA-

protein complexes. It is also the mechanism by which telomeres are maintained in alternative lengthening of telomeres (ALT) positive human cancers that lack telomerase (142). BIR can occur in a RAD51-dependent or independent manner, although both require RAD52. The former is initiated as in HR repair of two-ended DSBs. In yeast, the Pol32 subunit of Polδ is required for BIR but not for normal DNA replication (143). Interestingly, the mammalian ortholog of Pol32, POLD3, is essential for cell survival, which could hint that reliance upon BIR has increased with genome complexity (144). The mechanism for RAD51-independent BIR is much less understood. It is thought that RAD52 mediates annealing between the invading strand and exposed ssDNA, possibly in pre-existing D-loops, R-loops or even at stalled replication forks. Thus, this mode of BIR can lead to chromosomal translocations and copy number variations (65,141,145).

'Difficult to replicate' regions

There are certain regions of the genome that are particularly vulnerable to replication stress. This can be attributed to unique features characteristic of certain chromosomal loci such as repetitive DNA, chromatin condensation and high levels of gene expression, which can all present endogenous challenges to the progression and completion of DNA replication.

For example, common fragile sites (CFSs) are genomic regions that are naturally susceptible to DNA breakage and are frequently associated with chromosome rearrangements which drive cancer progression (146). These sites often reside in large transcribed regions where conflicts between replication and transcription machineries are frequent (147). There is also evidence to suggest that AT-rich DNA sequences contribute to the inherent fragility of certain fragile sites (148,149). Despite the

propensity of CFSs to experience replication stress, a lack of proximal replication origins has been reported for a subset of these sites, reducing the ability of replication forks to recover from stalling (150). Some CFSs are also late-replicating, further potentiating under-replication and chromosomal instability (151,152). A host of DDR factors have been implicated in maintaining genetic stability at CFSs including DSB repair factors, FA proteins as well as several helicases and nucleases (153–156).

Telomeres are the repetitive DNA regions found at the end of linear chromosomes and represent another specialised genomic environment. It is crucial that telomeres are protected from degradation and are adequately distinguished to avoid being erroneously mistaken for a DSB end which could lead to aberrant DDR activation and promote telomeric fusions (157). Telomeric DNA is comprised of short tandem repeats and an extensive G-rich 3' overhang at the very end of the chromosome. This region of ssDNA invades the telomeric duplex to form a lariat structure known as a T-loop, thus disguising the chromosome end (158). Telomeric repeats further engage a multiprotein complex called Shelterin which promotes T-loop formation and suppresses DDR activation (159,160). Thus, the unique configuration of telomeric DNA requires the action of specialised DDR factors to facilitate replication at these regions. Several helicases have been implicated in promoting efficient telomere replication including RTEL1 and Werner's syndrome (WRN) and BLM helicase (161–163). RTEL1 is required for T-loop disassembly as well as resolution of secondary structures known as G4 quadruplexes (161). Failure to dissolve the T-loop can result in inappropriate cleavage by structure-specific endonucleases which drives telomere shortening. As the replication fork moves through telomeric chromatin and approaches the T-loop, positively supercoiled DNA accumulates ahead of the fork which is relieved by the Shelterin subunit telomere repeat binding factor 2 (TRF2) along with TOPO II

(164,165). Interestingly, the fork remodeller SMARCAL1, which catalyses fork reversal, has also been shown to be important for telomeric DNA replication (166). Normally, SMARCAL1 activity is repressed by ATR kinase, but local DDR suppression may permit increased use of this tolerance mechanism, which may be relied upon more heavily given that use of other DDR mechanisms are restricted (167).

The response to replication stress is a complex process that involves the convergence of several arms of the DDR. Together these pathways ensure replication forks are processed and protected in the face of DNA damage, ultimately ensuring the timely and faithful completion of DNA replication. While barriers to replication can occur throughout the genome, there are certain regions which are inherently more difficult to replicate and particularly sensitive to replication stress. Understanding the unique compensatory mechanisms that exist at these regions to offset the inherent challenges they pose, could reveal hitherto, unappreciated nuances within the DDR. While CFSs and telomeres have received a notable amount of attention in the context of genetic stability, centromeres are complex and enigmatic regions of the genome that are relatively understudied. However, recent studies have begun to reveal some intriguing features of centromere stability that could have important implications for human disease.

1.6. Centromeres

Centromeres are specialised regions of eukaryotic genomes that are best known as the site of kinetochore assembly and spindle attachment during mitosis (168). Failure to establish functional kinetochores can result in chromosome segregation errors, thus centromeres play a central role in maintaining genome stability (169). Most centromeres are composed of highly repetitive DNA sequences. At human

centromeres, the basic repeat unit is a 171 bp AT-rich sequence termed alpha-satellite DNA (170,171). This monomer is repeated in a tandem, head-to-tail fashion to form extensive arrays of conserved higher order repeat (HOR) units spanning millions of bases. While HOR units display high levels of conservation, the primary sequences of the individual alpha-satellite monomers comprising them are significantly diverged (171).

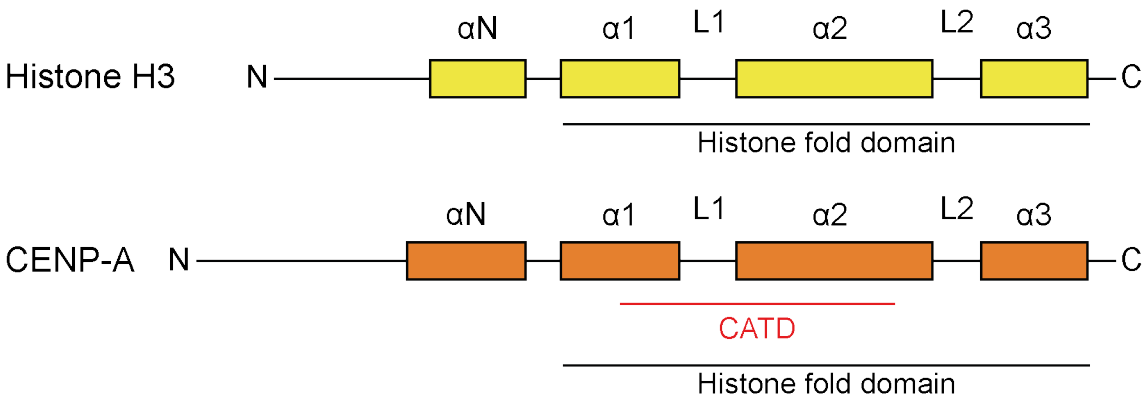


Figure 1-7: Structural features of CENP-A and histone H3

Schematic showing structural features of the centromeric histone H3 variant CENP-A compared to canonical H3. CATD = CENP-A targeting domain.

Although repetitive DNA is a common feature, centromeres are defined epigenetically by the presence of the centromeric histone H3 variant centromere protein A (CENP-A) (Figure1-7) (172,173). CENP-A nucleosomes possess distinctive structural features that are required for establishing and maintaining the unique chromatin environment at centromeres. For example, the first loop and second α-helix of the CENP-A histone fold domain is required for targeting the protein to centromeres, through an interaction with its own specific chaperone Holliday junction recognition protein (HJURP) (174). Thus, this region is known as the CENP-A targeting domain (CATD). The CATD is also known to mediate interactions with components of the constitutive centromere associated network (CCAN).

Constitutive centromere associated network

The CCAN is an extended network of 16 proteins that comprise the inner kinetochore and associate with CENP-A nucleosomes throughout interphase. The CCAN proteins are typically divided into 4 subcomplexes: CENP-LN, CENP-HIKM, CENP-TWSX and CENP-OPQUR, plus CENP-C which acts as a scaffold for the other components (Figure 1-8) (172).

CENP-C interacts with two regions of CENP-A: the CATD and the final six residues in its C-terminal tail as well as making contact with other histones within the CENP-A nucleosome (175). Studies have also shown that depleting CENP-C from cells results in a reduction of CENP-A from centromeres highlighting that it plays a key role in maintaining centromere identity (175). It has been suggested that association with CENP-C directly promotes the stability of CENP-A nucleosomes (176).

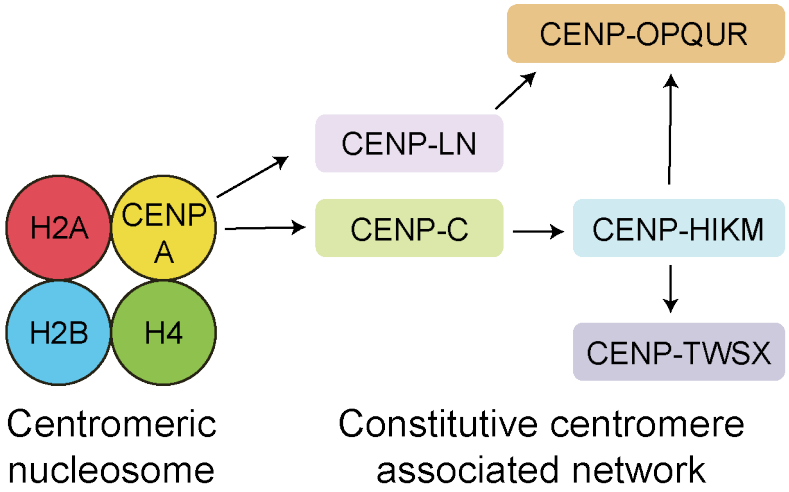


Figure 1-8: Constitutive centromere associated network

Schematic depicting physical relationships between CENP-A containing nucleosome and components of the CCAN.

The CENP-LN complex is recruited to CENP-A nucleosomes through a direct interaction between CENP-N and the CATD in CENP-A (177). CENP-N also makes extensive contact with the nucleosomal DNA which is important for stabilising its interaction with CENP-A as well as for its recruitment of CENP-L (178). CENP-LN also interacts with CENP-C through the C-terminus of CENP-L. While this interaction is dispensable for the recruitment of CENP-LN to CENP-A nucleosomes during interphase, it is required for retention of the complex in mitosis (179).

CENP-TW and CENP-SX are two heterodimeric subcomplexes which bear striking resemblance to canonical histones and together form a nucleosome-like heterotetramer (180). The formation of the CENP-TWSX complex, as well as its ability to bind and supercoil DNA, is critical to kinetochore assembly. Interestingly, unlike canonical nucleosomes, which negatively supercoil DNA, the CENP-TWSX tetramer promotes positive DNA supercoiling resulting in a unique topological arrangement of centromeric DNA (181). CENP-TWSX also interacts directly with the CENP-HIKM complex. CENP-T, specifically, is required for recruitment of CENP-HIKM subunits (177). Equally, CENP-HIKM has been shown to promote CENP-T stability demonstrating an interdependency between the two complexes (182). Furthermore, the CENP-HIKM complex is important for the centromeric recruitment of chromatin remodelling enzymes such as the FACT complex and CHD1 (183).

Finally, along with CENP-C and CENP-LN, CENP-HIKM is responsible for recruiting the CENP-OPQUR complex. The CENP-OPQUR complex is important for recruiting several factors involved in chromosome orientation and congression during mitosis. For example, CENP-Q is required for the recruitment of the motor protein CENP-E and phosphorylated CENP-U recruits polo-like kinase 1 (184,185). Both CENP-Q and

CENP-U can directly bind microtubules. CENP-Q, in particular, binds microtubules through its basic N-terminus (186).

Maintaining genetic stability at centromeres

Maintaining the integrity of centromeric DNA is an intrinsically challenging task due to its highly repetitive composition (187,188). Repetitive DNA sequences are difficult to replicate given their propensity to form secondary DNA structures, which can impede replication fork progression. Furthermore, in humans, centromeres are late-replicating regions which could potentially exacerbate these challenges by increasing the likelihood of under-replication (189). A genome-wide screen which mapped the location of replication stress induced DSBs in cancer cell lines, revealed that breaks occur more frequently within centromeric sequences (190). Given this inherent fragility, it is unsurprising that recent studies have begun to uncover unique regulatory processes that govern DNA replication and repair at these regions.

DNA replication at centromeres

An intriguing study by Aze et al, used bacterial artificial chromosomes (BACs) containing human chromosomal DNA in *Xenopus laevis* egg extracts to reconstitute centromeric DNA replication (191). They found that centromeric chromatin is enriched for DDR factors, like KU80, MRE11-RAD50, PARP1 and the mismatch repair factors MSH2/6 (191). DNA mismatch repair (MMR) is a highly conserved, post-replicative repair mechanism that removes nucleotides that have been misincorporated during DNA replication (192). In eukaryotes, mismatch recognition is carried out by the MSH2/6 or MSH2/3 heterodimers, depending on the extent of the mismatch, which form an ATP-dependent sliding clamp (192). Once the damage is detected, the ATPase activity of MSH2/3-6 is activated which induces conformational changes that

in turn promote recruitment of another heterodimer MLH1/PMS2, which promotes excision of the mispaired base(s) (192).

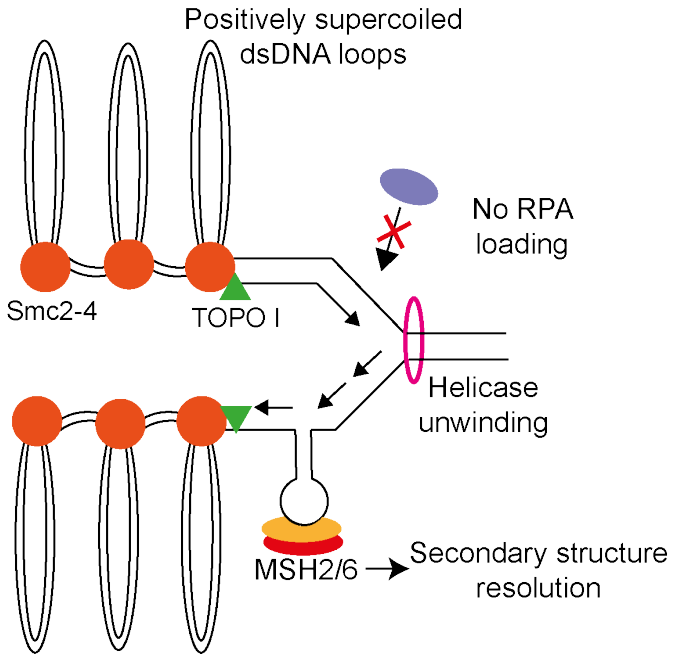


Figure 1-9: Model of centromeric DNA replication

Cartoon showing centromeric DNA replication based on the model proposed by Aze et al (191). ssDNA formation is limited by topological arrangement of DNA behind the replication fork. Thus, RPA loading does not occur, and ATR kinase is not activated. MSH2/6 are proposed to promote resolution of secondary DNA structures and facilitate DNA replication.

Aze et al demonstrated that MSH6 is required for centromeric replication in their system and propose that it counteracts secondary DNA structures to facilitate replication (Figure 1-9) (191). However, it is unclear whether this would be achieved through recruitment of canonical MMR factors. MSH2 was also enriched on centromeric chromatin in their system, but MSH3, MLH1 and PMS2 were not. Interestingly, a recent screen for factors involved in promoting CENP-A deposition and maintenance in human cells identified PMS2 as a contributing factor (193). Thus, the exact role of MMR proteins at centromeres remains to be elucidated but represents an exciting topic for future research.

Furthermore, Aze and colleagues showed that activation of ATR kinase is suppressed at centromeric chromatin, since ssDNA formation is constrained by the topological arrangement of centromeric DNA (191). Specifically, they show that centromeric DNA is actively arranged into positively supercoiled double-stranded DNA (dsDNA) loops, most likely behind replication forks in a TOPO I dependent manner that is possibly aided by the condensins structural maintenance of chromatin proteins 2-4 (SMC2-4) (Figure 1-9). While achieved through a distinct mechanism, this is somewhat reminiscent of DDR suppression at telomeres and could hint that centromeres, similarly, rely on alternative mechanisms for maintaining genetic stability.

Following completion of replication, centromeric DNA has a high propensity to become catenated leading to the formation of UFBs. These DNA structures arise physiologically between sister chromatids during mitosis but require timely resolution to avoid micronuclei formation, inherited DNA damage and chromosome segregation errors (119). UFBs originate from other unique genomic locations, albeit less frequently, such as telomeres and CFSs and are categorised accordingly (119). Central to centromeric UFB resolution is TOPO II, the inhibition of which increases the prevalence of these structures (194). Together with BLM helicase and the Snf2 ATPase PLK1-interacting checkpoint helicase (PICH), TOPO II promotes decatenation of UFBs (195,196). In addition to dsDNA catenanes, recombination intermediates can also be a pathological source of UFBs (197). Given that centromeric DNA is particularly vulnerable to replication stress and DNA breakage, replication and recombination intermediates may also contribute to the formation of centromeric UFBs.

DSB repair at centromeres

The inherent fragility of centromeres renders them vulnerable to DNA breakage. This may be in part reflected by the array of DSB repair factors found to be enriched on centromeric chromatin (191). Centromeres are embedded within wider regions of highly condensed and transcriptionally silent, heterochromatin, marked epigenetically by tri-methylation of H3 lysine 9 (H3K9me3) (198). This modification is responsible for the recruitment of heterochromatin protein 1 (HP1); a fundamental structural component of heterochromatin (199).

Several studies have made fascinating insights into how this unique chromatin environment influences DSB repair. Illegitimate recombination between homologous sequences on different chromosomes can drive chromosomal instability and cause diseases such as cancer (200). Thus, HR may seem like an unnecessarily risky method of DSB repair within heterochromatin, given that c-NHEJ would only cause small insertion/deletion events, which would likely have an insignificant impact on vast arrays of repetitive DNA. Yet studies have shown that HR does indeed contribute to DSB repair within heterochromatin. It has been demonstrated that DSBs occurring in *Drosophila* heterochromatin, are relocated outside of heterochromatin domains (201). This process occurs after DNA resection and is dependent on ATR kinase. Once the break has been relocated, RAD51 binds to the resected DNA and facilitates repair (201). It is proposed that the spatial disconnect between end resection and strand invasion limits illegitimate recombination between the same or similar DNA sequences.

Physical relocation of DSBs has also been observed during repair of heterochromatic DSBs in mammalian cells (202,203). One study specifically compared DSB repair in pericentromeric and centromeric sequences in mouse cells. Remarkably, they found

that the regulation of DSB repair in these two regions differed significantly. For pericentromeric sequences, there is a clear distinction on mode of repair depending on cell cycle stage (203). During G1 phase, c-NHEJ factors are recruited for DSB repair and only in S/G2 phase are breaks moved to the periphery for RAD51 loading and HR mediated repair (203). Conversely, centromeric breaks were observed to recruit c-NHEJ factors and relocate and engage RAD51 in both G1 and G2 phase. They further showed that a subset of DSBs fail to relocate to the heterochromatin periphery and that these breaks are engaged by RAD52, the key player in the mutagenic DSB repair pathway SSA (203). This highlights that error-prone pathways are operational at heterochromatin and perhaps serve as a backup when high fidelity repair cannot be completed.

It is known that mutation rates within heterochromatin are notably higher than in euchromatin (201). While error-free repair is clearly still achievable within these regions, this suggests that the use of mutagenic backup pathways is more prevalent within heterochromatin. It is likely that the consequence of using of these pathways is preferable to large scale genome rearrangements which could result from illegitimate recombination events. There are still many key questions surrounding the regulation of DSB repair within heterochromatin, generally and centromeres, specifically. Namely, how exactly do DSBs relocate and is it an active process catalysed by enzymatic activity? And, importantly, is this process conserved in human cells?

1.7. ERCC6L2: A member of the Snf2 ATPase family

Snf2 ATPases

Snf2 ATPases are a structurally and functionally diverse family of proteins that contain a highly conserved helicase-related ATPase domain (204,205). This catalytic core drives DNA translocase activity meaning that these proteins can track along DNA, which can result in direct or indirect manipulation of DNA-protein contacts or remodelling of specific DNA structures (206). Snf2 ATPases contain a variety of accessory domains which are important for chromatin recruitment, regulation of the ATPase domain and for mediating protein or DNA interactions. Some Snf2 ATPases contain additional domains with enzymatic activity, like ZRANB3 and HLTF which possess nuclease and E3 ubiquitin ligase domains, respectively (207,208).

The DNA translocase activity of Snf2 ATPases can drive several different types of DNA transactions. Chromatin remodelling is a key function of many Snf2 enzymes where their catalytic activity is used to reposition nucleosomes, incorporate/evict histones or exchange histone variants (209,210). Altering chromatin architecture is a fundamental step in all genetic events, as it permits the relevant machineries access to DNA and re-establishes the chromatin environment following completion of such processes. In addition to chromatin remodelling, Snf2 ATPases can also drive remodelling of specific DNA substrates such as the fork reversal enzymes SMARCAL1, ZRANB3 and HLTF, discussed in Section 1.5 (211). RAD54 also binds specifically to HJs during HR and promotes branch migration in an ATPase dependent manner (212). Certain Snf2 ATPases have also been reported to alter DNA topology, which could affect DNA-protein contacts or generate certain DNA conformations or substrates more amenable to further processing (213).

Snf2 ATPases are critical regulators of DNA replication, DNA repair, transcription and chromosome segregation (209). They are, therefore, key players in the maintenance of genome stability. Investigating uncharacterised members of the Snf2 ATPases could therefore uncover important factors in these fundamental processes.

ERCC6L2

Excision repair cross complementation group 6 like 2 (ERCC6L2) is a poorly characterised member of the Snf2 ATPase superfamily conserved in higher eukaryotes. It belongs to the ERCC6 subfamily in addition to ERCC6 or Cockayne Syndrome B (CSB) and ERCC6L or PICH (Figure 1-10A). Despite this subclassification, which is based on conservation within the ATPase domains, these proteins are structurally and functionally divergent. CSB plays a key role in TC-NER in which it uses its DNA translocase activity to remodel DNA at stalled RNA polymerase II complexes (214). It acts as an early sensor to the stalled polymerase and is required for the recruitment of downstream NER factors. The other ERCC6 subfamily member, PICH, as already discussed, is involved in the resolution of DNA structures known as UFBs which arise during mitosis (215).

In contrast to these two proteins, the role of ERCC6L2 is much less defined. Originally, *ERCC6L2* was annotated as a 712 amino acid protein containing an N-terminal Tudor domain and the ATPase catalytic core (216). Tudor domains typically bind methylated lysine and arginine residues on histone proteins. However, a binding partner for the Tudor domain in ERCC6L2 has yet to be identified. Reannotation of ERCC6L2 has since revealed that the primary isoform has an additional 850 amino acids at its C-terminus and arises from alternative splicing at exon 14 (217). Within this C-terminal region, although largely disordered, there is a patch of highly conserved residues

recognised by the Pfam domain database belonging to the 'VIGSSK' family (218). Interestingly, the 'VIGSSK' domain is unique to ERCC6L2 homologs suggesting that it may be important for regulating the specific activity of these proteins. Both the short and long ERCC6L2 isoforms are expressed ubiquitously and there is no evidence for either being expressed in a tissue specific manner (217).

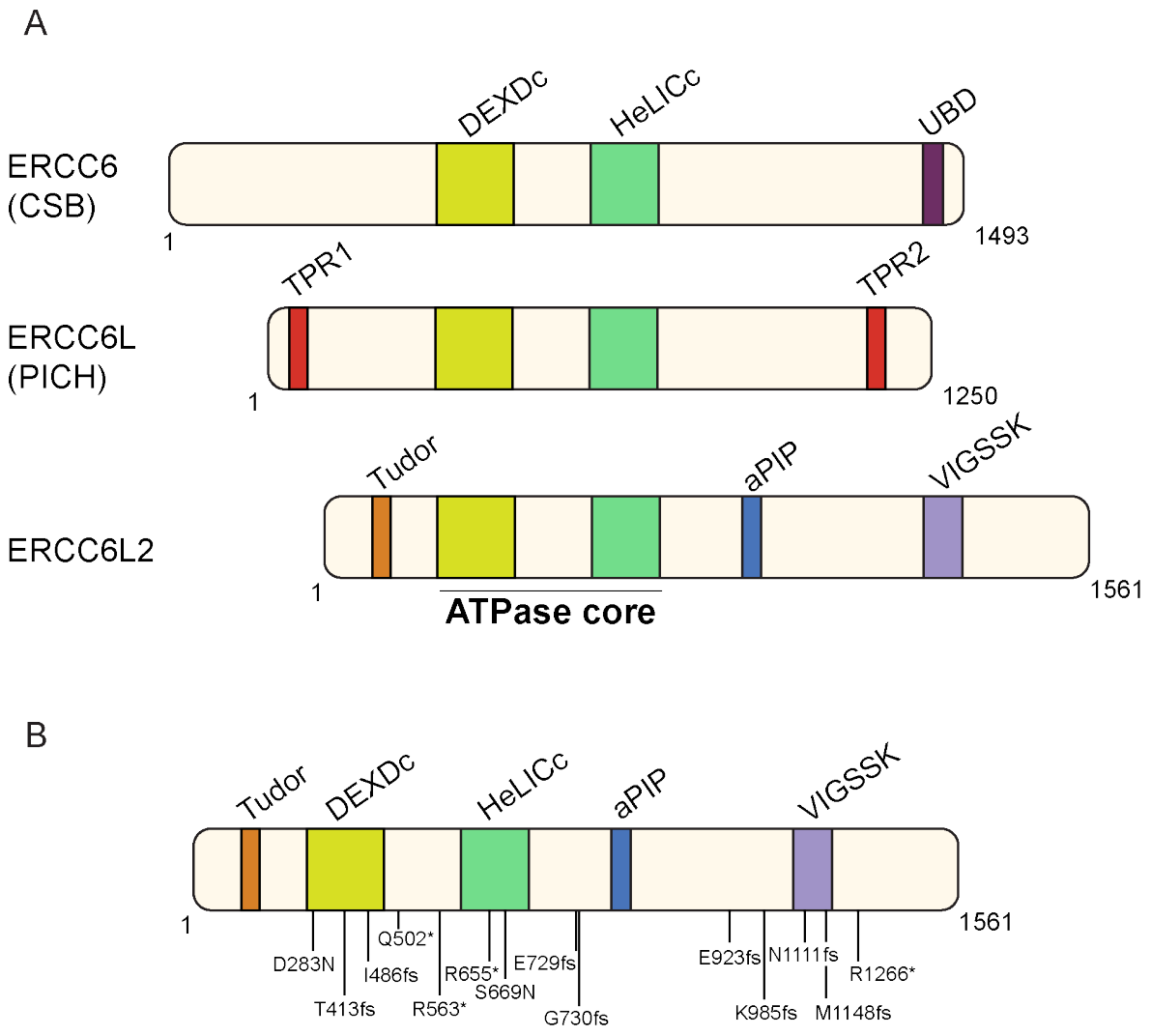


Figure 1-10: Schematic of ERCC6 Snf2 ATPase subfamily

A. Domain architecture of the three ERCC6 subfamily members. (UBD = ubiquitin binding domain. TPR = tetratrico peptide repeats. aPIP = atypical PCNA interaction protein). B. Disease-causing mutations in ERCC6L2.

ERCC6L2 deficiency results in a distinct subtype of inherited bone marrow failure (IBMF) additionally associated with delayed development and microcephaly (216,217,219–221). IBMF syndromes comprise a heterogeneous group of rare genetic blood disorders in which at least one of the three haemopoietic cell types are affected (222). Investigations into the molecular basis of these diseases has revealed that bone marrow failure can arise due to genetic perturbations in several distinct yet interconnected biological pathways (222). These include mutations in genes involved in the FA repair pathway, telomere maintenance and ribosome biogenesis (Table 1-1).

Inherited Bone Marrow Failure Syndrome	**Genetic cause**
Fanconi anaemia	FA/BRCA DNA repair genes
Dyskeratosis congenita	Telomere maintenance
Diamond Blackfan anaemia	Ribosome biogenesis and processing
Swachman Diamond syndrome	Ribosome biogenesis and processing

Table 1-1: Principle IBMF Syndromes

Several IBMF-associated mutations have been reported in the *ERCC6L2* gene (216,217,219–221). These mutations are found throughout the length of the protein and are mostly nonsense or frameshift mutations that result in truncated *ERCC6L2* variants (Figure 1-10B). Two homozygous missense mutations have also been identified at conserved residues within the catalytic core and thus are predicted to perturb the ATPase activity (220,223). Like many IBMF subtypes, *ERCC6L2*-deficiency can predispose humans to myelodysplastic syndrome (MDS) and acute myeloid leukaemia (AML). Seven of the eight reported deaths in patients suffering from *ERCC6L2* deficiency have been from AML, three of which progressed from MDS (220,223). However, the molecular basis of *ERCC6L2*-associated IBMF remains unclear.

Several studies have attempted to elucidate the function of ERCC6L2. One of the first reports into ERCC6L2 demonstrated that loss of the protein resulted in mild sensitivity to MMC and the alkylating agent irofulven in human A549 cells (216). They also showed that under irofulven treatment, ERCC6L2 depleted cells have increased levels of γH2AX compared to wildtype (WT) cells.

Similarly, in another study, Zhang and colleagues found that patient-derived fibroblasts are sensitive to MMC (217). It should be noted that *ERCC6L2*-deficient cells display mild sensitivity to MMC when compared to FA patient cells. *ERCC6L2*-deficient cells also do not exhibit the increased spontaneous chromosomal fragility observed in FA (217). The same study did demonstrate, however, that loss of ERCC6L2 function results in sensitivity to the radiomimetic phleomycin and IR, indicating a DSB repair defect. However, using *in vivo* reporter assays, it was shown that ERCC6L2 depletion did not significantly impair c-NHEJ or HR, the two canonical DSB repair pathways (217). This could suggest that ERCC6L2 only plays a minor role in one of these pathways, perhaps in specific circumstances.

More recently, another cohort of *ERCC6L2* mutations in patients with IBMF was identified (220). They showed that three patient-derived lymphoblastoid cell lines, harbouring different *ERCC6L2* mutations, are sensitive to the transcription inhibitors DRB and actinomycin D, in addition to irofulven, MMC and phleomycin. Following release from acute irofulven treatment, it was shown that the recovery rate of RNA synthesis is impaired in these cells and is accompanied by G2/M cell cycle arrest as well as elevated levels of DSBs (220). It was further demonstrated that these cells fail to efficiently repress transcription elongation upon DNA damage. Subsequently, increased levels of DNA:RNA hybrids were found in the patient cells. Based on these findings, the authors concluded that ERCC6L2 promotes resolution of DNA:RNA

hybrids, and thus protects against transcription associated DNA damage (220). Interestingly, this study also validated an interaction between ERCC6L2 and the DDR kinase DNA-PK. In line with the proposed hypothesis, it is speculated that this interaction could be key in preventing DNA damage arising from stalled RNA polymerase II complexes. However, it is important to consider that DNA-PK has a prominent role in DSB repair and given that *ERCC6L2*-deficient cells also appear to have a DSB repair defect, the significance of this interaction requires further investigation.

Recently, there have been three studies that collectively implicate ERCC6L2 in the DSB response, specifically c-NHEJ and CSR, in line with observations from earlier studies. Using a genotoxic screen, Liu et al revealed that *ERCC6L2* clusters with genes involved in c-NHEJ such as KU, DNA-PKcs and ligase IV (224). Upon depletion of ERCC6L2 from mouse B cells, rates of CSR are significantly reduced (224). Notably, this effect is comparable to loss of ligase IV. Interestingly, when breaks were introduced by the Cas9 nuclease, absence of ERCC6L2 had a less profound effect on CSR (224). This could imply that ERCC6L2 is required for physiological CSR but is necessary only in certain contexts for global repair of DSBs. As reported in previous studies, it was confirmed that recruitment of ERCC6L2 to sites of DNA damage is mediated through its C-terminus. This recruitment is not dependent on several DDR factors including KU, XLF, NBS1, H2AX or PARP1 (224). Recruitment of KU and XRCC4 to DNA damage, however, is moderately impacted by the absence of ERCC6L2 suggesting that it could promote complex assembly. Further analysis of the role of ERCC6L2 in CSR revealed that DNA end resection and microhomology usage is mildly increased in *ERCC6L2*-deficient cells (224). Usually, the vast majority (>90%) of programmed DSBs are repaired in a manner that results in deletion of the

intervening sequence between the recombining elements. However, intriguingly, loss of ERCC6L2 led to almost an equal ratio of inversion and deletion events at break sites indicating that ERCC6L2 is required for promoting directional repair of DSBs during CSR. Thus, it is proposed that ERCC6L2 may remove obstacles from DNA ends, such as nucleosomes or proteins, which may present a barrier to rapid end joining (224).

Another study from Olivieri and colleagues, similarly, identified ERCC6L2 in a genome-wide CRISPR screen as a c-NHEJ factor (225). *ERCC6L2*-deficient DT40 cells were shown to be sensitive to etoposide and bleomycin in cell proliferation assays. Furthermore, loss of ERCC6L2 causes a significant reduction in c-NHEJ efficiency (225). Following these observations, it was tested whether loss of ERCC6L2 could rescue PARP inhibitor sensitivity in *BRCA1/p53-null* cells, as has been demonstrated for some c-NHEJ promoting factors. Unlike 53BP1 and components of the Shieldin complex, loss of ERCC6L2 did not rescue PARP inhibitor sensitivity in *BRCA1/p53-null* RPE1 cells (225).

Most recently, Francica et al identified ERCC6L2 in a genome-wide screen which sought to uncover genes involved in mediating resistance to IR (226). Interestingly, in contrast to the findings from Olivieri and colleagues, they found that loss of ERCC6L2 does rescue sensitivity to PARP inhibition in *BRCA1/p53-null* mouse cells. Further characterisation revealed that loss of ERCC6L2 can partially restore HR in the absence of BRCA1 and can be reversed by inhibiting ATM kinase (226). Like previous studies, loss of ERCC6L2 was shown to increase levels of end resection and collectively, they suggest a role for ERCC6L2 in 53BP1 dependent c-NHEJ.

In addition to published findings, previous work from our lab has made several key insights into the cellular function of ERCC6L2. Clonogenic survival assays using

ERCC6L2 knockout (KO) U-2-OS cells revealed that loss of ERCC6L2 results in sensitivity to the TOPO II inhibitor etoposide as well as phleomycin (Dr Chris Carnie). Crucially, these phenotypes can be rescued by complementation with full-length ERCC6L2 but not with a catalytic dead mutant. Interestingly, evidence from our lab suggests that *ERCC6L2* knockout cells do not display sensitivity to the TOPO II inhibitor ICRF-193 (Dr Chris Carnie). Unlike etoposide, which prevents the re-ligation step during decatenation, ICRF-193 prevents TOPO II turnover (227). This suggests that the sensitivity to etoposide likely arises from a DSB defect as opposed to topological stress. Recruitment of ERCC6L2 to sites of DNA damage has also been shown to be dependent on the C-terminal half of the protein. Interestingly, when investigating recruitment of different C-terminal fragments, ERCC6L2$^{(1054\text{-}1247)}$, which includes the uncharacterised 'VIGSSK' domain, is rapidly recruited to sites of laser damage.

Another significant observation made by the lab is that ERCC6L2 interacts directly with proliferating cell nuclear antigen (PCNA), a core component of the replication machinery (Dr Chris Carnie). This is mediated through a novel PCNA binding motif in ERCC6L2 termed the atypical PCNA interacting protein (a-PIP) box. However, it was shown that this interaction is not required for damage recruitment of full-length ERCC6L2 and the a-PIP binding mutant is able to rescue etoposide and phleomycin sensitivities in the survival assays (Dr Chris Carnie). This indicates that other regions of ERCC6L2, such as the Tudor and VIGSSK domains, are sufficient for regulating its localisation and function at sites of damage. Nevertheless, the PCNA interaction does imply that ERCC6L2 is present at sites of DNA synthesis.

Perhaps the most intriguing discovery to emerge from our lab is that ERCC6L2 colocalises with centromeres. This finding was made following the observation that full-

length ERCC6L2 forms nuclear foci that do not overlap with PCNA. Like its recruitment to sites of DNA damage, the localisation of ERCC6L2 to centromeres is dependent on its C-terminus (Dr Chris Carnie). Interestingly, distinct portions of the C-terminus, ERCC6L2$^{(1054\text{-}1247)}$ and ERCC6L2$^{(1248\text{-}1561)}$ strongly localise, which could indicate a multifaceted basis of recruitment. It is still unclear how ERCC6L2 is recruited to centromeres or what its role at these unique genomic regions may be.

The aim of my project has been to build on the findings from our lab regarding the physiological function of ERCC6L2, using a mixture of cellular and biochemical approaches. I first aimed to improve understanding of the role of ERCC6L2 in the DDR by investigating its involvement in the DSB response. Next, given that ERCC6L2 interacts with PCNA, I explored the possibility of a role for ERCC6L2 in DNA replication. Finally, using biochemical approaches, I sought to characterise the enigmatic VIGSSK domain and gain more understanding of what role this small, yet important part of ERCC6L2 plays in both a DDR and centromeric context.

2. Materials and methods

Table 7-1: Antibodies

Target	Host	Dilution (WB)	Dilution (IF)	Manufacturer	Code
BRCA1	Mouse	1:500	N/A	Santa-Cruz	sc-6954
BrdU	Mouse	N/A	1:500	GE Healthcare	RPN202
DNA-PKcs	Rabbit	1:500	N/A	Abcam	ab174575
GFP	Mouse	1:2500	N/A	Santa Cruz	sc-9996
H3	Rabbit	1:25000	N/A	Millipore	06-755
His	Rabbit	1:1000	N/A	Abcam	ab9108
KAP1	Rabbit	1:1000	N/A	Bethyl	ABE1859
MSH6	Mouse	1:1000	N/A	BD Biosciences	610918
MSH3	Mouse	1:1000	N/A	BD Biosciences	611390
MSH2	Mouse	1:1000	N/A	BD Biosciences	556349
PMS2	Mouse	1:1000	N/A	Novus Biologicals	163C1251
pATM (s1981)	Rabbit	1:500	N/A	Abcam	ab81292
pCHK1 (s345)	Rabbit	1:500	N/A	Cell Signalling	2348
pCHK2 (t68)	Rabbit	1:500	N/A	Cell Signalling	2661
pKAP1 (s824)	Rabbit	1:1000	N/A	Bethyl	A300-767A
pRPA (s4/8)	Rabbit	1:1000	1:300	Bethyl	A300-245A
RAD52	Mouse	1:500	N/A	Santa Cruz	sc-365341
RPA32	Mouse	1:2000	1:300	Abcam	ab2175
β tubulin	Rabbit	1:10000	N/A	Abcam	ab15568
γH2AX (s139)	Mouse	1:1000	1:300	Millipore	05-636
Myc	Rabbit	1:1000	N/A	Abcam	ab9106
Streptavidin-HRP	Rabbit	1:1000	N/A	Abcam	ab191838
FLAG	Rabbit	1:2500	N/A	Sigma-Aldrich	F3165
Anti-mouse-HRP	Goat	1:2500	N/A	Dako	P0447
Anti-rabbit-HRP	Pig	1:2500	N/A	Dako	P0399
Anti-mouse-Alexa 488	Goat	N/A	1:500	Life Technologies	A28175
Anti-rabbit-Alexa 594	Goat	N/A	1:500	Life Technologies	A11037

Table 7-2: siRNA sequences

Target	Sequence	Supplier
Control	CGUACGCGGAAUACUUCGA(dTdT)	Dharmacon
BRCA1	GGAACCUGUCUCCACAAAG(dTdT)	Dharmacon

Table 7-3: Primers

Primer name	Sequence	Source
CENP-Q-EcoRI-F	GTTAGATTGAATTCATGAGCGGTAAAGCAAATGCC	Invitrogen
CENP-Q-XhoI-R	ACCTCAACTCGAGGCTGGCATCCAGTTTTTTGTAG	Invitrogen
CENP-Q-XhoI-N-67-R	ACCTCAACTCGAGTTATGCTGCGGTTTT	Invitrogen
CENP-Q-EcoRI-68-C-F	GTTAGATTGAATTCAGCAAACGTAAAACCTGGCAG	Invitrogen
VIGSSK K1143A F	CAGAATGTAATGGATCGAGCGCAGCTGAAAATCACATGAGCGC	Invitrogen
VIGSSK K1143A R	GCGCTCATGTGATTTTCAGCTGCGCTCGATCCAATTACATTCTG	Invitrogen
VIGSSK R1150A F	CTGAAAATCACATGAGCGCATGGGCAGCACATGACG	Invitrogen
VIGSSK R1150A R	CGTCATGTGCTGCCCATGCGCTCATGTGATTTTCAG	Invitrogen
VIGSSK K1213A F	TTTACATTTCAAATCCTGTAAACCAGGCGAAGAAAAAAGTCTACCATACAA	Invitrogen
VIGSSK K1213A R	GGTTTGTATGGTAGACTTTTTTCTTCGCCTGGTTTACAGGATTTGAAATGTAAA	Invitrogen
VIGSSK K1214A F	TTTCAAATCCTGTAAACCAGGCGGCGAAAAAGTCTACCATACAAACC	Invitrogen
VIGSSK K1214A R	GGTTTGTATGGTAGACTTTTTTCCCGCCTGGTTTACAGATTTGAAA	Invitrogen
VIGSSK K1215A F	TCCTGTAAACCAGGCGGCGGCAAAAGTCTACCATACAAAC	Invitrogen
VIGSSK K1215A R	GTTTGTATGGTAGACTTTTGCCGCCGCCTGGTTTACAGGA	Invitrogen
VIGSSK VF1156/1157AA F	GAGAAAACTGCTTCAACTCAGCTGCGTCATGTGCTGCCCATCGG	Invitrogen
VIGSSK VF1156/1157AA R	CCGATGGGCAGCACATGACGCAGCTGAGTTGAAGCGTTTTCTC	Invitrogen

Table 7-4: Plasmids

Plasmid	Vector	Source	Use
pGEX4T.1	pGEX4T.1	Dragana Ahel	Empty vector
pET28A	pET28A	Dragana Ahel	Empty vector
myc-BioID2-MCS	myc-BioID2-MCS	Kyle Roux (Addgene #74223) (228)	Empty vector
YFP-NLS	pN-YFP-DEST	Dragana Ahel	Mammalian overexpression
YFP-ERCC6L2 FL	pN-YFP-DEST	Dragana Ahel	Mammalian overexpression
YFP-ERCC6L2 N-term	pN-YFP-DEST	Dragana Ahel	Mammalian overexpression
YFP-ERCC6L2 C-term	pN-YFP-DEST	Dragana Ahel	Mammalian overexpression
YFP-VIGSSK	pN-YFP-DEST	Dragana Ahel	Mammalian overexpression
BioID-NLS	myc-BioID2-MCS	Dragana Ahel	Mammalian overexpression
BioID-VIGSSK	myc-BioID2-MCS	Dragana Ahel	Mammalian overexpression
BioID-VIGSSK V1156A F1157A	myc-BioID2-MCS	Mutagenesis	Mammalian overexpression
6xHis VIGSSK	pET28A	Dragana Ahel	*E. coli* expression and purification
6xHis VIGSSK K1143A R1150A	pET28A	Mutagenesis	*E. coli* expression and purification
6xHis VIGSSK K1213A K1214A K1215A	pET28A	Mutagenesis	*E. coli* expression and purification
GST-CENP-Q	pGEX4T.1	Restriction cloning	*E. coli* expression and purification
GST-CENP-Q N-67	pGEX4T.1	Restriction cloning	*E. coli* expression and purification
GST-CENP-Q 68-C	pGEX4T.1	Restriction cloning	*E. coli* expression and purification
His-MBP-ERCC6L2^{WT}-FLAG	pFastBac	Dragana Ahel	Sf9 virus generation
His-MBP-ERCC6L2^{WT}-FLAG	pFast Bac	Dragana Ahel	Sf9 virus generation

Table 7-5: Drug treatments

Drug name	Company	Stock solution
Etoposide	Abcam	20 mM in DMSO
Phleomycin	Sigma-Aldrich	20 µg/ ml in DMSO
Hydroxyurea	Sigma-Aldrich	0.5 mM in H_2O
Olaparib	Sigma-Aldrich	10 mM in DMSO
VE-821 (ATR inhibitor)	Abcam	10 µM in DMSO

2.1. Cell culture and cell lines

U-2-OS (ATCC HTB-96), HEK293T (ATCC CRL-11268) and RPE1 hTERT (ATCC CRL-4000) cells were cultured in T75 or T175 flasks (Corning). U-2-OS and HEK293T were maintained in DMEM (Sigma) with 10% FBS (Sigma) and 1% Penicillin-Streptomycin (Invitrogen). RPE1 cells were cultured in DMEM/F12 (Thermo) media supplemented with 10% FBS and 1% Penicillin-Streptomycin. Cells were kept in incubators at 37^{o}C and 5% CO_2. As required, cells were passaged or seeded by aspirating existing media and washing with 1X PBS. Cells were detached using TrypLE (Invitrogen) at 37^{o}C. Frozen cell stocks were prepared by detaching cells and collecting by centrifugation in a bench top centrifuge at 1000 rpm for 5 minutes. Cells were washed twice with 1X PBS and resuspended in 90% FBS plus 10% DMSO. Cells were aliquoted into cryogenic vials and immersed in ice in a polystyrene box and stored at -80^{o}C for short-term storage or liquid N_2 for long-term.

ERCC6L2 knockout (*ERCC6L2*^{KO}) cell lines were generated using CRISPR/Cas9 technology by Dr Chris Carnie. Briefly, two single-guide RNAs (sgRNAs) sequences targeting exon 2 and exon 3 of *ERCC6L2*, TATGGACACTACATCCATGGAGG and GCATAAAAAGGGAACTCGTGAGG, respectively (designed and tested by Dr. Andrew Bassett, formerly of Genome Engineering Oxford), were cloned into px459 (a gift from Feng Zhang, Addgene plasmid #48139). px459 constructs were transfected

into U-2-OS and hTERT-RPE1 cells and selected with puromycin (2 µg/ml) for 24h. Clonal cell lines were established from single cells and validated by sequencing. Due to the lack of an available antibody for ERCC6L2, the knockout cells were routinely checked by sequencing using *ERCC6L2* sequencing primers.

2.2. Cell cycle analysis by flow cytometry

Cells were seeded in 10 cm dishes and treated with 200 nM etoposide (Abcam) or 10 µg/ml phleomycin (Sigma) the following day for 48 hours. Cells were prepared for analysis using the BD Cycletest^{TM} Plus DNA kit. Sample preparation was conducted in a tissue culture hood. After two days of treatment, cells were collected and washed three times with Buffer Solution (PBS containing sodium citrate, sucrose and DMSO). Cells were counted using a haemocytometer and resuspended to give a final concentration of 1 x 10^6 cells/ml. For analysis, 5 x 10^5 cells/ml were used per sample. Cells were trypsinized and treated with RNase at room temperature. Cells were subsequently treated with ice-cold propidium iodide and transferred to 4^{o}C. Samples were filtered using a 40 µM cell strainer (Corning). Samples were acquired on the flow cytometer on the same day. Cell cycle analysis was performed on a Cytek DxP8 using the 488 nM laser. Gating was performed on live cells to exclude debris and dead cells (FSC/SSC plot). The cells from the first gate were plotted on a FSC/FSCW plot where another gate was applied to exclude doublets. This population of cells were plotted as a histogram using the BluFL2 channel. At least 30,000 events within this population were recorded.

2.3. Preparation of whole cell extracts

Following appropriate treatment, cells were harvested in trypsin and washed once with cold 1X PBS. Cell pellets were resuspended in warm 1% SDS lysis buffer (1% SDS, 10 mM Tris-HCl pH 8.5, 2 mM $MgCl_2$, protease inhibitor (PI) tablet (Roche)) and boiled for 3 minutes at 95^{o}C then allowed to cool to room temperature. Benzonase (Sigma) was then added to the lysates and left for at least 1 hour. Lysates were boiled for a further 5 minutes at 95^{o}C and then spun at 15,000 rpm for 10 minutes. Protein concentrations were determined by Bradford assay and samples were prepared for SDS-PAGE in 4X LDS buffer.

2.4. Western blotting

For electrophoresis, 1 mm 4-12% Bis-Tris Novex (Invitrogen) gels were run in a NuPAGE Novex mini or midi gel tank. Gels were run in either 1X NuPAGE MES or MOPS running buffer (Invitrogen), as appropriate. Wells were rinsed with running buffer using a syringe prior to loading. SeeBlue Plus2 Protein Standard (Life Technologies) was used as a molecular weight marker. Generally, gels were run at 150 V for 1 hour.

After SDS-PAGE, proteins were transferred to nitrocellulose membrane from a pre-prepared Transfer Pack (BioRad) using the semi-dry Trans-Blot Turbo Transfer system (BioRad). The Standard SD pre-programmed protocol was used to transfer proteins (30 minutes, 25 V; up to 1 A). Once complete, the membrane was moved to a plastic container and soaked briefly in Ponceau (Invitrogen) to check transfer efficiency. Membranes were then placed in a new plastic container with 5% milk (Marvel) in PBS-T (PBS + 0.1% Tween-20 (Sigma)). Membranes were blocked for at least 30 minutes

at RT with agitation. Primary antibodies were incubated in 5% BSA (Sigma) in PBS-T or 5% milk in PBS-T either for 1 hour at RT or overnight at 4°C with agitation. Following primary antibody incubation, membranes were washed 3 times in PBS-T for 5 minutes each. HRP-conjugated secondary antibodies were incubated in 5% milk in PBS-T for 1 hour at RT with agitation. Membranes were washed a further 3 times with PBS-T before detection.

Membranes were then transferred to a plastic folder stuck inside an Amersham Hypercasette (Fisher Scientific). Once the membrane was dry, a 1:1 mix of enhanced chemical luminescence (ECL) substrate (Pierce, Invitrogen) was added evenly to the membrane and left for 1 minute. The ECL substrate was removed and the plastic folder was folded over and secured with tape. In the darkroom, X-ray film (Fisher Scientific) was exposed to the membrane for various lengths of time, as appropriate. Films were then developed (Xograph).

2.5. Immunofluorescence

U-2-OS cells on glass coverslips were treated with stated concentrations of drugs. Cells were pre-extracted on ice for 3 minutes using pre-chilled 0.2% Triton-X. Cells were fixed in 4% formaldehyde for 10 minutes at room temperature. Cells were washed three times with 1X PBS and permeabilized for 5 minutes in PBS containing 0.5% Triton-X (Sigma). Cells were washed again with 1X PBS three times and blocked in 2% BSA for 30 minutes at room temperature. All antibodies were diluted in 2% BSA. Primary antibody incubations were performed for 1 hour at room temperature. Cells were washed three times with 1X PBS. Secondary antibody incubations were performed for 1 hour at room temperature. Coverslips were then washed three times with 1X PBS and incubated with DAPI (0.5 µg/ml) to stain DNA. Coverslips were

mounted onto slides using moviol. Cells were imaged on an Olympus BX61 microscope and a CoolSNAP HQ2 camera. Acquisition times and settings were adjusted in non-saturated conditions in the 12-bit dynamic range. Identical settings were applied to all images in one experiment.

2.6. Image analysis

Acquired images were processed and analysed using Cell Profiler software. The DAPI signal was used to generate a mask to define individual nuclei. To exclude unwanted objects, nuclei were also identified based on diameter in pixel units. For mean and total intensities, individual nuclei were defined as before. Mean and total intensities were measured per nucleus for the relevant channel. The data generated by Cell Profiler was exported to GraphPad Prism 8 to generate the relevant plots and statistically analyse the data.

2.7. Clonogenic survival assay

For clonogenic survival assays with BRCA1 depletion and olaparib treatment, U-2-OS cells were seeded from 10 cm dishes treated with control and BRCA1 siRNAs for 48 hours.

For siRNA depletion, 50 nM of siRNA and 20 µl RNAiMAX (Invitrogen) was used per plate. siRNA and RNAiMAX were incubated individually in 500 µl of OptiMEM (Thermo) for 5 minutes at room temperature. Mixtures were then combined, gently mixed and incubated for a further 15 minutes. The transfection mixture was then added dropwise to the cells. Approximately 6 hours later, transfection mixtures were removed from the cells and replaced with fresh media. siRNA depletion was repeated the following day.

Cells were seeded for the clonogenic survival assay 48 hours after the first transfection. 6000 cells were seeded per 6 cm dish in 5 ml of media. Cells were treated with the

appropriate concentration of olaparib (Sigma) the following day by replacing the existing media with drug containing media. Cells were fixed and stained after 15 days.

For clonogenic survival assays with hydroxyurea, U-2-OS and hTERT-RPE1 cells were seeded from confluent T75 flasks (Corning). 1500 cells were seeded per well of a 6-well plate in 3 ml of media. Cells were treated with the appropriate concentration of hydroxyurea (Sigma) the following day, by replacing the existing media with drug containing media. Cells were fixed and stained after 10 days. Each experiment was performed in technical triplicates, and data presented is from two independent repeats.

Once the cells were ready to be fixed, the media was removed, and cells were washed once with 1X PBS. 5 mg/ml crystal violet (Sigma) in 25% methanol was added to the plates and incubated for 30 minutes at room temperature. After 30 minutes, the crystal violet solution was removed, and plates were washed in ddH_2O. Plates were photographed using a Nikon D3200 DSLR camera.

Using ImageJ software, a macro was written to ensure consistent thresholding between each image. JPEG images were converted to 8-bit and thresholding adjusted to identify cell growth and minimise background. The ‘analyse particles’ function was used to quantify the area covered in each well/dish. Total well/dish area coverage was used for data analysis. Averages of surface area was calculated for the technical triplicates and clonogenic survival was calculated as a percentage of the untreated cells, individually for each cell line. To statistically analyse the data, GraphPad Prism 8 software was used to perform area under the curve (AUC) analysis for each biological replicate (each technical repeat for olaparib). AUC analysis was conducted to produce one value representing each experiment across the concentration range allowing statistical comparison between repeats, not just at individual concentrations. Average

AUC values for each cell line are presented as bar charts with SEM. Unpaired T-tests were performed on the means and p values represented as follows $P \leq 0.05$ = *, P< 0.005 = **, P< 0.0005 = ***.

2.8. Subcellular protein fractionation

U-2-OS cells were grown in 10 cm dishes and treated with 2 mM hydroxyurea (Sigma) for 24 hours. Cells were harvested in trypsin and washed once with cold 1X PBS. Subcellular protein fractionation was carried out using Subcellular Protein Fractionation Kit for Cultured Cells kit (Thermo). Briefly, cold cytoplasmic extraction buffer containing PIs was added to the cell pellets and incubated with mixing for 10 minutes at 4^{o}C. Samples were centrifuged for 5 minutes at 500 x g. The supernatant was removed and membrane extraction buffer with PIs was added to the pellet and vortexed for 5 seconds. The samples were incubated with mixing for 10 minutes at 4^{o}C. Samples were then centrifuged for 5 minutes at 3000 x g. The supernatants were removed and cold nuclear extraction buffer containing PIs was added to the pellets and vortexed for 15 seconds. Samples were incubated for 30 minutes with mixing at 4^{o}C. Samples were then centrifuged for 5 minutes at 5000 x g. The supernatants (soluble nuclear extracts) were carefully transferred to a pre-chilled tube and kept at 4^{o}C. Room temperature nuclear extraction buffer containing PIs, micrococcal nuclease and calcium chloride was added to the pellets. Samples were vortexed for 15 seconds and incubated at 37^{o}C for 10 minutes. After incubation, samples were vortexed for another 15 seconds and then centrifuged for 5 minutes at 16000 x g. The supernatants (chromatin bound nuclear extracts) were carefully transferred to a pre-chilled tube and kept at 4^{o}C. Soluble and chromatin bound nuclear extracts were prepared for western analysis in 4X LDS buffer. Western blotting was carried out as previously described.

2.9. Immunoprecipitation

HEK293T cells were grown to 70% confluence in 7 x 15 cm dishes (Corning) per construct. Cells were transfected with the relevant constructs in pN-YFP-DEST. 10 µg of DNA and 30 µl Lipofectamine 2000 was used per 15 cm plate. The DNA and lipofectamine 2000 were each incubated in 2 ml of serum free DMEM. After 5 minutes, the lipofectamine containing media was added dropwise to the DNA containing media, gently mixed and incubated for 15 minutes at room temperature. After incubation, the transfection mixture was added dropwise to the cells. 48 hours later, cells were detached mechanically, and the appropriate samples combined. Cells were centrifuged for 5 minutes at 500 x g and washed twice with 1X PBS. Cell pellets were placed on ice and lysed in 4 ml of cold lysis buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 0.5% Triton X-100, 5 mM $MgCl_2$ supplemented with a PI tablet (Roche) and 500U/ µl of benzonase). Lysates were sonicated at high amplitude 5 x 30 seconds and put on ice between rounds. Lysates were then placed on a rotor wheel at 15 rpm for 1 hour at 4^{o}C. Samples were then centrifuged at full speed at 4^{o}C for 20 minutes after which the supernatants were collected and transferred to new pre-chilled 15 ml Falcon tubes. 75 µl of input was collected and added to 25 µl of 4X LDS.

During centrifugation, GFP-Trap magnetic agarose beads (Chromotek) were equilibrated in lysis buffer (without PIs or benzonase). 100 µl of beads were used per condition. The appropriate volume of beads (~500 µl) was transferred to a microcentrifuge tube and separated from the solution using a magnetic rack. The solution was removed, and the beads were immediately resuspended in 1 ml of cold lysis buffer. The beads were then separated using the magnetic rack and the buffer removed. This was repeated twice more. The beads were finally resuspended in 500

µl of lysis buffer and 100 µl of beads was added to each sample. Samples were incubated with the beads overnight at 4°C on a rotor wheel (10 rpm).

The next day, beads were separated from the solution using a magnetic rack as before. 75 µl of the unbound sample was collected and added to 25 µl of 4X LDS. Beads were washed with lysis buffer 10 times. The beads were finally resuspended in 100 µl of 1X LDS and boiled for 5 minutes. Samples were placed on a magnetic rack and eluates were removed and transferred to a fresh tube. 5 µl of input and unbound samples and 20 µl of eluates were analysed by SDS-PAGE and western blotting as described previously.

2.10. Expression and purification of VIGSSK fragment

pET28A-ERCC6L2$^{(1054\text{-}1247)}$ was obtained from Dr. Dragana Ahel. Site directed mutagenesis was performed using a QuickChange Lightning Site-Directed mutagenesis kit (Agilent), according to the manufacturer's protocol. Mutagenesis primers were designed using the online Agilent QuickChange Primer Design tool. For transformation into *E. coli* (XL-Gold competent cells), see Section 2.13.

50 ng of plasmid DNA was used to transform Rosetta (DE3) competent *E. coli*. A single colony was used to inoculate 300 ml of LB supplemented with 30 µg/ml kanamycin and grown overnight at 37°C. Overnight culture was used to inoculate 3 L of fresh LB supplemented with 50 µg/ml kanamycin the next day. Cultures were grown in a shaking incubator at 37°C until OD_{600} = 0.6. Expression was induced with 1 mM IPTG (Sigma) for 4 hours at 30°C. Cells were harvested by centrifugation at 4500 rpm for 20 minutes at 4°C in a Beckman Avanti J-26S XP centrifuge and JLA 8.1 rotor.

Cells were lysed on ice with 50 ml lysis buffer containing 50 mM Tris-HCl pH 8, 500 mM NaCl, 10 mM imidazole, 1 mM DTT and 1 PI tablet (Roche). After pellets were thoroughly resuspended, 25 µl of benzonase (Sigma) was added. Lysates were left on rotor wheel rotating at 10 rpm for 20 minutes at 4^{o}C. Lysates were then sonicated 6 x 30 seconds with 30 seconds rests on ice between each round. Following sonication, samples were passed through a homogeniser 5 times at 15000 psi at 4^{o}C. The homogeniser was equilibrated in H_2O then lysis buffer before the sample was passed through. Lysates were clarified by centrifugation in a Beckman Avanti J-26S XP centrifuge and JA-17 rotor at 15000 rpm for 1 hour at 4^{o}C. Meanwhile, 1 ml of Ni-NTA beads (Qiagen) were washed with H_2O and equilibrated in Ni-NTA buffer (lysis buffer without PI tablet) in an Econo-Pac chromatography column (BioRad). Clarified lysate was applied to the column then beads were washed with 5 x 10 ml of Ni-NTA buffer Beads were washed once with 10 ml of Ni-NTA buffer containing 30 mM imidazole. The protein was eluted from the beads with 3 x 1 ml Ni-NTA buffer plus 500 mM imidazole. 30 µl of each fraction was mixed with 10 µl of 4X LDS and boiled at 95^{o}C for 5 minutes and assessed by SDS-PAGE. Fractions were pooled and concentrated then loaded onto a Superdex S-200 (16/600) column (GE Healthcare). The column was equilibrated in 2 column volumes (CV) of gel filtration (GF) buffer (25 mM Tris-HCl pH 8, 500 mM NaCl, 1 mM DTT). Protein was eluted isocratically in 1.5 CVs. Fraction collection started after 0.2 CVs and 1.8 ml fractions were collected. Peak fractions were assessed by SDS-PAGE as described previously. Fractions were pooled and concentrated using a 10 kDa concentrator that had been equilibrated in H_2O and GF buffer. Concentration of protein was determined by Bradford assay. Glycerol was added to protein at a final concentration of 10%. Protein was aliquoted, flash frozen in liquid N_2 and stored at -80^{o}C.

2.11. Preparation of radiolabelled substrates

200 nM of single-stranded DNA (5'-CTTCGTTGGAAACGGGATT TCTTCATTTCATGCTA-3') was labelled using [γ-^{32}P] ATP (6000 Ci/mmol, Easytides Perkin-Elmer) and T4 polynucleotide kinase (PNK) (NEB) in 10X PNK buffer. Reactions were incubated at 37^{o}C for 1 hour and quenched with 50 mM EDTA. Labelled oligonucleotide was purified over a G25 spin column (PD Spintrap, GE Healthcare) to remove unincorporated radioactivity. Double-stranded DNA substrate was prepared by mixing a 1.2-fold excess of the complementary unlabelled DNA oligonucleotide (5'-TAGCATGAAATGAAGAAATCCCGTTTCCAACGAAG-3') with the ^{32}P- labelled oligonucleotide. Annealing was performed in 10 mM Tris-HCl pH 7.5, 50 mM NaCl. Reactions were boiled at 95^{o}C for 5 minutes then allowed to cool slowly to room temperature.

2.12. Electrophoretic mobility shift assay

A range of concentrations of purified His-tagged VIGSSK fragment (ERCC6L2$^{(1054-1247)}$) (3 μM - 68 μM) were incubated with 2.5 nM of ^{32}P- labelled substrates in 10 μl of reaction buffer (50 mM Tris-HCl pH7.5, 5% glycerol, 5 mM EDTA, 1 mM DTT, 5% glycerol, 0.1 mg/ml BSA) for 30 minutes at room temperature. Reactions were resolved on 5% native polyacrylamide gels in 0.5X TBE buffer (45 mM Tris-borate, 1 mM EDTA; pH 7.5) at 4^{o}C at 100V. Gels were pre-run for 10 minutes. Resolved gels were removed from their case and placed on Whatman paper and wrapped in cling film. Gels were dried for 2 hours in a GelAir drying system (BioRad) and visualised by autoradiography.

2.13. BioID

HEK293T cells were grown to 70% confluence in 6 x 15 cm dishes per condition. Cells were transfected with myc-BioID2-MCS containing NLS and the VIGSSK fragment variants. For each dish, 10 µg of DNA was added to 2 ml of OptiMEM and 30 µl of Lipofectamine 2000 to another 2 ml of OptiMEM. Mixtures were left for 5 minutes at RT then combined and left for a further 20 minutes at RT. Meanwhile, 4 ml of media was removed from each plate. Transfection mixtures were added dropwise. 24 hours after transfection, the media containing transfection mixtures was removed and replaced with DMEM (10% FBS, P/S) containing 50 µM biotin (Sigma). 24 hours later cells were collected by aspirating existing media and detaching cells with 5 ml of 1X PBS containing 2 mM EDTA. Cell suspensions with the same conditions were combined a transferred to a falcon tube. Cells were centrifuged at 1500 rpm for 5 minutes in a bench top centrifuge. Supernatant was discarded and pellets washed by resuspending in 20 ml of 1X PBS followed by centrifugation. This wash step was repeated another two times. After the last wash, the supernatant was removed, and pellets were either stored at -80^{o}C or placed on ice and lysed.

Note all buffers from here on were made with HPLC grade water. LoBind Eppendorf tubes and filter tips were also used. Cells were lysed in 2 ml of cold hypotonic lysis buffer (10 mM HEPES pH 7.9, 60 mM KCl, 1.5 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, 0.075% NP40, 1 PI tablet (Roche), 1 mM PMSF). Lysates were left on a roller for 10 minutes at 4^{o}C. Lysates were then transferred to Eppendorf tubes (2 x 1 ml per sample). Samples were spun at 1200 rpm for 10 minutes at 4^{o}C and supernatants were discarded. Pellets were washed with 3 x 800 µl hypotonic lysis buffer without NP40. Nuclear pellets were resuspended in 6 ml (1 ml per dish) of nuclear extraction buffer

(50 mM Tris-HCl pH 7.4, 500 mM NaCl, 0.4% SDS, 5 mM EDTA, 1 mM DTT, 1 PI tablet (Roche), 2% Triton X-100, 1 mM PMSF). 1 ml was used to resuspend the pellet then transferred to 5 ml in a 15 ml falcon tube. Nuclear extracts were left on a roller for 30 minutes at 4^{o}C. 6 ml (equal volume) of cold 50 mM Tris-HCl pH 7.4 was added to extracts and vortexed briefly. Samples were sonicated for 10 seconds at 10% amplitude. 10 µl of benzonase was then added to each sample and left to digest for at least 1 hour on a roller at 4^{o}C. Samples were spun at 15000 rpm for 30 minutes at 4^{o}C. Meanwhile, 600 µl of Streptavidin magnetic dynabeads (Invitrogen) were prepared. A magnetic rack was used to separate beads from solution. The solution was removed using double tips and 1 ml of 50 mM Tris-HCl pH 7.4 was added immediately. This was repeated twice more. Beads were finally resuspended in 600 µl of 50 mM Tris-HCl pH 7.4. After centrifugation, supernatants were transferred to fresh tubes and 200 µl of beads were added to each sample. Samples were left incubating with beads overnight at 4^{o}C on a roller.

The next day, beads were separated from solution. Unbound sample was kept for testing. Beads were resuspended in 1 ml of wash buffer 1 (2% SDS) and transferred to an Eppendorf. Samples were left on rotor wheel for 8 minutes at RT. Note each wash step was performed for 8 minutes at RT on an orbital rotor. Wash buffers were removed by pipetting. Beads were washed once more with wash buffer 1. Beads were washed once with wash buffer 2 (50 mM HEPES pH 7.5, 500 mM NaCl, 0.1% deoxycholate, 1% Triton X-100, 1 mM DTT). Beads were washed once with wash buffer 3 (10 mM Tris-HCl pH 8, 250 mM LiCl, 0.5% NP-40, 0.5% deoxycholate, 1 mM EDTA). Beads were washed 4 times with wash buffer 4 (50 mM Tris-HCl pH 7.4, 50 mM NaCl). After the final wash, beads were resuspended in 550 µl of wash buffer 4. 60 µl of beads was combined with 20 µl 4X LDS, boiled for 5 minutes at 95^{o}C and

analysed by SDS-PAGE and Myc and streptavidin western blotting. The rest of the sample was analysed by liquid chromatography with tandem mass spectrometry (LC-MS/MS) following on bead tryptic digestion.

2.14. Expression and purification of GST-CENP-Q variants

Full-length and truncated CENP-Q were cloned into pGEX4T.1 by restriction cloning. Wildtype CENP-Q DNA sequence containing N-terminal EcoRI and C-terminal XhoI restriction sites was obtained using the GeneArt Synthesis service from Invitrogen. This construct was used as a template for polymerase chain reaction (PCR) to amplify the wildtype sequence and generate the truncated CENP-Q variants. PCR mixes contained 35 µl of H_2O, 10 µl 5X High fidelity buffer (NEB),1 µl (~30 ng) template DNA, 1 µl forward primer at 10 µM, 1 µl reverse primer at 10 µM, 1 µl dNTP (deoxyribose nucleotide triphosphate) mix, 1 µl $MgCl_2$ at 50 mM, and 1 µl Phusion High Fidelity Polymerase (NEB). The PCR reaction was carried out in a thermal cycler (Life Technologies) as follows: 1X 95°C for 3 minutes, 30X (95°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute per 1 kb) 1X 72°C for 10 minutes. PCR products were mixed with 6X loading dye (NEB) and analysed on a 1% agarose (Sigma) gel containing 0.0005% Ethidium bromide (Sigma) in 1X TBE (45 mM Tris-Borate (Trizma and Boric acid (Sigma)), 1 mM EDTA (Sigma)) buffer. Electrophoresis was performed at 125V for 30 minutes and DNA was visualised with a UV transilluminator (UVP). Bands were excised and DNA extracted using the Qiagen Gel Extraction Kit, according to the manufacturer's protocol. The PCR products were then digested with EcoRI (NEB) and XhoI (NEB) restriction enzymes by incubating the following mixture at 37°C for 1 hour: 20 µl of PCR product, 2.5 µl CutSmart buffer (NEB), 1 µl of each restriction enzyme. pGEX4.1 was also digested with EcoRI and XhoI by incubating the following

mixture at 37°C for 1 hour: 10 µl H_2O, 1 µg DNA, 2 µl CutSmart buffer, 1 µl of each enzyme. A QIAquick PCR Purification kit (Qiagen) was used to purify the digested PCR product. The digested pGEX4T.1 vector was run on a 1% agarose gel and the band corresponding to the digested vector was excised and gel extraction was performed. Ligation was performed by combining 150 ng of digested PCR product, 50 ng of digested vector, 5 µl 2X Quick Ligase buffer (NEB), 1.5 µl Quick Ligase (NEB). Reactions were left at room temperature for 2 hours. After ligation, 5 µl of the ligation reaction was used to transform 30 µl of XL-Gold competent *E. coli*. Cells were left to incubate with ligation mixtures for 15 minutes on ice. The transformation mixture was then heat shocked at 42°C for 45 seconds and returned to ice for 2 minutes. 300 µl of S.O.C medium (Invitrogen) was added to the mixtures which were then placed in a shaking incubator at 37°C for 1 hour. After incubation, cells were centrifuged at 2000 rpm for 3 minutes. 200 µl of supernatant was removed and the cells were gently resuspended in the remaining media and spread onto LB agar plates containing 100 µg/ml ampicillin. Plates were incubated overnight at 37°C. The following day, a single colony was used to inoculate 3 ml of LB supplemented with 100 µg/ml of ampicillin and grown overnight at 37°C in a shaking incubator. Plasmid DNA was purified the following morning using a QIAPrep Spin Miniprep Kit (Qiagen). Samples were then submitted to Source Bioscience for sequencing using pGEX 5' and 3' primers.

50 ng of pGEX4T.1- CENP-Q was used to transform 30 µl of competent Rosetta *E. coli* as previously described. A single colony was used to inoculate 300 ml of LB supplemented with 100 µg/ml ampicillin and grown overnight at 37°C. The overnight culture was used to inoculate 3 L of fresh LB supplemented with 100 µg/ml ampicillin the next day. Cultures were grown in a shaking incubator at 37°C until OD_{600} = 0.6. Expression was induced with 1 mM IPTG (Sigma) for 4 hours at 30°C. Cells were

harvested by centrifugation at 4500 rpm for 20 minutes at 4°C in a Beckman Avanti J-26S XP centrifuge and JLA 8.1 rotor. Pellets were stored at -80°C until lysis.

Cells were lysed on ice with 50 ml lysis buffer containing 1X PBS, 1 mM DTT and 1 PI tablet (Roche). Lysozyme (Sigma) was then added to a final concentration of 1 mg/ml. After pellets were thoroughly resuspended, 25 µl of benzonase (Sigma) was added. Lysates were left on rotor wheel rotating at 10 rpm for 20 minutes at 4°C. Lysates were then sonicated 5 x 20 seconds with 30 seconds rests on ice between each round. Following sonication, samples were passed through a homogeniser 5 times at 15000 psi at 4°C. The homogeniser was equilibrated in H_2O then lysis buffer before the sample was passed through. Lysates were clarified by centrifugation in a Beckman Avanti J-26S XP centrifuge and JA-17 rotor at 15000 rpm for 1 hour at 4°C. Meanwhile, 1 ml of glutathione sepharose beads (Invitrogen) were washed with H_2O and equilibrated in GST buffer (lysis buffer without PI tablet) in an Econo-Pac chromatography column (BioRad). Clarified lysate was applied to the column then beads were washed with 5 x 5 ml of GST buffer. After washing, the column was plugged and 1 ml of GST buffer was used to resuspend the beads. 75 µl of beads was mixed with 25 µl of 4X LDS and boiled at 95°C for 5-10 minutes, spun for 1 minute at 13000 rpm and assessed by SDS-PAGE.

2.15. GST-CENP-Q pulldown

Approximately 100 µl of GST-CENP-Q bound glutathione sepharose beads (Invitrogen) were transferred to Mini Bio-Spin chromatography columns (BioRad). Beads were washed in 1 ml binding buffer (50 mM Tris-HCl pH 8, 1 mM DTT, 50/150/300/500 mM NaCl). Columns were plugged and 200 µl of binding buffer containing His-VIGSSK fragment was added to the beads. Columns were capped and

left rotating for 2 hours at 4°C. Columns were unplugged and uncapped and flow-through was collected. Beads were washed 5 x 1 ml binding buffer. Columns were plugged and beads resuspended in 1 ml binding buffer then transferred to Eppendorfs. Beads were spun at 500 x g for 1 minute and supernatant was removed. Beads were resuspended in 100 µl 1X LDS and boiled for 5–10 minutes then spun for 1 minute at 13000 rpm and analysed by SDS-PAGE and anti-His western blotting.

2.16. Expression and purification of ERCC6L2

Wild-type and mutant ERCC6L2 were expressed and purified using the baculovirus expression system. pFastBac-ERCC6L2^{WT} and ERCC6L2^{K165A} were supplied by Dr Dragana Ahel. To generate viruses, Sf9 cells were transfected with pFastBac-ERCC6L2^{WT} and ERCC6L2^{K165A} to generate P0 virus. P0 virus was then used to infect Sf9 cells to obtain the P1 virus. Hi5 cells were infected with P1 viruses and harvested 48 hours post infection. Cells were resuspended in 10 mL cold lysis buffer (50 mM Tris-HCl 7.5, 2 M NaCl, 1 mM DDT, 10 % glycerol, 0.04% Triton-X, PI tablet (Roche), benzonase) at 4°C and clarified by centrifugation in a Beckman Avanti J-26S XP centrifuge and JA-17 rotor at 15000 rpm for 1 hour at 4°C. Clarified lysates were incubated with anti-FLAG M2 magnetic beads (Sigma) which were equilibrated in 50 mM Tris-HCl 7.5, 2 M NaCl and 1 mM DTT. Beads were extensively washed with the same buffer. Proteins were eluted with 50 mM Tris-HCl 7.5, 2 M NaCl, 2.5 mg/ml 3X FLAG peptide (Sigma). Fractions were analysed by SDS-PAGE and anti-FLAG western blotting.

3. Investigating the role of ERCC6L2 in the double-strand break response

3.1. Results

Previous studies, as well as work from our group, have shown that loss of ERCC6L2 results in sensitivity to DSB inducing agents (217,220). For example, in clonogenic survival assays, *ERCC6L2*-deficient cells show marked sensitivity to the TOPO II inhibitor etoposide and the radiomimetic phleomycin (Dr Chris Carnie) as well as IR (217). However, at the beginning of this study, the cellular role of ERCC6L2 was poorly understood and it was unclear as to why its absence causes such phenotypes and ultimately, how this contributes to *ERCC6L2*-associated IBMF.

I was keen to investigate the molecular basis of these sensitivities and firstly, aimed to assess cell cycle dynamics of WT and *ERCC6L2*^{KO} U-2-OS cell lines (generated by Dr Chris Carnie) following prolonged etoposide and phleomycin treatment. Cells were treated with either 200 nM etoposide or 10 µg/ml phleomycin for 48 hours, then fixed and stained with propidium iodide to measure DNA content. Cells were then analysed by flow cytometry.

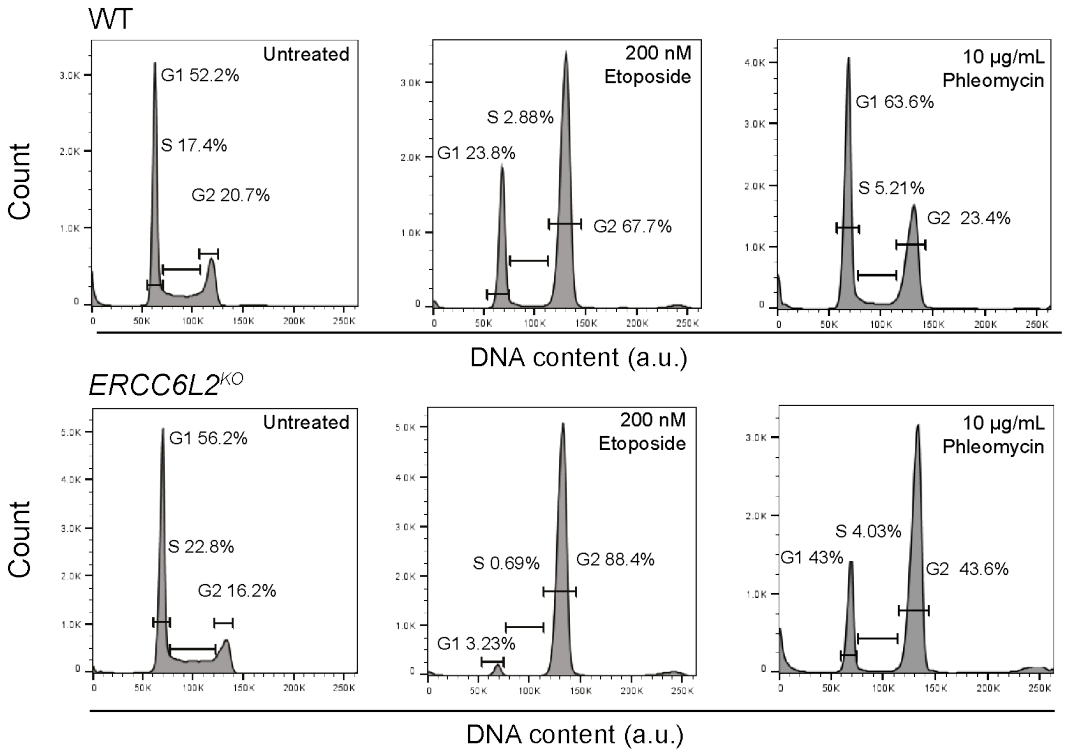


Figure 3-1: Cell cycle analysis of WT and *ERCC6L2*^{KO} cells

Cell cycle profiles of WT and *ERCC6L2*^{KO} U-2-OS cells treated for 48 hours with the stated doses of etoposide and phleomycin and stained with propidium iodide. Histograms are derived from flow cytometry analysis on cell populations after gating to exclude dead cells and doublets.

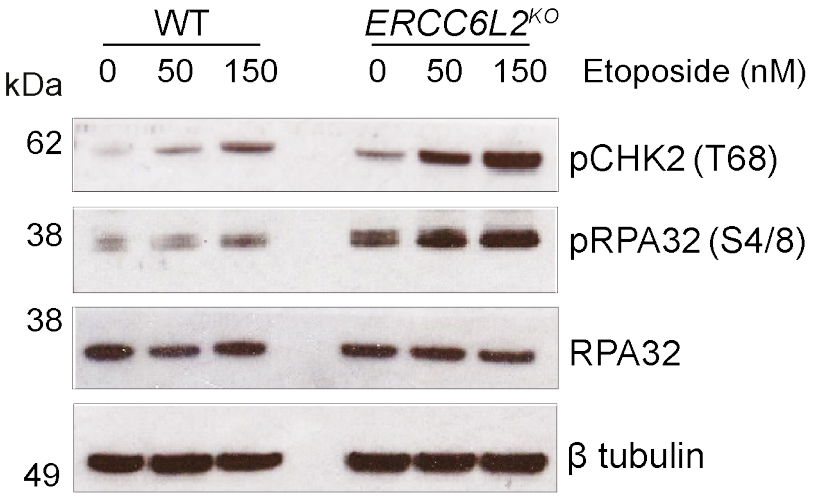
As shown in Figure 3-1, WT and *ERCC6L2*^{KO} cells exhibit similar cell cycle distributions in untreated conditions. Following treatment with etoposide and phleomycin however, a higher percentage of *ERCC6L2*^{KO} cells is observed in G2 phase. With etoposide treatment, an increased proportion of WT and *ERCC6L2*^{KO} cells accumulate in G2 phase, compared to untreated cells. However, this effect is markedly more dramatic in *ERCC6L2*^{KO} cells, with 88.4% of the population in G2 phase compared to 67.7% in the WT cell line. After phleomycin treatment, while the G2 population of the WT cells does not change significantly, the population of *ERCC6L2*^{KO} G2 cells increases almost 3-fold to 43.6% from 16.2%. It should be noted that comparable distributions were observed in two independent experiments. I speculated that a higher instance of G2

arrest reflects an accumulation of unresolved DNA damage due to the absence of ERCC6L2 and leads to activation of the G2/M checkpoint.

Following these observations, I next sought to examine active markers of DDR and checkpoint signalling in WT and *ERCC6L2*^{KO} cells. Whole cell extracts from WT and *ERCC6L2*^{KO} cells were prepared following 24 hours of etoposide and phleomycin treatment and levels of damage signalling proteins were assessed by western blotting (Figure 3-2).

The PI3K-like kinase ATM is central to DNA damage signal transduction in response to DSB formation. ATM is recruited to DSBs by the MRN complex where it undergoes autophosphorylation (serine 1981) and phosphorylates its effector checkpoint kinase CHK2 (threonine 68) which induces cell cycle arrest (29). DNA end resection at DSBs generates ssDNA, which is bound and protected by RPA. In turn, RPA is phosphorylated on multiple sites by the three apical DDR kinases (229). Thus, certain phospho-forms of RPA (pRPA) can be used as a readout for ssDNA formation that could be reflective of increased DNA end resection (230).

A



B

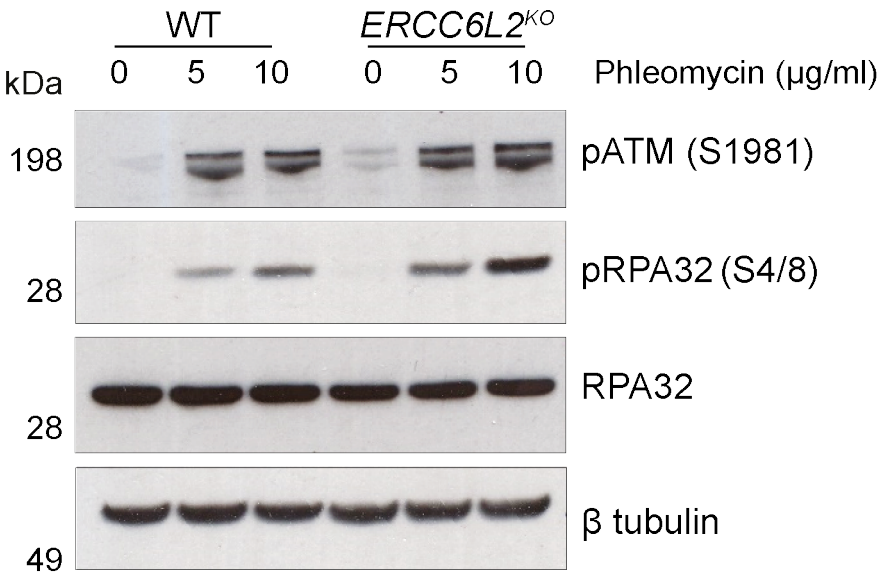


Figure 3-2: Analysis of DDR activation in WT and *ERCC6L2*^{KO} cells

A. WT and *ERCC6L2*^{KO} U-2-OS cells were treated with the indicated doses of etoposide for 24 hours. Whole cell extracts were analysed by western blotting with the indicated antibodies.

B. As in A, for phleomycin.

As shown in Figure 3-2A, ATM activation is signified through phosphorylation of CHK2 (t68) under etoposide treatment and is observed in both WT and *ERCC6L2*^{KO} U-2-OS cell lines. Levels of pCHK2 (t68) increase in response to etoposide treatment in a dose dependent manner but appear to be elevated in the *ERCC6L2*^{KO} cells. RPA phosphorylation is also detected in both cell lines in response to DNA damage but levels of pRPA (s4/8) are notably higher in *ERCC6L2*^{KO} cells.

Following phleomycin treatment, ATM activation is also observed in both cell lines through direct detection of ATM phosphorylation (s1981) (Figure 3-2B). Again, RPA phosphorylation (s4/8) is detected in response to phleomycin treatment in WT and *ERCC6L2*^{KO} cells. However, similarly to the etoposide treatments, phleomycin causes increased levels of pRPA (s4/8) in *ERCC6L2*^{KO} cells.

Together these results show that loss of ERCC6L2 does not appear to significantly affect the main damage signalling response to DSBs and, when considered with the cell cycle analysis in Figure 3-1, *ERCC6L2*^{KO} cells appear to possess a functional G2/M checkpoint. The observation that levels of pRPA (s4/8) are increased under prolonged treatment with both phleomycin and etoposide, suggests that DNA damage does accumulate in *ERCC6L2*^{KO} cells and hints towards a DSB repair defect.

Previous work from our group demonstrated that ERCC6L2 colocalises with centromeres during interphase suggesting that it may be important for maintaining genetic stability at these regions (Dr Chris Carnie). Thus, I was keen to investigate if the increase in DNA damage is related specifically to centromeres. It has been demonstrated that DSBs that occur within heterochromatin take over 8 hours to repair and depend on ATM kinase (231). Comparatively, DSBs that occur in less complex chromatin environments take significantly less time to repair. Initially, I measured

general levels of DNA damage following short (30 minutes) and long (8 hour) releases from phleomycin treatment to assess a potential role for ERCC6L2 in recovery from DNA damage. WT and *ERCC6L2*^{KO} cells were exposed to phleomycin for 30 minutes then allowed to recover in fresh media for either 30 minutes or 8 hours. Cells were pre-extracted, fixed and IF microscopy against γH2AX was performed followed by image analysis in Cell Profiler.

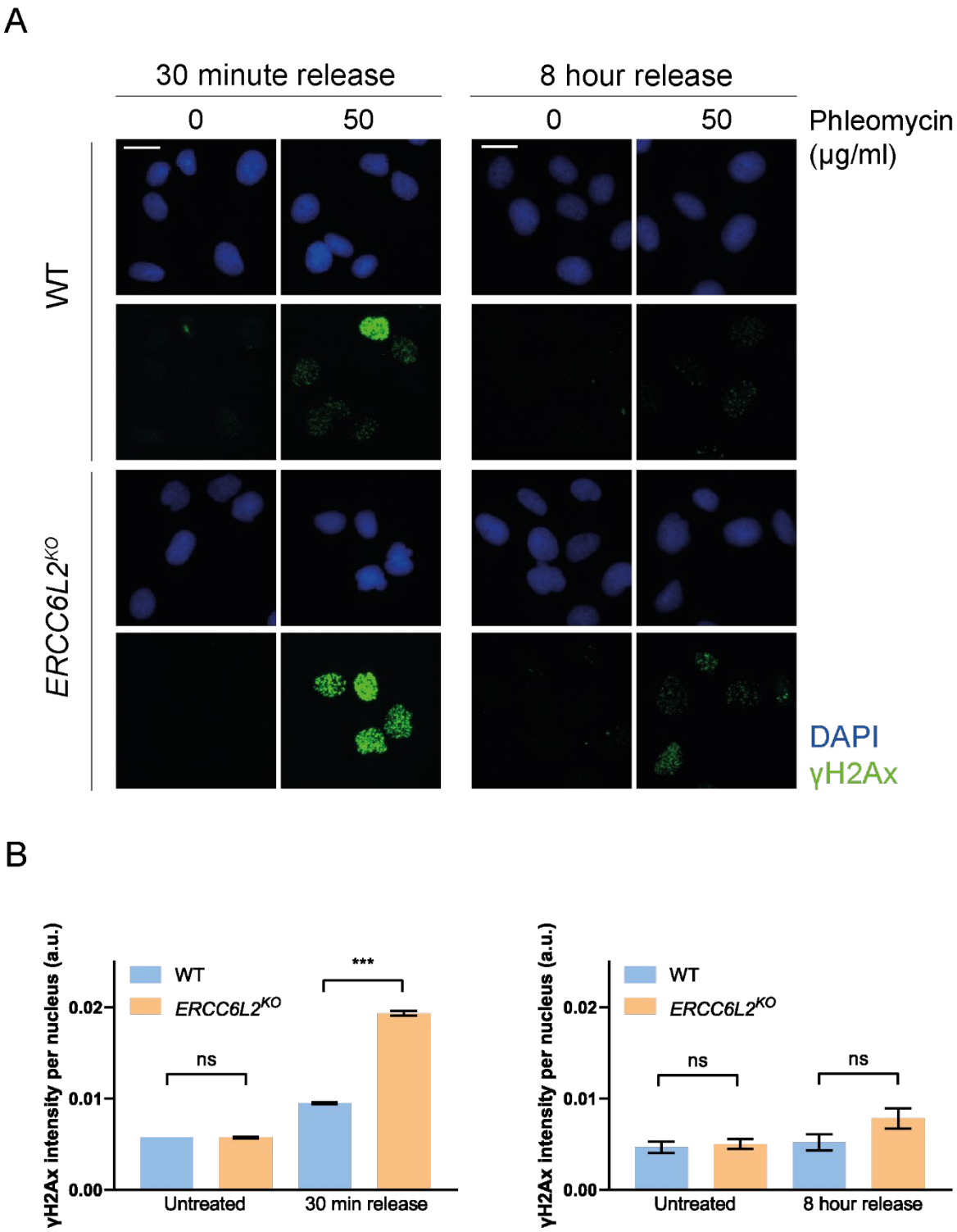


Figure 3-3: Assessing DNA damage recovery from acute phleomycin treatment

A. Representative images of WT and *ERCC6L2*^{KO} U-2-OS cells treated with 50 µg/ml phleomycin for 30 minutes then released for either 30 minutes or 8 hours. Cells were immunostained with γH2AX following pre-extraction and fixation. DNA was counterstained with DAPI. Scale bar, 20 µM. B. Bar plots represent average mean intensity of γH2AX per nucleus with SEM. ns = not significant; P < 0.001 = ***; unpaired t-test performed on means from two independent experiments with at least 250 nuclei analysed per condition. Mean intensities were measured using Cell Profiler analysis software.

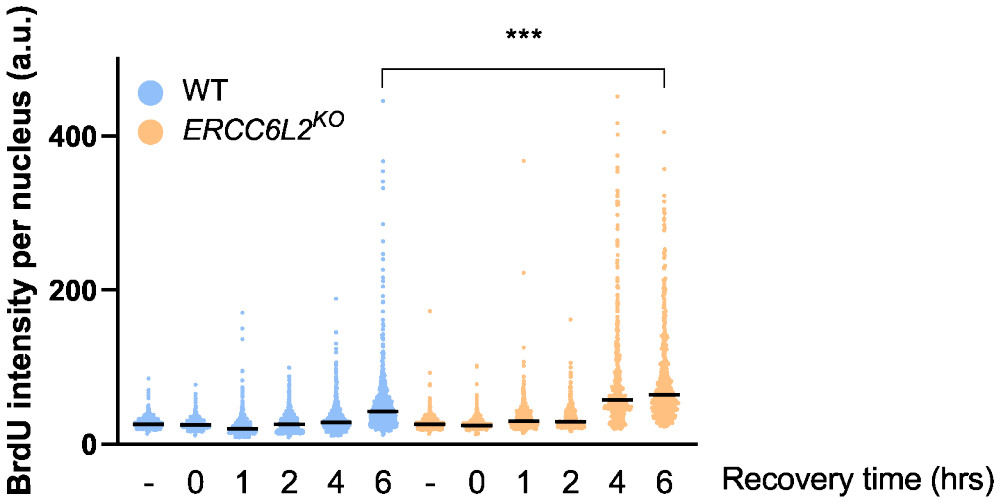
As shown in Figure 3-3, I found that the *ERCC6L2*^{KO} cells have significantly higher levels of γH2AX after 30 minutes of release, compared to WT cells. However, following 8 hours of recovery, damage levels in both cell lines had returned to amounts comparable to untreated conditions. The absence of damage response factors known to be required for DSB repair in regions of heterochromatin, such as ATM and Artemis, have been shown to cause increased levels of DSBs after such a release time (231). I also have preliminary evidence to suggest that colocalisation between γH2AX and the centromeric protein CENP-B, occurs at a similar rate to DNA damage induction, indicating a non-specific increase in centromeric DNA damage (data not shown). Collectively, this suggests that the role of ERCC6L2 in response to DNA damage is not specifically related to centromeres. Since γH2AX is not specific to DSBs, in the future it would be beneficial to quantify levels of additional markers such as 53BP1 or pKAP1 (s824) to check if similar observations are made.

It is important to note that the two major DSB repair pathways, c-NHEJ and HR, take varying lengths of time to complete. Using *in vivo* reporter assays, it has been shown that DSBs can be repaired via c-NHEJ in as little as 30 minutes (232). Comparatively, repair by HR takes much longer to complete (>7 hours). The observation that *ERCC6L2*^{KO} cells are able to recover from DNA damage after 8 hours of recovery but exhibit significantly higher levels at short release timepoints, may suggest that they have a defect in the rapid repair of DSBs by c-NHEJ.

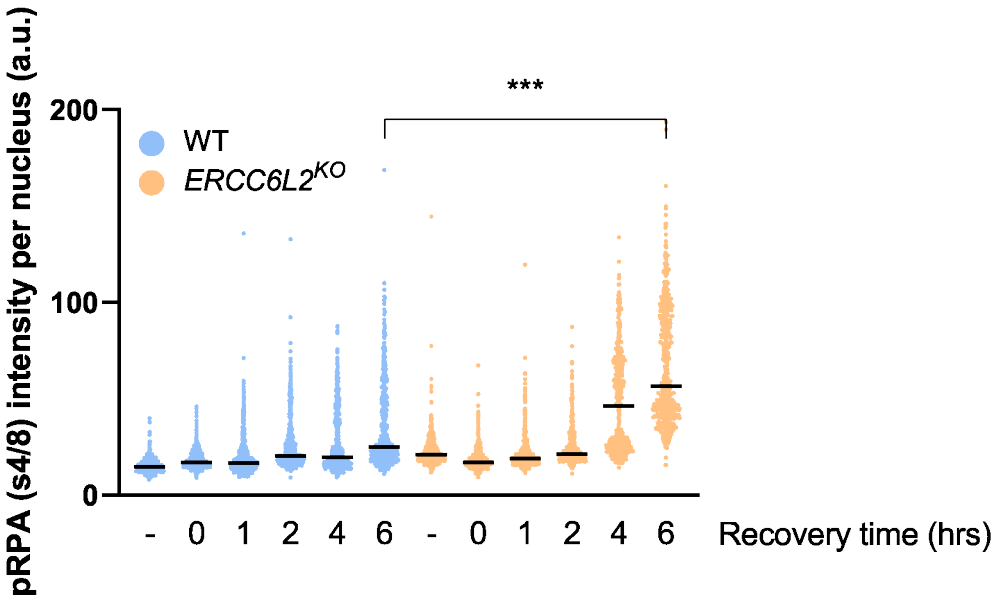
Figure 3-2 demonstrated that under prolonged etoposide and phleomycin treatment, *ERCC6L2*^{KO} cells accumulate higher levels of pRPA (s4/8). As previously mentioned, this could indicate that loss of ERCC6L2 leads to higher amounts of ssDNA that may reflect increased levels of DNA end resection, which is in line with the hypothesis that ERCC6L2 acts as a c-NHEJ promoting factor.

To further investigate whether ERCC6L2 plays a role in DNA end resection, I assessed the levels of resected DNA at various timepoints following release from phleomycin treatment in WT and *ERCC6L2*^{KO} cells. To quantify levels of ssDNA directly, I performed a native bromodeoxyuridine (BrdU) assay. Cells were incubated with the nucleoside analogue BrdU for 48 hours. Cells were then treated with phleomycin and allowed to recover for different lengths of time. Following pre-extraction and fixation, IF microscopy against BrdU was performed. I also quantified levels of pRPA (s4/8) as an additional readout for resected DNA.

A


***
WT
ERCC6L2KO
400
200
0
BrdU intensity per nucleus (a.u.)
- 0 1 2 4 6 - 0 1 2 4 6
Recovery time (hrs)

B


***
WT
ERCC6L2KO
200
100
0
pRPA (s4/8) intensity per nucleus (a.u.)
- 0 1 2 4 6 - 0 1 2 4 6
Recovery time (hrs)

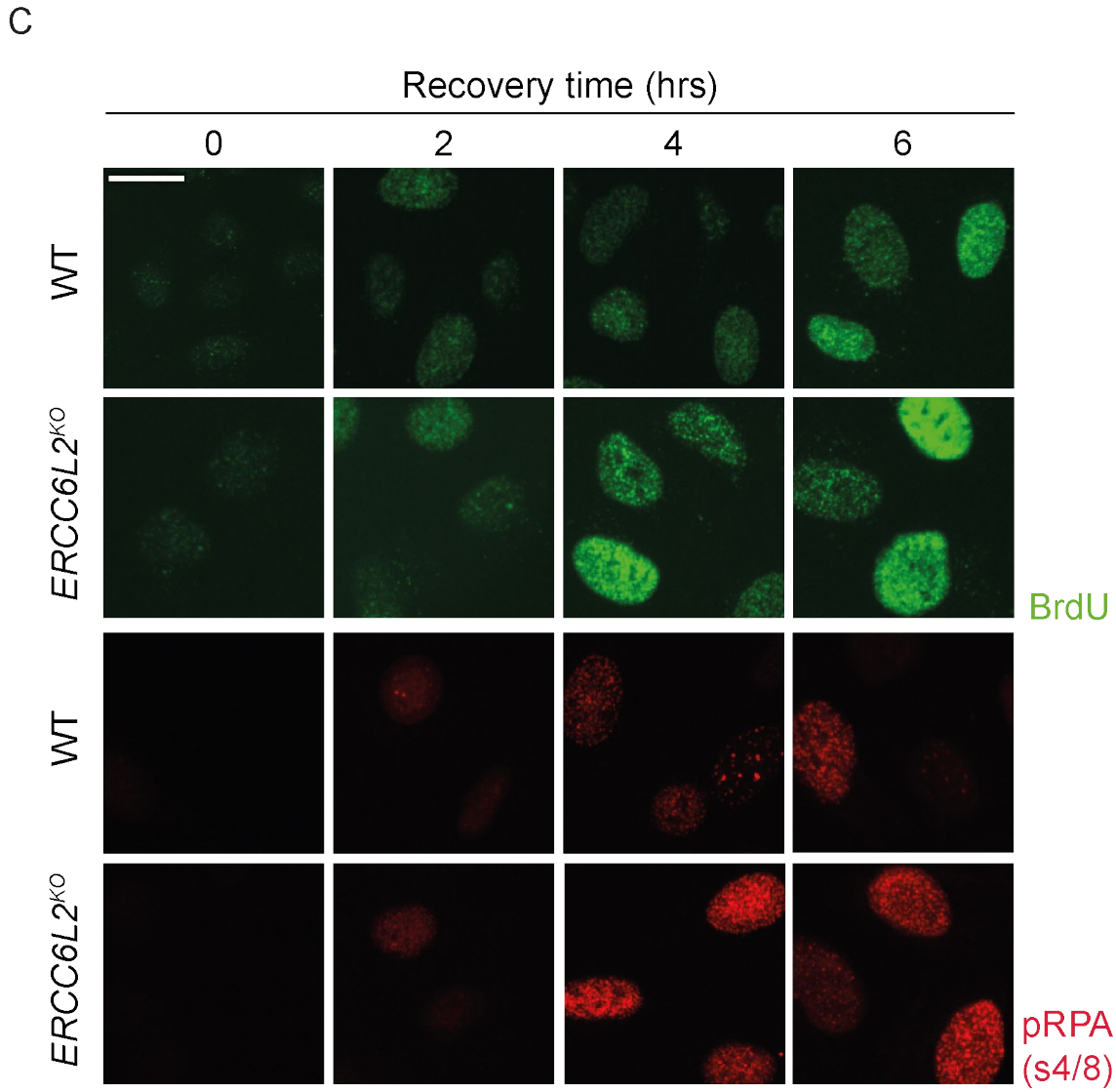


Figure 3-4: *ERCC6L2*^{KO} cells accumulate greater levels of ssDNA following phleomycin treatment

A. WT and *ERCC6L2*^{KO} U-2-OS cells were incubated for 48 hours with 10 μM BrdU. Cells were treated with 50 μg/ml phleomycin for 2 hours and released for the stated amounts of time. Following pre-extraction and fixation, cells were immunostained with anti-BrdU and anti-pRPA (s4/8). Total BrdU intensities per nucleus were quantified in Cell Profiler. Black bars indicate median values ($P < 0.0001$ = ***; Mann-Whitney test; N = > 500). B. As in A, for pRPA (s4/8) ($P < 0.0001$ = ***; Mann-Whitney test; N = > 500). At least 500 cells were analysed per condition. C. Representative images of cells analysed in A and B. Scale bar, 20 μM.

Intriguingly, I found that in the recovery period following treatment with phleomycin, *ERCC6L2*^{KO} cells showed higher levels of ssDNA, particularly at later timepoints (Figure 3-4). This hints towards a progressive accumulation of ssDNA. Experiments

from other group members have also shown that *ERCC6L2*^{KO} cells have increased RAD51 foci formation, compared to WT cells (Dr Dragana Ahel). This suggests that the absence of ERCC6L2 leads to an increase in resected DNA and the use of HR for repair of DSBs. In the future, it would be useful to assess repair kinetics using an additional direct method of ssDNA detection such as an alkaline COMET assay.

A common theme emerging from many studies is that loss of anti-resection factors can suppress PARP inhibitor sensitivity in *BRCA1*-null cells (75,77,96,233). PARP1 is a highly abundant protein with important roles in maintaining genome stability during SSB and DSB repair, as well as in replication fork protection (19). PARP inhibition leads to the formation of DSBs which depend on HR for repair (22). However, the exact nature of the molecular mechanisms underlying DSB induction by PARP inhibitors is still under intense investigation. Nevertheless, PARP inhibitors are commonly used to treat *BRCA1* and *BRCA2*-deficient cancers, exploiting their inability to repair DSBs via HR (22,234).

Several studies have demonstrated that *BRCA1*-deficient tumours can develop resistance to PARP inhibitors through loss or inactivation of pro-c-NHEJ factors, perhaps most notably 53BP1 and components of the Shieldin and CST complexes (75,76,96,235). Thus, I was keen to test if a similar phenotype could be observed in *ERCC6L2*^{KO} cells. To establish if loss of ERCC6L2 can suppress sensitivity to PARP inhibition in *BRCA1*-deficient cells, I performed clonogenic survival assays with WT and *ERCC6L2*^{KO} cells with the PARP inhibitor olaparib, following BRCA1 depletion.

Following BRCA1 depletion by siRNA (Figure 3-5A), cells were treated with a range of olaparib concentrations. As shown in Figure 3-5B, BRCA1 depletion caused a significant increase in sensitivity to olaparib, as expected. However, this is observed to

a similar extent for both WT and *ERCC6L2*^{KO} cells. Thus, I did not observe increased resistance to olaparib in the *ERCC6L2*^{KO} cells in the absence of BRCA1.

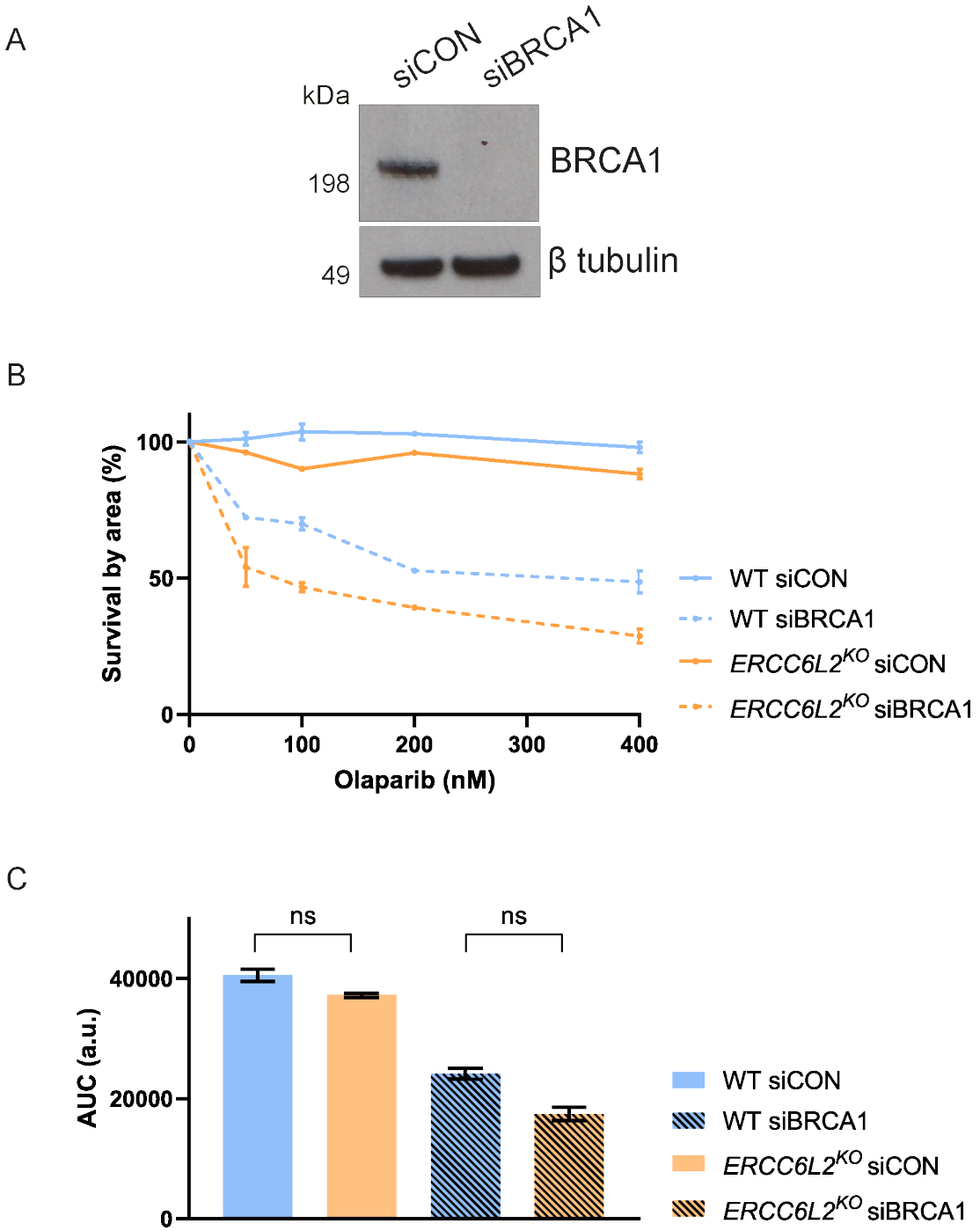


Figure 3-5: Loss of ERCC6L2 does not suppress PARP inhibitor sensitivity in the absence of BRCA1

A. Western blot analysis of BRCA1 levels in whole cell extracts prepared from U-2-OS cells transfected with BRCA1 siRNA. B. Clonogenic survival assay of WT and *ERCC6L2*^{KO} U-2-OS cells following BRCA1 depletion and olaparib treatment. Cells were transfected with control and BRCA1 siRNAs. 48 hours following transfection, cells were seeded in 6 cm dishes and treated with the stated doses of olaparib the following day. Cells were monitored over 10-14 days. Error bars represent SEM. Graphs are representative of three technical repeats. B. Area under the curve (AUC) calculations of A with Mann-Whitney test (ns = not significant). Error bars represent SEM.

3.2. Discussion

In this chapter I aimed to investigate and better understand the molecular basis of *ERCC6L2*^{KO} cell sensitivities to DSB inducers. Firstly, I showed that under etoposide and phleomycin treatment, *ERCC6L2*^{KO} cells accumulate in G2 phase to a greater extent than WT cells (Figure 3-1). I further demonstrated that this is likely owed to higher levels of unresolved DNA damage (Figure 3-2). I showed that *ERCC6L2*^{KO} cells eventually recover from acute phleomycin treatment but show significantly higher levels of DDR signalling after a short recovery period (Figure 3-3). Interestingly, following recovery from phleomycin treatment, *ERCC6L2*^{KO} cells progressively accumulate higher levels of ssDNA (Figure 3-4). Finally, I tested if loss of ERCC6L2 could suppress sensitivity to PARP inhibition in the absence of BRCA1. However, I was unable to detect any evidence of this in our cell lines (Figure 3-5).

The initial results from the cell cycle and DDR signalling analysis (Figure 3-1 and 3-2) suggest that loss of ERCC6L2 causes increased levels of DNA damage under etoposide and phleomycin treatment. Consequently, *ERCC6L2*^{KO} cells accumulate to a greater extent in G2 phase. Notably, under both treatments *ERCC6L2*^{KO} cells exhibited higher levels of pRPA (s4/8), which could indicate that there are increased levels of DNA end resection occurring in the absence of ERCC6L2 (230).

DNA end resection is a critical event in DSB repair since the initiation and extent to which it occurs determines which DSB repair pathway will be used. While c-NHEJ requires minimal end processing, HR involves extensive resection (236). It is known that c-NHEJ is a much faster and more efficient repair process than HR (232). Thus, the observation that *ERCC6L2*^{KO} cells have significantly higher levels of damage than WT cells after a short recovery could be indicative of a c-NHEJ defect (Figure 3-3).

The fact that *ERCC6L2*^{KO} cells can repair the damage following a longer release time suggests that cells lacking ERCC6L2 are still able to use HR for repair. This is supported by the lack of sensitivity observed towards olaparib in Figure 3-5 and the TOPO I inhibitor topotecan (Supp. Figure 7-1).

The hypothesis that ERCC6L2 has a role in promoting c-NHEJ is further strengthened by the increased levels of DDR signalling suggestive of ssDNA in *ERCC6L2*^{KO} cells following release from phleomycin treatment, particularly, at later timepoints (Figure 3-4). While the experiments differ slightly in concentrations and length of treatments, it is worth noting that while *ERCC6L2*^{KO} cells have comparable levels of γH2AX to WT cells following a long release from phleomycin treatment (Figure 3-3), they have markedly higher ssDNA levels after a similar release period. Other members of the lab have also observed similar effects, namely that levels of DSB markers such as γH2AX and MDC1 foci are largely unaffected, after at least 6 hours of recovery, yet levels of ssDNA are consistently higher in *ERCC6L2*^{KO} cells (Dr Dragana Ahel). This suggests that ERCC6L2 restrains DNA end processing and influences how at least a subset of DSBs are repaired.

Through the use of *in vivo* reporter assays, recent studies into the role of ERCC6L2 in DSB repair revealed that loss of ERCC6L2 leads to a deficiency in c-NHEJ and CSR (224–226). The impact of ERCC6L2 depletion on c-NHEJ was less significant than for core factors such as ligase IV and XRCC4 indicating that ERCC6L2 is not required for all DSB repair events. This observation contradicted an earlier report which found that loss of ERCC6L2 did not have significant effect on c-NHEJ (217). However, the use of a more sophisticated reporter assay that strictly requires repair mediated by c-NHEJ factors may have increased dependence on ERCC6L2 for repair (225). Interestingly, Olivieri and colleagues noted that loss of ERCC6L2 does not increase the sensitivity

of *XRCC4*^{KO} or *XLF*^{KO} cells to bleomycin treatment, further suggesting that ERCC6L2 influences or operates within the c-NHEJ pathway. The observation that ERCC6L2 promotes c-NHEJ align with our observation that its loss leads to increased levels of DNA end resection.

While core components of the c-NHEJ machinery are refractory to DNA end resection, the decision to promote or suppress this process, is governed mainly by BRCA1 and 53BP1, respectively (96). 53BP1 and BRCA1 mutually antagonise each other depending on cell cycle stage. In G1 phase, 53BP1 limits DNA resection through recruitment of effector proteins such as RIF1-Shieldin and CST complexes, as well as PTIP which actively counteract the process (70,74,75,237). BRCA1 is excluded from chromatin in a RIF1 dependent manner (79). As cells transition into S/G2, DNA end resection is promoted when 53BP1 is displaced by BRCA1. Extensive work has demonstrated that inactivation of 53BP1 and its downstream effectors can rescue PARP inhibitor sensitivity in *BRCA1*-deficient cancers (75–77). Notably, loss of core c-NHEJ factors such as ligase IV, do not mediate the same effect (96).

I tested if loss of ERCC6L2 could suppress sensitivity to PARP inhibition in the absence of BRCA1. However, I observed no difference in olaparib sensitivity between WT and *ERCC6L2*^{KO} cells when BRCA1 was depleted (Figure 3-5). In the future, it would be beneficial to conduct this experiment using 53BP1 depletion as a positive control to ensure that resistance to olaparib in the absence of BRCA1 is observable. Two recent publications, that propose a role for ERCC6L2 in c-NHEJ, also investigated this using similar cell growth assays. The first study by Olivieri et al, which conducted the experiment in *BRCA1/p53*-null RPE1 cells, found that olaparib sensitivity was not rescued upon ERCC6L2 depletion (225). However, Francica and colleagues found that ERCC6L2 loss did rescue olaparib sensitivity in *BRCA1/p53*-null mouse cells. This

followed the observation that *ERCC6L2* clustered strongly with components of the 53BP1-Shieldin repair axis in their genome wide screen that identified genes important in mediating resistance to IR (226). In this study, they used a *BRCA1*/*p53*-null cell line derived from a mouse mammary tumour (KB1P) and were able to demonstrate that this resistance is mediated by partial restoration of HR. This implies that ERCC6L2 functions as an anti-resection factor rather than a core c-NHEJ component.

The reason behind the differences between these studies and my own data is not obvious. Both studies depleted ERCC6L2 using lentiviral transduction in *BRCA1/p53*-deficient cells and used similar experimental set-ups, in terms of timeframe and olaparib concentrations. However, the study from Olivieri and colleagues, which did not observe olaparib resistance in the absence of ERCC6L2, used RPE1 cells. I also used human cells in my experiments, albeit a p53 proficient cancer derived cell line (U-2-OS) yet made similar observations. Interestingly, the initial screen that first identified the Shieldin components as mediators of PARP inhibitor resistance used a human *BRCA1* mutant cancer cell line (SUM149PT) and ERCC6L2 was not identified as a candidate (75). It would be intriguing to follow up these results by testing a range of cell lines to ascertain how genetic background or species of cell line influences the effect that *ERCC6L2* status has on cell survival in the absence of BRCA1.

Despite the differences between these studies and my own data, there is broad agreement that ERCC6L2 is involved in promoting c-NHEJ and its absence leads to an increase in DNA end resection. It is still unclear how exactly ERCC6L2 mediates this function and what context it may operate in. While biochemical studies are required to provide insight into the mechanistic function of ERCC6L2, it is interesting to speculate what role it could be playing at DSB sites.

A small number of interactors have been identified for ERCC6L2 including the recently characterised CYREN, which regulates c-NHEJ through an interaction with several DSB repair factors including the KU heterodimer, ATM, MRN and XLF (238,239). An interaction between ERCC6L2 and DNA-PK has also been described, which I was able to confirm (Supp. Figure 7-2) (220). Since ERCC6L2 is a predicted DNA translocase, it could act to influence the chromatin environment at DSB ends to facilitate repair. It is known that ERCC6L2 is recruited rapidly to DSB sites (Dr Chris Carnie), in a manner that is independent of KU, XLF, NBS1, γH2AX and PARP1 (224). A minor effect on recruitment of KU and XLF has been observed in the absence of ERCC6L2 (224). Thus, it has been proposed that ERCC6L2 may remodel DNA at DSB sites or perhaps remove proteins from DNA ends, to limit DNA end resection. In specific contexts, this could in turn promote efficient loading of core c-NHEJ factors and rapid end joining.

Given that we know ERCC6L2 colocalises with centromeres throughout interphase (Dr Chris Carnie), it is an intriguing prospect that its role in DSB repair relates to these regions. A study by Aze et al reconstituted human centromeric chromatin using bacterial artificial chromosomes (BACs), in the *Xenopus* egg extract system (191). Mass spectrometry analysis revealed that DSB repair factors, such as KU80, DNA-PK and MRE11/Rad50 are enriched on centromeric chromatin (191). Replication stress induced DSBs have also been shown to occur more frequently within centromeric DNA sequences (190). This may explain the constitutive presence of certain DSB repair proteins observed by Aze and colleagues (191). However, DSBs which arise from stalled or collapsed replication forks are strictly repaired by HR due to their one-ended nature (238,240). Given that ERCC6L2 appears to promote c-NHEJ and suppress DNA end resection, it is somewhat difficult to reconcile a role for it in the repair of replication associated breaks. One possibility is that ERCC6L2 functions to suppress

DNA end resection to ultimately prevent aberrant recombination events, which are a potential source of genome instability to arise from centromeres (200). During HR, the sister chromatid is typically used as a template for repair. However, there is evidence that repetitive sequences can use homologous and homeologous sequences on other chromosomes as templates for repair (241). This was only shown to occur for a small percentage (~2%) of repair events in wild-type cells indicating that such events are normally suppressed.

Repair of centromeric DSBs is a relatively understudied topic, especially in human cells. Studies conducted in Drosophila and mouse cells have revealed that DSBs are physically relocated away from dense regions of heterochromatin during S/G2 (201,203). DNA end resection occurs prior to relocation and RAD51 loading occurs following repositioning of the break, which is thought to limit illegitimate recombination. It is currently unknown whether such a process occurs in human cells. However, these observations are enough to demonstrate that the regulation of DSB repair within heterochromatin and at centromeres is distinct from other regions of the genome. While it seems likely that ERCC6L2 is involved in the global response to DSBs, it is tempting to speculate that the specific nature of its role could be related to chromatin context, which may not necessarily be limited to centromeres.

Overall, mounting evidence from our group along with published studies strongly suggests that ERCC6L2 functions in a c-NHEJ promoting capacity in the repair of DSBs. The observation that loss of ERCC6L2 leads to an increase in DNA end resection could indicate that it functions as an anti-resection factor rather than a core c-NHEJ component. Further work is required to determine the specific repair contexts in which ERCC6L2 functions and whether this is somehow linked to centromeres and indeed, other regions of the genome.

4. Exploring a role for ERCC6L2 in DNA replication

4.1. Results

While it has emerged over the course of this study that the primary function of ERCC6L2 is likely in the DSB response, I was intrigued by a number of observations which link ERCC6L2 to DNA replication. ERCC6L2 interacts with the sliding clamp PCNA through a novel PCNA interaction motif identified by a past member of our lab (Dr Chris Carnie). PCNA acts as a scaffold at replication forks, which engages a vast number of proteins such as core replication machinery components, chromatin remodelling enzymes and DNA repair factors (242). These interactions are tightly regulated to ensure faithful DNA replication. Notably, full-length ERCC6L2 does not colocalise with PCNA in unperturbed cells which could mean that ERCC6L2 is only present at a small subset of replication forks, which is not detectable by IF (Dr Chris Carnie). Colocalisation with PCNA is only observed in S phase cells when the C-terminal half (712-1561) of ERCC6L2, which is required for damage and centromere recruitment, is overexpressed (Dr Chris Carnie). This could mean that ERCC6L2 is only recruited to replication forks under periods of stress or to particular genomic regions that are inherently difficult to replicate.

In addition to the PCNA interaction, full-length ERCC6L2 also colocalises with centromeres in unperturbed conditions, throughout interphase (Dr Chris Carnie). The highly repetitive nature of centromeric DNA can challenge progression of the replication fork, thus retention of ERCC6L2 at centromeres during S phase could hint that it is involved in facilitating this process in some capacity.

Firstly, I reasoned that if ERCC6L2 is involved in either a global response to replication stress or DNA replication at specific genomic loci, its absence would cause sensitivity

to the replication inhibitor hydroxyurea (HU). To test this, I performed clonogenic survival assays using WT and *ERCC6L2*^{KO} U-2-OS and hTERT-RPE1 cells with a range of HU concentrations.

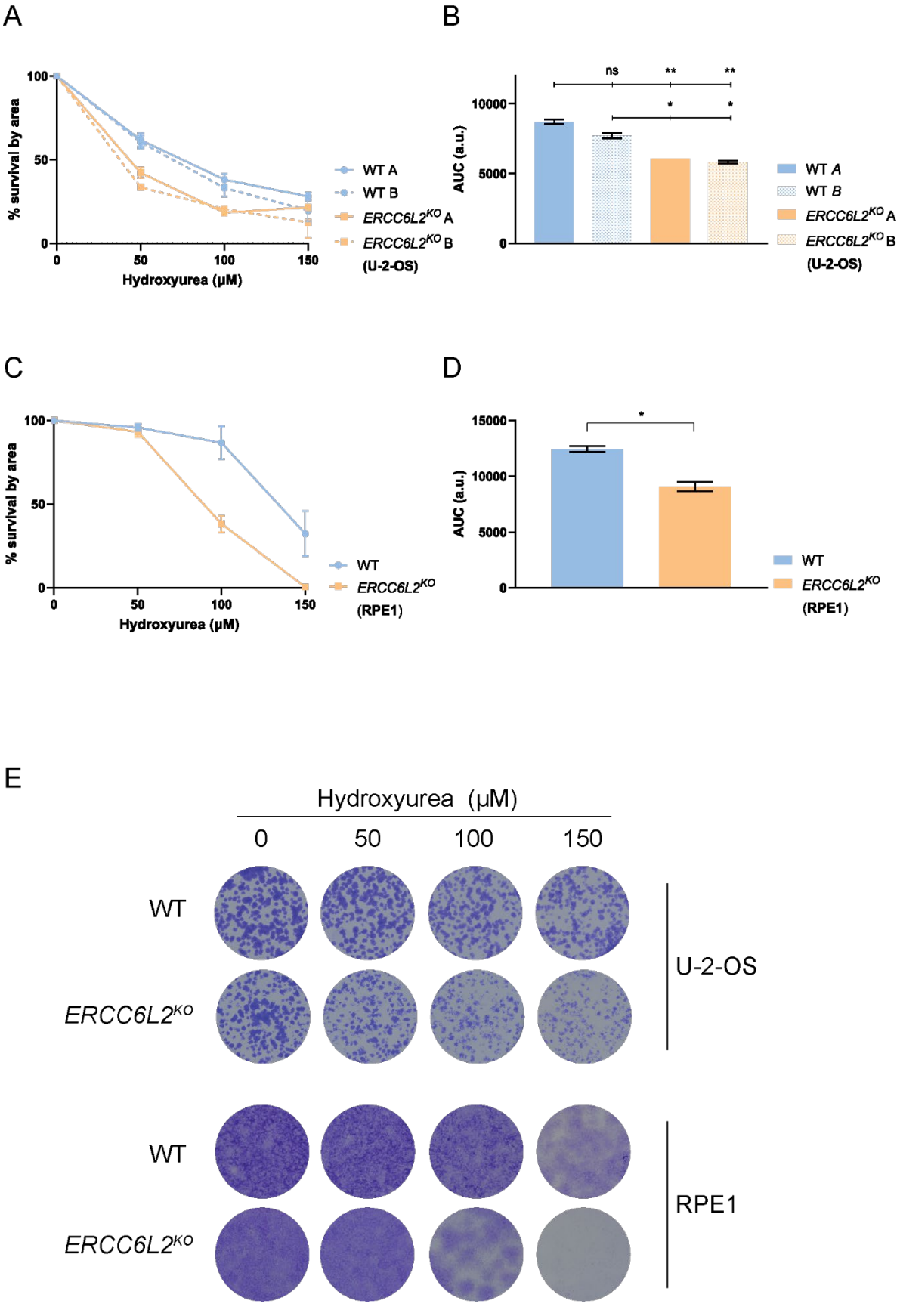

A
100
50
0
% survival by area
0
50
100
150
Hydroxyurea (μM)
WT A
WT B
ERCC6L2KO A
ERCC6L2KO B
(U-2-OS)
B
ns
**
**
*
*
10000
5000
0
AUC (a.u.)
WT A
WT B
ERCC6L2KO A
ERCC6L2KO B
(U-2-OS)
C
100
50
0
% survival by area
0
50
100
150
Hydroxyurea (μM)
WT
ERCC6L2KO
(RPE1)
D
*
15000
10000
5000
0
AUC (a.u.)
WT
ERCC6L2KO
(RPE1)
E
Hydroxyurea (μM)
0
50
100
150
WT
ERCC6L2KO
U-2-OS
WT
ERCC6L2KO
RPE1

Figure 4-1: *ERCC6L2*^{KO} cells exhibit mild sensitivity to HU-induced replication stress

A. Clonogenic survival assays performed with WT and *ERCC6L2*^{KO} U-2-OS cell lines treated with HU. Cells were seeded in 6-well plates and treated with a range of HU concentrations 24 hours later. Error bars represent SEM. Graphs are representative of two experimental repeats. B. Area under the curve (AUC) calculations of A with unpaired t-test performed on the means. Error bars represent SEM. C. As in A, with hTERT-RPE1 cells. D. As in B, for C. E. Representative images from clonogenic survival assays quantified in A and B. ns = not significant, $P \leq 0.05$ = *, $P < 0.005$ = **.

I tested two independent WT and *ERCC6L2*^{KO} U-2-OS cell lines and found that the *ERCC6L2*^{KO} cells show mild sensitivity to HU compared to the WT cells (Figure 4-1). The degree of sensitivity is comparable between these two *ERCC6L2*^{KO} cell lines. For the RPE1 *ERCC6L2*^{KO} cells, I also observed a significant difference in HU sensitivity, compared to WT cells, particularly at higher concentrations.

I had next planned to perform these experiments with complementation cell lines to see if these sensitivities can be reversed with stable overexpression of ERCC6L2. However, I had extensive problems with protein expression in these cell lines and therefore I was unable to test this with confidence. This may be due to the large size of ERCC6L2 or potential degradation at its unstructured C-terminus. It is also possible that ectopic expression of ERCC6L2 is detrimental to cells through negative interference with DNA end resection or replication. In the future, it may be beneficial to use an inducible expression system, which is easier to control and may mitigate overexpression-induced toxicity. Nevertheless, the fact that I was able to observe sensitivity across the different cell lines, suggests that loss of ERCC6L2 imparts some sensitivity to HU, and replication stress.

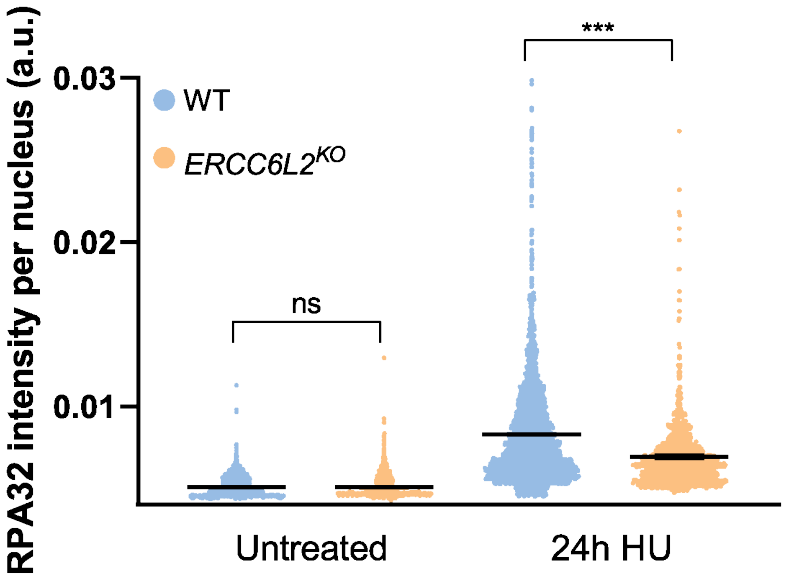
HU halts DNA replication by inhibiting ribonucleotide reductase; an enzyme involved in the production dNTPs (243). Depending on the length of treatment, HU will cause replication fork stalling, which after prolonged amounts of time (>4 hours) will begin to

collapse (42). Although replication fork collapse is not strictly defined, it is often associated with DSB formation. In such cases, DSBs can be generated actively by nuclease digestion of replication intermediates commonly found at stalled forks such as ssDNA or reversed forks (138,244). Alternatively, they can also form passively.

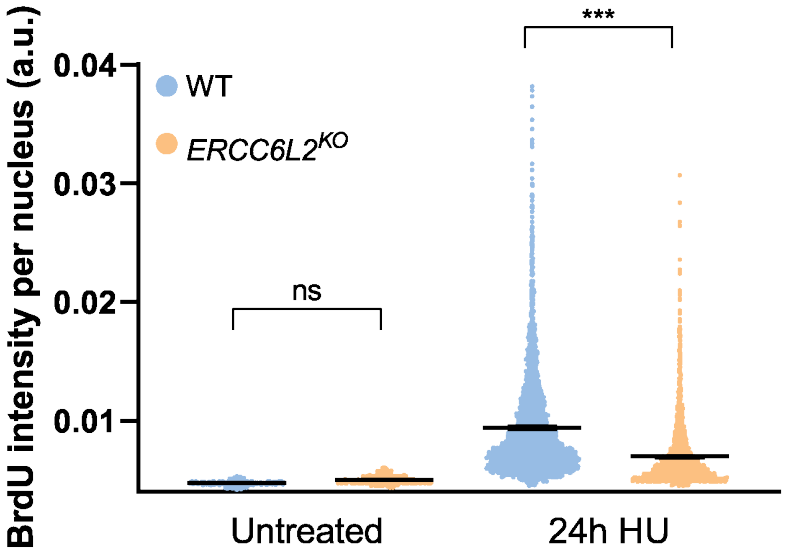
Collectively, published studies and my work from the previous chapter suggest that ERCC6L2 may play a role in DSB repair, where its loss is associated with an accumulation of ssDNA (224–226). Although during S phase DSBs are repaired mainly through HDR mechanisms, c-NHEJ is still active (238). At certain genomic loci, particularly repetitive regions, HR may also be a risky mode of repair in S/G2 due to the risk of chromosomal translocations. Furthermore, DSB repair at centromeres is poorly understood, especially in human cells and there is evidence to suggest that these processes are differentially regulated at centromeres in mammalian cells (203). Thus, it remains to be proven that ERCC6L2 is involved in the repair of S phase DSBs.

To investigate if the absence of ERCC6L2 resulted in similar phenotypes to those observed for DSB inducers, I chose to treat cells with HU for 24 hours, which causes prolonged fork stalling and DSB induction (245). I expected that if ERCC6L2 plays a similar role during DNA replication to what it does in DSB repair, then I would observe an increase in ssDNA formation under such conditions.

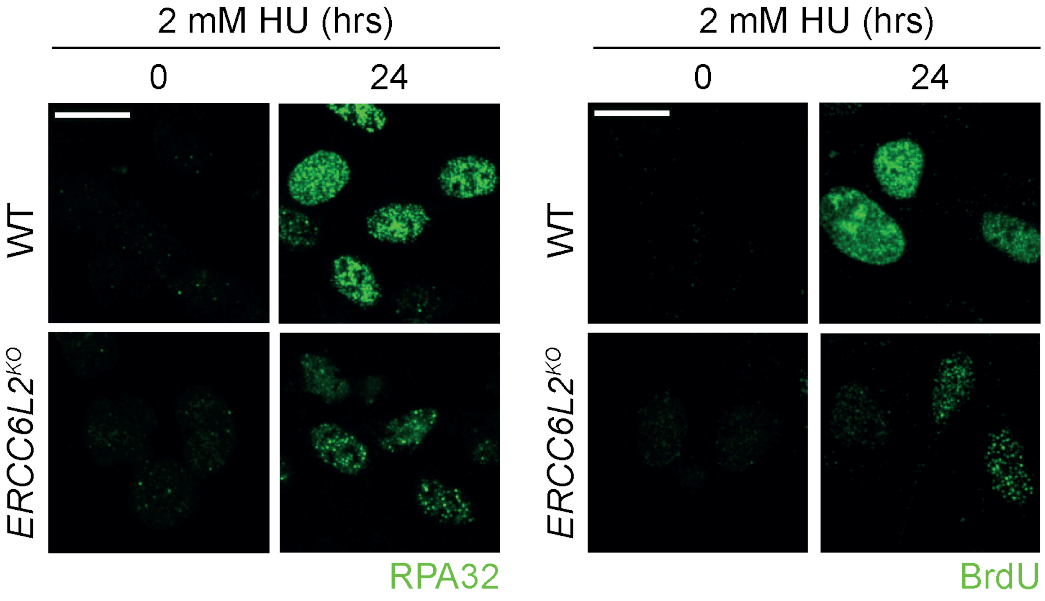
A


RPA32 intensity per nucleus (a.u.)
0.03
0.02
0.01
WT
ERCC6L2KO
ns
***
Untreated
24h HU

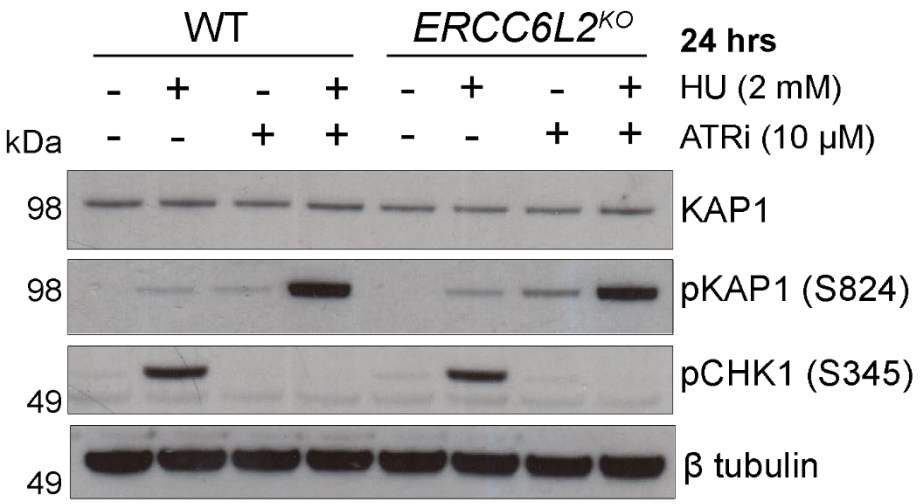
B


BrdU intensity per nucleus (a.u.)
0.04
0.03
0.02
0.01
WT
ERCC6L2KO
ns
***
Untreated
24h HU

C


2 mM HU (hrs)
0
24
2 mM HU (hrs)
0
24
WT
ERCC6L2KO
WT
ERCC6L2KO
RPA32
BrdU

D


WT
ERCC6L2KO
24 hrs
- + - + - + - +
HU (2 mM)
kDa - - + + - - + +
ATRi (10 μM)
98
KAP1
98
pKAP1 (S824)
49
pCHK1 (S345)
49
β tubulin

E

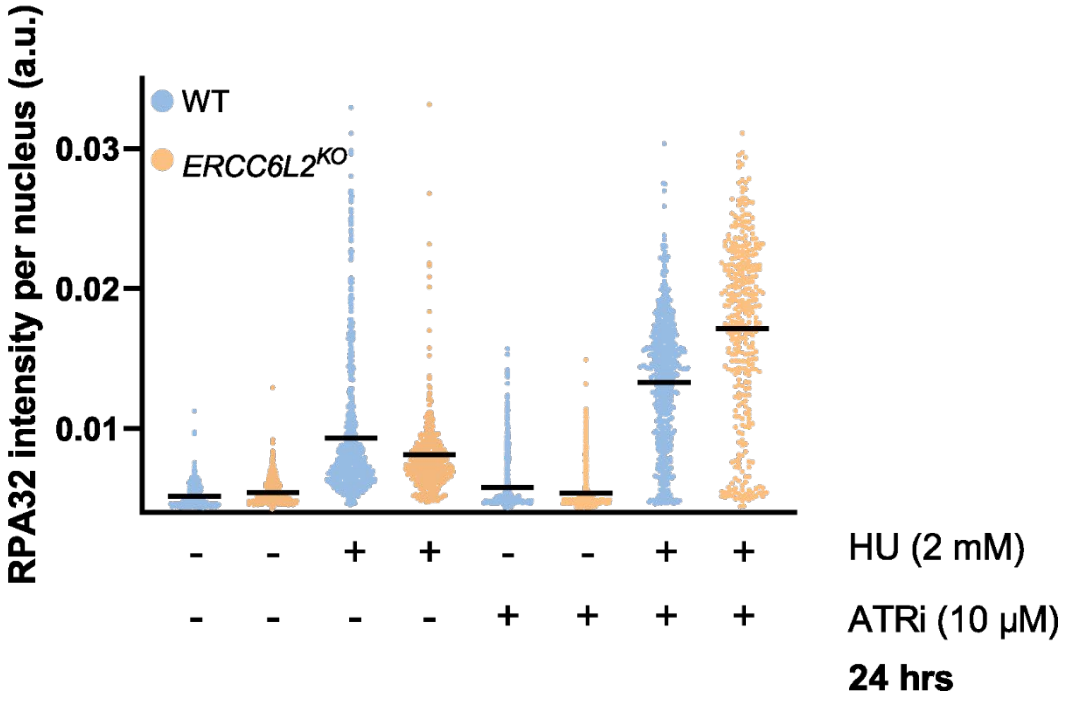

RPA32 intensity per nucleus (a.u.)
0.03
0.02
0.01
WT
ERCC6L2KO
- - + + - - + +
HU (2 mM)
- - - - + + + +
ATRi (10 μM)
24 hrs

Figure 4-2: ssDNA formation is limited in *ERCC6L2*^{KO} cells under HU treatment

A. WT and *ERCC6L2*^{KO} U-2-OS cells were treated with 2 mM HU for 24 hours. Following pre-extraction and fixation, cells were immunostained with anti-RPA32. Dot plots shows mean intensity of RPA32 per nucleus. Intensity of RPA32 was measured in at least 400 nuclei from three independent experiments. Bars represent means (ns = not significant; P < 0.0001 = ***; Mann-Whitney test; *N* = 1275). B. WT and *ERCC6L2*^{KO} U-2-OS cells were incubated for 48 hours with 10 μM BrdU. Cells were treated with 2 mM HU for 24 hours. Following pre-extraction and fixation, cells were immunostained with anti-BrdU. Dot plots show mean intensity of BrdU per nucleus. Intensity of BrdU was measured in at least 1400 nuclei from two independent experiments. Bars represent means (ns= not significant; P < 0.0001 = ***; Mann-Whitney test; *N* = 3039). C. Representative images of cells analysed in A and B. Scale bar, 20 μM. D. WT and *ERCC6L2*^{KO} U-2-OS cells were treated with the indicated doses of HU and ATRi (VE-821) for 24 hours. Whole cell extracts were analysed by western blotting with the indicated antibodies. E. WT and *ERCC6L2*^{KO} U-2-OS cells were treated with 2 mM HU for 24 hours +/- 10 μM ATRi (VE-821). Following pre-extraction and fixation, cells were immunostained with RPA32. Dot plots shows mean intensity of RPA32 per nucleus. Intensity of RPA32 was measured in at least 300 nuclei per condition.

To assess levels of ssDNA, I initially measured levels of chromatin bound RPA32 by IF microscopy. After 24 hours of HU treatment, cells were pre-extracted, fixed and immunostained with RPA32. As shown in Figure 4-2A, RPA intensity increases, following HU-induced replication stress in both WT and *ERCC6L2*^{KO} cells compared to untreated conditions. However, compared to WT cells, *ERCC6L2*^{KO} cells have significantly lower levels of chromatin bound RPA. To directly measure levels of ssDNA, I performed a native BrdU assay. As with RPA, I found that BrdU levels increase after HU treatment, but this occurs to a markedly lesser extent in the *ERCC6L2*^{KO} cells (Figure 4-2B).

The length of HU treatment used in this experiment means that ssDNA could be the product of different events at the replication fork. HU-induced replication fork stalling will result in uncoupling between the DNA polymerase and replicative helicase, which continues to unwind the DNA ahead of the fork (246). This is the major method for ATR kinase activation, which together with its binding partner, ATRIP, engages RPA coated ssDNA and phosphorylates its downstream targets. However, ssDNA can also be

found at replication intermediates such as reversed forks and ssDNA gaps (44). Finally, since this length of treatment would also lead to DSB formation, DNA end resection is another potential source of ssDNA.

Given that ssDNA is important for ATR activation, I tested to see if phosphorylation of downstream targets was affected. As shown in Figure 4-2D, robust phosphorylation of CHK1 (s345) is observed in both cell lines upon HU treatment. This suggests that while the levels of ssDNA are reduced in *ERCC6L2*^{KO} cells, they are still sufficient to activate ATR signalling. I then assessed the impact of ATR inhibition on levels of RPA combined with HU treatment.

As shown in Figure 4-2E, chromatin-associated RPA levels increase dramatically in both cell lines when HU treatment is combined with ATR inhibition, as expected. However, they increase by a much greater extent in *ERCC6L2*^{KO} cells. This may reflect an increase in fork collapse and DSB induction, since replication forks would no longer be stabilised or protected. This suggests that while *ERCC6L2*^{KO} cells have less ssDNA, global ATR activation is important for protecting against HU-induced replication stress in these cells.

Given that ATR activation still occurs in *ERCC6L2*^{KO} cells, it suggests that the process of replication uncoupling occurs sufficiently in the absence of ERCC6L2. However, it is possible that its absence may affect the process at a subset of replication forks, or it may limit the extent to which it occurs and consequently results in less ssDNA. For example, ERCC6L2 could be involved in removing proteins or secondary DNA structures which may prevent helicase progression.

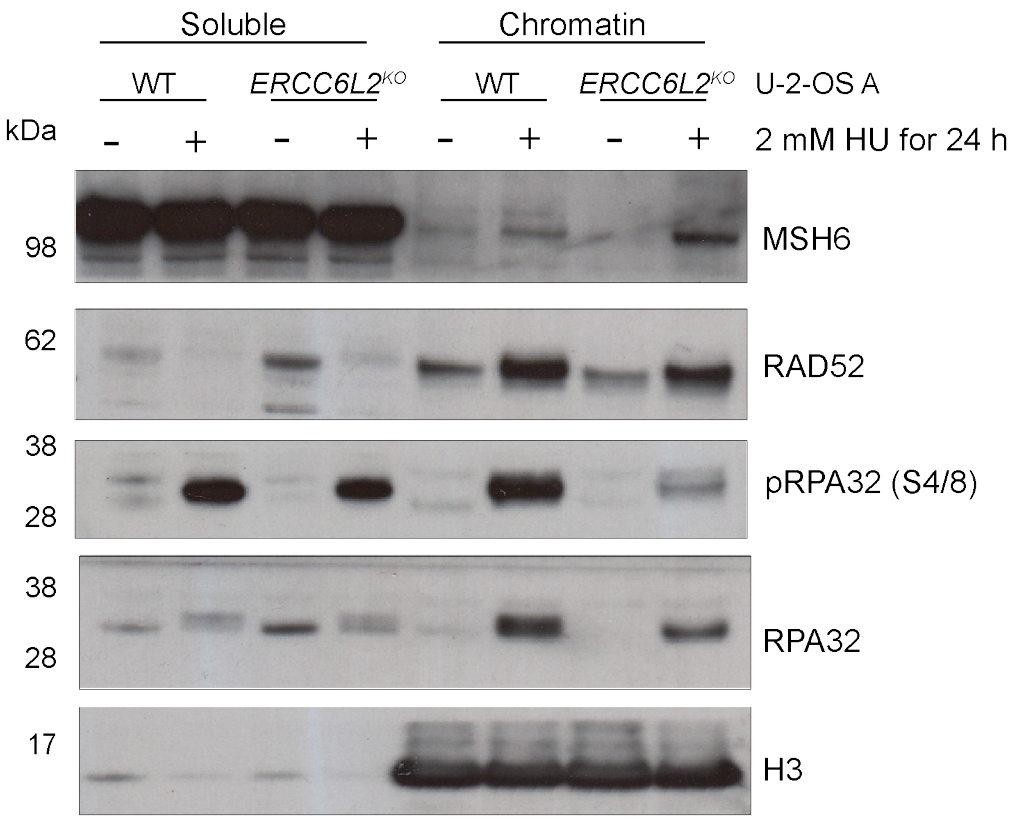
Secondary DNA structures are known to pose an inherent challenge to the replication of centromeric DNA so I hypothesised that ERCC6L2 may influence the presence of

such structures (187). A study by Aze and co-workers, which reconstituted centromeric DNA replication using human chromosomal DNA in BACs within the *Xenopus* egg extract system, revealed some intriguing and distinguishing features of this process (191). Namely, extensive ssDNA formation is limited by the topological arrangement of centromeric DNA and as a result, RPA loading occurs at significantly lower levels on centromeric chromatin. They further observed that RPA accumulation does not occur even in the presence of replication stress. Instead, the levels of the mismatch repair (MMR) factor MSH6, which they demonstrate is required for centromeric DNA replication in their system in unperturbed conditions, is further recruited in response to replication stress (191).

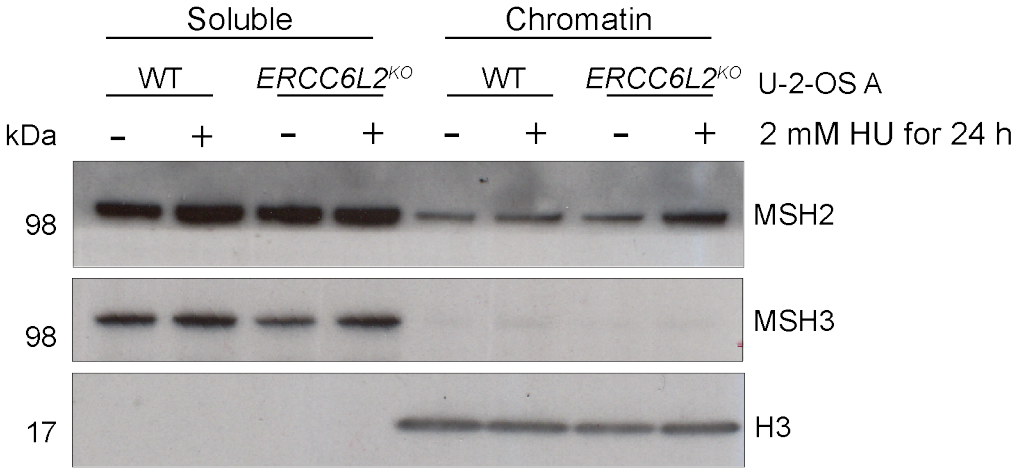
MSH6 is proposed to be required for secondary structure resolution during unperturbed centromeric DNA replication and it is suggested that replication stress would increase formation of these abnormal DNA structures leading to enhanced MSH6 recruitment. It is thought that short regions of ssDNA generated by replication uncoupling could form secondary DNA structures such as hairpin loops (191). Thus, it is also possible that formation of secondary DNA structures could reduce ssDNA levels.

Considering these observations, I hypothesised that if loss of ERCC6L2 was leading to an increase in secondary DNA structures, or simply more centromeric replication stress, then levels of chromatin bound MSH6 may be elevated in *ERCC6L2*^{KO} cells. To test this, I performed chromatin fractionation on WT and *ERCC6L2*^{KO} cells following treatment with HU and assessed protein levels by western blotting in soluble and chromatin bound nuclear extracts.

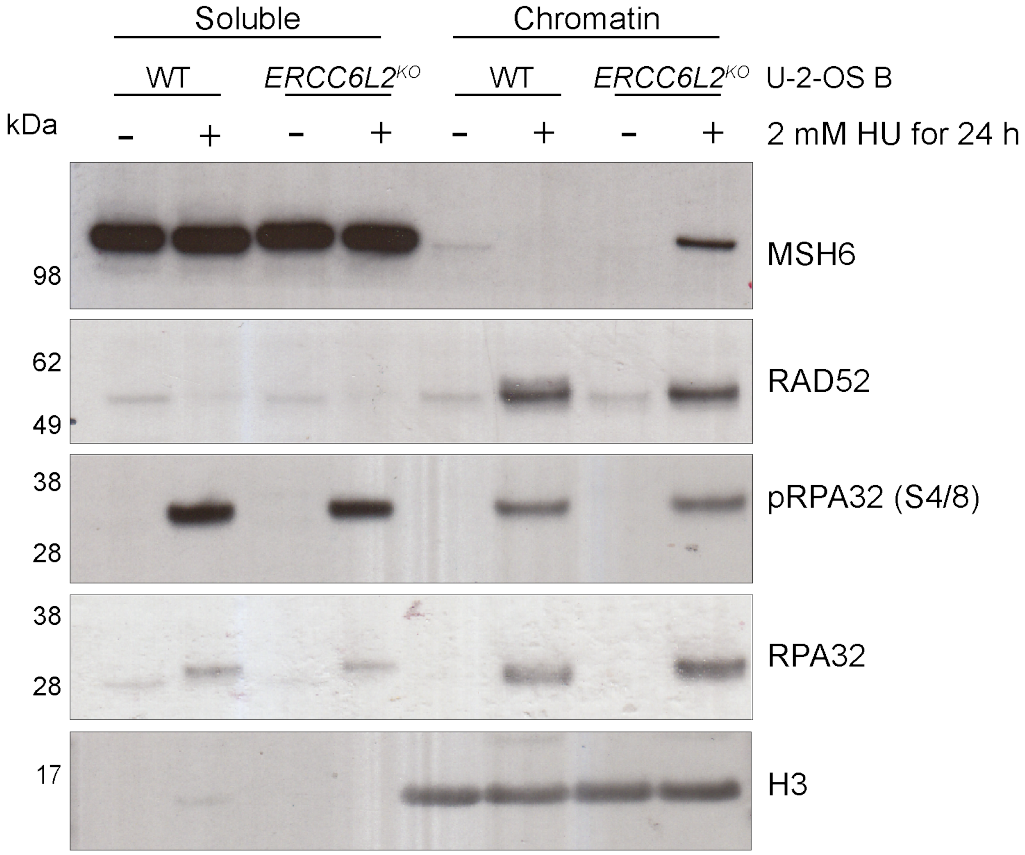
A


Soluble
Chromatin
WT
ERCC6L2KO
WT
ERCC6L2KO
U-2-OS A
kDa
- + - + - + - +
2 mM HU for 24 h
98
MSH6
62
RAD52
38
28
pRPA32 (S4/8)
38
28
RPA32
17
H3

B


Soluble
Chromatin
WT
ERCC6L2KO
WT
ERCC6L2KO
U-2-OS A
kDa
- + - + - + - +
2 mM HU for 24 h
98
MSH2
98
MSH3
17
H3

C


Soluble
Chromatin
WT
ERCC6L2KO
WT
ERCC6L2KO
U-2-OS B
kDa
- + - + - + - +
2 mM HU for 24 h
98
MSH6
62
49
RAD52
38
28
pRPA32 (S4/8)
38
28
RPA32
17
H3

D

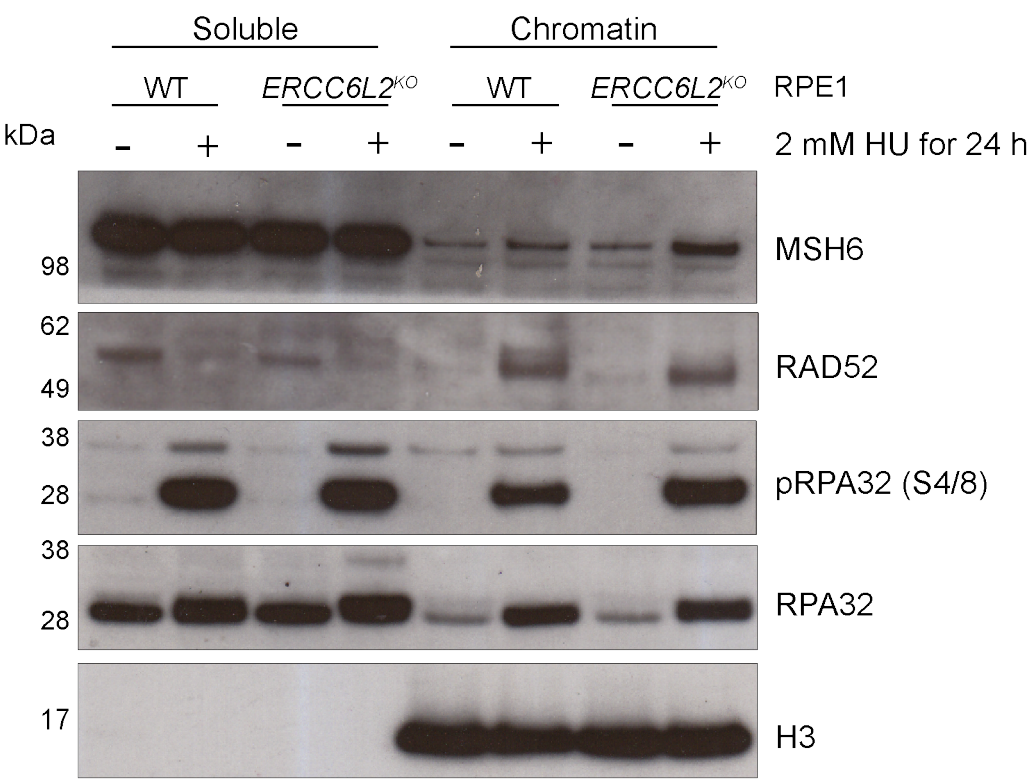

Soluble
Chromatin
WT
ERCC6L2KO
WT
ERCC6L2KO
RPE1
kDa
- + - + - + - +
2 mM HU for 24 h
98
MSH6
62
49
RAD52
38
28
pRPA32 (S4/8)
38
28
RPA32
17
H3

Figure 4-3: *ERCC6L2*^{KO} cells have increased levels of chromatin bound MSH2/6 under HU treatment

A. WT and *ERCC6L2*^{KO} U-2-OS cells were treated with 2 mM HU for 24 hours. Following treatment, cells were harvested, and chromatin fractionation was performed. Soluble and chromatin bound nuclear fractions were analysed by western blotting with the indicated antibodies. B. As in A. C. As in A for alternative *ERCC6L2*^{KO} U-2-OS cell line. D. As in A, for *ERCC6L2*^{KO} hTERT-RPE1 cells.

As shown in Figure 4-3A, western analysis of the soluble and chromatin bound nuclear fractions revealed that *ERCC6L2*^{KO} cells do indeed have increased levels of chromatin bound MSH6 under HU treatment. During mismatch repair MSH6 functions in a heterodimeric complex with MSH2 to resolve 1-2 bp nucleotide substitutions and small loops (247). MSH2 can also form a heterodimeric complex with MSH3 which resolves more extensive loop structures. Crucially, Aze et al found that MSH2 was also enriched on centromeric chromatin but not MSH3 (191). Considering this, I also assessed the levels of MSH2 and MSH3 following chromatin fractionation. I found that like MSH6, MSH2 was enriched on *ERCC6L2*^{KO} cell's chromatin under HU treatment but MSH3 was undetected in any of the chromatin bound samples (Figure 4-3B). I next aimed to validate if similar observations could be made in our other *ERCC6L2*^{KO} U-2-OS cell line and the *ERCC6L2*^{KO} RPE1 cell lines. As shown in Figure 4-3C-D, MSH6 is enriched in the chromatin bound fraction of all *ERCC6L2*^{KO} cell lines under HU treatment.

In these experiments, I also assessed the levels of other DDR factors including pRPA (s4/8) and RAD52. Phosphorylation of RPA on serine 4 and 8 is mediated predominantly by DNA-PKcs and is crucial for repair of replication associated DSBs (229). Thus pRPA (s4/8) is a good indicator of DSB formation. Notably, levels of pRPA (s4/8) are markedly different between all three cell lines. While the first *ERCC6L2*^{KO} U-2-OS cell line has less pRPA (s4/8) than WT cells (Figure 4-3A), the second appears to have similar levels to the WT cells (Figure 4-3C). In contrast, the *ERCC6L2*^{KO} RPE1

cells appear to have higher levels (Figure 4-3D). Differences between pRPA (s4/8) levels seem to be independent of chromatin bound RPA levels, at least in the U-2-OS cell lines, since I also verified that less chromatin associated RPA accumulation was occurring in the second set of *ERCC6L2*^{KO} U-2-OS cells (Supp. Figure 7-3). Further investigation is required to fully understand these differences and whether they are related to ERCC6L2 status.

I also assessed levels of RAD52 which is another ssDNA binding DDR protein. RAD52 has been shown to be important for restart of collapsed replication forks by mediating break induced replication (BIR) and its absence results in increased DNA damage under HU-induced replication stress (65). Although it appears that there is slightly less chromatin bound RAD52 in the *ERCC6L2*^{KO} cells, RAD52 is modified by ATR in response to replication stress (Figure 4-3A) (65). It is possible that there is less modified RAD52 in the *ERCC6L2*^{KO} cells. Interestingly, in the other *ERCC6L2*^{KO} U-2-OS and RPE1 cell lines, there appears to be less modified RAD52 (Figure 4-3C-D). It is not currently known how ATR mediated phosphorylation affects RAD52 function but it appears to be induced specifically in response to replication stress (65). Again, with the current information, it is not immediately clear how loss of ERCC6L2 could impact RAD52 function and it therefore warrants further investigation.

Given that MSH6 is enriched on chromatin in all three *ERCC6L2*^{KO} cells, I next aimed to investigate if it was occurring at centromeres specifically. To address this, I attempted IF microscopy to look for MSH6 colocalisation with centromeric markers. However, due to extensive antibody issues, I was unable to assess this reliably. In the future, a proximity ligation assay could be performed with MSH6 and a centromeric protein such as CENP-A to determine centromeric occupancy. While it will be useful to find out if MSH6 enrichment is occurring specifically at centromeres, or on a global

basis, it is an intriguing result, especially in the context of the observed limited ssDNA formation. It remains to be elucidated whether these two observations are linked, and if so whether they are a cause or a consequence of one another.

Interestingly, MSH6 along with other MMR proteins have been implicated in several DDR contexts, outside of the traditional MMR pathway. Many of these non-canonical roles are in pathways that have also been linked to ERCC6L2 function. For example, MMR proteins play critical roles during CSR where they process the nicked DNA, following cytidine deamination, into DSBs (84). MSH2/6 are also recruited rapidly to DSB sites where they counteract recombination between homeologous sequences (247–249). Given the similarity between the contexts in which MSH6 and ERCC6L2 are known to function as well as the enhanced enrichment of chromatin bound MSH6 in *ERCC6L2*^{KO} cells, I performed immunoprecipitation to test if ERCC6L2 associates with this protein *in vivo*. I overexpressed full-length ERCC6L2 and three truncated variants (Figure 4-4A) with an N-terminal yellow fluorescent protein (YFP) tag in HEK293T cells. I also used a YFP-tagged nuclear localisation sequence (NLS) of the SV40 Late T-antigen as a negative control. I isolated the YFP-tagged proteins using GFP-Trap affinity resin and performed a western blot against MSH6 on the eluates.

As shown in Figure 4-4B, a band corresponding to the expected size of MSH6 was detected to varying degrees in all the ERCC6L2 eluates and not in the NLS control. Interestingly, the interaction is relatively weak for full-length ERCC6L2. However, for the C-terminus and the VIGSSK fragment, the interaction is considerably stronger. Given that the ERCC6L2 C-terminus is the most poorly expressed out of the four ERCC6L2 constructs, it is unlikely that the strength of interaction is due to differences between expression levels. I also probed for the endonuclease PMS2 which functions downstream of MSH2/6 in the MMR pathway (Figure 4-4B). I failed to detect PMS2 in

any of the ERCC6L2 eluates, suggesting that association of ERCC6L2 with MSH6 is unlikely to be occurring in the context of active MMR.

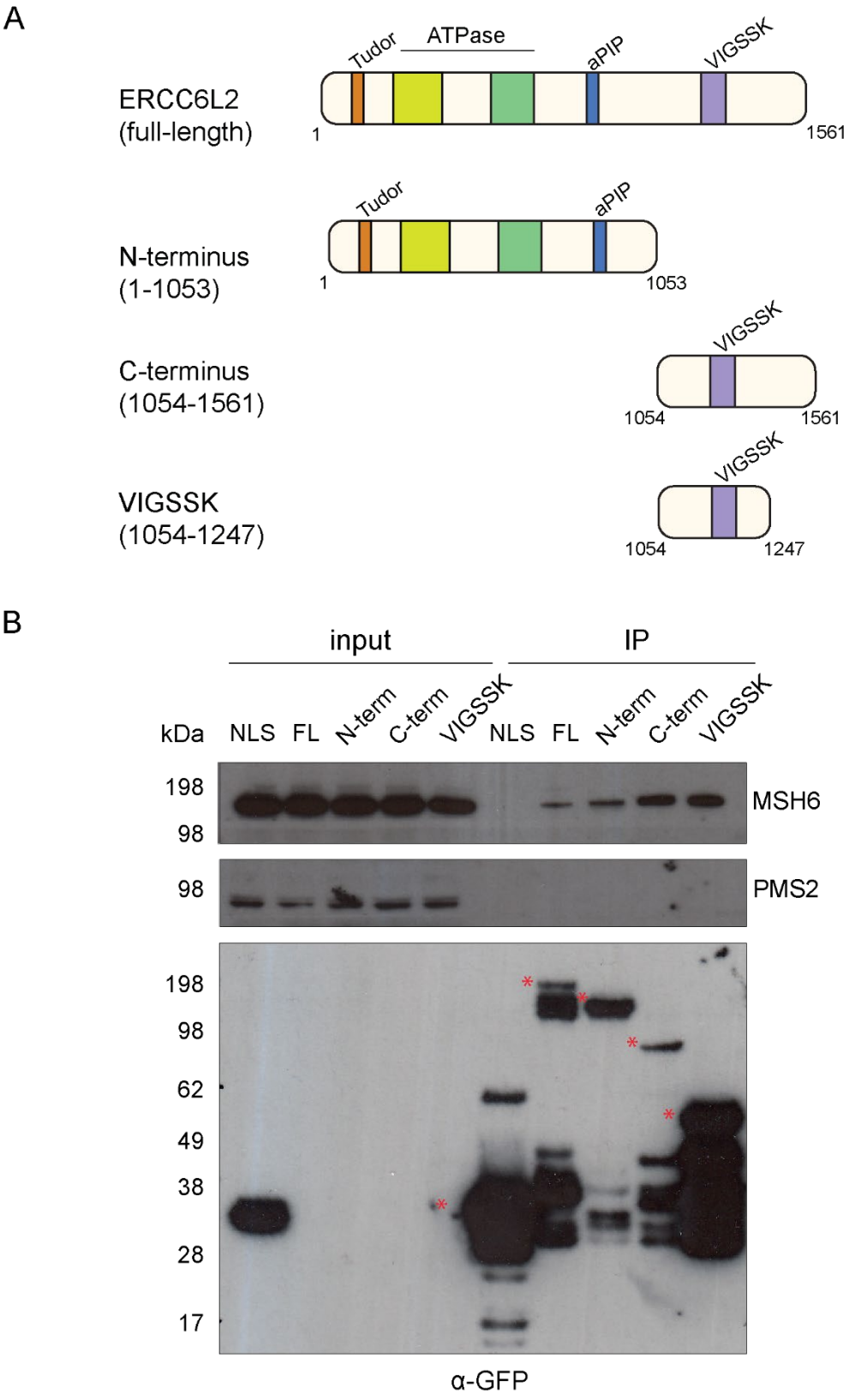


Figure 4-4: Overexpressed ERCC6L2 immunoprecipitates with MSH6

A. Schematic of ERCC6L2 variants used in the GFP-IP. B. Western blot analysis against MSH6, PMS2 and GFP in inputs and eluates from GFP immunoprecipitation performed on HEK293T cells expressing YFP-tagged ERCC6L2 variants (indicated by red stars). NLS = nuclear localisation sequence of the SV40 Late T-antigen.

4.2. Discussion

In this chapter, I set out to explore a possible role for ERCC6L2 in DNA replication and the response to replication stress. Firstly, I showed that *ERCC6L2*^{KO} U-2-OS and RPE1 cells exhibit mild sensitivity to the replication inhibitor HU (Figure 4-1). Further investigation revealed that *ERCC6L2*^{KO} cells likely form less ssDNA than WT cells under HU treatment (Figure 4-2A-B) via native BrdU and RPA32 retention analysis. However, ATR activation still occurs in the absence of ERCC6L2 (Figure 4-2D). When ATR inhibition was combined with HU treatment, ssDNA levels increased dramatically in the *ERCC6L2*^{KO} cells (Figure 4-2E). Under the same treatment conditions, I showed that the MMR repair factors MSH2/6 are enriched on chromatin in the *ERCC6L2*^{KO} cells (Figure 4-3). Finally, I showed that ectopically expressed ERCC6L2 associates with MSH6 *in vivo* (Figure 4-4).

ERCC6L2^{KO} U-2-OS cells display relatively mild sensitivity to HU when compared to the DSB inducers etoposide and phleomycin. This could suggest that if ERCC6L2 plays a role in either DNA replication or the response to replication stress, it may only operate in certain contexts. Full-length ERCC6L2 does not colocalise with PCNA in unperturbed cells. However, when the C-terminal half of ERCC6L2 (710-1561) is expressed, it overlaps with PCNA in S phase cells (Dr Chris Carnie). The PCNA interaction is also required for the recruitment of this fragment to sites of DNA damage. This suggests that the N-terminal region of ERCC6L2 is required for regulating recruitment of ERCC6L2 to sites of DNA synthesis, which may only be permitted during periods of stress. Since full-length ERCC6L2 is not visible at replication factories in unperturbed conditions but is instead found at centromeres, could suggest it is required for replication of these difficult-to-replicate regions in normal conditions.

Collectively, the results from the native BrdU assay and analysis of chromatin bound RPA32, suggest that *ERCC6L2*^{KO} cells have lower levels of ssDNA than WT cells following treatment with HU (Figure 4-2). As previously mentioned, ssDNA can form in a variety of contexts under HU-induced replication stress, thus it is important to consider how the absence of ERCC6L2 may limit such events. There is evidence to suggest that the presence of large dsDNA loops at centromeres can limit ssDNA formation (191). Thus, it is possible that similar complex topological arrangements constrain ssDNA formation at a subset of replication forks, not necessarily just at centromeres, where ERCC6L2 activity is required.

Replication fork reversal has been observed as a response to HU-induced replication stress (131). ssDNA stretches can be generated, in a controlled manner, at the regressed arms of reversed forks mediated by the concerted action of the helicase/nuclease DNA2 and WRN helicase (132). HR factors are associated with protecting the regressed arms of reversed forks from degradation. Since the role of ERCC6L2 in DSB repair is linked to c-NHEJ, it is unlikely that it is involved in protecting these structures from nucleolytic attack. However, three Snf2 ATPases have been shown to catalyse fork reversal *in vivo,* therefore it is possible that ERCC6L2 could perform a similar function (129,135,250,251). Local ATR suppression may mean that fork reversal occurs more frequently at centromeres since ATR activity normally suppresses fork reversal (144). Thus, even if ERCC6L2 is not responsible for catalysing fork reversal directly, it could be important for modifying the local chromatin environment at centromeres to facilitate the process.

Another possible, although speculative, explanation for the reduced levels of ssDNA markers is an increase in secondary DNA structures. It is possible that these arise in regions of repetitive ssDNA generated during replication uncoupling (191). It is also

possible that the increased presence of secondary DNA structures limits progression of the replicative helicase.

The actions of specialised helicases are required for replication through unusual DNA structures. For example, the RecQ helicases BLM and WRN are involved in G4 quadruplex unwinding at telomeres (252,253). RTEL1 is important for unwinding G4 quadruplexes at telomeres specifically but also during global DNA replication (161). The helicase and nuclease activity of DNA2 has been demonstrated to be required for removing secondary structures during centromeric DNA replication (254). Thus, it is possible that ERCC6L2 activity facilitates secondary structure resolution in concert with these factors. When HU was combined with ATR inhibition, a dramatic increase in RPA32 levels in *ERCC6L2*^{KO} cells was observed (Figure 4-2E). This may suggest that if there was an increase in secondary DNA structures, ATR protects them from degradation by structure-specific nucleases (255,256).

Chromatin fractionation revealed that under HU treatment, the MMR factors MSH2/6 are enriched on chromatin in *ERCC6L2*^{KO} cells (Figure 4-3). These proteins have been shown to be enriched on centromeric chromatin, with further accumulation observed under replication stress (191). Thus, this result may indicate that there is increased replication stress occurring at centromeres in *ERCC6L2*^{KO} cells. If so, it could imply that ERCC6L2 is involved directly in the resolution of secondary structures or that its activity is important upstream of secondary structure formation. However, it is equally possible that there is a global increase in secondary structures containing mispaired bases. While my efforts to ascertain if MSH2/6 enrichment occurs specifically at centromeres in *ERCC6L2*^{KO} cells were unsuccessful, in the future it could be beneficial to look for centromeric colocalisation with general markers of DNA damage. Chromatin immunoprecipitation could also be performed to test if MSH2/6 associate with

centromeric DNA sequences. Currently, the reasons behind MSH2/6 chromatin enrichment in the *ERCC6L2*^{KO} cells remain unknown. Nevertheless, this result does show that there is likely to be a general increase in mispaired DNA which can arise in a variety of DDR contexts.

In addition to secondary structures, mismatched DNA bases can occur within misaligned DNA sequences during DSB repair processes such as HR and SSA. In this context, MSH2/6 is proposed to promote rejection of non-identical sequences by recruiting helicases to unwind regions of heterology (247,257). Thus, it is possible that MSH2/6 could be in enriched due to an increase in misaligned sequences in *ERCC6L2*^{KO} cells. This is an interesting possibility considering that loss of ERCC6L2 causes higher levels of DNA end resection and increased use of HR for DSB repair (226).

The final result in this chapter demonstrated that ERCC6L2 associates with MSH6 *in vivo* (Figure 4-4). The interaction appears to be strongest with the C-terminus of ERCC6L2 when levels of expression are considered. Since this region has been shown to recruit to sites of laser damage, it could indicate that ERCC6L2 is recruited to DNA lesions where MSH6 is also important for repair. In the future, it will be interesting to investigate if these proteins act in the same pathway by depleting MSH6 from *ERCC6L2*^{KO} cells.

In this chapter, I sought to investigate if ERCC6L2 plays a role in the response to replication stress. My work suggests that loss of ERCC6L2 imparts some degree of sensitivity to HU-induced replication stress, which is associated with reduced levels of ssDNA markers and MMR factor chromatin enrichment. In the future, it would be useful to perform DNA fibres in the presence of replication stress to assess parameters such

as fork speed and asymmetry as well as inter-origin distance. Further investigation is required to understand if the role of ERCC6L2 during S phase is entirely distinct from its proposed function in DSB repair and whether it relates specifically to centromeres or is important on a genome-wide basis.

5. Investigating ERCC6L2's VIGSSK domain

5.1. Results

The C-terminal half of ERCC6L2 is critical to its cellular function as it is responsible for its recruitment to DNA damage, as well as its centromere localisation. Past work from our group revealed that a specific fragment within the C-terminus, ERCC6L2$^{(1054\text{-}1247)}$, is rapidly recruited to sites of damage and colocalises with centromeres (Dr Chris Carnie). This fragment contains an uncharacterised globular domain, recognised by the Pfam and NCBI conserved domain databases, called the VIGSSK domain (217,218,258). This domain is only found in ERCC6L2 homologs and does not obviously resemble domains found in any other proteins (217). This strongly suggests that this region is important and specific to the cellular role of ERCC6L2. It is still unclear how the VIGSSK domain contributes to the function of ERCC6L2, so I was motivated to investigate its biochemical properties and search for potential interaction partners.

Analysis of the primary amino acid sequence of the VIGSSK fragment (ERCC6L2$^{(1054\text{-}1247)}$) revealed several patches of positively charged residues (Figure 5-1A). We therefore hypothesised that the VIGSSK fragment may be able to bind DNA. To investigate this, I firstly expressed and purified the VIGSSK fragment from *E. coli*. The VIGSSK fragment was expressed with an N-terminal 6x His tag to allow for western blot validation.

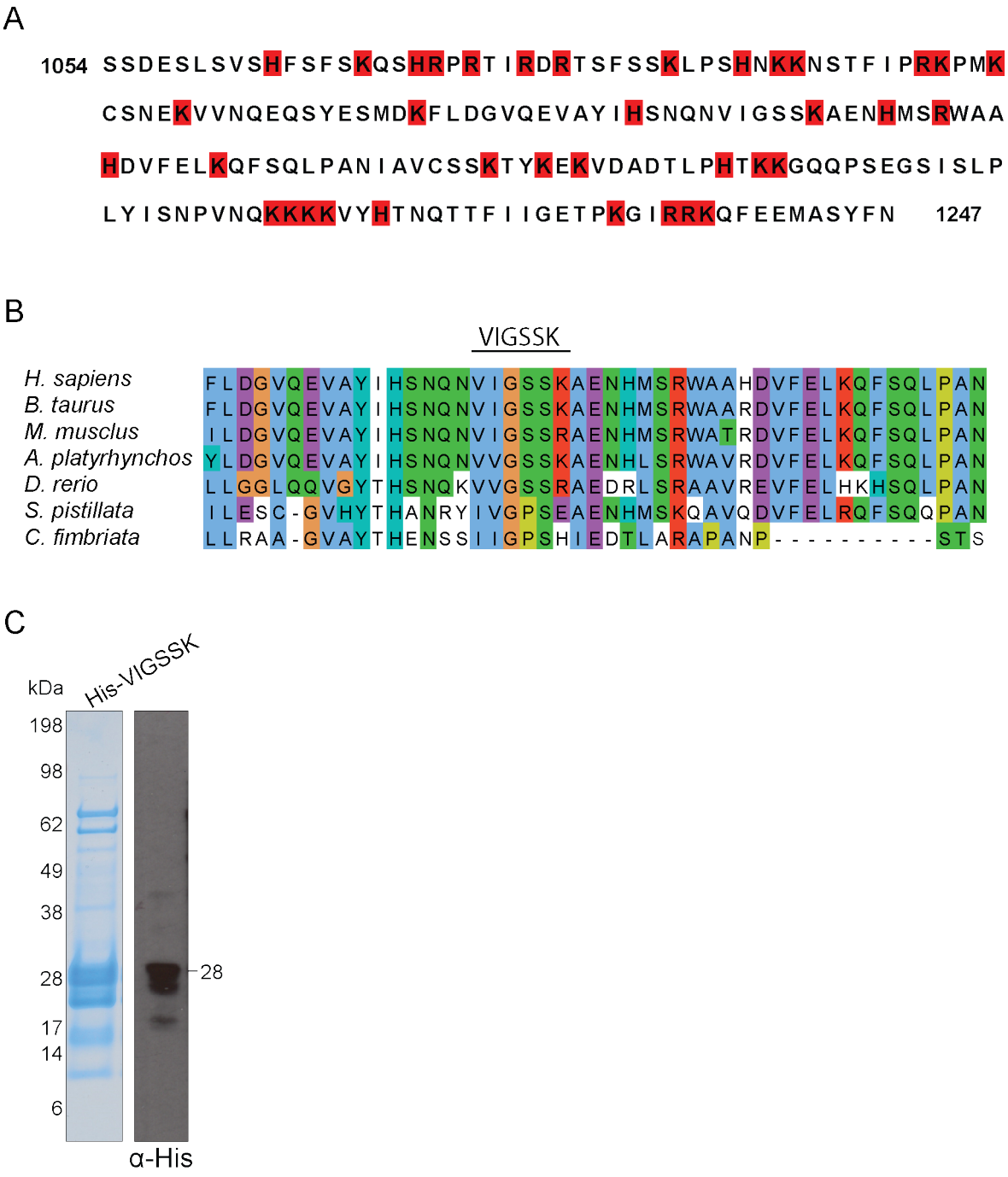


Figure 5-1: Purification of the His-VIGSSK fragment

A. Primary amino acid sequence of the VIGSSK fragment with positively charged residues highlighted in red. B. Multiple sequence alignment showing the highly conserved residues within the core of ERCC6L2 VIGSSK domain across various species. *H. sapiens* – aa1122-1168, *B. taurus* – aa1121-1167, *M. musculus* – aa1101-1147, *A. platyrhynchos* – aa1113-1159, *D. rerio* – aa871-917, *S. pistillata* – aa1146-1191 and *C. fimbriata* – aa847-882. Alignment was generated using Jalview with ClustalX colouring. C. Coomassie stained SDS-PAGE gel of recombinant His-VIGSSK fragment and corresponding anti-His western blot. The His-VIGSSK fragment was expressed and purified from *E. coli* using affinity and size exclusion chromatography.

After successfully purifying the His-VIGSSK fragment, I sought to test whether this region of ERCC6L2 is indeed capable of binding DNA. Thus, I performed an electrophoretic mobility shift assay (EMSA) using radiolabelled ssDNA and dsDNA substrates.

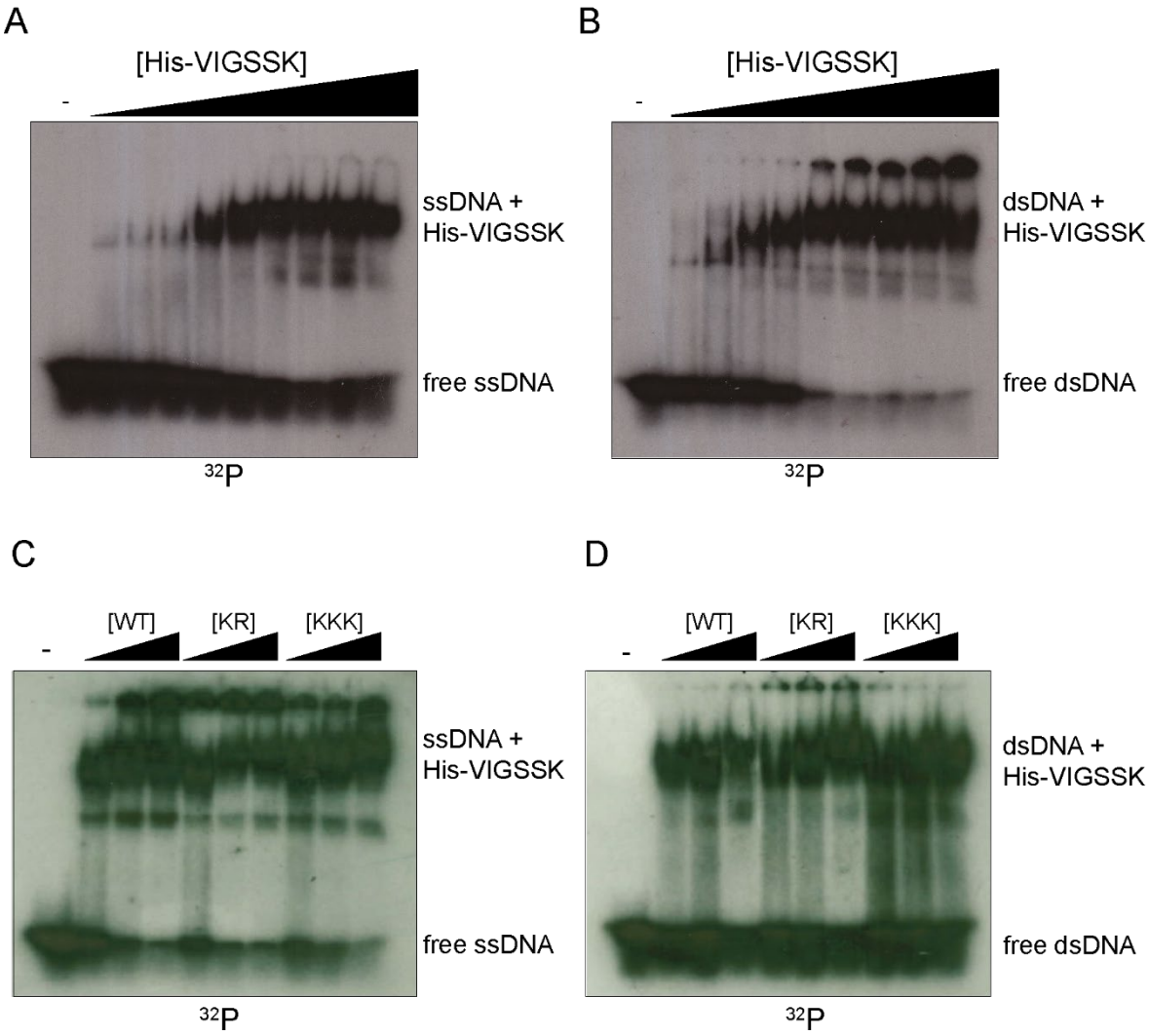


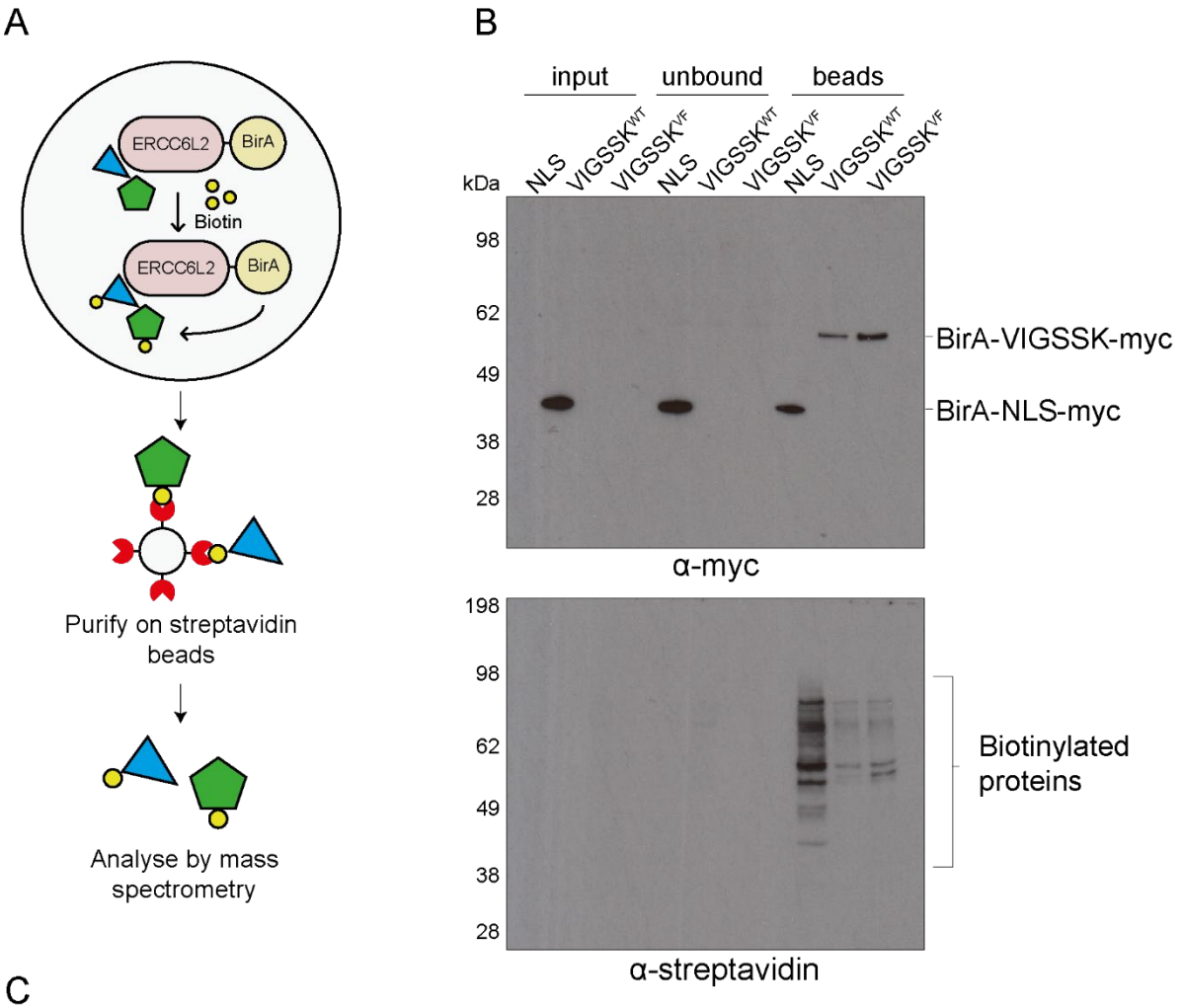
Figure 5-2: ERCC6L2's VIGSSK fragment binds ssDNA and dsDNA

A. EMSA with the His-VIGSSK fragment and radiolabelled ssDNA. His-VIGSSK fragment was incubated with increasing concentrations of protein (2.7, 4.1, 6.2, 9.2, 13.3, 20.8, 31, 41 and 68 μM) for 30 minutes at room temperature. Reaction mixtures were resolved by native PAGE and visualised by autoradiography. B. As in A, for dsDNA. C. EMSA with WT, K1143A R1150A (KR) and K1213A K1214A K1215A (KKK) His-VIGSSK fragments and radiolabelled ssDNA. His-VIGSSK fragments were incubated with increasing concentrations of protein (5, 10 and 20 μM). D. As in C, for dsDNA.

A range of concentrations of the His-VIGSSK fragment were incubated with either ssDNA or dsDNA and reactions were resolved by native PAGE and analysed by autoradiography. Figure 5-2A-B shows that the His-VIGSSK fragment of ERCC6L2 can bind both ssDNA and dsDNA *in vitro*. As the concentration of the protein increases, the shift, representing the formation of the nucleoprotein complex, becomes more pronounced with concominant reduction in the substrate DNA.

To further investigate the signifcance of this interaction, I aimed to identify key residues within the VIGSSK fragment responsible for mediating the DNA binding. I generated two mutants of the His-VIGSSK fragment: a double mutant (K1143A and R1150A) and a triple mutant (K1213A, K1214A amd K1215A). I expressed and purified these mutants alongside their WT counterpart and performed the EMSA. However, DNA binding was not abrogated by either set of mutations under my experimental conditons (Figure 5-2C-D). In the future, it would be useful to obtain the crystal structure of the VIGSSK fragment to gain insight into the DNA binding surface. While the observation that the VIGSSK fragment can bind DNA is an interesting one, until we identify a mutant which abrogates binding, its difficult to investigate the signifcance of this activity *in vivo*.

Next, I was keen to search for potential protein interactors of the VIGSSK fragment since it is unlikely that a DNA interaction would be solely responsible for driving the recruitment of this fragment to damage sites or centromeres. To investigate this, I performed BioID in HEK293T cells given that this would capture transient interactions that may be difficult to detect with a standard IP experiment. BioID involves fusing a protein of interest, in this case the VIGSSK fragment, to a promiscuous biotin ligase (BirA) that will modfiy endogenous proteins that come into close, although not necessarily direct, contact (Figure 5-3A) (259). Biotinylated proteins can then be isolated by affinity purification and analysed by mass spectrometry.



	Repeat 1			Repeat 2		
Protein ID	**NLS Peptides**	**VIGSSK Peptides**	**Coverage (%)**	**NLS Peptides**	**VIGSSK Peptides**	**Coverage (%)**
ERCC6L2	0	33	10.2	0	29	10.1
CENP-Q	0	5	28.4	0	4	24.3
Securin	0	5	31.7	0	3	21.8
PJA1	0	3	6.5	0	4	12
ZWINT	0	3	12.6	0	3	20.9
CENP-U	1	3	10.5	2	9	38.8
SENP3	0	3	6.3	3	7	16.9
RBM14	0	3	5.2	4	8	24.5

Figure 5-3: BioID with ERCC6L2's VIGSSK fragment

A. Schematic showing steps of BioID process. B. Myc and streptavidin western blots of Streptavidin pulldowns performed on nuclear extracts from HEK293T cells expressing the WT or a mutant (VF) VIGSSK fragment. NLS = nuclear localisation sequence of the SV40 Late T-antigen. C. Table of top candidate proteins from mass spectrometry analysis of two independent experiments. Coverage (%) relates to the amount of protein covered by the peptides identified.

The WT VIGSSK fragment and a potential binding mutant were overexpressed with the biotin ligase BirA fused to their N-termini. The proteins were also expressed with a C-terminal myc tag to assess expression by western blotting (Figure 5-3B). I prepared nuclear extracts from the cell lysates and pulled down the biotinylated proteins using streptavidin beads (Figure 5-3B). Samples were subsequently submitted for mass spectrometry analysis. Proteins were subject to on bead tryptic digestion and analysed by liquid chromatography with tandem mass spectrometry (LC-MS/MS) at the Advanced Proteomics Facility, Dept. of Biochemistry, University of Oxford.

In both experimental repeats, the top potential interaction partner for the VIGSSK fragment was centromere protein Q (CENP-Q) (Figure 5-3C). CENP-Q is a member of the constitutive centromere associated network (CCAN) which makes up the inner kinetochore (172). Proteins within the CCAN assemble at centromeres throughout interphase in a highly regulated manner. CENP-Q, specifically, exists in the CENP-O subcomplex along with CENP-O, CENP-P, CENP-R and CENP-U (260). CENP-U was also among the top candidates in my BioID screens (Figure 5-3C).

Securin was another relatively strong hit in both repeats. Securin is best known for its critical role during chromosome segregation, where it inhibits degradation of the cohesin complex (261,262). Securin has also been shown to interact with the c-NHEJ factor KU, which has interesting implications given the proposed role of ERCC6L2 in DSB repair (263). Furthermore, there is some evidence to suggest that depleting Securin from human cells impacts the degree of end processing during end joining (264).

Since CENP-Q was the strongest hit in each repeat, I chose to investigate whether the VIGSSK fragment could interact directly with this protein. To test this, I expressed and

purified CENP-Q from *E. coli* with an N-terminal GST-tag (Figure 5-4A). After successfully purifying GST-CENP-Q, I tested its ability to bind the VIGSSK fragment. To do this, I immobilised GST-CENP-Q on glutathione sepharose beads to use as bait for the recombinant His-VIGSSK fragment (Figure 5-1C), using GST as a negative control. The His-VIGSSK fragment was incubated with the GST proteins for at least 2 hours after which the beads were extensively washed. The beads were then boiled in LDS buffer and eluted proteins were analysed by an anti-His western blot.

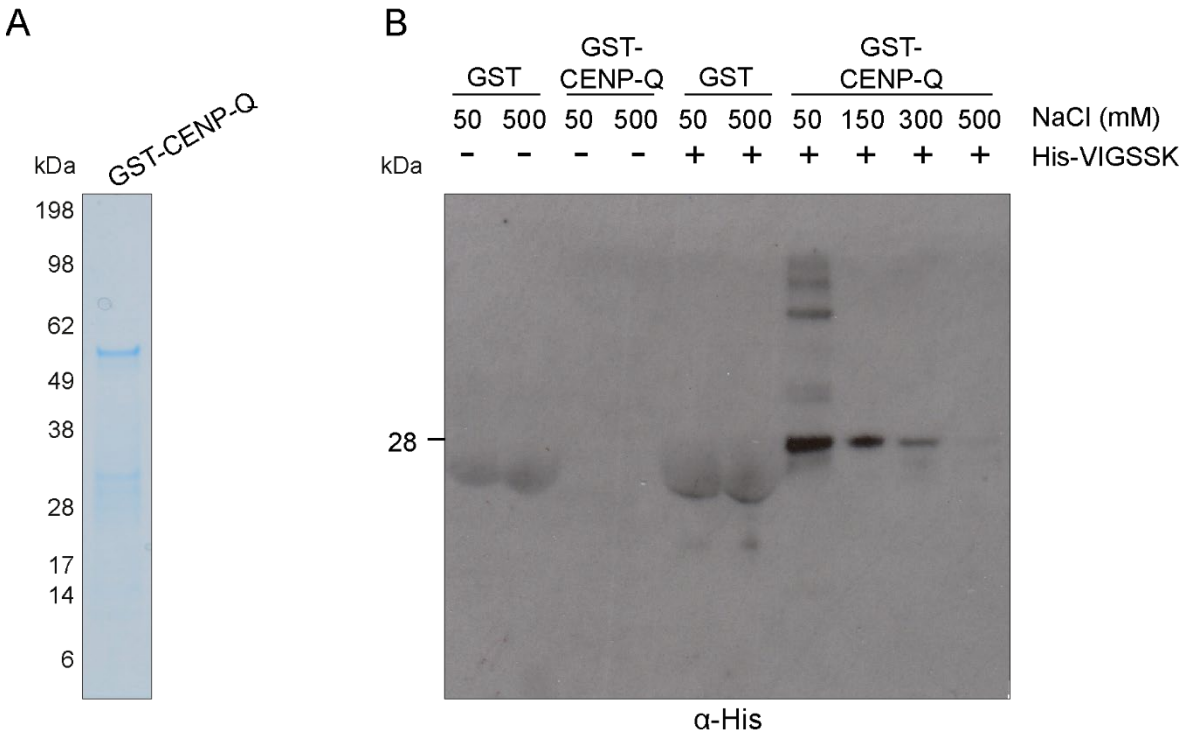


Figure 5-4: ERCC6L2's VIGSSK fragment interacts with CENP-Q

A. Coomassie stained SDS-PAGE gel of recombinant GST-CENP-Q. GST-CENP-Q was expressed and purified from *E. coli* using affinity chromatography. B. GST-CENP-Q was immobilised on glutathione sepharose beads and incubated with equal amounts of the recombinant His-VIGSSK fragment in various salt concentrations. Following incubation, beads were washed extensively. Samples were boiled in LDS loading buffer and analysed by an anti-His western blot.

As shown in Figure 5-4, the His-VIGSSK fragment was detected in the GST-CENP-Q sample but not in the GST negative control. Furthermore, the interaction is quite significantly weakened with increasing salt concentration. Crucially, the pulldown was

performed in the presence of benzonase, therefore DNA is not involved in mediating the interaction. This suggests that the His-VIGSSK fragment is retained through a direct interaction with CENP-Q. However, it will be important in the future to test different GST-tagged proteins to assess specificity of the interaction.

Shortly after these experiments were performed, a study was published revealing that while most of the residues in CENP-Q are involved in mediating interactions with other kinetochore proteins, the first 67 amino acids are involved in microtubule binding during mitosis (186). Thus, I reasoned that during interphase these residues may have other binding partners, potentially ERCC6L2. I expressed and purified two truncations of CENP-Q (CENP-$Q^{(1-67)}$ and CENP-$Q^{(68-268)}$) (Figure 5-5A), fused to GST and tested their ability to bind the His-VIGSSK fragment as before.

A

CENP-Q N-terminal fragment (aa1-67)

MSGKANASKKNAQQLKRNPKRKKDNEEVVLSENKVRNTVKK
NKNHLKDLSSEGQTKHTNLKHGKTAA

CENP-Q C-terminal fragment (aa68-268)

SKRKTWQPLSKSTRDHLQTMMESVIMTILSNSIKEKEEIQY
HLNFLKKRLLQQCETLKVPPKKMEDLTNVSSLLNMERARDK
ANEEGLALLQEEIDKMVETTELMTGNIQSLKNKIQILASEV
EEEEERVKQMHQINSSGVLSLPELSQKTLKAPTLQKEILAL
IPNQNALLKDLDILHNSSQMKSMSTFIEEAYKKLDAS

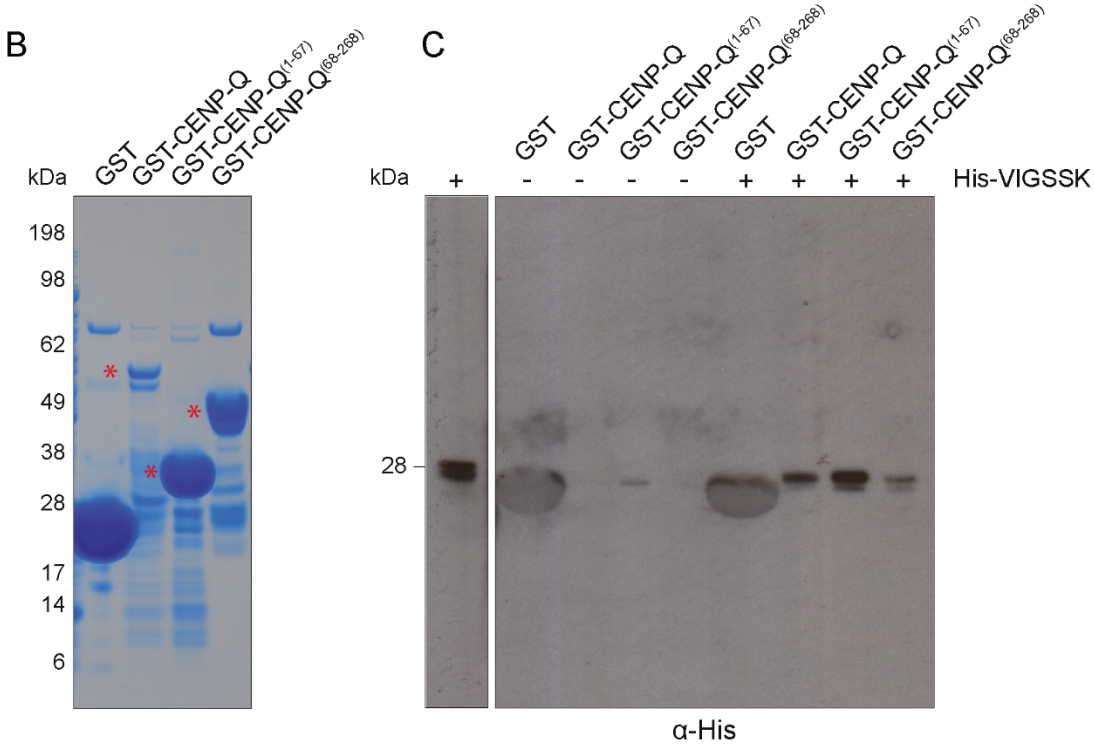


Figure 5-5: ERCC6L2's VIGSSK fragment interacts with CENP-Q's N-terminus

A. Amino acid sequences of CENP-Q N and C-terminal truncations used in B. B. Coomassie stained SDS-PAGE gel of recombinant GST-CENP-Q variants. GST-CENP-Q variants were expressed and purified from *E. coli* using affinity chromatography. Red stars indicate bands corresponding to each protein. C. Pulldowns were performed as previously described in the presence of 150 mM NaCl. Equal amounts of beads were used for each sample.

As before, the His-VIGSSK fragment was retained in the full-length GST-CENP-Q sample (Figure 5-5C). A strong interaction was observed between the His-VIGSSK fragment and the N-terminal portion of CENP-Q (GST-CENP-$Q^{(1-67)}$). However, the

expression of GST-CENP-Q$^{(1-67)}$ was significantly higher than GST-CENP-Q (Figure 5-5B). The C-terminal portion of CENP-Q (GST-CENP-Q$^{(68-208)}$) also expressed better than full-length CENP-Q. However, the interaction between the His-VIGSSK fragment and GST-CENP-Q$^{(68-268)}$ was significantly weaker (Figure 5-5C). This suggests that the VIGSSK fragment of ERCC6L2 interacts predominantly with the N-terminus of CENP-Q.

5.2. Discussion

In this chapter I set out to explore the function of the uncharacterised VIGSSK domain in ERCC6L2 using biochemical approaches. Briefly, I showed that the VIGSSK fragment binds both ssDNA and dsDNA (Figure 5-2). To search for protein interactors of the VIGSSK domain, I performed a BioID screen that identified members of the CCAN CENP-O subcomplex CENP-Q and CENP-U (Figure 5-3). Among these candidates, CENP-Q was consistently the strongest interaction candidate. I showed that the VIGSSK fragment of ERCC6L2 interacts directly with CENP-Q (Figure 5-4). This interaction appears to be mediated largely through the N-terminus of CENP-Q (Figure 5-5).

The VIGSSK fragment of ERCC6L2 is recruited strongly to sites of laser damage and colocalises with centromeres (Dr Chris Carnie) (216). Given the unique DNA configurations which form at sites of DNA damage and centromeres, it would be useful to test DNA binding with different types of DNA structures to assess if the VIGSSK fragment displays affinity for a particular substrate. For instance, centromeric DNA has the propensity to form secondary structures, such as hairpins and loops (187). I performed preliminary EMSAs with these types of substrates but observed binding comparable to ssDNA and dsDNA (data not shown). Many of these substrates have

stretches of either ssDNA or dsDNA thus it may be difficult to discern different degrees of affinity via this method. The additional two members of the ERCC6 subfamily can also bind DNA. Interestingly, the ATPase activity of CSB (ERCC6) is stimulated by a variety of DNA structures which contain dsDNA (265). PICH (ERCC6L) also displays extremely high binding affinity for dsDNA (215).

In addition to its suggested role in DSB repair, ERCC6L2 has been proposed to be important for removing R-loops during transcriptional elongation (220). R-loops are DNA: RNA hybrid nucleic acid structures consisting of three elements: the nascent RNA transcript that has become reannealed to its DNA template, and the displaced non-template DNA strand. Thus, it would be interesting to test if the VIGSSK fragment shows enhanced binding to these types of substrates.

In the future, it would be insightful to obtain the crystal structure of the VIGSSK domain, preferably in complex with a DNA substrate, to determine the binding surface and residues involved in mediating its interaction with DNA. I was unable to generate any VIGSSK mutants with abrogated DNA binding. It may be more effective to generate N- and C-terminal truncations of the VIGSSK fragment to try and narrow down the minimal DNA binding region. It would also be interesting to test if the VIGSSK fragment can bind DNA in complex with nucleosomes, particularly centromeric nucleosomes. ERCC6L2 may be considered unlikely to participate in chromatin remodelling since it exists in a subfamily with CSB and PICH, both of which exhibit little to no remodelling activity (215,265). Although, there is some indication that the nucleosome remodelling abilities of these proteins could be enhanced in certain circumstances. Among the wider Snf2 ATPase superfamily, DNA binding modules are common accessory domains which are critical for stimulating and directing the ATPase activity (204).

While this result has provided some meaningful insight, the full significance of the ability of the VIGSSK fragment to bind DNA cannot be properly realised until full-length ERCC6L2 is obtained and characterised. I attempted to optimise a strategy for purification of recombinant full-length ERCC6L2 and a predicted catalytically dead mutant (K165R), together with a past member from our lab (Dr Marek Sebesta). Using the baculovirus expression system, the proteins were expressed with an N-terminal 6 x His-MBP (maltose-binding protein) tag and a C-terminal 3X FLAG tag to counter solubility and degradation issues, respectively. Following elution from FLAG affinity resin with FLAG peptide, I obtained a small amount of full-length ERCC6L2, but further purification steps resulted in significant dilution of the protein (Supp. Figure 7-4). Elution from the FLAG affinity beads also requires very high concentrations of FLAG peptide and we have concerns that this contaminating peptide may interfere with biochemical assays. While this purification strategy is promising, the final steps still require some optimisation before we can use the protein in *in vitro* assays. If this can be achieved, the ATPase activity of ERCC6L2 could then be tested. Furthermore, the influence of DNA binding on the catalytic activity could be directly investigated, to see if it stimulates the ATPase activity in a general or substrate-specific sense.

ERCC6L2 colocalises with centromeres throughout interphase via its C-terminus (Dr Chris Carnie). However, it is not understood how this portion of ERCC6L2 mediates this recruitment. My BioID screen revealed that the VIGSSK fragment of ERCC6L2 appears to associate with components in the CENP-O subcomplex of the CCAN (Figure 5-3). Importantly, I was able to validate a direct interaction between the strongest candidate CENP-Q and ERCC6L2's VIGSSK fragment.

The CENP-O complex is a distal component of the CCAN and does not make direct contact with CENP-A nucleosomes (186). During mitosis, the CENP-O complex is

important for chromosome congression and recruiting additional components of the kinetochore (184,260). However, ERCC6L2 only colocalises with centromeres throughout interphase, implying it does not participate in chromosome segregation. In interphase, the CCAN assembles onto CENP-A nucleosomes to form the inner kinetochore. However, it is not clear if these proteins have additional roles outside of kinetochore assembly. It is known that the CENP-O complex assembles during S phase. In U-2-OS cells, it has been shown that CENP-Q abundance at kinetochores is at its highest during S phase and drops dramatically as cells enter G2 (266).

Interestingly, a recent study by Balmus and colleagues, which performed a genome-wide screen for factors involved in mediating toxic c-NHEJ events during S phase, found CENP-Q relatively high on their list of candidates; 56th out of a total of 18424 proteins (267). Although CENP-Q was below the false discovery rate threshold in this screen, ERCC6L2 also appeared at a similar position on the list, ranking slightly higher than CENP-Q at 49th. CENP-Q was also the highest centromeric protein by a considerable margin (267).

While these observations should not be overinterpreted, it is undeniably curious considering the role of ERCC6L2 in promoting c-NHEJ. This study specifically investigated genes which, when inactivated, conferred resistance to the TOPO I inhibitor topotecan (TPT) in an *ATM*-null background. The absence of ATM causes a delay in HDR-mediated repair of the single-ended DSBs induced by TPT leading to toxic c-NHEJ. Among the top hits were core c-NHEJ factors such as ligase IV, XRCC4 and XLF, as well as components of the BRCA1-A complex, which would normally act to limit DNA resection (267). Notably, inactivation of 53BP1 or its effector proteins does not suppress TPT or PARP inhibitor sensitivity in the absence of ATM, demonstrating that the mechanism underlying resistance is different from that observed in *BRCA1*-

deficient backgrounds. Although, ATM signalling is required for 53BP1 recruitment at two-ended DSBs, it is not yet known whether this occurs at single-ended DSBs. Nevertheless, the resistance mechanism appears to be mainly linked to proteins that are involved in the ligation step or proteins that normally act to repress resection independent of 53BP1.

If inactivation of ERCC6L2 is indeed capable of mediating TPT resistance in the absence of ATM, it suggests that it plays a supportive role in the final steps of c-NHEJ or acts to suppress resection, most likely, in a 53BP1-independent manner. It is interesting to consider how loss of CENP-Q could mediate such an effect. Given that CENP-Q has no known catalytic activity, it is unlikely that its absence directly influences events such as DNA end-resection or ligation. This could imply that it is important for mediating interactions with proteins either to promote their recruitment or retention at these sites. In the future, it will be important to assess if CENP-Q depletion affects centromere localisation of the VIGSSK fragment or full-length ERCC6L2.

The possibility that CENP-Q promotes recruitment of ERCC6L2 to centromeres is interesting. However, we know that the final 300 residues of ERCC6L2$^{(1248\text{-}1561)}$, which are largely disordered, also colocalise with centromeres suggesting there are multiple layers to the recruitment of ERCC6L2 to centromeres (Dr Chris Carnie). The fact that the C-terminal half of ERCC6L2$^{(701\text{-}1561)}$ is also required for recruitment to sites of DNA damage could mean that there is overlap between methods of damage and centromere recruitment. This is especially interesting given that there is evidence that centromeres are enriched with DDR factors (191).

While there are still many unanswered questions surrounding the role of the VIGSSK domain in ERCC6L2, the work in this chapter has provided new insight into its potential

functions, namely its ability to bind DNA and its interaction with CENP-Q. I have also identified additional interaction candidates such as CENP-U and Securin which merit further investigation. These observations provide a solid foundation for future studies into the regulation and recruitment of the VIGSSK domain and full-length ERCC6L2.

6. Discussion

Throughout this project, I have uncovered a series of interesting results on both a phenotypic and biochemical level that have brought us closer to understanding the physiological role of ERCC6L2. First, my results relating to the role of ERCC6L2 in the DSB response show that *ERCC6L2*^{KO} cells accumulate in G2 phase after prolonged DNA damage induced by the radiomimetic phleomycin and the TOPO II inhibitor etoposide (Figure 3-1 and 3-2). Following release from phleomycin treatment, *ERCC6L2*^{KO} cells progressively accumulate higher levels of ssDNA markers, hinting towards an anti-resection and c-NHEJ promoting function for ERCC6L2. However, unlike anti-resection factor 53BP1 and its effector proteins, loss of ERCC6L2 does not rescue PARP inhibitor sensitivity in the absence of BRCA1 (Figure 3-5). *ERCC6L2*^{KO} cells exhibit mild sensitivity to the replication inhibitor HU (Figure 4-1). Under HU-induced replication stress, *ERCC6L2*^{KO} cells likely accumulate less ssDNA than WT cells (Figure 4-2). This is also associated with MSH2/6 chromatin enrichment (Figure 4-3). These MMR factors have been shown to accumulate on centromeric chromatin and there is evidence that MSH6 depletion reduces the efficiency of centromeric DNA replication without any observable impact on non-centromeric sequences (191). Finally, the VIGSSK domain, found in the C-terminus of ERCC6L2, can bind ssDNA and dsDNA *in vitro* (Figure 5-2) and interacts directly with the centromere protein CENP-Q (Figure 5-4 and 5-5). Exploring the role of ERCC6L2 on a cellular and biochemical level has yielded a breadth of interesting and novel results. This has expanded our current understanding of how ERCC6L2 contributes to the maintenance of genome stability while providing a strong foundation for exciting follow-up studies.

Collectively, these observations could imply that ERCC6L2 operates on a multifunctional basis and is not strictly confined to one pathway. Along with evidence from our group, it has emerged from recent studies that ERCC6L2 is likely to function primarily in the DSB response, but there are certain lines of evidence that suggest it may have additional roles (220,224–226). My own work has linked ERCC6L2 function to DNA replication, while others have proposed a role for ERCC6L2 in transcription-coupled nucleotide excision repair (220). The range of processes that ERCC6L2 has been implicated in suggests that, mechanistically, it could be involved in a process that is common to the repair of different types of DNA lesions.

The localisation of ERCC6L2 to centromeres adds another layer of complexity when it comes to considering its cellular role. Centromeres are unique genomic loci characterised by repetitive DNA sequences and condensed chromatin (268). ERCC6L2 localises to these regions in unperturbed cells throughout interphase suggesting that it is important for maintaining centromere stability. It is emerging that DNA replication and repair at centromeres are regulated in specialised ways to compensate for the inherent challenges that these regions present (191,254,268). My findings and related hypotheses regarding the role of ERCC6L2 at centromeres will hopefully serve as a platform for further studies and may help to reveal novel regulatory mechanisms that contribute to the maintenance of centromere stability.

Overall, the results from this study provide useful insights into the cellular role of ERCC6L2. It is not immediately obvious how ERCC6L2 operates, in a mechanistic sense, therefore I will consider the possibilities based on evidence from findings from our group and the data in published studies.

6.1. ERCC6L2 in the double-strand break response

ERCC6L2 was linked to the repair of DSBs in early studies that demonstrated *ERCC6L2*-deficient patient derived fibroblasts display sensitivity to IR and the radiomimetic phleomycin (216,217). These observations were corroborated by work from our group, which also showed that *ERCC6L2*^{KO} cells are sensitive to phleomycin and the TOPO II inhibitor etoposide (Dr Chris Carnie). Etoposide traps TOPO II complexes on DNA that are encountered by replication or transcription machineries and lead to DSB formation. At low doses of etoposide, DSBs are predominantly replication-associated (12). Under prolonged treatment with low doses of phleomycin and etoposide, *ERCC6L2*^{KO} cells accumulate in G2 phase which is associated with increased levels of DNA damage (Figure 3-1 and 3-2). I also found that *ERCC6L2*^{KO} cells show significantly higher levels of DSBs after short releases from acute treatments of phleomycin (Figure 3-3). This suggests that DSB repair could be delayed at early timepoints in *ERCC6L2*^{KO} cells. At this stage, we also had preliminary evidence showing that *ERCC6L2*-deficient cells do not show sensitivity to TOPO I or PARP inhibitors, suggesting that ERCC6L2 does not play a direct role in HR (Figure 3-5 and Supp. Figure 7-1). Thus, we speculated that it could be involved in c-NHEJ. This notion was further supported by the observation that following phleomycin treatment, markers of ssDNA accumulate to a greater extent in the absence of ERCC6L2, suggesting it could be involved in suppressing DNA end resection (Figure 3-4).

Three recent studies have since shown that *ERCC6L2*-deficiency results in c-NHEJ and CSR defects (224–226). While these studies and our work are in broad agreement about the general role of ERCC6L2 in DSB repair, there are some important differences that could have major implications about the specific nature of ERCC6L2's

function. Given that ERCC6L2 is a predicted DNA translocase, it is unlikely that it acts as a physical block to DNA end resection, similar to the functions performed by components of the Shieldin complex. This is supported by the observation that ERCC6L2 does not accumulate at damage induced foci like SHLD2, SHLD3 and REV7 (75,76,269,270). Furthermore, tolerance to etoposide and phleomycin treatment is dependent on the catalytic activity of ERCC6L2 (Dr Chris Carnie). Together, this implies that ERCC6L2 may participate in events that prevent or counteract DNA end resection, in a manner that is dependent on its ATPase activity.

ERCC6L2 as a 'core' c-NHEJ factor

Olivieri and colleagues demonstrated that *ERCC6L2*^{KO} DT40 cells are sensitive to etoposide and bleomycin, although the effect of *ERCC6L2*-deficiency is less profound than loss of core c-NHEJ factors XLF and XRCC4 (225). They showed that ERCC6L2 depletion from *XRCC4*^{KO} cells does not enhance sensitivity to bleomycin or etoposide, implying ERCC6L2 operates epistatically with XRCC4. The same effect was observed with *XLF*^{KO} cells under bleomycin treatment. However, under etoposide treatment, *ERCC6L2*^{KO} *XLF*^{KO} double mutant cell lines are slightly more sensitive than *XLF*^{KO} cells (225). This may suggest that ERCC6L2 functions independently of XLF in the repair of TOPO II-linked DSBs.

XRCC4 and XLF form filaments with ligase IV at DNA ends flanking DSB sites and function to bridge the break and promote ligation (84). Super resolution microscopy has revealed that these filaments can engage in either an end to end or side to side configuration (271). It is proposed that the filaments fluctuate between these two different types of pairing to facilitate end processing and promote efficient re-joining.

Recently, ERCC6L2 has been shown to interact with splicing factor proline- and glutamine-rich (SFPQ) (226). Based on biochemical observations, SFPQ, together with a related protein non-POU domain-containing octamer-binding protein (NONO), has been proposed to form filamentous structures during c-NHEJ, which does not involve XLF. *In vitro,* SFPQ and NONO can functionally replace XLF and enhance ligase IV activity (272). SFPQ and NONO are not recruited to break sites via interactions with core c-NHEJ factors and there is evidence to suggest that they bind DNA away from the break site. Thus, it has been hypothesised that they may be involved in mediating an alternative, adjacent DNA end pairing mechanism, independently of XLF, to promote XRCC4 and ligase IV mediated end joining. The recruitment of ERCC6L2 to sites of DNA damage has also been shown to be independent of KU and XLF (224).

Currently, this mechanism of end joining is entirely speculative and has not been fully demonstrated *in vitro* or observed *in vivo.* Nevertheless, it is tempting to speculate that ERCC6L2 could be involved in such a process. The proposed adjacent pairing mechanism and the distal DNA binding of SFPQ and NONO from DNA ends, would mean that the DNA ends would not be constrained. This could mean the action of additional proteins is required to direct efficient end joining or facilitate processing of DNA ends. Furthermore, the observation that ERCC6L2 and XLF operate independently, to some extent, under etoposide treatment could hint ERCC6L2 is particularly important for the repair of TOPO II-linked DSBs (225). The presence of the TOPO II cleavage complex at the 5’ end of DSBs is refractory to rapid end ligation and requires further processing before the break can be repaired (273,274). ERCC6L2 therefore could be involved in supporting such an event.

However, it is entirely possible that this could also be due to a DSB repair independent function of ERCC6L2. The functional redundancy between XLF and ERCC6L2 during VDJ recombination, observed by Liu et al, could support this notion (224).

Is the function of ERCC6L2 dependent on 53BP1?

Currently, evidence surrounding the involvement of ERCC6L2 in the 53BP1 dependent anti-resection axis is conflicting. Loss of 53BP1, Shieldin and CST complex components have been demonstrated to rescue PARP inhibitor sensitivity in *BRCA1*-deficient cells (75,96,235). 53BP1 suppresses DNA resection through direct antagonism of BRCA1 and by recruiting the Shieldin and CST complexes to limit DNA end resection. In the absence of 53BP1 and its effectors, resection can be sufficiently restored to facilitate HR mediated repair of DSBs.

I found that loss of ERCC6L2 does not mediate this phenotype (Figure 3-5) which is in line with observations from Olivieri and colleagues (225). We also have preliminary evidence to suggest that depletion of 53BP1 and components of the Shieldin and CST complexes from *ERCC6L2*^{KO} cells have an additive effect in terms of ssDNA accumulation following phleomycin treatment (Dr Dragana Ahel). Collectively, this suggests that ERCC6L2 operates outside the 53BP1 dependent axis of anti-resection, at least in human cells. This does not necessarily mean they operate at different DSBs; it is possible that they act at different stages of the resection process similar to what has been observed between the core c-NHEJ factor KU and SHLD2 (275).

In contrast, Francica and colleagues found that the absence of ERCC6L2 does suppress PARP inhibitor sensitivity in *BRCA1*-deficient mouse mammary cells and show that DNA resection and HR are partially restored (226). Based on the similarity between these observations and those relating to 53BP1 and its effector proteins, they

propose that ERCC6L2 is likely to function through the 53BP1 dependent anti-resection axis. However, PARP inhibitor resistance in *BRCA1*-deficient cells is not mediated solely through loss of 53BP1 and its effectors. Loss of anti-resection factor DYNLL1 rescues PARP inhibitor sensitivity in *BRCA1*-deficient cells in a 53BP1 independent manner (276). Although DYNLL1 interacts with 53BP1, the mechanism of resistance is related to its function in inhibiting MRE11 mediated resection, in a manner that is not dependent on 53BP1. Thus, the observation that ERCC6L2 loss rescues PARP inhibitor sensitivity in *BRCA1*-deficient cells does not mean it is necessarily 53BP1 dependent.

It should be noted that, unlike DYNLL1, ERCC6L2 has not appeared as a candidate in screens for proteins involved in mediating PARP inhibitor resistance in *BRCA1*-deficient human cells (75,237). Interestingly, another study which screened for proteins involved in mediating resistance to TOPO I inhibitors in HR-deficient cells, which identified ERCC6L2 and CENP-Q as potential candidates, was conducted in *ATM*-null mouse embryonic stem cells (267). So far, loss of ERCC6L2 has not been shown to suppress PARP or TOPO I inhibitor sensitivity in HR-deficient human cells, only in mouse cells. Human and mouse ERCC6L2 homologs are almost identical to each other suggesting that functionally they are conserved. It could be that additional layers of redundancy in human cells prevent profound effects upon loss of ERCC6L2. There are also differences in chromatin organization between human and mouse cells which may account for such differences (277,278). The level of ERCC6L2's involvement during DSB repair may be influenced by, as yet, undetermined factors like chromatin environment or DNA end configuration at DSBs.

At present, evidence regarding the involvement of ERCC6L2 in the 53BP1 dependent axis of anti-resection is conflicting. The phenotypic similarities between cells lacking

ERCC6L2 and 53BP1 and its effectors are insufficient to make such a conclusion. However, it is important to consider that inactivation or loss of core c-NHEJ factors, like ligase IV and XRCC4, is not associated with rescuing PARP inhibitor resistance in *BRCA1*-deficient cells in any instance.

Insights from class switch and VDJ recombination defects

In addition to the repair of pathological DSBs, 53BP1 and core c-NHEJ factors are involved in the restoration of programmed DSBs during class-switch and VDJ recombination (89). These two processes are critical to the adaptive immune response by generating a vastly diverse assortment of antibodies through recombination of different genetic elements. Although mechanistically distinct, class-switch and VDJ recombination both favour deletional recombination, where the intervening DNA between the recombining elements is lost, rather than inverted (89,90).

Liu and colleagues showed that loss of ERCC6L2 has been shown to reduce the efficiency of these processes and is associated with an increase in inversion events during CSR (224). Loss of 53BP1 significantly reduces CSR efficiency. However, Liu and co-workers demonstrated that absence of 53BP1 leads to significantly higher levels of DNA resection than loss of ERCC6L2. Thus, it is suggested that absence of ERCC6L2 impacts the orientation of DSB repair which, unlike 53BP1, is not associated with very high levels of resection (224).

Recent studies have shown that chromatin architecture strongly influences orientation-specific CSR and VDJ recombination. Two compelling studies from Zhang et al propose that cohesin-mediated loop extrusion is the mechanism that underpins alignment of distal recombination elements and promotes deletional recombination (91,92). This process causes torsional stress, which is countered by the activity of

TOPO II (279). Cohesin is also enriched at centromeres and its activity has been shown to generate topological stress during DNA replication in yeast, which must be relieved by TOPO II (280). It has been shown that the action of TOPO II inhibitors like etoposide generate a high percentage of DSBs within cohesin-associated loop anchor regions (279).

Given that *ERCC6L2*^{KO} cells exhibit profound sensitivity to etoposide, it is an intriguing prospect that ERCC6L2 could promote genetic stability specifically within complex topological chromatin environments like those generated by cohesin. This could also explain why ERCC6L2 accumulates at centromeres in unperturbed conditions. Interestingly, an ERCC6L2 interactor screen performed by Tummala and colleagues found cohesin components in their list of candidates, but these have yet to be validated (220).

The precise nature of the function of ERCC6L2 in the response to DSBs remains to be elucidated. There is broad agreement that ERCC6L2 promotes c-NHEJ likely through anti-resection activity, which is likely to be independent of 53BP1. Complementary biochemical studies are needed to test the ATPase activity of ERCC6L2 and understand how it is regulated. If ERCC6L2 is involved in influencing chromatin structure or DNA topology to promote DSB repair, then it is entirely possible that it exerts such a function in other DDR contexts.

6.2. ERCC6L2 in DNA replication

Most of the available data regarding the cellular function of ERCC6L2 points towards a role in the DSB response. Yet, there are certain observations that suggest ERCC6L2 has alternative functions. ERCC6L2 interacts with the sliding clamp PCNA, which is a central component of the DNA replication machinery (Dr Chris Carnie). It is possible to

reconcile this interaction with a role in DSB repair such as during gap-filling synthesis. However, such a role remains entirely speculative and thus far, the PCNA interaction has not been linked to the role of ERCC6L2 in the DSB response.

In unstressed conditions, ERCC6L2 is found at centromeres throughout interphase. If ERCC6L2 is a strict c-NHEJ factor, it is somewhat surprising to observe that it is retained at these regions during S phase. However, it is still unclear how DSB repair is regulated at centromeres in human cells. Nevertheless, centromeres are inherently fragile regions of the genome due to their compact chromatin and repetitive DNA. It is emerging that DNA replication is differentially regulated at these regions and requires the action of non-canonical replication factors (191,254).

The observations made in this work provide preliminary evidence that ERCC6L2 could play a role in DNA replication or the response to replication stress. I observed that U-2-OS and RPE1 *ERCC6L2*^{KO} cells are mildly sensitive to the DNA replication inhibitor HU in clonogenic survival assays (Figure 4-1). The main priority moving forward will be to test if HU sensitivity can be reversed with ectopic expression of ERCC6L2. It will also be important to test if these phenotypes can be rescued with the ATPase and PCNA binding mutant.

I further observed that levels of ssDNA markers are reduced in *ERCC6L2*^{KO} cells under prolonged HU treatment (Figure 4-2A-B). However, ssDNA levels are sufficient to activate the ATR checkpoint (Figure 4-2D). While ATR functions locally at stalled forks by engaging ssDNA generated through replication uncoupling, its downstream effector kinase CHK1 diffuses throughout the nucleus to suppress origin firing globally and promote cell cycle arrest (50). If the lower levels of ssDNA observed in *ERCC6L2*^{KO} cells are associated with replication uncoupling, either due to helicase blocking or

formation of secondary structures at ssDNA between the uncoupled polymerase and helicase, it may only be occurring at a subset of replication forks.

In such a scenario, it is possible that ATR is locally suppressed at some replication forks. However, due to widespread fork stalling, ATR is activated adequately to activate CHK1 and promote the global response to replication stress. If unusual DNA structures or replication intermediates do occur more frequently in the absence of ERCC6L2, they may require resolution by structure specific nucleases. However, the activity of these proteins may be prevented through CHK1-mediated CDK inhibition leading to their persistence (281). This may explain why I observe such a large increase in RPA when ATR inhibition was combined with HU treatment (Figure 4-2E).

Currently, I cannot draw a conclusion as to why ssDNA levels are lower in the absence of ERCC6L2 when DNA replication is inhibited. However, considering other available data relating to the centromeric localisation of ERCC6L2 and MSH2/6 chromatin enrichment, as well as the functional contexts of its proposed role in DSB repair, I will offer some speculative possibilities.

Does the function of ERCC6L2 relate to DNA topology?

Aze and colleagues demonstrated that ATR kinase activation is suppressed at centromeric chromatin to facilitate DNA replication at these regions (191). There is evidence that DNA is arranged into large, positively supercoiled dsDNA loops at these regions, which occurs in the wake of DNA replication and in a manner that is dependent on the action of TOPO I (191). Preventing the formation of these structures by TOPO I inhibition results in RPA loading and ATR activation at centromeres in unperturbed conditions, which stalls DNA replication (191). It is thought that the presence of positively supercoiled DNA may limit helicase unwinding.

Interestingly, the closely related Snf2 ATPase, PICH (ERCC6L) is capable of influencing DNA topology during centromeric ultrafine bridge resolution in mitosis. Together with TOPO IIIA, PICH catalyses formation of positively supercoiled DNA. It is proposed that PICH achieves this by catalysing DNA loop extrusion (213). There are additional Snf2 ATPases that are capable of catalysing loop formation, however, unlike PICH, their activities result in negative supercoiling of the DNA (282,283). It is therefore tempting to speculate that ERCC6L2 could be required for aiding replication through complex topological environments. The notion that it could potentially catalyse loop extrusion is particularly interesting given that a similar type of DNA transaction occurs during repair of programmed DSBs. In the future, it would be interesting to see if ERCC6L2 associates with any topoisomerases, particularly during S phase, by performing immunoprecipitation followed by mass spectrometry analysis in the presence of replication stress.

It is possible that ERCC6L2 does not relieve topological stress directly, yet its absence may lead to an increase through indirect means, which therefore limits ssDNA formation. Topological stress can be generated when the replisome approaches stable protein-DNA complexes. There is evidence from studies in yeast that kinetochore proteins can limit replication fork progression and require removal by specialised helicases (284). I demonstrated that the VIGSSK fragment of ERCC6L2 interacts with the constitutive centromere protein CENP-Q (Figure 5-4 and 5-5), which is a distal component of the inner kinetochore. Thus, another possible function for ERCC6L2 could involve displacing proteins that impede helicase progression. In effect, ERCC6L2 could act as a helicase processivity factor. Interestingly, Tummala and colleagues identified components of the MCM helicase in their ERCC6L2 interaction screen, but these remain to be validated (216).

ERCC6L2 in the resolution of DNA secondary structures

MSH6 has been implicated in the replication of centromeric DNA where it is proposed to facilitate the process by resolving secondary DNA structures (191). I observed that chromatin bound levels of MSH2/6 in *ERCC6L2*^{KO} cells under HU-induced replication stress are elevated (Figure 4-3). However, I was unable to ascertain if this enrichment was occurring specifically at centromeres, therefore the reasons behind these observations remain speculative.

MSH2/6 accumulation could relate to an increase in DNA secondary structures either across the genome or at particular regions, like centromeres, where secondary structures containing mismatched DNA bases are enriched. This could suggest that ERCC6L2 plays a direct role in resolving secondary DNA structures. Resolution of secondary DNA structures is normally associated with DNA helicases which use ATP hydrolysis to unwind DNA (161,253,254). Thus, it is important to consider how ERCC6L2 could contribute to such a process, given that it is predicted to have DNA translocase activity. It is possible that ERCC6L2 could use its ATPase activity to move a DNA structure into a context that is more amenable for resolution by helicases. Alternatively, it could alter the local chromatin environment or topology of DNA surrounding secondary structures to either prevent their formation or promote their resolution.

Interestingly, work from Tummala and colleagues showed that DNA:RNA hybrids accumulate in the absence of ERCC6L2 and propose that together with DNA-PK, ERCC6L2 is important for their removal to facilitate transcriptional elongation (220). DNA:RNA hybrids, also known as R-loops, are highly stable nucleic acid structures composed of a newly formed RNA transcript that has reannealed to the DNA template

and the displaced single-stranded non-template strand. R-loops can pose a major threat to genome stability as they can impede transcription and replication fork progression (285). Thus, it is possible that ERCC6L2 has a role in aiding progression of transcription as well as replication machinery by directly resolving different types of unusual nucleic acid structures. It is also possible that ERCC6L2 counteracts chromatin or DNA configurations, which would increase the tendency for such structures to form.

The ssDNA component of R-loops is also vulnerable to nucleolytic attack which can result in DSB formation and a source of aberrant recombination events (285). ERCC6L2 has been recently shown to interact with SFPQ, which in addition to its role in DSB repair, suppresses R-loop formation at telomeres during replication (226,272,286). It is proposed that together with its binding partner NONO, SFPQ is important for recruiting factors that prevent R-loop formation, ultimately to suppress recombination between telomeric repeats (286). Although this function of SFPQ appears to be distinct from its role in DSB repair and has not been linked to ERCC6L2, it is curious that ERCC6L2 has been independently implicated in a similar process and accumulates at centromeres, which like telomeres, are prone to illegitimate recombination. In the future, it would be useful to measure R-loop levels in our *ERCC6L2*^{KO} cells and if so, see if they occur at specific regions of the genome. It would also be interesting to assess how their levels are affected under different genotoxic treatments.

Collectively, the observations that MSH2/6 and R-loops accumulate in the absence of ERCC6L2, albeit under different types of genotoxic stress, could suggest that ERCC6L2 has a general role in the resolution of unusual DNA structures or preventing their formation.

Does ERCC6L2 have an S phase role related to DSB repair?

DSBs which arise from stalled or collapsed replication forks are single-ended which means that use of c-NHEJ for repair could result in toxic chromosomal fusions (96,267). However, c-NHEJ is still used for repair of two-ended DSBs which are sustained in S phase, therefore it is possible that ERCC6L2 could function in such a capacity during DNA replication. I found that ssDNA formation in *ERCC6L2*^{KO} cells is significantly lower than WT cells under prolonged HU-induced replication stress (Figure 4-2). If ERCC6L2 operates in an anti-resection capacity during S phase, whether at DSBs or replication intermediates such as reversed forks, I expected an increase in ssDNA.

The observation, therefore, that ssDNA levels are lower in the *ERCC6L2*^{KO} cells could hint that the HU sensitivity is unrelated to the role that ERCC6L2 plays in DSB repair, given that this contrasts somewhat with what is observed upon treatment of *ERCC6L2*^{KO} cells with DSB inducers. Although it is entirely speculative, the reduction in ssDNA in the absence of ERCC6L2 could also indicate that there is less DNA resection taking place. However, this would imply that ERCC6L2 has both pro- and anti-resection functions. In the future, it would be worth investigating if ERCC6L2 is differentially regulated throughout the cell cycle. For instance, ERCC6L2 may undergo different post-translational modifications such as phosphorylation, which could be easily analysed by mass spectrometry of ERCC6L2 preparations from different cell cycle stages.

Another possibility could be that ERCC6L2 is involved in remodelling replication intermediates to produce cleavage substrates for nucleases. Moving forward, it will be important to thoroughly assess the levels of DSBs following treatment with HU, through

a physical method. A recently developed microscopy technique called STRIDE (SensiTive Recognition of Individual DNA Ends) could be particularly useful since this would allow for direct detection of DSBs by IF, where genomic location of the breaks could also be assessed (287). Furthermore, it would be beneficial to measure ssDNA and DSB levels following release from HU treatment to assess if *ERCC6L2*^{KO} cells behave normally after fork stalling has been relieved. DNA fibre assays could also be performed to assess the ability of stalled replication forks to restart after prolonged arrest in the absence of ERCC6L2.

With the current evidence, it is still unclear whether ERCC6L2 has a unique function during global DNA replication or at centromeres. In the future, it is hoped that purification of ERCC6L2 will allow us to understand the nature of its activity. This would allow us to make more informed predictions regarding what kind of role ERCC6L2 could play in DNA replication.

6.3. *ERCC6L2*-associated IBMF

ERCC6L2-deficiency is associated with a unique type of inherited bone marrow failure (IBMF) (216,217,219,220,223). IBMF syndromes are a heterogeneous group of disorders that can arise due to a broad range of genetic mutations (222). Notably, BMF has been linked to the deregulation of nuclear processes such as DNA repair, telomere maintenance and ribosome biogenesis (222). Most of the clinical mutations identified in *ERCC6L2* result in C-terminal truncated variants that are unable to localise to sites of DNA damage or centromeres. Immune defects have not been reported in IBMF patients harbouring *ERCC6L2* mutations despite the CSR deficiency observed in mice (224). *ERCC6L2*-deficiency has therefore been linked to maintenance of genome stability.

With the current evidence from our work and published studies, it is probable that the DSB repair defect observed in *ERCC6L2*-deficient cells contributes to *ERCC6L2*-associated IBMF. The core c-NHEJ factor DNA ligase IV has also been identified as a genetic cause of IBMF (223). Components of the Shieldin and CST complex, namely REV7 and CTC1, have also been linked to the disease. Given that loss of these factors cause an increase in DNA end resection, similar to ERCC6L2, highlights a possible link between defective end processing and IBMF (288,289).

Mutations in other DNA repair genes like RTEL1 have been identified as genetic causes underlying certain BMF syndromes (223). While RTEL1 is crucial for telomere maintenance, it also plays a global role during DNA replication (161). As discussed, it is possible that ERCC6L2 operates in different DDR contexts, and therefore its loss could impact a number of nuclear processes. Currently, it is likely that *ERCC6L2*-deficiency stems from a DSB repair defect. However, the possible links between ERCC6L2 and the regulation of DNA replication and transcription raise the possibility that disease-causing mutations in *ERCC6L2* could deregulate additional nuclear events.

6.4. Concluding remarks

The findings from this project have provided useful observations regarding the role of ERCC6L2 in DSB repair. My data have also provided the basis for some exciting new lines of investigation regarding the biochemical properties of the VIGSSK domain as well as a possible role for ERCC6L2 in DNA replication. Collectively, these results could help to further our understanding regarding the molecular basis of *ERCC6L2*-associated IBMF.

Loss of ERCC6L2 renders cells sensitive to DSB induction by phleomycin and the TOPO II inhibitor etoposide (Dr Chris Carnie) (216,217,225). *ERCC6L2*^{KO} U-2-OS cells accumulate in G2 phase of the cell cycle, as a result of increased DNA damage under treatment with phleomycin and etoposide (Figure 3-1 and 3-2). I further showed that following release from phleomycin treatment, *ERCC6L2*^{KO} U-2-OS cells accumulate higher levels of ssDNA indicating a DNA resection defect (Figure 3-4). This supports observations made by Francica and colleagues, which showed that loss of ERCC6L2 restores DNA resection in *BRCA1*-deficient cells (226). However, in contrast to this study, I did not find that loss of ERCC6L2 rescues sensitivity to PARP inhibitor in the absence of BRCA1 (Figure 3-5). Although, my findings were in line with observations by Olivieri and colleagues (225). As discussed, it is not immediately clear what underlies these differences, although it could be linked to using different species and/or cell lines. This represents an intriguing point for future investigation. If the basis of these differences is related to the species of cell lines used, it could reveal that the influence and importance the role ERCC6L2 plays in maintaining genome stability is determined by species-specific chromatin organisation.

There are multiple lines of evidence pointing towards a role for ERCC6L2 as a c-NHEJ promoting factor. In addition to ssDNA accumulation following phleomycin treatment (Figure 3-4), I found that after a short release from acute phleomycin treatment, *ERCC6L2*^{KO} cells have significantly more damage (Figure 3-3). This suggests that ERCC6L2 functions in a fast-acting mode of DSB repair.

It has since emerged that *ERCC6L2*-deficient cells harbour defects in c-NHEJ and CSR (224–226). However, how exactly ERCC6L2 mechanistically contributes to this process is unknown. There is evidence to suggest that ERCC6L2 functions at DNA ends due to the proteins it is known to interact with, such as CYREN and DNA-PK

(220,238). The epistasis observed between ERCC6L2 and XRCC4/XLF under phleomycin treatment could also indicate that it is involved in promoting the latter steps of end-joining (225). However, the absence of ERCC6L2 does not confer the same degree of sensitivity to etoposide or phleomycin as observed upon loss of core c-NHEJ factors like ligase IV and XRCC4, thus it may only be relevant at certain types of DSBs. Given that ERCC6L2 localises to centromeres, one possibility is that the involvement of ERCC6L2 in repair is determined by the local chromatin environment.

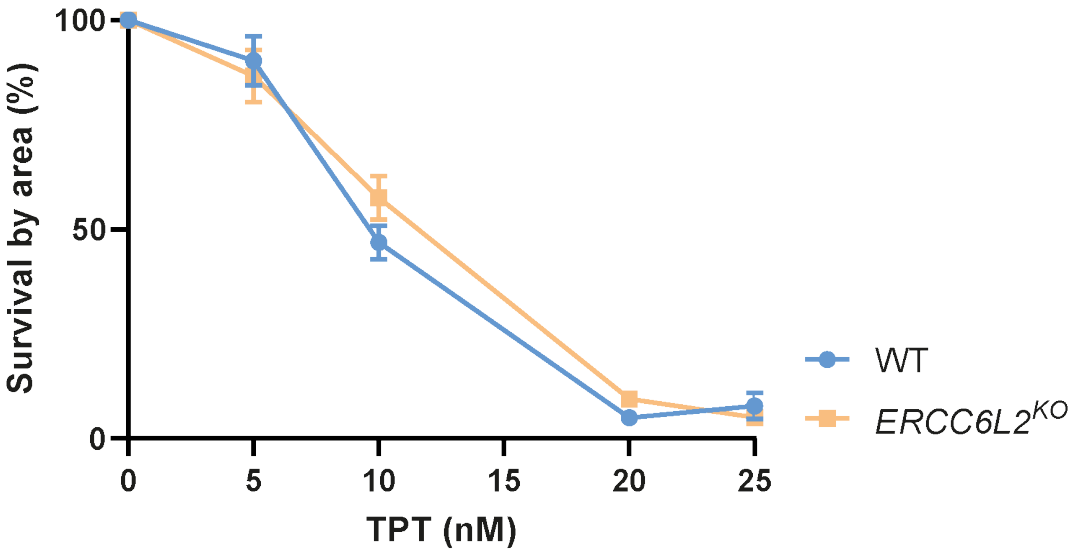
I explored a possible role for ERCC6L2 in DNA replication and revealed some interesting findings. *ERCC6L2*^{KO} cells appear to be mildly sensitive to HU-induced replication stress (Figure 4-1). However, future work is needed to fully verify that this phenotype is indeed arising from *ERCC6L2*-deficiency. Nevertheless, I found that *ERCC6L2*^{KO} cells have limited ssDNA formation following prolonged replication stress (Figure 4-2). Enrichment of mismatch repair factors including MSH6 was also observed on the chromatin of *ERCC6L2*^{KO} cells in the same conditions (Figure 4-3). My current hypothesis regarding a possible S phase role for ERCC6L2, is that it functions to facilitate replication through challenging chromatin environments, which may act as a barrier to helicase progression, potentially through resolving secondary structures, altering the DNA topology or removing proteins. While only biochemical data can confirm the nature of the mechanism of action of ERCC6L2, additional cellular work could be performed to provide more details about its potential S phase function. Digital droplet PCR, for example, could be used to assess replication timing at centromeres in *ERCC6L2*^{KO} cells (290). This would provide insight into whether ERCC6L2 has a specific role in the replication of centromeric DNA.

The observations regarding the uncharacterised C-terminal VIGSSK domain in ERCC6L2, namely that it can bind DNA and interact directly with the inner kinetochore

protein CENP-Q, provide interesting lines of further research. It would be useful to perform a comprehensive analysis of binding with a range of nucleic acid substrates. It will also be important to understand the significance of the interaction with CENP-Q and whether it is involved in mediating the recruitment of ERCC6L2 to centromeres.

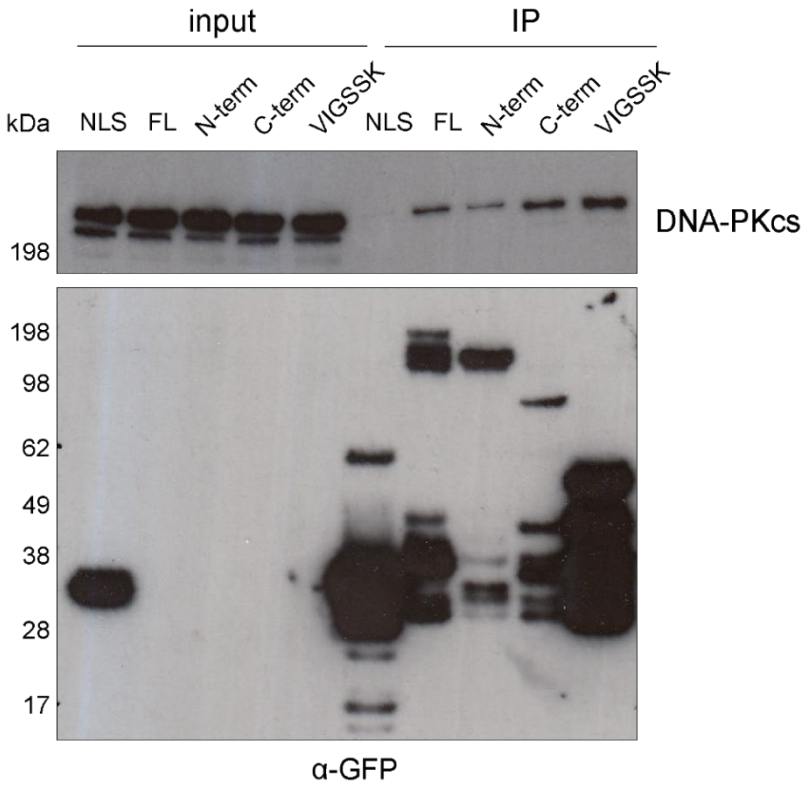
The work presented in this study has provided valuable insights into the physiological role of ERCC6L2. Understanding how ERCC6L2 functions mechanistically is now of utmost importance. Therefore, optimisation of an effective purification strategy for recombinant full-length ERCC6L2 is urgently needed, and we have already made significant steps towards this goal. Overall, my current hypothesis is that ERCC6L2 promotes a DNA transaction, such as DNA loop extrusion, chromatin remodelling or removal of non-histone proteins from DNA, which could be reconciled with roles in both DSB repair and DNA replication. While this remains a speculation, the sheer challenge of elucidating the cellular function of ERCC6L2 may be a simple reflection of its complicated molecular roles. Despite the challenging nature of this project, this work edges us closer to understanding how ERCC6L2 functions to preserve genome stability and how this role may relate to human disease.

7. Appendix



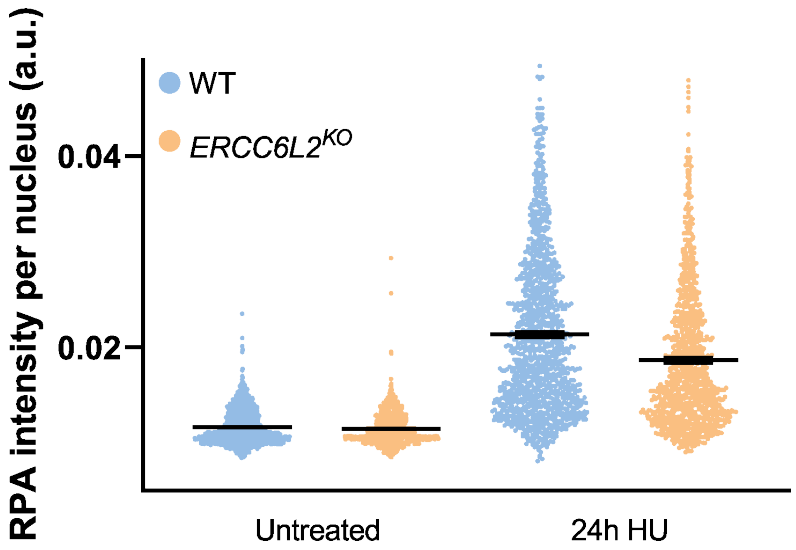
Supplementary figure 7-1: Clonogenic survival assay with topotecan

Clonogenic survival assay performed with WT and *ERCC6L2*^{KO} U-2-OS cell lines treated with topotecan (TPT). Cells were seeded in 6-well plates and treated with a range of TPT concentrations 24 hours later. Error bars represent SEM from three technical repeats.



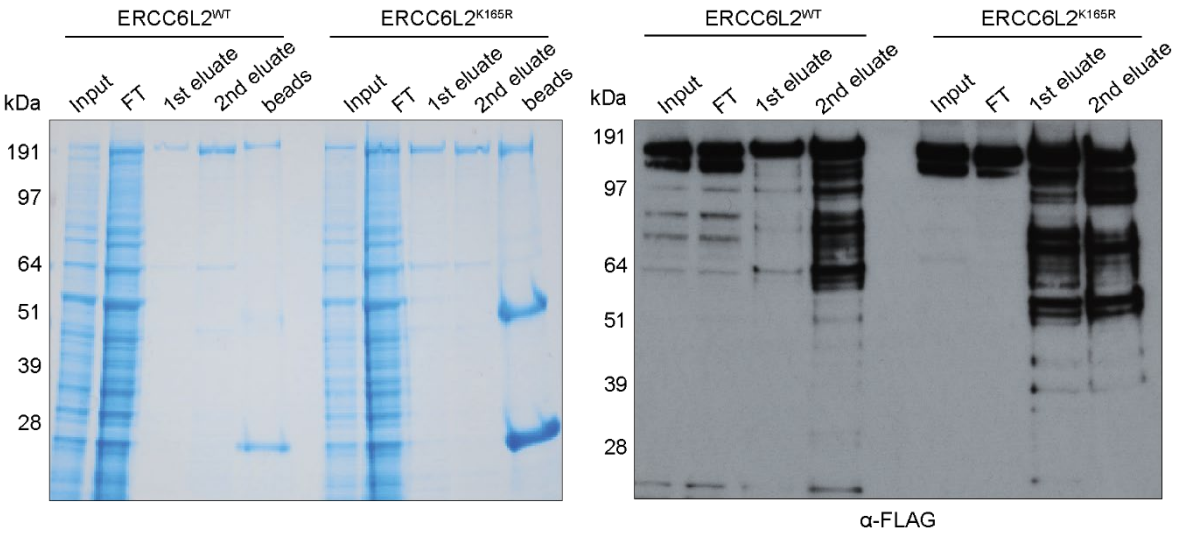
Supplementary figure 7-2: ERCC6L2 immunoprecipitates with DNA-PKcs

Western blot analysis against DNA-PKcs and GFP in inputs and eluates from GFP immunoprecipitation performed on HEK293T cells expressing the indicated YFP-tagged ERCC6L2 variants.



Supplementary figure 7-3: ssDNA formation is limited in *ERCC6L2*^{KO} cells under HU treatment

WT and *ERCC6L2*^{KO} U-2-OS cells were treated with 2 mM HU for 24 hours. Following pre-extraction and fixation, cells were immunostained with RPA32. Dot plots shows mean intensity of RPA per nucleus. Bars represent means with SEM. RPA32 intensity was measured in at least 1000 nuclei per condition.



Supplementary figure 7-4: Purification of ERCC6L2

Coomassie stained SDS-PAGE gel of fractions containing recombinant His-MBP-ERCC6L2^{WT}-FLAG and His-MBP-ERCC6L2^{K165R}-FLAG after elution from FLAG beads with FLAG peptide. After cell lysis and clarification, lysates were applied to FLAG beads and protein eluted 2 x 1 ml with 2.5 mg/ml FLAG peptide. Corresponding anti-FLAG western blot is also shown.

8. Bibliography

1. Giglia-Mari, G., Zotter A. and Vermeulen W. *DNA Damage Response.* Cold Spring Harb Perspect Biol, 2011. **3**: a000745

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

3. Chatterjee N. and Walker G.C. *Mechanisms of DNA damage, repair, and mutagenesis*. Environmental and Molecular Mutagenesis, 2017. **58**(5): p. 235–63.

4. Basu A.K. *DNA Damage, Mutagenesis and Cancer*. International Journal of Molecular Sciences, 2018. **19**(4): p. 970-83.

5. Lord C.J. and Ashworth A. *BRCAness revisited*. Nat Rev Cancer, 2016. **16**(2): p. 110–20.

6. De Bont R. and van Larebeke N. *Endogenous DNA damage in humans: a review of quantitative data*. Mutagenesis, 2004. **19**(3): p. 169–85.

7. Lindahl T. *Instability and decay of the primary structure of DNA*. Nature, 1993. **362**(6422): p. 709–15.

8. Krokan H.E. and Bjørås M. *Base Excision Repair.* Cold Spring Harb Perspect Biol, 2013. **5**(4): a012583.

9. Marteijn J.A. et al. *Understanding nucleotide excision repair and its roles in cancer and ageing.* Nature Reviews Molecular Cell Biology, 2014. **15**(7): p. 465–81.

10. de Boer J. and Hoeijmakers J.H.J. *Nucleotide excision repair and human syndromes*. Carcinogenesis, 2000. **21**(3): p. 453–60.

11. Vos S.M. et al. *All tangled up: how cells direct, manage and exploit topoisomerase function*. Nature Reviews Molecular Cell Biology, 2011. **12**(12): p. 827–41.

12. Tammaro M. et al. *Replication-Dependent and Transcription-Dependent Mechanisms of DNA Double-Strand Break Induction by the Topoisomerase 2-Targeting Drug Etoposide*. PLOS ONE, 2013. **8**(11): e79202.

13. Pommier Y. *Topoisomerase I inhibitors: camptothecins and beyond*. Nature Reviews Cancer, 2006. **6**(10): p. 789–802.

14. Zhao B. et al. *The molecular basis and disease relevance of non-homologous DNA end joining.* Nature Reviews Molecular Cell Biology, 2020. **21**(12): p. 765–81.

15. Krejci L. et al. *Homologous recombination and its regulation*. Nucleic Acids Research, 2012. **40**(13): p. 5795–818.

16. Mansour W.Y., Rhein T. and Dahm-Daphi J. *The alternative end-joining pathway for repair of DNA double-strand breaks requires PARP1 but is not dependent upon microhomologies*. Nucleic Acids Res, 2010. **38**(18): p. 6065–77.

17. Ochs F. et al. *53BP1 fosters fidelity of homology-directed DNA repair.* Nature Structural & Molecular Biology, 2016. **23**(8): p. 714–21.

18. Bhargava R., Onyango D.O. and Stark J.M. *Regulation of Single-Strand Annealing and its Role in Genome Maintenance*. Trends in Genetics, 2016. **32**(9): p. 566–75.

19. Ray Chaudhuri A. and Nussenzweig A. *The multifaceted roles of PARP1 in DNA repair and chromatin remodelling.* Nat Rev Mol Cell Biol, 2017. **18**(10): p. 610–21.

20. Abbotts R. and Wilson D.M. *Coordination of DNA single strand break repair*. Free Radical Biology and Medicine, 2017. **107**: p. 228–44.

21. Lord C.J. and Ashworth A. *PARP inhibitors: Synthetic lethality in the clinic*. Science, 2017. **355**(6330): p. 1152–8.

22. Pommier Y., O'Connor M.J. and Bono J de. *Laying a trap to kill cancer cells: PARP inhibitors and their mechanisms of action*. Science Translational Medicine, 2016. **8**(362): p. 17-23.

23. Huang Y. and Li L. *DNA crosslinking damage and cancer - a tale of friend and foe*. Transl Cancer Res, 2013. **2**(3): p. 144–54.

24. Deans A.J. and West S.C. *DNA interstrand crosslink repair and cancer.* Nature Reviews Cancer, 2011. **11**(7): p. 467–80.

25. Moldovan G-L. and D'Andrea A.D. *How the Fanconi Anemia Pathway Guards the Genome.* Annual Review of Genetics, 2009. **43**(1): p. 223–49.

26. Kastan M.B. and Bartek J. *Cell-cycle checkpoints and cancer.* Nature, 2004. **432**(7015): p. 316–23.

27. Satyanarayana A. and Kaldis P. *Mammalian cell-cycle regulation: several Cdks, numerous cyclins and diverse compensatory mechanisms.* Oncogene, 2009. **28**(33): p. 2925–39.

28. Blackford A.N. and Jackson S.P. *ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response*. Molecular Cell, 2017. **66**(6): p. 801–17.

29. Lee J-H. and Paull T.T. *ATM Activation by DNA Double-Strand Breaks Through the Mre11-Rad50-Nbs1 Complex.* Science, 2005. **308**(5721): p. 551–4.

30. McKinnon P.J. *ATM and the Molecular Pathogenesis of Ataxia Telangiectasia.* Annual Review of Pathology: Mechanisms of Disease, 2012. **7**(1): p. 303–21.

31. Shiloh Y. and Ziv Y. *The ATM protein kinase: regulating the cellular response to genotoxic stress, and more.* Nature Reviews Molecular Cell Biology, 2013. **14**(4): p. 197–210.

32. Stucki M. et al. *MDC1 Directly Binds Phosphorylated Histone H2AX to Regulate Cellular Responses to DNA Double-Strand Breaks.* Cell, 2005. **123**(7): p. 1213–26.

33. Kolas N.K. et al. *Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase.* Science, 2007. **318**(5856): p. 1637–40.

34. Doil C. et al. *RNF168 Binds and Amplifies Ubiquitin Conjugates on Damaged Chromosomes to Allow Accumulation of Repair Proteins.* Cell, 2009. **136**(3): p. 435–46.

35. Ahn J-Y. et al. *Threonine 68 Phosphorylation by Ataxia Telangiectasia Mutated Is Required for Efficient Activation of Chk2 in Response to Ionizing Radiation*. Cancer Res, 2000. **60**(21): p. 5934–6.

36. Falck J. et al. *The ATM–Chk2–Cdc25A checkpoint pathway guards against radioresistant DNA synthesis*. Nature, 2001. **410**(6830): p. 842–7.

37. Canman C.E. et al. *Activation of the ATM Kinase by Ionizing Radiation and Phosphorylation of p53*. Science, 1998. **281**(5383): p. 1677–9.

38. Yuan J. et al. *USP10 Regulates p53 Localization and Stability by Deubiquitinating p53*. Cell, 2010. **140**(3): p. 384–96.

39. Abbas T. and Dutta A. *p21 in cancer: intricate networks and multiple activities*. Nature Reviews Cancer, 2009. **9**(6): p. 400–14.

40. Aubrey B.J. et al. *How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression?* Cell Death & Differentiation, 2018. **25**(1): p. 104–13.

41. Chaturvedi P. et al. *Mammalian Chk2 is a downstream effector of the ATM-dependent DNA damage checkpoint pathway*. Oncogene, 1999. **18**(28): p. 4047–54.

42. Zeman M.K. and Cimprich KA. *Causes and consequences of replication stress.* Nature Cell Biology, 2014. **16**(1): p. 2–9.

43. Helmrich A. et al. *Transcription-replication encounters, consequences and genomic instability*. Nature Structural & Molecular Biology, 2013. **20**(4): p. 412–8.

44. Gaillard H., García-Muse T. and Aguilera A. *Replication stress and cancer*. Nature Reviews Cancer, 2015. **15**(5): p. 276–89.

45. Flynn R.L. and Zou L. *ATR: a master conductor of cellular responses to DNA replication stress*. Trends in Biochemical Sciences, 2011. **36**(3): p. 133–40.

46. Iyer D.R. and Rhind N. The Intra-S Checkpoint Responses to DNA Damage. Genes, 2017. **8**(2): 74; doi:10.3390

47. Cortez D. et al. *ATR and ATRIP: Partners in Checkpoint Signaling.* Science, 2001. **294**(5547): p. 1713–6.

48. Haahr P. et al. *Activation of the ATR kinase by the RPA-binding protein ETAA1*. Nature Cell Biology, 2016. **18**(11): p. 1196–207.

49. Zhao H. and Piwnica-Worms H. *ATR-Mediated Checkpoint Pathways Regulate Phosphorylation and Activation of Human Chk1.* Molecular and Cellular Biology, 2001. **21**(13): p. 4129–39.

50. Toledo L.I. et al. *ATR Prohibits Replication Catastrophe by Preventing Global Exhaustion of RPA*. Cell, 2013. **155**(5): p. 1088–103.

51. Saldivar J.C., Cortez D. and Cimprich K.A. *The essential kinase ATR: ensuring faithful duplication of a challenging genome*. Nat Rev Mol Cell Biol, 2017. **18**(10): p. 622–36.

52. Lecona E. and Fernandez-Capetillo O. *Targeting ATR in cancer.* Nature Reviews Cancer, 2018. **18**(9): p. 586–95.

53. Brown E.J. and Baltimore D. *ATR disruption leads to chromosomal fragmentation and early embryonic lethality*. Genes Dev, 2000. **14**(4): p. 397–402.

54. Gottlieb T.M. and Jackson S.P. *The DNA-dependent protein kinase: Requirement for DNA ends and association with Ku antigen*. Cell, 1993. **72**(1): p. 131–42.

55. Chan D.W. et al. *Autophosphorylation of the DNA-dependent protein kinase catalytic subunit is required for rejoining of DNA double-strand breaks.* Genes Dev, 2002. **16**(18): p. 2333–8.

56. Löbrich M. and Jeggo P. *A Process of Resection-Dependent Nonhomologous End Joining Involving the Goddess Artemis*. Trends in Biochemical Sciences, 2017. **42**(9): p. 690–701.

57. Goodarzi A.A. et al. *DNA-PK autophosphorylation facilitates Artemis endonuclease activity.* The EMBO Journal, 2006. **25**(16): p. 3880–9.

58. Buisson R. et al. *Distinct but Concerted Roles of ATR, DNA-PK, and Chk1 in Countering Replication Stress during S Phase*. Molecular Cell, 2015. **59**(6): p. 1011–24.

59. O'Driscoll M. et al. *DNA Ligase IV Mutations Identified in Patients Exhibiting Developmental Delay and Immunodeficiency.* Molecular Cell, 2001. **8**(6): p. 1175–85.

60. Murray J.E. et al. *Mutations in the NHEJ Component XRCC4 Cause Primordial Dwarfism.* The American Journal of Human Genetics, 2015. **96**(3): p. 412–24.

61. Sonoda E. et al. *Rad51-deficient vertebrate cells accumulate chromosomal breaks prior to cell death*. The EMBO Journal, 1998. **17**(2): p. 598–608.

62. Prakash R. et al. *Homologous Recombination and Human Health: The Roles of BRCA1, BRCA2, and Associated Proteins*. Cold Spring Harb Perspect Biol, 2015. **7**(4): a016600.

63. Decottignies A. *Alternative end-joining mechanisms: a historical perspective*. Front Genet, 2013. **4**: 10.3389.

64. Hanscom T. and McVey M. *Regulation of Error-Prone DNA Double-Strand Break Repair and Its Impact on Genome Evolution.* Cells, 2020. **9**(7):1657-80.

65. Sotiriou S.K. et al. *Mammalian RAD52 Functions in Break-Induced Replication Repair of Collapsed DNA Replication Forks.* Mol Cell, 2016. **64**(6): p. 1127–34.

66. Symington L.S. and Gautier J. *Double-Strand Break End Resection and Repair Pathway Choice*. Annual Review of Genetics, 2011. **45**(1): p. 247–71.

67. Chapman J.R., Taylor M.R.G. and Boulton S.J. *Playing the End Game: DNA Double-Strand Break Repair Pathway Choice*. Molecular Cell, 2012. **47**(4): p. 497–510.

68. 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.

69. Panier S. and Boulton S.J. *Double-strand break repair: 53BP1 comes into focus.* Nature Reviews Molecular Cell Biology, 2014. **15**(1): p. 7–18.

70. Chapman J.R. et al. *RIF1 Is Essential for 53BP1-Dependent Nonhomologous End Joining and Suppression of DNA Double-Strand Break Resection.* Molecular Cell, 2013. **49**(5): p. 858–71.

71. Munoz I.M et al. *Phospho-epitope binding by the BRCT domains of hPTIP controls multiple aspects of the cellular response to DNA damage.* Nucleic Acids Research, 2007. **35**(16): p. 5312–22.

72. Mirman Z. et al. *53BP1–RIF1–shieldin counteracts DSB resection through CST- and Polα-dependent fill-in*. Nature, 2018. **560**(7716): p. 112–6.

73. Zimmermann M. et al. *53BP1 Regulates DSB Repair Using Rif1 to Control 5' End Resection.* Science, 2013. **339**(6120): p. 700–4.

74. Callen E. et al. *53BP1 Mediates Productive and Mutagenic DNA Repair through Distinct Phosphoprotein Interactions*. Cell, 2013. **153**(6): p. 1266–80.

75. Dev H. et al. *Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells.* Nature Cell Biology, 2018. **20**(8): p. 954–65.

76. Ghezraoui H. et al. *53BP1 cooperation with the REV7–shieldin complex underpins DNA structure-specific NHEJ*. Nature, 2018. **560**(7716): p. 122–7.

77. Barazas M. et al. *The CST Complex Mediates End Protection at Double-Strand Breaks and Promotes PARP Inhibitor Sensitivity in BRCA1-Deficient Cells*. Cell Rep, 2018. **23**(7): p. 2107–18.

78. Sartori A.A. et al. *Human CtIP promotes DNA end resection.* Nature, 2007. **450**(7169): p. 509–14.

79. Escribano-Díaz C. et al. *A Cell Cycle-Dependent Regulatory Circuit Composed of 53BP1-RIF1 and BRCA1-CtIP Controls DNA Repair Pathway Choice*. Molecular Cell, 2013. **49**(5): p. 872–83.

80. Grundy G.J. et al. *One ring to bring them all—The role of Ku in mammalian non-homologous end joining*. DNA Repair, 2014. **17**: p. 30–8.

81. McElhinny S.A.N. et al. *Ku Recruits the XRCC4-Ligase IV Complex to DNA Ends.* Mol Cell Biol, 2000. **20**(9): p. 2996–3003.

82. Liu S. et al. *C-terminal region of DNA ligase IV drives XRCC4/DNA ligase IV complex to chromatin*. Biochemical and Biophysical Research Communications, 2013. **439**(2): p. 173–8.

83. Ropars V. et al. *Structural characterization of filaments formed by human Xrcc4–Cernunnos/XLF complex involved in nonhomologous DNA end-joining*. PNAS, 2011. **108**(31): p. 12663–8.

84. Hammel M. et al. *XRCC4 Protein Interactions with XRCC4-like Factor (XLF) Create an Extended Grooved Scaffold for DNA Ligation and Double Strand Break Repair.* J Biol Chem, 2011. **286**(37): p. 32638–50.

85. Ochi T. et al. *PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair.* Science, 2015. **347**(6218): p. 185–8.

86. Ma Y. et al. *A Biochemically Defined System for Mammalian Nonhomologous DNA End Joining.* Molecular Cell, 2004. **16**(5): p. 701–13.

87. McElhinny S.A.N. et al. *A Gradient of Template Dependence Defines Distinct Biological Roles for Family X Polymerases in Nonhomologous End Joining.* Molecular Cell, 2005. **19**(3): p. 357–66.

88. Chi X., Li Y. and Qiu X. *V(D)J recombination, somatic hypermutation and class switch recombination of immunoglobulins: mechanism and regulation.* Immunology, 2020. **160**(3): p. 233–47.

89. Stavnezer J., Guikema J.E.J. and Schrader C.E. *Mechanism and Regulation of Class Switch Recombination*. Annual Review of Immunology, 2008. **26**(1): p. 261–92.

90. Schatz D.G. and Ji Y. *Recombination centres and the orchestration of V(D)J recombination*. Nature Reviews Immunology, 2011. **11**(4): p. 251–63.

91. Zhang X. et al. *Fundamental roles of chromatin loop extrusion in antibody class switching.* Nature, 2019. **575**(7782): p. 385–9.

92. Zhang Y. et al. *The fundamental role of chromatin loop extrusion in physiological V(D)J recombination.* Nature, 2019. **573**(7775): p. 600–4.

93. Rooney S., Chaudhuri J. and Alt F.W. *The role of the non-homologous end-joining pathway in lymphocyte development*. Immunological Reviews, 2004. **200**(1): p. 115–31.

94. Bothmer A. et al. *Regulation of DNA End Joining, Resection, and Immunoglobulin Class Switch Recombination by 53BP1.* Molecular Cell, 2011. **42**(3): p. 319–29.

95. Bothmer A et al. *53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination.* Journal of Experimental Medicine, 2010. **207**(4): p. 855–65.

96. Bunting S.F. et al. *53BP1 Inhibits Homologous Recombination in Brca1-Deficient Cells by Blocking Resection of DNA Breaks*. Cell, 2010. **141**(2): p. 243–54.

97. Mao Z. et al. *DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells.* Cell Cycle, 2008. **7**(18): p. 2902–6.

98. Symington L.S. *Mechanism and regulation of DNA end resection in eukaryotes*. Critical Reviews in Biochemistry and Molecular Biology, 2016. **51**(3): p. 195–212.

99. Liu J. et al. *Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA*. Nature Structural & Molecular Biology, 2010. **17**(10): p. 1260–2.

100. Zhang F. et al. *PALB2 Links BRCA1 and BRCA2 in the DNA-Damage Response*. Current Biology, 2009. **19**(6): p. 524–9.

101. San Filippo J., Sung P. and Klein H. *Mechanism of Eukaryotic Homologous Recombination.* Annual Review of Biochemistry, 2008. **77**(1): p. 229–57.

102. Barber L.J. et al. *RTEL1 Maintains Genomic Stability by Suppressing Homologous Recombination.* Cell, 2008. **135**(2): p. 261–71.

103. Simandlova J. et al. *FBH1 Helicase Disrupts RAD51 Filaments in Vitro and Modulates Homologous Recombination in Mammalian Cells.* J Biol Chem, 2013. **288**(47): p. 34168–80.

104. Schwendener S.et al. *Physical Interaction of RECQ5 Helicase with RAD51 Facilitates Its Anti-recombinase Activity*. J Biol Chem, 2010. **285**(21): p. 15739–45.

105. Matos J. and West S.C. *Holliday junction resolution: Regulation in space and time.* DNA Repair, 2014. **19**: p. 176–81.

106. Singh T.R. et al. *BLAP18/RMI2, a novel OB-fold-containing protein, is an essential component of the Bloom helicase–double Holliday junction dissolvasome*. Genes Dev, 2008. **22**(20): p. 2856–68.

107. Bussen W. et al. *Holliday Junction Processing Activity of the BLM-Topo IIIα-BLAP75 Complex*. J Biol Chem, 2007. **282**(43): p. 31484–92.

108. Wyatt H.D.M. et al. *Coordinated Actions of SLX1-SLX4 and MUS81-EME1 for Holliday Junction Resolution in Human Cells*. Molecular Cell, 2013. **52**(2): p. 234–47.

109. Simsek D. and Jasin M. *Alternative end-joining is suppressed by the canonical NHEJ component Xrcc4–ligase IV during chromosomal translocation formation.* Nat Struct Mol Biol, 2010. **17**(4): p. 410-416.

110. Wang M. et al. *PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways*. Nucleic Acids Res, 2006. **34**(21): p. 6170–82.

111. Deng S.K. et al. *RPA antagonizes microhomology-mediated repair of DNA double-strand breaks*. Nat Struct Mol Biol, 2014. **21**(4): p. 405–12.

112. Ahrabi S. et al. *A role for human homologous recombination factors in suppressing microhomology-mediated end joining*. Nucleic Acids Res, 2016. **44**(12): p. 5743–57.

113. Mateos-Gomez P.A. et al. *The helicase domain of Polθ counteracts RPA to promote alt-NHEJ.* Nature Structural & Molecular Biology, 2017. **24**(12): p. 1116–23.

114. Lu G. et al. *Ligase I and ligase III mediate the DNA double-strand break ligation in alternative end-joining.* Proc Natl Acad Sci U S A, 2016. **113**(5): p. 1256–60.

115. Lieberman R. et al. *Functional characterization of RAD52 as a lung cancer susceptibility gene in the 12p13.33 locus.* Molecular Carcinogenesis, 2016. **55**(5): p. 953–63.

116. Nogueira A. et al. *RAD52 Functions in Homologous Recombination and Its Importance on Genomic Integrity Maintenance and Cancer Therapy.* Cancers, 2019. **11**(11): 1622. https://doi.org/10.3390/cancers11111622

117. Blow J.J. and Gillespie P.J. *Replication licensing and cancer — a fatal entanglement?* Nature Reviews Cancer, 2008. **8**(10): p. 799–806.

118. Moyer S.E., Lewis P.W. and Botchan M.R. *Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase.* PNAS, 2006. **103**(27): p. 10236–41.

119. Liu Y. et al. *The origins and processing of ultra fine anaphase DNA bridges*. Current Opinion in Genetics & Development, 2014. **26**: p. 1–5.

120. Lukas C. et al. *53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress.* Nature Cell Biology, 2011. **13**(3): p. 243–53.

121. Spies J. et al. *53BP1 nuclear bodies enforce replication timing at under-replicated DNA to limit heritable DNA damage*. Nature Cell Biology, 2019. **21**(4): p. 487–97.

122. Minocherhomji S. et al. *Replication stress activates DNA repair synthesis in mitosis*. Nature, 2015. **528**(7581): p. 286–90.

123. McCulloch S.D. and Kunkel T.A. *The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases.* Cell Research, 2008. **18**(1): p. 148–61.

124. Powers K.T. and Washington M.T. *Eukaryotic translesion synthesis: Choosing the right tool for the job.* DNA Repair, 2018. **71**: p. 127–34.

125. Branzei D. *Ubiquitin family modifications and template switching*. FEBS Letters, 2011. **585**(18): p. 2810–7.

126. Anand R.P. et al. *Chromosome rearrangements via template switching between diverged repeated sequences*. Genes Dev, 2014. **28**(21): p. 2394–406.

127. Mourón S. et al. *Repriming of DNA synthesis at stalled replication forks by human PrimPol.* Nature Structural & Molecular Biology, 2013. **20**(12): p. 1383–9.

128. Quinet A., Lemaçon D. and Vindigni A. *Replication Fork Reversal: Players and Guardians*. Molecular Cell, 2017. **68**(5): p. 830–3.

129. Kolinjivadi A.M. et al. *Smarcal1-Mediated Fork Reversal Triggers Mre11-Dependent Degradation of Nascent DNA in the Absence of Brca2 and Stable Rad51 Nucleofilaments.* Molecular Cell, 2017. **67**(5): p. 867-881.

130. Vujanovic M. et al. *Replication Fork Slowing and Reversal upon DNA Damage Require PCNA Polyubiquitination and ZRANB3 DNA Translocase Activity.* Molecular Cell, 2017. **67**(5): p. 882-890.

131. Zellweger R. et al. *Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells.* J Cell Biol, 2015. **208**(5): p. 563–79.

132. Thangavel S. et al. *DNA2 drives processing and restart of reversed replication forks in human cells.* J Cell Biol, 2015. **208**(5): p. 545–62.

133. Mijic S. et al. *Replication fork reversal triggers fork degradation in BRCA2-defective cells*. Nature Communications, 2017. **8**(859). https://doi.org/10.1038/s41467-017-01164-5.

134. Lemaçon D. et al. *MRE11 and EXO1 nucleases degrade reversed forks and elicit MUS81-dependent fork rescue in BRCA2-deficient cells.* Nature Communications, 2017. **8**(860). https://doi.org/10.1038/s41467-017-01180-5

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

136. 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.

137. Cortez D. *Preventing replication fork collapse to maintain genome integrity.* DNA Repair, 2015. **32**: p. 149–57.

138. Hanada K. et al. *The structure-specific endonuclease Mus81 contributes to replication restart by generating double-strand DNA breaks*. Nature Structural & Molecular Biology, 2007. **14**(11): p. 1096–104.

139. Pepe A. and West S.C. *MUS81-EME2 Promotes Replication Fork Restart*. Cell Reports, 2014. **7**(4): p. 1048–55.

140. Llorente B., Smith C.E. and Symington L.S. *Break-induced replication: What is it and what is it for?* Cell Cycle, 2008. **7**(7): p. 859–64.

141. Kramara J., Osia B. and Malkova A. *Break-Induced Replication: The Where, The Why, and The How.* Trends in Genetics, 2018. **34**(7): p. 518–31.

142. Dilley R.L. et al. *Break-induced telomere synthesis underlies alternative telomere maintenance.* Nature, 2016. **539**(7627): p. 54–8.

143. Lydeard J.R. et al. *Break-induced replication and telomerase-independent telomere maintenance require Pol32.* Nature, 2007. **448**(7155): p. 820–3.

144. Kolinjivadi A.M. et al. *Moonlighting at replication forks – a new life for homologous recombination proteins BRCA1, BRCA2 and RAD51*. FEBS Letters, 2017. **591**(8): p. 1083–100.

145. Deem A. et al. *Break-Induced Replication Is Highly Inaccurate.* PLOS Biology, 2011. **9**(2): e1000594.

146. Glover T.W., Wilson T.E. and Arlt M.F. *Fragile sites in cancer: more than meets the eye.* Nat Rev Cancer, 2017. **17**(8): p. 489–501.

147. Helmrich A., Ballarino M. and Tora L. *Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes.* Molecular Cell, 2011. **44**(6): p. 966–77.

148. Yu S. et al. *Human Chromosomal Fragile Site FRA16B Is an Amplified AT-Rich Minisatellite Repeat*. Cell, 1997. **88**(3): p. 367–74.

149. Zhang H. and Freudenreich C.H. *An AT-Rich Sequence in Human Common Fragile Site FRA16D Causes Fork Stalling and Chromosome Breakage in S. cerevisiae*. Molecular Cell, 2007. **27**(3): p. 367–79.

150. Miotto B., Ji Z. and Struhl K. *Selectivity of ORC binding sites and the relation to replication timing, fragile sites, and deletions in cancers*. PNAS, 2016. **113**(33): p. 4810–9.

151. Le Beau M.M. et al. *Replication of a Common Fragile Site, FRA3B, Occurs Late in S Phase and is Delayed Further Upon Induction: Implications for the Mechanism of Fragile Site Induction*. Human Molecular Genetics, 1998. **7**(4): p. 755–61.

152. Palakodeti A. et al. *The role of late/slow replication of the FRA16D in common fragile site induction.* Genes, Chromosomes and Cancer, 2004. **39**(1): p. 71–6.

153. Wang H.et al. *CtIP Maintains Stability at Common Fragile Sites and Inverted Repeats by End Resection-Independent Endonuclease Activity.* Molecular Cell, 2014. **54**(6): p. 1012–21.

154. Madireddy A. et al. *FANCD2 Facilitates Replication through Common Fragile Sites*. Molecular Cell, 2016. **64**(2): p. 388–404.

155. Naim V. et al. *ERCC1 and MUS81–EME1 promote sister chromatid separation by processing late replication intermediates at common fragile sites during mitosis.* Nature Cell Biology, 2013. **15**(8): p. 1008–15.

156. Wang H. et al. *BLM prevents instability of structure-forming DNA sequences at common fragile sites.* PLOS Genetics, 2018. **14**(11): e1007816.

157. Rai R. et al. *The function of classical and alternative non-homologous end-joining pathways in the fusion of dysfunctional telomeres*. The EMBO Journal, 2010. **29**(15): p. 2598–610.

158. Maestroni L., Matmati S. and Coulon S. *Solving the Telomere Replication Problem.* Genes, 2017. **8**(55): doi:10.3390.

159. Denchi E.L. and de Lange T. *Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1*. Nature, 2007. **448**(7157): p. 1068–71.

160. Doksani Y. et al. *Super-Resolution Fluorescence Imaging of Telomeres Reveals TRF2-Dependent T-loop Formation.* Cell, 2013. **155**(2): p. 345–56.

161. Vannier J-B. et al. *RTEL1 Is a Replisome-Associated Helicase That Promotes Telomere and Genome-Wide Replication*. Science, 2013. **342**(6155): p. 239–42.

162. Crabbe L. et al. *Defective Telomere Lagging Strand Synthesis in Cells Lacking WRN Helicase Activity*. Science, 2004. **306**(5703): p. 1951–3.

163. Barefield C. and Karlseder J. *The BLM helicase contributes to telomere maintenance through processing of late-replicating intermediate structures.* Nucleic Acids Research, 2012. **40**(15): p. 7358–67.

164. Amiard S. et al. *A topological mechanism for TRF2-enhanced strand invasion.* Nature Structural & Molecular Biology, 2007. **14**(2): p. 147–54.

165. Ye J. et al. *TRF2 and Apollo Cooperate with Topoisomerase 2α to Protect Human Telomeres from Replicative Damage*. Cell, 2010. **142**(2): p. 230–42.

166. Poole L.A. et al. *SMARCAL1 maintains telomere integrity during DNA replication*. Proc Natl Acad Sci USA, 2015. **112**(48): p. 14864–9.

167. Couch F.B. et al. *ATR phosphorylates SMARCAL1 to prevent replication fork collapse*. Genes Dev, 2013. **27**(14): p. 1610–23.

168. Musacchio A. and Desai A. *A Molecular View of Kinetochore Assembly and Function.* Biology, 2017. **6**(1):5. doi: 10.3390/biology6010005.

169. Westhorpe F.G. and Straight A.F. *Functions of the centromere and kinetochore in chromosome segregation*. Current Opinion in Cell Biology, 2013. **25**(3): p. 334–40.

170. Willard H.F. and Waye J.S. *Chromosome-specific subsets of human alpha satellite DNA: Analysis of sequence divergence within and between chromosomal subsets and evidence for an ancestral pentameric repeat.* Journal of Molecular Evolution, 1987. **25**(3): p. 207–14.

171. Sullivan L.L., Chew K. and Sullivan B.A. *α satellite DNA variation and function of the human centromere.* Nucleus, 2017. **8**(4): p. 331–9.

172. Foltz D.R. et al. *The human CENP-A centromeric nucleosome-associated complex.* Nature Cell Biology, 2006. **8**(5): p. 458–69.

173. Hasson D. et al. *The octamer is the major form of CENP-A nucleosomes at human centromeres*. Nat Struct Mol Biol, 2013. **20**(6): p. 687–95.

174. Foltz D.R. et al. *Centromere-Specific Assembly of CENP-A Nucleosomes Is Mediated by HJURP.* Cell, 2009. **137**(3): p. 472–84.

175. Carroll C.W., Milks K.J. and Straight A.F. *Dual recognition of CENP-A nucleosomes is required for centromere assembly.* Journal of Cell Biology, 2010. **189**(7): p. 1143–55.

176. Falk S.J. et al. *CENP-C reshapes and stabilizes CENP-A nucleosomes at the centromere.* Science, 2015. **348**(6235): p. 699–703.

177. Gascoigne K.E. et al. *Induced Ectopic Kinetochore Assembly Bypasses the Requirement for CENP-A Nucleosomes*. Cell, 2011. **145**(3): p. 410–22.

178. Tian T. et al. *Molecular basis for CENP-N recognition of CENP-A nucleosome on the human kinetochore*. Vol. 28, Cell Research, 2018. **28**: p. 374–378.

179. McKinley K.L. et al. *The CENP-L-N Complex Forms a Critical Node in an Integrated Meshwork of Interactions at the Centromere-Kinetochore Interface*. Molecular Cell, 2015. **60**(6): p. 886–98.

180. Nishino T. et al. *CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold.* Cell, 2012. **148**(3): p. 487–501.

181. Takeuchi K. et al. *The centromeric nucleosome-like CENP–T–W–S–X complex induces positive supercoils into DNA.* Nucleic Acids Research, 2014. **42**(3): p. 1644–55.

182. Basilico F. et al. *The pseudo GTPase CENP-M drives human kinetochore assembly.* eLife, 2014. **3**: e02978.

183. Okada M. *CENP-H–containing Complex Facilitates Centromere Deposition of CENP-A in Cooperation with FACT and CHD1*. MBoC, 2009. **20**(18): p. 3986–95.

184. Bancroft J. et al. *Chromosome congression is promoted by CENP-Q- and CENP-E-dependent pathways*. Journal of Cell Science, 2015. **128**(1): p. 171–84.

185. Kang Y.H. et al. *Mammalian Polo-like Kinase 1-dependent Regulation of the PBIP1-CENP-Q Complex at Kinetochores.* J Biol Chem, 2011. **286**(22): p. 19744–57.

186. Pesenti M.E. et al. *Reconstitution of a 26-Subunit Human Kinetochore Reveals Cooperative Microtubule Binding by CENP-OPQUR and NDC80.* Molecular Cell, 2018. **71**(6): p. 923-939.

187. Barra V. and Fachinetti D. *The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA.* Nature Communications, 2018. **9**: 4340. https://doi.org/10.1038/s41467-018-06545-y.

188. McKinley K.L. and Cheeseman I.M. T*he molecular basis for centromere identity and function.* Nature Reviews Molecular Cell Biology, 2016. **17**(1): p. 16–29.

189. Hagen K.G.T. et al. *Replication timing of DNA sequences associated with human centromeres and telomeres*. Molecular and Cellular Biology, 1990. **10**(12): p. 6348–55.

190. Crosetto N. et al. *Nucleotide-resolution DNA double-strand breaks mapping by next-generation sequencing.* Nat Methods, 2013. **10**(4): p. 361–5.

191. Aze A. et al. *Centromeric DNA replication reconstitution reveals DNA loops and ATR checkpoint suppression*. Nature Cell Biology, 2016. **18**(6): p. 684–91.

192. Li G-M. *Mechanisms and functions of DNA mismatch repair*. Cell Research, 2008. **18**(1): p. 85–98.

193. Mitra S. et al. *Genetic screening identifies a SUMO protease dynamically maintaining centromeric chromatin.* Nature Communications, 2020. **11**:501. https://doi.org/10.1038/s41467-019-14276-x.

194. Chan K-L., North P.S. and Hickson I.D. *BLM is required for faithful chromosome segregation and its localization defines a class of ultrafine anaphase bridges*. The EMBO Journal, 2007. **26**(14): p. 3397–409.

195. Rouzeau S. et al. *Bloom's Syndrome and PICH Helicases Cooperate with Topoisomerase IIα in Centromere Disjunction before Anaphase*. PLOS ONE, 2012. **7**(4): e33905.

196. Nielsen C.F. et al. *PICH promotes sister chromatid disjunction and co-operates with topoisomerase II in mitosis*. Nature Communications, 2015. **6**:8962. https://doi.org/10.1038/ncomms9962

197. Chan Y.W., Fugger K. and West S.C. *Unresolved recombination intermediates lead to ultra-fine anaphase bridges, chromosome breaks and aberrations*. Nature Cell Biology, 2018. **20**(1): p. 92–103.

198. Nakayama J. et al. *Role of Histone H3 Lysine 9 Methylation in Epigenetic Control of Heterochromatin Assembly*. Science, 2001. **292**(5514): p. 110–3.

199. Lomberk G., Wallrath L. and Urrutia R. *The Heterochromatin Protein 1 family.* Genome Biol, 2006. **7**: 228. https://doi.org/10.1186/gb-2006-7-7-228.

200. Giunta S. and Funabiki H. *Integrity of the human centromere DNA repeats is protected by CENP-A, CENP-C, and CENP-T.* Proc Natl Acad Sci USA, 2017. **114**(8): p. 1928–33.

201. Chiolo I. et al. *Double-Strand Breaks in Heterochromatin Move Outside of a Dynamic HP1a Domain to Complete Recombinational Repair*. Cell, 2011. **144**(5): p. 732–44.

202. Jakob B. et al. *DNA double-strand breaks in heterochromatin elicit fast repair protein recruitment, histone H2AX phosphorylation and relocation to euchromatin.* Nucleic Acids Res, 2011. **39**(15): p. 6489–99.

203. Tsouroula K. et al. *Temporal and Spatial Uncoupling of DNA Double Strand Break Repair Pathways within Mammalian Heterochromatin.* Molecular Cell, 2016. **63**(2): p. 293–305.

204. Flaus A. et al. *Identification of multiple distinct Snf2 subfamilies with conserved structural motifs.* Nucleic Acids Res, 2006. **34**(10): p. 2887–905.

205. Narlikar G.J., Sundaramoorthy R. and Owen-Hughes T. *Mechanisms and Functions of ATP-Dependent Chromatin-Remodeling Enzymes*. Cell, 2013. **154**(3): p. 490–503.

206. Singleton M.R., Dillingham M.S. and Wigley D.B. *Structure and Mechanism of Helicases and Nucleic Acid Translocases*. Annual Review of Biochemistry, 2007. **76**(1): p. 23–50.

207. Weston R., Peeters H. and Ahel D. *ZRANB3 is a structure-specific ATP-dependent endonuclease involved in replication stress response.* Genes Dev, 2012. **26**(14): p. 1558–72.

208. Motegi A. et al. *Polyubiquitination of proliferating cell nuclear antigen by HLTF and SHPRH prevents genomic instability from stalled replication forks.* PNAS, 2008. **105**(34): p. 12411–6.

209. Ryan D.P. and Owen-Hughes T. *Snf2-family proteins: chromatin remodellers for any occasion*. Current Opinion in Chemical Biology, 2011. **15**(5): p. 649–56.

210. Li M. et al. *Mechanism of DNA translocation underlying chromatin remodelling by Snf2.* Nature, 2019. **567**(7748): p. 409–13.

211. Joseph S.A. et al. *Time for remodeling: SNF2-family DNA translocases in replication fork metabolism and human disease.* DNA Repair, 2020. **95**: e102943.

212. Bugreev D.V., Mazina O.M. and Mazin A.V. *Rad54 protein promotes branch migration of Holliday junctions*. Nature, 2006. **442**(7102): p. 590–3.

213. Bizard A.H. et al. *PICH and TOP3A cooperate to induce positive DNA supercoiling*. Nature Structural & Molecular Biology, 2019. **26**(4): p. 267–74.

214. Vermeulen W. and Fousteri M. *Mammalian transcription-coupled excision repair.* Cold Spring Harbor Perspectives in Biology, 2013. **5**: a012625.

215. Biebricher A. et al. *PICH: A DNA Translocase Specially Adapted for Processing Anaphase Bridge DNA.* Molecular Cell, 2013. **51**(5): p. 691–701.

216. Tummala H. et al. *ERCC6L2 mutations link a distinct bone-marrow-failure syndrome to DNA repair and mitochondrial function.* American Journal of Human Genetics, 2014. **94**(2): p. 246–56.

217. Zhang S. et al. *A nonsense mutation in the DNA repair factor Hebo causes mild bone marrow failure and microcephaly*. Journal of Experimental Medicine, 2016. **213**(6): p. 1011–28.

218. El-Gebali S. et al. The Pfam protein families database in 2019. Nucleic Acids Research, 2019. **47**: p. 427–32.

219. Shabanova I. et al. *ERCC6L2 -associated inherited bone marrow failure syndrome.* Molecular Genetics & Genomic Medicine, 2018. **6**(3): p. 463–8.

220. Tummala H. et al. *Genome instability is a consequence of transcription deficiency in patients with bone marrow failure harboring biallelic ERCC6L2 variants*. Proceedings of the National Academy of Sciences of the United States of America, 2018. **115**(30): p. 7777–82.

221. Douglas S.P.M. et al. *ERCC6L2 defines a novel entity within inherited acute myeloid leukemia*. Blood. American Society of Hematology, 2019. **133**: p. 2724–2728.

222. Dokal I. and Vulliamy T. *Inherited bone marrow failure syndromes.* Haematologica; Ferrata Storti Foundation, 2010. **95**: p. 1236–1240.

223. Bluteau O. et al. *A landscape of germ line mutations in a cohort of inherited bone marrow failure patients*. Blood, 2018. **131**(7): p. 717–32.

224. Liu X. et al. *ERCC6L2 promotes DNA orientation-specific recombination in mammalian cells.* Cell Research, 2020. **30**(9): p. 732–44.

225. Olivieri M. et al. *A Genetic Map of the Response to DNA Damage in Human Cells.* Cell, 2020. **182**(2): p. 481-496.e21.

226. Francica P. et al. *Functional Radiogenetic Profiling Implicates ERCC6L2 in Non-homologous End Joining.* Cell Reports, 2020. **32**(8): e108068.

227. Larsen A.K., Escargueil A.E. and Skladanowski A. *Catalytic topoisomerase II inhibitors in cancer therapy.* Pharmacology & Therapeutics, 2003. **99**(2): p. 167–81.

228. Kim D.I. et al. *An improved smaller biotin ligase for BioID proximity labeling.* MBoC, 2016. **27**(8): p. 1188–96.

229. Liu S. et al. *Distinct roles for DNA-PK, ATM and ATR in RPA phosphorylation and checkpoint activation in response to replication stress*. Nucleic Acids Research, 2012. **40**(21): p. 10780–94.

230. Maréchal A. and Zou L. *RPA-coated single-stranded DNA as a platform for post-translational modifications in the DNA damage response.* Cell Res, 2015. **25**(1): p. 9–23.

231. Goodarzi A.A., Jeggo P. and Lobrich M. *The influence of heterochromatin on DNA double strand break repair: Getting the strong, silent type to relax*. DNA Repair, 2010. **9**(12): p. 1273–82.

232. Mao Z. et al. *Comparison of nonhomologous end joining and homologous recombination in human cells*. DNA Repair, 2008. **7**(10): p. 1765–71.

233. Jaspers J.E. et al. *Loss of 53BP1 Causes PARP Inhibitor Resistance in Brca1-Mutated Mouse Mammary Tumors*. Cancer Discov, 2013. **3**(1): p. 68–81.

234. Lord C.J. and Ashworth A. *Targeted therapy for cancer using PARP inhibitors*. Current Opinion in Pharmacology, 2008. **8**(4): p. 363–9.

235. Barazas M. et al. *The CST Complex Mediates End Protection at Double-Strand Breaks and Promotes PARP Inhibitor Sensitivity in BRCA1-Deficient Cells.* Cell Reports, 2018. **23**(7): p. 2107–18.

236. Huertas P. *DNA resection in eukaryotes: deciding how to fix the break.* Nature Structural & Molecular Biology, 2010. **17**(1): p. 11–6.

237. Noordermeer S.M. et al. *The shieldin complex mediates 53BP1-dependent DNA repair.* Nature, 2018. **560**(7716): p. 117–21.

238. 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. 548–52.

239. Hung P.J. et al. *MRI Is a DNA Damage Response Adaptor during Classical Non-homologous End Joining*. Molecular Cell, 2018. **71**(2): p. 332-342.

240. Shibata A. *Regulation of repair pathway choice at two-ended DNA double-strand breaks.* Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 2017. **803–805**: p. 51–5.

241. Richardson C., Moynahan M.E. and Jasin M. *Double-strand break repair by interchromosomal recombination: suppression of chromosomal translocations*. Genes Dev, 1998. **12**(24): p. 3831–42.

242. Mailand N., Gibbs-Seymour I. and Bekker-Jensen S. *Regulation of PCNA–protein interactions for genome stability*. Nature Reviews Molecular Cell Biology, 2013. **14**(5): p. 269–82.

243. Yarbro J.W. *Mechanism of action of hydroxyurea*. Semin Oncol, 1992. **19**: p. 1–10.

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

245. 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.

246. Byun T.S. et al. *Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint*. Genes Dev, 2005. **19**(9): p. 1040–52.

247. Tham K-C., Kanaar R. and Lebbink J.H.G. *Mismatch repair and homeologous recombination.* DNA Repair, 2016. **38**: p. 75–83.

248. Stavnezer J. and Schrader C.E. *Mismatch repair converts AID-instigated nicks to double-strand breaks for antibody class-switch recombination*. Trends in Genetics, 2006. **22**(1): p. 23–8.

249. Chakraborty U. and Alani E. *Understanding how mismatch repair proteins participate in the repair/anti-recombination decision*. FEMS Yeast Res, 2016. **16**(6): fow071. doi: 10.1093/femsyr/fow071.

250. Bai G. et al. *HLTF Promotes Fork Reversal, Limiting Replication Stress Resistance and Preventing Multiple Mechanisms of Unrestrained DNA Synthesis.* Molecular Cell, 2020. **78**(6): p. 1237-1251.

251. Ciccia A. et al. *Polyubiquitinated PCNA Recruits the ZRANB3 Translocase to Maintain Genomic Integrity after Replication Stress.* Molecular Cell, 2012. **47**(3): p. 396–409.

252. Sun H. et al. *The Bloom's Syndrome Helicase Unwinds G4 DNA*. J Biol Chem, 1998. **273**(42): p. 27587–92.

253. Opresko P.L. et al. *POT1 Stimulates RecQ Helicases WRN and BLM to Unwind Telomeric DNA Substrates*. J Biol Chem, 2005. **280**(37): p. 32069–80.

254. Li Z. et al. *hDNA2 nuclease/helicase promotes centromeric DNA replication and genome stability*. The EMBO Journal, 2018. **37**(14): e96729.

255. Segurado M. and Diffley J.F.X. *Separate roles for the DNA damage checkpoint protein kinases in stabilizing DNA replication forks.* Genes Dev, 2008. **22**(13): p. 1816–27.

256. Cotta-Ramusino C. et al. *Exo1 Processes Stalled Replication Forks and Counteracts Fork Reversal in Checkpoint-Defective Cells.* Molecular Cell, 2005. **17**(1): p. 153–9.

257. Sugawara N. et al. *Heteroduplex rejection during single-strand annealing requires Sgs1 helicase and mismatch repair proteins Msh2 and Msh6 but not Pms1*. PNAS, 2004. **101**(25): p. 9315–20.

258. Marchler-Bauer A. et al. *CDD: NCBI's conserved domain database.* Nucleic Acids Res, 2015. **43**: p. 222-226.

259. Roux K.J., Kim D.I. and Burke B. *BioID: A Screen for Protein-Protein Interactions.* Current Protocols in Protein Science, 2013. **74**(1): p. 1-14.

260. Hori T. et al. *CENP-O Class Proteins Form a Stable Complex and Are Required for Proper Kinetochore Function.* MBoC, 2007. **19**(3): p. 843–54.

261. Jallepalli P.V. et al. *Securin Is Required for Chromosomal Stability in Human Cells*. Cell, 2001. **105**(4): p. 445–57.

262. Waizenegger I.C. et al. *Regulation of Human Separase by Securin Binding and Autocleavage*. Current Biology, 2002. **12**(16): p. 1368–78.

263. Romero F. et al. *Human securin, hPTTG, is associated with Ku heterodimer, the regulatory subunit of the DNA-dependent protein kinase*. Nucleic Acids Research, 2001. **29**(6): p. 1300–7.

264. Bernal J.A. et al. *Proliferative potential after DNA damage and non-homologous end joining are affected by loss of securin.* Cell Death & Differentiation, 2008. **15**(1): p. 202–12.

265. Lake R.J. and Fan H-Y. *Structure, function and regulation of CSB: A multi-talented gymnast.* Mechanisms of Ageing and Development, 2013. **134**(5): p. 202–11.

266. Eskat A. et al. *Step-Wise Assembly, Maturation and Dynamic Behavior of the Human CENP-P/O/R/Q/U Kinetochore Sub-Complex.* PLOS ONE, 2012. **7**(9): e44717.

267. 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**: 87. https://doi.org/10.1038/s41467-018-07729-2.

268. Barra V. and Fachinetti D. *The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA.* Nature Communications, 2018. **9**: p. 1–17.

269. Boersma V. et al. *MAD2L2 controls DNA repair at telomeres and DNA breaks by inhibiting 5' end-resection*. Nature, 2015. **521**(7553): p. 537.

270. Xu G. et al. *REV7 counteracts DNA double-strand break resection and affects PARP inhibition*. Nature, 2015. **521**(7553): p. 541–4.

271. Reid D.A. et al. *Organization and dynamics of the nonhomologous end-joining machinery during DNA double-strand break repair.* Proc Natl Acad Sci USA, 2015. **112**(20): p. 2575–84.

272. Jaafar L. et al. *SFPQ•NONO and XLF function separately and together to promote DNA double-strand break repair via canonical nonhomologous end joining.* Nucleic Acids Research, 2017. **45**(4): p. 1848–59.

273. Gómez-Herreros F. et al. *TDP2–Dependent Non-Homologous End-Joining Protects against Topoisomerase II–Induced DNA Breaks and Genome Instability in Cells and In Vivo.* PLOS Genetics, 2013. **9**(3): e1003226.

274. Quennet V. et al. *CtIP and MRN promote non-homologous end-joining of etoposide-induced DNA double-strand breaks in G1.* Nucleic Acids Research, 2011. **39**(6): p. 2144–52.

275. Gao S. et al. *An OB-fold complex controls the repair pathways for DNA double-strand breaks.* Nature Communications, 2018. **9**:3925. https://doi.org/10.1038/s41467-018-06407-7.

276. He Y.J. et al. *DYNLL1 binds to MRE11 to limit DNA end resection in BRCA1-deficient cells.* Nature, 2018. **563**(7732): p. 522–6.

277. Gilbert N. and Allan J. *Distinctive higher-order chromatin structure at mammalian centromeres.* PNAS, 2001. **98**(21): p. 11949–54.

278. Guenatri M. et al. *Mouse centric and pericentric satellite repeats form distinct functional heterochromatin*. J Cell Biol, 2004. **166**(4): p. 493–505.

279. Canela A. et al. *Genome Organization Drives Chromosome Fragility.* Cell, 2017. **170**(3): p. 507-521.

280. Minchell N.E., Keszthelyi A. and Baxter J. *Cohesin Causes Replicative DNA Damage by Trapping DNA Topological Stress.* Molecular Cell, 2020. **78**(4): p. 739-751.

281. Dehé P-M. and Gaillard P-HL. *Control of structure-specific endonucleases to maintain genome stability*. Nature Reviews Molecular Cell Biology, 2017. **18**(5): p. 315–30.

282. Lia G. et al. *Direct Observation of DNA Distortion by the RSC Complex*. Molecular Cell, 2006. **21**(3): p. 417–25.

283. Prasad T.K. et al. *A DNA-translocating Snf2 Molecular Motor: Saccharomyces cerevisiae Rdh54 Displays Processive Translocation and Extrudes DNA Loops*. Journal of Molecular Biology, 2007. **369**(4): p. 940–53.

284. Ivessa A.S. et al. *The Saccharomyces cerevisiae Helicase Rrm3p Facilitates Replication Past Nonhistone Protein-DNA Complexes*. Molecular Cell, 2003. **12**(6): p. 1525–36.

285. Skourti-Stathaki K. and Proudfoot N.J. *A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression.* Genes Dev, 2014. **28**(13): p. 1384–96.

286. Petti E. et al. *SFPQ and NONO suppress RNA:DNA-hybrid-related telomere instability*. Nature Communications, 2019. **10**:1001. https://doi.org/10.1038/s41467-019-08863-1.

287. Kordon M.M. et al. *STRIDE—a fluorescence method for direct, specific in situ detection of individual single- or double-strand DNA breaks in fixed cells*. Nucleic Acids Research, 2020. **48**(3): e14–e14.

288. Bluteau D. et al. *Biallelic inactivation of REV7 is associated with Fanconi anemia*. J Clin Invest, 2017. **127**(3): e1117-1121.

289. Shen W. et al. *Impact of germline CTC1 alterations on telomere length in acquired bone marrow failure.* British Journal of Haematology, 2019. **185**(5): p. 935–9.

290. Batrakou D.G., Heron E.D. and Nieduszynski C.A. *Rapid high-resolution measurement of DNA replication timing by droplet digital PCR.* Nucleic Acids Research, 2018. **46**(19): e112. doi: 10.1093/nar/gky590.