Investigating the role of ERCC6L2 in the DNA damage response



Christopher J. Carnie

Lincoln College
Sir William Dunn School of Pathology

University of Oxford

Thesis submitted for the degree of DPhil in Molecular and Cell Biology in Health and Disease
Trinity Term 2018

Abstract:

Background: ERCC6L2 is a SNF2 ATPase found mutated in a small number of bone marrow failure (BMF) patients. Studies identifying ERCC6L2 mutations have suggested a role for this uncharacterised protein in the DNA damage response (DDR), but its function is still unclear. **Aims:** The purpose of my project was to ascertain whether or not ERCC6L2 is involved in specific arms of the DDR and to investigate the contributions of putative functional domains to such a role. **Results:** Overexpressed ERCC6L2 localises to DNA lesions induced by laser microirradiation. This occurs in a manner dependent on ERCC6L2's C-terminus, which contains no characterised functional domains and is lost in a number of BMF-causing mutations. Concomitantly, a range of clinical mutants of ERCC6L2 fail to localise to DNA damage sites, perhaps explaining a DDR defect in ERCC6L2-mutant patients. By establishing ERCC6L2$^{-/-}$ cell lines and performing clonogenic survival assays, I found that ERCC6L2 is necessary for U2OS cells' tolerance of both phleomycin and etoposide. Notably, ERCC6L2's role in the response to these drugs is dependent on its ATPase activity, which is likely lost in several BMF-causing point mutations. While phleomycin and etoposide are typically considered to cause DNA double strand breaks (DSBs), studies published while this one was underway have demonstrated that the two primary means of repairing DSBs are not affected by the loss of ERCC6L2. As such, I sought to further characterise the basis of this sensitivity. Performing cell cycle analysis by flow cytometry, we found that ERCC6L2$^{-/-}$ cells treated with etoposide display a profound defect in DNA replication. This defect is rescuable by overexpression of WT ERCC6L2, but not by overexpression of a catalytic-dead mutant. Furthermore, ERCC6L2 interacts directly with PCNA, suggesting that the replication defect in ERCC6L2$^{-/-}$ cells treated with etoposide might relate directly to DNA replication. I found that ERCC6L2's C-terminus contains a novel type of atypical PCNA interaction motif. By co-crystallisation of PCNA with this peptide we have noted that the mechanism of this interaction is similar to that other canonical PCNA interaction motifs, despite differences in primary sequence and spacing. However, it seems that ERCC6L2's role at DNA damage sites and/or during DNA replication is dependent on more than just its PCNA interaction and that some redundancy or cooperation exists. A PCNA-binding mutant of ERCC6L2 is still recruited to DNA lesions, while overexpressing this mutant in ERCC6L2$^{-/-}$ cells still provides a partial rescue of DNA synthesis and survival in etoposide-treated cells. Interestingly, ERCC6L2 colocalises with the centromeric histone variant CENP-A through a portion of its C-terminus throughout interphase, including during centromere replication. **Conclusions:** ERCC6L2 is involved in the DDR and its function relates to genotoxic stress caused by etoposide and phleomycin. The ATPase activity of ERCC6L2 is important for its role in the DDR, and elements of its C-terminus are necessary for its localisation to both DNA lesions and centromeres. ERCC6L2's direct interaction with PCNA suggests a role related to DNA replication. Considering the DNA replication defect in etoposide-treated ERCC6L2$^{-/-}$ cells and ERCC6L2's PCNA interaction, as well as the propensity for DNA secondary structure formation at centromeres, our working model for ERCC6L2 function is that this protein facilitates replication through centromeres – and perhaps other genomic loci – by relaxing replication-perturbing DNA structures. *In vitro* investigations into ERCC6L2 function will be required to test this model fully. These findings might provide valuable insights into ERCC6L2 function and the basis of ERCC6L2 deficiency disorders.

Acknowledgements

Firstly, I would like to thank my supervisor, Dragana Ahel, for the opportunity to join in her group and work on this project as well as the EPA Trust for funding this DPhil. I feel I have developed as a scientist and learnt a lot from Dragana during this time and owe a great deal of gratitude for her input into this process.

On a personal and professional level, I have also benefitted immensely from the supportive and friendly environment of the Dunn School, particularly on OMPI1. The quality of people I've met here has formed the basis of what has been a really enriching experience. I'd particularly like to thank several colleagues who have become friends. Chris, Marek, Adam, Lucy, Nina and Io Nam, thank you for the laughs, science and advice!

The concept of a social toll of DPhil/PhD studies is now established to the point of cliché. As such there are many people whose support and friendship has helped keep me sane at times. I'd like to thank Anna, Ben, Lisa, David, Greg, Ed and Phil for reminding me, each in their own way, that there's more to life than failed experiments.

The nature versus nurture debate is one which has existed for some time, despite the fairly obvious fact that the development of a human is influenced heavily by both. Since my parents Jill and Bob are responsible for both my genetic makeup and the environment in which I was brought up, I'd like to thank them for their direct and indirect contributions to this experience.

Last, but certainly not least, I'd like to thank my partner Neele for the weeks of love, patience, support and scientific input to the writing of this thesis. Her input can't be overstated and I hate to think of the state this thesis would be in – much less the state I would be in – without her impressive efforts. Thank you!

1 Introduction 11

1.1 The DNA damage response 11

1.2 DNA lesions: what are they and how do they occur? 12

1.3 Sensing and responding to DNA damage 15

1.4 The DDR and the cell cycle 17

1.4.1 The G1/S checkpoint 19
1.4.2 The intra-S checkpoint 20
1.4.3 The G2/M checkpoint 21
1.4.4 The DDR in mitosis 22
1.4.5 Replication and recombination intermediates in mitosis 23

1.5 DSB repair 26

1.5.1 Non-homologous end-joining 26
1.5.2 Homologous recombination 30

1.6 Replication stress: a major source of genomic instability 32

1.7 The influence of genomic location on the DDR 40

1.7.1 Telomeres and the DDR 41
1.7.2 CFSs and the DDR 42
1.7.3 Centromeres and the DDR 45

1.8 ERCC6L2: a novel SNF2 ATPase in the DDR 47

1.8.1 SNF2 ATPases: a diverse family of molecular motors 47
1.8.2 ERCC6L2 49

2 Materials and Methods 55

2.1 Mammalian cell culture 55

2.2 Cloning and site-directed mutagenesis 55

2.3 Live cell imaging and laser microirradiation 59

2.4 Immunofluorescence 60

2.5 Establishing knockout cell lines by CRISPR/Cas9 61

2.6 Validating ERCC6L2$^{-/-}$ U2OS cell lines 63

2.7 Establishing stable overexpression cell lines 64

2.8 Western blotting 67

2.9 Clonogenic survival assays 69

2.10 Cell cycle analysis by flow cytometry 70

2.11 Immunoprecipitation 71

2.12 PCNA purification 74

2.13 Streptavidin pulldown of biotinylated peptides 76

2.14 Co-crystallisation of PCNA with peptides and structure resolution 77

3 ERCC6L2 is involved in the DNA damage response 78

3.1 Results 78

3.2 Discussion 90
4 Exploring the physiological role of ERCC6L2 98
4.1 Results 98
4.2 Discussion 112
5 Characterising the ERCC6L2-PCNA interaction 122
5.1 Results 122
5.2 Discussion 137
6 Discussion 143
6.1 ERCC6L2 as a DSB repair factor 145
6.1.1 ERCC6L2 deficiency phenotypes 145
6.1.2 ERCC6L2: a DSB repair factor at centromeres? 149
6.1.3 The ERCC6L2-PCNA interaction: relevant to the DSB response? 151
6.2 ERCC6L2 as a resolver of DNA secondary structures 153
6.3 Concluding remarks 158
7 Appendix 162
8 Bibliography 168

Table of figures:

Figure 1-1: Major pathways of double strand break repair 29
Figure 1-2: Replication factories and forks 33
Figure 1-3: PCNA and ZRANB3 PIP-box co-crystal structure with APIM overlaid 36
Figure 1-4. Domain structures of ERCC6 family SNF2 ATPases. 48
Figure 1-5: ERCC6L2 domain structure showing BMF-causing mutations 50
Figure 3-1: ERCC6L2 is recruited to DNA lesions 79
Figure 3-2: Clinical mutants of ERCC6L2 tested in this study 80
Figure 3-3: Clinical mutants of ERCC6L2 are deficient in recruitment to DNA damage sites 82
Figure 3-4: Generation of ERCC6L2$^{-/-}$ U2OS cell lines by CRISPR/Cas9 83
Figure 3-5: ERCC6L2$^{-/-}$ U2OS cells are sensitive to genotoxic insult 86
Figure 3-6: Establishing ERCC6L2 complementation cell lines 87
Figure 3-7: Overexpression of ERCC6L2 rescues ERCC6L2$^{-/-}$ cells' tolerance of etoposide and phleomycin in a manner dependent on its ATPase activity 90
Figure 4-1: Etoposide treatment impairs DNA synthesis in ERCC6L2$^{-/-}$ cells 99
Figure 4-2: ERCC6L2 expression restores DNA synthesis under etoposide treatment in ERCC6L2-deficient cells in a manner dependent on its ATPase activity 101
Figure 4-3: The C-terminus of ERCC6L2 is recruited to laser-induced DNA lesions 103
Figure 4-4: ERCC6L2 interacts with PCNA 104
Figure 4-5: ERCC6L2 colocalises with centromeres through its C-terminus 108
Figure 4-6: ERCC6L2 localises to centromeres throughout interphase 109
Figure 4-7: Clinically relevant ERCC6L2 truncations display defective centromere localisation 111
Figure 5-1: ERCC6L2 interacts with PCNA through a novel PCNA binding site 123
Figure 5-2: Identifying ERCC6L2's PCNA binding site 127
Figure 5-3: Co-crystal structure of PCNA with the ERCC6L2 atypical PIP-box 130
Figure 5-4: The ERCC6L2-PCNA interaction is dispensable for ERCC6L2 recruitment to DNA lesions 132
Figure 5-5: The ERCC6L2-PCNA interaction is dispensable for etoposide and phleomycin tolerance 135
Figure 5-6: ERCC6L2 Q798A partially restores DNA synthesis in ERCC6L2-deficient cells treated with etoposide 136

Figure 6-1 Working hypothesis of ERCC6L2 function.....155

Abbreviations

Abbreviation	long name
53BP1	p53-binding protein 1
a-EJ	alternative end-joining
ACD	adrenocortical dysplasia homologue
ADH5	alcohol dehydrogenase 5
ALC1	amplified in liver cancer 1
AML	acute myeloid leukaemia
ATM	ataxia telangiectasia-mutated
ATR	ataxia telangiectasia and Rad3-related
ATRIP	ATR interacting protein
BIR	break-induced replication
BLM	bloom syndrome protein
BMF	bone marrow failure
BRCA1	breast cancer type 1 susceptibility protein
BRCA2	breast cancer type 2 susceptibility protein
c-NHEJ	canonical non-homologous end-joining
C-UFBs	Centromeric UFBs
CCAN	constitutive centromere-associated network
CDK	cyclin-dependent kinase
CENP-A	centromere protein A
CFSs	Common fragile sites
CHK1	checkpoint kinase 1
CHK2	checkpoint kinase 2
CO	crossovers
CPT	camptothecin
CSB	Cockayne Syndrome protein B
DDR	DNA damage response
DDT	DNA damage tolerance
DNA	deoxyribonucleic acid
DNA-PKcs	DNA protein kinase catalytic subunit
DSB	double strand break
ERCC1-XPF	xeroderma pigmentosum group F)
ERCC6L2	excision repair cross-complementation group 6-like 2
EXO1	exonuclease 1
FA	Fanconi anaemia
FEN1	flap endonuclease 1
FS-UFBs	Fragile site UFBs
G4s	G-quadruplexes
GCRs	gross chromosomal rearrangements
GG	global genome
Hebo	helicase mutated in bone marrow failure
HJ	Holliday junction

HLTF	helicase-like transcription factor
HORs	higher-order repeats
HR	homologous recombination
HR-UFBs	homologous recombination UFBs
ICL	inter- or intra-strand crosslink
IR	ionising radiation
LIG1	DNA ligase I
LIG3	DNA ligase III
LIG4	DNA ligase IV
MDC1	mediator of damage checkpoint protein 1
MiDAS	mitotic DNA synthesis
MMC	mitomycin C
MMR	mismatch repair
MRN	Mre11-Rad50-NBS1 complex
NCO	noncrossover
NER	nucleotide excision repair
NHEJ	non-homologous end-joining
PALB2	partner and localiser of BRCA2
PARP1	poly(ADP-ribose) polymerase 1
PARylation	poly(ADP-ribose)ylation
PCNA	proliferating cell nuclear antigen
PI3K	phosphatidyl-inositol-3-kinase
PICH	PLK1-interacting checkpoint helicase/ERCC6L
PIP-box	PCNA-interaction peptide
PLK1	polo-like kinase 1
PNK	polynucleotide kinase/phosphatase
POLD3	DNA polymerase δ subunit 3
POT1	protection of telomeres protein 1
PTM	post-translational modification
RAP1	repressor/activator protein 1
RMI1/2	RecQ-mediated genome instability protein
RNF8	ring finger protein 8
ROS	reactive oxygen species
RPA	replication protein A
RTEL1	regulator of telomere elongation helicase 1
seDSB	single-ended DSB
sgRNA	single-guide RNAs
SHPRH	SNF2 histone linked PHD RING helicase
SNF2	sucrose non-fermentable 2
SSA	single-strand annealing
SSB	single strand break
ssDNA	single stranded DNA
SSEs	structure-selective endonucleases
T-UFBs	telomeric UFBs
TC	transcription coupled
TIN2	TRF1-interacting nuclear factor 2
TLS	translesion synthesis

TOPBP1	topoisomerase II-binding protein
TOPO1	DNA topoisomerases I
TOPO2	DNA topoisomerases II
TRF1	telomere repeat factor 1
TRF2	telomere repeat factor 2
TS	template switching
UFB	ultrafine anaphase bridges
UV	ultraviolet light
WT	wild type
XLF	XRCC4-like factor
XRCC4	X-ray repair cross-complementing protein 5
γH2AX	S139 phosphorylation of the histone variant H2AX
MMS	methyl methanesulphonate
NLS	nuclear localisation sequence
CRISPR	clustered regularly interspersed short palindromic repeats
KO	knock out
PI	propidium iodide
EdU	ethynyl deoxyuridine
ChIP	chromatin immunoprecipitation
a-PIP	atypical PIP-box
PFGE	pulsed-field gel electrophoresis
pol δ	DNA polymerase δ
qPCR	quantitative PCR
Pen/Strep	penicillin/streptomycin
DMSO	dimethyl sulfoxide
dNTP	deoxyribonucleotide triphosphate
min	minutes
x g	multiplied by acceleration of gravity on earth, 9.8 m/s^2
FBS	foetal bovine serum
gDNA	genomic DNA
PCR	polymerase chain reaction
ddH_2O	double-distilled water
DTT	dithiothreitol
PVDF	polyvinylidine fluoride
PBS	phosphate-buffered saline
IP	immunoprecipitation
U/µl	units per microlitre
µl	microlitre
sec	seconds
M	molar (mol/l)
l	litre
°C	degree celsius
ml	mililitre
LB	Lysogeny broth
IPTG	isopropyl β-D-1-thiogalactopyranoside
Ni-NTA	nickel-nitrilotriacetic acid
Tris	trisaminomethane

pH	$-\log_{10}([H^+])$ where H^+ concentration is measured in M
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis

1 Introduction

1.1 The DNA damage response

It is estimated that every day, each cell in the human body experiences tens of thousands of individual DNA lesions [1]. DNA lesions, which can occur in a range of different forms, can pose challenges to cells and have the potential to cause mutations and chromosomal rearrangements. However, cells have evolved a range of complex, interconnected pathways which function in what is known as the DNA damage response (DDR). Once elicited, the DDR will dictate the final consequence of a DNA lesion. The DDR can lead to successful repair of a lesion and restoration of a cell's replicative potential, but it may also signal a cell to stop dividing and enter senescence, or may even signal for cell death.

Defects in the mechanisms by which cells respond to and repair DNA damage can have disastrous consequences for human health, with a variety of diseases and cancer predispositions arising from mutations in genes whose products are involved in the DDR [1-3]. Further to the disease-causing potential of a defective DDR, a range of cancer therapies in clinical use either directly induce DNA damage or exploit specific DDR deficiencies in some tumours [2, 4, 5]. It is hoped that research into specific DDR pathways and their relationships to one another will not only yield new insights into the basis of diseases involving DNA repair defects, but also uncover new therapeutic opportunities for the treatment of these diseases.

1.2 DNA lesions: what are they and how do they occur?

A variety of different kinds of DNA lesions can be caused by a wide range of endogenous and exogenous genotoxic insults [2]. Exogenous sources of DNA damage include ultraviolet (UV) A and B light, ionising radiation (IR), naturally occurring radon gas, and chemical mutagens such as aflatoxins and those released by the burning of tobacco [2, 6].

UV causes a range of bulky DNA lesions including 6-4 photoproducts, cyclobutane-pyrimidine dimers and oxidative base damage. Such lesions are usually repaired by global genome (GG-) or, in transcriptionally-active regions, transcription coupled (TC-) nucleotide excision repair (NER) [7]. These pathways culminate in the excision of a short stretch of nucleotides containing the lesion, leaving a single stranded DNA (ssDNA) gap for gap-filling DNA synthesis using the undamaged DNA strand as a template [7]. Should bulky lesions such as those caused by UV exposure fail to be repaired by the GG- or TC-NER machinery, they can pose problems for cells actively replicating their genomes in S-phase, causing replication stress – a potentially dangerous phenomenon for a replicating cell which is discussed in more detail in Section 1.6. Briefly, these bulky lesions pose a roadblock to high-fidelity DNA polymerases, which cannot accommodate or synthesise past the damage [8]. The damage can therefore cause replication fork slowing, stalling or collapse, potentially compromising genome stability. Potentially mutagenic DNA damage tolerance mechanisms have evolved to allow replication past these lesions, leaving nicks, gaps and ssDNA behind the replication fork for normal DNA repair mechanisms to resolve. Therefore, the means by which repair and replication factors are regulated in the locale of replication forks is critical for the maintenance of genome stability in replicative cells.

Like UV, ionising radiation can cause multiple different types of DNA lesion, most notably double strand breaks (DSBs), a highly toxic lesion in which the DNA double helix is severed to expose two DNA ends. Typically, cells would repair DSBs by one of two major pathways: non-homologous end-joining (NHEJ) and homologous recombination (HR). However, on occasion alternative methods of DSB resolution are utilised, including alternative end-joining (a-EJ) and single-strand annealing (SSA), although these alternative pathways are considered relatively mutagenic [9-11]. DSB repair is described in more detail in Section 1.5.

Certain therapeutics, such as cisplatin and mitomycin C, function by inducing DNA lesions, forming inter-strand and/or intra-strand crosslinks (ICLs). These crosslinks form bulky structures, which as like UV-induced damage pose a significant challenge to replicative DNA polymerases and thus cause replication stress. The Fanconi anaemia (FA) pathway functions to repair ICLs and is so named because loss-of-function mutations in the genes of this pathway cause FA, a disease characterised by chromosomal aberrations, developmental problems, bone marrow failure and a predisposition to cancer [12].

In addition to environmental mutagens, DNA lesions can arise from normal and malfunctioning cellular processes. Of particular relevance to genome stability is DNA replication. If a replication fork collides with an unrepaired SSB or single-stranded DNA (ssDNA) gap, it can expose a 'single-ended' DSB (seDSB), resulting in a collapsed replication fork. Critically, these DSBs being single-ended means cannot be repaired by canonical NHEJ and are instead proposed to be repaired by break-induced replication (BIR) followed by HR [13, 14]. Malfunctions of the DNA topoisomerases I and II can leave also behind single and double strand breaks, respectively [15-17]. When bound to DNA, Topoisomerase I (TOPO1) nicks a single DNA strand, while each monomer of a

TOPO2 dimer cleaves a DNA strand. Normally, these enzymes would religate the breaks they induced, but if this religation fails or is inhibited, replication runoff of TOPO1 or TOPO2 can expose single-ended or double-ended DSBs, respectively [15-17].

Chemical insults arising from normal metabolism can also threaten genome integrity. Recently, it has become clear that the formation of endogenous aldehydes as by-products of multiple metabolic pathways can form adducts with DNA, which require the action of FA proteins for its repair. Deficiencies in FA proteins or the aldehyde-metabolising alcohol dehydrogenase 5 (ADH5) have been shown to render endogenous aldehydes highly mutagenic [18-20]. Furthermore, endogenous aldehydes can lead to the formation of DNA replication-perturbing DNA-protein crosslinks (DPCs), which require the activity of the protease SPRTN/DVC1 for their repair [21, 22]. Recent findings suggest that aldehydes constitute the major source of DNA lesions arising from multitude of carcinogenic compounds found in tobacco smoke [23].

With a diverse array of different DNA lesions originating from both exogenous sources and endogenous sources such as aldehydes and reactive oxygen species, it is unsurprising that cells have evolved a complex web of interconnected DNA repair pathways in order to overcome genotoxic challenges. Many of these genotoxic challenges appear to be particularly acute in the context of DNA replication, and it is becoming increasingly apparent that this process contains a number of pitfalls for genome stability. As our understanding of the DDR has increased at a mechanistic level, new technologies have also allowed the identification of previously unrecognised connections between individual pathways, as well as highlighting novel factors with as-yet uncharacterised functions in these processes. It is hoped therefore that better characterisation of such factors may reveal new therapeutic opportunities for the treatment of diseases which relate to altered cell cycle behaviours. These diseases

include both cancers and neurodegenerative disorders; while uncontrolled cell proliferation can contribute to tumourigenesis, defects in similar pathways can lead to premature ageing phenotypes. This point is underlined by the fact that multiple disorders caused my mutations in DNA repair genes both predispose to cancer and cause neurodegenerative phenotypes, such as ataxia telangiectasia [24].

1.3 Sensing and responding to DNA damage

A proliferative cell must maintain its genome integrity and replicate its genome faithfully in order to ensure the viability and functionality of daughter cells. In a multicellular eukaryote, the proliferation of cells bearing chromosomal abnormalities such as amplifications, deletions, translocations, or even small point mutations in certain genes could have disastrous consequences for the individual. Gain-of-function mutations in oncogenes or loss-of-function mutations in tumour suppressor genes can push a cell towards transformation into a cancer cell [3]. Furthermore, larger-scale genomic rearrangements may have a dramatic impact on this process, leading to up- or down-regulation of factors important for regulating cell behaviour. Therefore, it is essential that cells harbouring potentially harmful DNA lesions or chromosomal aberrations are prevented from dividing and propagating their dangerous genomes. Indeed, in non-transformed cells, sensing of DNA damage at any stage in the cell cycle will induce the DDR to limit progression until the damage is resolved or is deemed irreparable and cell death or senescence is initiated [25, 26].

Immediate sensing of DSBs induces activation of the PI3K-like (phosphatidyl-inositol-3-kinase) kinases ATM (ataxia telangiectasia-mutated). ATM recruitment to the damage site, leads to its autophosphorylation and phosphorylation of other proteins needed to recruit necessary factors and reinforce the signal to initiate the repair process.

Resection from a DSB as the upstream even committing a the lesion to repair by HR (described in Section 1.5.2) exposes ssDNA, which is rapidly coated by RPA. ATR (ataxia telangiectasia and Rad3-related) and its interaction partner ATRIP are recruited to RPA-coated ssDNA. A 'second wave' of phosphorylation events by ATM and ATR then occurs, activating a large range of proteins including the checkpoint kinases CHK1 (by ATR) and CHK2 (by ATM), which are critical for cell cycle arrest in response to DNA damage [27-29].

Classically, the ATR-CHK1 signalling axis has been associated with the cellular response to DNA crosslinks, UV damage, stalled replication forks and to resected DSBs. ATR and its interaction partner ATRIP (ATR interacting protein) sense replication protein A (RPA)-coated ssDNA, leading to its activation by topoisomerase II-binding protein (TOPBP1) and the heterotrimeric 9-1-1 complex (Rad9-Rad1-Hus1). This in turn leads to ATR's activation by phosphorylation of CHK1 at S317 and S345 [30]. On the other hand, the ATM-CHK2 signalling cascade mainly happens in response to DSBs.

Upon sensing of a DSB by the Mre11-Rad50-NBS1 complex (MRN), ATM is rapidly recruited to the break site and activated [31]. ATM is then able to phosphorylate its many substrates, including CHK2 at multiple sites, enabling CHK2 multimerisation and autophosphorylation in order to further increase its activity [30].

Interestingly, there is some degree of crosstalk between the ATR- and ATM-dependent pathways. For example, both ATM and ATR, as well as the related PI3K-like kinase DNA-PKcs (DNA protein kinase catalytic subunit), can phosphorylate CHK2 *in vitro*. Furthermore, ATM-stimulated resection of DSBs has been shown to lead to ATR and subsequent CHK1 activation, as mentioned previously [32-35]. While the crosstalk between these pathways is now established, precisely how the complex interplay of

different DDR signalling cascades can influence one another to mediate changes in cell behaviour is not completely understood.

1.4 The DDR and the cell cycle

Replication stress, which consists of aberrant replication fork dynamics such as slowing or stalling, occurs while cells are replicating their genomes in S-phase and exemplifies how the cell cycle can influence genome instability. The choice between DSB repair pathways is also heavily influenced by cell cycle stage. Understanding the nature of differential maintenance of genome stability throughout the cell cycle could be critical in understanding the basis of diseases related to DDR defects.

Cell cycle progression is the process by which a cell prepares itself for division. In order for this to occur without the loss of genetic material, the cell must first replicate its entire genome faithfully. G1 phase is the stage in which cells have not yet begun to replicate their DNA. Some cells transition from G1 into what is termed 'G0', a dormant, nonproliferative state. Accumulation of cyclin A promotes entry into S-phase, when cells then undertake DNA replication. Following the completion of replication and the resolution of any lingering problems that occurred during replication in G2 phase, cells enter M (mitosis) phase and divide, yielding two daughter cells [36, 37]. Given that cell cycle dysregulation is a major factor in cellular transformation and tumour progression, the regulatory mechanisms which govern a cell's behaviour through the cell cycle are of great interest to both basic and clinical research.

In a population of proliferative cells, a DNA lesion could occur at any stage of the cell cycle, whether from endogenous or exogenous sources. Of particular relevance to human biology are genotoxins which selectively target S-phase cells, once they ostensibly committed to dividing. Such genotoxins – such as cisplatin and mitomycin C

– can prove extremely useful in the clinic against cancers, but also may be particularly harmful to other highly proliferative, non-cancerous tissues such as bone marrow and thus cause severe side effects. Interestingly, recent findings suggest that the PARP1 inhibition causes replication fork speed to increase, presumably compromising the fidelity of DNA replication in BRCA-deficient cancers [38]. However, it is an important question in human health to understand the means by which non-transformed cells respond to genotoxic insults and are kept from dividing. Cell cycle checkpoints have evolved to halt the cell cycle if a potentially problematic DNA lesion is detected. A simplified cartoon of DNA damage-induced cell cycle checkpoint signalling is shown in Figure 1-1.

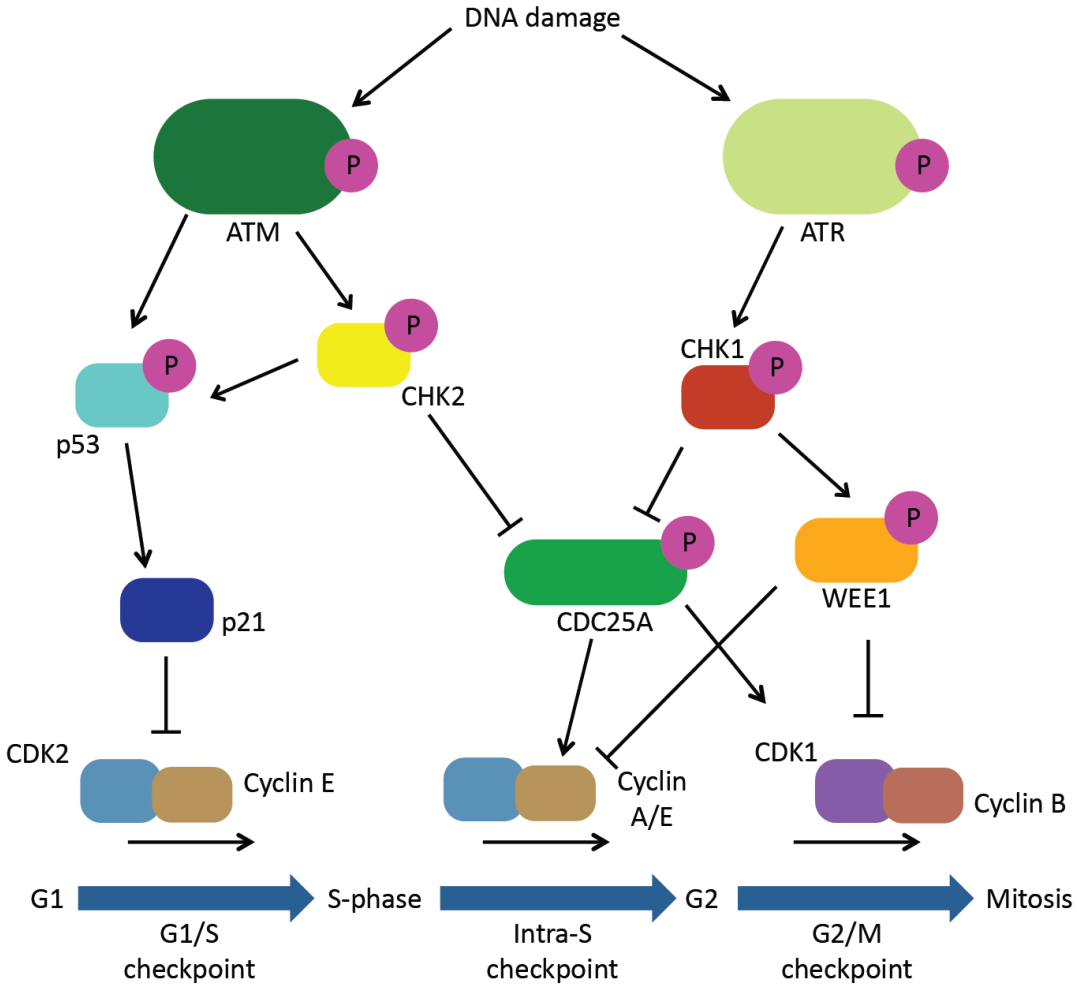


Figure 1-1 Cell cycle checkpoint signalling by the ATM-CHK2 and ATR-CHK1 axes

Cartoon showing the basic principles of checkpoint signalling downstream of DNA damage-induced activation of ATM and ATR.

1.4.1 The G1/S checkpoint

In response to DNA damage in G1, ATM-mediated phosphorylation and activation of CHK2 or ATR-mediated activation of CHK1 promotes the stabilisation of the transcription factor p53, while both ATM and ATR also phosphorylate and activate p53 directly on Ser15. Stabilisation and activation of p53 leads to the induction of a wide range of DDR response genes, including the critical cell cycle regulator p21 [39, 40]. As a result, p21 accumulates and binds to cyclin-CDK (cyclin-dependent kinase) complexes, preventing their promotion of cell cycle progression. Notably, this arm of

the G1 checkpoint is entirely dependent on both p53 and p21 [41-43]. This disadvantage of this mechanism for cell cycle arrest is that it is relatively slow, relying on the action of the transcription and translation machineries to produce enough p21 to counter cyclin-CDK activity. However, ATM-mediated activation of the p38 MAPK pathway mitigates this somewhat by stabilising p21 transcripts [44-46]. In additional to this slow arm of the G1 checkpoint, the activity of the ATM, CHK2 and p38 kinases also promote degradation of cyclin D, thus restricting entry into S-phase [44, 46-50].

1.4.2 The intra-S checkpoint

S-phase is a difficult time for a dividing cell; the entirety of its genome must be copied faithfully and packaged into chromatin amenable to chromosome segregation and cell division. The replication machinery itself can make mistakes and incorporate incorrect nucleotides or even sometimes ribonucleotides, sometimes requiring the action of downstream pathways to follow behind the replication fork and correct such errors [51]. Furthermore, relatively innocuous nicks in DNA if not recognised and repaired could result in collapsed replication forks and the generation of seDSBs. S-phase represents a fascinating cell cycle stage in terms of the DDR; not only is replication stress a potential threat, but DSB repair is mediated by one of several competing pathways with differential regulation through the cell cycle, and factors integral to the replication fork itself can act as primary sensors and upstream mediators of the repair of other lesions.

Like G1, S-phase has its own checkpoint, which acts to halt cell cycle progression upon sensing of DNA damage; this is termed the intra-S checkpoint. The intra-S checkpoint relies on the activity of the kinase WEE1, which upon phosphorylation (activation) by CHK1, phosphorylates and inactivates CDK2, preventing further cell cycle progression

[52, 53]. Furthermore, CHK1 and CHK2 phosphorylate and inactivate the CDK2 phosphatase CDC25A, preventing its antagonism of WEE1 activity [54-56].

1.4.3 The G2/M checkpoint

The G2/M transition is regulated by factors which influence the accumulation of active CDK1-cyclin B complexes, which are required for mitotic entry and progression. During interphase, CDK-cyclin B accumulates at centrosomes in the cytoplasm, where CDK1 is inactivated by phosphorylation on Tyr15 by WEE1. Similarly to the intra-S checkpoint, this activity of WEE1 is antagonised by CDC25A's phosphatase activity, which removes the inhibitory phosphorylation mark on CDK1 Tyr15. CHK1 and CHK2 exert their control over the G2/M transition by phosphorylating and inactivating CDC25A at multiple sites [57].

As with the intra-S checkpoint, the kinase activity of WEE1 is central to the G2/M checkpoint, which prevents cells harbouring extensive DNA damage from entering mitosis [44, 58]. Failure to halt such cells could lead to defective chromosome segregation and aneuploidy, which can play a role in cellular transformation [44, 59]. The G2/M checkpoint differs from the G1/S checkpoint, in that while the G1/S checkpoint depends solely on ATM-CHK2 signalling in response to DSBs, the kinase activities of ATR and CHK1 are essential for maintenance of the G2/M checkpoint [27, 60-62]. However, ATM-CHK2 still appears to contribute to acute cell cycle arrest in G2 after sensing of damage [63, 64]. Considering the increase in the number of different repair pathways employed by cells during S-phase and G2, it is unsurprising that multiple different signalling pathways can feed into cell cycle arrest at the G2/M checkpoint. Understanding how these pathways influence each other in the multitude

of genetic backgrounds observed in cancer cells could be critical to our understanding of genomic instability in cancer.

1.4.4 The DDR in mitosis

During mitosis, the canonical response to DSBs is suppressed. While the upstream kinase ATM is still activated following DSB induction, CHK2 is not activated and various other DNA repair factors are not recruited to the damage site [65, 66]. Furthermore, CHK1 activity is suppressed by the degradation of the ATR-CHK1 bridging factor Claspin, suggesting that once mitosis is initiated, activation of the checkpoint kinases is undesirable. This may be because CHK1/2 would counteract the CDK1 activity required for mitotic progression [44, 66-69]. In addition to limiting effects on the checkpoint kinases during mitosis, the activity of key DDR mediators is suppressed at damage sites. Polo-like kinase 1 (PLK1) plays an important role in mitosis, regulating a range of other kinases which together mediate several key events including spindle assembly, the activation of the anaphase-promoting complex and in promoting the removal of cohesins from chromosome arms [70]. In addition to these roles, PLK1 phosphorylates the critical NHEJ factor 53BP1 to limit its mobilisation to damage sites, while CDK1 phosphorylates the DNA damage-responsive E3 ubiquitin ligase RNF8 (ring finger protein 8) to inhibit its establishment of K63-linked polyubiquitin chains at DNA lesions [66, 71, 72]. PLK1 and CDK1 also suppress NHEJ through inactivating phosphorylation of the key NHEJ scaffold protein XRCC4 during mitosis [73]. However, the activity of PLK1 is antagonised by both ATM and ATR-dependent signalling mechanisms [74]. Interestingly, it has been suggested recently that ATR performs a function in mitosis which is distinct from its role in DNA repair and replication stress. Through the formation of centromeric R-loops and Aurora A kinase activity, ATR is

recruited to mitotic centromeres where it is proposed to phosphorylate and activate CHK1 at centromeres, which promotes faithful chromosome segregation by maintaining Aurora B activity [75].

Overall, although the precise details of how the DDR is regulated during mitosis are far from complete, it certainly appears that there is a significant alteration to the mounting of a normal DNA damage response, both in terms of checkpoint activation and mobilisation of DNA repair factors to damage sites.

1.4.5 Replication and recombination intermediates in mitosis

In order for faithful segregation of chromosomes to occur during anaphase, any physical connections between sister chromatids must be disassembled in a timely manner. Such connections could be proteinaceous, as in the case of cohesion mediated by the critical scaffold Cohesin [76], or could involve DNA structures resulting from replication or recombination intermediates. Such DNA 'bridges' can lead to the formation of micronuclei, which upon rupture are proposed to contribute to metastasis of cancer cells through the activation of cytoplasmic DNA-sensing pathways [77, 78].

DNA connections between sister chromatids form thin strands between segregating chromosomes in anaphase and are thus referred to as ultrafine anaphase bridges (UFBs). These can be visualised by immunofluorescent staining against BLM, RPA or the SNF2 ATPase PICH (PLK1-interacting checkpoint helicase) [79-81].

UFBs are classified based on their origin, and it is generally accepted that there are three main classes of UFB, with an additional one identified more recently. Centromeric UFBs (C-UFBs) result from double-stranded DNA catenanes which can by enriched for by topoisomerase II inhibition. C-UFBs are very common occurrences in anaphase, forming between centromeres meaning they are identifiable by positive

immunostaining of centromeric markers such as the centromeric histone variant CENP-A (centromere protein A) at either end of the bridge [82-84]. Fragile site UFBs (FS-UFBs) form between common fragile sites (regions of the genome prone to breakage, described in Section 1.7.2) in anaphase and are far less common than C-UFBs, but are inducible by replication stress conditions and are thought to form between sites of under-replicated DNA. FS-UFBs can be visualised by the formation of pairs of FANCD2 or FANCI foci at the UFB termini [85-89]. Telomeric UFBs (T-UFBs) are coated by the BLM helicase, form between telomere ends and are thought to result from similar problems in replication as FS-UFBs, since with their repetitive sequences telomeres are considered difficult to replicate. Like FS-UFBs, T-UFBs are induced by replication stress conditions such as aphidicolin treatment [84, 89-92]. Homologous recombination UFBs (HR-UFBs) were identified more recently [93]. Covalently-linked DNA intermediates form between sister chromatids in homologous recombination, and in cells lacking the nucleases GEN1 and MUS81, RPA-coated ssDNA UFBs form following BLM-mediated dissolution of these structures [93]. Given that chromosome segregation defects have a dramatic impact on genome stability, it is important to understand the mechanisms in place for safeguarding this process.

PICH is a critical player in UFB recognition and resolution and was first identified as an interaction partner of the mitotic kinase PLK1. It is excluded from the nucleus during interphase, but then associates with centromeres following the breakdown of the nuclear envelope. PICH recognises UFBs and appears to support the recruitment of other UFB-processing proteins to these structures [79, 81, 89, 94]. PICH is a member of the SNF2 family of ATPases, possessing the characteristic ATPase motor domain, which couples ATP hydrolysis with DNA translocase activity (Figure 1-6) [95, 96]. Interestingly, the increased tension under which UFBs are placed by the mitotic spindle

appears to increase the affinity of PICH for UFBs, whose timely resolution appears to rely on the PICH's ATPase activity [79, 82, 96].

The BTR complex, as discussed further in 1.5.2, is comprised of the BLM helicase, Toposiomerase IIIα and RMI1/2, dissolves Holliday junctions in S/G2. In the event that this dissolution fails or other replication or recombination structures such as D-loops (DNA triplexes formed by RAD51-mediated strand invasion) persist, the structure-selective endonucleases (SSEs) MUS81-EME1, SLX1-SLX4 and GEN1 act to remove these structures [97-105].

SLX1-SLX4 and MUS81-EME1 cooperate in this process, forming a complex along with XPF-ERCC1 called the SMX complex [106, 107]. This trinuclease complex forms during prometaphase in response to phosphorylation of SLX4 and MUS81 by PLK1 and CDK1, respectively [107, 108]. The formation of this complex, thus limited to mitotic cells, can cleave any recombination or replication intermediates which have persisted from S/G2 into early mitosis [106]. On the other hand, GEN1 acts later in mitosis. Normally restricted to the cytoplasm, GEN1 can only access mitotic chromosomes after the breakdown of the nuclear envelope when it can perform a similar function to the SMX complex in terms of resolving difficult recombination or repair intermediates [109, 110].

Persistent intermediates from DNA replication and repair can pose a problem for segregation. These structures can lead to the formation of UFBs in anaphase between difficult-to-replicate chromosome regions or between sites of repair intermediates in the absence of the nucleases that would normally cleave them. How replication stress or DNA damage might affect genome stability through predominant effects on certain genomic locations is still a matter of close investigation.

1.5 DSB repair

As mentioned previously, DNA double strand breaks (DSBs) are highly toxic DNA lesions which are usually repaired by one of the two major DSB repair pathways: non-homologous end-joining and homologous recombination, summarised in Figure 1-2. The choice between these two pathways is largely dependent on the cell cycle and can have profound effects on genome stability.

1.5.1 Non-homologous end-joining

Repair of a DSB in G1 occurs primarily through non-homologous end-joining (NHEJ). Canonical NHEJ (c-NHEJ) begins with the binding of the KU70/80 heterodimer to blunt DNA ends, protecting them from resection and tethering them to keep them in close proximity [111, 112]. These dsDNA-KU70/80 complexes recruit the PI3K-family kinase DNA-PKcs, which acts to promote blunt-ended ligation by DNA ligase IV (LIG4), aided by the scaffold proteins XLF (XRCC4-like factor) and XRCC4 (X-ray repair cross-complementing protein 5) [113-115]. A critical mediator of this process is 53BP1 (p53-binding protein 1), which interacts with DSBs at chromatin in a manner dependent on phosphorylation of the histone variant H2AX (referred to as γH2AX), binding through its tandem Tudor domains to methylated histones at DSB sites, and its interaction partner MDC1 (mediator of damage checkpoint protein 1). 53BP1 acts to promote c-NHEJ by suppressing resection at DSBs in G1 [116-118], thus preventing the initiation of competing resection-mediated pathways [119]. Upon entry to S-phase, several mechanisms appear to promote nucleolytic processing of DNA ends and subsequent out competition of NHEJ by HR, as will be discussed in 1.5.2.

Despite the importance of c-NHEJ, an alternative end-joining (a-EJ) pathway was proposed to exist following the observation that yeast deficient in KU70 or LIG4 still exhibit residual end-joining activity, which appeared to utilise short regions of microhomology [120, 121]. Similar observations have been made since in a range of other species – including humans – and it appears that microhomology is often used in a-EJ [10, 122-124].

Although the exact mechanisms of a-EJ are poorly understood, assuming similarity in principle to c-NHEJ, it has been suggested that a-EJ probably comprises three broad stages. First, recognition and tethering of DNA ends in a KU-independent manner. Second, processing of these DNA ends in order to render them ligatable. Third, ligation of processed DNA ends independently of LIG4 – presumably by DNA Ligase I or III (LIG1 or LIG3) in mammals [122].

Multiple studies have suggested a role for PARP1 (poly(ADP-ribose) polymerase 1) in a-EJ. PARP1 is a well-established SSB/DSB sensor, catalysing poly(ADP-ribose)ylation (PARylation) of both itself and a range of other proteins including histones at sites of DNA damage to mediate the recruitment of DNA repair and chromatin remodelling factors [122, 125-128]. In a-EJ, PARP1 tethers DSB ends together antagonistically to KU70/80, although in most cellular settings the extremely high affinity of KU70/80 for DNA ends means that PARP1 is usually outcompeted. Indeed, KU70/80 inhibits PARP1 recruitment to DSBs, PAR synthesis and exposure of ssDNA [122, 129-131]. In addition to its tethering function, PARP1 may also play a role in directly or indirectly recruiting other important a-EJ factors to the damage site, such as XRCC1, LIG3, PNK (polynucleotide kinase/phosphatase) and the MRN complex [122, 124, 129, 131, 132].

DNA end processing is the next critical step for a-EJ. Similarly to the initial steps of HR, short ssDNA stretches are exposed at the break site by the action of CtIP and the

MRN complex, which act in a 5’ to 3’ manner. This means that unlike c-NHEJ, a-EJ, while active in G1, increases in S/G2 along with HR [122, 133-136]. Following resection, gap-filling DNA synthesis is thought to be mediated by DNA polymerases λ and θ, while further processing of the DNA ends by the nucleases FEN1 (flap endonuclease 1) and ERCC1-XPF (xeroderma pigmentosum group F) has also been suggested [122, 129, 137-143]. Finally, phosphorylation of 5’ DNA ends by PNK in a manner dependent on XRCC1 and PARP1 renders the DNA ends ligatable [122, 129, 132, 143].

While LIG4 is critical for c-NHEJ, it appears to have no role in a-EJ, suggesting that one of the other two mammalian DNA ligases - LIG1 or LIG3 - performs the ligation step in a-EJ [122, 144]. LIG3 is proposed to be the dominant ligase of the two, favouring microhomology-based a-EJ, while in the absence of LIG3, LIG1 can take over and can act independently of microhomology. However, relatively little is known about the regulation of these subpathways, which could act differently in distinct contexts [122-124, 129, 145-150].

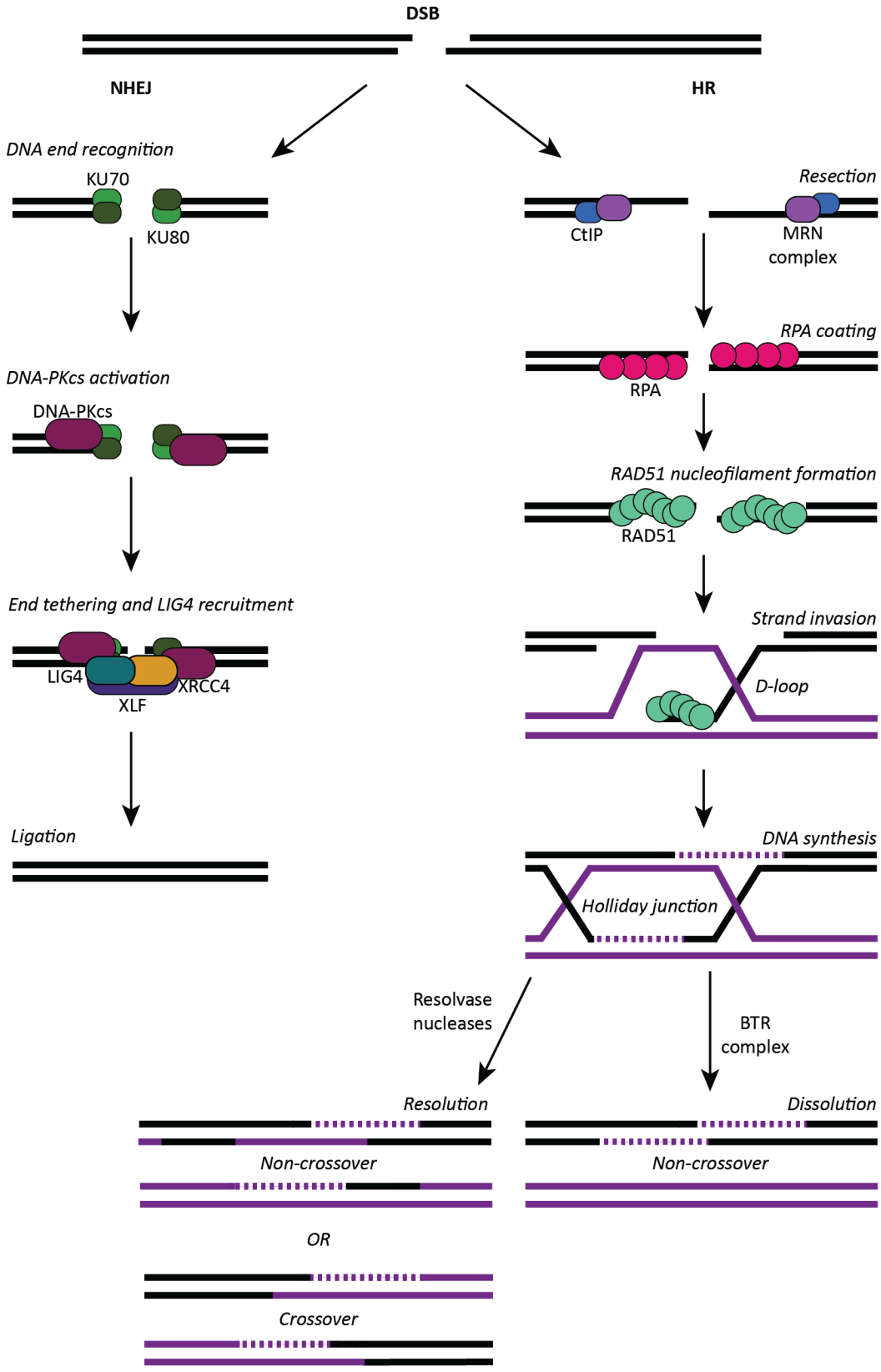


Figure 1-2: Major pathways of double strand break repair

Diagram of nonhomologous end-joining (NHEJ) and homologous recombination (HR), the two major pathways used to repair double strand breaks. HR can result in the formation of non-crossover or crossover products, depending on how Holliday junctions are disassembled.

1.5.2 Homologous recombination

While a DSB occurring in G1 is repaired predominantly by c-NHEJ due to tight restriction of resection, S-phase cells also permit the activity of the nucleases required for resection, enabling HR and other resection-mediated pathways to come into play [119]. To initiate HR, extensive resection occurs at the DSB site, mediated by EXO1 (exonuclease 1), DNA2, the MRN complex and the endonuclease CtIP, in a manner dependent on phosphorylation of CtIP by CDK2 at the G1-S transition. As such, HR can only be initiated during S-phase or G2 when there is a sister template available for faithful repair [151-153]. The tumour suppressor BRCA1 (breast cancer type 1 susceptibility protein) is another critical positive regulator of HR, promoting resection by antagonising the interaction of 53BP1 with damaged chromatin and thus relieving 53BP1-mediated suppression of resection [152]. Resection from the damage site exposes ssDNA for RPA coating, which leads to activation of the ATR-CHK1 pathway alongside ATM-CHK2 pathway [44, 154].

In addition to its upstream role in promoting HR by counteracting 53BP1, BRCA1 also acts after resection to recruit BRCA2 via PALB2 to RPA-coated ssDNA, where they cooperate to exchange RPA for the recombinase RAD51. RAD51 binding to DNA forms a nucleoprotein filament capable of strand invasion [155-159]. PALB2 and BRCA2 are thought to then facilitate strand invasion of the RAD51 nucleoprotein filament into a sister chromatid, leading to the formation of a DNA joint molecule called a D-loop. Anti-recombinases such as RTEL1 (regulator of telomere elongation helicase 1) can function at this stage to antagonise strand invasion events [160, 161]. Persistence of a recombinogenic DNA joint molecule gives rise to the formation of a four-way repair intermediate known as a Holliday junction (HJ). In this structure, the 3’ end of the

invading (damaged) DNA strand can act as a primer for DNA synthesis, extending a D-loop structure complementary to the undamaged template [162-167].

Following repair synthesis, tight regulation of HJ disassembly is critical. In meiosis, crossovers (CO) are a desirable outcome of inter-homologue exchanges in order to drive genetic diversification within the germ line, but in non-meiotic cells, COs could cause gross chromosomal rearrangements or cause loss of heterozygosity of mutated tumour suppressor genes. Therefore, noncrossover (NCO) recombination products are favoured in non-meiotic cells [161, 168, 169]. In humans, the BTR complex, composed of BLM (Bloom syndrome protein), Topoisomerase IIIα, RMI1 and RMI2 (RecQ-mediated genome instability protein), acts to migrate along and decatenate HJs in a process termed HJ dissolution, giving rise to NCO recombination products [105, 161, 170-172]. Should the BTR complex fail to resolve an HJ, this could pose a major problem for chromosome segregation. To combat this, a host of nucleases (including SLX1-SLX4, MUS81 and GEN1) have evolved to resolve persistent HJs, although they can give rise to both CO and NCO products. Therefore, their activity is tightly restricted outside of mitosis.

The availability of HR as means of repairing DSBs in S-phase gives rise to a complex competition between repair pathways, which is still under intense investigation. To add further depth to this complexity, normal replication comes with its own challenges to genome stability. Genotoxic insult can cause widespread challenges to replication forks, and the precise nature of their interplay with DNA lesions can have drastic consequences for cell fate.

1.6 Replication stress: a major source of genomic instability

The phenomenon of replication stress, which is emerging as a complex and dangerous source of genome instability in S-phase, has been proposed as an emerging hallmark of cancer [173, 174]. While difficult to define precisely, replication stress is a state in which a cell is experiencing abnormal levels of replication fork slowing or stalling, or reduced DNA synthesis. These events often – although not always – result in the accumulation of ssDNA due to uncoupling of the replicative helicase from the replicative polymerase [51, 175]. The causes of replication stress, which can occur locally or across the entirety of an S-phase nucleus, include by-products of cellular metabolism such as aldehydes, DNA lesions caused by UV light, chemical mutagens, collisions with transcription machinery and even particular DNA structures encountered by the replication machinery [51, 176-178]. These include secondary structures formed in GC-rich sequences, G-quadruplexes (G4s), repetitive sequences and R-loops formed as a result of replication-transcription collisions [51, 179-184]. It has been suggested that chemical stabilisation of difficult-to-replicate sequences or the loss of factors which would resolve them, can contribute to toxic replication stress [51].

Eukaryotic DNA replication is initiated upon firing of 1000s of 'licensed' origins, the timing of which is tightly regulated with some origins firing early and some firing late. Two replication forks move bidirectionally away from a licensed origin and meet others to complete replication across the entire genome. However, not all origins fire; many lie dormant and can be activated under replication stress in order to help complete replication [51, 185-188]. Sites of on-going DNA replication are relatively easy to visualise by immunofluorescence against proteins involved in replication, since replication forks cluster together into 'replication factories' [189]. This, along with a simplified cartoon of a replication fork, is shown in Figure 1-3.

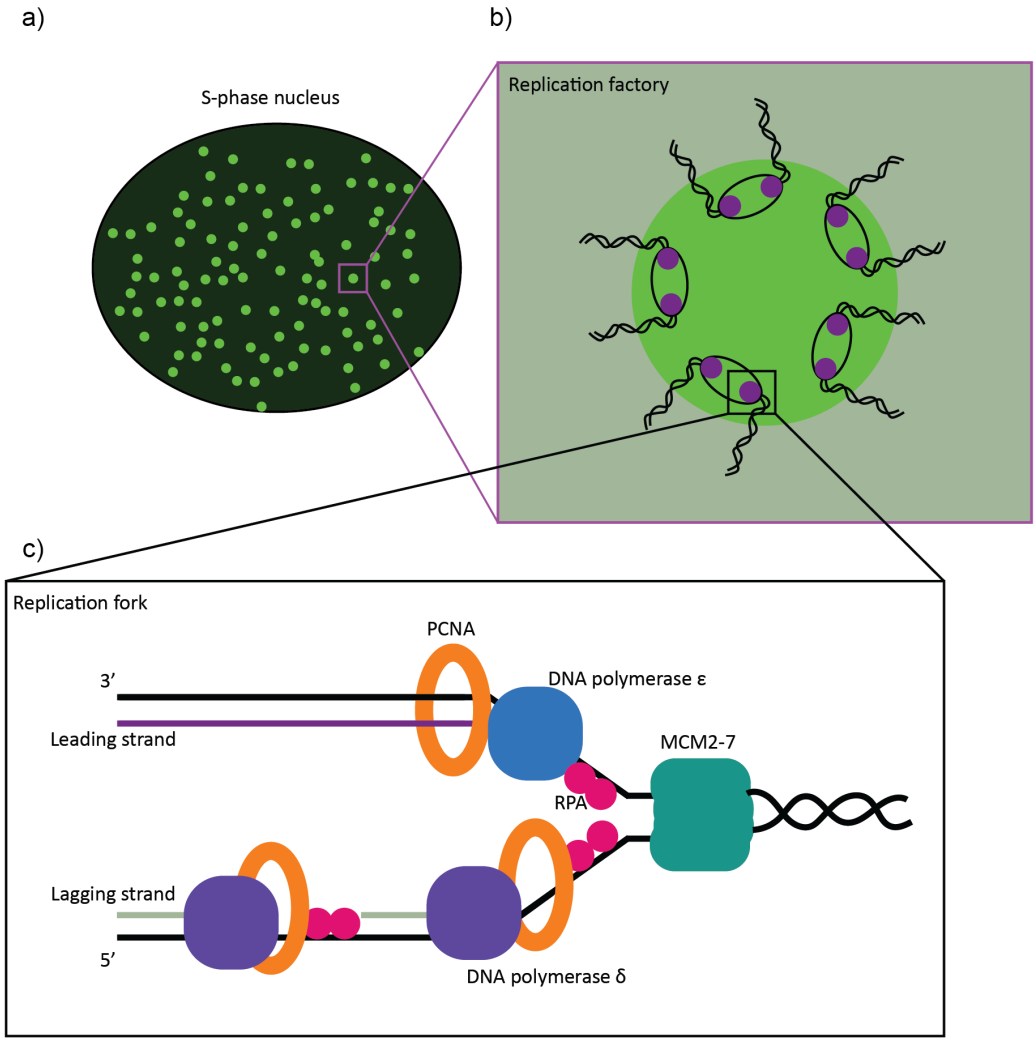


Figure 1-3: Replication factories and forks

a) Representation of an S-phase nucleus immunostained against a DNA replication factor such as PCNA. b) Representation of a DNA replication factory, comprising multiple sites of on-going bi-directional replication. c) Cartoon of a replication fork. The replicative helicase (MCM2-7) unwinds the double helix ahead of the leading strand polymerase DNA pol ε and the lagging strand polymerase DNA pol δ, both of which are stabilised at the replication fork by PCNA. RPA coats singe-stranded DNA to protect it from breakage.

As noted, replication stress often results in the exposure of stretches of ssDNA, which is coated by RPA, protecting it to some extent from passive breakage. This results in the recruitment and activation of the ATR-CHK1 pathway, which halts cell cycle progression and suppresses the firing of late-replicating origins [29, 51, 154, 190-193]. In addition to suppressing firing of late-replicating origins, ATR signalling also promotes the firing of dormant origins near stalled forks as well as stabilising existing stalled forks [51, 186-188, 194]. Overall, these downstream effects of ATR signalling

enable the completion of replication across challenging sequences or through sites bearing lesions before the cell proceeds through S/G2 and into mitosis.

In some cases – usually in the presence of a fork-blocking lesion – DNA damage tolerance (DDT) or translesion synthesis (TLS) may be used to continue replication across a damaged site. DDT involves repriming of DNA synthesis downstream of the fork-blocking lesion, leaving behind it a stretch of ssDNA for post-replicative repair [195, 196]. In contrast, TLS continues replication through the lesion, but the high-fidelity replicative polymerase must be exchanged for a more flexible TLS polymerase such as Rev1, pol η or pol κ which can accommodate the lesion but is more likely to leave behind replication errors [195-197]. This is of clinical relevance in the case of xeroderma pigmentosum-variant (XP-V), in which loss-of-function mutations in *POLH* – the gene encoding DNA polymerase η – sensitise skin cells to UV-induced DNA damage and predispose patients to skin cancers [198, 199]. A further, error-free pathway is known as template switching (TS). This poorly understood homology-directed DDT mechanism allows replication to proceed by switching from the damaged DNA strand to the undamaged one. However, TS leaves in its wake small gaps, which must later be repaired separately using the newly synthesised strand as a template [196, 200, 201].

Central to the regulation of these pathways is the sliding clamp PCNA (proliferating cell nuclear antigen), which encircles DNA and acts as a scaffold for the recruitment of DNA polymerases to an unperturbed replication fork as well as the recruitment of large number of different DNA replication and repair proteins in a context-specific manner [197]. In order to encircle DNA, PCNA forms a homotrimer with a central pore of around 35Å and a net positive charge, enabling it to slide along DNA independently of sequence. Each monomer consists of two globular domains connected by an unstructured loop known as the inter-domain connecting loops (IDCL) [197, 202, 203].

PCNA-protein interactions are of critical importance to DNA replication and repair, and a large number of proteins contain sequences matching PCNA interaction motifs. The best studied of these motifs, found in at least 200 proteins, is the PCNA interacting peptide (PIP) box, bearing the signature Q-x-x-[VILM]-x-x-[FY], although atypical variants on this core sequence have been identified [197, 204]. The first well-characterised interaction between PCNA and a PIP-box was that with the PIP-box of p21, which binds to PCNA at the IDCL but also extends to contact both of the globular domains [197, 205]. Interestingly, despite the sequence similarity between the p21 PIP-box and those of both the DNA pol δ subunit p66 and the flap endonuclease FEN1, the p21 PIP-box interacts with PCNA with roughly a 200-fold greater affinity. This suggests a degree of sequence hierarchy between PIP-boxes, as well as underlining p21's role as a negative regulator of cell cycle progression [197, 205, 206]. Another well known but less-studied PCNA interaction motif is the APIM (AlkB homologue 2 PCNA interacting motif), bearing the signature [KR]-[FYW]-[LIVA]-[LIVA]-[KR], which is also found in a large number of proteins [207]. Despite the differences in sequence between PIP-boxes and APIM motifs, recent findings indicate that they show similar binding properties. An APIM-containing peptide from the SNF2 ATPase ZRANB3 was found to bind the same site on PCNA's IDCL as a PIP-box-containing peptide also from ZRANB3 [208]. With a very large pool of identified and possible PCNA interactors residing in the nucleus and only three PIP- or APIM-binding sites per homotrimer, it seems likely that a number of factors beyond simple sequence hierarchy must regulate PCNA-protein interactions. The interaction between PCNA and the PIP-box and APIM of ZRANB3 is shown in Figure 1-4, using the structures published in [208].

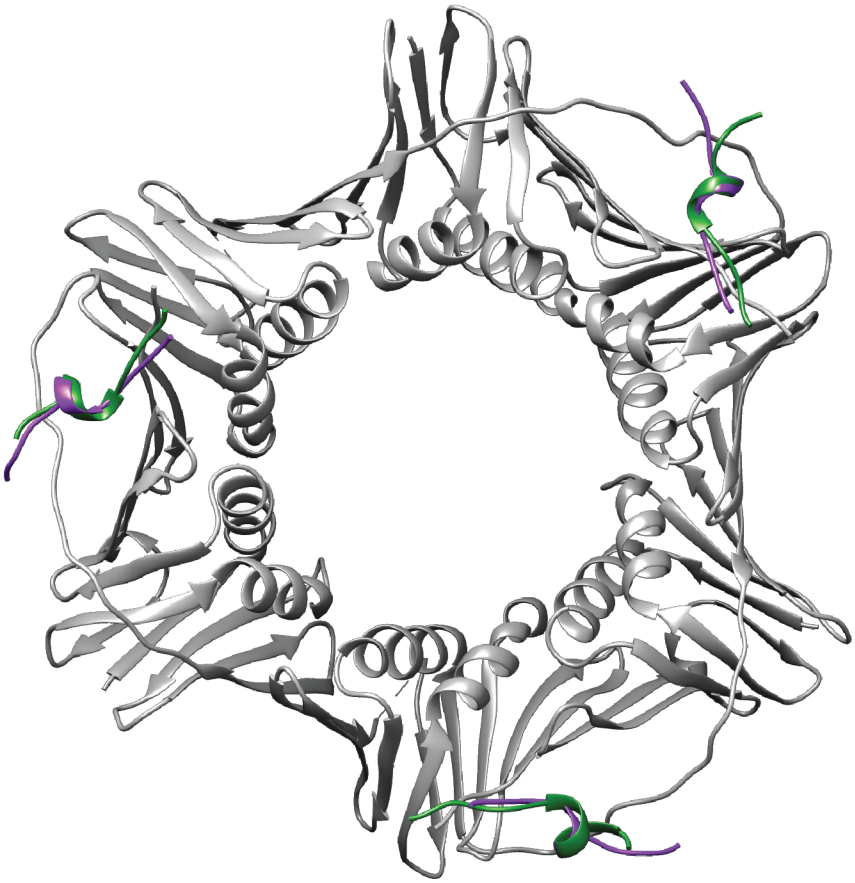


Figure 1-4: PCNA and ZRANB3 PIP-box co-crystal structure with APIM overlaid

The co-crystal structure of PCNA (dark grey) with ZRANB3 PIP-box peptides (purple). The APIM of ZRANB3 (green) is overlaid. These structures are published in [208].

Post-translational modifications (PTMs) of PCNA such as ubiquitylation or sumoylation add a layer of regulation to PCNA-directed recruitment of proteins to regulation forks. Mono-ubiquitylation of PCNA at K164 by Rad18 signals for recruitment of a TLS polymerase, while extension of this in a Lys-63-linked manner by Rad5 in yeast initiates the template switching DDT pathway [209-211]. This is summarised in Figure 1-5. TS has mostly been studied in yeast, which utilise Rad5 for this pathway. It is only more recently that human homologues of Rad5 have been discovered. The SNF2 family ATPases HLTF (helicase-like transcription factor) and SHPRH (SNF2 histone linked PHD RING helicase) are now thought to be the mammalian homologues of Rad5, sharing with it 39% and 21% sequence similarity, respectively [212-216]. Functional evidence also exists supporting this functional homology. Both

SHPRH and HLTF have been shown to interact with the human E3 enzyme Rad18 and polyubiquitylate mono-Ub-PCNA *in vitro*, while also promoting PCNA polyubiquitylation in mammalian cells [212-216]. Furthermore, there is *in vitro* evidence that HLTF can reverse fork-like DNA structures as well as promoting strand invasion and D-loop formation, consistent with some of the fork manipulations likely to be required during TS [217, 218]. Why higher eukaryotes appear to possess two homologues of *S. cerevisiae* Rad5 is unclear. However, given that higher eukaryotes have more complex genomes and aberrant regulation of template switching has been associated with complex chromosomal rearrangements in humans [219-223], it might be that higher eukaryotes have evolved as-yet unrecognised nuances in this process.

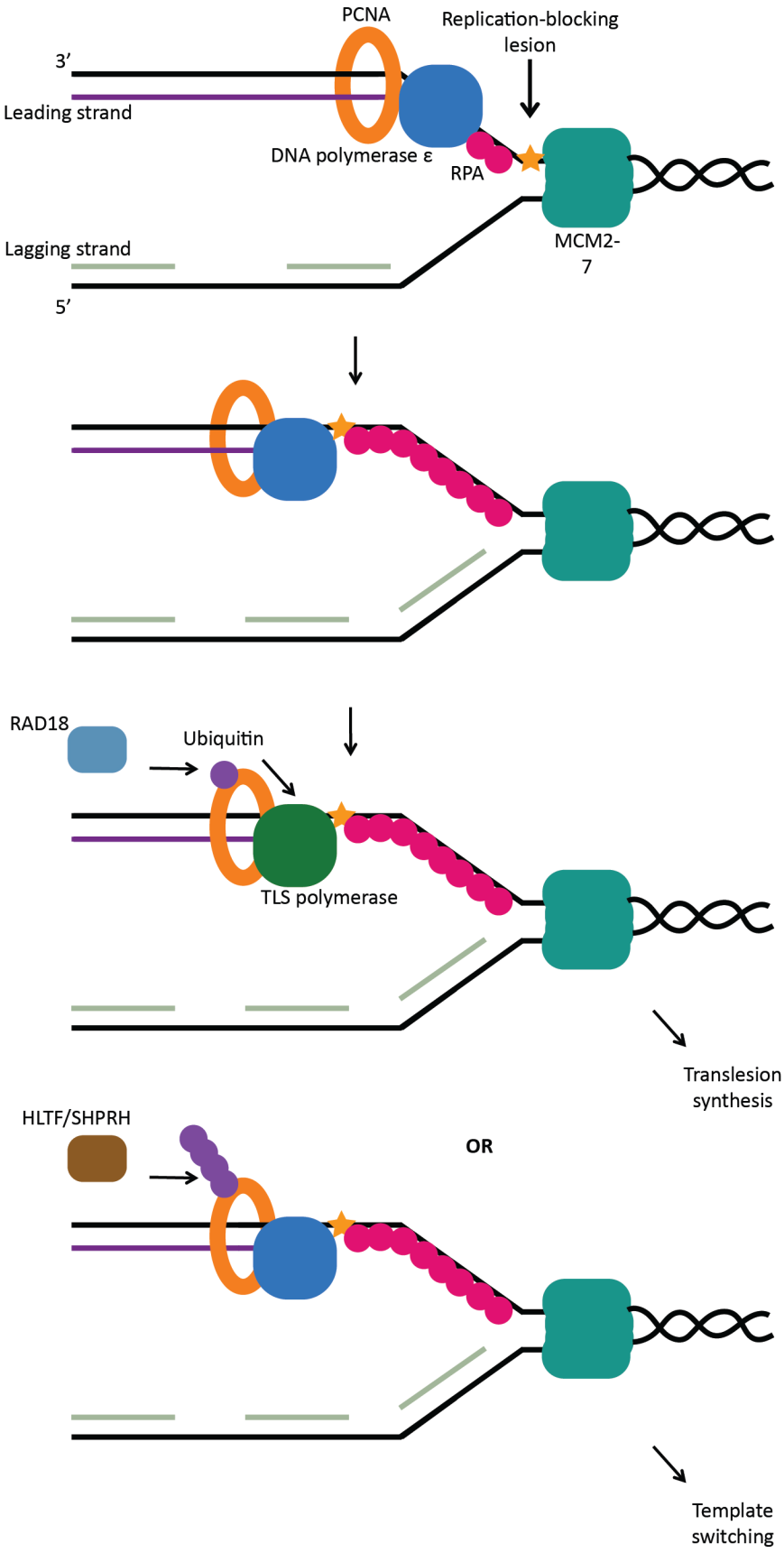


Figure 1-5 PCNA ubiquitylation in DNA damage tolerance pathway choice

Another SNF2 ATPase – ZRANB3 – also appears to be an important interactor of PCNA, with its depletion sensitising cells to a range of genotoxic insults. Unlike any TLS polymerases, ZRANB3 interacts specifically with polyubiquitylated PCNA through PCNA-binding (PIP and APIM) and ubiquitin-binding (NZF) motifs. This interaction targets ZRANB3 to stalled forks where it performs a structure-specific nuclease activity, cleaving flap-like DNA structures [224-226]. More recent studies demonstrate that this nuclease activity is stimulated by its interaction with PCNA *in vitro* [208].

Should the replication stress response fail to resolve persistent replication challenges, fork collapse can occur, resulting in DSB formation and ATM/DNA-PK activation; the response to DSBs is described in Section 1.5 [51, 227, 228]. DSB formation during replication stress is thought to occur through both passive breakage of vulnerable ssDNA at stalled forks, and through the action of misregulated nucleases such as MUS81 if the ATR pathway is defective [51, 196, 229].

Interestingly, the genome contains hotspots for replication-coupled genomic instability. Common fragile sites (CFSs) are more sensitive than most regions of the genome to even mild replication stress (see Section 1.7.2) and the precise nature of the replication stress response appears to be unusual at these sites [230]. While ATR activity helps to stabilise stalled forks at CFSs, it doesn't appear to halt the cell cycle, suggesting that at these sites some degree of stalling and even breakage may be tolerated [51, 231]. It has been suggested that the sensitivity of CFSs to replication stress can be explained by a relative dearth of replication origins, limiting the cell's ability to rescue stalled forks simply by firing nearby dormant origins [51]. Notably, ssDNA found at these stalled forks at CFSs do not appear to break passively, but may instead form unusual replication or repair intermediates which are cleaved by endonucleases such as MUS81 or ERCC1. While such events remain poorly understood,

it is thought that this controlled cleavage may be preferable to passive breakage in terms of maintaining genomic integrity [51, 87, 88]. Overall, the concept of genomic location having a strong influence on replication and the local replication stress response varying somewhat from held norms is an exciting avenue of current research and may have a dramatic impact on our understanding of replication stress in cancer.

1.7 The influence of genomic location on the DDR

Genomic location can have a strong influence on the occurrence, tolerance or repair of DNA damage. As mentioned in 1.6, collisions between replication and transcription machinery could feasibly be more common in transcriptionally active regions. The effects of this may be exacerbated in regions with scarce origins of replication since the firing of dormant forks is not an option, as is proposed to be the case for common fragile sites (CFSs) [51]. The chromatin environment also has a dramatic impact on the DDR. Not only do a range of chromatin remodelling factors and histone chaperones play essential roles in the DDR, but also chromatin compaction state is proposed to have a strong influence on repair pathway kinetics and therefore pathway choice [128, 232-236]. Incorrect pathway choice and regulation, particularly when concerning recombination-mediated repair at repetitive sequences, can lead to gross chromosomal rearrangements (GCRs) such as translocations and large insertions or deletions [237-240].

Telomeres and centromeres are both chromosomal structures bearing highly repetitive DNA sequences and thus constitute a major source of potential GCRs if DNA repair or replication are misregulated in these sites [239, 240].

1.7.1 Telomeres and the DDR

Telomeres constitute a unique environment in terms of DNA replication and the DDR, and as mentioned in 1.4.5, replication stress can cause the formation of T-UFBs in anaphase. Telomeres are repetitive sequences found at the ends of chromosomes with the sequence TTAGGG in vertebrates with a 3' G-rich overhang and are bound by a large protein complex called Shelterin [241-244].

Shelterin is important for suppressing particular DDR features at telomeres, including ATM/ATR-mediated checkpoint activation and preventing ligation of DNA ends at telomeres, thus avoiding end-to-end fusions of entire chromosomes [241]. The Shelterin complex consists of 6 protein subunits: TRF1 (telomere repeat factor 1), TRF2 (telomere repeat factor 2), POT1 (protection of telomeres protein 1), TIN2 (TRF1-interacting nuclear factor 2), RAP1 (repressor/activator protein 1) and ACD (adrenocortical dysplasia homologue) [241]. These proteins have a range of roles in telomere maintenance, which have much to do with suppressing erroneous DDR activation at these sites. For example, TRF2 inhibits activation of ATM at telomeres, while POT1 binds telomeric ssDNA to prevent its coating by RPA and subsequent ATR activation [241, 245-251].

A key function of TRF2 is the formation of unique DNA structures at telomeres called T-loops. T-loops are formed when the 3' ssDNA overhang invades internal telomeric repeats, forming a protective structure at chromosome ends which limits degradation and masks them from DNA damage sensing proteins [252-254]. Despite the beneficial nature of T-loops, they pose a problem for replication machinery during S-phase and thus must be transiently dissolved in order for replication to proceed through telomeric repeats. This T-loop dissolution activity is performed by the helicase RTEL1 (regulator of telomere elongation helicase 1) [255]. In the absence of RTEL1, the

nuclease activity of SLX1-SLX4 is employed to excise persistent T-loops. However, this cleavage causes notable telomere shortening [256]. RTEL1 also seems to have a role in unwinding G-quadruplexes at telomeres, which in the absence of RTEL1 will also hinder telomere replication. Notably, RTEL1's activity on G4s also appears to be important genome-wide for efficient replication through these structures [256, 257].

Telomere biology highlights the concept that specialised or unusual DNA structures may form at certain genomic locations and that the regulation of such structures with respect to DNA replication is unique to those locations. We may therefore also expect to see specific features of DDR regulation that are distinct at these locations.

As might be expected from the DDR-suppressive nature of the Shelterin complex, DNA lesions at telomeres appear to be repaired less efficiently than elsewhere in the genome [258, 259]. However, precise details of telomeric DNA repair are elusive and contradictory, and it may be that hitherto unrecognised effects of proliferative state or genotoxic dose currently cloud the picture [260-262]. However, multiple reports demonstrate a suppression of NHEJ at telomeres and it has been suggested that in proliferative cells, HR is used to repair DSBs at telomeres, while in senescent cells experiencing persistent DDR activation, DNA repair is suppressed [248, 262-265].

1.7.2 CFSs and the DDR

Common fragile sites (CFSs) are genomic loci that display a relatively high incidence of chromosomal aberrations such as deletions, translocations and duplications, and genomic instability arising from CFSs has been implicated in tumourigenesis [266-270].

Three major models, which are not mutually exclusive, exist to explain the basis of genomic instability at CFSs. First, the repetitive AT-rich sequences found in CFSs are flexible and thought to form structures which are problematic for the progression of

the replicative DNA polymerases. As such, CFSs can be considered sites prone to localised replication stress in the absence of additional insults [270-274]. Second, collisions between transcription and replication machineries are thought to give rise to stabilised RNA/DNA hybrid structures and prevent fork progression. This is exacerbated in the context of the third model, which describes the relative paucity of DNA replication origins in CFSs. As a result, stalled forks cannot be rescued by the firing of nearby dormant origins, as could happen elsewhere in the genome [271, 272]. Under-replicated DNA arising from the various challenges to replication at CFSs can cause defects in chromosome segregation during anaphase and subsequent mitotic DNA breakage. This leads to the establishment of a DNA damage response in the daughter G1 cells, in which nuclear bodies of 53BP1 form around the damaged CFSs [275]. With a variety of processes and factors influencing DNA replication and genome stability at CFSs, it is likely that a multitude of different enzymatic activities are required for faithful replication and repair at CFSs, including helicases, nucleases, translocases and TLS polymerases.

CFSs are replicated late in S-phase, and mild replication stress can induce chromosomal breaks or gaps at CFSs, which are visible on metaphase chromosome spreads; this is referred to as CFS 'expression' [276]. CFS expression is thought to be generated early in mitosis by the activity of MUS81-EME1 in the SMX trinuclease complex. Interestingly, the resulting DNA gaps are filled in by mitotic DNA synthesis (MiDAS), seemingly dependent on DNA polymerase δ subunit 3 (POLD3) and perhaps mediated via RAD52-dependent break-induced replication [107, 277, 278]. In the event of replication fork reversal upon severe or prolonged fork stalling at CFSs, RAD51 coats and protects the DNA exposed by the reversed forks. It has recently been shown that

the ATP-dependent helicase RECQ5L acts to remove RAD51 from these stabilised forks in order to render them accessible to MUS81-EME1 activity [279-281].

There is also evidence that proteins of the Fanconi anaemia (FA) pathway contribute to maintenance of faithful DNA replication at CFSs, with absence of FA proteins causing an increased frequency of chromosome breaks at these sites [282]. The FA pathway is best known for its role in the repair of DNA interstrand crosslinks (ICLs), with patients deficient in FA proteins exhibiting developmental problems, bone marrow failure and cancer predisposition [283, 284]. In addition to their role in ICL repair and DSB repair (as is the case for BRCA2, which is also known as FANCD1), there is now much evidence that a number of FA proteins have additional roles. The FA pathway is activated by fork stalling under conditions of replication stress, and FANCD2 has been shown to stabilise stalled forks [176, 280, 285, 286]. FANCD2 assists in the resolution of replication intermediates arising from under-replication at CFSs in G2/M, in concert with the activity of the BLM helicase. More recent findings suggest that FANCD2 also influences the resolution of DNA/RNA hybrids at CFSs to enable continued replication [86, 287]. Additionally, in a manner independent of other members of the FA pathway, the ATP-dependent DNA translocase FANCM was recently demonstrated to remodel replication forks at CFSs which bear problematic secondary structures, rendering them amenable to replication [288].

Understanding the basis of replication fork dynamics at CFSs and how replication is achieved faithfully across these sites could have a profound impact on our understanding of genome stability as a whole. Like telomeres, CFSs constitute genomic loci which during replication appear to exhibit features specific to them. It remains possible that insights gained into DNA replication and repair at telomeres and CFSs might pave the way for new discoveries relevant to other genomic structures, such as

centromeres. Altered regulation of DNA replication and repair at centromeres is relatively poorly understood but is beginning to attract more attention [238, 289, 290].

1.7.3 Centromeres and the DDR

The centromere is the constriction point of a chromosome, acting as the point from which kinetochores are established in order for chromosome segregation to take place during anaphase [291-293]. Centromeric DNA is highly repetitive, and although the precise sequence is poorly conserved, its repetitive nature is conserved from fungi to human [294-296]. The modular monomeric unit of centromeric DNA repeats is the α-satellite, which in humans is 171bp long and can be found in arrays which form higher-order repeats (HORs) up to 5Mb in length [295, 297, 298]. Rather than being defined as a distinct genomic locus by DNA sequence, centromeres are defined epigenetically by the centromere-specific presence in nucleosomes of the histone 3 variant CENP-A. CENP-A is the basis of the constitutive centromere-associated network (CCAN) of proteins and is critical for both the establishment of kinetochores in mitosis and faithful replication of centromeric DNA [299-303].

It is well established that repetitive DNA sequences can pose localised challenges to normal DNA replication through the formation of structures that may hinder normal progress of the replication fork; centromeres are no exception [51, 290, 297, 298, 304-307]. The challenging nature of replication across the secondary structures formed at centromeric chromatin makes it likely that specialised factors and mechanisms act to ensure faithful DNA replication across these regions, although relatively little is known about the precise details of centromere replication [290, 297, 298, 308].

Recently, Aze and colleagues reconstituted centromere replication from *Xenopus laevis* egg extracts and gained several significant new insights into the biology of

centromere replication. They found that not only are some factors specifically enriched at replicating centromeres, including the DDR factors PARP1, XRCC1, XRCC5, MSH2 and MSH6, but other factors commonly associated with events at the replication fork were underrepresented, including RPA proteins and TOPBP1 [290]. Importantly, this points to altered regulation of DNA replication-related processes at centromeres. Indeed, even after treatment with the replication stress-inducing drug aphidicolin, activation of the ATR-CHK1 signalling pathway was suppressed [290, 309]. Instead, the DNA mismatch repair proteins MSH2 and MSH6 were upregulated at centromeric chromatin under aphidicolin treatment, and depletion of MSH6 from the *Xenopus* egg extracts impaired DNA replication at centromeres [290]. However, the precise function of MSH2-6 during DNA replication at centromeres is still unknown.

Another DNA repair factor with a recently recognised role at centromeres is DNA2, known for its end resection activity during homologous recombination and its role in nucleolytic processing of DNA flaps, ssDNA and secondary structures [308, 310-316]. In addition to its nuclease activity, DNA2 exhibits helicase activity, although the physiological role of this helicase activity has remained unknown [317-319]. A recent study found that DNA2 localises to centromeres and seems to support efficient replication at these sites. Interestingly, this function was dependent on both the helicase activity and the nuclease activity, leading the authors to propose that the helicase activity of DNA2 serves to unwind challenging structures to then enable nucleolytic processing [308].

Overall, it has become clear that the dense and repetitive nature of centromeric chromatin poses a challenge for both normal and perturbed DNA replication. Given the importance of centromere integrity during mitosis, errors during centromeric replication could have a dramatic impact on genome stability. While it is accepted that

there must be specialised regulation of replication- and repair-related processes, much is yet to be understood.

1.8 ERCC6L2: a novel SNF2 ATPase in the DDR

1.8.1 SNF2 ATPases: a diverse family of molecular motors

SNF2 (sucrose non-fermentable 2) ATPases are a large family of ATPases characterised by the presence of a highly conserved catalytic core, which couples ATP hydrolysis to DNA translocase activity. Outside the catalytic core, SNF2 ATPases possess a diverse range of functional domains, which provide layers of regulation to individual ATPases. These additional domains allow the direction SNF2 ATPases to specific genomic locations or support activation only in certain cellular contexts [95, 320-323]. These proteins participate in a broad range of DNA-related processes, including transcription, DNA replication, DNA repair, the oxidative stress response and chromosome segregation. Examples of SNF2 ATPases with important roles in genome stability and DNA replication include SHPRH, HLTF, ALC1, ZRANB3, CSB (also known as ERCC6) and PICH (also known as ERCC6L) [7, 77, 84, 128, 197, 225, 324]. Domain structures of the SNF2 ATPases ERCC6, ERCC6L, ERCC6L2 and ZRANB3 are shown in Figure 1-6.

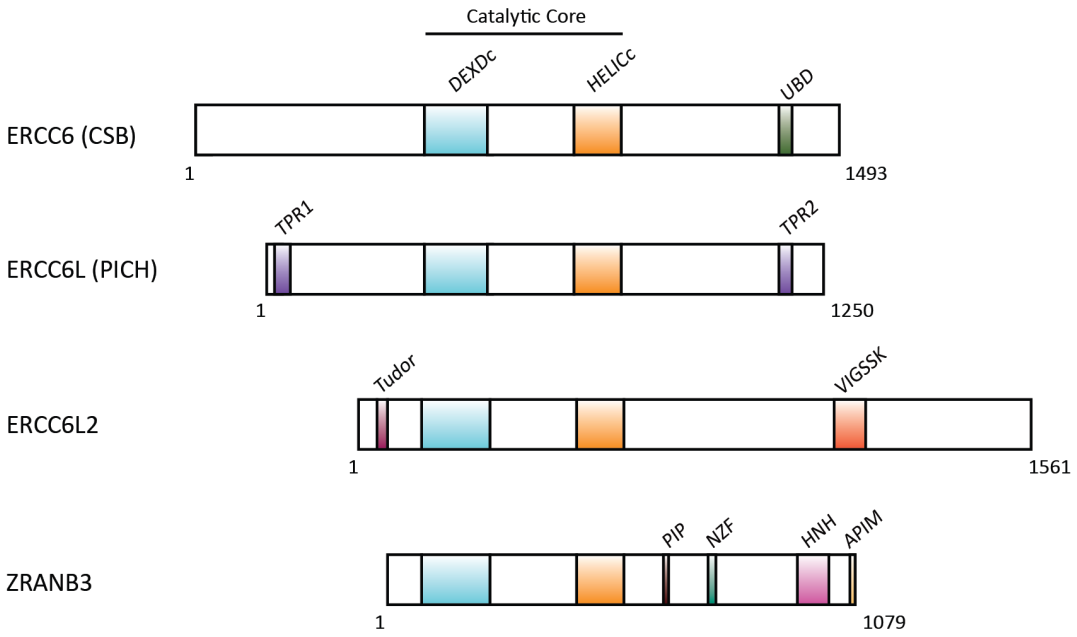


Figure 1-6. Domain structures of selected SNF2 ATPases.

The DEXDc and HELICc domains make up the characteristic SNF2 ATPase core. To the C-terminus of ERCC6 a ubiquitin binding domain (UBD) is found. ERCC6L contains two tetratricopeptide repeats (TPR1/2), one at its N-terminus and one at its C-terminus. ERCC6L2 contains an N-terminal Tudor domain and a VIGSSK domain in its C-terminal region. The VIGSSK domain is an uncharacterised domain identified by the NCBI Conserved Domain Database and Pfam. ZRANB3 contains two PCNA interaction sites (the PIP-box and APIM motif), a ubiquitin-binding NZF domain and an HNH nuclease domain.

As discussed previously in Section 1.6, the interface between DNA replication and the DNA damage response represents a complex situation, and dysregulation of the events occurring in this context can have dramatic consequences. Notably, a multitude of SNF2 ATPases seem to function in the context of DNA replication under genotoxic stress. Importantly, a range of enzymatic activities has also been ascribed to these proteins in addition to ATP hydrolysis. For example, ZRANB3 exhibits nuclease activity [208, 224-226], SHPRH and HTLF are E3 ubiquitin ligases [212, 213, 215, 325], and SMARCAL1 performs an annealling helicase activity [326-328]. The variety of biochemical functions ascribed to SNF2 ATPases coupled to an ATP hydrolysis-driven translocase activity makes the study of these proteins a potentially rich source of revealing insights into genome stability.

Given the range of processes in which SNF2 ATPases are involved, it is unsurprising that mutations and expression abnormalities in a number of these proteins are associated with several disease states, including bone marrow failure, Cockayne syndrome, alpha-thalassemia syndrome and cancer [329-332]. However, despite strong clinical implications, a number of SNF2 ATPases remain poorly understood. As such, studying the cellular and biochemical characterisation of such SNF2 ATPases remains an attractive means of understanding the malfunctions responsible for these diseases.

1.8.2 ERCC6L2

ERCC6L2 (excision repair cross-complementation group 6-like 2) is a poorly understood SNF2 ATPase and *ERCC6L2* has been identified as the gene causative mutation in 22 patients (to date) suffering bone marrow failure (BMF), inherited in an autosomal recessive manner [332-337]. BMF syndromes are relatively diverse but are characterised by a deficiency in mature blood cell production from their precursor cells. These disorders can occur in one or two of the three haematopoetic lineages (erythroid, myeloid and karyocytic) or result in trilineage failure (pancytopenia) [338, 339]. BMF syndromes are associated with several cellular phenotypes, including DNA repair defects (as in the case of Fanconi anaemia), defective telomere maintenance (dyskeratosis congenita) and impaired ribosome biogenesis (Shwachan-Diamond syndrome) [338-342].

Although Fanconi anaemia – a relatively well-known DNA repair disorder – causes BMF and as such is worthy of comparison to ERCC6L2 deficiency disorders, *ERCC6L2*-mutated patient cells do not appear to display the elevated chromosomal breakage patterns common in FA, suggesting that ERCC6L2 performs a function distinct from members of the FA pathway [332-337, 343]. BMF patients including those with FA are

often considered predisposed to the development of acute myeloid leukaemia (AML) [344]. Thus far, two ERCC6L2-deficient patients have died from AML and another has died from Epstein-Barr virus-associated lymphoma [333, 336]. One of the patients who developed fatal AML – which progressed from myelodysplastic syndrome (MDS), a disorder which dramatically sensitises patients to the development of poor-prognosis AML – has two sisters, both of whom are compound heterozygotes for *ERCC6L2* mutations both have MDS [336]. Another patient reported in the same study is homozygous for the K985fs mutation and has both trisomy 20 and monosomy 7, although one of this patient's siblings who is also homozygous for this frameshift is currently asymptomatic at age 3y [336].

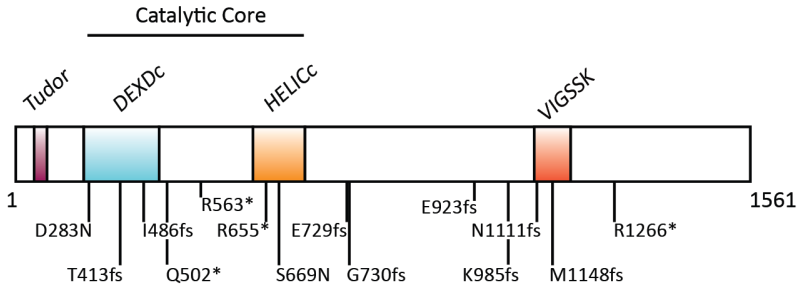


Figure 1-7: ERCC6L2 domain structure showing BMF-causing mutations

At its N-terminus, ERCC6L2 contains a Tudor domain. The DEXDc and HELICc domains make up the SNF2 ATPase catalytic core and between them bear the only non-frameshifting disease-causing point mutations. Towards the C-terminus is the VIGSSK domain, identified by the NCBI Conserved Domain Database and Pfam. One disease-causing frameshifting mutation (R1266*) occurs C-terminal to all annotated domains. Modified from [336].

The catalytic core of ERCC6L2, characteristic of SNF2 ATPases, bears similarity to those of ERCC6 (CSB) and ERCC6-like (PICH), although the accessory domains flanking this conserved region are entirely different (Figure 1-7), suggesting that ERCC6L2 may perform a distinct function from either of these two proteins. At its N-terminus, ERCC6L2 contains a Tudor domain. Tudor domains are typically responsible for

binding proteins bearing specific methylation patterns on lysine or arginine residues [345]. C-terminal to the Tudor domain lies the SNF2 catalytic core, composed of DEXDc and HELICc domains. In the C-terminal half of the protein there are no characterised protein domains, although an uncharacterised domain termed 'VIGSSK' by the NCBI conserved domain database and 'HEBO' by Zhang et al after their proposed protein name of Hebo (helicase mutated in bone marrow failure) is found towards the C-terminus, along with a number of patches of sequence conservation [337, 346]. From the identities of specific disease-causing mutations identified in *ERCC6L2* (shown in Figure 1-7), it appears that the C-terminal half of the protein may be important in its function. While the only disease-causing point mutations identified to date lie within or near the SNF2 ATPase core and as such are likely to disrupt the protein's catalytic activity, most mutants are frame-shifting and give rise to truncated protein products. It should be noted, however, that most patients are compound heterozygote for these mutations. Each individual disease-causing mutation identified to date is listed in Table 1-1. Several of these truncated protein products still include an intact SNF2 core, indicating that the C-terminus imparts critical functionality to the protein. Zhang et al named ERCC6L2 'Hebo' (helicase mutated in bone marrow failure) and named the VIGSSK domain HEBO, which is not obviously present in other proteins, finding this domain conserved from human to *Neurospora crassa* [337]. However, since *bona fide* helicase activity of ERCC6L2 has not been demonstrated *in vitro* and its closest catalytic relatives are DNA translocases rather than helicases, I shall continue to refer to the protein as ERCC6L2 and the HEBO domain as VIGSSK.

Table 1-1: Individual *ERCC6L2* mutations identified in bone marrow failure patients

Mutation (DNA)	Effect (aa)	Outcome (predicted)	Patients	Refs
G847A	D283N	Point mutant, ATPase dead	1	[333]
1236-1239delAACA	T413Cfs	Truncated, ATPase dead	1	[332]

1457del	I486fs	Truncated, ATPase dead	2	[334]
C1504T	Q502*	Truncated, ATPase dead	2	[333]
C1687T	R563*	Truncated, ATPase dead	1	[335]
C1963T	R655*	Truncated, ATPase dead	4	[332, 333, 337]
G2006A	S669N	Point mutant, unclear outcome	1	[336]
2187delG	E729fs	Truncated	2	[333]
2189delG	G730Dfs	Truncated	1	[336]
2767delG	E923Rfs	Truncated	1	[336]
2952_2956delAAAAG	K985Hfs	Truncated	1	[336]
3333-3336delTCAA	N1111Kfs	Truncated	1	[336]
3442_3443delAT	M1148Efs	Truncated	1	[336]
C3796T	R1266*	Truncated	3	[333, 336]

ERCC6L2 was first identified in 2014, when Tummala and colleagues reported its mutation as a genetic cause for BMF, reporting severe truncating mutations. The authors demonstrated mild sensitivity of ERCC6L2-depleted A549 cells to the DNA and protein alkylating agent irofulven and to a lesser extent the interstrand DNA crosslinker mitomycin C (MMC) [332]. Furthermore, A549 cells transfected with siRNA against ERCC6L2 showed increased fluorescence intensity of the DNA damage marker γH2AX following irofulven treatment, leading the authors to propose an unknown role for ERCC6L2 in the DDR [332].

In 2016, Zhang et al reported another BMF case caused by a mutation yielding the same protein product as one of the mutations reported by Tummala et al in 2014 [332, 337]. Zhang et al showed that overexpressed GFP-ERCC6L2 mobilises to sites of DNA damage induced by laser microirradiation. This study went on to further strengthen the case for ERCC6L2 as a DDR factor, demonstrating mild MMC sensitivity in patient-derived cell lines and strong sensitivity to the SSB/DSB-inducing drug phleomycin, as well as ionising radiation (IR). Notably, the authors reported no sensitivity of these cell

lines to PARP inhibition, UV light or methyl methanesulphonate (MMS) [337]. Taken together, these results might indicate a role for ERCC6L2 in the DSB response. However, unlike fibroblasts derived from a patient lacking the c-NHEJ protein XLF, ERCC6L2-mutant fibroblasts exhibited no deficiency in c-NHEJ. Additionally, U2OS cells depleted of ERCC6L2 using RNAi were proficient in homologous recombination, unlike cells depleted of RAD51 [337]. Given no apparent defect in either of the primary DSB repair pathways, it seems unlikely that ERCC6L2 is directly involved in a canonical DSB response.

A more recent publication from Tummala et al, published soon after the work described in this study, demonstrated an interaction between ERCC6L2 and the critical PI3K-family kinase DNA-PK, as well as proposing an ERCC6L2 interactome based on immunoprecipitation and mass spectrometry experiments. This interactome, although as-yet unvalidated, contains a number of potentially interesting candidate ERCC6L2 interaction partners involved in processes including RNA binding and transport, DNA repair, mitosis and mitochondrial function. Interestingly, a number of replisome components were also among the list of candidate ERCC6L2 interactors, including members of the MCM helicase complex (MCM2, MCM3, MCM6, MCM7) and cohesin components (SMC1A, SMC3). Furthermore, the authors identified a number of further *ERCC6L2* mutations and tested sensitivity of three of these patients' cell lines to various drugs. All three cell lines exhibited mild MMC and irofulven sensitivity, as well as sensitivity to the transcription elongation inhibitor DRB. The three patient-derived cell lines also exhibited increased presence of nuclear 53BP1 foci following extended irofulven treatment (24h + 24h release). ERCC6L2-deficient cell lines were also shown to exhibit reduced revcovery of RNA synthesis after irofulven treatment, as well as reduced phosphorylation of the RNA polymerase II CTD, suggesting a transcription

defect in these cells. These observations, along with an elevated presence of nuclear RNA/DNA hybrids in patient cells, led the authors to propose a role for ERCC6L2 in transcription-coupled DNA damage repair and preventing the accumulation of genotoxic R-loops [336]. While such a claim is still largely speculative and more evidence at both a cellular and biochemical level would be necessary to draw such a conclusion, some interesting insights have been gained by this group into an angle of possible ERCC6L2 function which may inform future work.

Tummala et al suggest that ERCC6L2 could be participate in TC-NER and that the loss of ERCC6L2 at sites of stalled RNA polymerase II leads to the accumulation of toxic R-loops and a persistent level of DNA damage, possibly explaining the genotoxin sensitivities described by both Zhang et al and Tummala et al [332, 336, 337]. This is in contrast to the interpretations of Zhang et al, whose suggest a role for ERCC6L2 in some non-canonical or context-specific role in the response to DSBs [337].

It is important to note, however, that a mechanistic function for ERCC6L2 based upon its translocase activity may allow it to participate in more than one repair pathway, meaning these models are not mutually exclusive. A better understanding of ERCC6L2's role in the DDR at a cellular and biochemical level may explain the observations about ERCC6L2 function that have been made to date.

My project's purpose was to shed new light on ERCC6L2's role in the DDR, with two core aims. Firstly, I aimed to establish whether or not ERCC6L2 plays a role in the DDR and identify candidate genome stability functions for ERCC6L2. Secondly, I aimed to explore ERCC6L2 at a protein level, identifying domains important for its function in order to better understand how the protein might fit into its cellular role.

2 Materials and Methods

For a full list of oligonucleotides, antibodies, peptides and plasmids used, refer to Section 7.

2.1 Mammalian cell culture

Running stocks of U2OS (ATCC HTB-96) and HEK 293T (ATCC CRL-11268) cells were maintained in T75 flasks (Corning) in standard culture medium (DMEM (Sigma-Aldrich) + 10% FBS (Sigma-Aldrich) and Penicillin/Streptomycin (Pen/Strep) (Life Technologies)) at 100U/ml and 100µg/ml respectively, and stored in incubators kept at 37°C with 5% CO_2. Upon reaching confluence or near-confluence, cells were passaged by aspirating growth medium, washing once with PBS (Sigma-Aldrich) and detaching in TrypLE Express (Life Technologies) at 37°C. Once cells had detached, they were resuspended in standard culture medium, passaged as necessary and seeded as appropriate in standard culture medium. For short-term storage, cells were detached in Tryple-E and resuspended in standard culture medium as when passaging, then centrifuged in a bench-top centrifuge (Thermo Scientific) at 800 x *g* for 5 minutes (min) and resuspended in an appropriate volume of 90% FBS/10% Dimethyl sulfoxide (DMSO) (Insight Biotechnology) and aliquoted into cryogenic vials (Starlabs). Vials were then placed on ice inside a styrofoam box which was placed in a -80°C freezer. Vials were transferred to liquid nitrogen for long-term storage.

2.2 Cloning and site-directed mutagenesis

50µl PCR (polymerase chain reaction) reactions were prepared as follows: 33µl ddH_2O, 10µl 5X HF buffer (NEB), 1µl $MgCl_2$ (NEB), 1µl dNTP (deoxyribonucleotide

triphosphate) mix (1:1:1:1 mix of individual dNTPs from Bioline), 1µl forward primer at 10µM, 1µl reverse primer at 10µM, 2µl genomic deoxyribonucleic acid (gDNA) template, 1µl Phusion High-Fidelity DNA polymerase (NEB).

PCR for cloning purposes was performed in a thermal cycler (Life Technologies) as follows: 1x 95°C 3min, 30x (95°C 30sec, 53°C 30sec/500bp to be amplified, 72°C 2min), 1x 72°C 10min, and hold at 4°C. Following PCR, 37µl of PCR product was mixed with 10µl DNA loading dye (Promega) and loaded onto a 1% agarose (Sigma-Aldrich) gel in 1X TBE (45mM Tris-borate (Tris base and boric acid, both from Sigma-Aldrich)), 1mM EDTA (Applichem)) with 0.005% EtBr (Sigma-Aldrich). Electrophoresis was performed for 25min at 150V, after which EtBr staining in the gel was visualised using a UV transilluminator (UVP). DNA bands visible under UV light were excised and DNA was extracted using a Gel Extraction Kit (Qiagen), as per the manufacturer's instructions.

For Gateway cloning, primers were designed against the sequence of interest with an *aatB* site on the 5' and 3' ends of the forward and reverse primers, respectively. PCR was performed, and the PCR product was resolved on a 1% agarose gel, excised and purified as outlined above. A BP reaction was then performed to recombine the purified linear *attB* PCR product with the *attP*-containing pDONR-221 vector (Invitrogen). This was carried out by preparing a reaction containing the Gateway BP Clonase II Enzyme mix (Invitrogen) as per the manufacturer's instructions. The BP reaction was incubated at RT for 1h and the entire mix was then transformed into competent DH5α *E.coli* cells (Invitrogen). To transform *E.coli*, the BP reaction was mixed gently into 45µl competent cells and incubated on ice for 15min. The transformation mix was then heat shocked at 42°C for 45sec, then placed back on ice for 45sec. 250µl S.O.C medium (Invitrogen) was then added to the mix, which was placed at 37°C with agitation for 1h. Following recovery, *E.coli* were centrifuged at RT for 1min at 1000 x *g*. 150µl supernatant was

then removed, then bacteria were resuspended in the remaining liquid and spread on LB-agar plates containing 30μg/ml kanamycin (Fisher Scientific). Plates were placed at 37°C overnight to allow colonies to grow. 24h later, colonies were picked and used to inoculate cultures of 3ml LB + 30μg/ml kanamycin, which were then incubated overnight at 37°C with agitation. The next morning, plasmid DNA was purified using a QIAPrep Spin Miniprep Kit (Qiagen). Samples were then submitted to Source Bioscience for Sanger sequencing using M13F and M13R primers. Sequencing results were analysed to verify successful cloning of the desired insert into pDONR-221. 1.5μl of the validated construct was then mixed gently with 1.5μl pN-YFP-DEST vector and 1.5μl LR Clonase II (Invitrogen), and incubated at RT for 1h. Transformation and plating of DH5α cells was performed as described above, although selection was performed using 100μg/ml ampicillin (Life Technologies). 3ml cultures and subsequent minipreps were performed as described above, and resultant plasmid DNA was sequenced by Source Bioscience using JYB-F and JYB-R sequencing primers. For a full list of plasmids and primers used, see Table 7-1 and Table 7-2. LR reactions were also used to clone pLX304 constructs; the sequencing primers used for these constructs were JYB-F and WPRE-R.

For restriction cloning into pEXPR-IBA105 (IBA Life Sciences), pEXPR-IBA105 was first digested with EcoRI (NEB) and XhoI (NEB) by preparing a reaction containing 10μl H_2O, 1μg pEXPR-IBA105, 2μl Cutsmart buffer (NEB) and 1μl of each restriction enzyme. This reaction was incubated at 37°C for 1h and then run on a 1% agarose gel as described previously, with 1μg of undigested vector also loaded for comparison. The band containing digested vector was then excised and gel purification was performed as described previously. Primers were designed for amplification of the desired region of ERCC6L2 with an EcoRI site added to the forward primer and an XhoI site added to

the reverse primer and PCR was performed as described above, using pDONR-221-ERCC6L2 as a template, which was obtained from Dr. Dragana Ahel. Following amplification by PCR, the product was resolved on a 1% agarose gel, excised and purified as described previously. The purified PCR product was then subjected to restriction digestion with EcoRI and XhoI by mixing the following reaction and placing it at 37°C for 1h: 20µl purified PCR product, 2.5µl Cutsmart buffer and 1µl of each restriction enzyme. Digested PCR product was purified using a QIAquick PCR Purification kit (Qiagen) as per the manufacturer's instructions. A ligation reaction was then prepared with 50ng digested vector, 150ng digested insert, 5µl Quick Ligase buffer (NEB), 1.5µl Quick Ligase (NEB) and left for 1h at RT. Following ligation, 45µl XL-Gold competent *E.coli* (Agilent) was transformed with 3.5µl of the ligation reaction as described previously and transformed bacteria were spread on LB-agar plates containing 100µg/ml ampicillin. The LB-agar plate was then placed at 37°C overnight. The next evening, colonies were picked and grown overnight in 3ml LB + 100µg/ml ampicillin with agitation. Minipreps were then performed as described previously and purified plasmid DNA was sent for Sanger sequencing by Source Bioscience, using Strepseq F and Strepseq R as sequencing primers.

Site-directed mutagenesis was performed, as per the manufacturer's instructions, using a QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent). Mutagenesis primers were designed using the Agilent QuikChange Primer Design web tool. This program is freely accessible at www.genomics.agilent.com/primerDesignProgram.jsp. PCR conditions were performed as follows: 1x 95°C 2min, 18x (95°C 20sec, 60°C 10sec, 68°C 30sec/500bp of plasmid to be amplified), 1x 68°C 5min, and hold at 4°C. After PCR, amplification was verified by running a small amount of the PCR product on a 1% agarose gel containing EtBr, as outlined above. The remainder of the PCR product was

then subjected to 1h treatment with 1.5µl DPN1, as per the manufacturer's instructions and 3.5µl of this final volume was then transformed into XL-Gold competent *E.coli*, also supplied with the QuikChange Lightning Site-Directed Mutagenesis kit.

2.3 Live cell imaging and laser microirradiation

For laser microirradiation experiments, U2OS cells were seeded in glass-bottomed 24-well plates (Greiner Bio-one). Upon reaching approximately 80% confluence, cells were transfected with YFP constructs as appropriate. For transfection, 1µg plasmid DNA and 6µl Lipofectamine-2000 (Invitrogen) were added to 50µl serum-free DMEM, tapped gently to mix and incubated at RT for 15min. 200µl serum-free DMEM was then added dropwise to the transfection mix, which was tapped gently. 60µl of this final mix was then added to each well to a total of up to four wells. During incubation, 100µl growth media was removed from each well to be transfected. Following transfection, cells were incubated at 37°C and 5% CO2 for 24-36h before laser microirradiation. 16-24h prior to microirradiation, BrdU (Sigma-Aldrich) was added to wells to a final concentration of 10µM. Laser microirradiation was performed with a 405nm laser using an Olympus FV1200 scanning confocal microscope, with a cellVivo chamber (Olympus) attached to maintain the cells and apparatus at 37°C and 5% CO_2. Imaging was performed using a 60X or 100X objective lens and images were collected continually for up to three minutes. For quantification of recruitment to laser-induced DNA damage, FIJI software was used to measure fluorescence intensity at both the laser track and another non-damaged area of the nucleus throughout the time course. Relative fluorescence intensity at damage sites was then calculated at each time point as a percentage of the signal present in the non-damaged area. For quantification, data

from at least 10 cells were collected per experiment. At least three independent experiments were performed; data shown are from one representative experiment.

2.4 Immunofluorescence

U2OS cells on glass coverslips (Scientific Laboratory Supplies) were transfected using Lipofectamine-2000 as described previously with appropriate YFP constructs, as appropriate. 24h post-transfection, cells were washed with PBS and incubated in pre-extraction buffer (10mM PIPES (Sigma-Aldirch), 300mM sucrose (VWR), 3mM $MgCl_2$ (Sigma-Aldrich), 20mM NaCl (Honeywell), 0.5% Triton-X (Sigma-Aldrich)) at -20°C for 5 minutes. In experiments displayed in Figure 3-3, Figure 4-6 and Figure 4-7, pre-extraction prior to fixation was not performed. Next, cells were washed with PBS and fixed with 4% formaldehyde (Sigma-Aldrich) for 10 minutes. Cells were washed with PBS and permeabilised in PBS + 0.5% Triton-X for 5 minutes, then blocked in 2% BSA (Sigma-Aldrich) for 30 minutes. Primary antibody incubations were performed for 1h followed by 3 PBS washes. Secondary antibodies were added, incubated for 1h and washed with PBS. Coverslips were mounted onto slides using ProLong gold anti-fade reagent containing DAPI (Sigma-Aldrich) and analysed by confocal microscopy.

For visualisation of nascent DNA synthesis, a Click-iT EdU Alexa Fluor 594 imaging kit (Life Technologies) was used and I followed the manufacturer's protocol. Briefly, U2OS cells growing in a glass-bottomed 24-well plate were transfected with YFP-ERCC6L2 as described previously. 36h later, EdU was added to the cells to a final concentration of 10μM and cells were incubated at 37°C with 5% CO2 for 15min. Next, pre-extraction and fixation were performed as described above. Following fixation, cells were washed and permeabilised as described above. Next, cells were blocked in 2% BSA and then incubated in Click-iT reaction cocktail with Alexa Fluor 594, prepared as per the

manufacturer's instructions, for 30min at RT and protected from light. Following the Click-iT reaction, cells were washed once with 2% BSA and incubated with anti-CENP-A antibody (Abcam) diluted 1:500 in 2% BSA at RT for 1h, protected from light. After primary antibody incubation, cells were washed three times in PBS and then incubated with Alexa Fluor 405 conjugated anti-mouse antibody (Life Technologies) for 1h at RT, protected from light. Cells were then washed three times in PBS and kept in 1ml of PBS for confocal microscopy using a 60X objective lens on an Olympus FV1200.

For a full list of antibodies used in immunofluorescence experiments, see Table 7-3.

2.5 Establishing knockout cell lines by CRISPR/Cas9

Single guide RNAs (sgRNAs) targeting ERCC6L2 were designed and cloned into px459 by Dr. Andrew Bassett, formerly of Genome Engineering Oxford (GEO). Details of sgRNA sequences used can be found in Section 7.

U2OS cells were grown to roughly 70% confluence in 10cm culture dishes (Corning). Cells were then transfected with px459 empty vector (Addgene #62988 was gift from Feng Zhang), px459-ERCC6L2sg1, px459-ERCC6L2sg2 or left untransfected. At this scale, transfections were performed as follows. First – for each transfectant – one mix of 500µl serum-free DMEM + 16µl Lipofectamine-2000 and one mix of 500µl serum-free DMEM + 4µg px459 plasmid were prepared. After incubation for 15min at RT, these mixes were combined, mixed gently by tapping and incubated for a further 15min at RT. Meanwhile, 1ml growth medium was removed from each 10cm dish to be transfected. After the incubation, transfection mixes were added drop wise to their appropriate dishes, which were agitated gently to ensure even distribution of transfectant. Dishes were returned to the incubator and kept at 37°C and 5% CO_2 for 24h.

24h after transfection, were washed and detached using TrypLE Express as described in Section 2.1, then resuspended thoroughly in 10ml normal growth medium supplemented with 3μg/ml puromycin (Invivogen) and returned to the incubator. 24h after selection, it was verified that all untransfected cells had died and that some transfected cells in each dish survived puromycin selection. Then, puromycin-containing growth medium was aspirated, cells were washed once in 10ml normal growth medium, and then given a media change into another 10ml normal growth medium. Cells were then returned to the incubator and allowed to recover from selection for 48h.

In order to establish clonal cell lines, cells were seeded at low density in 96-well plates (Corning). To achieve this, cells were first detached in TrypLE Express as described in Section 2.1 48h after being taken out of selection. Detached cells were resuspended in 10ml normal growth medium. 1ml of this suspension was diluted 10X in normal growth medium, while the remainder was centrifuged, frozen down and stored at -80°C as described in Section 2.1. Next, cell counting was performed on the 10X diluted cell suspension using a Moxi Type S cell counter (VWR). Using the resultant cell counts, 20ml cell suspensions in normal growth medium were then prepared at concentrations of 5, 10, 30 and 60 cells/ml. From each cell suspension, 2 96-well plates were then seeded with 100μl cell suspension per well, giving 2 96-well plates seeded at 0.5, 1, 3 and 6 cells/well for each puromycin-selected cell line. 96-well plates were returned to the incubator and 24h later each well was topped up with another 100μl normal growth medium.

Four days after seeding, 96-well plates were checked under a bright field microscope to identify wells containing single colonies of cells; such wells were marked as such, while wells containing more than one colony were also labelled. Colonies were left to

grow for 3 weeks, during which time each well containing a clonal colony was checked on a weekly basis and growth medium was replenished as necessary.

Upon reaching confluence, colonies were detached as described in Section 2.1 and moved from wells of 96-well plates to wells of 24-well plates (Corning) in ample growth medium. Clonal cell lines were monitored and growth medium was replenished when required. Upon reaching confluence, clonal cell lines in 24-well format were detached as described in Section 2.1 and seeded in 5ml of normal growth medium in T25 flasks (Corning).

Upon reaching confluence, cell lines growing in T25s were detached as described in Section 2.1, and 90% of the cell suspension was frozen down and stored at -80°C as described in Section 2.1. The remaining 10% of cell suspension was re-seeded in a T25 flask and grown until confluent. When confluent, cells were washed once in 1X PBS and detached in 1X PBS + 10mM EDTA at 37°C for 10min. Cells were then centrifuged for 5min at 800 x *g* in a bench-top centrifuge, the supernatant was discarded and cell pellets were stored at -80°C prior to genome DNA extraction and sequencing, as described below.

2.6 Validating ERCC6L2$^{-/-}$ U2OS cell lines

Cell pellets collected as described in Section 2.1 from candidate ERCC6L2$^{-/-}$ U2OS cell lines were thawed on ice and then resuspended in 120µl squishing buffer (10mM Tris-HCl pH 8 (Life Technologies), 1mM EDTA, 25mM NaCl, 200µg/ml Proteinase K (Invitrogen)). Samples were incubated at 65°C for 30min, and then placed at 95°C for 2min to inactivate Proteinase K and stored at 4°C. 2µl of the resultant gDNA extract was used as a template for PCR, using Gateway-compatible primers specific to the genomic

sequence of ERCC6L2 to amplify a region of 400-500bp spanning the sgRNA target site. Primer names and sequences are shown in Table 7-2.

PCR and Gateway cloning was performed as described in Section 2.2; DNA products were cloned into pDONR-221 and transformed into competent DH5α cells as described in Section 2.2. 10 colonies from each candidate ERCC6L2$^{-/-}$ cell line were picked and overnight cultures were inoculated in 3ml LB + 30μg/ml kanamycin. Plasmid DNA was then purified using a QIAPrep Spin Miniprep Kit (Qiagen). Samples were then submitted to Source Bioscience for Sanger sequencing using the M13F primer.

Sequencing results were aligned to the WT genomic DNA sequence of ERCC6L2 and examined for frame-shifting insertion or deletion mutations (INDELs). Cell lines whose genomic DNA sequencing results demonstrated frame-shifting INDELs in all sequencing reads were considered successful knockout cell lines. Cell lines with more than two different genomic sequences at the target site were not considered as useful cell lines and were omitted from further experiments.

2.7 Establishing stable overexpression cell lines

Desired overexpression constructs were generated using Gateway cloning as described in Section Cloning and site-directed mutagenesis. First, pDONR-221 constructs were cloned without a stop codon, followed by an LR Clonase II reaction to move the insert into pLX304 (Addgene #25890 was a gift from David Root). For the empty vector (EV) control, the nuclear localisation sequence (NLS) of the SV40 large T antigen was cloned into pLX304. Maxipreps of pLX304 constructs were prepared using a HiSpeed Plasmid Maxi Kit (Qiagen) as per the manufacturer's instructions.

To produce lentiviruses containing pLX304 constructs, HEK 293Ts were grown to roughly 70% confluence in 10cm dishes in standard culture medium and transfected

with the desired pLX304 construct along with pCMV-VSVG (Addgene #8454 was a gift from Bob Weinberg) and pCMV-dr8.2-dpr (Addgene #8455 was a gift from Bob Weinberg). 2 10cm dishes of HEK 293Ts were seeded per pLX304 construct. Per 10cm dish, two 500µl mixes were made consisting of serum-free DMEM plus either 20µl Lipofectamine-2000 or 1µg pCMV-VSVG, 5µg pCMV-dr8.2dpr and 5µg pLX304 construct. After 15min incubation at RT, these mixes were combined drop-wise, mixed gently and left for another 15min incubation at RT, during which time 1ml of normal growth medium was removed from each 10cm dish of HEK 293Ts. Following the incubation, the transfection mix was added drop-wise to the HEK 293Ts, after which the dishes were agitated gently and returned to the incubator. 12-15h later, the growth medium was removed from the HEK 293Ts and replaced with 8ml fresh growth medium. Target U2OS cells (2 ERCC6L2$^{-/-}$ and 2 WT control cell lines) were then seeded in 10cm dishes in normal growth medium to be roughly 70% confluent the following day; one extra 10cm dish per cell line was seeded as a control for successful selection. 24h later, lentivirus-containing media were removed from HEK 293Ts and filtered through a 0.45µm filter (Generon) into a 50ml tube. Protamine sulphate (Sigma-Aldrich) was added to lentivirus-containing media at a final concentration of 8µg/ml. The growth medium was then removed from each 10cm dish of U2OS cells, and the appropriate lentivirus and protamine sulphate containing growth medium was added to the cells. U2OS cells were then returned to the incubator and dishes containing 293Ts were sterilised and discarded.

48h after lentiviral transduction, the growth medium was removed from each dish of U2OS cells, and cells were detached as described in Section Mammalian cell culture. Cells were then resuspended in normal growth medium + 10µg/ml blasticidin

(Invivogen) and returned to the incubator. Blasticidin-containing medium was replaced every 48h until all untransduced cells had died under selection – approximately 6 days.

Once untransduced cells had died, transduced cell lines were removed from selection and allowed to recover in normal growth medium until confluent. At this stage, cells were detached from 10cm dishes as described in Section 2.1 and re-seeded in T175 flasks (Corning). Upon reaching confluence, cells were detached again as described in Section 2.1. 20% of the resultant cell suspension was frozen down in FBS + 10% DMSO as described in Section 2.1 and stored at -80°C. The remaining 80% of cells were centrifuged for 5min at 800 x *g*, after which the supernatant was discarded and cells were washed in 10ml PBS. Cells were then centrifuged again for 5min at 800 x *g*, the supernatant was discarded and the cell pellets were frozen at -80°C ahead of lysis for Western blotting.

To lyse cells for Western blotting, pellets were allowed to thaw on ice, before being resuspended thoroughly in 250µl cell lysis buffer (50mM Tris pH 8, 150mM NaCl, 1mM DTT (Fisher Scientific), 1% Triton-X, supplemented with 1 protease inhibitor tablet per 30ml total volume (Roche) and 500U/µl Benzonase (Sigma-Aldrich)). Samples were then sonicated thoroughly (4 rounds for 30sec each on full power using a bench-top sonicator (MSE Soniprep 150)) and placed on a rotor wheel at 4°C for 45min. Lysates were then centrifuged in a bench-top microcentrifuge (Eppendorf) at 15,000 x *g* for 20min at 4°C. The supernatant was then moved to a fresh 1.5ml microcentrifuge tube (Starlabs) and 70µl NuPAGE LDS sample buffer (Invitrogen) + 200mM DTT. Samples were then boiled at 95°C for 3min and stored at -20°C before Western blot analysis using antibodies against β-tubulin and the V5 tag, as described in Section 2.8.

2.8 Western blotting

For Details of antibodies used, see Section 7.

A NuPAGE Novex mini gel tank (Invitrogen) was assembled with a 10- or 15-well 1.0mm 4-12% Bis-Tris Novex gradient gel (Invitrogen) as appropriate, according to the manufacturer's instructions. 500ml 1X NuPAGE MOPS running buffer (20X from Invitrogen) in ddH_2O was prepared, which was used to fill the central chamber of the gel tank, after which the rest of the buffer was added to the outer chamber. Cell lysates or IP samples in NuPAGE LDS sample buffer + 200mM DTT were centrifuged briefly in a bench-top microcentrifuge prior to loading. Wells of 10-well gels were loaded with 36µl of each sample, while wells of 15-well gels were loaded with 20µl of each sample. 10µl of SeeBlue Plus2 Protein Standard (Life Technologies) was also added to one well as a molecular weight marker. SDS-PAGE was performed at 200V for 40min using a Bio-Rad power pack.

Following electrophoresis, protein was transferred onto a 9cm x 7cm PVDF membrane (Sigma-Aldrich) pre-soaked in 100% methanol (Sigma-Aldrich), using a wet transfer system (Bio-Rad) as per the manufacturer's instructions. Transfer was performed for 1h at 100V in 1X NuPAGE transfer buffer (20X from Invitrogen) with 20% methanol, with a cold pack placed inside the transfer tank. Next, the membrane was blocked for 30min in 30ml 5% skimmed milk (Marvel) in PBS-T (PBS + 0.1% Tween-20 (Sigma-Aldrich)), in a plastic tray with agitation. Meanwhile, primary antibodies were diluted in 10ml 2% skimmed milk in PBS-T, allowing 10ml per membrane.

After blocking, 5% milk in PBS-T was decanted and stored for short-term reuse at 4°C, and 10ml primary antibody in PBS-T was added to each membrane, as appropriate. Primary antibody incubation was performed for 1h at RT or overnight at 4°C with

agitation. Meanwhile, HRP-conjugated secondary antibodies (Dako) were prepared as required in 10ml 2% skimmed milk in PBS-T per membrane. For a full list of primary and secondary antibodies used for Western blotting, see Table 7-3.

Following primary antibody incubation, primary antibody mixes were decanted back into tubes and kept at -20°C for short-term reuse. Three 5min washes with 20ml PBS-T were then performed. After the final wash, 10ml secondary antibody solution in 2% skimmed milk in PBS-T was then added to each membrane as appropriate and incubated for 1h at RT with agitation. Three 5min washes with 20ml PBS-T were then performed as before.

For detection, a plastic film was first secured using adhesive tape in an Amersham Hypercassette (Fisher Scientific). Next, a 1:1 mix of enhanced chemiluminescence (ECL) reagents A and B (Life Technologies) was prepared, allowing 4ml per membrane. The membrane was removed from PBS-T and touched onto tissue paper to remove excess liquid, then placed protein side up on the plastic film. The ECL reagent mix was then added to the membrane, distributed evenly and left for 1min. Excess ECL reagent was then blotted off using paper towels and a second piece of plastic film was placed on top of the membrane. Excess ECL reagent and air bubbles were squeezed out from the plastic film and blotted away using paper towels. The second piece of plastic film was then secured using adhesive tape. In the dark, photographic film (Sigma-Aldrich) was place on top of the membrane-containing plastic folder and the cassette was closed and left for 30 seconds. The film was then developed using a developer (Xograph), and exposures for other time points were performed and developed as necessary.

2.9 Clonogenic survival assays

U2OS cells were grown to roughly 70% confluence, then detached using TrypLE Express and seeded at a density of 1000 cells per 6cm dish (Corning) or 700 cells per well of a 6-well plate (Cornign). For clonogenic survival assays using phleomycin (Sigma-Aldrich), etoposide (Abcam) or mitomycin C (MMC, Sigma-Aldrich), cells were seeded in a total volume of 5ml growth medium containing the appropriate concentration of drug in a 6cm dish. For dishes with MMC treatments, MMC-containing media was aspirated 24h after seeding, cells were washed once with 5ml of normal growth medium and which was replaced with another 5ml growth medium. Cells in 6cm dishes were fixed and stained for quantification 12 days after seeding. For assays using etoposide, cells were seeded in 6-well plates and then treated with the appropriate concentration the following day. After 11 days, cells grown in 6-well plates were fixed and stained. Each experiment was performed in technical triplicates, and data shown represent combined data from at least three independent experiments.

For fixing and staining cells prior to quantification, growth medium was decanted, cells were washed gently in 1X PBS, and incubated at room temperature in (5mg/ml) crystal violet (Sigma-Aldrich) in 25% methanol for 30min for simultaneous fixation and colony staining. For quantification, pictures were taken of dishes using a Nikon D3200. A stand for the camera and homemade frame were used to ensure that dishes or plates were fixed in place and that within each experiment, dishes/plates were always in the same place in an image.

Using FIJI software, a macro was recorded to ensure that thresholding and quantification methods used between images was the same. JPEGs were converted to 8-bit, then used to generate a binary image in which the threshold settings maximised colony visibility without background signal appearing inside dishes/wells. FIJI's

'Analyze Particles' function was then used to quantify colonies judged to consist of at least 50 cells within as large a circle as possible, made to fit within the visible walls of the wells/dishes. With our setup, we defined a minimum colony size of $65px^2$. Particle analysis yielded quantifications of both total well area covered by cell growth and the number of individual colonies meeting our size criteria. For subsequent data analysis, values obtained from technical triplicates were first averaged. Next, clonogenic potential was calculated as a percentage of growth compared to untreated cells within each cell line. This quantification was performed on both colony number and total growth area, to ensure that trends between cell lines were similar and control for under-counting resulting from merged colonies. Data displayed represent colony numbers. In order to perform statistical testing on clonogenic survival assay quantifications, GraphPad Prism 7 software was used to calculate the area under the curve (AUC) for each biological replicate based on the average percentage survival for each technical replicate within one experiment. Average AUC values for cell line treated with etoposide or phleomycin were then displayed as bar charts with error bars denoting SEM. One-way ANOVA was performed on the AUC values and p values were represented as follows: $p \leq 0.05$ = *, $p \leq 0.01$ = **, $p \leq 0.001$ = ***, $p \leq 0.0001$ = ****.

2.10 Cell cycle analysis by flow cytometry

U2OS cells were grown to around 70% confluence in 10cm dishes and treated with varying doses of etoposide for 48h. Using a Click-iT Plus Alexa Fluor 647 Flow Cytometry A kit (Life Technologies), cells were pulse-labelled with EdU for 15min or 3h and pr°Cessed for flow cytometry by labelling of incorporated EdU with Alexa Fluor 647 and labelling DNA with propidium iodide (BD Biosciences) as per the manufacturer's instructions.

Flow cytometry was performed on a Cytek DxP, using the BluFL2 channel (488nm laser, 585/45 filter) to detect propidium iodide and the RedFL1 channel (640nm laser, 661/16 filter) to detect Alexa Fluor 647. Gating was applied to exclude dead cells and doublets from the analysis, and 50,000 cells within the gating parameters were analysed per condition. For the experiments displayed in Figure 4-1, Figure 4-2 and Figure 5-6, raw data were collected by Dr. Dragana Ahel. Data were analysed using FlowJo software (FlowJo).

2.11 Immunoprecipitation

HEK 293T cells were grown in DMEM + 10% FBS + Pen/Strep to roughly 80% confluence in 5 15cm dishes (Corning) per condition. Cells were then transfected with the appropriate construct in pEXPR-IBA105; per 15cm dish, 10µg DNA and 200µl Lipofectamine-2000 was used. For each construct, mixes were made of 10ml serum-free DMEM + 50µg DNA and 10ml serum-free DMEM + 200µl Lipofectamine-2000. After incubation at RT for 15min, these mixes were combined and mixed gently, then incubated at RT for another 15min. Meanwhile, 4ml growth media was removed from each 15cm dish to be transfected. After the 15min incubation, 4ml of the DNA/Lipofectamine-2000 mix in serum-free DMEM was added to each 15cm dish. 36h later, cells were washed once in PBS and then detached in 7ml PBS per 15cm dish. Cell suspensions from replicate 15cm dishes were combined and centrifuged at 800 x *g* for 5min in a bench-top centrifuge. Following centrifugation, supernatant was aspirated and cell pellets were resuspended in 30ml PBS as a wash step. This centrifugation and wash step was repeated once more, followed by a final centrifugation at 800 x *g* for 5min. The supernatant was removed and the cell pellet was frozen at -80°C for

immunoprecipitation (IP) the following day, or placed on ice and lysed immediately if IP was being performed on the same day.

To lyse cells, they were resuspended well in 3ml ice-cold lysis buffer (50mM Tris pH 8, 150mM NaCl, 1mM DTT, 1% Triton-X, supplemented with 1 protease inhibitor tablet per 30ml total volume (Roche) and 500U/µl Benzonase (Sigma-Aldrich)). Samples were then sonicated thoroughly (4 rounds for 30sec each on full power) and placed on ice between rounds of sonication. Following sonication, samples were placed on a rotor wheel at 4°C for 45min to ensure digestion of genomic DNA. Meanwhile, an Econo-Pac empty chromatography column (Bio-Rad) was equilibrated in ice-cold IP buffer (50mM Tris pH 8, 150mM NaCl, 1mM DTT, 1% Triton-X), allowing the buffer to run through. 600µl StrepTactin sepharose beads 50% (IBA Life Sciences) per sample were then added to the column, added to which was 10 volumes of IP buffer, which was allowed to drip through the column. Once the IP buffer had dripped through the column, a stopper was placed on the bottom. Beads were then resuspended in 500µl IP buffer per sample and distributed evenly between Poly-Prep chromatography columns (Bio-Rad) (one column per sample); this should give a bed volume of approximately 300µl beads in each column. Excess buffer was allowed to drip through the columns, before stoppers were placed on the bottom. After the 45min incubation step at 4°C, samples were distributed between 2ml microcentrifuge tubes (Starlabs) and centrifuged in a bench-top microcentrifuge at 13,000 x *g* for 20min.

Following centrifugation, the supernatants from each sample were combined in ice and 250µl was removed and stored in a 1.5ml microcentrifuge tube (Starlabs) as an input control. 70µl of each input sample was mixed with 30µl NuPAGE LDS sample buffer + 200mM DTT, boiled at 95°C for 3min and stored at -20°C. Remaining inputs were snap-frozen on dry ice and stored at -80°C. The remaining supernatants were

then added to their respective StrepTactin Sepharose-containing chromatography columns and the caps were added. Stoppered and capped columns containing StrepTactin beads and cell lysates were then placed on a rotor wheel at 4°C for 1h.

After incubation, columns were removed from the rotor wheel and, on ice, placed in 15ml tubes (Falcon) with the stoppers and caps removed for collection of the flow-through. 250µl of each flow-through was snap-frozen on dry ice and stored at -80°C in case pulldown efficiency needed to be tested. Next, columns were placed in a rack over a large beaker at 4°C for waste collection, and washed thoroughly by gravity flow; for each column, 3 full column volume (roughly 14ml) washes followed by 7 5ml washes were performed with ice-cold IP buffer. During the washes, 1 empty Mini Bio-Spin column (Bio-Rad) per sample was prepared per sample for elution by equilibration with 500µl ice-cold IP buffer; buffer was allowed to drain by gravity flow.

After the washes, stoppers were placed back on the bottoms of each Poly-Prep column and beads were resuspended in 500µl IP buffer, and then moved into the appropriate Mini Bio-Spin column; excess buffer was allowed to drain through the Mini Bio-Spin column. This step was repeated once to maximise recovery of beads. Once excess IP buffer had drained through, stoppers were placed back on the bottoms of the Mini Bio-Spin columns. 250µl ice-cold 1X Elution Buffer diluted in IP buffer was added to each column (10X Elution Buffer: 250mM biotin (Sigma-Aldrich) in Tris pH 8; final pH 8 is necessary for biotin solubility and was achieved by the addition of NaOH). Caps were then added to the columns and elution was performed on a rotor wheel at 4°C for 1h. Following elution, caps and stoppers were removed from columns, which were then placed in 1.5ml microcentrifuge tubes for eluate collection by gravity flow. 70µl of each eluate was mixed with 30µl NuPAGE LDS sample buffer + 200mM DTT and boiled for

3min in preparation for Western blotting and stored at -20°C. The remaining eluates were snap-frozen in dry ice and stored at -80°C.

Following eluate collection, stoppers were replaced on the Mini Bio-Spin columns. Beads were then resuspended in 500µl IP buffer and moved to a clean 1.5ml microcentrifuge tube. This step was repeated to maximise bead recovery. Once beads had settled in the microcentrifuge tubes, most of the excess IP buffer was pipetted out carefully, avoiding disturbance of the beads and leaving a small volume of buffer behind. Care was taken to ensure that the volume of buffer left behind was the same between samples. Next, beads were resuspended thoroughly in 80µl NuPAGE LDS sample buffer + 200mM DTT per 250µl bead bed volume, boiled at 95°C for 5min and stored at -20°C. Analysis of the IP was then performed by Western blotting, as described previously in Section 2.8.

2.12 PCNA purification

pET28a-6xHis-PCNA obtained from Dr. Dragana Ahel was transformed into competent Rosetta *E.coli* on an LB-agar plate with 30µg/ml kanamycin and grown overnight at 37°C. The following day, 50ml LB + 30µg/ml kanamycin was inoculated with a single Rosetta colony and grown overnight at 37°C with agitation. 16h later, the 50ml overnight culture was added to 450ml fresh LB and grown at 30°C until an optical density (OD) at 600nM of 0.7-0.8 was reached. Then, PCNA expression was induced with 0.4mM IPTG (Fisher Scientific) for 3h at 30°C. The 500ml culture was then centrifuged at 4,500 rpm at 4°C in a Beckman Avanti J-26S XP centrifuge using the JLA 10.500 rotor (Beckman Coulter). Following centrifugation, the supernatant was removed, sterilised and discarded, and the bacterial pellet was collected into a 50ml tube (Falcon) and frozen at -20°C for lysis the following day.

The bacterial pellet was thawed in 25ml ice-cold Ni-NTA lysis buffer (50mM Tris pH 8, 500mM NaCl, 10mM imidazole (Sigma-Aldrich), 4mM β-mercaptoethanol (Sigma-Aldrich) with 1 protease inhibitor tablet per 50ml buffer) on a rotor wheel at 4°C. Once the pellet had thawed, the cell suspension was resuspended thoroughly until smooth and homogenous, at which point 3ml Bug Buster (Merck) and 4µl Benzonase (Sigma-Aldrich) were added. The cell suspension was then sonicated for 6 rounds at 30sec each, being placed on ice between for 30sec between each round. The cell suspension was then placed on a rotor wheel at 4°C for 45min, and then centrifuged at 15,000 rpm for 30min at 4°C in a Beckman Avanti J-26S XP centrifuge using a JA-17 rotor.

During centrifugation, 1.2ml Ni-NTA bead suspension (Qiagen) (roughly 0.6ml bed volume) was added to an Econo-Pac empty chromatography column (Bio-Rad) and equilibrated at 4°C by allowing 25ml ice-cold Ni-NTA buffer (50mM Tris pH 8, 500mM NaCl, 10mM imidazole, 4mM β-mercaptoethanol) to run through. After centrifugation and bead equilibration, the supernatant was added to the Ni-NTA resin and allowed to run through by gravity flow at 4°C. PCNA-bound Ni-NTA beads were then washed twice at 4°C with 25ml ice-cold Ni-NTA buffer, which was allowed to drain by gravity flow. Beads were then washed again at 4°C in Ni-NTA + 20mM imidazole, and eluted from in 3ml Ni-NTA buffer + 500mM imidazole into a 15ml tube. 30µl eluate was mixed with 15µl LDS sample buffer + 200mM DTT, boiled at 95°C for 3min and resolved by SDS-PAGE as described in Section 2.8. After running, the SDS-PAGE gel was placed in 20ml Instant Blue (Stratech Scientific) stain for 20min and then destained in ddH_2O for 20min to assess the purification. The remaining eluate was dialysed overnight at 4°C in SnakeSkin 3.5kDa cut-off dialysis tubing (Fisher Scientific), with 50mM Tris pH 8, 100mM NaCl and 1mM DTT.

2.13 Streptavidin pulldown of biotinylated peptides

Custom synthetic biotinylated peptides were purchased from (Eurogentech) and dissolved in 50mM Tris pH 8, 150mM NaCl to a concentration of 2mM and stored at -20°C. See Table 7-4 for full sequence details of peptides used. Working solutions were prepared by diluting peptides 10X to 200nM in ddH_2O. Streptavidin Dynabeads (Life Technologies) were resuspended gently but thoroughly, and 100µl of the resultant suspension was transferred to one 1.5ml microcentrifuge tube for each peptide to be tested, plus one for a no-peptide control, and kept on ice. Using a magnetic rack, beads were separated from solution, the supernatant was removed and beads were resuspended in 500µl ice-cold wash buffer (50mM Tris pH 8, 150mM NaCl, 1mM DTT, 0.2% Triton-X). This was step was repeated once more, before 10µl of each 200nM peptide solution was added to the appropriate tube, with 10µl ddH_2O added to the no-peptide control tube. Samples were then placed on a rotor wheel at 4°C for 1h to allow peptides to bind to the beads. Using the magnetic rack as before, beads were then washed a total of 5 times with 1ml ice-cold wash buffer.

A working solution of PCNA was then prepared, first diluting 50µl of recombinant, His-tagged PCNA purified as described in Section 2.12, in 5ml wash buffer, which was then mixed gently by inverting the tube several times. The wash buffer was then removed from tubes using a magnetic rack as before, and beads were resuspended in 500µl wash buffer + PCNA. Samples were then placed on a rotor wheel at 4°C for 1h. 10x 1ml washes were then performed in wash buffer, using the magnetic rack to separate beads from the suspension as before. After washing, the final 1ml wash buffer was removed and beads were resuspended in 80µl wash buffer. 40µl LDS sample buffer + 200mM DTT was then added to each suspension and mixed thoroughly. Beads were

then boiled at 95°C for 5min. After boiling, SDS-PAGE and Western blotting against PCNA were performed as described in Section 2.8.

2.14 Co-crystallisation of PCNA with peptides and structure resolution

Untagged PCNA in pET28a was expressed in Rosetta *E.coli* cells for 4h at 30°C in 0.4mM IPTG by Dr. Marek Šebesta. Cell lysis was performed in cell lysis buffer (50mM Tris pH 8, 150mM NaCl, 1mM DTT, 2mg/ml lysozyme (Sigma-Aldrich), protease inhibitor cocktail and 25U/ml Benzonase). PCNA purification was then performed by Dr. Šebesta as described in [208].

Untagged PCNA was co-crytallised with a peptide containing the a-PIP of ERCC6L2 by Dr. Šebesta, who also co-crystallised PCNA with a peptide containing the canonical PIP-box of ZRANB3 (PDB code: 5MLO) as shown in [208]. PCNA (7.5mg/ml in 20mM Tris pH 8, 50mM NaCl, 1mM DTT) was co-crystallised at 20°C with ERCC6L2 a-PIP peptide (5mg/ml in 20mM Tris pH 8, 50mM NaCl, 1mM DTT) in 0.1M Bis-Tris 5.5, 25% w/v PEG 3350. Data acquisition on these crystals was performed by Dr. Antonio Ariza, who also solved the final structures to a resolution of 2.73Å.

Chimera software (UCSF) was used to visualise the PCNA-ERCC6L2 (aPIP) structure, which was overlaid with the co-crystal structure of PCNA with the ZRANB3 PIP-box.

3 ERCC6L2 is involved in the DNA damage response

3.1 Results

Previous studies showing that cell lines derived from ERCC6L2-deficient patients have proposed a possible role for ERCC6L2 in the DNA damage response [332, 336, 337]. However, much is yet to be understood about the basis of ERCC6L2's cellular function. In order to determine whether ERCC6L2 is recruited to sites of DNA damage and eventually explore the domain requirements of this recruitment (see Section 4), I employed live cell imaging and laser microirradiation, an established method for real-time imaging of proteins' recruitment to a broad range of DNA lesions [347]. While these studies were on-going, another group published similar observations [337].

N-terminally tagged YFP-ERCC6L2 localised to the nucleus and was recruited efficiently to sites of laser-induced DNA damage in U2OS cells pretreated with BrdU for 16h, in contrast to the YFP-tagged nuclear localisation sequence (NLS) of the SV40 Late T-antigen (Figure 3-1). This suggested to us that ERCC6L2 might function at DNA damage sites.

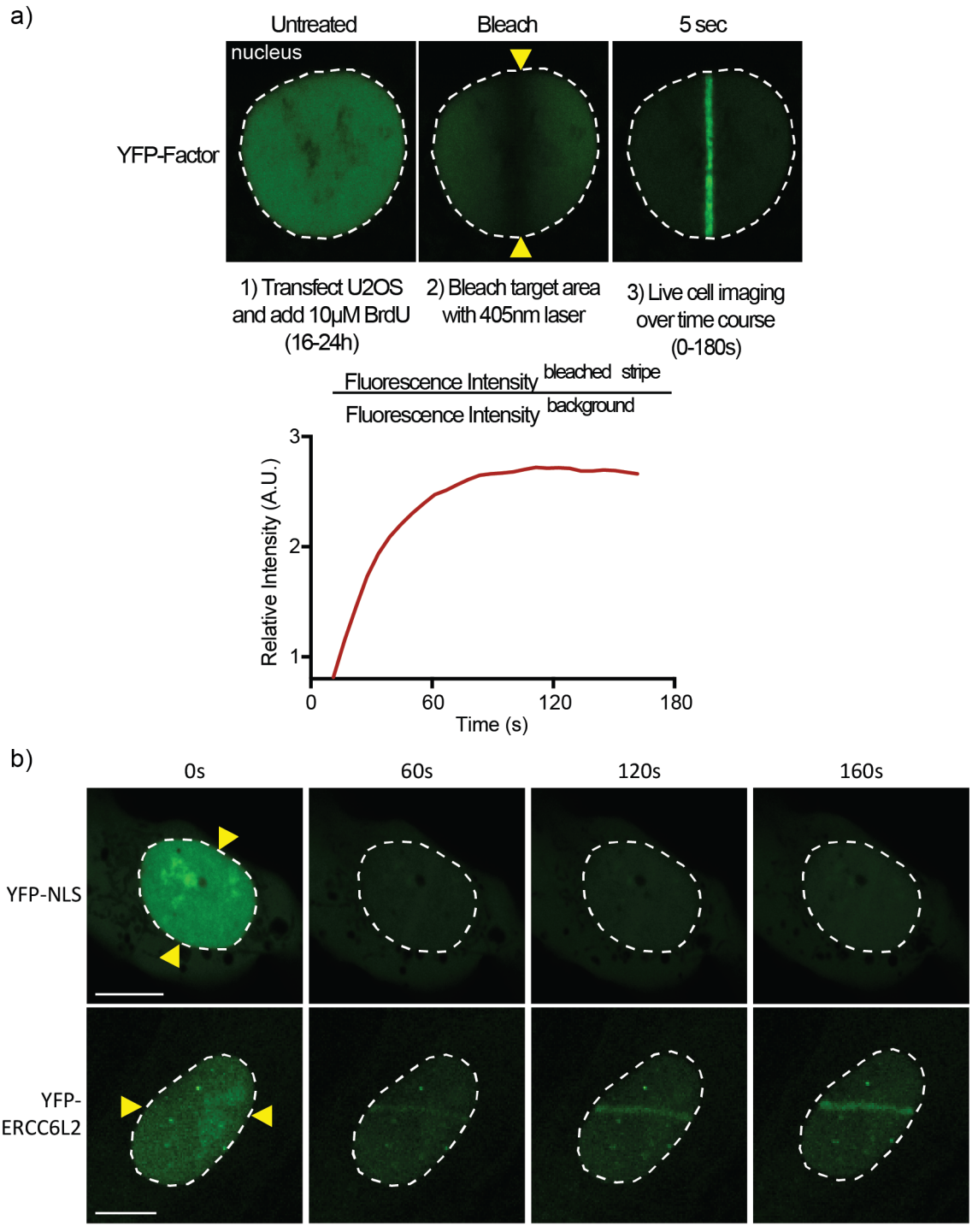


Figure 3-1: ERCC6L2 is recruited to DNA lesions

a) Procedure for analysing recruitment to DNA damage using laser microirradiation. U2OS cells transiently expressing a YFP-tagged protein are treated with BrdU for 16h. A 405nm laser is directed to specific nuclear sites to induce localised DNA damage and the recruitment of the YFP-tagged protein to the bleached region is imaged in real time. Recruitment can be quantified by calculating percentage fluorescence intensity of the targeted region versus unbleached nuclear YFP signal. In this example, YFP-ALC1 is used. b) Laser microirradiation on U2OS cells

expressing YFP-tagged ERCC6L2 or a nuclear localisation signal (NLS). Yellow triangles indicate the microirradiation target region and dotted while lines indicate the periphery of nuclei. Scale bars = 10μm.

Given the multitude of different disease-causing truncation mutants of ERCC6L2 [332-337], I cloned a range of these mutants into the same N-terminal YFP-tagging expression vector and performed laser microirradiation, followed by fixation and immunofluorescence against the DNA damage marker γH2AX. The clinical ERCC6L2 mutants tested in this study are shown in Figure 3-2.

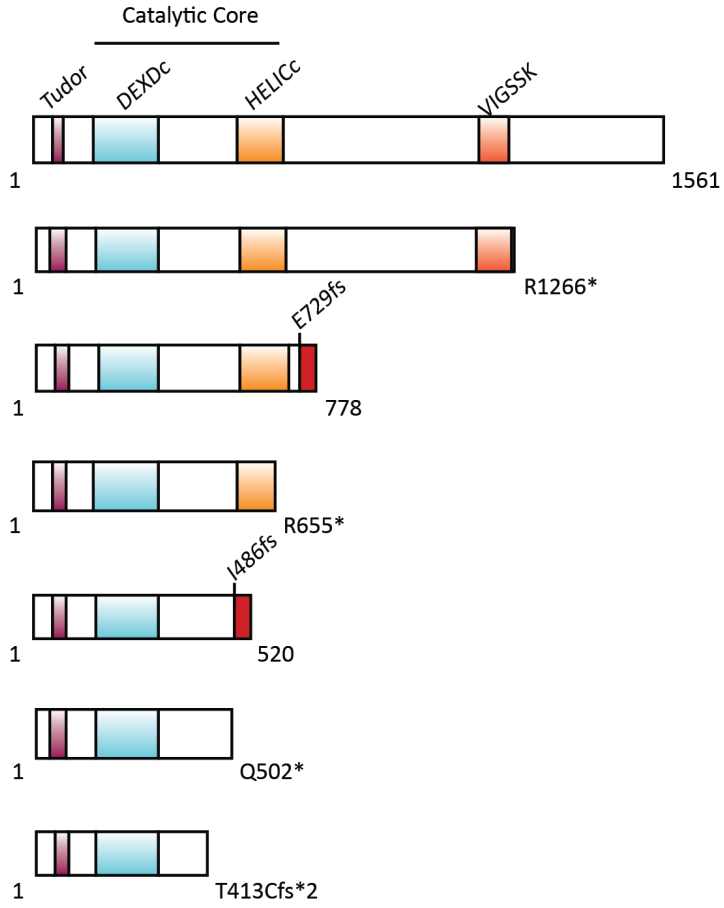


Figure 3-2: Clinical mutants of ERCC6L2 tested in this study

While WT ERCC6L2 mobilised to damage sites and colocalised with γH2AX, all of the mutants tested exhibited dramatically weakened recruitment to DNA lesions and, aside form the R1266* mutant, showed mislocalisation to the cytoplasm (Figure 3-3).

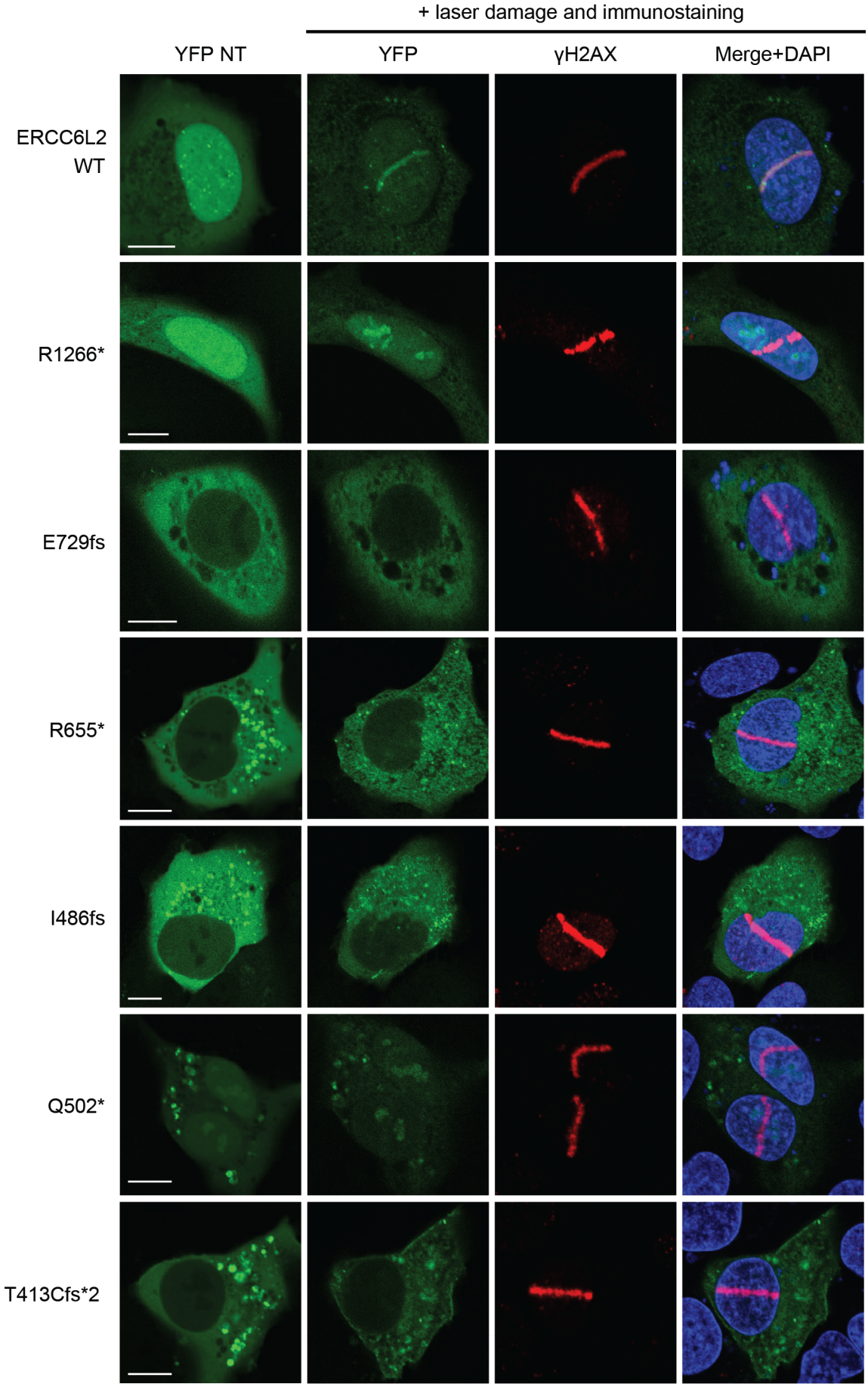

+ laser damage and immunostaining
YFP NT
YFP
γH2AX
Merge+DAPI
ERCC6L2 WT
R1266*
E729fs
R655*
I486fs
Q502*
T413Cfs*2

Figure 3-3: Clinical mutants of ERCC6L2 are deficient in recruitment to DNA damage sites

Laser microirradiation on U2OS cells expressing YFP-tagged ERCC6L2 and a range of clinically relevant mutants. Cells were fixed 30min following laser damage and immunostained for the DNA damage marker γH2AX. Scale bars = 10μm.

Considering ERCC6L2 as a likely novel DDR factor, I next sought to establish knockout (KO) cell lines deficient in ERCC6L2 by using CRISPR (clustered regularly interspersed short palindromic repeats)/Cas9 technology [348, 349]. Single-guide RNAs (sgRNAs) were designed and tested by Dr. Andrew Bassett, formerly of Genome Engineering Oxford. I used two different sgRNAs in the Cas9-expressing px459 vector, one targeting exon 2 and the other targeting exon 3 of ERCC6L2.

Due to the absence of a commercially available antibody against ERCC6L2, I validated these cell lines by cloning a region of genomic DNA spanning the targeted exons into pDONR-221, transforming DH5α *E. coli* cells with the construct and analysing at least 10 resultant colonies by Sanger sequencing following plasmid DNA extraction. This approach allowed us to detect frame-shifting INDELs (insertions or deletions) in the target exons.

I successfully obtained two clonally derived ERCC6L2$^{-/-}$ U2OS cell lines from each sgRNA which should yield severely truncated protein products (Figure 3-4). In parallel, I generated a number of control U2OS cell lines subjected to transfection with the empty px549 vector (without a sgRNA) and the same process for establishing clonal cell lines. The resulting four control cell lines are referred to as Ctrl 1-4, and the resulting knockout cell lines are referred to as ERCC6L2$^{-/-}$-1A, ERCC6L2$^{-/-}$-1B, ERCC6L2$^{-/-}$-2A and ERCC6L2$^{-/-}$-2B. In this nomenclature, the number associated with the cell line indicates the sgRNA used (1 or 2) and the subsequent letter refers to the clonal cell line derived (A or B). The cell lines generated in this study are summarised in Table 3-1.

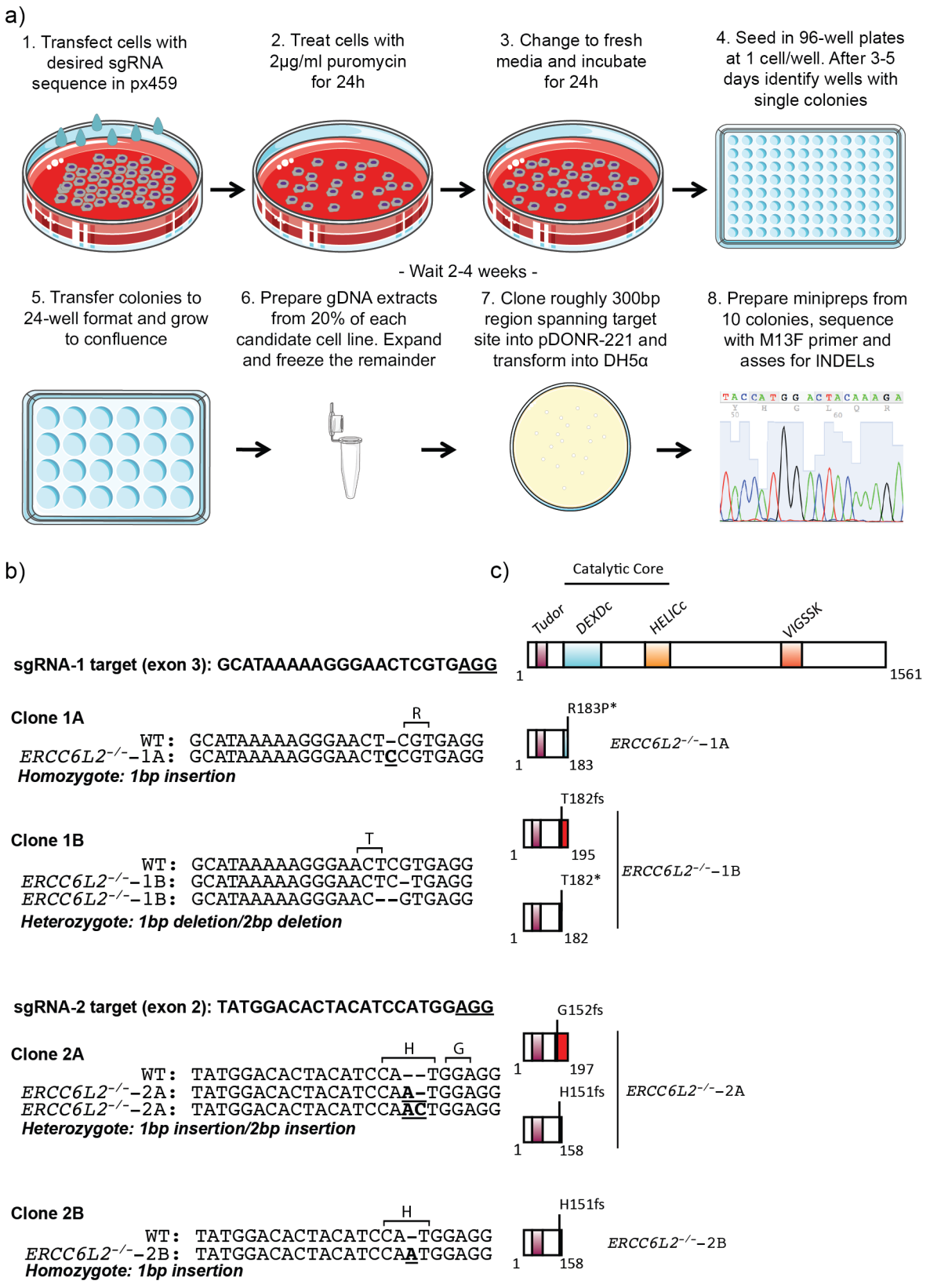


Figure 3-4: Generation of ERCC6L2$^{-/-}$ U2OS cell lines by CRISPR/Cas9

a) Strategy for generation of knockout cell lines over the course of around 6 weeks. Cells were transfected with sgRNA of choice in px459 and 24h later placed under puromycin selection for 24h. Following selection and recovery, cells were grown as single cell colonies in 96-well plates. Genomic DNA (gDNA) extracts were prepared and a region surrounding the targeted exon was cloned into pDONR-221 for validation by Sanger sequencing. sgRNAs were

designed by Dr. Andrew Bassett. It took around 6 weeks to generate isogenic cell lines ready for collection of gDNA samples. b) Sanger sequencing results from established ERCC6L2$^{-/-}$ U2OS cells. c) Predicted protein products resulting from INDELs identified in (b).

To establish whether or not our ERCC6L2$^{-/-}$ cell lines are sensitive to genotoxic insult, I performed clonogenic survival assays with a range of DNA-damaging drugs, initially using a single control cell line and a single knockout cell line. ERCC6L2$^{-/-}$-1A cells showed only mild sensitivity at a high dose of MMC (Figure 3-5). Other genotoxic treatments to which ERCC6L2$^{-/-}$ cells are not sensitive include MMS, UV and cisplatin. The topoisomerase I inhibitor camptothecin was also tested, however there was some variability between WT control cell lines in the tolerance of camptothecin, making it difficult to draw conclusions regarding a possible sensitivity of knockout cells to this drug. However, early results indicated that these cells displayed profound sensitivity to both phleomycin and the topoisomerase II (TOPO2) inhibitor etoposide when compared to control cell lines. Etoposide was tested because previous studies have established sensitivity of ERCC6L2-deficient cells to phleomycin and ionising radiation [336, 337]. These treatments should both cause DSBs among other lesions, while etoposide causes a degree of topological stress in addition to DSBs, which it generates mostly in a DNA replication-coupled manner [350]. I expanded this experiment to include another control cell line (Ctrl 2), as well as the other three ERCC6L2 knockout cell lines generated. In comparison to the two control U2OS cell lines (Ctrl 1 and Ctrl 2), the ERCC6L2$^{-/-}$-1A, -2A and -2B cell lines all displayed strong sensitivity to phleomycin and etoposide. ERCC6L2$^{-/-}$-1B appeared somewhat less sensitive to phleomycin than the other knockout cell lines and slightly less sensitive to etoposide than ERCC6L2$^{-/-}$-1A (Figure 3-5).

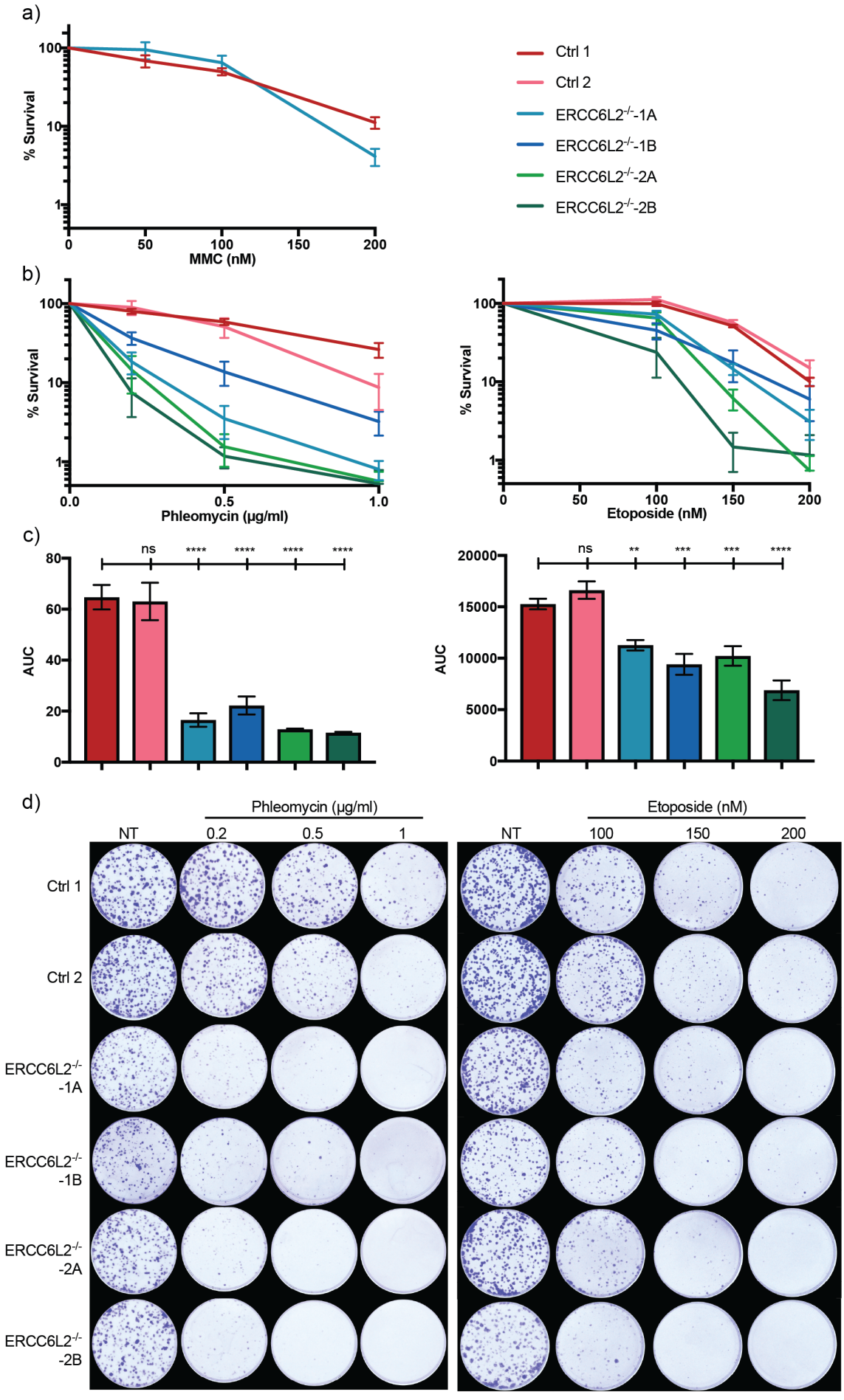

a)
% Survival
MMC (nM)
Ctrl 1
Ctrl 2
ERCC6L2-/--1A
ERCC6L2-/--1B
ERCC6L2-/--2A
ERCC6L2-/--2B
b)
% Survival
Phleomycin (µg/ml)
% Survival
Etoposide (nM)
c)
AUC
ns
AUC
ns
d)
Phleomycin (µg/ml)
NT
0.2
0.5
1
Etoposide (nM)
NT
100
150
200
Ctrl 1
Ctrl 2
ERCC6L2-/--1A
ERCC6L2-/--1B
ERCC6L2-/--2A
ERCC6L2-/--2B

Figure 3-5: ERCC6L2$^{-/-}$ U2OS cells are sensitive to genotoxic insult

a) Clonogenic survival assays on Ctrl 1 and ERCC6L2$^{-/-}$-1A cells treated with mitomycin C. b) Clonogenic survival assays performed on the expanded set of cell lines (Ctrl 1, Ctrl 2, ERCC6L2$^{-/-}$-1A, ERCC6L2$^{-/-}$-1B, ERCC6L2$^{-/-}$-2A and ERCC6L2$^{-/-}$-2B) using phleomycin and etoposide treatments. c) Area under the curve (AUC) calculations of the clonogenic survival assays quantified in b) with one-way ANOVA performed. d) Representative images from the clonogenic survival assays quantified in b). Clonogenic survival assays were fixed and stained after 12 days of growth in 6cm dishes or 10 days growth in 6-well plates. Error bars ± SEM. $p \leq 0.05$ = *, $p \leq 0.01$ = **, $p \leq 0.001$ = ***, $p \leq 0.0001$ = ****.

To eliminate the possibility that the sensitivities observed arose from off-target effects of the CRISPR/Cas9 system, I next sought to complement the KO cell lines with both WT ERCC6L2 and a mutant predicted to eliminate its ATPase activity (K165R), rendering the protein catalytic-dead. Equivalent mutations of this conserved lysine residue in other SNF2 ATPases has been demonstrated to abrogate their catalytic activities, including SNF2 ATPases studied in our group [128, 225]. Using a lentivirus-compatible mammalian expression vector (pLX304), I established stable cell lines overexpressing ERCC6L2 WT or ERCC6L2 K165R with a C-terminal V5 tag in the U2OS Ctrl 1, Ctrl 2, ERCC6L2$^{-/-}$-1A and ERCC6L2$^{-/-}$-2A backgrounds (Figure 3-6). It should be noted that the level of ERCC6L2-V5 expression does vary across the cell lines generated. An overview of the cell lines generated in this work is provided in Table 3-1.

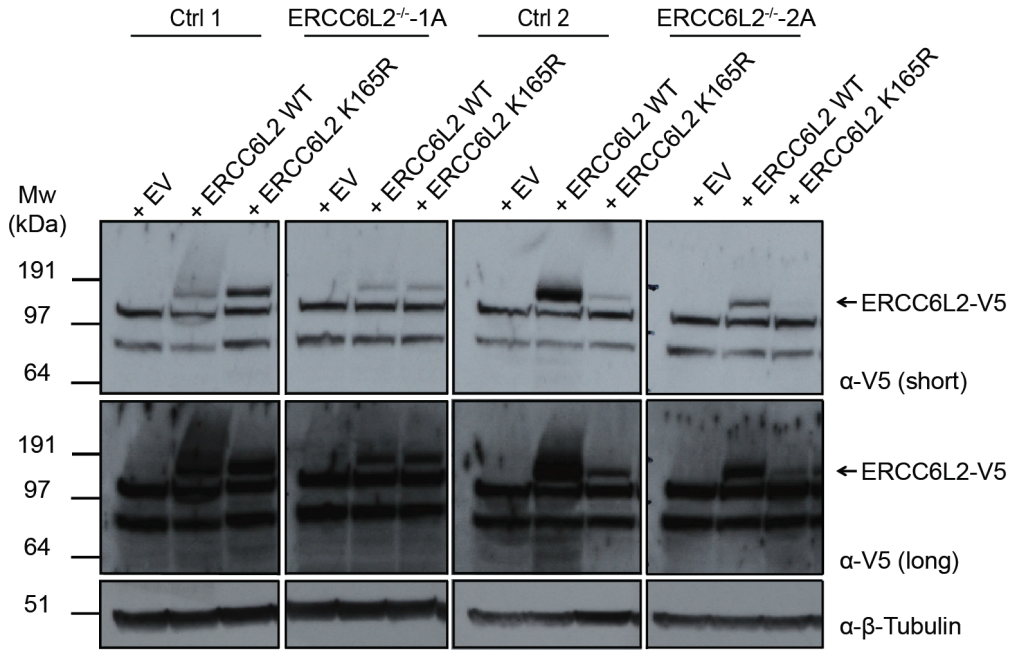


Figure 3-6: Establishing ERCC6L2 complementation cell lines

Western blot of complementation cell lines derived from Ctrl 1, Ctrl 2, ERCC6L2$^{-/-}$-1A and ERCC6L2$^{-/-}$-2A cell lines. β-tubulin was used as a loading control.

Table 3-1: overview of generated cell lines

Clonally derived U2OS cell lines generated in this study (see chapter 3.1 and 5.1) are listed below, listing the sgRNA used at the CRISPR stage and whether and with which ERCC6L2 genotype cell lines were complemented. ERCC6L2 forms overexpressed are WT, K165R (catalytic-dead) and Q798A (PCNA-binding null mutant; see Section 5).

U2OS cell line name	**CRISPR sgRNA**	**Observed mutations (allele 1/2)**	**Overexpressed construct**
Ctrl 1	none	none	EV ERCC6L2 WT ERCC6L2 K165R ERCC6L2 Q798A
Ctrl 2	none	none	EV ERCC6L2 WT ERCC6L2 K165R ERCC6L2 Q798A
Ctrl 3	none	none	Not attempted
Ctrl 4	none	none	Not attempted
ERCC6L2$^{-/-}$-1A	sgRNA-1 (exon 3)	R183P* (homozygote)	EV ERCC6L2 WT ERCC6L2 K165R ERCC6L2 Q798A
ERCC6L2$^{-/-}$-1B	sgRNA-1 (exon 3)	T182fs/T182*	Not attempted
ERCC6L2$^{-/-}$-2A	sgRNA-2 (exon 2)	G152fs/H151fs	EV ERCC6L2 WT

			ERCC6L2 K165R ERCC6L2 Q798A
ERCC6L2$^{-/-}$-2B	sgRNA-2 (exon 2)	H151fs (homozygote)	Not attempted

In clonogenic survival assays, the complementation cell lines (+EV, +ERCC6L2 WT, + ERCC6L2 K165R) derived from Ctrl 1 or Ctrl 2 cells exhibited etoposide and phleomycin sensitivities similar to the parental cell lines not infected with any lentivirus. While ERCC6L2$^{-/-}$-1A cells still displayed sensitivity to phleomycin and etoposide, complementation of these cells with WT ERCC6L2 gave partial or complete rescue of clonogenic potential in response to these drugs, respectively (Figure 3-7). Similarly, WT ERCC6L2 overexpression in the ERCC6L2$^{-/-}$-2A background restored clonogenic potential to a level similar to that of Ctrl 2-derived control cell lines under treatment with both drugs (Figure 3-7). Notably, ERCC6L2 K165R overexpression only slightly increased clonogenic potential of ERCC6L2$^{-/-}$-1A cells under phleomycin treatment, and did not rescue ERCC6L2$^{-/-}$-1A clonogenic potential under etoposide treatment. The observation for the ATPase-dead K165R mutant in ERCC6L2$^{-/-}$-2A cells is similar, with overexpression of ERCC6L2 K165R in this background failing to improve tolerance to etoposide or phleomycin when compared to the dramatic effect of expressing the WT protein (Figure 3-7). It should be noted that the difference in area under the curve (AUC) under phleomycin treatment between ERCC6L2$^{-/-}$ parental cell lines and those overexpressing WT ERCC6L2 is statistically significant. In contrast, this phenotype rescue under etoposide treatment is not statistically significant, perhaps due to the less dramatic sensitivity of ERCC6L2$^{-/-}$ cells to etoposide when compared to WT cells (Figure 3-7).

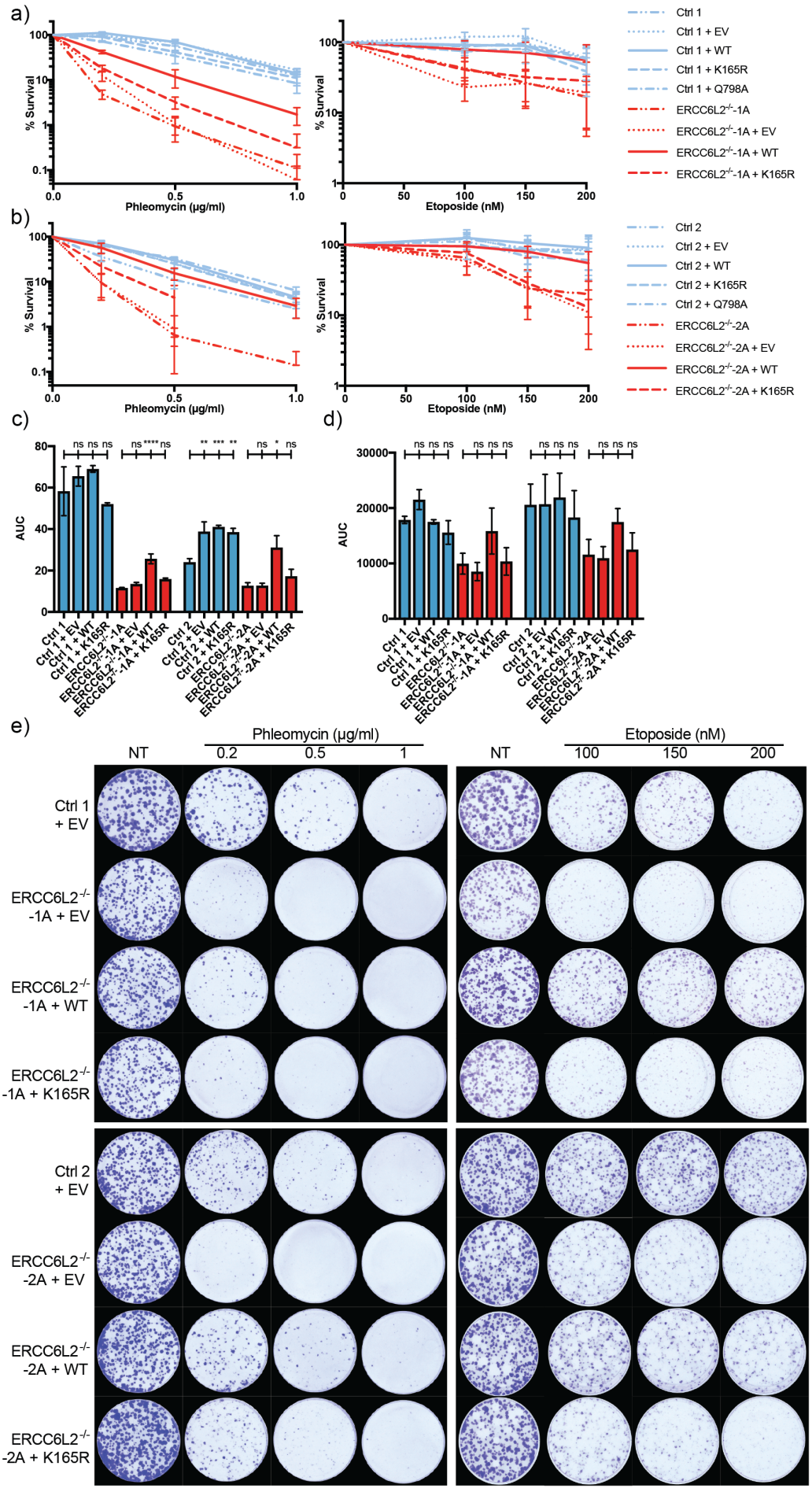

a)
% Survival
Phleomycin (µg/ml)
Etoposide (nM)
Ctrl 1
Ctrl 1 + EV
Ctrl 1 + WT
Ctrl 1 + K165R
Ctrl 1 + Q798A
ERCC6L2-/--1A
ERCC6L2-/--1A + EV
ERCC6L2-/--1A + WT
ERCC6L2-/--1A + K165R
b)
Ctrl 2
Ctrl 2 + EV
Ctrl 2 + WT
Ctrl 2 + K165R
Ctrl 2 + Q798A
ERCC6L2-/--2A
ERCC6L2-/--2A + EV
ERCC6L2-/--2A + WT
ERCC6L2-/--2A + K165R
c)
AUC
d)
e)
Phleomycin (µg/ml)
NT
0.2
0.5
1
Etoposide (nM)
100
150
200
Ctrl 1 + EV
ERCC6L2-/- -1A + EV
ERCC6L2-/- -1A + WT
ERCC6L2-/- -1A + K165R
Ctrl 2 + EV
ERCC6L2-/- -2A + EV
ERCC6L2-/- -2A + WT
ERCC6L2-/- -2A + K165R

Figure 3-7: Overexpression of ERCC6L2 rescues ERCC6L2$^{-/-}$ cells' tolerance of etoposide and phleomycin in a manner dependent on its ATPase activity

a) Clonogenic survival assays of Ctrl 1- and ERCC6L2$^{-/-}$-1A-derived complementation cell lines from Fig treated with phleomycin and etoposide. b) Clonogenic survival assays of Ctrl 2- and ERCC6L2$^{-/-}$-2A-derived complementation cell lines from Fig treated with phleomycin and etoposide. c) Area under the curve (AUC) calculated on phleomycin clonogenic survival assays from a) and b) with one-way ANOVA performed. d) Area under the curve (AUC) calculated on etoposide clonogenic survival assays from a) and b) with one-way ANOVA performed e) Representative images from the assays performed in a) and b). Error bars ± SEM. $p \leq 0.05$ = *, $p \leq 0.01$ = **, $p \leq 0.001$ = ***, $p \leq 0.0001$ = ****.

3.2 Discussion

In this section, I have demonstrated that ERCC6L2 likely plays a role in the DNA damage response. I have shown that ERCC6L2 is recruited to DNA damage sites induced by laser microirradiation. Furthermore, I show that clinical mutants frequently display mislocalisation and a DNA damage recruitment defect. Finally, ERCC6L2 is required for U2OS cells' tolerance of phleomycin and etoposide in a manner dependent on its ATPase activity.

ERCC6L2 is recruited rapidly to sites of DNA damage induced by laser microirradiation, implying the existence of signalling mechanisms that direct ERCC6L2 to DNA lesions (Figure 3-1). This is in agreement with data published while this work was on-going [337]. Importantly, a range of clinically relevant truncations of ERCC6L2 displayed deficiencies in recruitment to laser-induced DNA damage when nuclear, and often mislocalised from the nucleus to the cytoplasm (Figure 3-3). Given that the R1266* mutant fails to localise to DNA damage sites, it might be that the region C-terminal to this contributes to ERCC6L2's recruitment to DNA lesions. Broadly, this indicates a likely role for ERCC6L2 in the DDR and a possible DDR defect in ERCC6L2-mutant cells. However, more in-depth studies of the cellular behaviour of clinical ERCC6L2 truncations would be well served by forcing cells to retain these fragments in

nuclei, given the complications arising from mislocalisation. However, this is unlikely to be as simple as adding a nuclear localisation sequence (NLS) to the cloned constructs, since other data (Figure 4-3) indicate that cells may still export truncated NLS-bearing ERCC6L2 fragments. The use of an inhibitor of nuclear protein export such as leptomycin B may facilitate this experiment [351, 352]. Thorough extensions to these experiments using knockdown of upstream DNA damage sensors such as ATM, ATR, DNA-PK and the MRN complex could reveal insights into the mechanisms of ERCC6L2 recruitment to DNA damage sites. ERCC6L2 is still recruited to laser-induced DNA damage sites after pre-treatment with the inhibitors KU-55933, VE-821, NU-7441, mirin and olaparib, which inhibit ATM, ATR, DNA-PKcs, MRE11 and PARP1/2, respectively. However, with the exception of olaparib, which abrogated the recruitment of ALC1 to laser-induced DNA damage as published [128], the rest of these experiments were performed without positive controls for the action of these inhibitors.

While the finding that ERCC6L2 is recruited to laser-induced DNA damage constitutes evidence for a role in the DDR, the disadvantage of this technique is that it can cause a variety of DNA lesions. This experiment therefore bears the limitation that, at least in our case, it cannot be used to identify a specific response pathway in which ERCC6L2 operates. For this reason, we aimed to generate ERCC6L2 knockout U2OS cells, as such cell lines would enable us to search for sensitivity phenotypes which may be specific to certain drugs, and as such gain insights into the pathway(s) in which ERCC6L2 might function.

Using CRISPR/Cas9 technology to target exons 2 or 3 with two different sgRNAs, I generated four independent ERCC6L2$^{-/-}$ cell lines, two from each sgRNA. These cell lines were validated by Sanger sequencing of the genomic DNA spanning the target exons. It is known that exon skipping can occur across a damaged site, yielding in-

frame transcripts which may give rise to partially-functional proteins [353]. By sequencing of cDNA prepared from RNA extracts, I identified the above-mentioned four cell lines, which showed no evidence of exon skipping. It is to be kept in mind that we cannot exclude the possibility that on occasion some cells are still capable of generating a protein with a small deletion. In this case, however, any protein product should be rendered catalytically inactive, since both exons 2 and 3 are required for the portion of ERCC6L2 mRNA corresponding to the DEXDc domain of the catalytic core (Figure 3-4).

Clonogenic survival assays using the ERCC6L2 knockout cell lines generated revealed further insights into a possible DDR role for ERCC6L2. These cells demonstrated increased sensitivity to the DSB-inducing drugs phleomycin and etoposide. Importantly, I also identified DNA damaging agents to which ERCC6L2$^{-/-}$ cells do not exhibit increased sensitivity compared to parental control cells, including MMC, cisplatin, UV and MMS.

ERCC6L2$^{-/-}$-1A, a U2OS-derived knockout, did not show marked sensitivity to mitomycin C in comparison to the WT control cell line Ctrl 1, except for a mild sensitivity at the highest dose tested (200nM, 24h) (Figure 3-5a); this suggests that ERCC6L2 is not involved directly in the repair of DNA interstrand crosslinks. This finding is in contrast with two studies which have been published during this project: one in which siRNA-mediated depletion of ERCC6L2 in A549 cells led to very mild MMC sensitivity [332], and another in which a patient-derived fibroblast cell line deficient in functional ERCC6L2 exhibited some sensitivity to MMC [337]. However, the increased sensitivity observed in this second study was small in comparison to that seen in fibroblasts from a patient with Fanconi anaemia [337]. This suggests that the MMC sensitivity of ERCC6L2-deficient cells observed in this study may arise from a secondary DNA repair deficiency, distinct from classical ICL repair. Notably, patient

cells do not tend to exhibit the elevated chromosomal breakage phenotype characteristic of Fanconi anaemia [332, 337].

A possible role for ERCC6L2 in the response to SSBs or DSBs was suggested by the observation that the four ERCC6L2 knockout cell lines tested were all sensitive to phleomycin when compared to two WT control cell lines (Figure 3-5 b/c). Our data are supported by the fact that other groups have also noted this phenotype in patient-derived cell lines in studies published during this project [336, 337]. Phleomycin is a radiomimetic classically considered to induce SSBs and DSBs by direct chemical attack of the DNA backbone [354, 355]. However, more recent insights into its mechanism of action demonstrate that in addition to SSBs and DSBs, phleomycin (and to a lesser extent ionising radiation) can induce complex DNA lesions through the formation of deoxycytosine adducts. Repair of these adducts takes a relatively long period of time in human cells with a half-life of around 10h, suggesting that they require more intricate repair mechanisms than 'clean' DSBs [356, 357].

Other studies demonstrate no sensitivity of ERCC6L2-deficient cells to the SSB-inducing drug MMS [337, 358]; observations in agreement with preliminary data of my own. As such, it seems that ERCC6L2-deficient cells are sensitive to either DSBs or the complex, bulky lesions formed by aldehyde by-products of radiomimetic- or IR-induced damage. However, Zhang and co-workers identified no defect of ERCC6L2-deficient cells in either of the primary DSB repair pathways. Using a plasmid-based (pRec-SJ) assay for precise V(D)J recombination in fibroblasts from healthy individuals or derived from the ERCC6L2-deficient patient they identified, Zhang et al noted no V(D)J recombination defect in patient cells, in contrast to cells derived from a Cernunnos patient deficient in the critical c-NHEJ factor XLF [337]. The authors showed no HR defect in cells after transient knockdown of ERCC6L2, in contrast to cells depleted of

the critical recombinase RAD51 [337]. Zhang et al sowed this using the DR-GFP assay, in which cells possess two non-functional GFP coding sequences. One of these sequences has a small deletion which renders it unable to produce GFP and in the place of the deleted region has an ISce-I restriction site. The other sequence contains the coding sequence missing from the ISce-I site-containing sequence, so upon ISce-I-mediated DSB induction, HR-mediated repair reconstitutes a functional GFP sequence which can be detected and quantified by flow cytometry. Taken together, the observations that ERCC6L2 depletion compromises neither HR (measured by the DR-GFP assay) and NHEJ (measured by a V(D)J recombination assay), suggest that ERCC6L2 does not function in canonical double strand break repair [337].

In addition to phleomycin sensitivity, the set of ERCC6L2 knockout cell lines exhibited sensitivity to the TOPO2 inhibitor etoposide (Figure 3-5 b/c). Etoposide traps TOPO2 on nucleolytically processed DNA before religation of cleaved DNA ends, leading to DSB exposure following transcription or replication runoff [350]. It should also be noted that etoposide, as a TOPO2 inhibitor, may also cause problems for cells aside from DSBs, such as topological stress and associated perturbation of DNA replication.

Our data (Figure 3-5, Figure 3-7), taken alongside the observations by several groups published during this study [336, 337], show that ERCC6L2 deficiency increases sensitivity to several of DSB-inducing genotoxic insults. However, the lesions induced by the range of genotoxic insults to which ERCC6L2-deficient cells are sensitive (phleomycin, etoposide and IR) do not exclusively cause DSBs [332, 336, 337, 356, 357]. Furthermore, given the aforementioned findings by Zhang et al that ERCC6L2 does not seem to be essential for HR or c-NHEJ [337], the precise basis of ERCC6L2's DNA damage sensitivity is not immediately obvious and requires additional exploration. If the basis of this sensitivity is rooted in sensitivity to broad-spectrum

replication stress, it would be expected that ERCC6L2-deficient cells are sensitive to replication stress-inducing agents such as hydroxyurea (HU), aphidicolin, or the TOPO1 inhibitor camptothecin, although camptothecin also causes seDSBs. I have attempted to address this question, although the results are not clear-cut. Upon initial testing by clonogenic survival assay, ERCC6L2$^{-/-}$ cells seemed sensitive to HU, however this sensitivity was not reversed by overexpression of WT ERCC6L2, unlike the observations for phleomycin and etoposide shown in Figure 3-7. Currently, the question of camptothecin sensitivity is also unclear, although it seems possible that ERCC6L2$^{-/-}$ cells are indeed sensitive to this agent. In total I have tested all four WT cell lines and all four ERCC6L2-/- cell lines for camptothecin sensitivity by clonogenic survival assays, but the four control cell lines' tolerance of this agent was too variable to draw a strong conclusion. I am currently performing assays with the ERCC6L2-overexpressing cell lines to get a better idea of whether or not the possible camptothecin sensitivity I have observed is ERCC6L2-dependent.

The process of deriving knockout cell lines by CRISPR/Cas9 involves expressing a nuclease and establishing clonal cell lines from single-cell colonies could lead to off-target effects of Cas9 and the propagation of undesirable variation between cell lines. As such, I aimed to verify that the DNA damage sensitivity phenotypes observed in ERCC6L2$^{-/-}$ cells was a result of ERCC6L2 loss by establishing stable overexpression cell lines in the WT and knockout backgrounds and repeating the clonogenic survival assays. In this way, I aimed to determine whether or not the ERCC6L2 deficiency phenotypes observed in this study were caused directly by the loss of ERCC6L2.

Overexpression of WT ERCC6L2 in an ERCC6L2$^{-/-}$ background restores some viability of these cells when treated with phleomycin or etoposide (Figure 3-7), suggesting that the sensitivity phenotypes observed are caused by the loss of ERCC6L2. It has been

noted in other SNF2 ATPases that mutation of a conserved lysine in the catalytic site to arginine abrogates ATPase activity [128, 225]. Importantly, expression of the predicted ATPase-dead K165R mutant failed to restore viability to the same extent (Figure 3-7), demonstrating that ERCC6L2's catalytic activity is a necessary aspect of U2OS cells' tolerance of etoposide and phleomycin.

In ERCC6L2$^{-/-}$-1A-derived cell lines, overexpression of WT ERCC6L2 and the K165R mutant is very similar (Figure 3-6), and the K165R mutant rescues viability under both etoposide and phleomycin treatment to a much lesser extent than WT ERCC6L2 (Figure 3-7). However, in other backgrounds the relative expression level of different ERCC6L2 constructs does differ, meaning that further experiments must be interpreted with caution (Figure 3-6). It may be that only a small level of ERCC6L2 expression is necessary for the rescue of these DNA damage sensitivity phenotypes and it would be worthwhile for future experiments to determine relative ERCC6L2 transcript levels in the complemented cell lines.

For the first time, these results demonstrate the importance of ERCC6L2's ATPase activity in the DNA damage response (Figure 3-7), and support the notion that point mutants affecting this activity can cause the bone marrow failure phenotypes observed in patients harbouring such mutations [333, 336]. However, many *ERCC6L2* mutations in BMF patients do not likely disrupt ERCC6L2's ATPase activity directly, lying C-terminal to the catalytic core. All of these mutations are truncating mutations, however, rather than point mutants [332-337]. This suggests that functional elements C-terminal to the catalytic core – including C-terminal to the predicted NLS, which lies at amino acids 1004-1014 – are necessary for ERCC6L2's function. Based on findings in Figure 3-3 and Figure 4-7, this supporting role of C-terminal sequence elements might be important for ERCC6L2 recruitment to sites where its action is required.

Overall, this chapter demonstrates that ERCC6L2 functions in the DNA damage response. At first glance, it seems that this function relates somehow to double strand breaks. However, findings from an independent group that were published during this study show that there is no major mechanistic DSB repair defect in ERCC6L2-deficient cells [337], although an effect of ERCC6L2 loss on minor DSB repair pathways has not been tested. Furthermore, the DSB-inducing agents tested might also be causing other types of genotoxic stress, which could explain both the sensitivity of ERCC6L2$^{-/-}$ cells to these drugs, and the lack of any observable defect in NHEJ or HR [337]. In addition to a sensitivity of ERCC6L2$^{-/-}$ cells to genotoxic stress, I have found that ERCC6L2 is recruited to DNA damage induced by laser microirradiation, and shown that even clinical mutants of this protein which do localise to the nucleus are deficient in full recruitment to laser-induced DNA damage. Similar findings were published by another group during the course of this study, which showed that ERCC6L2 is recruited to laser-induced DNA damage and that the C-terminal half of ERCC6L2 supports this recruitment [337]. I have further underlined the physiological importance of ERCC6L2 in the DNA damage response, showing that an ATPase-dead mutant of ERCC6L2 is incapable of rescuing viability of ERCC6L2$^{-/-}$ cells treated with phleomycin or etoposide.

Taken together, these findings suggest that ERCC6L2 participates in the DNA damage response, and that although the precise nature of its function is not clear, it seems plausible that a DDR defect could contribute to bone marrow failure phenotypes in ERCC6L2-deficient patients. This statement is in broad agreement with suggestions by other groups made during the course of this study [332, 336, 337].

4 Exploring ERCC6L2 function at the cellular level

4.1 Results

We next set out to assess the cell cycle distribution of ERCC6L2$^{-/-}$ U2OS cells under etoposide and phleomycin treatment. Preliminary experiments using the DNA stain propidium iodide (PI) suggested an accumulation of ERCC6L2$^{-/-}$ cells in G2 (not shown). We sought to perform a more thorough analysis, using a Click-iT EdU (ethynyl deoxyuridine) Alexa Fluor 647 assay kit for flow cytometry, in which a 15 minute pulse of EdU incorporation is used to label cells actively undergoing DNA synthesis in addition to PI staining. This experiment demonstrated a dose-dependent loss of DNA replication in both Ctrl 2 and ERCC6L2$^{-/-}$-2A cells treated with etoposide for 48h, an effect which was more pronounced in ERCC6L2$^{-/-}$-2A cells (Figure 4-1). Raw data were collected by Dr. Dragana Ahel. Using the highest dose of etoposide (150nM) for 48h and increasing the EdU incorporation time to 3h to allow comparison between cell lines across a longer time period within S-phase, cell cycle analysis was performed on the sets of four knockout and four control U2OS cell lines described in Table 3-1. Indeed, each of the four knockout cell lines (ERCC6L2$^{-/-}$-1A, -1B, -2A and -2B) exhibited markedly less EdU incorporation under etoposide treatment than the four WT control cell lines tested (Figure 4-1).

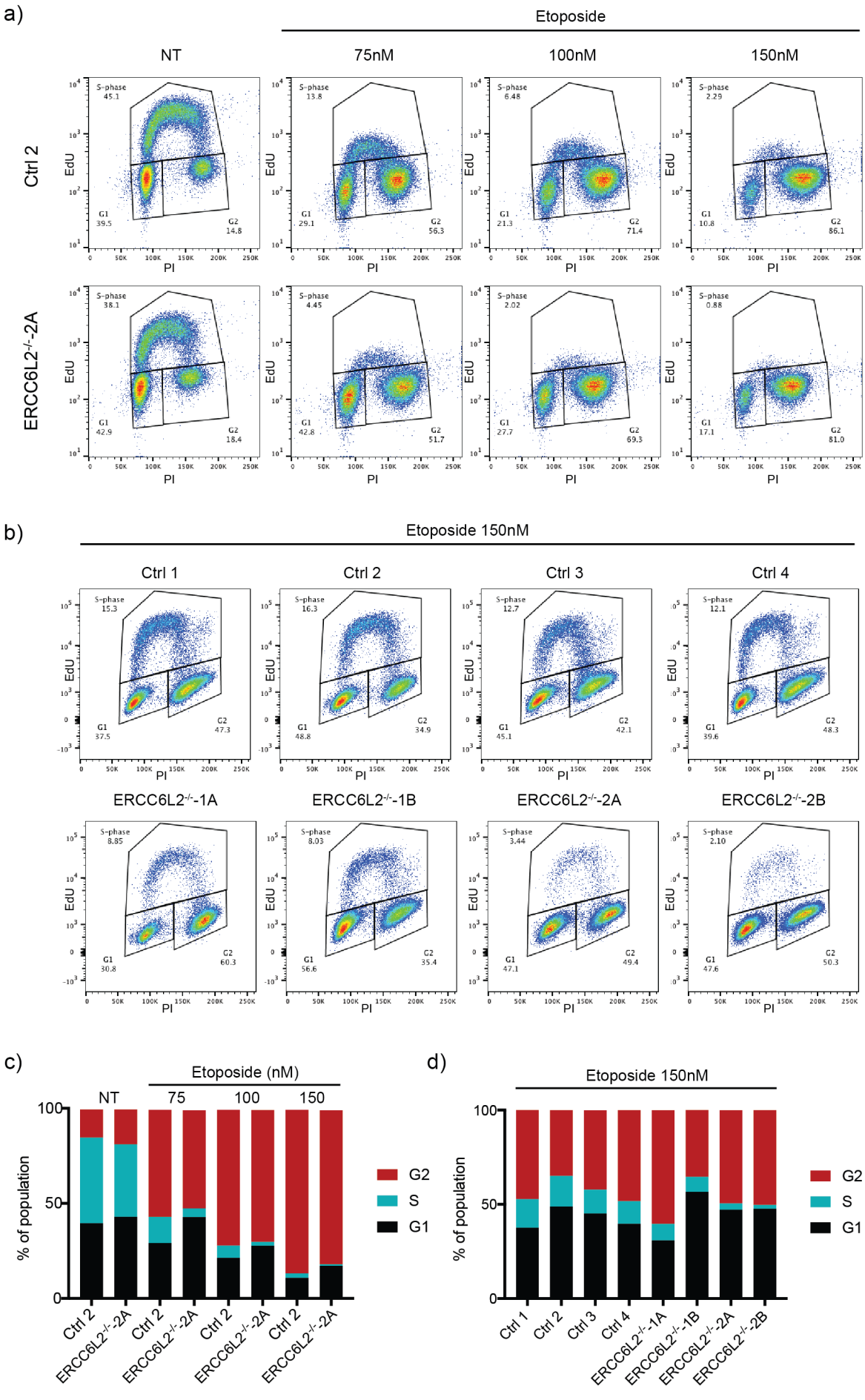

a)
Etoposide
NT
75nM
100nM
150nM
Ctrl 2
ERCC6L2-/--2A
EdU
PI
S-phase 45.1
G1 39.5
G2 14.8
S-phase 13.8
G1 29.1
G2 56.3
S-phase 6.48
G1 21.3
G2 71.4
S-phase 2.29
G1 10.8
G2 86.1
S-phase 38.1
G1 42.9
G2 18.4
S-phase 4.45
G1 42.8
G2 51.7
S-phase 2.02
G1 27.7
G2 69.3
S-phase 0.88
G1 17.1
G2 81.0
b)
Etoposide 150nM
Ctrl 1
Ctrl 2
Ctrl 3
Ctrl 4
ERCC6L2-/--1A
ERCC6L2-/--1B
ERCC6L2-/--2A
ERCC6L2-/--2B
S-phase 15.3
G1 37.5
G2 47.3
S-phase 16.3
G1 48.8
G2 34.9
S-phase 12.7
G1 45.1
G2 42.1
S-phase 12.1
G1 39.6
G2 48.3
S-phase 8.85
G1 30.8
G2 60.3
S-phase 8.03
G1 56.6
G2 35.4
S-phase 3.44
G1 47.1
G2 49.4
S-phase 2.10
G1 47.6
G2 50.3
c)
Etoposide (nM)
NT
75
100
150
% of population
100
50
0
G2
S
G1
Ctrl 2
ERCC6L2-/--2A
d)
Etoposide 150nM
Ctrl 1
Ctrl 2
Ctrl 3
Ctrl 4
ERCC6L2-/--1A
ERCC6L2-/--1B
ERCC6L2-/--2A
ERCC6L2-/--2B

Figure 4-1: Etoposide treatment impairs DNA synthesis in ERCC6L2$^{-/-}$ cells

a) Ctrl 2 and ERCC6L2$^{-/-}$-2A U2OS cells treated with increasing doses of etoposide for 48h followed by EdU incorporation for 15min still in the presence of etoposide and cell cycle analysis by flow cytometry using propidium iodide as a DNA stain. b) Ctrl 1-4 and the set of four ERCC6L2 knockout U2OS cell lines treated with 150nM etoposide for 48h and subjected to cell cycle analysis using Click-iT EdU labelling of nascent DNA. EdU labelling was performed for 3h, still in the presence of etoposide. c-d) Graphical representation of cell cycle distributions quantified in a) and b), respectively. This experiment is representative of more than three independent experiments.

As with the clonogenic survival assays, we next repeated the cell cycle analysis assays using the set of ERCC6L2-complemented cell lines generated previously (Figure 3-6 and Table 3-1). As in Figure 4-1, cells were treated with 150nM etoposide for 48h and EdU incorporation was performed for 3h. While ERCC6L2 expression had little effect on Ctrl 2 U2OS cells, WT ERCC6L2 successfully rescued the DNA replication defect in ERCC6L2$^{-/-}$-2A cells, an effect that was not observed with expression of the K165R ATPase-dead mutant (Figure 4-2).

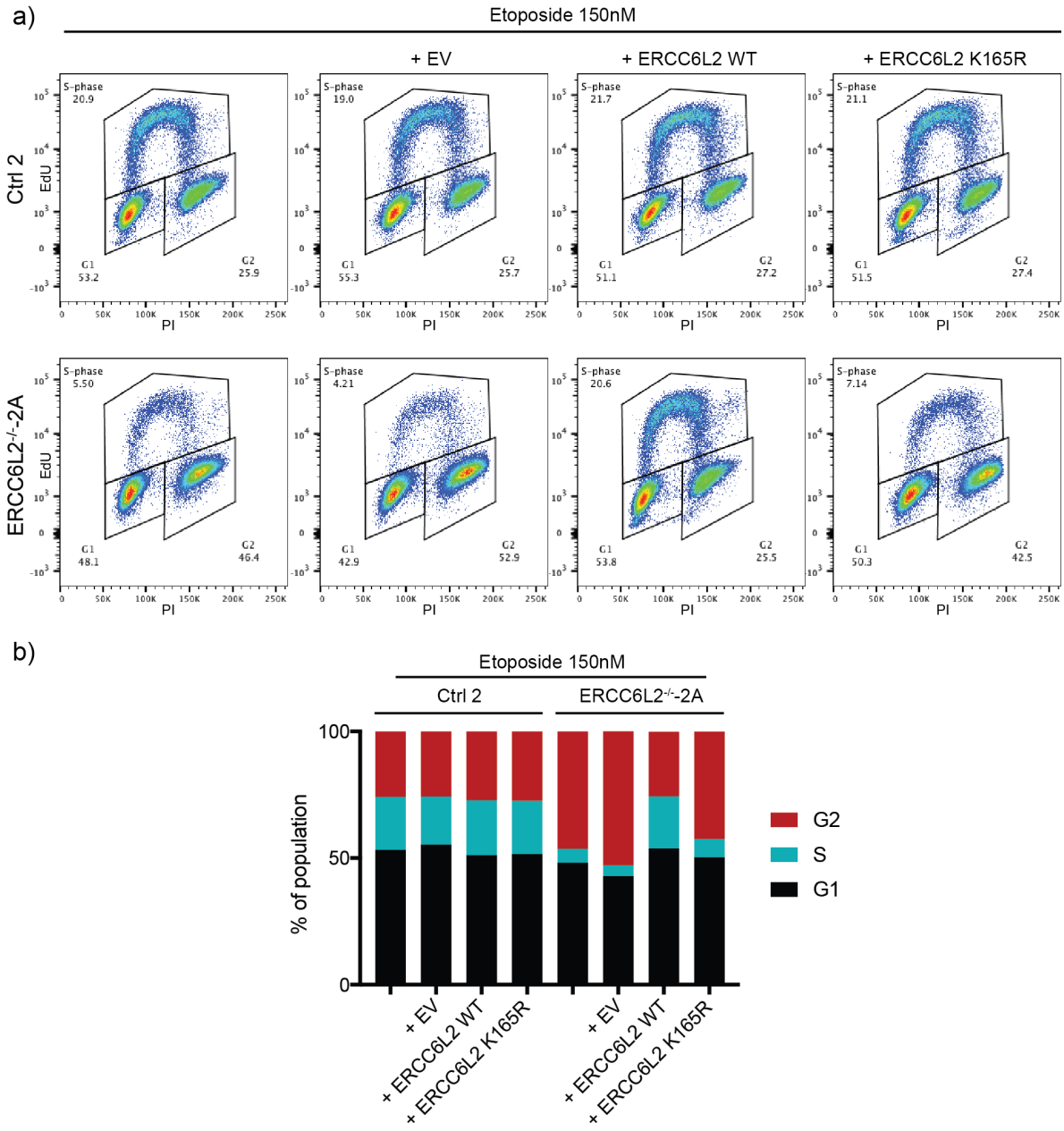


Figure 4-2: ERCC6L2 expression restores DNA synthesis under etoposide treatment in ERCC6L2-deficient cells in a manner dependent on its ATPase activity

a) Ctrl 2- and ERCC6L2$^{-/-}$-2A-derived complementation cell lines from Figure 3-6 subjected to etoposide treatment and cell cycle analysis as in Figure 4-1b. b) Graphical representation of cell cycle distributions quantified in a). Raw data were collected by Dr. Dragana Ahel. This experiment is representative of more than three independent experiments.

Having confirmed ERCC6L2 as a DNA damage-responsive protein and successfully established cellular phenotypes of ERCC6L2-deficient cells, I next aimed to explore the protein itself and identify regions of importance to its function aside from the catalytic core. The recruitment of the critical DSB repair factor 53BP1 to DSBs is thought to

involve the interaction of its tandem Tudor domains with dimethylated lysine 20 of histone H4 [117, 359]. The importance of 53BP1's tandem Tudor domains is highlighted by the more recent discovery of TIRR (Tudor interacting repair regulator), a protein which prevents 53BP1 from associating with H4K20me2 by binding the Tudor domain of 53BP1 directly, impairing its recruitment to DNA lesions [360, 361]. Given the relevance of 53BP1's tandem Tudor domains for 53BP1 recruitment to DNA damage sites, we reasoned that the Tudor domain of ERCC6L2 might mediate its recruitment to DNA lesions. To test this hypothesis, I cloned an ERCC6L2 mutant truncated from the N-terminus, lacking the Tudor domain, into the pDEST-YFP expression vector and assessed its recruitment to laser-induced DNA damage. However, the overexpressed ERCC6L2 fragment (aa94-1561) was still recruited efficiently to DNA lesions (Figure 4-3). In order to narrow down a region of ERCC6L2 which might be important for its recruitment to DNA damage sites, I next cloned N- and C-terminal fragments (aa2-711 and 712-1561, respectively) into the pDEST-YFP vector. The SV40 Large T-antigen NLS was also added to ERCC6L2 (2-711), as ERCC6L2 (712-1561) contains a predicted NLS which we expected to be responsible for ERCC6L2's nuclear localisation. Interestingly, while YFP-NLS-ERCC6L2 (2-711) was still mislocalised to the cytoplasm despite the NLS added and didn't mobilise to DNA damage sites, YFP-ERCC6L2 (712-1561) was recruited strongly and rapidly to laser damage tracks (Figure 4-3). This implies the presence of a portion (or portions) within ERCC6L2's C-terminus which may be important for its recruitment to DNA damage sites.

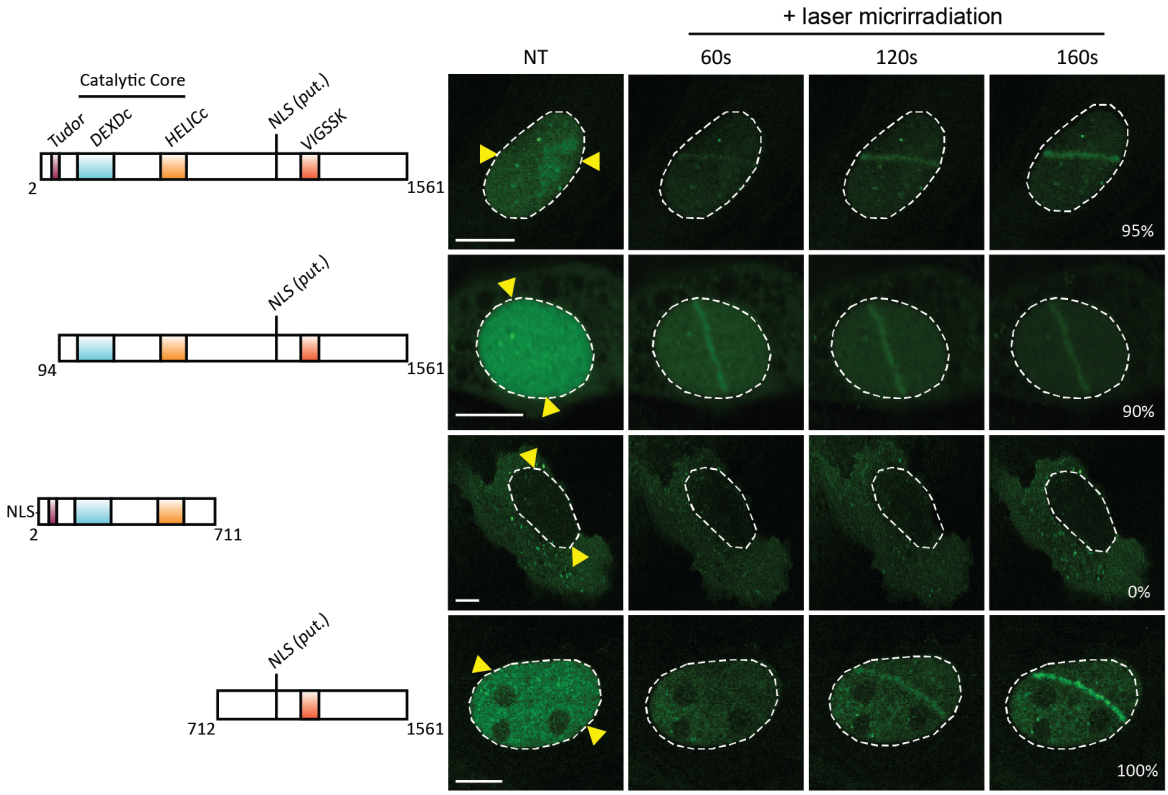


Figure 4-3: The C-terminus of ERCC6L2 is recruited to laser-induced DNA lesions

YFP-ERCC6L2 constructs (WT, (94-1561), NLS-(2-711), (712-1561)) expressed in U2OS cells and assessed for DNA damage recruitment by laser microirradiation following BrdU incorporation. Domain maps of the ERCC6L2 constructs expressed are shown alongside, including the putative NLS identified (aa1004-1014). Yellow triangles indicate the microirradiation target region and dotted while lines indicate the periphery of nuclei. The SV40 large T antigen NLS was added to ERCC6L2 (2-711). Scale bars = 10µm.

Interestingly, YFP-ERCC6L2 (712-1561) forms nuclear foci in roughly 20% of U2OS cells. Given that PCNA foci are visible in S-phase nuclei in clusters of replication forks – referred to as replication factories [189] – we considered it possible that the foci we observed were sites of on-going DNA replication. Indeed, YFP-ERCC6L2 (712-1561) foci colocalise with PCNA, indicating the presence of this fragment at replication factories (Figure 4-4). In contrast, YFP-ERCC6L2 WT does form nuclear foci, but these foci do not colocalise with PCNA and are visible in the vast majority of cells after pre-extraction of soluble nuclear proteins (Figure 4-4).

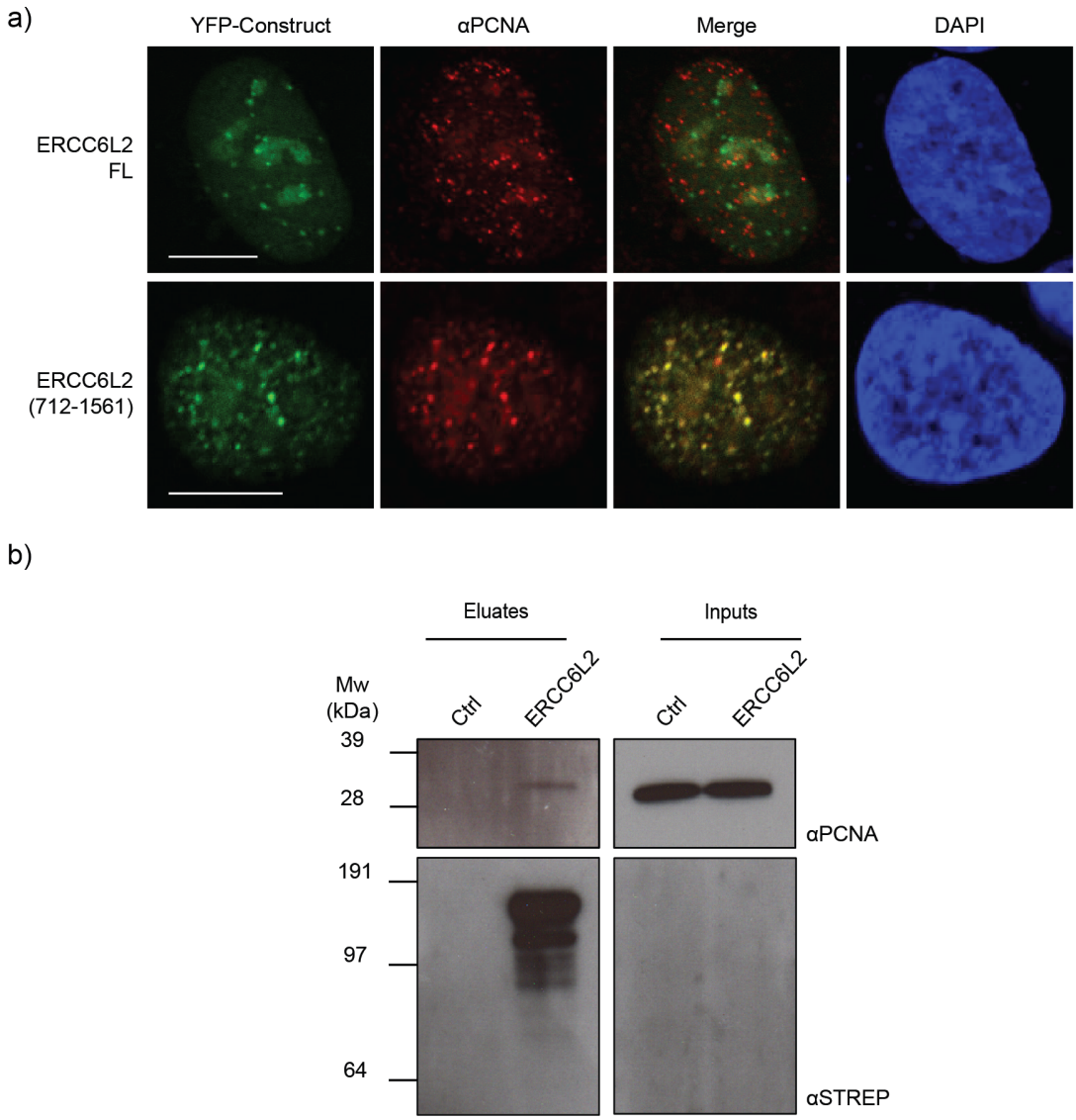


Figure 4-4: ERCC6L2 interacts with PCNA

a) Immunofluorescence against PCNA of unperturbed U2OS cells expressing YFP-ERCC6L2 or YFP-ERCC6L2 C-terminus (aa712-1561) subjected to pre-extraction and fixation. DNA was counterstained using DAPI. b) Western blot analysis against the Twin-Strep tag or PCNA in eluates and inputs from Strep-Tactin immunoprecipitation performed on 293T cells transiently expressing TwinStrep-ERCC6L2 (full-length). Scale bars = 10μm.

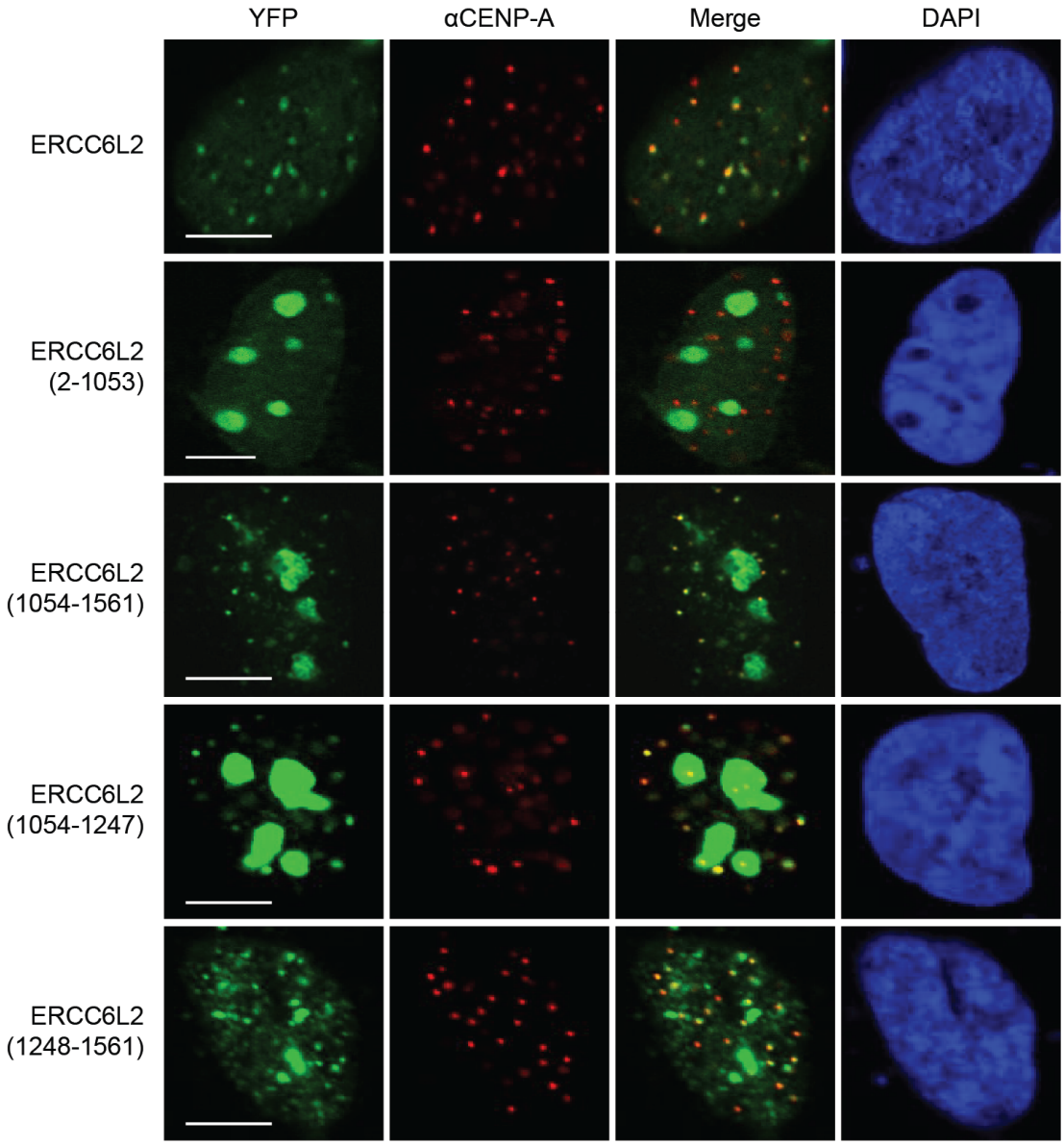
Given the colocalisation of PCNA with the C-terminus of ERCC6L2 in S-phase cells but importantly not FL ERCC6L2, I reasoned that ERCC6L2 and PCNA might interact only in relatively rare cases rather than at every replication fork, and that we may be able to detect such an interaction by immunoprecipitation. I overexpressed full length ERCC6L2 with an N-terminal Twin-Strep-tag in HEK 293Ts and pulled down Twin-Strep-ERCC6L2 using Strep-Tactin resin and performed a Western blot for PCNA on the

eluates. Indeed, a band migrating at the expected molecular weight of PCNA (29kDa) appeared in the eluate from Twin-Strep-ERCC6L2 cells, but not in the eluate from untransfected control cells (Figure 4-4). The observation that this interaction appears to be fairly weak, in combination with the fact that FL ERCC6L2 does not colocalise with PCNA in unperturbed cells, suggests that ERCC6L2 may interact with PCNA in specific contexts but at too few replisomes to visualise this by immunofluorescence.

Having noted the formation of chromatin-bound YFP-ERCC6L2 foci in all untreated interphase U2OS cells which do not colocalise with PCNA, we reasoned that identifying the nature of these non-replisome foci might help us to understand ERCC6L2's cellular function. By performing pre-extraction and immunofluorescence against the centromeric histone H3 variant CENP-A on U2OS cells overexpressing YFP-ERCC6L2, we identified CENP-A as a colocalisation partner for ERCC6L2, indicating that ERCC6L2 localises to centromeres in U2OS cells (Figure 4-5). To explore the basis of this localisation behaviour, I cloned a truncated form of ERCC6L2 (aa2-1053). This truncated protein no longer localised to centromeres, while the deleted region (aa1054-1561) did colocalise with CENP-A, indicating that ERCC6L2's centromeric localisation likely occurs through its C-terminus (Figure 4-5). In order to assess whether or not the VIGSSK domain is responsible for this localisation, I cloned and expressed a region containing the VIGSSK domain (aa1054-1247) and the region C-terminal to it (aa1248-1561). However, both of these fragments colocalised with CENP-A (Figure 4-5), suggesting that multiple interaction modules in the C-terminus might contribute to ERCC6L2's localisation to centromeres. Interestingly, each of the truncated ERCC6L2 fragments tested in this experiment exhibited strong localisation to what appear to be nucleoli; an observation I have made with many ERCC6L2 constructs

including – occasionally – full length ERCC6L2. However, YFP expressed with an N-terminal NLS also exhibited similar – albeit less dramatic – behaviour (Figure 3-1).

a)


YFP
αCENP-A
Merge
DAPI
ERCC6L2
ERCC6L2 (2-1053)
ERCC6L2 (1054-1561)
ERCC6L2 (1054-1247)
ERCC6L2 (1248-1561)

b)

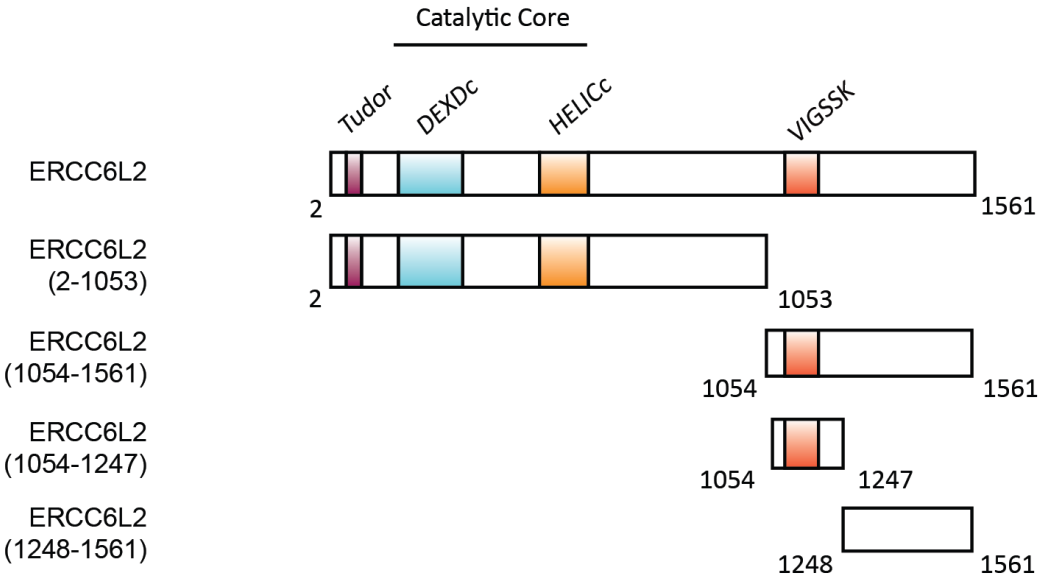

Catalytic Core
Tudor
DEXDc
HELICc
VIGSSK
ERCC6L2
2
1561
ERCC6L2 (2-1053)
2
1053
ERCC6L2 (1054-1561)
1054
1561
ERCC6L2 (1054-1247)
1054
1247
ERCC6L2 (1248-1561)
1248
1561

Figure 4-5: ERCC6L2 colocalises with centromeres through its C-terminus

a) CENP-A immunofluorescence on U2OS cells expressing a range of YFP-tagged ERCC6L2 constructs following pre-extraction and fixation. DAPI was used to stain DNA. b) Domain maps of the ERCC6L2 constructs expressed and analysed in a). Scale bars = 10μm.

Given the dynamic nature of chromatin, I next aimed to find out whether ERCC6L2's localisation at centromeres was confined to a particular phase of the cell cycle. First, I expressed YFP-ERCC6L2 and performed immunofluorescence against the constitutive centromere-associated network (CCAN) protein CENP-B and Cyclin A, which should not be present in G1 cells and should give strong staining in G2 cells [362, 363]. YFP-ERCC6L2 colocalised with CENP-B in both cyclin A-negative and strongly cyclin A-positive cells (Figure 4-6), indicating that ERCC6L2 can colocalise with centromeres in both the G1 and G2 phases of the cell cycle. Next, to test whether or not ERCC6L2 localises to centromeres during S-phase, I performed a 15 minute pulse of EdU incorporation followed by a Click-iT reaction to fluorescently label newly synthesised DNA in YFP-ERCC6L2 expressing cells, followed by immunofluorescence against CENP-A. YFP-ERCC6L2 still colocalised with CENP-A in EdU-positive (S-phase) cells, also showing at least partial colocalisation with both CENP-A and EdU in some cases (Figure 4-6). Finally, given the importance of CENP-A in kinetochore establishment [299-302], I performed immunofluorescence against CENP-A on YFP-ERCC6L2-expressing U2OS cells and searched for mitotic cells. In prometaphase cells, YFP-ERCC6L2 no longer appeared to colocalise with CENP-A (Figure 4-6).

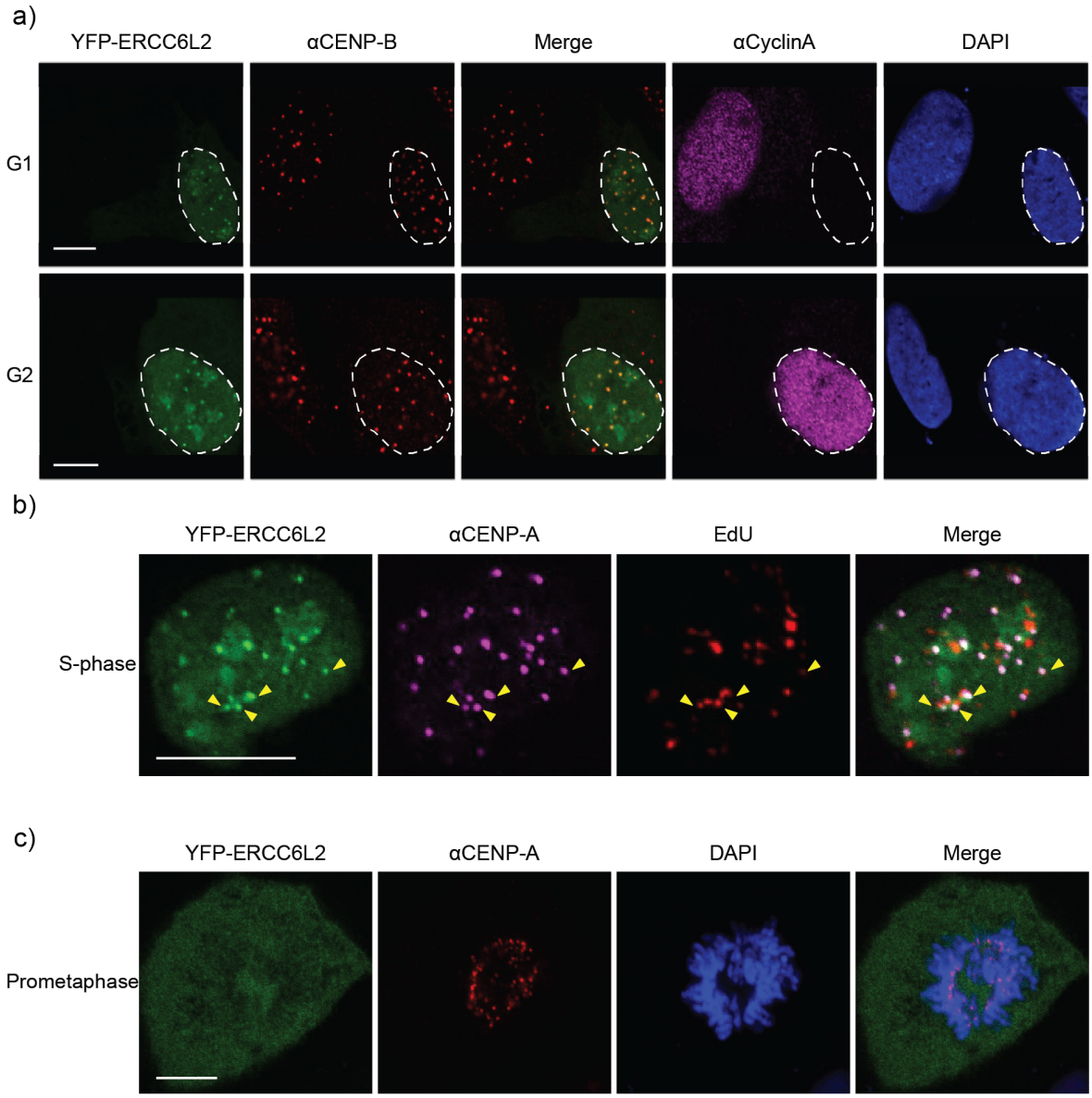


Figure 4-6: ERCC6L2 localises to centromeres throughout interphase

a) Immunofluorescence against CENP-B and Cyclin A on unperturbed U2OS cells expressing YFP-ERCC6L2 following pre-extraction and fixation. DAPI was used to stain DNA and dashed white lines indicate the periphery of YFP-ERCC6L2-containing nuclei. Cyclin A immunostaining was used to distinguish G1 from G2 cells. b) YFP-expressing U2OS cell subjected to Click-iT EdU labelling and detection, pre-extraction, fixation and CENP-A immunostaining. Yellow triangles indicate YFP-ERCC6L2 foci which colocalise with both CENP-A and EdU. c) Prometaphase cell expressing YFP-ERCC6L2 fixed without pre-extraction and immunostained for CENP-A. DAPI was used to stain DNA. Scale bars = 10μm.

Previously, we had noted mislocalisation and deficient recruitment to DNA damage sites of clinically relevant ERCC6L2 mutants (Figure 3-3). In order to test whether these mutants also lose centromere localisation, I expressed the YFP-tagged mutants in U2OS cells and performed immunofluorescence against CENP-A. As expected given the

mislocalisation, all of the cytoplasmic ERCC6L2 mutants tested, as well as the nuclear mutant (R1266*), failed to colocalise with CENP-A (Figure 4-7).

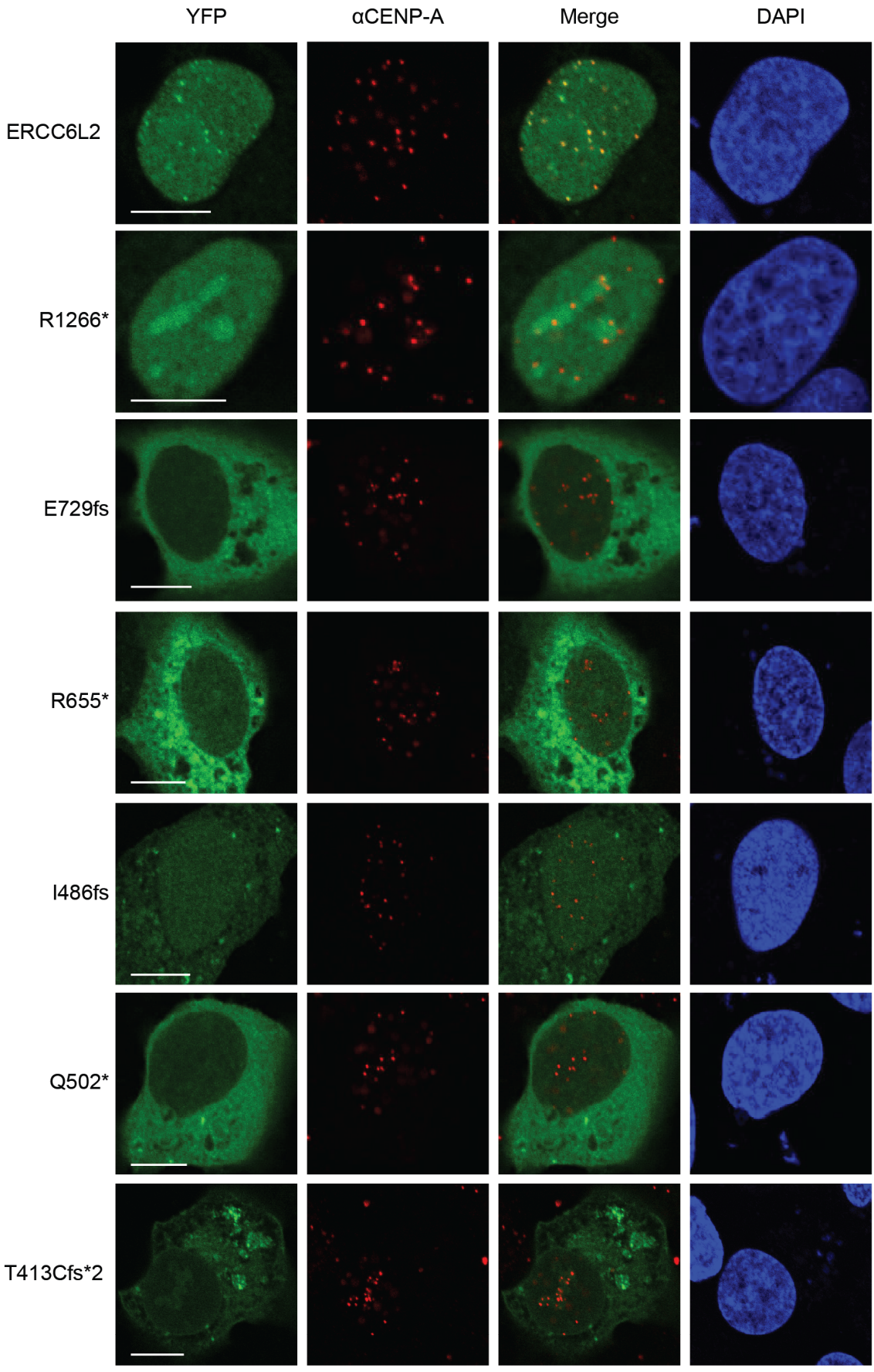


Figure 4-7: Clinically relevant ERCC6L2 truncations display defective centromere localisation

Immunofluorescence against CENP-A in unperturbed U2OS cells expressing YFP-tagged WT ERCC6L2 and a range of predicted protein products of ERCC6L2 mutants identified in bone marrow failure patients with DAPI used to stain DNA. Domain maps of the ERCC6L2 constructs expressed are shown in Figure 3-2. Scale bars = 10μm.

4.2 Discussion

In this Chapter, I set out to explore the function of ERCC6L2 at the cellular level. To do this, I first aimed to mimic genotoxic conditions to which ERCC6L2-deficient cells are sensitive and look for phenotypes at the cellular level. Because most DNA lesions caused by etoposide are replication-coupled and by this point I had discovered the ERCC6L2-PCNA interaction, I performed preliminary experiments in which I treated ERCC6L2-deficient cells with low doses of etoposide for long time points (48h) and noticed that the nuclei of ERCC6L2$^{-/-}$ cells under etoposide treatment appeared considerably larger than those of parental control cells. This suggested altered cell cycle distribution of knockout cells under these conditions.

Data collected by Dr. Dragana Ahel demonstrate that under etoposide treatment, ERCC6L2-deficient U2OS cells exhibit a dramatic DNA replication defect compared to WT U2OS cells, as measured by EdU incorporation and flow cytometry (Figure 4-1). While this effect can be rescued by overexpression of WT ERCC6L2, it is not rescued by expression of the ATPase-dead K165R mutant (Figure 4-2), suggesting that ERCC6L2's influence on DNA replication under etoposide treatment is dependent on its catalytic activity. This is in agreement with results from Section 3.1 (Figure 3-7), which demonstrate the importance of ERCC6L2's ATPase activity in the response to etoposide and phleomycin. This DNA replication defect under etoposide treatment could partially explain the loss of clonogenic potential of ERCC6L2$^{-/-}$ cells treated with this drug (Figure 3-5). A worthwhile extension to these assays would be to test the effect of the

TOPO1 inhibitor camptothecin and ICRF-193, a TOPO2 inhibitor which does not cause DSBs, on DNA replication in WT and ERCC6L2$^{-/-}$ cells.

Importantly, given that a multitude of disease-associated ERCC6L2 mutations would yield catalytically-dead forms of the protein (Table 1-1), the DNA replication defect we observe in ERCC6L2-deficient cells under genotoxic stress could constitute a clinically relevant phenotype. Interesting insights into the function of ERCC6L2's C-terminus could be gained by complementing ERCC6L2$^{-/-}$ cells with the R1266* mutant. This mutant protein still exhibits nuclear localisation and likely possesses ATPase activity but is not recruited to sites of DNA damage (Figure 3-3) or to centromeres (Figure 4-7).

Other findings in this Chapter demonstrate a novel interaction between ERCC6L2 and PCNA (discussed further below), implying that ERCC6L2's cellular function may relate to DNA replication. However, if ERCC6L2$^{-/-}$ cells fail to repair etoposide-induced DNA damage efficiently, the persistence of damage could induce broad-spectrum cell cycle arrest. As such, it may be important going forward to assess the activation of different cell cycle checkpoints under these conditions. These experiments are currently being undertaken by other members of the group.

Protein domains outside the ATPase domain of SNF2 ATPases are often essential for directing their catalytic activity to certain locations or for mediating their action within particular cellular contexts [95]. Considering the multitude of clinically relevant ERCC6L2 mutations which do not render ERCC6L2 catalytically inactive but yield protein products truncated from the C-terminus (Table 1-1), as-yet uncharacterised functional domains are likely to exist in this portion of the protein.

In order to identify regions of ERCC6L2 important for its function aside from its catalytic core, I performed laser microirradiation and live cell imaging in WT U2OS cells overexpressing YFP-tagged ERCC6L2 fragments of interest. Like WT ERCC6L2, a

truncation mutant lacking the Tudor domain, as well as a C-terminal fragment of ERCC6L2 (aa712-1561) both localised to the nucleus and were recruited to sites of DNA damage. In contrast, the N-terminal half of ERCC6L2 mislocalised to the cytoplasm even after the addition of an NLS and failed to mobilise to DNA lesions (Figure 4-3). These observations suggest that the Tudor domain alone is not required for ERCC6L2's recruitment to DNA damage sites, and that recruitment is instead mediated by some functional element of the C-terminal portion of the protein, which contains an uncharacterised domain (VIGSSK) and a multitude of conserved regions. Importantly, this strengthens the suggestion from Figure 3-3 that ERCC6L2 mutants with C-terminal truncations might exhibit defective recruitment to DNA damage sites.

During the course of this study, another group published similar findings showing that the annotated 'short' isoform of ERCC6L2, roughly corresponding to our N-terminal construct, localises to the cytoplasm while the full-length isoform is nuclear [337]. Zhang and co-workers also demonstrated the importance of ERCC6L2's C-terminus in cells' tolerance of phleomycin [337], but we show that the C-terminus can mobilise to DNA lesions independently of the N-terminus (Figure 4-3). Taken together, these findings suggest that the C-terminus contains some kind of interaction domain responsible for this recruitment.

An important observation made during the course of these experiments was that the C-terminal half of ERCC6L2 (but not the full-length protein) forms PCNA-colocalising foci in around 20% of nuclei, indicating its localisation to DNA replication factories (Figure 4-4a). This led us to find for the first time that WT ERCC6L2 interacts with PCNA, a scaffold protein critical for mediating the recruitment of DNA replication and repair factors to replication forks in a highly context-specific manner (Figure 4-4b). Taken alongside the observation that ERCC6L2-deficient cells exhibit a strong DNA

replication defect under mild etoposide treatment (Figure 4-1), this interaction implies a possible function for ERCC6L2 at the replication fork. For further studies into the importance of the ERCC6L2-PCNA interaction, performing cell synchronisation to enrich for S-phase cells followed by mass spectrometry experiments such as BioID may enable the identification of transient ERCC6L2 interactors at replisomes and shed light on ERCC6L2's role at the replication fork. Notably, recent findings reported by Tummala et al included a candidate, albeit unvalidated, interactome of ERCC6L2 which did not include PCNA but did include a range of other proteins associated with DNA replication, including members of the MCM replicative helicase and cohesin complexes [336]. While this candidate interactome may represent an interesting starting point for further interaction studies, collectively these data support a function of ERCC6L2 relating to DNA replication.

The fact that full-length ERCC6L2 interacts with PCNA but does not appear to colocalise with it in unperturbed cells, suggests one of several possibilities. First, that ERCC6L2 and PCNA interact at a replication fork only when a cell is experiencing some kind of localised replication stress, and that this can occur in the absence of exogenous genotoxic insult. Second, that ERCC6L2 interacts with PCNA during normal DNA replication at a subset of replication forks at specific genomic loci where replication is challenging, such as sites prone to forming replication-perturbing secondary structures. In either case, it must be assumed that within one replication factory, only a minority of replication forks bear ERCC6L2, otherwise ERCC6L2-PCNA colocalisation would be visible by conventional light microscopy in unperturbed cells.

In either scenario, the fact that ERCC6L2's C-terminus colocalises readily with PCNA while the full-length protein does not, would imply some degree of regulation placed upon ERCC6L2 to prevent illegitimate replication fork localisation, which is lost when

only the C-terminus is expressed. For example, some structural element of ERCC6L2 might mask the interaction site or sites responsible for this localisation, providing a layer of autoinhibition. Similar autoinhibition (although of the ATPase activity) has been noted for the *Schizosaccharomyces pombe* homologue of ERCC6L2's relative CSB. This protein (Rhp26) possesses a conserved leucine latch motif, which through folding of a hinge region interacts with its own ATPase domain, inhibiting ATP hydrolysis [364]. In order to understand ERCC6L2's role in concert with PCNA, it will be necessary to understand how ERCC6L2's PCNA interaction is stimulated, and under which conditions this occurs. Furthermore, exploring whether the ERCC6L2-PCNA interaction is enriched at certain genomic loci could yield valuable insights into the cellular importance of this interaction.

While ERCC6L2 colocalisation with PCNA was not observed in untreated U2OS cells, I did notice that ERCC6L2 forms nuclear foci which do not colocalise with PCNA (Figure 4-4a). These foci are more easily visible following pre-extraction of soluble nuclear protein prior to fixation, and it seemed likely that identifying the nature of these foci could be important in understanding ERCC6L2's function. By co-immunofluorescence with a wide range of factors, we observed that these ERCC6L2 foci colocalise with the centromeric histone variant CENP-A, demonstrating that ERCC6L2 is enriched at centromeres (Figure 4-5).

Considering that several clinical ERCC6L2 mutations are unlikely to impair its catalytic activity but instead cause C-terminal truncations (Table 1-1), I next aimed to identify portions of ERCC6L2 that might contribute to its centromere localisation.

Experiments aimed at characterising the basis of this localisation highlighted a C-terminal region as being responsible for ERCC6L2's centromere enrichment. However, my experiments imply that multiple parts of ERCC6L2's C-terminus might contribute to

centromere localisation, including the as-yet uncharacterised VIGSSK domain (Figure 4-5). In accordance with this, clinically relevant forms of ERCC6L2 do not colocalise with centromeres, including the R1266* mutant, which unlike other mutants tested is retained in nuclei (Figure 4-7). These observations suggest that the C-terminus of ERCC6L2 might contain multiple sites which direct ERCC6L2 to centromeres. One notable observation from these experiments, along with many others during the course of this study, is the apparent nucleolar localisation of a wide range of ERCC6L2 fragments, including full length ERCC6L2. While it cannot be excluded that this is an artefact of overexpression, since NLS-YFP exhibits similar behaviour (Figure 3-1), it remains possible that ERCC6L2 functions in nucleoli. Within the framework of our working hypothesis for ERCC6L2 function, discussed in Section 6.2, in which ERCC6L2 facilitates replication or transcription through repetitive DNA sequences which can form secondary DNA structures and impede these processes, a role for ERCC6L2 in nucleoli remains entirely plausible. Notably, Shwachman-Diamond syndrome patients suffer bone marrow failure phenotypes as a result of defective ribosome biogenesis while also being predisposed to both MDS and AML [341, 342, 365]. Given the similarities between this disease and the ERCC6L2 deficiency disorders described [332-337], it would be interesting for further work to explore the possibility that rDNA transcription and ribosome biogenesis are impaired in ERCC6L2-deficient cells. Indeed, recent findings by Tummala et al pointing to a defect in transcription recovery in ERCC6L2-deficient cells might reflect a defect in rDNA transcription, since rDNA is highly transcriptionally active [336].

The importance of centromeres in establishing the kinetochore during mitosis is well established, while recent studies indicate altered regulation of DNA replication specifically at centromeres [290]. As such, I considered it important to assess the cell

cycle context of ERCC6L2's centromere localisation, aiming to determine in which stage or stages this occurs.

Experiments using markers of cell cycle stage demonstrated that ERCC6L2 colocalises with centromeres throughout interphase but not during mitosis (Figure 4-6). Notably, this suggests that ERCC6L2 does not function directly in chromosome segregation at the kinetochore. These results therefore imply that ERCC6L2's role at centromeres occurs specifically during interphase. The fact that ERCC6L2 remains at these sites during their replication might suggest that the ERCC6L2-PCNA interaction comes into play in this context, given the importance of PCNA in regulating the localisation of many different proteins at replication forks. Such proteins include DNA polymerases ε and δ during normal replication, TLS polymerases during translesion synthesis and CAF1, WSTF and SMARCAD1 during chromatin establishment behind a replication fork [197].

In broad terms, a role for ERCC6L2 in some aspect of DNA replication is supported by two major findings. Firstly, its interaction with PCNA (Figure 4-4) and secondly the defective DNA synthesis noted in ERCC6L2-deficient cells treated with etoposide (Figure 4-1), a drug to which these cells are sensitive (Figure 3-5). Taking into account the localisation of ERCC6L2 to centromeres even during replication at these sites (Figure 4-6b), the hypothesis that ERCC6L2 functions in centromere replication is an attractive one. Due to its repetitive nature, centromeric DNA is more prone to the formation of secondary structures. Such structures can block or slow the progression of replicative DNA polymerases [51, 290, 297, 298, 304-306]. Indeed, it has been shown recently that centromeric replication utilises an array of proteins not usually associated with non-centromeric replication, including MSH2-MSH6, the MRN complex and XRCC1 [290]. As such, it could be that ERCC6L2 functions in relaxing secondary structures, which may otherwise disrupt the progression of DNA polymerases.

If ERCC6L2 does indeed play a facilitative role in centromere DNA replication, this could have important implications for genome stability in cells lacking ERCC6L2. One might expect that in ERCC6L2-deficient cells, under-replicated centromeric DNA could lead to increasing genomic instability with multiple rounds of cell division. This might manifest itself by the presence of micronuclei resulting from lagging chromosomes mis-segregating during anaphase, or an increased level of ultrafine anaphase bridge formation or persistence. These experiments are currently underway by other members of the group, and preliminary results do suggest an increase in UFB persistence and micronuclei formation even in untreated ERCC6L2$^{-/-}$ cells.

Since defects in centromere replication could constitute dangerous phenotypes for both cells and patients, it would be interesting to assess the rate of centromeric DNA replication in ERCC6L2$^{-/-}$ cells in future studies, using centromere-specific SMARD (single-molecule analysis of replicated DNA), as described recently [308].

As a point of speculation, if ERCC6L2 is involved in resolving DNA structures that form at centromeric repeats, it could be that similar structures also form at other locations or as intermediates in some DNA repair pathways. This could explain ERCC6L2's recruitment to DNA damage sites and role in the DDR. Indeed, ERCC6L2 would not be alone in such a multi-faceted role, with a host of other proteins such as DNA2 and BLM performing similar functions relevant to multiple aspects of genome maintenance [81, 90, 308, 366].

On the surface, it might seem that this interpretation doesn't explain the presence of ERCC6L2 at centromeres during G1, or its poorly understood association with the transcription machinery and RNA-DNA hybrids [336]. However, it is known that transcription at centromeric DNA is required for stable deposition of nascent CENP-A into centromeric nucleosomes [367]. Since the persistence of bulky DNA structures can

cause problems for transcription machinery as well as replication forks [368], ERCC6L2 might function to resolve such structures ahead of both DNA and RNA polymerases. Indeed, depletion of the helicase/nuclease DNA2, which was recently shown to unwind and cleave DNA secondary structures at centromeres, leads to reduced CENP-A deposition at centromeres [308], suggesting that factors which process problematic DNA structures at centromeres can influence multiple aspects of centromere maintenance.

Given the range of factors which act to prevent persistent DNA structures from causing severe chromosome segregation defects in mitosis, it could well be that functional redundancy exists between ERCC6L2 and other factors such as DNA2 or the nucleases GEN1, MUS81 or SLX1-SLX4. Depleting or inhibiting these factors in our ERCC6L2 knockout cell lines might reveal as-yet undiscovered genome stability- or centromere-related phenotypes. Furthermore, preliminary experiments from another member of the group (Lucy Armstrong) suggest that a fragment of ERCC6L2 (aa1054-1257), which contains the VIGSSK and colocalises with CENP-A in U2OS cells, can bind dsDNA and ssDNA *in vitro*. Assessing whether or not this fragment binds other substrates bearing secondary structure with better affinity might give us some insight into ERCC6L2's role at a mechanistic level. It could also be possible that ERCC6L2 plays a role at other repetitive genomic loci such as common fragile sites; chromatin immunoprecipitation (ChIP) experiments aimed at identifying loci of ERCC6L2 occupancy might answer such a question. However, a reliable ChIP-grade ERCC6L2 antibody is not yet commercially available. The development of a high quality antibody against ERCC6L2 would enable us to perform a number of other experiments, which thus far have been challenging to tackle. Although we attempted to generate a polyclonal ERCC6L2 antibody, raised against the Tudor domain, this antibody appeared

to recognise the antigen we had sent but failed to bind overexpressed ERCC6L2, even following immunoprecipitation of tagged and overexpressed ERCC6L2.

Considering a possible role for ERCC6L2 at centromeres, it would be of interest to compare centromere stability in ERCC6L2$^{-/-}$ or patient-derived ERCC6L2-deficient cell lines with cell lines deficient in DNMT3B. The loss of DNMT3B causes ICF syndrome (immunodeficiency-centromeric instability-facial abnormalities syndrome 1), the cellular pathologies of which include centromere instability [369]. Examining centromere instability fluorescence in-situ hybridisation and/or CENP-A staining in interphase and metaphase cells derived from ICF or ERCC6L2-deficient BMF patients could reveal whether or not centromere instability is a clinically relevant feature of ERCC6L2 deficiency disorders.

Overall, the findings presented in this chapter might point to an interesting function for ERCC6L2 in genome stability. ERCC6L2 interacts with PCNA and colocalises with centromeres, and its knockout exacerbates the DNA replication deficiency caused by etoposide treatment.

Our working hypothesis is that ERCC6L2 functions somehow in facilitating centromere replication or preserving centromere integrity; this line of research represents an exciting opportunity for gaining important new insights into genome stability. Other members of the group are currently undertaking experiments aimed at exploring this hypothesis further. Given that a role for ERCC6L2 in the replication of centromeres or across challenging DNA structures is likely to involve its PCNA interaction, I next sought to better understand the nature of this interaction.

5 Characterising the ERCC6L2-PCNA interaction

5.1 Results

Results shown in Section 4.1 (Figure 4-4) demonstrated an interaction between ERCC6L2 and PCNA, a protein which plays a critical role at the replication fork, mediating the recruitment of a wealth of different factors to sites of DNA synthesis in a context-dependent manner [197]. Such interactions usually occur through a motif called a PIP-box, bearing the signature Q-x-x-[VILM]-x-x-[FY]-[FY] [197].

In order to gain a better understanding of the mechanistic basis of ERCC6L2's interaction with PCNA, I sought to characterise this interaction at the molecular level. Examination of ERCC6L2's primary amino acid sequence revealed a conserved sequence matching that of a canonical PIP-box towards the C-terminus (Figure 5-1b). In order to test the hypothesis that this was a PCNA binding site, I first purified recombinant His-tagged PCNA by Ni-NTA from *E. coli* (Figure 5-1c). Next, I immobilised synthetic biotinylated peptides containing the ERCC6L2 candidate PIP-box or a predicted binding-null mutant on Streptavidin beads and used these peptides as bait for the recombinant PCNA. I also used the *bona fide* PIP-box of the SNF2 ATPase ZRANB3 as a positive control and a confirmed binding-null mutant as a negative control [225]. After washing the beads thoroughly and boiling them in SDS-containing loading dye, I performed a Western blot against PCNA. While the WT PIP-box of ZRANB3 bound PCNA strongly and no binding was detected for the mutant ZRANB3 PIP-box, no PCNA binding by the ERCC6L2's candidate PIP-box was detected (Figure 5-1d).

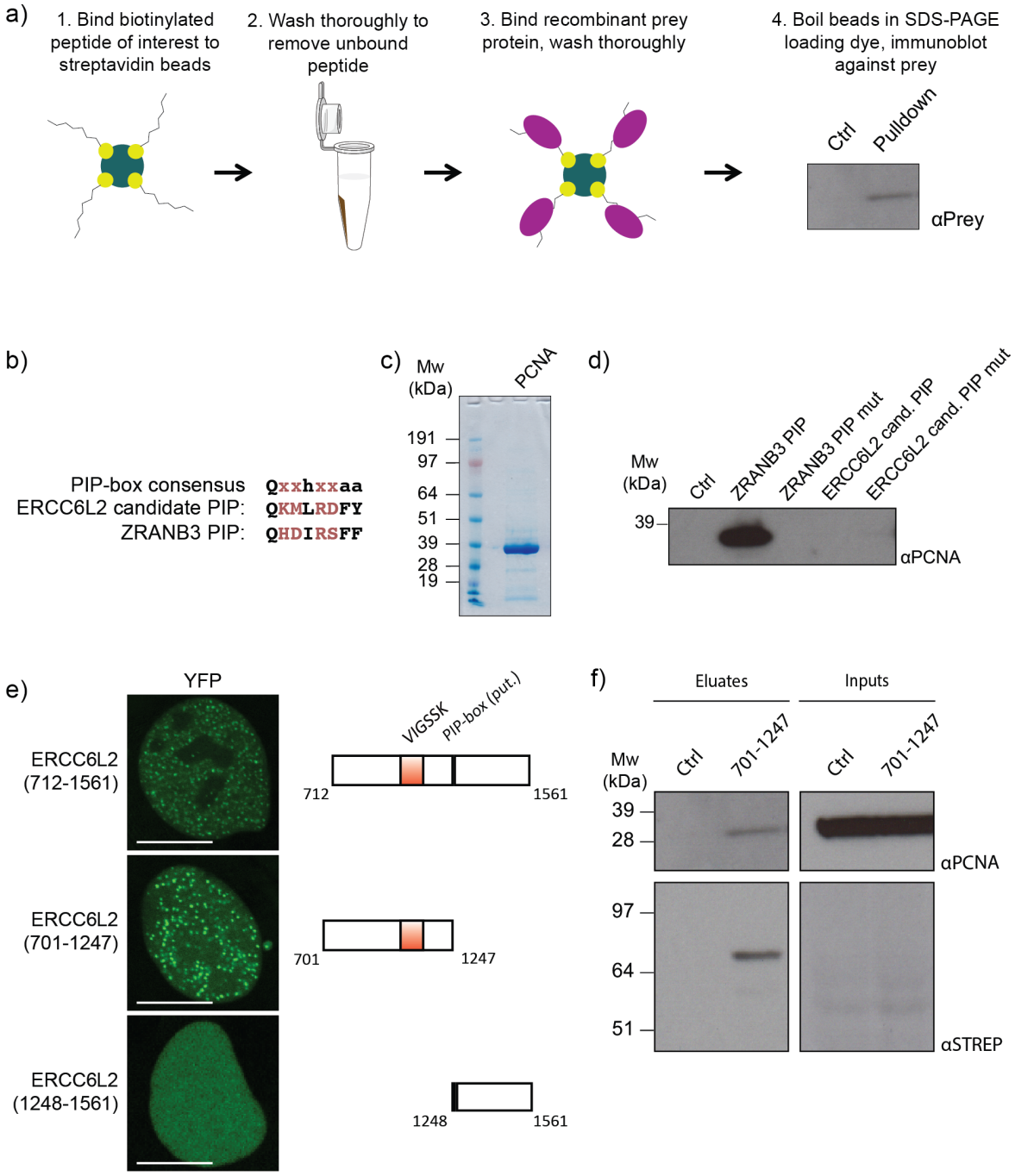


Figure 5-1: ERCC6L2 interacts with PCNA through a novel PCNA binding site

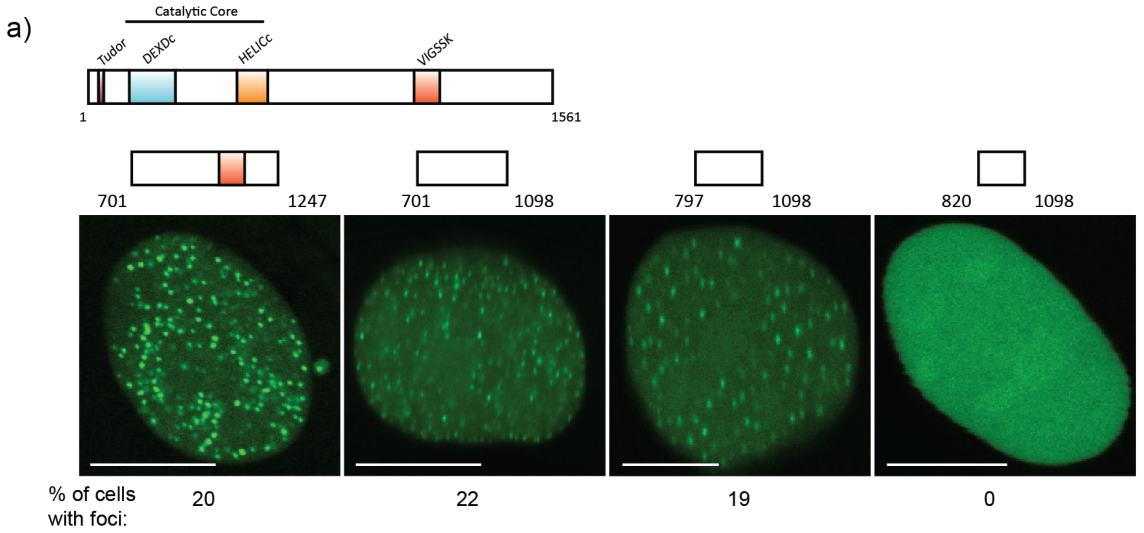
a) Strategy for assessing *in vitro* interaction of recombinant PCNA with synthetic biotinylated peptides of interest. Biotinylated peptides were immobilised on streptavidin dynabeads, washed thoroughly, and incubated with recombinant PCNA. Following PCNA binding, beads were washed thoroughly and boiled in SDS-PAGE loading dye, which was then subjected to Western blot analysis against PCNA to assess peptide-PCNA binding. b) A candidate PIP-box identified in ERCC6L2's C-terminus (aa1267-1274) aligned to the consensus PIP-box sequence and the canonical PIP-box of ZRANB3 (aa519-526). c) Coomassie blue staining of recombinant PCNA purified from *E. coli* by Ni-NTA. d) *In vitro* pulldown of recombinant PCNA using the approach outlined in a), performed on peptides containing the ERCC6L2 candidate PIP-box and a predicted binding-null mutant. Peptides containing WT or a binding-null mutant of ZRANB3's *bona fide* PIP-box were used as positive and negative controls, respectively. e) Representative images of nuclear foci formation of C-terminal YFP-tagged ERCC6L2 constructs expressed in unperturbed U2OS cells. Domain maps of the expressed constructs are displayed alongside. f) Strep and PCNA Western blot performed on inputs and

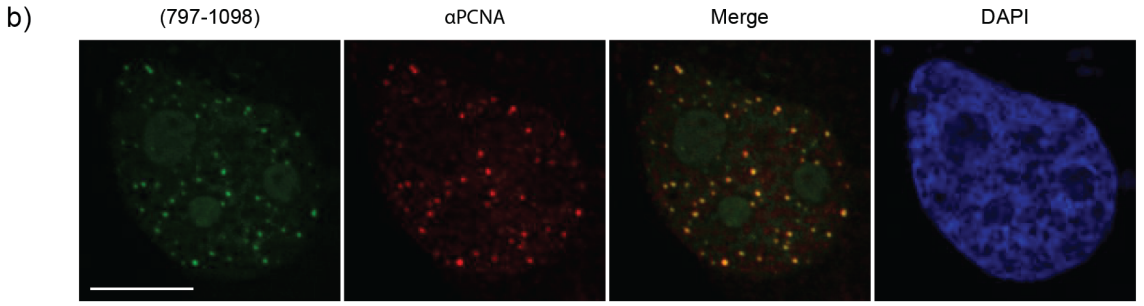
eluates from a StrepTactin immunoprecipitation performed on 293T cells expressing Twin-Strep ERCC6L2 (701-1247) or left untransfected. Scale bars = 10μm.

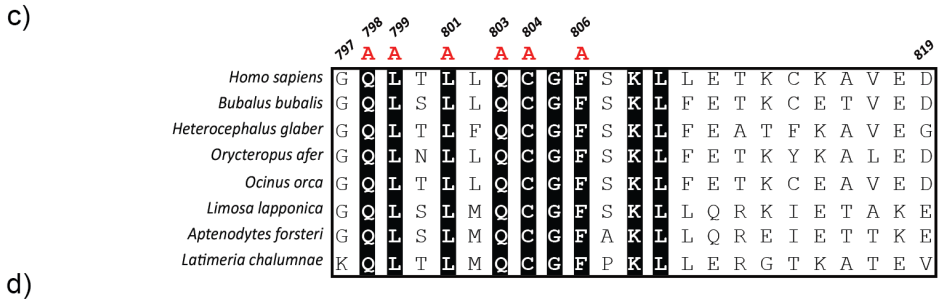
I next considered the possibility that ERCC6L2's interaction PCNA is mediated by a non-canonical PCNA binding site, and cloned fragments of the C-terminal portion of the protein that had colocalised with PCNA; these fragments consisted of amino acids 701-1247 and 1248-1561. I expressed YFP-tagged forms of these fragments in U2OS cells and noticed that the 701-1247 fragment formed nuclear foci reminiscent of PCNA foci in around 20% of cells. In contrast, the 1248-1561 fragment – which contains the candidate canonical PIP-box – did not form such foci (Figure 5-1e). Furthermore, when I performed immunoprecipitation on the 701-1247 fragment with a Twin-Strep tag, I saw that this fragment interacts with PCNA (Figure 5-1f). Since there is no canonical PIP-box sequence within this fragment, these results suggested to us that ERCC6L2 interacts through a novel class of binding site.

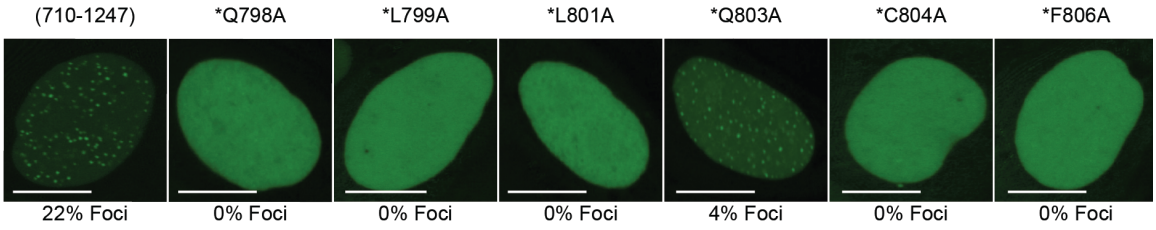
To map the PCNA interaction site, my approach consisted of making YFP-tagged truncations from the core 701-1247 fragment to remove patches of conserved amino acids, and then using nuclear foci formation in U2OS cells as a readout for a probable interaction-proficient fragment. Fragments 701-1098 and 797-1098 both still formed PCNA-like foci, but the 820-1098 fragment did not form nuclear foci in any YFP-positive cells (Figure 5-2a). Furthermore, the foci formed by the 797-1098 fragment colocalised with PCNA (Figure 5-2b). This suggested that the conserved sequence we had identified between G797 and D820 might contain a PCNA interaction site. Next, in the original 701-1247 fragment, I mutated six conserved residues we predicted to participate in such an interaction, and assessed foci formation of these YFP-tagged mutants. Five of the mutants (Q798A, L799A, L801A, C804A and F806A) failed to form nuclear foci, while one mutant (Q803A) formed foci in a reduced percentage of cells (4%) compared to the wild-type fragment (20%) (Figure 5-2d). In immunoprecipitations of the 701-

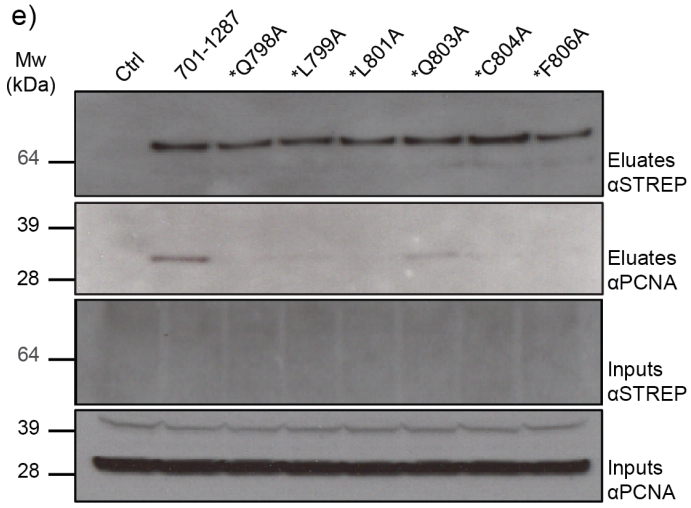
1247 fragment and its mutants, four of the mutants when Twin-Strep tagged failed to interact with PCNA, (Q798A, L801A, C804A, F806A) whereas Q803A still displayed a weak interaction. A very weak signal appeared to indicate reduced affinity of the L799A mutant for PCNA (Figure 5-2e).


a)
Catalytic Core
Tudor
DEXDc
HELICc
VIGSSK
1
1561
701
1247
701
1098
797
1098
820
1098
% of cells with foci:
20
22
19
0


b)
(797-1098)
αPCNA
Merge
DAPI


c)
797
798
799
801
803
804
806
819
A A A A A
Homo sapiens G Q L T L L Q C G F S K L L E T K C K A V E D
Bubalus bubalis G Q L S L L Q C G F S K L F E T K C E T V E D
Heterocephalus glaber G Q L T L F Q C G F S K L F E A T F K A V E G
Orycteropus afer G Q L N L L Q C G F S K L F E T K Y K A L E D
Ocinus orca G Q L T L L Q C G F S K L F E T K C E A V E D
Limosa lapponica G Q L S L M Q C G F S K L L Q R K I E T A K E
Aptenodytes forsteri G Q L S L M Q C G F A K L L Q R E I E T T K E
Latimeria chalumnae K Q L T L M Q C G F P K L L E R G T K A T E V


d)
(710-1247)
*Q798A
*L799A
*L801A
*Q803A
*C804A
*F806A
22% Foci
0% Foci
0% Foci
0% Foci
4% Foci
0% Foci
0% Foci


e)
Mw (kDa)
Ctrl
701-1287
*Q798A
*L799A
*L801A
*Q803A
*C804A
*F806A
64
39
28
64
39
28
Eluates αSTREP
Eluates αPCNA
Inputs αSTREP
Inputs αPCNA

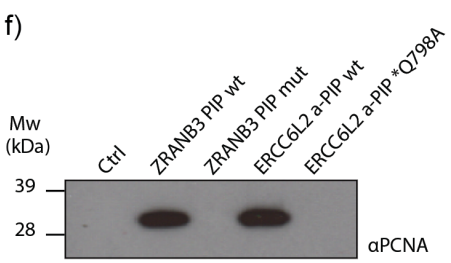

f)
Mw (kDa)
Ctrl
ZRANB3 PIP wt
ZRANB3 PIP mut
ERCC6L2 a-PIP wt
ERCC6L2 a-PIP *Q798A
39
28
αPCNA

Figure 5-2: Identifying ERCC6L2's PCNA binding site

a) U2OS cells expressing YFP-ERCC6L2 (701-1247) and smaller constructs with conserved regions removed, assessed for the formation of nuclear foci. Domain maps of ERCC6L2 constructs are shown above. b) Immunofluorescence against PCNA in U2OS cells expressing YFP-ERCC6L2 (797-1098). DAPI was used to counterstain DNA. c) Multiple sequence alignment of ERCC6L2 aa797-819. Completely conserved residues are highlighted in black. Residues mutated to alanine for subsequent validation experiments are denoted with red 'A's. d) WT YFP-ERCC6L2 (701-1247) and the point mutants indicated in c) assessed for their ability to form nuclear foci in unperturbed U2OS cells following transient transfection. The percentage of YFP-positive nuclei displaying YFP foci formation is indicated underneath the image for each construct. e) Western blot against Strep and PCNA of StrepTactin immunoprecipitations from 293T cells expressing WT TwinStrep-ERCC6L2 (701-1247) and the range of point mutants indicated in c). f) *In vitro* pulldown of PCNA as described in Figure 5-1a, using biotinylated peptides containing either the candidate WT atypical PIP-box of ERCC6L2 or the predicted binding-null Q798A mutant. The ZRANB3 WT and mutant PIP-box peptides from Figure 5-1d were used as positive and negative controls for PCNA binding, respectively. Scale bars = 10μm.

Taken together, these results indicated that we had likely located a novel PCNA binding site in ERCC6L2. In order to confirm whether or not this interaction was truly direct, I performed an *in vitro* pulldown experiment with biotinylated peptides as described in Figure 5-1a, again using WT and mutant PIP-box peptides from ZRANB3 as positive and negative controls, respectively. The biotinylated ERCC6L2 atypical PIP (a-PIP) interacted strongly with PCNA, while a peptide bearing a sequence change corresponding to the Q798A mutation failed to interact (Figure 5-2e). This demonstrates that this portion of ERCC6L2 binds PCNA directly.

Recently, work from our group demonstrated by co-crystallisation that peptides containing either of the two PCNA interaction sites of the SNF2 ATPase ZRANB3 – the PIP-box and a C-terminal APIM motif – both bind the same site on the inter-domain connecting loop of PCNA [208]. I next considered it worthwhile to determine whether or not the a-PIP of ERCC6L2 binds to PCNA in a similar manner to PIP-boxes or APIM motifs. Dr. Marek Šebesta from our group co-crystallised PCNA with a synthetic ERCC6L2 peptide containing the a-PIP identified in Figure 5-2. Crystal mounting and data collection on the resultant crystals were performed by Dr. Antonio Ariza (of the Ivan Ahel group at the Sir William Dunn School of Pathology), who then solved the co-

crystal structure to a resolution of 2.73Å. As might be expected from PIP-boxes and APIMs binding the same site on PCNA [208], the ERCC6L2 a-PIP bound PCNA at the same site as the PIP-box of ZRANB3 (Figure 5-3a). Notably, the side groups of conserved residues predicted from immunoprecipitation experiments (Figure 5-2c) to participate in the ERCC6L2-PCNA interaction (Q798, L799, L801, C804, F806, shown in orange) face towards the binding surface on PCNA, while a conserved residue that appears dispensable for the interaction (Q803, shown in cyan) points away from the interaction site (Figure 5-3).

It was shown by our group recently that the canonical PIP-box of ZRANB3 binds PCNA in a similar manner to the previously characterised interactions between PCNA and the PIP-boxes of FEN1 and p21. This is based on a number of hydrogen bonds between both side chain and main chain atoms of the PIP-boxes and PCNA [206, 208]. The main chain carbonyl group of G797 in ERCC6L2 forms a hydrogen bond with the main chain amine of I255 in PCNA (Figure 5-3c). The amine group on ERCC6L2's Q798 side chain forms a hydrogen bond with the main chain carbonyl of PCNA's A252, while ERCC6L2 L799 forms a hydrogen bond with the main chain carbonyl of PCNA's P253 (Figure 5-3c). Main chain hydrogen bonds also form between the amine of ERCC6L2 L801 and the carbonyl of PCNA H44, and the amine of ERCC6L2 S807 and carbonyl of PCNA G127 (Figure 5-3d). Beyond the hydrogen bonds formed between the ERCC6L2 a-PIP and PCNA, the residues from L801-C804 form a 3_{10} helix, which along with F806 form the 'hydrophobic plug' characteristic of PIP-box and APIM motif interactions with PCNA (Figure 5-3b/e) [206, 208]. The 3_{10} helix, as in other 3_{10} helices of PIP-boxes and APIM motifs [206, 208], appears to be stabilised by several side chain-main chain and main chain-main chain hydrogen bonds forming. The Q798 side chain amine forms an H-bond with the main chain carbonyl of L799, while T800 appears to play an important

stabilising role in this structure. T800 forms several structural H-bonds: from its main chain amine to the side chain carbonyl of Q803, from its side chain hydroxyl and its main chain carbonyl both to the main chain amine of Q803, and finally from its main chain carbonyl to the C804 side chain thiol. Within the 3_{10} helix, L801's main chain carbonyl forms H-bonds with the main chain amines of both C804 and F806, perhaps stabilising this side of the helix.

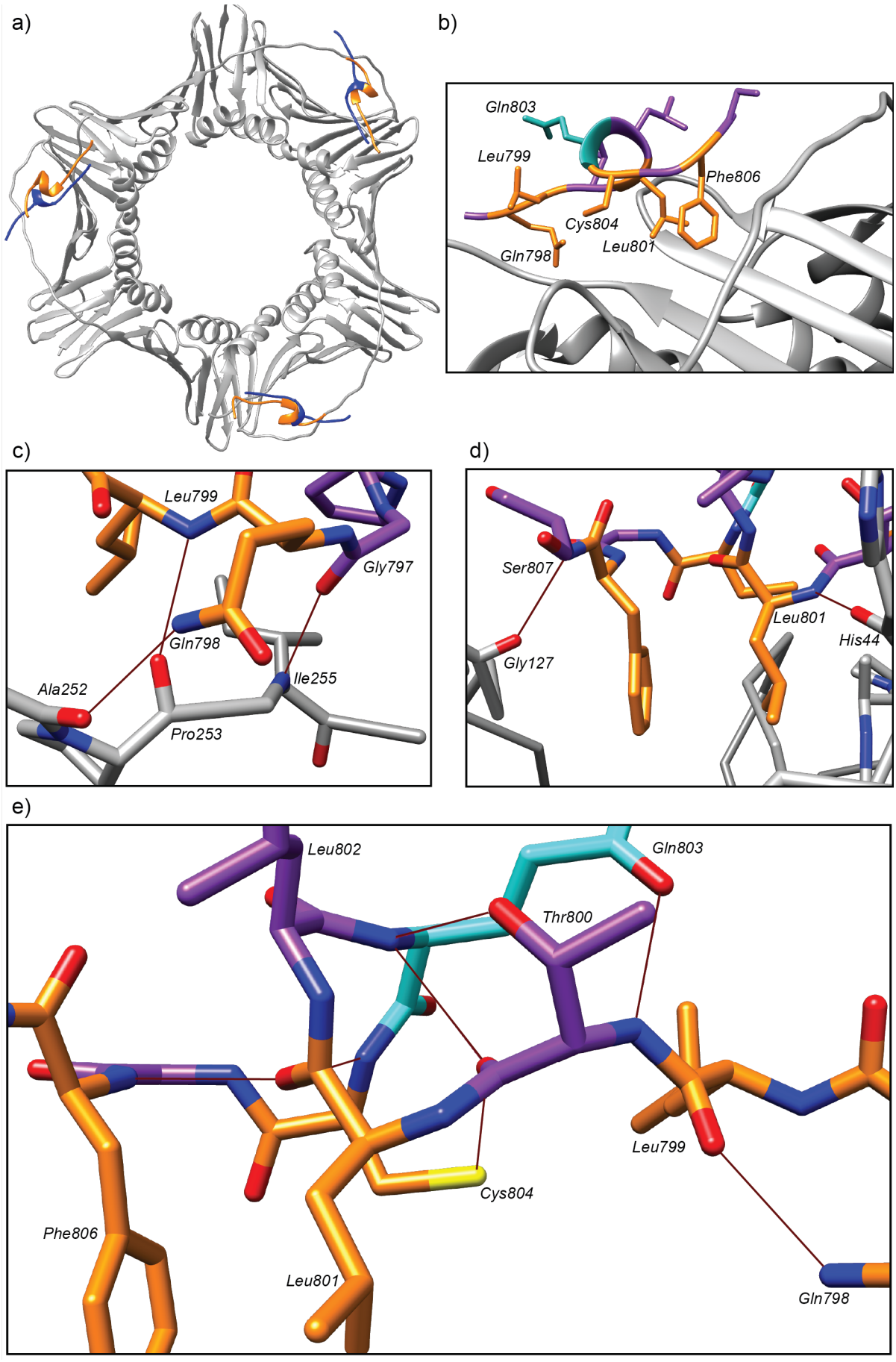


Figure 5-3: Co-crystal structure of PCNA with the ERCC6L2 atypical PIP-box

a) Co-crystal structure of PCNA (grey) with the ERCC6L2 a-PIP (orange) peptide, with the ZRANB3 PIP-box (blue) from a previously-published co-crystal structure overlaid [208]. b) Closer view of the interaction between the ERCC6L2 a-PIP and PCNA. ERCC6L2 residues predicted by previous experiments (Figure 5-2c) to be critical for the ERCC6L2-PCNA interaction are shown in orange, while the Q803 residue predicted to be of lesser importance is shown in cyan. Residues whose importance to the ERCC6L2-PCNA interaction has not yet been tested are shown in purple. c) Interaction analysis showing hydrogen bonds between G797, Q798 and L799 and PCNA's PIP-box binding surface at I255, A252 and P253, respectively. d) Interaction analysis showing hydrogen bonds between L801 and S807 of ERCC6L2 and H44 and G127 of PCNA, respectively. e) Intra-molecule analysis of the ERCC6L2 a-PIP showing hydrogen bonds between residues of the a-PIP centred around the 3_{10} helix. Hydrogen bonds are shown in dark red, nitrogen atoms in blue, oxygen atoms in bright red and sulphur in yellow. PCNA-ERCC6L2 a-PIP co-crystals were generated by Dr. Marek Šebesta, while data collection and structure solution was performed by Dr. Antonio Ariza.

Armed with knowledge of ERCC6L2's PCNA interaction site, we next aimed to explore the physiological role of this interaction by assessing its contribution to our observations on ERCC6L2 function in a range of the assays described so far. PCNA plays a critical role in mediating the recruitment of a wide range of factors to the replication fork during both normal replication and under conditions of DNA damage or replication stress [197]. Considering the fact that ERCC6L2 does not appear to colocalise with PCNA in unperturbed cells (Figure 4-4a), we reasoned that PCNA might be responsible for mediating ERCC6L2's recruitment to DNA damage sites. In U2OS cells, the YFP-(701-1247) fragment which interacts with PCNA (Figure 5-1) recruits strongly to sites of laser-induced DNA damage, while recruitment of the Q798A mutant fragment is considerably weaker (Figure 5-4a). This suggests that the recruitment of this fragment to DNA lesions is mediated by the PCNA interaction. I next tested the effect of the Q798A mutant on DNA damage recruitment of full-length ERCC6L2. Unexpectedly, ERCC6L2 Q798A was recruited to DNA damage sites with similar kinetics to the WT protein (Figure 5-4b), suggesting that redundant or cooperating mechanisms exist for ERCC6L2's DNA damage recruitment.

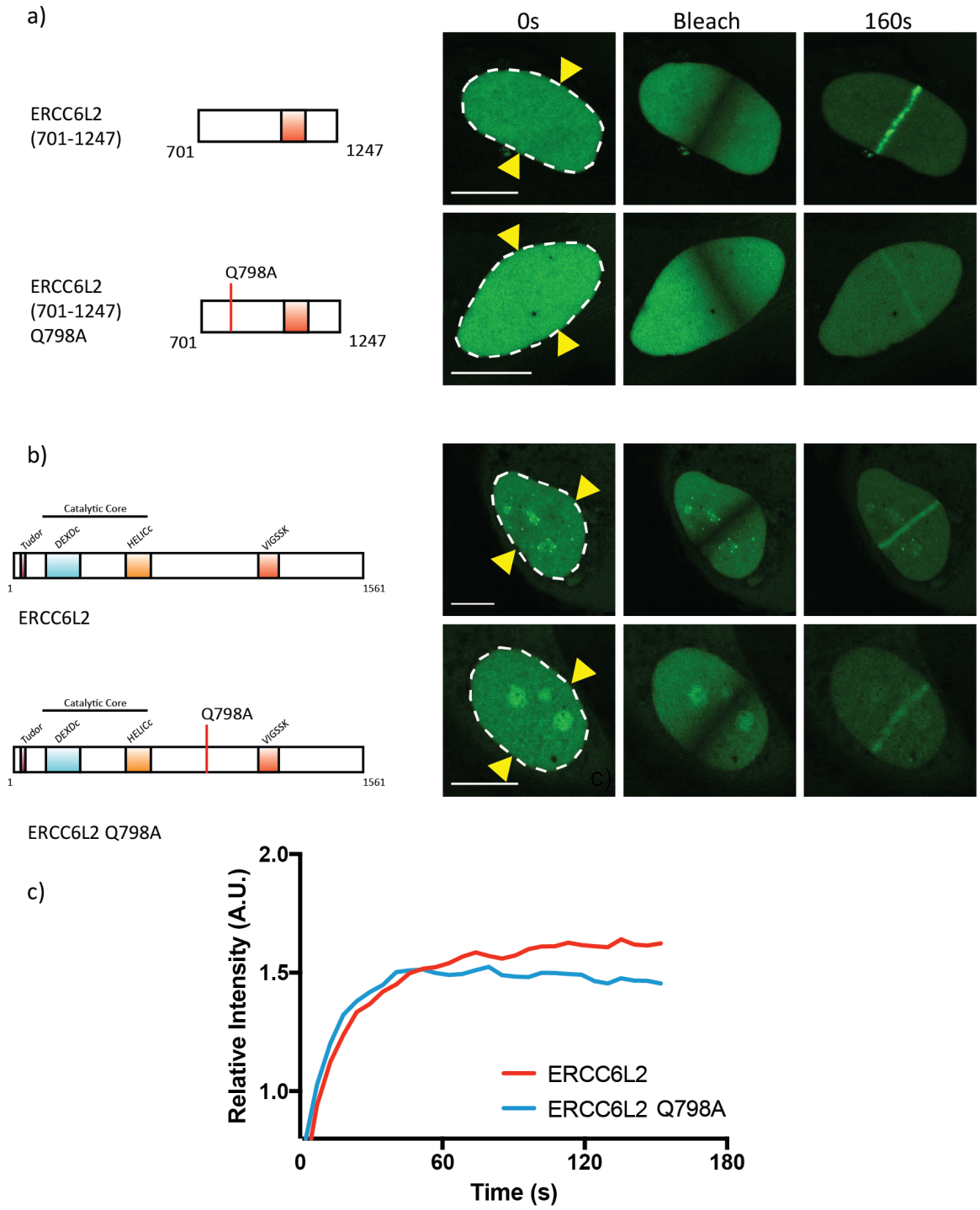


Figure 5-4: The ERCC6L2-PCNA interaction is dispensable for ERCC6L2 recruitment to DNA lesions

a) Laser microirradiation on U2OS cells expressing YFP-ERCC6L2 (701-1247) WT or Q798A. Domain maps of the ERCC6L2 constructs expressed are shown alongside. b) Laser microirradiation on U2OS cells expressing YFP-ERCC6L2 WT or Q798A. Yellow triangles indicate regions targeted by laser microirradiation and dashed white lines denote the periphery of YFP-positive nuclei. c) Quantification of YFP recruitment to DNA damage sites displayed in b) (n = 20 cells). Scale bars = 10μm.

Despite the Q798A mutation having little effect on full-length ERCC6L2's immediate recruitment to DNA damage sites (Figure 5-4b/c), it remains conceivable that the

ERCC6L2-PCNA interaction still has relevance to its function at these sites, given that YFP-(701-1247) is recruited to DNA damage sites in a manner dependent on its PCNA interaction (Figure 5-4a).

In parallel to the generation of complementation cell lines discussed in Section 3.1 (Figure 3-6), I established ERCC6L2 KO cells expressing C-terminally V5-tagged ERCC6L2 Q798A. Figure 5-5a confirms expression of ERCC6L2 Q798A-V5 alongside the previously shown overexpression in the other ERCC6L2 complementation cell lines (Figure 3-6). Next, I performed clonogenic survival assays with this cell line alongside those displayed previously (Figure 5-5b/c, a subset of data also shown in Figure 3-7). Under phleomycin treatment, Ctrl 1- and Ctrl 2-derived cell lines all behaved similarly. In contrast, while ERCC6L2$^{-/-}$-1A + K165R gave only a weak improvement in clonogenic potential compared to parental cells or empty vector control cells, ERCC6L2$^{-/-}$-1A + Q798A cells showed phleomycin tolerance comparable to that of ERCC6L2$^{-/-}$-1A + WT cells (Figure 5-5b). A similar effect is observed in the ERCC6L2$^{-/-}$-2A-derived cell lines. While ERCC6L2$^{-/-}$-2A and ERCC6L2$^{-/-}$-2A + EV cells are highly sensitive to phleomycin and ERCC6L2$^{-/-}$-2A + K165R cells display only marginally better tolerance, the ERCC6L2$^{-/-}$-2A + WT and + Q798A cell lines retain similar clonogenic potential to Ctrl 2-derived cell lines under phleomycin treatment (Figure 5-5c). ERCC6L2$^{-/-}$-1A + K165R cells did not tolerate etoposide appreciably better than ERCC6L2$^{-/-}$-1A or ERCC6L2$^{-/-}$-1A + EV. However, complementation of ERCC6L2$^{-/-}$-1A cells with ERCC6L2 Q798A restored clonogenic potential to similar levels as both ERCC6L2$^{-/-}$-1A + WT cells and the range of Ctrl 1-derived cell lines under etoposide treatment (Figure 5-5b). In contrast, the rescue of ERCC6L2$^{-/-}$-2A cells' clonogenic potential by ERCC6L2 Q798A under etoposide treatment was much weaker compared to ERCC6L2$^{-/-}$-2A or ERCC6L2$^{-/-}$-2A + EV cell lines, only rescuing any tolerance at the highest dose (Figure 5-5c).

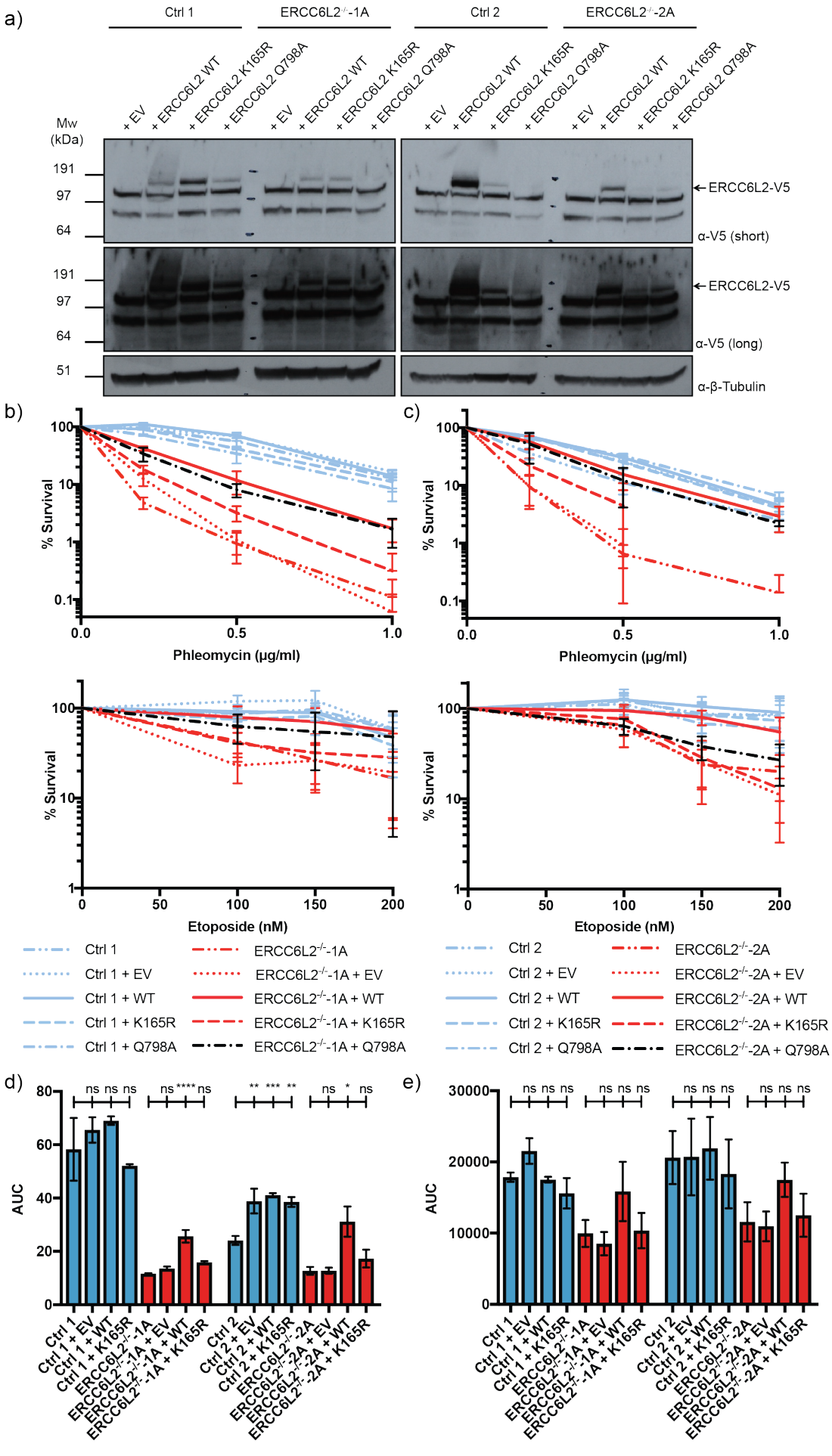

a)
Ctrl 1
ERCC6L2-/--1A
Ctrl 2
ERCC6L2-/--2A
+ EV
+ ERCC6L2 WT
+ ERCC6L2 K165R
+ ERCC6L2 Q798A
Mw (kDa)
191
97
64
51
ERCC6L2-V5
α-V5 (short)
α-V5 (long)
α-β-Tubulin
b)
c)
% Survival
Phleomycin (µg/ml)
Etoposide (nM)
Ctrl 1
Ctrl 1 + EV
Ctrl 1 + WT
Ctrl 1 + K165R
Ctrl 1 + Q798A
ERCC6L2-/--1A
ERCC6L2-/--1A + EV
ERCC6L2-/--1A + WT
ERCC6L2-/--1A + K165R
ERCC6L2-/--1A + Q798A
Ctrl 2
Ctrl 2 + EV
Ctrl 2 + WT
Ctrl 2 + K165R
Ctrl 2 + Q798A
ERCC6L2-/--2A
ERCC6L2-/--2A + EV
ERCC6L2-/--2A + WT
ERCC6L2-/--2A + K165R
ERCC6L2-/--2A + Q798A
d)
e)
AUC
ns ns ns ns **** ns
** *** ** ns * ns

Figure 5-5: The ERCC6L2-PCNA interaction is dispensable for etoposide and phleomycin tolerance

a) Western blot of complementation cell lines derived from Ctrl 1, Ctrl 2, ERCC6L2$^{-/-}$-1A and ERCC6L2$^{-/-}$-2A cell lines. Long and short exposures for the V5 Western blot are shown. β-tubulin was used as a loading control. b) Clonogenic survival assays on Ctrl 1- and ERCC6L2$^{-/-}$-1A-derived complementation cell lines from a) treated with phleomycin or etoposide. c) Clonogenic survival assays on Ctrl 2- and ERCC6L2$^{-/-}$-2A-derived complementation cell lines from a) treated with phleomycin or etoposide. d) AUC calculations on phleomycin clonogenics from b) and c) with one-way ANOVA performed. e) AUC calculations on etoposide clonogenics from b) and c) with one-way ANOVA performed. Error bars ± SEM. $p \leq 0.05$ = *, $p \leq 0.01$ = **, $p \leq 0.001$ = ***, $p \leq 0.0001$ = ****.

Considering the importance of PCNA in DNA replication [197], we next sought to determine whether or not the ERCC6L2-PCNA interaction influences DNA synthesis under etoposide treatment. While expression of WT ERCC6L2 in ERCC6L2$^{-/-}$-2A cells rescues DNA synthesis under etoposide treatment as shown previously (Figure 4-2), data collected by Dr. Dragana Ahel show that ERCC6L2 Q798A is only capable of supporting a partial rescue, possibly suggesting that the Q798A mutant has compromised ERCC6L2 function in this context (Figure 5-6). However, variation of expression levels between WT and mutant ERCC6L2 (Figure 5-5a) prevents any firm conclusion. However, expression of ERCC6L2 Q798A in ERCC6L2$^{-/-}$-1A cells is barely visible in comparison to WT and K165 expression (Figure 5-5a), yet this expression level is sufficient to rescue cells' tolerance of both phleomycin and etoposide in clonogenic survival assays to the same extent as WT ERCC6L2 (Figure 5-5b). It should be noted that while area under the curve (AUC) differences between parental ERCC6L2$^{-/-}$ cell lines and those expressing WT ERCC6L2 are statistically significant (as calculated by one-way ANOVA) in phleomycin experiments (Figure 5-5d), this difference is not statistically significant in etoposide experiments (Figure 5-5e). While the biological result is clear, this discrepancy at the quantification level could have arisen from the fact that in comparison to parental WT control cell lines, parental ERCC6L2$^{-/-}$ cell lines are less sensitive to etoposide than to phleomycin in fold-difference terms in these experiments (Figure 5-5).

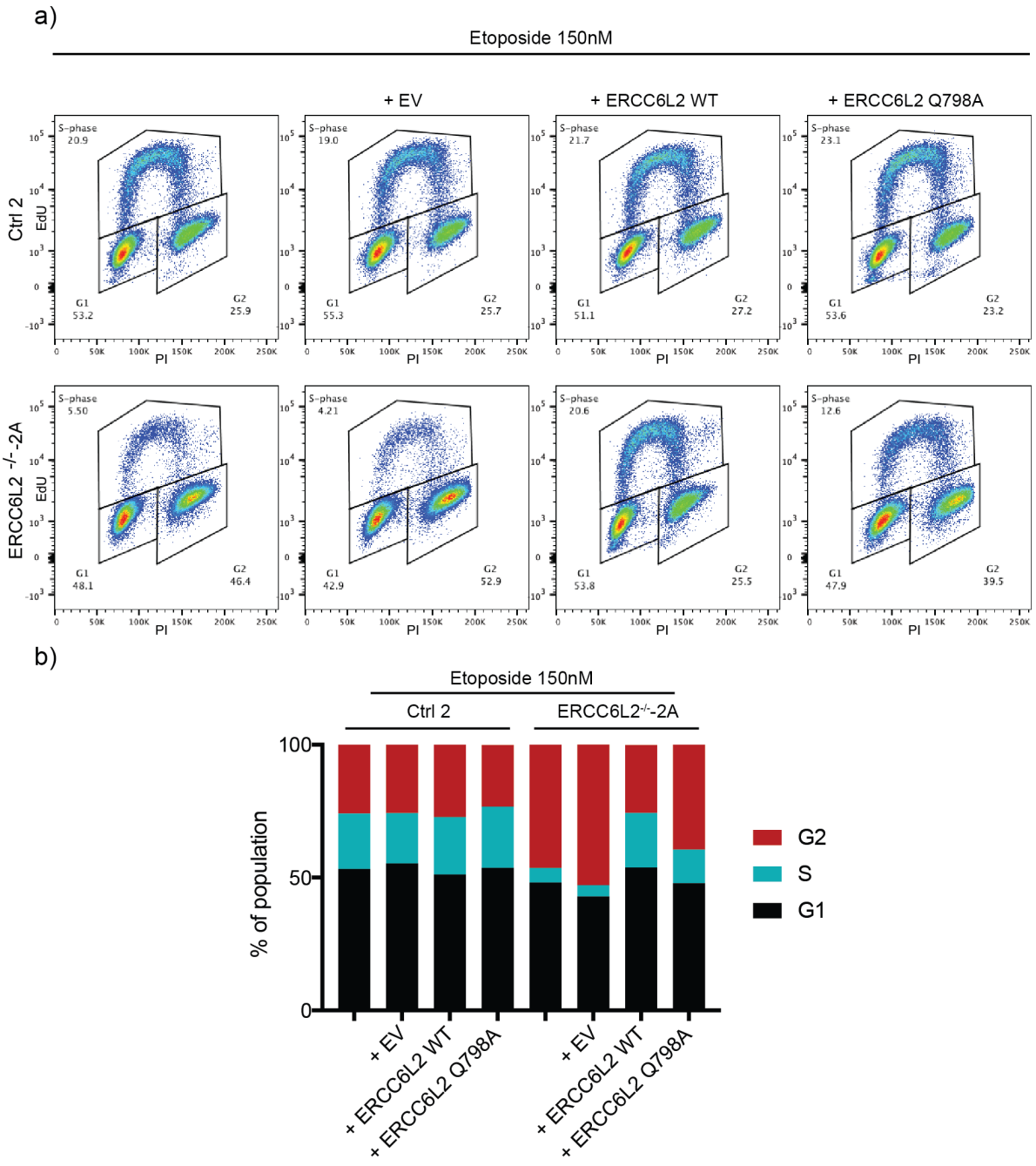


Figure 5-6: ERCC6L2 Q798A partially restores DNA synthesis in ERCC6L2-deficient cells treated with etoposide

a) Ctrl 2- and ERCC6L2$^{-/-}$-2A-derived complementation cell lines from Figure 5-5a treated with 150nM etoposide for 48h and subjected to cell cycle analysis by flow cytometry as in Figure 4-1b. b) Graphical representation of cell cycle distributions quantified in a). Raw Data were collected by Dr. Dragana Ahel.

Taken together, these results suggest that ERCC6L2's PCNA interaction is less important than ERCCL2's ATPase activity for its response to etoposide and phleomycin.

5.2 Discussion

In this Chapter, I set out to explore the relationship between ERCC6L2 and its interaction partner PCNA. PCNA has a critical role at the replication fork, mediating the recruitment of a wide range of factors to both unperturbed and challenged replication forks, as well as in gap-filling DNA synthesis following DNA repair pathways [197, 370, 371]. Given the importance of PCNA in this range of contexts related to genome stability and the localisation of ERCC6L2 to centromeres during their replication, we reasoned that ERCC6L2's interaction with PCNA might be of central importance to its cellular function. Briefly, I identified ERCC6L2's atypical PCNA binding site and confirmed that a peptide bearing this site binds PCNA directly *in vitro*. Furthermore, I have found that the ERCC6L2 a-PIP binds to PCNA in a similar manner to PIP-boxes and APIM motifs. Finally, I found that disrupting the ERCC6L2's PCNA interaction on its own does not cause severe cellular phenotypes in the context of previous experiments. This finding suggests that other as-yet uncharacterised elements of ERCC6L2 may provide a degree of redundancy or cooperativity with the PCNA interaction.

To search for a likely ERCC6L2-PCNA binding site, I analysed ERCC6L2's primary amino acid sequence and identified a conserved site matching the consensus sequence of a PCNA-binding motif, known as a PIP-box. However, a peptide from ERCC6L2 spanning this site failed to interact with recombinant PCNA *in vitro* (Figure 5-1d). Since ERCC6L2 contains no other sequence matching a canonical PIP-box, this result indicated that ERCC6L2's PCNA interaction occurs through a non-canonical interaction site.

To explore the C-terminus of ERCC6L2 further, I expressed a fragment of the C-terminus not containing the PIP-box-like sequence in U2OS cells and noticed that it formed PCNA-like nuclear foci in around 20% of cells (Figure 5-1e). Considering this a

likely indication that this fragment interacts with PCNA, I performed immunoprecipitation from HEK 293T cells, confirming this interaction (Figure 5-1f). This indicated to us that ERCC6L2's interaction with PCNA does indeed occur through a novel type of PCNA interaction site. As such, I next aimed to make use of these observations to narrow down the precise PCNA interaction site in ERCC6L2.

Using the formation of nuclear foci as a readout for likely PCNA interaction, I expressed various truncated ERCC6L2 fragments lacking different patches of conserved residues. This approach led me to identify a new candidate PCNA interaction site not dissimilar to a canonical PIP-box. This atypical PIP-box features different spacing from canonical PIP-boxes and has only one aromatic residue compared to the characteristic pair (Figure 5-2b). Mutation of some conserved residues within this site abolished the interaction of the 701-1247 ERCC6L2 fragment with PCNA, as well as its propensity to form nuclear foci (Figure 5-2c/d).

I next sought to confirm that this site was responsible for a direct interaction with PCNA. A biotinylated ERCC6L2 peptide bearing this site interacted directly with recombinant PCNA *in vitro*, while the mutation of a glutamine residue in this peptide corresponding to position 798 in full-length ERCC6L2 abolished this interaction (Figure 5-2e). This experiment confirmed that ERCC6L2 interacts directly with PCNA through this site, and gave us an insight into a possible similarity between this a-PIP and canonical PIP-boxes, which bear a glutamine residue at their C-terminal ends. In order to gain a better understanding of this interaction at the molecular level and compare it with PCNA-PIP-box interactions, we next aimed to co-crystallise PCNA with the a-PIP identified.

Generation of co-crystals from purified PCNA and a peptide containing the a-PIP and subsequent data collection and structure solution were performed by Dr. Marek

Šebesta (from our group) and Dr. Antonio Ariza (from the Ivan Ahel group, Sir William Dunn School of Pathology, University of Oxford), respectively. Analysis of this structure revealed that the ERCC6L2 a-PIP binds the same surface on PCNA as both the canonical PIP-box of ZRANB3 and the ZRANB3 APIM motif, another atypical PCNA interaction motif (Figure 5-3a) [208]. Notably, hydrogen bonds important for the interaction appear to be conserved between the ERCC6L2 a-PIP and ZRANB3's PIP-box. These bonds include ERCC6L2-PCNA contacts, as is the case between ERCC6L2 Q798 and PCNA A252, as well as H-bonds internal to the ERCC6L2 peptide between amino acid side chains and the main chain, which likely stabilise the 3_{10} helix necessary to form a hydrophobic plug (Figure 5-3). This hydrophobic plug is an important feature of PCNA interaction sites, as noted previously in canonical PIP-boxes and APIMs [206, 208].

These findings suggest that although these motifs differ in primary sequence, it is the three dimensional structure of the PCNA interaction motif which plays the critical role in protein interactions with PCNA. With these insights, it might be possible to predict novel PCNA interaction partners and thus identify new players in DNA replication and repair. Indeed, a total of 28 of human proteins contain peptides matching the candidate consensus sequence Q-[VILM]-x-[VILM]-x-x-C-x-[FY], while 136 proteins contain sequences matching a motif of Q-x x-[VILM]-x-x-C-x-[FY], as identified by use of the ExPASy ScanProsite webtool. Of these include multiple candidate PCNA interactors which may be of interest to follow-up studies, including BAZ2A, MCMDC2, MSH2 and DNA polymerase delta subunit 2.

Overall, these results demonstrate the identification of a novel type of atypical PCNA interaction motif based on primary amino sequence and a characterisation of the mechanistic basis of this interaction. ERCC6L2's PCNA interaction site (a-PIP) interacts with the same site on PCNA as canonical PIP-boxes, suggesting that the competition for

binding sites on a PCNA homotrimer is shared between ERCC6L2 and other canonical PIP- and APIM-containing proteins. This in turn implies that layers of as-yet unidentified regulation may mediate the binding of ERCC6L2 to PCNA and out-competition of other factors. Characterising the regulatory mechanism(s) governing ERCC6L2's interaction with PCNA and the functional insights this provides may be crucial in fully understanding the basis of ERCC6L2's cellular role and pathologies resulting from its deficiency.

The next raft of experiments aimed to explore the role of ERCC6L2's PCNA interaction in the range of cellular observations we had made about ERCC6L2 function. The approach consisted of abrogating ERCC6L2's PCNA interaction and examining the effect of this on ERCC6L2's recruitment to DNA lesions as well as its ability to rescue DNA damage sensitivity phenotypes in ERCC6L2$^{-/-}$ cells.

ERCC6L2 (701-1247), which interacts with PCNA and forms PCNA-colocalising foci in untreated cells, was not recruited to DNA lesions after mutation of the Q798 residue (Figure 5-4a). This suggested that the PCNA interaction is important for the recruitment of this portion of ERCC6L2 to DNA lesions. However, when I assessed the ability of full-length ERCC6L2 bearing the Q798A mutation to mobilise to DNA damage sites, I noticed that it was still recruited rapidly to these sites (Figure 5-4b). These results did however suggest that ERCC6L2 Q798A might be slightly less stable at DNA damage sites over a longer time period than the one tested. Although this is a preliminary observation, further experiments to determine the time of ERCC6L2 retention at DNA lesions compared to the PCNA-binding mutant would be worthwhile. Together, these results suggest that other interactions may provide a layer of redundancy or cooperation for ERCC6L2 recruitment. Interaction studies on other

regions of interest within ERCC6L2, such as the Tudor domain and the VIGSSK domain, might reveal new interaction partners and help explain this.

At a phenotype level, I first expressed ERCC6L2 Q798A-V5 in ERCC6L2$^{-/-}$ cell lines (Figure 5-5a), as described previously (Figure 3-6). I then aimed to assess the ability of this mutant protein in comparison to WT ERCC6L2 to rescue the etoposide and phleomycin sensitivities of knockout cells, along with the rescue of DNA synthesis under etoposide treatment.

Mutating ERCC6L2's a-PIP did not abrogate clonogenic potential in complemented ERCC6L2$^{-/-}$ cells in most circumstances, although ERCC6L2$^{-/-}$-2A cells complemented with the Q798A mutant did not tolerate etoposide appreciably better than the ATPase-dead K165R mutant (Figure 5-5). Furthermore, ERCC6L2$^{-/-}$-2A cells expressing ERCC6L2 Q798A did not display a full rescue of DNA synthesis under etoposide treatment, unlike ERCC6L2$^{-/-}$-2A cells expressing WT ERCC6L2 (Figure 5-6). Differences in expression levels in these complemented cell lines still complicate the picture (Figure 5-5), but it may be that although the PCNA interaction may support and contribute to ERCC6L2's cellular role, other features of ERCC6L2 are sufficient to support some function in our experimental conditions.

Overall, we have gained valuable insights into the mechanistic basis of the interaction between ERCC6L2 and PCNA, demonstrating that this occurs through a new type of atypical PIP-box. This a-PIP binds to PCNA at the same site as canonical PIP-boxes and APIM motifs, and our co-crystal structure underlines the point that although these PCNA interaction motifs differ considerably in spacing and sequence, it is the three-dimensional structure they form which determines their binding to PCNA. Thus far we have been unable to observe dramatic phenotypes associated with the loss of ERCC6L2's PCNA interaction, and it seems likely that there may be a degree of

redundancy in this respect between PCNA and other aspects of ERCC6L2 regulation. Given the importance of PCNA in a range of contexts at the replication fork, it will be important for future studies to determine the contribution of the PCNA interaction to ERCC6L2's cellular function. Such studies should note the recent publication of a candidate ERCC6L2 interactome, which while not yet validated, contains a range of proteins relating to DNA replication [336].

6 Discussion

Through the course of this project we have made a number of observations which may be important in understanding the cellular function of ERCC6L2. These include the sensitivity of ERCC6L2-deficient cells to genotoxic drugs, a DNA replication defect in these cells in response to etoposide treatment, the recruitment of ERCC6L2 to laser-induced DNA lesions and centromeres, and the interaction of ERCC6L2 with the DNA replication scaffold PCNA. Notably, ERCC6L2's PCNA interaction occurs through a novel class of PCNA binding motif. Importantly, ERCC6L2's ATPase activity appears to be important for its role in the DDR, while its C-terminus mediates its centromeric localisation and its recruitment to sites of laser-induced DNA damage.

A range of clinical mutants of ERCC6L2 have been reported, either resulting in severe truncations from the C-terminus sometimes including a probable loss of ATPase activity, or resulting in amino acid changes in the ATPase domain which likely render it catalytically inactive (Table 1-1) [332-337]. Notably, all of the mutant protein products tested appear defective in recruitment to DNA damage sites, localisation to centromeres, and often localisation in the nucleus. Mutations causing truncated, catalytically dead ERCC6L2 are likely to render patients' cells sensitive to genotoxic stress, as suggested by our data showing that a loss of ERCC6L2's ATPase activity is sufficient to cause cells' sensitivity to phleomycin and etoposide (Figure 3-7). However, further insights can be gained from considering the other mutants tested. While ERCC6L2 E729fs does possess an intact catalytic core, like the larger truncations it is mislocalised to the cytoplasm (Figure 3-3, Figure 4-7). This would suggest that C-terminal to E729 is a nuclear localisation sequence (NLS) which supports ERCC6L2's nuclear localisation, and that the lack of nuclear ERCC6L2 can contribute to clinical phenotypes in ERCC6L2-mutant patients. The lack of nuclear localisation of this mutant

can explain its deficiencies in recruitment to both DNA lesions (Figure 3-3) and centromeres (Figure 4-7). In contrast, ERCC6L2 R1266* is – like the WT protein – localised to the nucleus, suggesting that ERCC6L2's NLS lies in the region between E729 and R1266. Notably, ERCC6L2 R1266* fails to mobilise to DNA lesions and to centromeres (Figure 3-3, Figure 4-7). This would suggest that the C-terminus of ERCC6L2 is important for its localisation to these sites. As such, it could be expected that the host of new mutants identified since the time of these experiments [336] are also defective in centromere and DNA damage localisation. Testing their expression might help narrow down the NLS responsible for ERCC6L2's nuclear expression pattern.

Overall, our findings relating to ERCC6L2 function in the context of genome stability might contribute to the basis of ERCC6L2-deficient pathologies. It has been proposed that ERCC6L2 participates in the response to DSBs based on phenotypic evidence, however more direct functional assays suggest that this is not the case [337]. Additionally, Tummala et al (2018) propose a role for ERCC6L2 in transcription-coupled nucleotide repair [336]. However, this is unlikely to be the case, since I have preliminary but convincing data showing that ERCC6L2$^{-/-}$ cells are not sensitive to UV, which would cause a range lesions requiring repair by TC-NER.

When our data are taken into consideration another picture emerges, in which several lines of evidence point to a role relating to DNA replication and suggest there might be some context-specific aspect of this function based on genomic location. Below, I will comment on some of the possible interpretations of the data available to us, from both this project and previous studies.

6.1 ERCC6L2 as a DSB repair factor

6.1.1 ERCC6L2 deficiency phenotypes

Human patients with possible function-compromising mutations in *ERCC6L2* suffer a bone marrow failure (BMF) disorder distinct from Fanconi anaemia, a common form of BMF characterised by genomic instability and cancer predisposition [338, 365, 372, 373]. Proteins mutated in FA are considered to be members of the FA pathway, which is important for the repair of DNA interstrand crosslinks. In addition to their role in ICL repair, a number of FA proteins are heavily involved in other aspects of the DDR, including DSB repair in the case of BRCA1, BRCA2, RAD51 and PALB2 [374]. A whole exome sequencing study on a cohort of BMF patients identified *ERCC6L2* mutations as the probable genetic cause in 7 of these patients out of 86 (8.1%) [333]. Interestingly, while this study also identified a number of disease-causing mutations in genes unrelated to DNA repair, several other DNA repair genes were identified with probable disease-causing mutations, including LIG4, RUNX1, RTEL1 and ATR [333]. Together, these clinical observations point towards a possible role for ERCC6L2 in maintaining genomic stability. Although the precise function for ERCC6L2 in genome stability is still unclear, it has been proposed that ERCC6L2 functions in the response to double strand breaks [337].

Based on several lines of phenotypic evidence, it is tempting to suggest a role for ERCC6L2 in the response to DSBs. Zhang and colleagues reported that fibroblasts derived from an ERCC6L2-deficient patient are sensitive to ionising radiation, as well as the DSB-inducing drug phleomycin [337]. This finding was further supported by a later publication from an independent group using cell lines derived from three other ERCC6L2-deficient patients [336, 337]. Independently, we have found that

CRISPR/Cas9-generated U2OS cells harbouring severe truncating mutations in ERCC6L2 are sensitive to phleomycin (Figure 3-5b). Additionally, these ERCC6L2 knockout U2OS cells are sensitive to etoposide (Figure 3-5b), a TOPO2 inhibitor that exposes DSBs following replication or transcription runoff of chromatin-trapped TOPO2 [350, 375, 376].

Consistent with an influence on cell survival and proliferation under etoposide treatment, our experiments demonstrate a clear reduction of DNA synthesis in ERCC6L2$^{-/-}$ cells treated for 48h with a low dose (150nM) of etoposide, indicating impaired DNA replication under these conditions (Figure 4-1b). Importantly, ERCC6L2's catalytic activity appears to be critical for its role in the response to etoposide, as noted in both clonogenic survival assays (Figure 3-7) and in cell cycle analysis (Figure 4-2).

Previous studies suggest that while high doses of etoposide are sufficient to cause transcription-mediated DSB exposure, low doses (<10μM) cause DSB exposure only through replication runoff and not through transcription [350, 375, 376]. As such, it is tempting to suggest that under our etoposide treatment conditions, it is the replication-dependent exposure of DSBs which explains both the loss of ERCC6L2-deficient cells actively undergoing DNA synthesis, and the sensitivity of these cells to prolonged treatment (Figure 3-5b, Figure 4-1b). However, studies relating etoposide dose to mode of toxicity have thus far used varying etoposide doses at short time points [350, 375, 376]. In comparison to these studies, we have used very low doses of etoposide but for more extended periods. A detailed assessment of G1/S, intra-S and G2/M checkpoint status, as well as addressing precisely when in the cell cycle ERCC6L2 KO cells die due to etoposide treatment, might reveal useful insights into ERCC6L2's influence on etoposide toxicity. An important extension to these experiments will be to confirm

whether or not ERCC6L2-deficient cells are sensitive to the TOPO1 inhibitor camptothecin; this agent induces replication-coupled seDSBs that require function HR for their repair, and as such HR-deficient cells display camptothecin hypersensitivity.

In addition to sensitivity phenotypes towards DSB-inducing agents, previous studies have reported that ERCC6L2-deficient cells are sensitive to mitomycin C and irofulven [332, 336, 337], although our knockout cell lines exhibit only very mild sensitivity to MMC at high doses (Figure 3-5a). While classically associated with ICL repair, MMC can have other genotoxic effects, and DSBs have been shown to arise during normal MMC repair [377]. Irofulven, however, is a DNA and protein alkylating agent suggested to require transcription-coupled NER for repair of its lesions since it induces stalling of RNA pol II at active transcription sites [378-380]. This has led Tummala and colleagues to propose a role for ERCC6L2 in TC-NER [336]. However, irofulven treatment has also been shown to induce DSB formation [381-383], leaving the possibility that some aspect of DSB repair is defective in ERCC6L2-deficient cells and that this defect is behind the sensitivity of these cells to irofulven, or even that some other form of genotoxic stress such as replication stress could be occurring under irofulven treatment. This point is not addressed by Tummala et al [336]. At this stage, given incomplete understanding of irofulven's precise mechanism(s) of action and a lack of mechanistic insights into ERCC6L2's function, there is little support for a role for ERCC6L2 in TC-NER.

While cells deficient in ERCC6L2 appear to be sensitive to the induction of DSBs, the picture becomes more complex when considering the range of cellular observations relating to ERCC6L2 that have been made, which raise doubts for a role in DSB repair.

ERCC6L2 does not appear to function in either of the major DSB repair pathways: nonhomologous end-joining and homologous recombination [337]. Zhang et al found

no defect in NHEJ in fibroblasts from an *ERCC6L2*-mutant patient and no defect in HR following ERCC6L2 knockdown, a finding further supported by the absence of sensitivity of patient-derived cells to PARP inhibition [337]. However, it should be noted that the assays used by Zhang et al to assess NHEJ and HR efficiency utilise restriction enzyme activity to introduce DSBs in reporter plasmids; as such these lesions are not necessarily identical to physiological DSBs.

Tummala and co-workers noted an interaction between GFP-tagged ERCC6L2 and DNA-PK, and suggested that RNA/DNA hybrids increase in ERCC6L2-deficient cells [336]. These findings led to the suggestion that ERCC6L2 cooperates with DNA-PK (which has a long-established role in transcriptional regulation [384-386]) and transcription machinery to resolve R-loops, partly due to a high number of RNA-binding proteins appearing in an immunoprecipitation/mass spectrometry (IP/MS) experiment aimed at identifying ERCC6L2 interactors [336]. However, this dataset has not yet been validated and it should be noted that RNA-binding proteins are common contaminants in IP/MS experiments. Tummala et al do demonstrate defective recovery of RNA synthesis in ERCC6L2-deficient cells following irofulven treatment [336]. A role for ERCC6L2 which affects transcription rates is an interesting possibility which merits further investigation, particularly in the context of rDNA considering the nucleolar localisation of several C-terminal ERCC6L2 fragments (Figure 4-5). Thorough analysis by IP/MS of ERCC6L2's interactions under both unperturbed and DNA-damaging conditions might shed new light on ERCC6L2's function under treatments such as etoposide and phleomycin.

When my data are taken into consideration, further open questions present themselves. Firstly, if ERCC6L2 is a DSB repair factor, why might it be desirable to a cell for it to localise to centromeres in unperturbed cells? Secondly, how might ERCC6L2's

interaction with PCNA come into play in the context of a DSB? Below, I will consider the possible relevance of these observations in relation to DSB repair.

6.1.2 ERCC6L2: a DSB repair factor at centromeres?

The localisation of ERCC6L2 to centromeres in unperturbed interphase cells is an intriguing observation (Figure 4-5). Centromeres constitute a unique environment for DNA replication and repair, with highly repetitive sequences packaged into dense heterochromatin specially designated as centromeric by the deposition of the histone H3 variant CENP-A into nucleosomes [291-293].

Interestingly, mutation rates in heterochromatin are higher than in euchromatin, implying that chromatin state may affect either the kinetics of certain DNA repair pathways or influence the choice between such pathways [236, 237]. It has been proposed that relaxation of heterochromatin is an important prerequisite for efficient DNA repair [387, 388]. However, it appears that there is some benefit in retaining heterochromatin and limiting the activity of the homologous recombination – thought to be an error-free process – between repetitive sequences [389]. This restriction of HR at repetitive sequences in heterochromatin might suppress erroneous recombination between such sequences could lead to gross chromosomal rearrangements (GCRs). As such, the use of lower-fidelity alternative DSB repair pathways such as alternative end-joining (a-EJ) or single-strand annealling (SSA) to prevent GCRs could explain the relatively higher mutation rates in heterochromatin [390, 391].

A host of new insights into DNA repair in heterochromatic regions have been gained in recent years. DSBs' relocation to the periphery of dense heterochromatic domains has been observed in *Drosophila* and mammalian cells, and it has been proposed that this process constitutes a layer of regulation to provide repetitive damaged sequences

with spatial separation from other repeats of the same sequence to prevent illegitimate recombination events [392, 393]. This relocation appears to require resection and ATR activity, but recruitment of the RAD51 recombinase necessary for strand invasion can only occur after relocation. Occlusion of RAD51 from the cores of dense heterochromatin domains might thus provide additional suppression of potentially deleterious recombination between repetitive sequences [392].

More recently, stark differences between DSB repair mechanisms have been noted between centromeric and its surrounding pericenctromeric heterochromatin [238]. Using an elegant approach based on targeting the Cas9 nuclease to either centromeric or pericentromeric sequences in mouse NIH 3T3 cells, Tsouroula and co-workers induced DSBs specifically at centromeric or pericentromeric DNA. They found that while the upstream DSB end binding factor KU80 was recruited to both regions irrespective of cell cycle stage, RAD51 was recruited only at the periphery of heterochromatin after relocation. Notably, RAD51 was recruited to centromeric DSBs at the periphery of heterochromatin irrespective of cell cycle stage, although this increased in G2. In contrast, RAD51 recruitment to pericentromeric DSBs only occurred at post-replicative chromatin after relocation [238]. It therefore seems likely that the regulation of DSB repair at centromeres bears distinctions from other genomic loci, and possible that ERCC6L2 might function in the DDR at these sites. Indeed, a DSB repair role at specific sites might explain why ERCC6L2-deficient cells exhibited no overt defect in HR or NHEJ [337].

While the mechanisms governing the differential regulation between centromeres and even neighbouring pericentromeric heterochromatin are not understood, it could be that they constitute an interesting new avenue of research. How ERCC6L2 might fit into this picture is currently unclear. Given that several lines of evidence from multiple

groups have established ERCC6L2 as a factor impacting genome stability, investigating ERCC6L2's role specifically in the DDR at centromeres could provide exciting new insights into genome maintenance.

Given ERCC6L2's localisation to centromeres throughout interphase (Figure 4-5, Figure 4-6), it seems likely that ERCC6L2's function is important at centromeres even in untreated cells. It is possible that treatment with etoposide and phleomycin poses particular challenges to these structures and that the loss of ERCC6L2 sensitises particular genomic regions to these drugs. Establishing cell lines similar to those used by Tsouroula et al [238] in which we can induce centromeric DSBs in ERCC6L2 proficient and deficient backgrounds, could allow us to examine the role of ERCC6L2 in the molecular events of DSB repair at these sites.

Aside from double strand breaks, it might be expected that if ERCC6L2 has a role in DNA repair which is particularly important at centromeres, then a loss of centromere integrity could be observed in ERCC6L2$^{-/-}$ cells treated with genotoxic agents. This hypothesis could be tested by using pulsed-field gel electrophoresis (PFGE) followed by Southern blotting against centromeric sequences from WT or ERCC6L2-deficient cells before and after treatment, an approach which has been reported previously [394].

6.1.3 The ERCC6L2-PCNA interaction: relevant to the DSB response?

As a critical scaffold functioning at the replication fork and during DNA repair, the sliding clamp PCNA is critical for DNA replication and genome stability [197]. For the first time, we have demonstrated that ERCC6L2 interacts with PCNA (Figure 4-4b). Moreover, this occurs through a novel interaction motif, potentially opening the door for the identification of new PCNA interactors (Figure 5-2b). Given reports of ERCC6L2-deficient cells' sensitivity to DSB-inducing agents [332, 336, 337] in addition to similar

findings from my own experiments (Figure 3-5b), it might seem possible that the ERCC6L2-PCNA interaction has relevance in the DSB response.

PCNA's role in DSB repair is best characterised in terms of events relating to homologous recombination. PCNA appears to perform an important function in the latter stages of HR, functioning in DNA polymerase δ (pol δ)-dependent DNA synthesis that follows strand invasion [395, 396]. Furthermore, it has been suggested that PCNA also acts to stimulate the extension of D-loop intermediates in HR, in a manner dependent on its interaction partner pol δ [397]. Interestingly, although pol δ is considered the canonical DNA polymerase functioning in HR, *in vitro* reconstitution of D-loop extension with purified proteins demonstrates a degree of redundancy between polymerases; in addition to pol δ, the TLS polymerases κ and η can perform D-loop extension [398]. It has been suggested that PCNA might recruit different DNA polymerases to HR intermediates under certain circumstances and thus influence the outcome of HR [398]. This is a concept with which strong parallels can be drawn to the choice between DNA damage tolerance pathways in the replication stress response.

Aside from its role in HR, there exist associations – albeit poorly understood ones – between PCNA and other DSB repair pathways. An ionising radiation-stimulated interaction between PCNA and the KU70/80 heterodimer has been reported [399], and a possible role for PCNA in a minor end joining-related pathway has been suggested [370]. Interestingly, DNA ligase I, which may perform a function in a-EJ, interacts with PCNA in normal DNA replication and this interaction is important for joining Okazaki fragments produced by lagging strand replication [400]. Additionally, an interaction between *S. cerevisiae* pol32 (DNA pol δ subunit 3) and PCNA appears to be important for break-induced replication (BIR) in the repair of single-ended DSBs [401].

ERCC6L2 is still recruited to sites of laser-induced DNA damage when its interaction with PCNA is abolished by the Q798A mutation (Figure 5-4b). Given that PCNA interactions are often critical for recruiting different proteins to replication forks both in unperturbed conditions and under DNA damage [197], ERCC6L2 recruitment to DNA damage sites might occur through more than one mechanism. On a phenotypic level, expression of ERCC6L2 Q798A provides at least a partial rescue of ERCC6L2 KO cells' tolerance to both etoposide and phleomycin (Figure 5-5b/c), as well as partially (but not completely) rescuing DNA synthesis in ERCC6L2$^{-/-}$-2A cells treated with etoposide (Figure 5-6). Therefore, it seems plausible that the ERCC6L2-PCNA interaction plays a role in the response to phleomycin or etoposide which is supported by functional interactions between ERCC6L2 of other aspects of the DDR. A comprehensive analysis of the ERCC6L2 interactome under genotoxic conditions might reveal how ERCC6L2 is recruited to and stabilised at damage sites and give us a better understanding of the precise context in which ERCC6L2 functions.

Overall, considering that the strongest associations between the DSB response and PCNA relate to DNA replication in the context of repair intermediates, the ERCC6L2-PCNA interaction might point to a role for ERCC6L2 in the manipulation of these structures. Such a role could facilitate DNA synthesis at these difficult sites and could fit with ERCC6L2's presence at centromeres, given the tendency of repetitive centromeric DNA to form similarly replication-perturbing structures.

6.2 ERCC6L2 as a resolver of DNA secondary structures

The presence of ERCC6L2 at centromeres (Figure 4-5) raises intriguing questions about its cellular function. This observation might suggest that something about the repetitive and highly specialised genomic location that is the centromere requires

ERCC6L2 activity for some aspect of its biology. The fact that we observe ERCC6L2 localisation to centromeres throughout interphase in even undamaged cells implies that ERCC6L2's presence at these sites might either safeguard them from genomic instability somehow, or that ERCC6L2 is required for the tolerance of potentially genotoxic events which occur more frequently at centromeres than other genomic loci. However, since ERCC6L2 remains at centromeres during their replication (Figure 4-6) and interacts with PCNA (Figure 4-4), a critical regulator of DNA replication, it seems likely that ERCC6L2 functions somehow in DNA replication at these sites.

Centromeres are composed of higher order repeats of highly repetitive sequences, which can form replication-impeding, hairpin-containing secondary structures [294-298]. The secondary structures that can form at repetitive sequences can constitute sources of genomic instability, and it is emerging that that specialised mechanisms have evolved to enable efficient DNA replication and repair at centromeres, although the precise details of these mechanisms are far from understood [238, 290, 304-306, 308].

The notion that ERCC6L2 functions in DNA replication across sites bearing polymerase-perturbing secondary structures is supported by the observation that ERCC6L2 interacts directly with PCNA (Figure 5-2) and that its knockout induces a DNA replication defect in etoposide-treated cells (Figure 4-1). Notably, this effect of ERCC6L2 on DNA replication under genotoxic stress is dependent on its ATPase (and therefore likely translocase) activity (Figure 4-2).

Given that ERCC6L2 is not visible in unperturbed replication factories (Figure 4-4), such a role in DNA replication may be limited to a subset of replication forks, perhaps defined by genomic location, localised replication stress or both. It is therefore tempting to speculate that ERCC6L2 promotes efficient centromeric DNA replication by

alleviating the barrier to replication caused by bulky secondary structures at these sites (Figure 6-1).

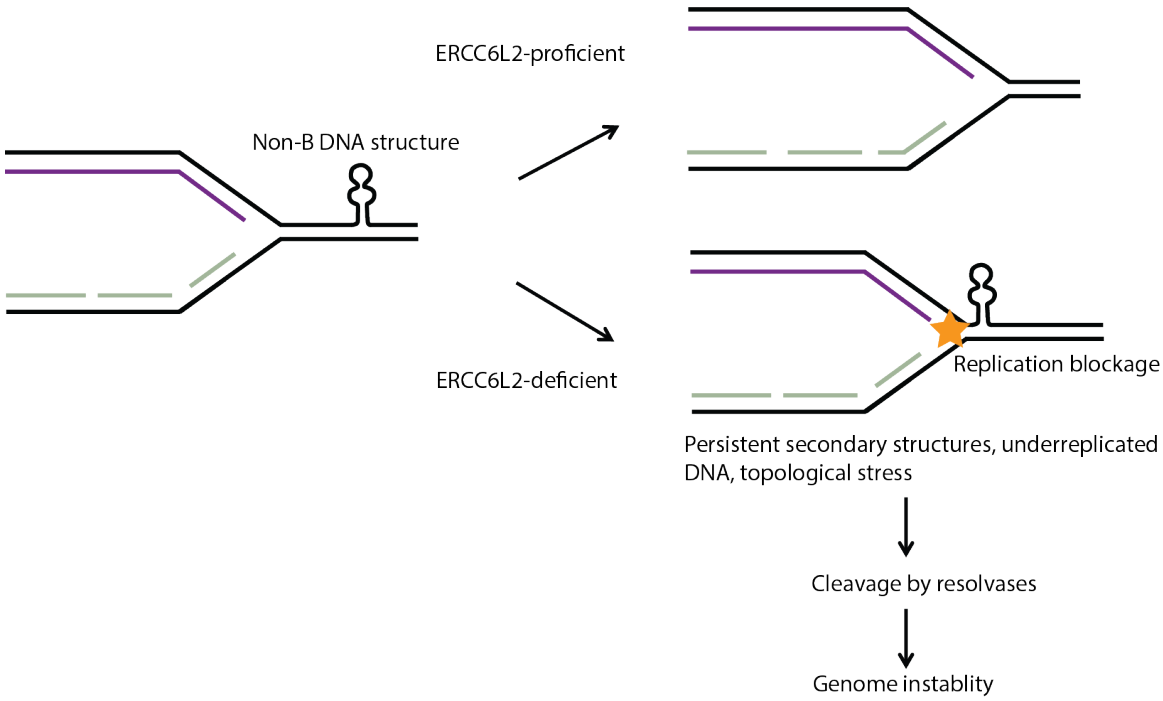


Figure 6-1 Working hypothesis of ERCC6L2 function

Our working model of a role for ERCC6L2 at centromeres. We speculate that non-B DNA structures formed at repetitive sequences such as centromeres are resolved ahead of replication forks in a manner involving ERCC6L2. While I have placed this in the context of replication, it remains entirely possible that ERCC6L2 performs a similar function ahead of transcription and in late-stage DNA repair; these models are not necessarily mutually exclusive.

Although ERCC6L2 localises to centromeres during their replication (Figure 4-6b), it also present at these sites during G1 and G2 (Figure 4-6a). This might suggest – although not necessarily – suggest some other relevance for ERCC6L2 function at centromeres. Interestingly, it is thought that secondary structures which form at repetitive sequences could act as blockages to transcription machinery as well as replication forks [402-405]. Since transcription is known to occur at centromeres in G1 and G2 despite its repressive chromatin state [367, 406], it could be that ERCC6L2's activity relieves the secondary structures formed at centromeric DNA to facilitate both transcription and replication through these sequences, depending on the context. Considering that centromeric transcription has been associated with the deposition of

new CENP-A at centromeric chromatin [367], it could be interesting to test whether new CENP-A deposition is affected in ERCC6L2 KO cells.

Interestingly, a similar mode of ERCC6L2 function has been suggested recently [336]. Tummala and colleagues proposed that, in concert with DNA-PK, ERCC6L2 functions in the removal of transcription-blocking RNA/DNA hybrids, and that the persistence of these structures leads to defective transcription-coupled DNA repair [336]. However, this study did not address the possibility that their experimental conditions were inducing DSBs – or other genotoxic insults such as replication stress – and direct *in vitro* evidence for such a function is still lacking. We hope that with the successful purification of recombinant ERCC6L2 by other members of the group, such assays will be within reach. However, it should be noted that a function for ERCC6L2 in resolving RNA/DNA hybrids is not mutually exclusive with the resolution of other DNA structures. Other factors with similar functions have been reported to show a degree of substrate flexibility, including DNA2 [308, 314], BLM [407, 408] and RTEL1 [255, 256, 409].

Given our findings, it could be interesting to examine the nature of the RNA/DNA hybrids reported by Tummala et al to be elevated in ERCC6L2-deficient cells [336]. Specifically, since our data might point to a function for ERCC6L2 at genomic loci prone to the formation of DNA secondary structures, it would be interesting to assess the rate of RNA/DNA hybrid appearance at regions such as centromeres, common fragile sites and telomeres. Furthermore, confirming whether or not reported interaction between ERCC6L2 and DNA-PK [336] is real or artificial might allow us to explore the physiological importance of this interaction in the context of our findings.

If ERCC6L2 does indeed function in the removal of DNA secondary structures and/or RNA/DNA hybrids, this raises important questions about genomic loci aside from

centromeres which may be affected by ERCC6L2 deficiency. Chromatin immunoprecipitation (ChIP) of overexpressed ERCC6L2 and quantitative PCR (qPCR) of regions of interest might reveal whether or not ERCC6L2 is enriched at genomic loci other than centromeres, such as CFSs. Additionally, techniques such as DRIP-seq using the S9.6 antibody against RNA/DNA hybrids may reveal whether or not specific genomic loci are more prone to R-loop formation in ERCC6L2-deficient cells than in WT cells. Additionally, discovering whether or not G-quadruplex processing is affected by ERCC6L2 could be achieved by assessing DNA replication rates and cell viability under treatment with G4-stabilising drugs such as pyridostatin [410].

Finally, it might be expected that the activity of nucleases known to cleave potentially deleterious DNA structures during early mitosis such as MUS81, SLX1-SLX4 and GEN1 compensates for the continued presence of such structures into mitosis. Knocking out or inhibiting these nucleases in our ERCC6L2-deficient cell lines might reveal new phenotypes, such as the persistence of lagging chromosomes or UFBs in anaphase, or micronuclei resulting from chromosome segregation defects in mitosis. A detailed analysis of UFBs is particularly appealing, since it could reveal deeper insights into ERCC6L2's function. An increase in the frequency of centromeric UFBs might indicate the persistence of dsDNA catenanes at centromeres, while an increase in FS-UFBs, characterised by the presence of twin FANCD2/FANCI foci would suggest an under-replication phenotype, as mentioned in Section 1.4.5.

Overall, it remains a possibility – albeit a largely speculative one – that ERCC6L2 functions in the dissolution or resolution of DNA structures which can form at certain sites in unperturbed cells, and perhaps near sites of DNA damage during DNA repair processes. Should this be the case, it could have important implications for the understanding of bone marrow failure phenotypes caused by ERCC6L2 deficiency.

6.3 Concluding remarks

This project has yielded a number of important new insights, which may prove critical to understanding the cellular role of ERCC6L2 and the pathological basis of ERCC6L2 deficiencies.

ERCC6L2$^{-/-}$ U2OS cells are sensitive to the induction of double strand breaks and exhibit markedly impaired DNA synthesis when treated with the topoisomerase II inhibitor etoposide (Figure 3-5b, Figure 4-1). This corroborates findings that cell lines derived from ERCC6L2-deficient BMF patients are sensitive to genotoxic insult [332, 336, 337]. Notably, the loss of ERCC6L2's ATPase activity by a single point mutation is sufficient to cause these phenotypes (Figure 3-7, Figure 4-2), demonstrating for the first time that ATP hydrolysis by ERCC6L2 is an important feature of the DNA damage response. In accordance with this, I have shown that ERCC6L2 is recruited to sites of DNA damage induced by laser microirradiation through its C-terminus, which is lost in a large number of clinically relevant ERCC6L2 mutants (Figure 3-1, Figure 3-3, Figure 4-3). Importantly, all clinically relevant truncation mutants of ERCC6L2 tested show defective recruitment to DNA lesions, implying that a DDR deficiency might underlie ERCC6L2 deficiency disorders (Figure 3-3). In addition to its recruitment to DNA lesions, I have shown that ERCC6L2 localises to centromeres even in unperturbed interphase U2OS cells and that this centromere targeting occurs through its C-terminus (Figure 4-5). As might be expected given frequent mislocalisation, the ERCC6L2 truncation mutants tested failed to colocalise with centromere markers (Figure 4-7). Notably, even the ERCC6L2 R1266* mutant, which does localise to the nucleus, still fails to localise to centromeres (Figure 4-7), suggesting that this behaviour is driven through the C-terminus. The finding that ERCC6L2 may function at centromeres opens new avenues of investigation into the basis of the BMF pathologies of ERCC6L2-mutant

patients. Finally, I have found for the first time that ERCC6L2 interacts with PCNA through a novel type of atypical PIP-box (Figure 5-2). While the precise contribution of this interaction to ERCC6L2's cellular function is still unclear, it seems likely that this interaction will play a role in ERCC6L2 function, based on the importance of PCNA in regulating proteins' localisation to replication forks. Identifying the exact contextual importance of the ERCC6L2-PCNA interaction might prove a valuable target for mechanistic investigations in follow-up studies.

These results when taken alongside past observations regarding ERCC6L2's function point to several possible interpretations as to its cellular role. On a broad level, the sensitivity of ERCC6L2-deficient cells to drugs which induce double strand breaks points to a function in DSB repair (Figure 3-5b) [332, 336, 337]. However, no defect in the major DSB repair pathways has been identified under ERCC6L2 depletion [337], suggesting that any role in the DSB response might be indirect or compensated to some extent by alternative pathways. A thorough investigation of the impact of ERCC6L2 deficiency on the wide range of competing DSB repair pathways may reveal which DNA repair mechanism(s) involve ERCC6L2. An alternative – although not necessarily mutually exclusive – interpretation of all the available data is that ERCC6L2 functions in the resolution of potentially genotoxic DNA structures such as G-quadruplexes, RNA/DNA hybrids, hairpin structures or D-loops. Since these structures can occur at some loci in unperturbed cells and may form as intermediates in DNA repair processes, this interpretation offers an attractive explanation for the observed DDR defects in ERCC6L2 cells along with ERCC6L2's localisation at centromeres. However, *in vitro* investigations into such activities using purified ERCC6L2 will be necessary conclude this firmly, since the actions of a multitude of other proteins such as BLM, RTEL1,

DNA2, SLX1-SLX4, MUS81 and GEN1 might mask some cellular phenotypes of ERCC6L2 loss.

A number of immediate questions over ERCC6L2 function are attractive for follow-up studies and may yield great benefits to our understanding of this protein. These include structure-function studies such as identifying the binding partner of ERCC6L2's Tudor domain and characterising the function of the VIGSSK domain. Characterising the basis of ERCC6L2's recruitment to DNA damage sites might be highly beneficial, as well as identifying how ERCC6L2 localises to centromeres at a mechanistic level. Exploring the functional link between ERCC6L2's role in the DDR and its role at centromeres could be critical to understanding ERCC6L2 and pathologies associated with its deficiency.

While ERCC6L2 mutations have been identified in only a small number of BMF patients, it could be important for the field – and the patients – that their long-term health is monitored closely. Loss-of-function mutants in many DNA repair factors can cause pathological effects, many of which predispose patients to cancer [2]. While two patients so far have died as a result of AML, the first patients identified were only reported in 2014 and it remains possible that as these patients age, new pathological features will reveal themselves. One of the patients who died from AML has two siblings who currently have MDS, likely predisposing them to the development of AML [336].

Overall, this project has made strong steps toward an understanding of ERCC6L2 function at the cellular level. It is our interpretation of the available data that ERCC6L2 functions at the interface between secondary DNA structures and DNA replication, and that such a role supports replication across sites bearing these structures, such as centromeres. This function might explain the DNA damage sensitivity noted in ERCC6L2-deficient cells, but further *in vitro* experiments using purified ERCC6L2 would

be required to confirm this hypothesis. It is hoped that follow-up studies at the mechanistic and physiological level will yield insights into the basis of pathologies resulting from ERCC6L2 deficiency and into the regulation of the DNA damage response as a whole.

7 Appendix

Table 7-1: Plasmids

Plasmid	Vector	Source	F primer	R primer	Notes/uses
pDONR-221	pDONR-221	Invitrogen	n/a	n/a	Empty vector
pN-YFP-DEST	pN-YFP-DEST	Invitrogen	n/a	n/a	Empty vector
pLX304	pLX304	David Root (Addgene #25890) [411]	n/a	n/a	Empty vector
pEXPR-IBA105	pEXPR-IBA105	IBA Life Sciences	n/a	n/a	Empty vector
pET28a	pET28a	Dragana Ahel	n/a	n/a	Empty vector
px459	pSpCas9(BB)-2A-Puro (PX459)	Feng Zhang (Addgene #48139) [412]	n/a	n/a	Empty vector
pCMV-VSVG	pCMV-VSVG	Bob Weinberg (Addgene #8454) [413]	n/a	n/a	Lentiviral envelope plasmid
pCMV-dr8.2-dpr	pCMV-dr8.2-dpr	Bob Weinberg (Addgene #8455) [413]	n/a	n/a	Lentiviral packaging plasmid
NLS-p221	pDONR-221	Cloned (Gateway)	NLS-GTWF	NLS-GTWR	Gateway donor construct
ERCC6L2-p221	pDONR-221	Dragana Ahel	n/a	n/a	Gateway donor construct
ERCC6L2-Y67L-p221	pDONR-221	Mutagenesis	Y67LmutF	Y67LmutR	Gateway donor construct
ERCC6L2-Q798A-p221	pDONR-221	Mutagenesis	Q798AmutF	Q798AmutR	Gateway donor construct
ERCC6L2-d93-p221	pDONR-221	Cloned (Gateway)	d93-GTWF	ERCC6L2-FL-GTWR	Gateway donor construct
ERCC6L2-NLS-Nterm-p221	pDONR-221	Cloned (Gateway)	ERCC6L2-Nterm-NLS-GTWF	ERCC6L2-Nterm-GTWR	Gateway donor construct
ERCC6L2-Cterm-p221	pDONR-221	Dragana Ahel	ERCC6L2-Cterm-GTWF	ERCC6L2-FL-GTWR	Gateway donor construct
ERCC6L2-2-1053-p221	pDONR-221	Cloned (Gateway)	ERCC6L2-GTWF	ERCC6L2-F5N-GTWR	Gateway donor construct
ERCC6L2-1054-Cterm-NLS-p221	pDONR-221	Cloned (Gateway)	ERCC6L2-F5C-NLS-GTWF	ERCC6L2-FL-GTWR	Gateway donor construct
ERCC6L2-1054-1247-NLS-p221	pDONR-221	Cloned (Gateway)	ERCC6L2-F5C-NLS-GTWF	ERCC6L2-F5-GTWR	Gateway donor construct
ERCC6L2-1248-Cterm-NLS-p221	pDONR-221	Cloned (Gateway)	ERCC6L2-F4-NLS-GTWF	ERCC6L2-FL-GTWR	Gateway donor construct
ERCC6L2-701-1247-p221	pDONR-221	Cloned (Gateway)	ERCC6L2-Chalf-GTWF	ERCC6L2-FL-GTWR	Gateway donor construct
ERCC6L2-R1266*-p221	pDONR-221	Mutagenesis	c3796tmutF	c3796tmutR	Gateway donor construct
ERCC6L2-E729fs-p221	pDONR-221	Mutagenesis	2187delgmutF	2187delgmutR	Gateway donor construct
ERCC6L2-R655*-p221	pDONR-221	Mutagenesis	ERCC6L2-C>AmutF	ERCC6L2-C>AmutR	Gateway donor construct
ERCC6L2-I486fs-p221	pDONR-221	Mutagenesis	c1458delmutF	c1458delmutR	Gateway donor construct
ERCC6L2-Q502*-p221	pDONR-221	Mutagenesis	c1504tmutF	c1504tmutR	Gateway donor construct
ERCC6L2-T413Cfs*2-p221	pDONR-221	Mutagenesis	ERCC6L2-delmutF	ERCC6L2-delmutR	Gateway donor construct
ERCC6L2-NOSTOP-p221	pDONR-221	Cloned (Gateway)	ERCC6L2-GTWF	ERCC6L2-NOSTOP-GTWR	Gateway donor construct
ERCC6L2-K165R-NOSTOP-p221	pDONR-221	Mutagenesis	K165mutF	K165mutR	Gateway donor construct
ERCC6L2-Q798A-NOSTOP-p221	pDONR-221	Mutagenesis	Q798AmutF	Q798AmutR	Gateway donor construct
ERCC6L2-701-	pDONR-221	Cloned (Gateway)	ERCC6L2-Chalf-	ERCC6L2-	Gateway donor

1098-p221			GTWF	F5N+1-GTWR	construct	
ERCC6L2-797-1098-p221	pDONR-221	Cloned (Gateway)	ERCC6L2-Ncut2-NLS-GTWF	ERCC6L2-F5N+1-GTWR	Gateway construct	donor
ERCC6L2-820-1098-p221	pDONR-221	Cloned (Gateway)	ERCC6L2-Ncut2/3-NLS-GTWF	ERCC6L2-F5N+1-GTWR	Gateway construct	donor
ERCC6L2-701-1247-Q798A-p221	pDONR-221	Mutagenesis	Q798AmutF	Q798AmutR	Gateway construct	donor
ERCC6L2-701-1247-L799-p221	pDONR-221	Mutagenesis	L799AmutF	L799AmutR	Gateway construct	donor
ERCC6L2-701-1247-L801A-p221	pDONR-221	Mutagenesis	L801AmutF	L801AmutR	Gateway construct	donor
ERCC6L2-701-1247-Q803A-p221	pDONR-221	Mutagenesis	Q803AmutF	Q803AmutR	Gateway construct	donor
ERCC6L2-701-1247-C804A-p221	pDONR-221	Mutagenesis	C804AmutF	C804AmutR	Gateway construct	donor
ERCC6L2-701-1247-F806A-p221	pDONR-221	Mutagenesis	F806AmutF	F806AmutR	Gateway construct	donor
YFP-NLS	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-Y67L	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-Q798A	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-d93	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-NLS-Nterm	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-Cterm	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-2-1053	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-1054-Cterm-NLS	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-1054-1247-NLS	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-1248-Cterm-NLS	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-701-1247	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-R1266*	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-E729fs	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-R655*	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-I486fs	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-Q502*	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-T413Cfs*2	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-701-1098	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-797-1098	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-820-1098	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-701-1098-Q798A	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression	
YFP-ERCC6L2-	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian	

701-1098-L799A					overexpression
YFP-ERCC6L2-701-1098-L801A	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression
YFP-ERCC6L2-701-1098-Q803A	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression
YFP-ERCC6L2-701-1098-C804A	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression
YFP-ERCC6L2-701-1098-F806A	pN-YFP-DEST	Cloned (Gateway)	n/a	n/a	Mammalian overexpression
ERCC6L2-sgRNA1	px459	Andrew Bassett	n/a	n/a	sgRNA target: GCATAAAAAGGGAACTCGTGAGG
ERCC6L2-sgRNA2	px459	Andrew Bassett	n/a	n/a	sgRNA target: TATGGACACTACATCCATGGAGG
EV-PLX304	pLX304	Cloned (Gateway)	n/a	n/a	Stable overexpression (cloned from NLS-p221)
ERCC6L2-WT-V5	pLX304	Cloned (Gateway)	n/a	n/a	Stable overexpression
ERCC6L2-K165R-V5	pLX304	Cloned (Gateway)	n/a	n/a	Stable overexpression
ERCC6L2-Q798A-V5	pLX304	Cloned (Gateway)	n/a	n/a	Stable overexpression
TwinStrep-ERCC6L2	pEXPR-IBA105	Cloned (Restriction)	ERCC6L2-EcoRI-F	ERCC6L2-XhoI-R	Mammalian overexpression
TwinStrep-ERCC6L2-701-1247	pEXPR-IBA105	Cloned (Restriction)	ERCC6L2-F5-EcoRI-F	ERCC6L2-F5-XhoI-R	Mammalian overexpression
TwinStrep-ERCC6L2-701-1247-Q798A	pEXPR-IBA105	Cloned (Restriction)	ERCC6L2-F5-EcoRI-F	ERCC6L2-F5-XhoI-R	Mammalian overexpression
TwinStrep-ERCC6L2-701-1247-L799A	pEXPR-IBA105	Cloned (Restriction)	ERCC6L2-F5-EcoRI-F	ERCC6L2-F5-XhoI-R	Mammalian overexpression
TwinStrep-ERCC6L2-701-1247-L801A	pEXPR-IBA105	Cloned (Restriction)	ERCC6L2-F5-EcoRI-F	ERCC6L2-F5-XhoI-R	Mammalian overexpression
TwinStrep-ERCC6L2-701-1247-Q803A	pEXPR-IBA105	Cloned (Restriction)	ERCC6L2-F5-EcoRI-F	ERCC6L2-F5-XhoI-R	Mammalian overexpression
TwinStrep-ERCC6L2-701-1247-C804A	pEXPR-IBA105	Cloned (Restriction)	ERCC6L2-F5-EcoRI-F	ERCC6L2-F5-XhoI-R	Mammalian overexpression
TwinStrep-ERCC6L2-701-1247-F806A	pEXPR-IBA105	Cloned (Restriction)	ERCC6L2-F5-EcoRI-F	ERCC6L2-F5-XhoI-R	Mammalian overexpression
YFP-ALC1	pN-YFP-DEST	Dragana Ahel	n/a	n/a	Mammalian overexpression
6xHis-PCNA	pET28a	Dragana Ahel	n/a	n/a	PCNA expression and purification

Table 7-2: Oligonucleotides

Primer name	Sequence	Source
M13F	TGTAAAACGACGGCCAGT	Source Bioscience
M13R	CAGGAAACAGCTATGACC	Source Bioscience
JYBF	CGCAAATGGGCGGTAGGCGTG	Eurofins
JYBR	CCTCATTTTATTAGGAAAGG	Eurofins
WPRER	CATAGCGTAAAAGGAGCAACA	Invitrogen
StrepseqF	GAGAACCCACTGCTTACTGGC	Eurofins
StrepseqR	TAGAAGGCACAGTCGAGG	Eurofins
ERCC6L2-Seq1	AAAGGGAACTCGTGAGG	Eurofins

ERCC6L2-Seq2	CCTAAGAAGGAAGACCG	Eurofins
ERCC6L2-Seq2/3	TTACCGACGACTTGATGG	Eurofins
ERCC6L2-Seq3	CAAGGGTCTTGTCTTACG	Eurofins
ERCC6L2-Seq4	CACTGATGACAAAAGAAATGG	Eurofins
ERCC6L2-Seq5	GTCTACCATACAAACCAG	Eurofins
sg1-gDNA-GTWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCAAATCTCTTTGTACATACTTTGTC	Eurofins
sg1-gDNA-GTWR	GGGGACCACTTTGTACAAGAAAGCTGGGTATAGATTAAAACCTCAAAGCAAAACC	Eurofins
sg2-gDNA-GTWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGTTAAAGATTGCCCTAGGAATC	Eurofins
sg2-gDNA-GTWR	GGGGACCACTTTGTACAAGAAAGCTGGGTATTATCCACTCATGTACACAGTAAAGTG	Eurofins
NLS-GTWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGCGAGCCCCTAAGAAAAAGCGGAAGGTG GGCGGCTGATACCCAGCTTTCTTGTACAAAGTGGTCCCC	Eurofins
NLS-GTWR	GGGGACCACTTTGTACAAGAAAGCTGGGTATCAGCCGCCCACCTTCCGCTTTTTCTTAGGGG CTCGCATGAAGCCTGCTTTTTTGTACAAACTTGTCCCC	Eurofins
d93-GTWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGCGAGCCCCTAAGAAAAAGCGGAAGGTG GGCGGCCTTATATTTGATGATGAAGATTTAGAAAAAC	Eurofins
ERCC6L2-GTWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCAGCCGGGCTCGGC	Eurofins
ERCC6L2-Nterm-NLS-GTWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGCGAGCCCCTAAGAAAAAGCGGAAGGTG GGCGGCCAGCCGGGCTCGGC	Eurofins
ERCC6L2-Nterm-GTWR	GGGGACCACTTTGTACAAGAAAGCTGGGTATCACACCTCCAGGATGTCC	Eurofins
ERCC6L2-FL-GTWR	GGGGACCACTTTGTACAAGAAAGCTGGGTATTATGTGGTACTCTGTGTATTGG	Eurofins
ERCC6L2-NOSTOP-GTWR	GGGGACCACTTTGTACAAGAAAGCTGGGTATGTGGTACTCTGTGTATTGGTTG	Invitrogen
ERCC6L2-F5N-GTWR	GGGGACCACTTTGTACAAGAAAGCTGGGTATCATGAATAATCATCAATAAACTGAGAG	Eurofins
ERCC6L2-F5C-NLS-GTWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGCGAGCCCCTAAGAAAAAGCGGAAGGTG GGCGGCTCCTCAGATGAGAGTTTATCCG	Eurofins
ERCC6L2-F4-NLS-GTWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGCGAGCCCCTAAGAAAAAGCGGAAGGTG GGCGGCTCGTCTTCTGTAAACGAATTTGC	Eurofins
ERCC6L2-Chalf-GTWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCAAGGGTCTTGTCTTACGAAG	Eurofins
ERCC6L2-F5-GTWR	GGGGACCACTTTGTACAAGAAAGCTGGGTATCAGTTAAAATAAGAGGCCATTTCTTC	Eurofins
ERCC6L2-F5N+1-GTWR	GGGGACCACTTTGTACAAGAAAGCTGGGTATCATGGAATAAAAGTGCTATTTTTCTTATTATG	Eurofins
ERCC6L2-Ncut2-NLS-GTWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGGACAGCTTACCTTACTCCAG	Eurofins
ERCC6L2-Ncut2/3-NLS-GTWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAGTGATGGAAATACTGCCTCTG	Eurofins
ERCC6L2-C>AmutF	GAACTGTGGAGGAAATCATGTATTTATGACAGATATACAAGCAGC	Eurofins
ERCC6L2-C>AmutR	GCTGCTTGTATATCTGTCATAAATACATGATTTCCTCCACAGTTC	Eurofins
ERCC6L2-delmutF	CCAGAAAGCTGTCTATCAGTGTTAGAAACAGAGGAC	Eurofins
ERCC6L2-delmutR	GTCCTCTGTTTCTAACACTGATAGACAGCTTTCTGG	Eurofins
Y67LmutF	CCTTTCTTGAAAATCTGCTAATAAGACGACTGCAAATGACTTGCCA	Eurofins
Y67LmutR	TGGCAAGTCATTTGCAGTCGTCTTATTAGCAGATTTTCAAGAAAGG	Eurofins
K165mutF	GTGATGACATGGGACTTGGAAGAACAGTACAGGTTATTTCA	Eurofins
K165mutR	TGAAATAACCTGTACTGTTCTTCCAAGTCCCATGTCATCAC	Eurofins
Q798AmutF	GCTTCCAGCTCTCCAGGAGCGCTTACCTTACTCC	Eurofins
Q798AmutR	GGAGTAAGGTAAGCGCTCCTGGAGAGCTGGAAGC	Eurofins
L799AmutF	CCAGCTCTCCAGGACAGGCTACCTTACTCCAGTGTG	Eurofins

L799AmutR	CACACTGGAGTAAGGTAGCCTGTCCTGGAGAGCTGG	Eurofins
L801AmutF	TCTCCAGGACAGCTTACCGCACTCCAGTGTGGTTTCTC	Eurofins
L801AmutR	GAGAAACCACACTGGAGTGCGGTAAGCTGTCCTGGAGA	Eurofins
Q803AmutF	AGGACAGCTTACCTTACTCGCGTGTGGTTTCTCGAAATTG	Eurofins
Q803AmutR	CAATTTCGAGAAACCACACGCGAGTAAGGTAAGCTGTCCT	Eurofins
C804AmutF	GACAGCTTACCTTACTCCAGGCTGGTTTCTCGAAATTGCTTG	Eurofins
C804AmutR	CAAGCAATTTCGAGAAACCAGCCTGGAGTAAGGTAAGCTGTC	Eurofins
F806AmutF	GCTTACCTTACTCCAGTGTGGTGCCTCGAAATTGCTTGAAACAAAA	Eurofins
F806AmutR	TTTTGTTTCAAGCAATTTCGAGGCACCACACTGGAGTAAGGTAAGC	Eurofins
ERCC6L2-EcoRI-F	AATCTAAGAATTCGCAGCCGGGCTCGGC	Eurofins
ERCC6L2-XhoI-R	AATCTAACTCGAGTTATGTGGTACTCTGTGTATTGG	Eurofins
ERCC6L2-F5-EcoRI-F	AATCTAAGAATTCGCAAGGGTCTTGTCTTACGAAG	Eurofins
ERCC6L2-F5-XhoI-R	AATCTAACTCGAGTCAGTTAAAATAAGAGGCCATTTCTTC	Eurofins
c3796tmutF	AACAATGCCACATCAGAAGAATGACAGAAAATGCTAAGAG	Invitrogen
c3796tmutR	CTCTTAGCATTTTCTGTCATTCTTCTGATGTGGCATTGGT	Invitrogen
g847amutF	AATGGTCAGCTGTCATTGTGAATGAAGCTCATAGAATCAAG	Invitrogen
g847amutR	CTTGATTCTATGAGCTTCATTCACAATGACAGCTGACCATT	Invitrogen
c1458delmutF	TAGTACTTCCAAACAACAGGAAACACTTATAAAAGGATATGTGATCA	Eurofins
c1458delmutR	TGATCACATATCCTTTTATAAGTGTTTCCTGTTGTTTGGAAGTACTA	Eurofins
2187delgmutF	AGCCACAACATGGTTGAAAGAGGACCTCCAGC	Invitrogen
2187delgmutR	GCTGGAGGTCCTCTTTCAACCATGTTGTGGCT	Invitrogen
c1504tmutF	CAGATTCCCAGATTTTGTGTAGAAAAGCAAAGATGCAGC	Invitrogen
c1504tmutR	GCTGCATCTTTGCTTTTCTACACAAAATCTGGGAATCTG	Invitrogen
c1963tmutF	GAACTGTGGAGGAAATCATGTATTTATGACAGATATACAAGCAGC	Invitrogen
c1963tmutR	GCTGCTTGTATATCTGTCATAAATACATGATTTCCTCCACAGTTC	Invitrogen

Table 7-3: Antibodies

Antibody target	Manufacturer	Product Code	Host Species	Dilution (WB)	Dilution (IF)
β-tubulin	Abcam	ab6046	Rabbit	1:5000	n/a
CENP-A	Abcam	ab13939	Mouse	n/a	1:500
CENP-B	Abcam	ab25734	Rabbit	n/a	1:300
Cyclin A	Santa Cruz	sc-271682	Mouse	n/a	1:300
γH2AX (S139)	Millipore	05-636	Mouse	n/a	1:500
PCNA	Abcam	ab18197	Rabbit	1:1000	1:500
Strep tag	IBA Life Sciences	2-1507-001	Mouse	1:2000	n/a
V5 tag	Abcam	ab9116	Rabbit	1:1000	n/a
Anti-Mouse-HRP	Dako	P0447	Goat	1:5000	n/a
Anti-Mouse-Alexa405	Life Technologies	A31553	Goat	n/a	1:500
Anti-Mouse-Alexa594	Life Technologies	A11032	Goat	n/a	1:500
Anti-Mouse-Alexa647	Life Technologies	A32728	Goat	n/a	1:500
Anti-Rabbit-HRP	Dako	P0399	Pig	1:5000	n/a
Anti-Rabbit-Alexa555	Life Technologies	A32732	Goat	n/a	1:500
Anti-Rabbit-Alexa594	Life Technologies	A11037	Goat	n/a	1:500

Table 7-4: Peptides

Peptide	Sequence (aa)	Source
ERCC6L2 PIP WT	Biotin-TNATSEERQKMLRDFYASQYPE	Eurogentech
ERCC6L2 PIP mut	Biotin-TNATSEERAKMLRDAAASQYPE	Eurogentech

ZRANB3 PIP WT	Biotin-FTHFEKEKQHDIRSFFVPQPKK	Eurogentech
ZRANB3 PIP mut	Biotin-FTHFEKEKAHDIRSAAVPQPKK	Eurogentech
ERCC6L2 a-PIP WT	Biotin-SKASSSPGQLTLLQCGFSKLLE	Eurogentech
ERCC6L2 a-PIP mut	Biotin-SKASSSPGALTLLQCGFSKLLE	Eurogentech

8 Bibliography

1. Lindahl, T. and D.E. Barnes, *Repair of endogenous DNA damage.* Cold Spring Harb Symp Quant Biol, 2000. **65**: p. 127-134.
2. Jackson, S.P. and J. Bartek, *The DNA-damage response in human biology and disease.* Nature, 2009. **461**(7267): p. 1071-1078.
3. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
4. Farmer, H., et al., *Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy.* Nature, 2005. **434**: p. 917-921.
5. Bryant, H.E., et al., *Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase.* Nature, 2005. **434**(7035): p. 913-7.
6. Wogan, G.N., et al., *Environmental and chemical carcinogenesis.* Semin Cancer Biol, 2004. **14**(6): p. 473-486.
7. Marteijn, J.A., et al., *Understanding nucleotide excision repair and its roles in cancer and ageing.* Nat Rev Mol Cell Biol, 2014. **15**(7): p. 465-81.
8. Shah, P. and Y.Y. He, *Molecular regulation of UV-induced DNA repair.* Photochem Photobiol, 2015. **91**(2): p. 254-64.
9. Mansour, W.Y., et al., *Hierarchy of nonhomologous end-joining, single-strand annealing and gene conversion at site-directed DNA double-strand breaks.* Nucleic Acids Res, 2008. **36**(12): p. 4088-98.
10. Wang, H., et al., *Biochemical evidence for Ku-independent backup pathways of NHEJ.* Nucleic Acids Res, 2003. **31**(18): p. 5377-88.
11. Kabotyanski, E.B., et al., *Double-strand break repair in Ku86- and XRCC4-deficient cells.* Nucleic Acids Res, 1998. **26**(23): p. 5333-42.
12. Kim, H. and A.D. D'Andrea, *Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway.* Genes & Dev, 2012. **26**: p. 1393-1408.
13. Kraus, E., W.Y. Leung, and J.E. Haber, *Break-induced replication: a review and an example in budding yeast.* Proc Natl Acad Sci U S A, 2001. **98**(15): p. 8255-62.
14. Helleday, T., *Pathways for mitotic homologous recombination in mammalian cells.* Mutat Res, 2003. **532**(1-2): p. 103-15.
15. Deweese, J.E. and N. Osheroff, *The DNA cleavage reaction of topoisomerase II: wolf in sheep's clothing.* Nucleic Acids Res, 2009. **37**(3): p. 738-48.
16. Li, T.K. and L.F. Liu, *Tumor cell death induced by topoisomerase-targeting drugs.* Annu Rev Pharmacol Toxicol, 2001. **41**: p. 53-77.
17. Pourquier, P. and Y. Pommier, *Topoisomerase I-mediated DNA damage.* Adv Cancer Res, 2001. **80**: p. 189-216.
18. Rosado, I.V., et al., *Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway.* Nat Struct Mol Biol, 2011. **18**(12): p. 1432-4.
19. Langevin, F., et al., *Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice.* Nature, 2011. **475**(7354): p. 53-8.
20. O'Brien, P.J., A.G. Siraki, and N. Shangari, *Aldehyde sources, metabolism, molecular toxicity mechanisms, and possible effects on human health.* Crit Rev Toxicol, 2005. **35**(7): p. 609-62.
21. Vaz, B., et al., *Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair.* Mol Cell, 2016. **64**(4): p. 704-719.
22. Stingele, J., et al., *Mechanism and Regulation of DNA-Protein Crosslink Repair by the DNA-Dependent Metalloprotease SPRTN.* Mol Cell, 2016. **64**(4): p. 688-703.

23. Weng, M.W., et al., *Aldehydes are the predominant forces inducing DNA damage and inhibiting DNA repair in tobacco smoke carcinogenesis.* Proc Natl Acad Sci U S A, 2018.
24. Savitsky, K., et al., *A single ataxia telangiectasia gene with a product similar to PI-3 kinase.* Science, 1995. **268**(5218): p. 1749-53.
25. Bartek, J. and J. Lukas, *Mammalian G1- and S-phase checkpoints in response to DNA damage.* Curr Opin Cell Biol, 2001. **13**(6): p. 738-47.
26. Zhou, B.B. and S.J. Elledge, *The DNA damage response: putting checkpoints in perspective.* Nature, 2000. **408**(6811): p. 433-9.
27. Liu, Q., et al., *Chk1 is an essential kinase that is regulated by Atr and required for the G(2)/M DNA damage checkpoint.* Genes & Dev, 2000. **14**(12): p. 1448-1459.
28. Matsuoka, S., M. Huang, and S.J. Elledge, *Linkage of ATM to cell cycle regulation by the Chk2 protein kinase.* Science, 1998. **282**(5395): p. 1893-7.
29. Marechal, A. and L. Zou, *DNA damage sensing by the ATM and ATR kinases.* Cold Spring Harb Perspect Biol, 2013. **5**(9).
30. Stracker, T.H., T. Usui, and J.H.J. Petrini, *Taking the time to make important decisions: The checkpoint effector kinases Chk1 and Chk2 and the DNA damage response.* DNA Repair, 2009. **8**(9): p. 1047-1054.
31. Lee, J.H. and T.T. Paull, *Activation and regulation of ATM kinase activity in response to DNA double-strand breaks.* Oncogene, 2007. **26**(56): p. 7741-8.
32. Gatei, M., et al., *Ataxia-telangiectasia-mutated (ATM) and NBS1-dependent phosphorylation of Chk1 on Ser-317 in response to ionizing radiation.* J Biol Chem, 2003. **278**(17): p. 14806-11.
33. Garcia-Muse, T. and S.J. Boulton, *Distinct modes of ATR activation after replication stress and DNA double-strand breaks in Caenorhabditis elegans.* EMBO J, 2005. **24**(24): p. 4345-55.
34. Jazayeri, A., et al., *ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks.* Nat Cell Biol, 2006. **8**(1): p. 37-45.
35. Stiff, T., et al., *Nbs1 is required for ATR-dependent phosphorylation events.* EMBO J, 2005. **24**(1): p. 199-208.
36. Chibazakura, T., et al., *Cyclin A promotes S-phase entry via interaction with the replication licensing factor Mcm7.* Mol Cell Biol, 2011. **31**(2): p. 248-55.
37. Vermeulen, K., D.R. Van Bockstaele, and Z.N. Berneman, *The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer.* Cell Prolif, 2003. **36**(3): p. 131-49.
38. Maya-Mendoza, A., et al., *High speed of fork progression induces DNA replication stress and genomic instability.* Nature, 2018. **559**(7713): p. 279-284.
39. Harper, J.W., et al., *The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases.* Cell, 1993. **75**(4): p. 805-16.
40. Harper, J.W., et al., *Inhibition of cyclin-dependent kinases by p21.* Mol Biol Cell, 1995. **6**(4): p. 387-400.
41. Deng, C., et al., *Mice lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control.* Cell, 1995. **82**(4): p. 675-84.
42. Hirao, A., et al., *Chk2 is a tumor suppressor that regulates apoptosis in both an ataxia telangiectasia mutated (ATM)-dependent and an ATM-independent manner.* Mol Cell Biol, 2002. **22**(18): p. 6521-32.
43. Kastan, M.B., et al., *Participation of p53 protein in the cellular response to DNA damage.* Cancer Res, 1991. **51**(23 Pt 1): p. 6304-11.
44. Shaltiel, I.A., et al., *The same, only different - DNA damage checkpoints and their reversal throughout the cell cycle.* J Cell Sci, 2015. **128**(4): p. 607-20.

45. Lafarga, V., et al., *p38 Mitogen-activated protein kinase- and HuR-dependent stabilization of p21(Cip1) mRNA mediates the G(1)/S checkpoint.* Mol Cell Biol, 2009. **29**(16): p. 4341-51.
46. Reinhardt, H.C., et al., *p53-deficient cells rely on ATM- and ATR-mediated checkpoint signaling through the p38MAPK/MK2 pathway for survival after DNA damage.* Cancer Cell, 2007. **11**(2): p. 175-89.
47. Agami, R. and R. Bernards, *Distinct initiation and maintenance mechanisms cooperate to induce G1 cell cycle arrest in response to DNA damage.* Cell, 2000. **102**(1): p. 55-66.
48. Deckbar, D., et al., *The limitations of the G1-S checkpoint.* Cancer Res, 2010. **70**(11): p. 4412-21.
49. Mailand, N., et al., *Rapid destruction of human Cdc25A in response to DNA damage.* Science, 2000. **288**(5470): p. 1425-9.
50. Santra, M.K., N. Wajapeyee, and M.R. Green, *F-box protein FBXO31 mediates cyclin D1 degradation to induce G1 arrest after DNA damage.* Nature, 2009. **459**(7247): p. 722-5.
51. Zeman, M.K. and K.A. Cimprich, *Causes and consequences of replication stress.* Nat Cell Biol, 2014. **16**(1): p. 2-9.
52. Chow, J.P., et al., *Differential contribution of inhibitory phosphorylation of CDC2 and CDK2 for unperturbed cell cycle control and DNA integrity checkpoints.* J Biol Chem, 2003. **278**(42): p. 40815-28.
53. Watanabe, N., M. Broome, and T. Hunter, *Regulation of the human WEE1Hu CDK tyrosine 15-kinase during the cell cycle.* EMBO J, 1995. **14**(9): p. 1878-91.
54. Beck, H., et al., *Regulators of cyclin-dependent kinases are crucial for maintaining genome integrity in S phase.* J Cell Biol, 2010. **188**(5): p. 629-38.
55. Hughes, B.T., et al., *Essential role for Cdk2 inhibitory phosphorylation during replication stress revealed by a human Cdk2 knockin mutation.* Proc Natl Acad Sci U S A, 2013. **110**(22): p. 8954-9.
56. O'Connell, M.J., et al., *Chk1 is a wee1 kinase in the G2 DNA damage checkpoint inhibiting cdc2 by Y15 phosphorylation.* EMBO J, 1997. **16**(3): p. 545-54.
57. Hochegger, H., S. Takeda, and T. Hunt, *Cyclin-dependent kinases and cell-cycle transitions: does one fit all?* Nat Rev Mol Cell Biol, 2008. **9**(11): p. 910-6.
58. Hirai, H., et al., *Small-molecule inhibition of Wee1 kinase by MK-1775 selectively sensitizes p53-deficient tumor cells to DNA-damaging agents.* Mol Cancer Ther, 2009. **8**(11): p. 2992-3000.
59. Gordon, D.J., B. Resio, and D. Pellman, *Causes and consequences of aneuploidy in cancer.* Nat Rev Genet, 2012. **13**(3): p. 189-203.
60. Brown, E.J. and D. Baltimore, *Essential and dispensable roles of ATR in cell cycle arrest and genome maintenance.* Genes Dev, 2003. **17**(5): p. 615-28.
61. Shibata, A., et al., *Role of ATM and the damage response mediator proteins 53BP1 and MDC1 in the maintenance of G(2)/M checkpoint arrest.* Mol Cell Biol, 2010. **30**(13): p. 3371-83.
62. Takai, H., et al., *Aberrant cell cycle checkpoint function and early embryonic death in Chk1(-/-) mice.* Genes Dev, 2000. **14**(12): p. 1439-47.
63. Painter, R.B. and B.R. Young, *Radiosensitivity in ataxia-telangiectasia: a new explanation.* Proc Natl Acad Sci U S A, 1980. **77**(12): p. 7315-7.
64. Rainey, M.D., et al., *Transient inhibition of ATM kinase is sufficient to enhance cellular sensitivity to ionizing radiation.* Cancer Res, 2008. **68**(18): p. 7466-74.

65. Giunta, S., R. Belotserkovskaya, and S.P. Jackson, *DNA damage signaling in response to double-strand breaks during mitosis.* J Cell Biol, 2010. **190**(2): p. 197-207.
66. van Vugt, M.A., et al., *A mitotic phosphorylation feedback network connects Cdk1, Plk1, 53BP1, and Chk2 to inactivate the G(2)/M DNA damage checkpoint.* PLoS Biol, 2010. **8**(1): p. e1000287.
67. Mailand, N., et al., *Destruction of Claspin by SCFbetaTrCP restrains Chk1 activation and facilitates recovery from genotoxic stress.* Mol Cell, 2006. **23**(3): p. 307-18.
68. Mamely, I., et al., *Polo-like kinase-1 controls proteasome-dependent degradation of Claspin during checkpoint recovery.* Curr Biol, 2006. **16**(19): p. 1950-5.
69. Peschiaroli, A., et al., *SCFbetaTrCP-mediated degradation of Claspin regulates recovery from the DNA replication checkpoint response.* Mol Cell, 2006. **23**(3): p. 319-29.
70. Combes, G., et al., *Playing polo during mitosis: PLK1 takes the lead.* Oncogene, 2017. **36**(34): p. 4819-4827.
71. Orthwein, A., et al., *Mitosis inhibits DNA double-strand break repair to guard against telomere fusions.* Science, 2014. **344**(6180): p. 189-93.
72. Lee, D.H., et al., *Dephosphorylation enables the recruitment of 53BP1 to double-strand DNA breaks.* Mol Cell, 2014. **54**(3): p. 512-25.
73. Terasawa, M., A. Shinohara, and M. Shinohara, *Canonical non-homologous end joining in mitosis induces genome instability and is suppressed by M-phase-specific phosphorylation of XRCC4.* PLoS Genet, 2014. **10**(8): p. e1004563.
74. Joukov, V. and A. De Nicolo, *Aurora-PLK1 cascades as key signaling modules in the regulation of mitosis.* Sci Signal, 2018. **11**(543).
75. Kabeche, L., et al., *A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation.* Science, 2018. **359**(6371): p. 108-114.
76. Uhlmann, F., *SMC complexes: from DNA to chromosomes.* Nat Rev Mol Cell Biol, 2016. **17**(7): p. 399-412.
77. Chan, K.L. and I.D. Hickson, *New insights into the formation and resolution of ultra-fine anaphase bridges.* Semin Cell Dev Biol, 2011. **22**(8): p. 906-12.
78. Bakhoum, S.F., et al., *Chromosomal instability drives metastasis through a cytosolic DNA response.* Nature, 2018. **553**(7689): p. 467-472.
79. Baumann, C., et al., *PICH, a centromere-associated SNF2 family ATPase, is regulated by Plk1 and required for the spindle checkpoint.* Cell, 2007. **128**(1): p. 101-14.
80. Chan, K.L. and I.D. Hickson, *On the origins of ultra-fine anaphase bridges.* Cell Cycle, 2009. **8**(19): p. 3065-6.
81. Chan, K.L., P.S. North, and I.D. Hickson, *BLM is required for faithful chromosome segregation and its localization defines a class of ultrafine anaphase bridges.* EMBO J, 2007. **26**(14): p. 3397-409.
82. Nielsen, C.F., et al., *PICH promotes sister chromatid disjunction and co-operates with topoisomerase II in mitosis.* Nat Commun, 2015. **6**: p. 8962.
83. Wang, L.H., et al., *Persistence of DNA threads in human anaphase cells suggests late completion of sister chromatid decatenation.* Chromosoma, 2008. **117**(2): p. 123-35.
84. Liu, Y., et al., *The origins and processing of ultra fine anaphase DNA bridges.* Curr Opin Genet Dev, 2014. **26**: p. 1-5.
85. Chan, K.L., et al., *Replication stress induces sister-chromatid bridging at fragile site loci in mitosis.* Nat Cell Biol, 2009. **11**(6): p. 753-60.

86. Naim, V. and F. Rosselli, *The FANC pathway and BLM collaborate during mitosis to prevent micro-nucleation and chromosome abnormalities.* Nat Cell Biol, 2009. **11**(6): p. 761-8.
87. Naim, V., et al., *ERCC1 and MUS81-EME1 promote sister chromatid separation by processing late replication intermediates at common fragile sites during mitosis.* Nat Cell Biol, 2013. **15**(8): p. 1008-15.
88. Ying, S., et al., *MUS81 promotes common fragile site expression.* Nat Cell Biol, 2013. **15**(8): p. 1001-7.
89. Sarlos, K., et al., *Knotty Problems during Mitosis: Mechanistic Insight into the Processing of Ultrafine DNA Bridges in Anaphase.* Cold Spring Harb Symp Quant Biol, 2017. **82**: p. 187-195.
90. Barefield, C. and J. Karlseder, *The BLM helicase contributes to telomere maintenance through processing of late-replicating intermediate structures.* Nucleic Acids Res, 2012. **40**(15): p. 7358-67.
91. Maciejowski, J., et al., *Chromothripsis and Kataegis Induced by Telomere Crisis.* Cell, 2015. **163**(7): p. 1641-54.
92. Nera, B., et al., *Elevated levels of TRF2 induce telomeric ultrafine anaphase bridges and rapid telomere deletions.* Nat Commun, 2015. **6**: p. 10132.
93. Chan, Y.W., K. Fugger, and S.C. West, *Unresolved recombination intermediates lead to ultra-fine anaphase bridges, chromosome breaks and aberrations.* Nat Cell Biol, 2018. **20**(1): p. 92-103.
94. Hengeveld, R.C., et al., *Rif1 Is Required for Resolution of Ultrafine DNA Bridges in Anaphase to Ensure Genomic Stability.* Dev Cell, 2015. **34**(4): p. 466-74.
95. Singleton, M.R., M.S. Dillingham, and D.B. Wigley, *Structure and mechanism of helicases and nucleic acid translocases.* Annu Rev Biochem, 2007. **76**: p. 23-50.
96. Biebricher, A., et al., *PICH: a DNA translocase specially adapted for processing anaphase bridge DNA.* Mol Cell, 2013. **51**(5): p. 691-701.
97. Chen, X.B., et al., *Human Mus81-associated endonuclease cleaves Holliday junctions in vitro.* Mol Cell, 2001. **8**(5): p. 1117-27.
98. Ciccia, A., A. Constantinou, and S.C. West, *Identification and characterization of the human mus81-eme1 endonuclease.* J Biol Chem, 2003. **278**(27): p. 25172-8.
99. Fekairi, S., et al., *Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases.* Cell, 2009. **138**(1): p. 78-89.
100. Munoz, I.M., et al., *Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair.* Mol Cell, 2009. **35**(1): p. 116-27.
101. Svendsen, J.M., et al., *Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair.* Cell, 2009. **138**(1): p. 63-77.
102. Ip, S.C., et al., *Identification of Holliday junction resolvases from humans and yeast.* Nature, 2008. **456**(7220): p. 357-61.
103. Rass, U., et al., *Mechanism of Holliday junction resolution by the human GEN1 protein.* Genes Dev, 2010. **24**(14): p. 1559-69.
104. Chan, Y.W. and S. West, *GEN1 promotes Holliday junction resolution by a coordinated nick and counter-nick mechanism.* Nucleic Acids Res, 2015. **43**(22): p. 10882-92.
105. Wu, L. and I.D. Hickson, *The Bloom's syndrome helicase suppresses crossing over during homologous recombination.* Nature, 2003. **426**(6968): p. 870-4.
106. Wyatt, H.D., et al., *The SMX DNA Repair Tri-nuclease.* Mol Cell, 2017. **65**(5): p. 848-860 e11.
107. Wyatt, H.D., et al., *Coordinated actions of SLX1-SLX4 and MUS81-EME1 for Holliday junction resolution in human cells.* Mol Cell, 2013. **52**(2): p. 234-47.

108. Duda, H., et al., *A Mechanism for Controlled Breakage of Under-replicated Chromosomes during Mitosis.* Dev Cell, 2017. **40**(4): p. 421-422.
109. Matos, J., et al., *Regulatory control of the resolution of DNA recombination intermediates during meiosis and mitosis.* Cell, 2011. **147**(1): p. 158-72.
110. Chan, Y.W. and S.C. West, *Spatial control of the GEN1 Holliday junction resolvase ensures genome stability.* Nat Commun, 2014. **5**: p. 4844.
111. Mimori, T. and J.A. Hardin, *Mechanism of interaction between Ku protein and DNA.* J Biol Chem, 1986. **261**(22): p. 10375-9.
112. Costantini, S., et al., *Interaction of the Ku heterodimer with the DNA ligase IV/Xrcc4 complex and its regulation by DNA-PK.* DNA Repair (Amst), 2007. **6**(6): p. 712-22.
113. Sibanda, B.L., et al., *Crystal structure of an Xrcc4-DNA ligase IV complex.* Nat Struct Biol, 2001. **8**(12): p. 1015-9.
114. Grawunder, U., et al., *Requirement for an interaction of XRCC4 with DNA ligase IV for wild-type V(D)J recombination and DNA double-strand break repair in vivo.* J Biol Chem, 1998. **273**(38): p. 24708-14.
115. Grawunder, U., D. Zimmer, and M.R. Lieber, *DNA ligase IV binds to XRCC4 via a motif located between rather than within its BRCT domains.* Curr Biol, 1998. **8**(15): p. 873-6.
116. Bothmer, A., et al., *53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination.* J Exp Med, 2010. **207**(4): p. 855-65.
117. Botuyan, M.V., et al., *Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair.* Cell, 2006. **127**(7): p. 1361-73.
118. Lu, R. and G.G. Wang, *Tudor: a versatile family of histone methylation 'readers'.* Trends Biochem Sci, 2013. **38**(11): p. 546-55.
119. Daley, J.M. and P. Sung, *53BP1, BRCA1, and the choice between recombination and end joining at DNA double-strand breaks.* Mol Cell Biol, 2014. **34**(8): p. 1380-8.
120. Boulton, S.J. and S.P. Jackson, *Identification of a Saccharomyces cerevisiae Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance.* Nucleic Acids Res, 1996. **24**(23): p. 4639-48.
121. Wilson, T.E., U. Grawunder, and M.R. Lieber, *Yeast DNA ligase IV mediates non-homologous DNA end joining.* Nature, 1997. **388**(6641): p. 495-8.
122. Frit, P., et al., *Alternative end-joining pathway(s): bricolage at DNA breaks.* DNA Repair (Amst), 2014. **17**: p. 81-97.
123. Verkaik, N.S., et al., *Different types of V(D)J recombination and end-joining defects in DNA double-strand break repair mutant mammalian cells.* Eur J Immunol, 2002. **32**(3): p. 701-9.
124. Della-Maria, J., et al., *Human Mre11/human Rad50/Nbs1 and DNA ligase IIIalpha/XRCC1 protein complexes act together in an alternative nonhomologous end joining pathway.* J Biol Chem, 2011. **286**(39): p. 33845-53.
125. Hassa, P.O., et al., *Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going?* Microbiol Mol Biol Rev, 2006. **70**(3): p. 789-829.
126. Caldecott, K.W., *Single-strand break repair and genetic disease.* Nat Rev Genet, 2008. **9**(8): p. 619-31.
127. Haince, J.F., et al., *PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites.* J Biol Chem, 2008. **283**(2): p. 1197-208.

128. Ahel, D., et al., *Poly(ADP-ribose)-dependent regulation of DNA repair by the chromatin remodeling enzyme ALC1.* Science, 2009. **325**(5945): p. 1240-3.
129. Audebert, M., B. Salles, and P. Calsou, *Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining.* J Biol Chem, 2004. **279**(53): p. 55117-26.
130. 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.
131. Cheng, Q., et al., *Ku counteracts mobilization of PARP1 and MRN in chromatin damaged with DNA double-strand breaks.* Nucleic Acids Res, 2011. **39**(22): p. 9605-19.
132. Audebert, M., et al., *Involvement of polynucleotide kinase in a poly(ADP-ribose) polymerase-1-dependent DNA double-strand breaks rejoining pathway.* J Mol Biol, 2006. **356**(2): p. 257-65.
133. Huertas, P., *DNA resection in eukaryotes: deciding how to fix the break.* Nat Struct Mol Biol, 2010. **17**(1): p. 11-6.
134. Truong, L.N., et al., *Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells.* Proc Natl Acad Sci U S A, 2013. **110**(19): p. 7720-5.
135. Geuting, V., C. Reul, and M. Lobrich, *ATM release at resected double-strand breaks provides heterochromatin reconstitution to facilitate homologous recombination.* PLoS Genet, 2013. **9**(8): p. e1003667.
136. Wu, W., et al., *Repair of radiation induced DNA double strand breaks by backup NHEJ is enhanced in G2.* DNA Repair (Amst), 2008. **7**(2): p. 329-38.
137. Chang, H.H.Y., et al., *Non-homologous DNA end joining and alternative pathways to double-strand break repair.* Nat Rev Mol Cell Biol, 2017. **18**(8): p. 495-506.
138. Wyatt, D.W., et al., *Essential Roles for Polymerase theta-Mediated End Joining in the Repair of Chromosome Breaks.* Mol Cell, 2016. **63**(4): p. 662-673.
139. Nick McElhinny, S.A., et al., *A gradient of template dependence defines distinct biological roles for family X polymerases in nonhomologous end joining.* Mol Cell, 2005. **19**(3): p. 357-66.
140. Crespan, E., et al., *Microhomology-mediated DNA strand annealing and elongation by human DNA polymerases lambda and beta on normal and repetitive DNA sequences.* Nucleic Acids Res, 2012. **40**(12): p. 5577-90.
141. Ahmad, A., et al., *ERCC1-XPF endonuclease facilitates DNA double-strand break repair.* Mol Cell Biol, 2008. **28**(16): p. 5082-92.
142. Tseng, H.M. and A.E. Tomkinson, *Processing and joining of DNA ends coordinated by interactions among Dnl4/Lif1, Pol4, and FEN-1.* J Biol Chem, 2004. **279**(46): p. 47580-8.
143. Weinfeld, M., et al., *Tidying up loose ends: the role of polynucleotide kinase/phosphatase in DNA strand break repair.* Trends Biochem Sci, 2011. **36**(5): p. 262-71.
144. Lieber, M.R., *The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway.* Annu Rev Biochem, 2010. **79**: p. 181-211.
145. Wang, H., et al., *DNA ligase III as a candidate component of backup pathways of nonhomologous end joining.* Cancer Res, 2005. **65**(10): p. 4020-30.
146. Liang, L., et al., *Human DNA ligases I and III, but not ligase IV, are required for microhomology-mediated end joining of DNA double-strand breaks.* Nucleic Acids Res, 2008. **36**(10): p. 3297-310.

147. Simsek, D., et al., *DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation.* PLoS Genet, 2011. **7**(6): p. e1002080.
148. Caldecott, K.W., *XRCC1 and DNA strand break repair.* DNA Repair (Amst), 2003. **2**(9): p. 955-69.
149. Han, L., W. Mao, and K. Yu, *X-ray repair cross-complementing protein 1 (XRCC1) deficiency enhances class switch recombination and is permissive for alternative end joining.* Proc Natl Acad Sci U S A, 2012. **109**(12): p. 4604-8.
150. Paul, K., et al., *DNA ligases I and III cooperate in alternative non-homologous end-joining in vertebrates.* PLoS One, 2013. **8**(3): p. e59505.
151. Huertas, P. and S.P. Jackson, *Human CtIP mediates cell cycle control of DNA end resection and double strand break repair.* J Biol Chem, 2009. **284**(14): p. 9558-65.
152. Chapman, J.R., M.R. Taylor, and S.J. Boulton, *Playing the end game: DNA double-strand break repair pathway choice.* Mol Cell, 2012. **47**(4): p. 497-510.
153. Ira, G., et al., *DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1.* Nature, 2004. **431**(7011): p. 1011-7.
154. Zou, L. and S.J. Elledge, *Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes.* Science, 2003. **300**(5625): p. 1542-8.
155. Xia, B., et al., *Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2.* Mol Cell, 2006. **22**(6): p. 719-29.
156. Zhang, F., et al., *PALB2 functionally connects the breast cancer susceptibility proteins BRCA1 and BRCA2.* Mol Cancer Res, 2009. **7**(7): p. 1110-8.
157. Zhang, F., et al., *PALB2 links BRCA1 and BRCA2 in the DNA-damage response.* Curr Biol, 2009. **19**(6): p. 524-9.
158. Sy, S.M., M.S. Huen, and J. Chen, *PALB2 is an integral component of the BRCA complex required for homologous recombination repair.* Proc Natl Acad Sci U S A, 2009. **106**(17): p. 7155-60.
159. 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): p. a016600.
160. Barber, L.J., et al., *RTEL1 maintains genomic stability by suppressing homologous recombination.* Cell, 2008. **135**(2): p. 261-71.
161. Matos, J. and S.C. West, *Holliday junction resolution: regulation in space and time.* DNA Repair (Amst), 2014. **19**: p. 176-81.
162. Morrical, S.W., *DNA-pairing and annealing processes in homologous recombination and homology-directed repair.* Cold Spring Harb Perspect Biol, 2015. **7**(2): p. a016444.
163. Holliday, R., *A mechanism for gene conversion in fungi.* Genet Res, 2007. **89**(5-6): p. 285-307.
164. San Filippo, J., P. Sung, and H. Klein, *Mechanism of eukaryotic homologous recombination.* Annu Rev Biochem, 2008. **77**: p. 229-57.
165. Heyer, W.D., K.T. Ehmsen, and J. Liu, *Regulation of homologous recombination in eukaryotes.* Annu Rev Genet, 2010. **44**: p. 113-39.
166. Bzymek, M., et al., *Double Holliday junctions are intermediates of DNA break repair.* Nature, 2010. **464**(7290): p. 937-41.
167. Bell, L. and B. Byers, *Occurrence of crossed strand-exchange forms in yeast DNA during meiosis.* Proc Natl Acad Sci U S A, 1979. **76**(7): p. 3445-9.
168. Petronczki, M., M.F. Siomos, and K. Nasmyth, *Un menage a quatre: the molecular biology of chromosome segregation in meiosis.* Cell, 2003. **112**(4): p. 423-40.

169. Bishop, D.K. and D. Zickler, *Early decision; meiotic crossover interference prior to stable strand exchange and synapsis.* Cell, 2004. **117**(1): p. 9-15.
170. Ellis, N.A., et al., *The Bloom's syndrome gene product is homologous to RecQ helicases.* Cell, 1995. **83**(4): p. 655-66.
171. Cejka, P., et al., *Rmi1 stimulates decatenation of double Holliday junctions during dissolution by Sgs1-Top3.* Nat Struct Mol Biol, 2010. **17**(11): p. 1377-82.
172. Ray, J.H. and J. German, *Bloom's syndrome and EM9 cells in BrdU-containing medium exhibit similarly elevated frequencies of sister chromatid exchange but dissimilar amounts of cellular proliferation and chromosome disruption.* Chromosoma, 1984. **90**(5): p. 383-8.
173. Macheret, M. and T.D. Halazonetis, *Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress.* Nature, 2018. **555**(7694): p. 112-116.
174. Macheret, M. and T.D. Halazonetis, *DNA replication stress as a hallmark of cancer.* Annu Rev Pathol, 2015. **10**: p. 425-48.
175. Pacek, M. and J.C. Walter, *A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication.* EMBO J, 2004. **23**(18): p. 3667-3676.
176. Petermann, E., et al., *Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair.* Mol Cell, 2010. **37**(4): p. 492-502.
177. Shi, Y., et al., *Histone demethylation mediated by the nuclear amine oxidase homolog LSD1.* Cell, 2004. **119**(7): p. 941-53.
178. Brooks, P.J. and J.A. Theruvathu, *DNA adducts from acetaldehyde: implications for alcohol-related carcinogenesis.* Alcohol, 2005. **35**(3): p. 187-93.
179. McMurray, C.T., *Mechanisms of trinucleotide repeat instability during human development.* Nat Rev Genet, 2010. **11**(11): p. 786-99.
180. Kim, J.C. and S.M. Mirkin, *The balancing act of DNA repeat expansions.* Curr Opin Genet Dev, 2013. **23**(3): p. 280-8.
181. Bochman, M.L., K. Paeschke, and V.A. Zakian, *DNA secondary structures: stability and function of G-quadruplex structures.* Nat Rev Genet, 2012. **13**(11): p. 770-80.
182. Helmrich, A., et al., *Transcription-replication encounters, consequences and genomic instability.* Nat Struct Mol Biol, 2013. **20**(4): p. 412-8.
183. Bermejo, R., M.S. Lai, and M. Foiani, *Preventing replication stress to maintain genome stability: resolving conflicts between replication and transcription.* Mol Cell, 2012. **45**(6): p. 710-8.
184. Aguilera, A. and T. Garcia-Muse, *R loops: from transcription byproducts to threats to genome stability.* Mol Cell, 2012. **46**(2): p. 115-24.
185. Masai, H., et al., *Eukaryotic chromosome DNA replication: where, when, and how?* Annu Rev Biochem, 2010. **79**: p. 89-130.
186. Woodward, A.M., et al., *Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress.* J Cell Biol, 2006. **173**(5): p. 673-83.
187. Ge, X.Q., D.A. Jackson, and J.J. Blow, *Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress.* Genes Dev, 2007. **21**(24): p. 3331-41.
188. McIntosh, D. and J.J. Blow, *Dormant origins, the licensing checkpoint, and the response to replicative stresses.* Cold Spring Harb Perspect Biol, 2012. **4**(10).
189. Frouin, I., et al., *DNA replication: a complex matter.* EMBO Rep, 2003. **4**(7): p. 666-70.

190. 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.
191. MacDougall, C.A., et al., *The structural determinants of checkpoint activation.* Genes Dev, 2007. **21**(8): p. 898-903.
192. Nam, E.A. and D. Cortez, *ATR signalling: more than meeting at the fork.* Biochem J, 2011. **436**(3): p. 527-36.
193. Labib, K. and G. De Piccoli, *Surviving chromosome replication: the many roles of the S-phase checkpoint pathway.* Philos Trans R Soc Lond B Biol Sci, 2011. **366**(1584): p. 3554-61.
194. Petermann, E. and T. Helleday, *Pathways of mammalian replication fork restart.* Nat Rev Mol Cell Biol, 2010. **11**(10): p. 683-7.
195. Elvers, I., et al., *UV stalled replication forks restart by re-priming in human fibroblasts.* Nucleic Acids Res, 2011. **39**(16): p. 7049-57.
196. Lopes, M., M. Foiani, and J.M. Sogo, *Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions.* Mol Cell, 2006. **21**(1): p. 15-27.
197. Mailand, N., I. Gibbs-Seymour, and S. Bekker-Jensen, *Regulation of PCNA-protein interactions for genome stability.* Nat Rev Mol Cell Biol, 2013. **14**(5): p. 269-82.
198. Masutani, C., et al., *The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta.* Nature, 1999. **399**(6737): p. 700-4.
199. Johnson, R.E., et al., *hRAD30 mutations in the variant form of xeroderma pigmentosum.* Science, 1999. **285**(5425): p. 263-5.
200. Lehman, A.R., *Postreplication repair of DNA in ultraviolet-irradiated mammalian cells.* J Mol Biol, 1972. **66**: p. 319-337.
201. Heller, R.C. and K.J. Marians, *Replication fork reactivation downstream of a blocked nascent leading strand.* Nature, 2006. **439**(7076): p. 557-62.
202. Indiani, C. and M. O'Donnell, *The replication clamp-loading machine at work in the three domains of life.* Nat Rev Mol Cell Biol, 2006. **7**(10): p. 751-61.
203. Ivanov, I., et al., *Proliferating cell nuclear antigen loaded onto double-stranded DNA: dynamics, minor groove interactions and functional implications.* Nucleic Acids Res, 2006. **34**(20): p. 6023-33.
204. Moldovan, G.L., B. Pfander, and S. Jentsch, *PCNA, the maestro of the replication fork.* Cell, 2007. **129**(4): p. 665-79.
205. Gulbis, J.M., et al., *Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA.* Cell, 1996. **87**(2): p. 297-306.
206. Bruning, J.B. and Y. Shamoo, *Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1.* Structure, 2004. **12**(12): p. 2209-19.
207. Gilljam, K.M., et al., *Identification of a novel, widespread, and functionally important PCNA-binding motif.* J Cell Biol, 2009. **186**(5): p. 645-54.
208. Sebesta, M., et al., *Structural insights into the function of ZRANB3 in replication stress response.* Nat Commun, 2017. **8**: p. 15847.
209. Hoege, C., et al., *RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO.* Nature, 2002. **419**(6903): p. 135-41.
210. Carlile, C.M., et al., *Synthesis of free and proliferating cell nuclear antigen-bound polyubiquitin chains by the RING E3 ubiquitin ligase Rad5.* J Biol Chem, 2009. **284**(43): p. 29326-34.
211. Parker, J.L. and H.D. Ulrich, *Mechanistic analysis of PCNA poly-ubiquitylation by the ubiquitin protein ligases Rad18 and Rad5.* EMBO J, 2009. **28**(23): p. 3657-66.

212. Motegi, A., et al., *Polyubiquitination of proliferating cell nuclear antigen by HLTF and SHPRH prevents genomic instability from stalled replication forks.* Proc Natl Acad Sci U S A, 2008. **105**(34): p. 12411-6.
213. Motegi, A., et al., *Human SHPRH suppresses genomic instability through proliferating cell nuclear antigen polyubiquitination.* J Cell Biol, 2006. **175**(5): p. 703-8.
214. Unk, I., et al., *Human HLTF functions as a ubiquitin ligase for proliferating cell nuclear antigen polyubiquitination.* Proc Natl Acad Sci U S A, 2008. **105**(10): p. 3768-73.
215. Unk, I., et al., *Human SHPRH is a ubiquitin ligase for Mms2-Ubc13-dependent polyubiquitylation of proliferating cell nuclear antigen.* Proc Natl Acad Sci U S A, 2006. **103**(48): p. 18107-12.
216. MacKay, C., R. Toth, and J. Rouse, *Biochemical characterisation of the SWI/SNF family member HLTF.* Biochem Biophys Res Commun, 2009. **390**(2): p. 187-91.
217. Chiu, R.K., et al., *Lysine 63-polyubiquitination guards against translesion synthesis-induced mutations.* PLoS Genet, 2006. **2**(7): p. e116.
218. Blastyak, A., et al., *Role of double-stranded DNA translocase activity of human HLTF in replication of damaged DNA.* Mol Cell Biol, 2010. **30**(3): p. 684-93.
219. Carvalho, C.M., et al., *Complex rearrangements in patients with duplications of MECP2 can occur by fork stalling and template switching.* Hum Mol Genet, 2009. **18**(12): p. 2188-203.
220. Lee, J.A., C.M. Carvalho, and J.R. Lupski, *A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders.* Cell, 2007. **131**(7): p. 1235-47.
221. Zhang, F., et al., *The DNA replication FoSTeS/MMBIR mechanism can generate genomic, genic and exonic complex rearrangements in humans.* Nat Genet, 2009. **41**(7): p. 849-53.
222. Llorente, B., C.E. Smith, and L.S. Symington, *Break-induced replication: what is it and what is it for?* Cell Cycle, 2008. **7**(7): p. 859-64.
223. Smith, C.E., B. Llorente, and L.S. Symington, *Template switching during break-induced replication.* Nature, 2007. **447**(7140): p. 102-5.
224. Ciccia, A., et al., *Polyubiquitinated PCNA recruits the ZRANB3 translocase to maintain genomic integrity after replication stress.* Mol Cell, 2012. **47**(3): p. 396-409.
225. Weston, R., H. Peeters, and D. Ahel, *ZRANB3 is a structure-specific ATP-dependent endonuclease involved in replication stress response.* Genes Dev, 2012. **26**(14): p. 1558-72.
226. Yuan, J., G. Ghosal, and J. Chen, *The HARP-like domain-containing protein AH2/ZRANB3 binds to PCNA and participates in cellular response to replication stress.* Mol Cell, 2012. **47**(3): p. 410-21.
227. Chanoux, R.A., et al., *ATR and H2AX cooperate in maintaining genome stability under replication stress.* J Biol Chem, 2009. **284**(9): p. 5994-6003.
228. Ciccia, A. and S.J. Elledge, *The DNA damage response: making it safe to play with knives.* Mol Cell, 2010. **40**(2): p. 179-204.
229. Sogo, J.M., M. Lopes, and M. Foiani, *Fork reversal and ssDNA accumulation at stalled replication forks owing to checkpoint defects.* Science, 2002. **297**(5581): p. 599-602.
230. Debatisse, M., et al., *Common fragile sites: mechanisms of instability revisited.* Trends Genet, 2012. **28**(1): p. 22-32.

231. Koundrioukoff, S., et al., *Stepwise activation of the ATR signaling pathway upon increasing replication stress impacts fragile site integrity.* PLoS Genet, 2013. **9**(7): p. e1003643.
232. De Koning, L., et al., *Histone chaperones: an escort network regulating histone traffic.* Nat Struct Mol Biol, 2007. **14**(11): p. 997-1007.
233. Deem, A.K., X. Li, and J.K. Tyler, *Epigenetic regulation of genomic integrity.* Chromosoma, 2012. **121**(2): p. 131-51.
234. Luijsterburg, M.S. and H. van Attikum, *Chromatin and the DNA damage response: the cancer connection.* Mol Oncol, 2011. **5**(4): p. 349-67.
235. Ransom, M., B.K. Dennehey, and J.K. Tyler, *Chaperoning histones during DNA replication and repair.* Cell, 2010. **140**(2): p. 183-95.
236. Soria, G., S.E. Polo, and G. Almouzni, *Prime, repair, restore: the active role of chromatin in the DNA damage response.* Mol Cell, 2012. **46**(6): p. 722-34.
237. Roberts, S.A. and D.A. Gordenin, *Hypermutation in human cancer genomes: footprints and mechanisms.* Nat Rev Cancer, 2014. **14**(12): p. 786-800.
238. Tsouroula, K., et al., *Temporal and Spatial Uncoupling of DNA Double Strand Break Repair Pathways within Mammalian Heterochromatin.* Mol Cell, 2016. **63**(2): p. 293-305.
239. Putnam, C.D., T.K. Hayes, and R.D. Kolodner, *Specific pathways prevent duplication-mediated genome rearrangements.* Nature, 2009. **460**(7258): p. 984-9.
240. Sasaki, M., J. Lange, and S. Keeney, *Genome destabilization by homologous recombination in the germ line.* Nat Rev Mol Cell Biol, 2010. **11**(3): p. 182-95.
241. de Lange, T., *Shelterin: the protein complex that shapes and safeguards human telomeres.* Genes Dev, 2005. **19**(18): p. 2100-10.
242. Greider, C.W. and E.H. Blackburn, *Identification of a specific telomere terminal transferase activity in Tetrahymena extracts.* Cell, 1985. **43**(2 Pt 1): p. 405-13.
243. Harley, C.B., A.B. Futcher, and C.W. Greider, *Telomeres shorten during ageing of human fibroblasts.* Nature, 1990. **345**(6274): p. 458-60.
244. Shippen-Lentz, D. and E.H. Blackburn, *Functional evidence for an RNA template in telomerase.* Science, 1990. **247**(4942): p. 546-52.
245. Celli, G.B. and T. de Lange, *DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion.* Nat Cell Biol, 2005. **7**(7): p. 712-8.
246. Denchi, E.L. and T. de Lange, *Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1.* Nature, 2007. **448**(7157): p. 1068-71.
247. Karlseder, J., et al., *p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2.* Science, 1999. **283**(5406): p. 1321-5.
248. Palm, W. and T. de Lange, *How shelterin protects mammalian telomeres.* Annu Rev Genet, 2008. **42**: p. 301-34.
249. Sfeir, A. and T. de Lange, *Removal of shelterin reveals the telomere end-protection problem.* Science, 2012. **336**(6081): p. 593-7.
250. Takai, K.K., et al., *Telomere protection by TPP1/POT1 requires tethering to TIN2.* Mol Cell, 2011. **44**(4): p. 647-59.
251. van Steensel, B., A. Smogorzewska, and T. de Lange, *TRF2 protects human telomeres from end-to-end fusions.* Cell, 1998. **92**(3): p. 401-13.
252. Doksani, Y., et al., *Super-resolution fluorescence imaging of telomeres reveals TRF2-dependent T-loop formation.* Cell, 2013. **155**(2): p. 345-356.
253. Griffith, J.D., et al., *Mammalian telomeres end in a large duplex loop.* Cell, 1999. **97**(4): p. 503-14.

254. O'Sullivan, R.J. and J. Karlseder, *Telomeres: protecting chromosomes against genome instability.* Nat Rev Mol Cell Biol, 2010. **11**(3): p. 171-81.
255. Sarek, G., et al., *TRF2 recruits RTEL1 to telomeres in S phase to promote t-loop unwinding.* Mol Cell, 2015. **57**(4): p. 622-635.
256. Vannier, J.B., et al., *RTEL1 dismantles T loops and counteracts telomeric G4-DNA to maintain telomere integrity.* Cell, 2012. **149**(4): p. 795-806.
257. 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.
258. Kruk, P.A., N.J. Rampino, and V.A. Bohr, *DNA damage and repair in telomeres: relation to aging.* Proc Natl Acad Sci U S A, 1995. **92**(1): p. 258-62.
259. Petersen, S., G. Saretzki, and T. von Zglinicki, *Preferential accumulation of single-stranded regions in telomeres of human fibroblasts.* Exp Cell Res, 1998. **239**(1): p. 152-60.
260. Fumagalli, M., et al., *Telomeric DNA damage is irreparable and causes persistent DNA-damage-response activation.* Nat Cell Biol, 2012. **14**(4): p. 355-65.
261. Sun, L., et al., *Targeted DNA damage at individual telomeres disrupts their integrity and triggers cell death.* Nucleic Acids Res, 2015. **43**(13): p. 6334-47.
262. Mao, P., et al., *Homologous recombination-dependent repair of telomeric DSBs in proliferating human cells.* Nat Commun, 2016. **7**: p. 12154.
263. Cho, N.W., et al., *Interchromosomal homology searches drive directional ALT telomere movement and synapsis.* Cell, 2014. **159**(1): p. 108-121.
264. Tarsounas, M., et al., *Telomere maintenance requires the RAD51D recombination/repair protein.* Cell, 2004. **117**(3): p. 337-47.
265. Verdun, R.E. and J. Karlseder, *The DNA damage machinery and homologous recombination pathway act consecutively to protect human telomeres.* Cell, 2006. **127**(4): p. 709-20.
266. Chesi, M., et al., *Frequent dysregulation of the c-maf proto-oncogene at 16q23 by translocation to an Ig locus in multiple myeloma.* Blood, 1998. **91**(12): p. 4457-63.
267. Finnis, M., et al., *Common chromosomal fragile site FRA16D mutation in cancer cells.* Hum Mol Genet, 2005. **14**(10): p. 1341-9.
268. O'Keefe, L.V. and R.I. Richards, *Common chromosomal fragile sites and cancer: focus on FRA16D.* Cancer Lett, 2006. **232**(1): p. 37-47.
269. Arlt, M.F., A.M. Casper, and T.W. Glover, *Common fragile sites.* Cytogenet Genome Res, 2003. **100**(1-4): p. 92-100.
270. Glover, T.W., *Common fragile sites.* Cancer Lett, 2006. **232**(1): p. 4-12.
271. 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.* Hum Mol Genet, 1998. **7**(4): p. 755-61.
272. Lucas, I., et al., *High-throughput mapping of origins of replication in human cells.* EMBO Rep, 2007. **8**(8): p. 770-7.
273. Shah, S.N., et al., *DNA structure and the Werner protein modulate human DNA polymerase delta-dependent replication dynamics within the common fragile site FRA16D.* Nucleic Acids Res, 2010. **38**(4): p. 1149-62.
274. Zhang, H. and C.H. Freudenreich, *An AT-rich sequence in human common fragile site FRA16D causes fork stalling and chromosome breakage in S. cerevisiae.* Mol Cell, 2007. **27**(3): p. 367-79.
275. Lukas, C., et al., *53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress.* Nat Cell Biol, 2011. **13**(3): p. 243-53.

276. Durkin, S.G. and T.W. Glover, *Chromosome fragile sites.* Annu Rev Genet, 2007. **41**: p. 169-92.
277. Bhowmick, R., S. Minocherhomji, and I.D. Hickson, *RAD52 Facilitates Mitotic DNA Synthesis Following Replication Stress.* Mol Cell, 2016. **64**(6): p. 1117-1126.
278. Minocherhomji, S., et al., *Replication stress activates DNA repair synthesis in mitosis.* Nature, 2015. **528**(7581): p. 286-90.
279. Schlacher, K., et al., *Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11.* Cell, 2011. **145**(4): p. 529-42.
280. Schlacher, K., H. Wu, and M. Jasin, *A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2.* Cancer Cell, 2012. **22**(1): p. 106-16.
281. Di Marco, S., et al., *RECQ5 Helicase Cooperates with MUS81 Endonuclease in Processing Stalled Replication Forks at Common Fragile Sites during Mitosis.* Mol Cell, 2017. **66**(5): p. 658-671 e8.
282. Howlett, N.G., et al., *The Fanconi anemia pathway is required for the DNA replication stress response and for the regulation of common fragile site stability.* Hum Mol Genet, 2005. **14**(5): p. 693-701.
283. Howlett, N.G., et al., *Biallelic inactivation of BRCA2 in Fanconi anemia.* Science, 2002. **297**(5581): p. 606-9.
284. Hirsch, B., et al., *Association of biallelic BRCA2/FANCD1 mutations with spontaneous chromosomal instability and solid tumors of childhood.* Blood, 2004. **103**(7): p. 2554-9.
285. Karanja, K.K., et al., *Preventing over-resection by DNA2 helicase/nuclease suppresses repair defects in Fanconi anemia cells.* Cell Cycle, 2014. **13**(10): p. 1540-50.
286. Taniguchi, T., et al., *S-phase-specific interaction of the Fanconi anemia protein, FANCD2, with BRCA1 and RAD51.* Blood, 2002. **100**(7): p. 2414-20.
287. Madireddy, A., et al., *FANCD2 Facilitates Replication through Common Fragile Sites.* Mol Cell, 2016. **64**(2): p. 388-404.
288. Wang, H., et al., *The concerted roles of FANCM and Rad52 in the protection of common fragile sites.* Nat Commun, 2018. **9**(1): p. 2791.
289. Zeitlin, S.G., et al., *Double-strand DNA breaks recruit the centromeric histone CENP-A.* Proc Natl Acad Sci U S A, 2009. **106**(37): p. 15762-7.
290. Aze, A., et al., *Centromeric DNA replication reconstitution reveals DNA loops and ATR checkpoint suppression.* Nat Cell Biol, 2016. **18**(6): p. 684-91.
291. Cleveland, D.W., Y. Mao, and K.F. Sullivan, *Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling.* Cell, 2003. **112**(4): p. 407-21.
292. McKinley, K.L. and I.M. Cheeseman, *The molecular basis for centromere identity and function.* Nat Rev Mol Cell Biol, 2016. **17**(1): p. 16-29.
293. Verdaasdonk, J.S. and K. Bloom, *Centromeres: unique chromatin structures that drive chromosome segregation.* Nat Rev Mol Cell Biol, 2011. **12**(5): p. 320-32.
294. Choo, K.H., *Centromere DNA dynamics: latent centromeres and neocentromere formation.* Am J Hum Genet, 1997. **61**(6): p. 1225-33.
295. Fukagawa, T. and W.C. Earnshaw, *The centromere: chromatin foundation for the kinetochore machinery.* Dev Cell, 2014. **30**(5): p. 496-508.
296. Steiner, F.A. and S. Henikoff, *Diversity in the organization of centromeric chromatin.* Curr Opin Genet Dev, 2015. **31**: p. 28-35.
297. Schueler, M.G., et al., *Genomic and genetic definition of a functional human centromere.* Science, 2001. **294**(5540): p. 109-15.

298. Rudd, M.K., G.A. Wray, and H.F. Willard, *The evolutionary dynamics of alpha-satellite.* Genome Res, 2006. **16**(1): p. 88-96.
299. Barnhart, M.C., et al., *HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore.* J Cell Biol, 2011. **194**(2): p. 229-43.
300. Guse, A., et al., *In vitro centromere and kinetochore assembly on defined chromatin templates.* Nature, 2011. **477**(7364): p. 354-8.
301. Hori, T., et al., *The CCAN recruits CENP-A to the centromere and forms the structural core for kinetochore assembly.* J Cell Biol, 2013. **200**(1): p. 45-60.
302. Mendiburo, M.J., et al., *Drosophila CENH3 is sufficient for centromere formation.* Science, 2011. **334**(6056): p. 686-90.
303. Giunta, S. and H. Funabiki, *Integrity of the human centromere DNA repeats is protected by CENP-A, CENP-C, and CENP-T.* Proc Natl Acad Sci U S A, 2017. **114**(8): p. 1928-1933.
304. Mirkin, S.M., *Expandable DNA repeats and human disease.* Nature, 2007. **447**(7147): p. 932-40.
305. Zhao, J., et al., *Non-B DNA structure-induced genetic instability and evolution.* Cell Mol Life Sci, 2010. **67**(1): p. 43-62.
306. Grady, D.L., et al., *Highly conserved repetitive DNA sequences are present at human centromeres.* Proc Natl Acad Sci U S A, 1992. **89**(5): p. 1695-9.
307. Zhu, L., S.H. Chou, and B.R. Reid, *A single G-to-C change causes human centromere TGGAA repeats to fold back into hairpins.* Proc Natl Acad Sci U S A, 1996. **93**(22): p. 12159-64.
308. Li, Z., et al., *hDNA2 nuclease/helicase promotes centromeric DNA replication and genome stability.* EMBO J, 2018. **37**(14).
309. Errico, A. and V. Costanzo, *Mechanisms of replication fork protection: a safeguard for genome stability.* Crit Rev Biochem Mol Biol, 2012. **47**(3): p. 222-35.
310. Bae, S.H., et al., *RPA governs endonuclease switching during processing of Okazaki fragments in eukaryotes.* Nature, 2001. **412**(6845): p. 456-61.
311. Copeland, W.C. and M.J. Longley, *DNA2 resolves expanding flap in mitochondrial base excision repair.* Mol Cell, 2008. **32**(4): p. 457-8.
312. Kao, H.I., J.L. Campbell, and R.A. Bambara, *Dna2p helicase/nuclease is a tracking protein, like FEN1, for flap cleavage during Okazaki fragment maturation.* J Biol Chem, 2004. **279**(49): p. 50840-9.
313. Lee, C.H., et al., *The N-terminal 45-kDa domain of Dna2 endonuclease/helicase targets the enzyme to secondary structure DNA.* J Biol Chem, 2013. **288**(13): p. 9468-81.
314. Lin, W., et al., *Mammalian DNA2 helicase/nuclease cleaves G-quadruplex DNA and is required for telomere integrity.* EMBO J, 2013. **32**(10): p. 1425-39.
315. Stewart, J.A., J.L. Campbell, and R.A. Bambara, *Dna2 is a structure-specific nuclease, with affinity for 5'-flap intermediates.* Nucleic Acids Res, 2010. **38**(3): p. 920-30.
316. Duxin, J.P., et al., *Okazaki fragment processing-independent role for human Dna2 enzyme during DNA replication.* J Biol Chem, 2012. **287**(26): p. 21980-91.
317. Pinto, C., et al., *Human DNA2 possesses a cryptic DNA unwinding activity that functionally integrates with BLM or WRN helicases.* Elife, 2016. **5**.
318. Budd, M.E. and J.L. Campbell, *A yeast gene required for DNA replication encodes a protein with homology to DNA helicases.* Proc Natl Acad Sci U S A, 1995. **92**(17): p. 7642-6.

319. Budd, M.E., W.C. Choe, and J.L. Campbell, *DNA2 encodes a DNA helicase essential for replication of eukaryotic chromosomes.* J Biol Chem, 1995. **270**(45): p. 26766-9.
320. Flaus, A., et al., *Identification of multiple distinct Snf2 subfamilies with conserved structural motifs.* Nucleic Acids Res, 2006. **34**(10): p. 2887-905.
321. Flaus, A. and T. Owen-Hughes, *Mechanisms for ATP-dependent chromatin remodelling: the means to the end.* FEBS J, 2011. **278**(19): p. 3579-95.
322. Hauk, G. and G.D. Bowman, *Structural insights into regulation and action of SWI2/SNF2 ATPases.* Curr Opin Struct Biol, 2011. **21**(6): p. 719-27.
323. Hopfner, K.P., et al., *Swi2/Snf2 remodelers: hybrid views on hybrid molecular machines.* Curr Opin Struct Biol, 2012. **22**(2): p. 225-33.
324. Velez-Cruz, R. and J.M. Egly, *Cockayne syndrome group B (CSB) protein: at the crossroads of transcriptional networks.* Mech Ageing Dev, 2013. **134**(5-6): p. 234-42.
325. Lin, J.R., et al., *SHPRH and HLTF act in a damage-specific manner to coordinate different forms of postreplication repair and prevent mutagenesis.* Mol Cell, 2011. **42**(2): p. 237-49.
326. Bansbach, C.E., et al., *The annealing helicase SMARCAL1 maintains genome integrity at stalled replication forks.* Genes Dev, 2009. **23**(20): p. 2405-14.
327. Ciccia, A., et al., *The SIOD disorder protein SMARCAL1 is an RPA-interacting protein involved in replication fork restart.* Genes Dev, 2009. **23**(20): p. 2415-25.
328. Yusufzai, T. and J.T. Kadonaga, *HARP is an ATP-driven annealing helicase.* Science, 2008. **322**(5902): p. 748-50.
329. Kokavec, J., et al., *Chromatin remodeling and SWI/SNF2 factors in human disease.* Front Biosci, 2008. **13**: p. 6126-34.
330. Picketts, D.J., et al., *ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome.* Hum Mol Genet, 1996. **5**(12): p. 1899-907.
331. Mallery, D.L., et al., *Molecular analysis of mutations in the CSB (ERCC6) gene in patients with Cockayne syndrome.* Am J Hum Genet, 1998. **62**(1): p. 77-85.
332. Tummala, H., et al., *ERCC6L2 mutations link a distinct bone-marrow-failure syndrome to DNA repair and mitochondrial function.* Am J Hum Genet, 2014. **94**(2): p. 246-56.
333. Bluteau, O., et al., *A landscape of germ line mutations in a cohort of inherited bone marrow failure patients.* Blood, 2018. **131**(7): p. 717-732.
334. Jarviaho, T., et al., *Bone marrow failure syndrome caused by homozygous frameshift mutation in the ERCC6L2 gene.* Clin Genet, 2018. **93**(2): p. 392-395.
335. Shabanova, I., et al., *ERCC6L2-associated inherited bone marrow failure syndrome.* Mol Genet Genomic Med, 2018. **6**(3): p. 463-468.
336. Tummala, H., et al., *Genome instability is a consequence of transcription deficiency in patients with bone marrow failure harboring biallelic ERCC6L2 variants.* Proc Natl Acad Sci U S A, 2018. **115**(30): p. 7777-7782.
337. Zhang, S., et al., *A nonsense mutation in the DNA repair factor Hebo causes mild bone marrow failure and microcephaly.* J Exp Med, 2016. **213**(6): p. 1011-28.
338. Dokal, I. and T. Vulliamy, *Inherited bone marrow failure syndromes.* Haematologica, 2010. **95**(8): p. 1236-40.
339. Young, N.S., A. Bacigalupo, and J.C. Marsh, *Aplastic anemia: pathophysiology and treatment.* Biol Blood Marrow Transplant, 2010. **16**(1 Suppl): p. S119-25.

340. Garcia, C.K., W.E. Wright, and J.W. Shay, *Human diseases of telomerase dysfunction: insights into tissue aging.* Nucleic Acids Res, 2007. **35**(22): p. 7406-16.
341. Wong, C.C., et al., *Defective ribosome assembly in Shwachman-Diamond syndrome.* Blood, 2011. **118**(16): p. 4305-12.
342. Rujkijyanont, P., et al., *Bone marrow cells from patients with Shwachman-Diamond syndrome abnormally express genes involved in ribosome biogenesis and RNA processing.* Br J Haematol, 2009. **145**(6): p. 806-15.
343. Kottemann, M.C. and A. Smogorzewska, *Fanconi anaemia and the repair of Watson and Crick DNA crosslinks.* Nature, 2013. **493**(7432): p. 356-63.
344. Collins, J. and I. Dokal, *Inherited bone marrow failure syndromes.* Hematology, 2015. **20**(7): p. 433-4.
345. Lasko, P., *Tudor domain.* Curr Biol, 2010. **20**(16): p. R666-7.
346. Marchler-Bauer, A., et al., *CDD: NCBI's conserved domain database.* Nucleic Acids Res, 2015. **43**(Database issue): p. D222-6.
347. Limoli, C.L. and J.F. Ward, *A new method for introducing double-strand breaks into cellular DNA.* Radiat Res, 1993. **134**(2): p. 160-9.
348. Cong, L., et al., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**(6121): p. 819-23.
349. Jinek, M., et al., *RNA-programmed genome editing in human cells.* Elife, 2013. **2**: p. e00471.
350. 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): p. e79202.
351. Kudo, N., et al., *Leptomycin B inhibition of signal-mediated nuclear export by direct binding to CRM1.* Exp Cell Res, 1998. **242**(2): p. 540-7.
352. Wolff, B., J.J. Sanglier, and Y. Wang, *Leptomycin B is an inhibitor of nuclear export: inhibition of nucleo-cytoplasmic translocation of the human immunodeficiency virus type 1 (HIV-1) Rev protein and Rev-dependent mRNA.* Chem Biol, 1997. **4**(2): p. 139-47.
353. Mou, H., et al., *CRISPR/Cas9-mediated genome editing induces exon skipping by alternative splicing or exon deletion.* Genome Biol, 2017. **18**(1): p. 108.
354. Sleigh, M.J., *The mechanism of DNA breakage by phleomycin in vitro.* Nucleic Acids Res, 1976. **3**(4): p. 891-901.
355. Tounekti, O., et al., *The ratio of single- to double-strand DNA breaks and their absolute values determine cell death pathway.* Br J Cancer, 2001. **84**(9): p. 1272-9.
356. Regulus, P., et al., *Oxidation of the sugar moiety of DNA by ionizing radiation or bleomycin could induce the formation of a cluster DNA lesion.* Proc Natl Acad Sci U S A, 2007. **104**(35): p. 14032-7.
357. Cadet, J., T. Douki, and J.L. Ravanat, *Oxidatively generated base damage to cellular DNA.* Free Radic Biol Med, 2010. **49**(1): p. 9-21.
358. Fujiwara, Y., *Postreplication repair of alkylation damage to DNA of mammalian cells in culture.* Cancer Res, 1975. **35**(10): p. 2780-9.
359. Huyen, Y., et al., *Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks.* Nature, 2004. **432**: p. 406-411.
360. Drane, P., et al., *TIRR regulates 53BP1 by masking its histone methyl-lysine binding function.* Nature, 2017. **543**(7644): p. 211-216.

361. Zhang, A., et al., *The p53-binding protein 1-Tudor-interacting repair regulator complex participates in the DNA damage response.* J Biol Chem, 2017. **292**(16): p. 6461-6467.
362. Coverley, D., H. Laman, and R.A. Laskey, *Distinct roles for cyclins E and A during DNA replication complex assembly and activation.* Nat Cell Biol, 2002. **4**(7): p. 523-8.
363. Henglein, B., et al., *Structure and cell cycle-regulated transcription of the human cyclin A gene.* Proc Natl Acad Sci U S A, 1994. **91**(12): p. 5490-4.
364. Wang, L., et al., *Regulation of the Rhp26ERCC6/CSB chromatin remodeler by a novel conserved leucine latch motif.* Proc Natl Acad Sci U S A, 2014. **111**(52): p. 18566-71.
365. Wilson, D.B., et al., *Inherited bone marrow failure syndromes in adolescents and young adults.* Ann Med, 2014. **46**(6): p. 353-63.
366. Ke, Y., et al., *PICH and BLM limit histone association with anaphase centromeric DNA threads and promote their resolution.* EMBO J, 2011. **30**(16): p. 3309-21.
367. Bobkov, G.O.M., N. Gilbert, and P. Heun, *Centromere transcription allows CENP-A to transit from chromatin association to stable incorporation.* J Cell Biol, 2018. **217**(6): p. 1957-1972.
368. Wang, G. and K.M. Vasquez, *Effects of Replication and Transcription on DNA Structure-Related Genetic Instability.* Genes (Basel), 2017. **8**(1).
369. Hagleitner, M.M., et al., *Clinical spectrum of immunodeficiency, centromeric instability and facial dysmorphism (ICF syndrome).* J Med Genet, 2008. **45**(2): p. 93-9.
370. Pospiech, H., A.K. Rytkonen, and J.E. Syvaoja, *The role of DNA polymerase activity in human non-homologous end joining.* Nucleic Acids Res, 2001. **29**(15): p. 3277-88.
371. Strzalka, W. and A. Ziemienowicz, *Proliferating cell nuclear antigen (PCNA): a key factor in DNA replication and cell cycle regulation.* Ann Bot, 2011. **107**(7): p. 1127-40.
372. Shimamura, A. and B.P. Alter, *Pathophysiology and management of inherited bone marrow failure syndromes.* Blood Rev, 2010. **24**(3): p. 101-22.
373. Townsley, D.M., B. Dumitriu, and N.S. Young, *Bone marrow failure and the telomeropathies.* Blood, 2014. **124**(18): p. 2775-83.
374. Michl, J., J. Zimmer, and M. Tarsounas, *Interplay between Fanconi anemia and homologous recombination pathways in genome integrity.* EMBO J, 2016. **35**(9): p. 909-23.
375. Zhang, A., et al., *A protease pathway for the repair of topoisomerase II-DNA covalent complexes.* J Biol Chem, 2006. **281**(47): p. 35997-6003.
376. Azarova, A.M., et al., *Roles of DNA topoisomerase II isozymes in chemotherapy and secondary malignancies.* Proc Natl Acad Sci U S A, 2007. **104**(26): p. 11014-9.
377. Lee, Y.J., et al., *An in vivo analysis of MMC-induced DNA damage and its repair.* Carcinogenesis, 2006. **27**(3): p. 446-53.
378. Jaspers, N.G., et al., *Anti-tumour compounds illudin S and Irofulven induce DNA lesions ignored by global repair and exclusively processed by transcription- and replication-coupled repair pathways.* DNA Repair (Amst), 2002. **1**(12): p. 1027-38.
379. Koeppel, F., et al., *Irofulven cytotoxicity depends on transcription-coupled nucleotide excision repair and is correlated with XPG expression in solid tumor cells.* Clin Cancer Res, 2004. **10**(16): p. 5604-13.

380. Escargueil, A.E., et al., *Influence of irofulven, a transcription-coupled repair-specific antitumor agent, on RNA polymerase activity, stability and dynamics in living mammalian cells.* J Cell Sci, 2008. **121**(Pt 8): p. 1275-83.
381. Wang, J., et al., *ATM-dependent CHK2 activation induced by anticancer agent, irofulven.* J Biol Chem, 2004. **279**(38): p. 39584-92.
382. Wiltshire, T., et al., *BRCA1 contributes to cell cycle arrest and chemoresistance in response to the anticancer agent irofulven.* Mol Pharmacol, 2007. **71**(4): p. 1051-60.
383. Woynarowski, J.M., et al., *Effects on DNA integrity and apoptosis induction by a novel antitumor sesquiterpene drug, 6-hydroxymethylacylfulvene (HMAF, MGI 114).* Biochem Pharmacol, 1997. **54**(11): p. 1181-93.
384. Finnie, N., et al., *Transcription factor phosphorylation by the DNA-dependent protein kinase.* Biochem Soc Trans, 1993. **21**(4): p. 930-5.
385. Jackson, S., T. Gottlieb, and K. Hartley, *Phosphorylation of transcription factor Sp1 by the DNA-dependent protein kinase.* Adv Second Messenger Phosphoprotein Res, 1993. **28**: p. 279-86.
386. Kuhn, A., et al., *DNA-dependent protein kinase: a potent inhibitor of transcription by RNA polymerase I.* Genes Dev, 1995. **9**(2): p. 193-203.
387. Goodarzi, A.A., et al., *ATM signaling facilitates repair of DNA double-strand breaks associated with heterochromatin.* Mol Cell, 2008. **31**(2): p. 167-77.
388. Noon, A.T., et al., *53BP1-dependent robust localized KAP-1 phosphorylation is essential for heterochromatic DNA double-strand break repair.* Nat Cell Biol, 2010. **12**(2): p. 177-84.
389. Guirouilh-Barbat, J., et al., *Is homologous recombination really an error-free process?* Front Genet, 2014. **5**: p. 175.
390. Brunet, E., et al., *Chromosomal translocations induced at specified loci in human stem cells.* Proc Natl Acad Sci U S A, 2009. **106**(26): p. 10620-5.
391. Stark, J.M., et al., *Genetic steps of mammalian homologous repair with distinct mutagenic consequences.* Mol Cell Biol, 2004. **24**(21): p. 9305-16.
392. 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.
393. 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.
394. Rief, N. and M. Lobrich, *Efficient rejoining of radiation-induced DNA double-strand breaks in centromeric DNA of human cells.* J Biol Chem, 2002. **277**(23): p. 20572-82.
395. Wang, X., et al., *Role of DNA replication proteins in double-strand break-induced recombination in Saccharomyces cerevisiae.* Mol Cell Biol, 2004. **24**(16): p. 6891-9.
396. Li, X., et al., *PCNA is required for initiation of recombination-associated DNA synthesis by DNA polymerase delta.* Mol Cell, 2009. **36**(4): p. 704-13.
397. Li, J., D.L. Holzschu, and T. Sugiyama, *PCNA is efficiently loaded on the DNA recombination intermediate to modulate polymerase delta, eta, and zeta activities.* Proc Natl Acad Sci U S A, 2013. **110**(19): p. 7672-7.
398. Sebesta, M., et al., *Role of PCNA and TLS polymerases in D-loop extension during homologous recombination in humans.* DNA Repair (Amst), 2013. **12**(9): p. 691-8.

399. Balajee, A.S. and C.R. Geard, *Chromatin-bound PCNA complex formation triggered by DNA damage occurs independent of the ATM gene product in human cells.* Nucleic Acids Res, 2001. **29**(6): p. 1341-51.
400. Montecucco, A., et al., *DNA ligase I is recruited to sites of DNA replication by an interaction with proliferating cell nuclear antigen: identification of a common targeting mechanism for the assembly of replication factories.* EMBO J, 1998. **17**(13): p. 3786-95.
401. Lydeard, J.R., et al., *Break-induced replication and telomerase-independent telomere maintenance require Pol32.* Nature, 2007. **448**(7155): p. 820-3.
402. Mirkin, S.M., *DNA structures, repeat expansions and human hereditary disorders.* Curr Opin Struct Biol, 2006. **16**(3): p. 351-8.
403. Bacolla, A. and R.D. Wells, *Non-B DNA conformations as determinants of mutagenesis and human disease.* Mol Carcinog, 2009. **48**(4): p. 273-85.
404. Wang, G. and K.M. Vasquez, *Non-B DNA structure-induced genetic instability.* Mutat Res, 2006. **598**(1-2): p. 103-19.
405. Belotserkovskii, B.P., et al., *Mechanisms and implications of transcription blockage by guanine-rich DNA sequences.* Proc Natl Acad Sci U S A, 2010. **107**(29): p. 12816-21.
406. Chan, F.L. and L.H. Wong, *Transcription in the maintenance of centromere chromatin identity.* Nucleic Acids Res, 2012. **40**(22): p. 11178-88.
407. Chang, E.Y., et al., *RECQ-like helicases Sgs1 and BLM regulate R-loop-associated genome instability.* J Cell Biol, 2017. **216**(12): p. 3991-4005.
408. van Wietmarschen, N., et al., *BLM helicase suppresses recombination at G-quadruplex motifs in transcribed genes.* Nat Commun, 2018. **9**(1): p. 271.
409. Frizzell, A., et al., *RTEL1 inhibits trinucleotide repeat expansions and fragility.* Cell Rep, 2014. **6**(5): p. 827-35.
410. Zimmer, J., et al., *Targeting BRCA1 and BRCA2 Deficiencies with G-Quadruplex-Interacting Compounds.* Mol Cell, 2016. **61**(3): p. 449-460.
411. Yang, X., et al., *A public genome-scale lentiviral expression library of human ORFs.* Nat Methods, 2011. **8**(8): p. 659-61.
412. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-2308.
413. Stewart, S.A., et al., *Lentivirus-delivered stable gene silencing by RNAi in primary cells.* RNA, 2003. **9**(4): p. 493-501.