“MOLECULAR CHARACTERISATION OF HELQ HELICASE’S ROLE IN DNA REPAIR AND GENOME STABILITY”

Rafal Lolo

University College London

and

Cancer Research UK London Research Institute

PhD Supervisor: Simon Boulton

A thesis submitted for the degree of

Doctor of Philosophy

University College London

September 2018

Declaration

I Rafal Lolo confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Maintenance of genome stability is a critical condition that ensures that daughter cells acquire an accurate copy of the genetic information from the parental cell. DNA replication stress that arises from blocked replication forks, can be a major challenge to genome integrity. Cells have therefore developed complex mechanisms to detect and deal with the replication-associated DNA damage. Intra-S-phase ATR checkpoint, FA pathway and RAD51 paralog BCDX2 complex together constitute key components of the replication stress response system that is essential to sense, repair and restart damaged forks.

Previous studies in *D. melanogaster* and *C. elegans* have positioned HELQ as an important factor in DNA damage repair and maintaining genome stability. In this work I develop and combine biochemical assays, proteomic studies, mouse model and molecular biology tools to further characterise HELQ function in DNA repair and genome stability. I establish that HELQ plays a pivotal role in the replication stress response in mammalian cells. By developing a system in which I was able to pull down tagged HELQ and subject it to Mass Spectrometry analysis I identified its molecular partners and showed that HELQ interacts with, and interfaces between, the central FANCD2/FANCI heterodimer and the downstream RAD51 paralog BCDX2 complex. From mechanistical point of view, interaction with BCDX2 complex was the most interesting one and I took comprehensive experimental approach to identify the nature of HELQ-BCDX2 binding. I was able to show that this interaction does not need any mediating factors and is most likely a direct one. Despite many attempts and trying various tools I was not able to identify interacting motif neither in HELQ nor RAD51 paralogs, I identified however a candidate amino acid stretch within RAD51B as potential candidate. This works requires further validation.
Additionally, I show that HELQ is actively recruited to chromatin upon exposure to DNA damaging agents and this chromatin enrichment is ATR-dependent. Once present on chromatin HELQ performs its function to promote Homologous Recombination-depended repair as shown by persistence of chromatin RAD51 and RPA32 protein in HELQ deficient cells as well as decreased HR-dependent repair measured by DR-GFP assay. To better understand how HELQ promotes HR-

dependent ICL repair I decided to set up immune-depletion based ICL-repair system in *Xenopus laevis* egg extract (developed by Johannes Walter group). Although important progress has been made and I produced several valuable tools I did not manage to validate the system due to time limitations.

Lastly, our data positions HELQ as an ovarian cancer susceptibility gene with low penetrance effect. This makes HELQ and interesting object of familial cancer screens and potentially opens an exciting therapeutic approach based on PARP inhibition.

Impact statement

Genetic material of all cells is constantly subjected to physical and chemical damaging factors. If unrepaired or repaired illegitimately damaged DNA causes genomic instability and accumulation of mutations which constitute the hallmarks of cancer. Organisms from archaea to vertebrates have evolved elaborated mechanism allowing to detect and repair damaged DNA in efficient and timely manner. Research performed in the field of DNA damage response have provided fundamental insights into both processes driving cancer formation, and ways that cancer biology can be exploited for therapeutic purposes. One of the best examples of such approach is Fanconi Anemia, an extremely rare genetic disorder characterised by congenital defects, bone marrow failure and predisposition to cancer. Research done in the field of Fanconi Anemia contributed immeasurably to better understanding of how DNA damage is formed and repaired, in particular to our better understanding of most widely used chemotherapeutics - regimens based on ICL inducing agents.

In the presented study we show that HELQ is an important DNA damage repair factor acting in response to replication fork stalling agents such as ICL-inducing drugs and replication inhibitors. Our work led to detailed identification of interacting network of proteins that HELQ work in response to DNA damage. Interestingly almost all of HELQ identified interacting partners, if mutated lead to increased risk of cancer development placing the protein in the cancer suppressor network. Indeed, in case of HelQ mutated mice we also observed increased incidence of cancer, predominantly ovarian carcinoma. This phenotype is in line with our observation that the role of HELQ in maintaining genome stability can be contributed to promotion of efficient Homologous Recombination (HR) repair of stalled replication forks. The link between germline and somatic mutations in genes responsible for HR and increased risk of ovarian and breast cancers is well established and our data indicate that HELQ is a good candidate gene for screening in ovarian cancer.

Importantly, HELQ implication in ovarian cancer can potentially open targeted treatment options for patients caring pathogenic mutations in this gene. DNA

damage response inhibition is a novel anti-cancer approach which could revolutionise the treatment of cancer. It works by exploiting DNA damage response defects which are specific to cancer cells and kills them whilst sparing normal cells. Cancer cells with a DNA damage response deficiency are heavily reliant on remaining ‘back up’ DNA repaired systems which can be efficiently blocked by small molecule inhibitors causing synthetic lethality specifically in cancer cells. One of first and best studied example of such system is olaparib - PARP-1 inhibitor. In our work we established that HELQ deficient cells are hypersensitive to olaparib which potentially opens treatment option for patients with HELQ-mutated cancers.

To summarize, we established that HELQ plays a pivotal role in the replication stress response, interfacing between the central FANCD2/FANCI heterodimer and the downstream RAD51 paralog BCDX2 complex, which is critical for HR repair. Additionally, we showed that HELQ is a tumor suppressor and a good candidate for screening in ovarian cancer patients for clinical trial of PARP-1 inhibitors.

Acknowledgement

First, I would like to thank my supervisor Simon Boulton for giving me the opportunity to work in his lab and for guiding me throughout my PhD. I am particularly grateful for the patience and openness Simon has given me in times when I needed it most.

I would like to thank all members of the Boulton's Lab for their friendship, guidance and support throughout the years we spent together. My special thanks go to Carrie who I had the great pleasure to share the HelQ story with. Carrie's patience, experience and advice were invaluable and catalysed my own growth and development. Thank you, Carrie!

I would also like to express my gratitude to Costanzo's Group for their help and guidance in my work with *Xenopus* model.

Many people have made Clare Hall a very special place and the time spent there an important life lesson. Thank you all.

Last but not least, I would like to thank my Mother for the endless reserves of faith and confidence in me, in particular in those moments when I deserved it least.

Table of Contents

Abstract **3**
Impact statement **5**
Acknowledgement **7**
Table of Contents **8**
Table of figures **11**
List of tables **15**
Abbreviations **16**
Chapter 1.Introduction **19**
1.1 Overview of DNA Damage. **19**
1.1.1 Cellular consequences of DNA damage 19
1.1.2 DNA Repair Mechanisms 21
1.2 Formation and repair of DNA Interstrand Cross-Links (ICLs). **30**
1.2.1 ICL-inducing agents. 31
1.2.2 ICL Repair mechanisms 37
1.2.3 Recognition of ICL 41
1.3 The role of HELQ helicase in ICL repair **67**
Chapter 2.Materials & Methods **70**
2.1 Enzymes and reagents **70**
2.1.1 Enzymes 70
2.1.2 General reagents 70
2.1.3 Antibodies 71
2.2 Buffers and solutions **72**
2.2.1 Media and protein buffers 72
2.2.2 DNA buffers 72
2.3 Bacterial strains **73**
2.3.1 Bacterial strains 73
2.3.2 Transformation of chemically competent cells 73
2.4 Cell lines and cell culture methods **73**
2.4.1 siRNA transfection 74
2.4.2 Cell line derivation (work done with/ by C. Adelman) 78
2.4.3 Clonogenic survival assay 79
2.5 Gel electrophoresis **80**
2.5.1 SDS-polyacrylamide gel electrophoresis (PAGE) 80
2.5.2 Agarose gel electrophoresis 81
2.5.3 InstantBlue staining 81
2.5.4 SYPRO® Ruby staining 81
2.6 General methods of DNA and protein manipulation **82**
2.6.1 DNA concentration determination 82
2.6.2 PCR for cloning 82
2.6.3 Gateway recombinational cloning 82
2.6.4 Ethanol precipitation 82
2.6.5 Protein concentration determination 83
2.6.6 Western blotting 83
2.7 Baculovirus and insect cells **83**
2.7.1 Production of the bacmid DNA 83

2.8 Peptide arrays – BCDX2-HELQ interaction site mapping 84
2.8.1 Peptide arrays-based mapping of HELQ-FLAG binding site within BCDX2 complex 84
2.8.2 BCDX2 complex binding site within HELQ studies by peptide arrays 86
Chapter 3.HELQ is a DNA response protein that promotes RAD51 paralogue-dependent repair of DSBs 87
3.1 HELQ interacts with known DNA repair factors..... 87
3.1.1 HELQ expression systems 88
3.1.2 HELQ expression validation 90
3.1.3 HELQ purification procedure – small scale 90
3.1.4 HELQ purification procedure – large scale. 93
3.1.5 NIH-3T3 HelQ-GFP Mass Spectrometry analysis 94
3.1.6 Human 293 HELQ-FLAG Mass Spectrometry analysis. 96
3.1.7 Validation of the interactions identified by Mass Spectrometry..... 98
3.2 HELQ specifically interacts with BCDX2 but not CX3 of the RAD51 paralogs complexes. 103
3.2.1 Endogenous CX3 complex IP 103
3.2.2 HELQ – BCDX2 *in vitro* binding assay 105
3.3 Discussion 110
Chapter 4.Examining HELQ role in replication stress. 112
4.1.1 HELQ subcellular localization studies (Immunofluorescence)..... 112
4.1.2 HELQ subcellular localization studies (Fractionation) 115
4.1.3 iPOND-based studies of HELQ dynamics at replication forks. 120
4.2 HELQ functions in the DNA Interstrand Cross Links repair in parrarel to the FANCD2-FANCI. 125
4.2.1 HelQ mouse model generation (Carrie Adelman)..... 125
4.2.2 HelQ confers crosslinking agents resistance. 127
4.2.3 HelQ is dispensable for Intra-S-phase checkpoint activation. 128
4.2.4 HelQ is dispensable for FANCD2-FANCI ubiquitination. 130
4.3 HELQ is involved in the Homologous Recombination. 131
4.4 Discussion 148
Chapter 5.HELQ-BCDX2 interaction mapping. 150
5.1 HELQ and BCDX2 stability studies. 150
5.2 HELQ and BCDX2 interaction site analysis 151
5.2.1 Interaction mapping - peptide arrays. 151
5.3 Discussion 167
Chapter 6.HELQ functional studies using *Xenopus laevis* egg extract immunodepletion system. 171
6.1 *X. laevis* egg extract immunodepletion experimental system to study the role of HELQ in DNA interstrand cross links. 171
6.2 Generation of *X. laevis* HelQ expression constructs. 172
6.2.1 Determining the *X.laevis* HelQ cDNA sequence 173
6.2.2 *De novo* synthesis of codon optimized *X.laevis* HelQ cDNA and expression of the full-length protein in insect High5 cell line. 174
6.2.3 Expression of the fragments of *X. laevis* HelQ in *E.coli*..... 177
6.2.4 *X.laevis* HelQ tertiary structure prediction 177
6.2.5 Cloning, expression and purification of *X. laevis* HelQ immunogenic fragments. 180

6.2.6 Antibodies directed against *X.laevis HelQ*- validation 183
6.2.7 Generation of *X. laevis* HelQ antibodies with antigenic peptides. ... 187
6.2.8 Anti-*X.laevis* HelQ peptide antibodies testing. 190
6.3 Discussion 192
Chapter 7.Discussion 194
Reference List 204

Table of figures

Figure 1-1 Outcomes of accumulated DNA damage. 20
Figure 1-2 Common types of damage and corresponding repair pathways. 21
Figure 1-3 Schematic representation of MMR. 23
Figure 1-4 Schematic representation of BER. 25
Figure 1-5 Schematic representation of GG-NER and TC-NER. 29
Figure 1-6 Schematic representation of Intrastrand Crosslink (left) and Interstrand Crosslink of DNA (right) 30
Figure 1-7 Schematic representation of interstrand crosslink (ICL). 31
Figure 1-8 Cisplatin-induced ICL. 34
Figure 1-9 Mitomycin C induced ICL. 35
Figure 1-10 Schematic representation of differences in ICL repair during and outside of S phase 40
Figure 1-11 Schematic representation of FA role in ICL repair. 49
Figure 1-12 Architecture of the FA complexes including FAAPs (Fanconi anemia-associated proteins). 54
Figure 3-1 HELQ Purification Scheme. 87
Figure 3-2 Schematic representation of HELQ expression constructs. 89
Figure 3-3 HELQ expression validation. 91
Figure 3-4 Exemplary HELQ IP optimization - small scale. 92
Figure 3-5 SYPRO Ruby-stained SDS-PAGE gels with large scale HELQ immunoprecipitation and copurified proteins. 94
Figure 3-6 Mass Spectrometry results for HELQ-FLAG IP from human 293 FlpIn cells. 97
Figure 3-7 RAD51B western blot signal improvement over time. 99
Figure 3-8 Western blot validation of HELQ interactions identified by Mass Spectrometry. 100
Figure 3-9 Western blot analysis of RAD51C IP 101
Figure 3-10 IP of endogenous HELQ and RAD51C from 293 cells. 103
Figure 3-11 Endogenous XRCC3 IP. 104
Figure 3-12 Schematic representation of HELQ-FLAG and BCDX2 complex in vitro binding assay. 106

Figure 3-13 Recombinant BCDX2 complexes. 107
Figure 3-14 HELQ-FLAG - BCDX2 binding in vitro assay. 109
Figure 4-1 Endogenous HELQ IF (C. Adelman) 113
Figure 4-2 IF of HeLa cells transiently expressing HELQ-GFP (C. Adelman) 114
Figure 4-3 U2OS and NIH-3T3 HelQ-GFP cell lines fractionation results. 116
Figure 4-4 Western blot analysis of NIH-3T3 HelQ-GFP fractionation. 118
Figure 4-5 Western blot and densitometry analysis of DNA damage induced HelQ-GFP chromatin recruitment. 119
Figure 4-6 Identification of Proteins On Nascent DNA (iPOND) Methodology 122
Figure 4-7 HELQ is recruited to stalled replication forks as shown by iPOND 123
Figure 4-8 Schematic representation of HelQ$^{\Delta C/\Delta C}$ deletion introduced in the mouse model. 126
Figure 4-9 HelQ$^{\Delta C/\Delta C}$ cells sensitivity towards DNA damage. (C. Adelman) 128
Figure 4-10 Checkpoint activation status in mouse HelQ$^{\Delta C/\Delta C}$ cells 129
Figure 4-11 Checkpoint activation status in USOS cells depleted of HELQ. 130
Figure 4-12 Fanconi anemia pathway activation in the absence of HelQ. 131
Figure 4-13 MMC induced RAD51 chromatin recruitment in HELQ depleted cells. 132
Figure 4-14 MMC induced RAD51 foci formation in HELQ depleted U2OS cells. 133
Figure 4-15 MMC induced RPA32 foci formation in HELQ depleted U2OS cells. 134
Figure 4-16 HELQ deficiency affects MMC induced RAD51 chromatin dynamics. 136
Figure 4-17 MMC induced RAD51 nuclear foci dynamics in mouse HelQ$^{\Delta C/\Delta C}$ cells 138
Figure 4-18 MMC induced γH2AX nuclear foci dynamics in mouse HelQ$^{\Delta C/\Delta C}$ cells 139
Figure 4-19 MMC induced RAD51 nuclear foci dynamics in HELQ depleted human U2OS cells. 140
Figure 4-20 MMC induced γH2AX nuclear foci dynamics in HELQ depleted human U2OS cells. 141
Figure 4-21 Homologous Recombination frequencies in DR-GFP reporter cells depleted for HELQ. 143
Figure 4-22 Homologous Recombination frequencies in DR-GFP reporter cells depleted for HELQ – transfection efficiency normalization. 144

Figure 4-23 Homologous Recombination frequencies in DR-GFP reporter cells depleted for HELQ (Alessandro Sartori).....146
Figure 4-24 HELQ depleted U2OS cells are sensitive to PARP inhibitor, Olaparib. 147
Figure 5-1 HELQ and BCDX2 stability in U2OS cells.....150
Figure 5-2 HELQ-BCDX2 interaction site mapping by peptide array - schematic.153
Figure 5-3 Peptide array-based mapping of BCDX2-HIS binding site within HELQ protein – detection with anti RAD51C antibodies. 154
Figure 5-4 Peptide array - based mapping of BCDX2 complex binding site within HELQ. 156
Figure 5-5 Peptide array-based mapping of BCDX2-HIS binding site within HELQ protein – detection with anti HIS antibodies. 157
Figure 5-6 Control HLEQ membrane. 158
Figure 5-7 Peptide array-based mapping of HELQ-FLAG binding site within BCDX2 complex (RAD51B, RAD51C). 159
Figure 5-8 Peptide array-based mapping of HELQ-FLAG binding site within BCDX2 complex (RAD51D, XRCC2).....160
Figure 5-9 Negative control - XRCC3 peptide array. 161
Figure 5-10 Peptide array-based mapping of HELQ-FLAG binding site within BCDX2 complex using anti HELQ antibodies for detection. 162
Figure 5-11 RAD51D peptide pull down from HEK293 HELQ-FLAG 164
Figure 5-12 Substitution analysis of HELQ-FLAG binding to RAD51D.....167
Figure 5-13 Substitution analysis of HELQ-FLAG binding to RAD51D Ponceau staining. 167
Figure 6-1 Schematic representation of *Xenopus* egg extract preparation and replication dependent ICL repair plasmid model. 172
Figure 6-2 Obtaining *Xenopus laevis* HelQ cDNA sequence. 174
Figure 6-3 *X. laevis* HelQ expression construct - schematic. 175
Figure 6-4 Analysis of *X. laevis* expression in the High5 cell line by SDS-Page (A) and Western blot (B).....177
Figure 6-5 *X.laevis* HelQ sequence alignment with known proteins from PDB database: 5 hits with highest scores. 178
Figure 6-6 *X.laevis HelQ* three-dimensional structure prediction.....179
Figure 6-7 *X. laevis* HelQ Fragment 3-HIS expression in *E.coli* 181

Figure 6-8 *X.laevis* HelQ fragment 3 purification.....182
Figure 6-9 *X. laevis* HelQ antibodies testing. 185
Figure 6-10 *X. laevis* HelQ immunodepletion testing. 186
Figure 6-11 *X.laevis* HelQ structure prediction with two antigenic peptides mapped. 188
Figure 6-12 Response of different bleeds to the antigenic peptides as measured by ELISA 190
Figure 6-13 *X.laevis* HelQ IP from egg extract with antibodies raised with immunogenic peptides. 191

List of tables

Table 1-1 Summary of FA and FA-like genes. 66
Table 2-1 List of cell lines used in this study.....73
Table 3-1 Optimized HELQ IP conditions - small scale. 92
Table 3-2 List of most abundant proteins copurified with HelQ-GFP (mouse NIH - 3T3 cell line). The list lacks respective numbers of unique peptides. 95
Table 3-3 List of potential binding partners of mouse HelQ-GFP. The list lacks respective numbers of unique peptides.....96
Table 4-1 Comparison of HELQ and FANCD2 model mouse phenotypes. 127
Table 6-1 Characterization of the *X.laevis* HelQ fragments.....180
Table 6-2 *X.laevis* immunogenic peptides used for raising antibodies 188

Abbreviations

ALT	alternative lengthening of telomeres
APBs	ALT-associated PML bodies
ATP	adenosine triphosphate
BAC	bacterial artificial chromosome
BER	base excision repair
bp	base pair
BrdU	bromodeoxyuridine
BS	Bloom's syndrome
CFS	common fragile site
CO-FISH	chromosome orientation-FISH
CPT	camptothecin
CS	Cockayne syndrome
dGTP	deoxyguanosine triphosphate
dHJ	double HJ
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphate
DSB	double-strand break
DSBR	double-strand break repair
dsDNA	double-stranded DNA
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
ES cells	embryonic stem cells
FA	Fanconi anaemia
FBS	fetal bovine serum
FISH	fluorescence *in situ* hybridisation
G1-phase	gap1 phase
G2-phase	gap2 phase
GFP	green fluorescent protein
HDX-MS	hydrogen-deuterium exchange mass spectrometry
HhH	helix-hairpin-helix

Hi5	High five
HJ	Holliday junction
HR	homologous recombination
hr	hour
HtH	helix-turn-helix
HU	hydroxyurea
ICL	interstrand cross-link
IF	immunofluorescence
IPOND	isolation of proteins at nascent DNA
IPTG	isopropyl-β-D-thiogalactoside
IR	ionising radiation
kDa	kilo Daltons
KIN	karyomegalic interstitial nephritis
LOH	loss of heterozygosity
M phase	mitotic phase
MEF	mouse embryonic fibroblast
MMC	mitomycin C
MMS	methyl methanesulfonate
MS	Mass Spectrometry
NCBI	National Centre for Biotechnology Information
NER	nucleotide excision repair
NHEJ	non-homologous end joining
NP-40	nonidet P40
nt	nucleotide
OB	oligonucleotide/oligosaccharide binding
OD	optical density
P. furiosus	*Pyrococcus furiosus*
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffer saline
PCR	polymerase chain reaction
PI	propidium iodide
PML	promyelocytic leukaemia

PSI-BLAST	Position Specific Iterated–Basic Local Alignment Search Tool
Q-FISH	quantitative-FISH
RF	replication fork
ROS	reactive oxygen species
RT	room temperature
RT-PCR	reverse transcriptase-PCR
s	second
S-phase	synthesis phase
S. cerevisiae	*Saccharomyces cerevisiae*
S. pombe	*Schizosaccharomyces pombe*
SCE	sister chromatid exchange
SDS	sodium dodecyl sulfate
SDSA	synthesis-dependent strand annealing
Sf9	*Spodoptera frugiperda 9*
siRNA	small interfering RNA
ssDNA	single-stranded DNA
TEMED	N,N,N',N'-(tetramethylethylenediamine)
Tris	Tris (hydroxymethyl) aminomethane
U	unit
UV	ultra violet
V	Volts
W	Watts
WCE	whole cell extract
WRN	Werner helicase
WS	Werner syndrome
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
XL-MS	protein cross-linking coupled with mass spectrometry
XP	xeroderma pigmentosum

Chapter 1. Introduction

1.1 Overview of DNA Damage.

DNA is the exclusive repository of genetic information thus preserving its stability and integrity is of much greater importance than that of RNA, proteins or other cellular components. Maintaining genome stability and introduction of genetic diversity to drive evolution are two contradictive processes that require an extremely complex regulation. Organisms from bacteria to mammals have developed a range of mechanisms that govern the molecular response to lesions that occur within the DNA. There are sensors that detect damaged or mismatched DNA, systems that signal the presence of the lesions and finally there are elaborated repair mechanisms that remove the error in a tightly regulated fashion.

DNA is constantly exposed to various chemical and physical conditions that damage the DNA double helix. Although it was initially believed that exogenous sources constitute the primary cause of mutagenic lesions, it has been proposed that endogenous factors are equally or more significant (Jackson and Loeb, 2001). By-products of cellular metabolism such as reactive oxygen species (ROS) (Pavlov et al., 1994) or naturally occurring deamination of DNA bases (Lindahl, 1993) are examples of intrinsic sources of damage. Additionally the nature of DNA polymerase enzymes can lead to the introduction of mis-incorporated nucleotides in each replication cycle (Echols and Goodman, 1991).

Exogenous mutagens can be divided into physical and chemical. Physical sources are primarily radiation such as ultraviolet (UV) that induces formation of covalent bonds between adjacent bases or ionizing radiation that triggers formation of ROS. Chemical mutagens react with DNA and proteins. One of the best-known chemical modifiers are alkylating agents that covalently attach methyl or ethyl moieties to DNA.

1.1.1 Cellular consequences of DNA damage

It is estimated that every day in each of the 10^{13} cells of the human body thousands to millions of DNA lesions occur. Fortunately, DNA repair mechanisms effectively remove the vast majority of damage. However, some of the lesions may stay

unrepaired for longer with ramifications for a range of cellular functions (Figure 1-1). For example, most of the lesion types can block transcription and replication – key processes involving DNA. If the amount of damage exceeds the repair capability of the cell, then unrepaired lesions can accumulate causing mutations. There are three main scenarios for a cell that lose the ability to effectively remove the damage (Figure 1-1):

- Cells undergo senescence; become permanently dormant and all division process are halted (Collado et al., 2005).
- Cells undergo programmed death (apoptosis). Accumulation of excessive DNA damage can trigger p53 dependent apoptosis.
- Cells undergo malignant transformation. Accumulation of oncogenic mutations can lead to development of cancer (Hanahan and Weinberg, 2011)

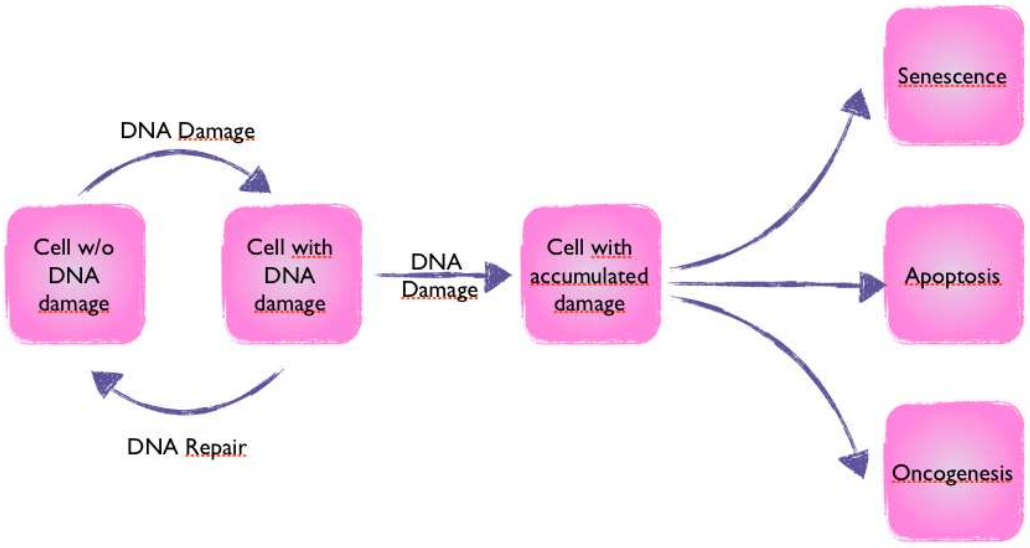


Figure 1-1 Outcomes of accumulated DNA damage.

Damaged DNA is sensed and repaired by DDR mechanisms. If the repair mechanisms are compromised or the load of the damage exceeds the repair capacity, then lesions accumulate, and mutations can arise. Depending on the extent of damage, cells may become senescent, trigger apoptosis or undergo oncogenic transformation.

1.1.2 DNA Repair Mechanisms

In order to preserve genomic integrity and compensate for different type of damage, cells have evolved a wide range of different DNA repair mechanisms. The way cells respond to DNA damage depends on many factors, the most important being the type of the lesion and cell cycle stage. Interestingly, although there are some significant differences in the repair mechanisms between organisms, the principles remain the same.

In mammalian cells there are at least five distinguishable DNA repair mechanisms that are responsible for sensing and removing damage (Figure 1-2). Their activation depends on the type of damage and cell cycle stage. In the past decade, it has become evident that there is considerable cross-talk between the repair mechanisms and together they work as a DNA repair network rather than a number of linear, independent pathways

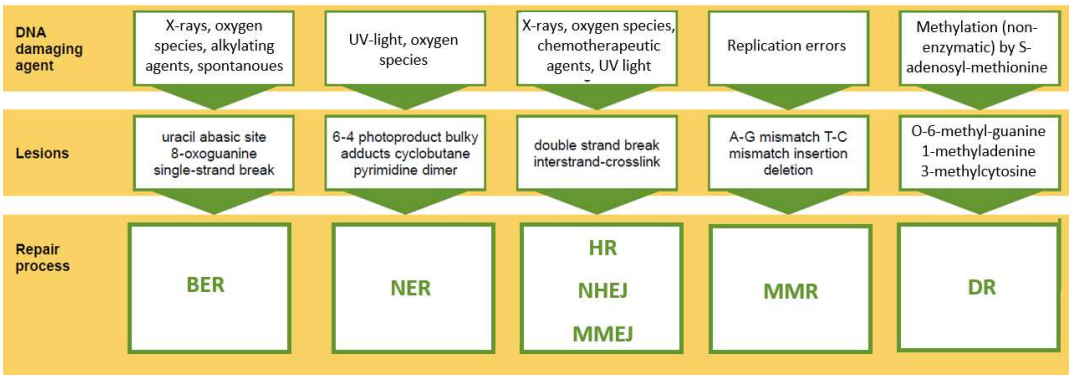


Figure 1-2 Common types of damage and corresponding repair pathways.

The table shows common damaging agents, DNA lesion types caused by the damaging agents and corresponding detection and repair mechanisms. Although it is possible to dissect separate repair pathways, a cross-talk between different branches is possible (i.e. NER can be involved in ICLs repair). BER – Base Excision Repair, NER – Nucleotide Excision Repair, HR – Homologous Recombination, NHEJ – Non-Homologous-End Joining, MMEJ – Microhomology-Mediated-End-Joining, MMR – Mismatch repair, Direct Reversal.

1.1.2.1 Mismatch repair (MMR)

Although replicative DNA polymerases are characterized by high fidelity with only one nucleotide in every 10^7 events mis-incorporated into DNA, the size of the human

genome ($3x10^9$ bp) is large enough to produce hundred to thousands of mismatches in each replication cycle (McCulloch and Kunkel, 2008) (Echols and Goodman, 1991). The main role of MMR is to sense and remove mis-incorporated nucleotides that escaped the proofreading activity of DNA polymerases, which increases the overall fidelity of the replication process. Additionally, components of the pathway help to remove insertions and deletions created by DNA polymerase slippage when replicating repetitive sequences.

The pivotal role MMR plays in maintaining genome stability is underlined by the fact that cells with compromised MMR pathway exhibit a strong mutator phenotype characterized by significantly elevated mutations rates and microsatellite instability (Li, 2008). Germline mutations in MMR genes cause Lynch syndrome – a condition with significant cancer predisposition (Hewish et al., 2010; Rodriguez-Bigas et al., 1997). From a mechanistic point of view, the pathway consists of three major steps: (i) detection where mismatches are recognized, excision of the strand containing the mis-incorporated nucleotides and a synthesis step where specialised DNA polymerases fill the incised gap (Fukui, 2010) . Each of the steps is carried out by a separate set of proteins.

In brief, the MutSα complex (consisting of MSH2-MSH6) is responsible for recognising mismatches and small indels whereas MutSβ (MSH2-MSH3) preferentially binds to larger indels (Jiricny, 2006). Following the MutS-DNA lesion binding, another major complex is recruited to the lesion. In mammalian cells the most significant MutL homolog complex is MutLα consisting of MLH1 and PMS2. To a lesser extend two other homologs have been found to play a role in MMR - MutLβ, PMS1 and MutLγ. The main role of the MutL complex is to couple the recognition step with the downstream incision event (Kunkel & Erie, 2015). First MutLα via its PCNA/replication factor C (RFC) mediated exonuclease activity generates a single stranded nick that is used as a starting point for the main exonuclease. In humans, degradation of the lesion-containing strand is performed by the exonuclease 1 enzyme (Exo1) (Galio et al., 1999; Kadyrov et al., 2006; Tran et al., 2004). In the final step, the single stranded gap is filled in by δ polymerase and

the remaining nick is sealed thanks to DNA ligase I activity (Constantin et al., 2005; Longley et al., 1997; Zhang et al., 2005).

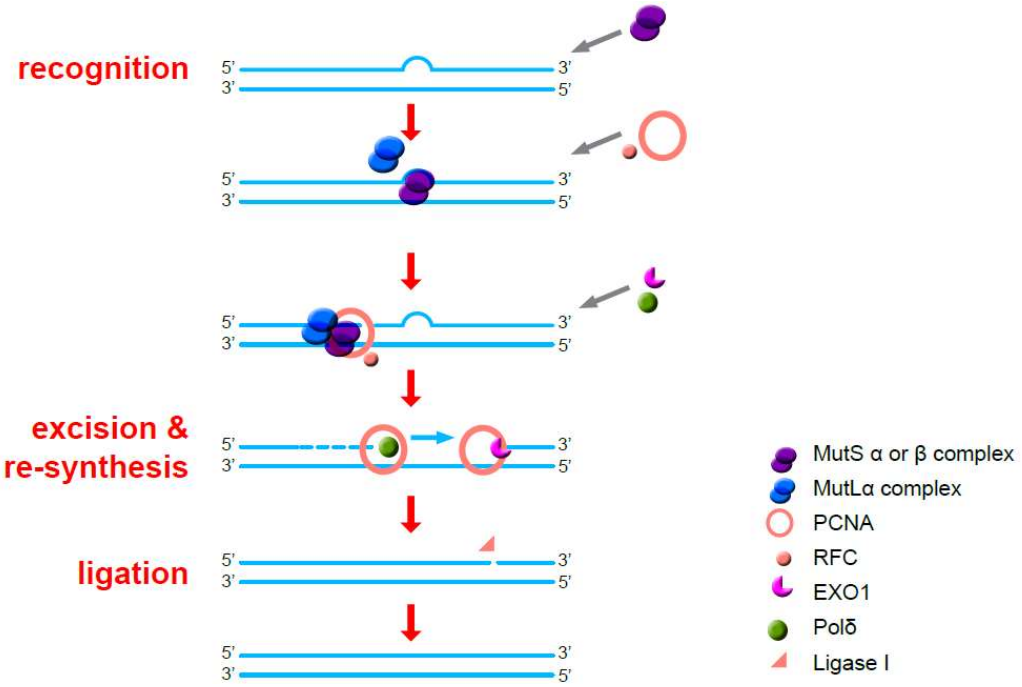


Figure 1-3 Schematic representation of MMR. Schematic representation of MMR. Mismatched base pair is recognized by MutSα or MutSβ. MutLα (PMS2 and MLH1 heterodimer) is recruited to the damaged DNA and forms a complex with MutS. PCNA stimulates endonuclease activity of the PMS2 component of MutL α, that subsequently creates a nick at 5′ to the mismatch. EXO1 nuclease is recruited to extend the nick in 5' to 3' direction. Pol δ stimulated by PCNA fills the gap and Ligase I seals the remaining nick.

1.1.2.2 Base Excision Repair (BER)

BER is a predominant mechanism responsible for repair of DNA lesions that affect individual bases but do not significantly change the local structure of the DNA double helix. Similar to MMR, BER consists of lesion detection, excision and synthesis events but unlike with MMR, there are several entry points depending on the type of damage that has been sensed. Another difference from MMR is that DNA glycosylases that detect the lesion also have enzymatic activity that allows excision of the affected bases. To ensure high specificity, this dual activity comes at a price. DNA Glycosylases are only able to detect a single type of lesion, so cells need to

express a number of different glycosylases in order to recognize and remove a spectrum of different lesions (Dizdaroglu, Coskun, & Jaruga, 2017).

Upon recognition, DNA glycosylase cleaves the glycosidic bond between the base and deoxyribose leading to formation of an apurinic/apyrimidinic site (AP site). As a matter of fact, this BER intermediate is also a very abundant type of damage resulting from spontaneous DNA deamination (Wallace, 2014). In the next step, APE1 endonuclease recognizes the AP site and cleaves 5' to the lesion generating a standard 3'OH substrate for the DNA re-synthesis event (Mol, Izumi, Mitra, & Tainer, 2000). Interestingly, some of the DNA glycosylases can cleave the AP sites in a mechanism known as β-elimination to produce 5' phosphate ends and 3'-phospho-α, β-unsaturated aldehyde. Successive β and δ elimination catalysed by NEIL1 and NEIL2 enzymes turns the 3'-phospho-α, β-unsaturated aldehyde into a 3' phosphate. In order to complete repair of the lesion, both phosphate ends need to be converted into standard 3'OH and 5' phosphates that are substrate for DNA polymerase and ligase. This step can be carried out by a number of enzymes depending on the nature of the cleavage that led to formation of the AP site.

Depending on whether one or many nucleotides are incorporated the final synthesis and ligation steps are divided into two sub-pathways: short-patch and long-patch BER respectively (Fortini and Dogliotti, 2007). Short-patch BER is by far the most frequently occurring mechanism and involves single-nucleotide gap filling performed by DNA Polymerase β and ligation with either DNA ligase I or a XRCC1-ligase III complex. Long-patch BER is initiated by lesions refractory to the lyse activity of Polymerase β and requires involvement of replication factors. In short, PCNA together with Polymerase β, δ or ε elongate the 3'OH terminus displacing the 5'lesion. The displacement flap sequence is then removed by FEN1 endonuclease followed by ligase-I dependent sealing of the downstream-pushed nick. (Sattler et al., 2003).

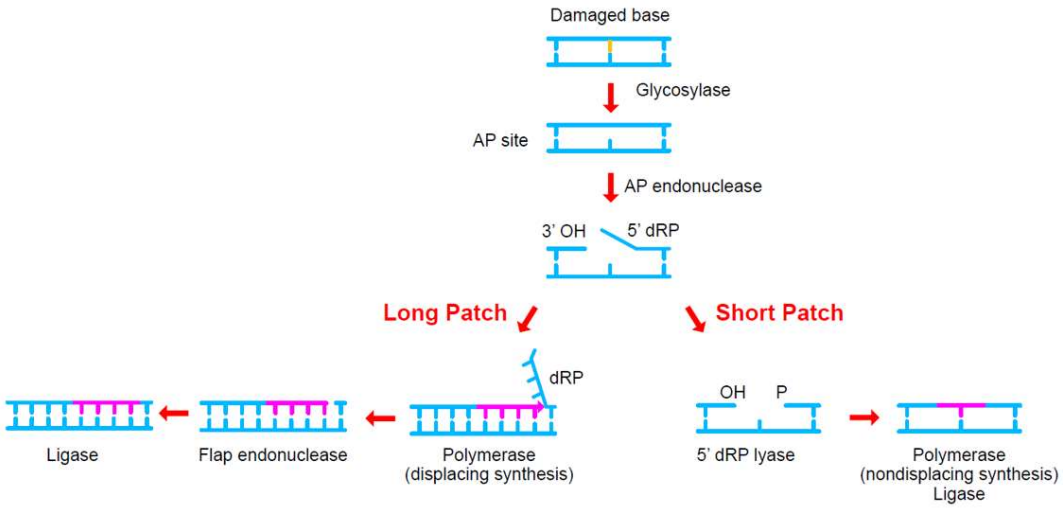


Figure 1-4 Schematic representation of BER.

Damaged base is recognised and cleaved off by damage-specific glycosylase leading to formation of apurinic/apyrimidinic site (AP site). AP-site is then digested 5' of damage by AP endonuclease generating 3' hydroxyl (3' OH) adjacent to a 5' deoxyribosephosphate (dRP). Short-patch BER involves formation of standard phosphate at 5' end and single-nucleotide gap filling performed by DNA Polymerase β followed by ligation. Lesions refractory to the activity of Polymerase β undergo long-patch BER in which 3' end is elongated by PCNA and polymerases which in turn displaces 5' lesion. The flap sequence is then digested by flap FEN1 endonuclease and remaining nick is ligated completing repair cycle.

1.1.2.3 Nucleotide Excision Repair (NER)

Although mechanistically reminiscent of BER, Nucleotide Excision Repair is a much more complex process involving over 30 different proteins. It is also a more diverse mechanism than BER targeted at repair of a wide variety of lesions locally distorting the double helix. The biological significance of NER is highlighted by the fact that defects in the pathway lead to several hereditary disorders such as *Xeroderma pigmentosum* or Cockayne syndrome. Both diseases are characterized by oversensitivity to UV radiation but important differences in symptoms spectrum are also present: Cockayne syndrome patients suffer from neurodegeneration whereas individuals with *Xeroderma pigmentosum* have elevated incidence of cancer . (Cleaver, 2005; Cleaver, Lam, & Revet, 2009)

The main type of damage targeted by NER are UV-induced pyrimidine dimers and DNA intrastrand crosslinks. NER is initiated by a recognition event followed by

localized melting of the region adjacent to the lesion, excision of a single stranded fragment and then ligation. Two mechanistically similar sub-pathways can be distinguished – global genome NER (GG-NER) and transcription-coupled NER (TC-NER). GG-NER is responsible for repair processes across the entire genome whereas TC-NER is activated in regions being transcribed. Both processes are related and differ mostly by the recognition step. For GG-NER the pivotal role in sensing the lesion is played by the XPC/RAD23B/CETN2 protein complex (Sugasawa et al., 1998). Crystallographic studies of yeast XPC orthologue, Rad4 bound to DNA containing a cyclobutane pyrimidine dimer (CPD) revealed the mechanism in which XPC recognizes the lesion. Upon recruitment XPC inserts its β-hairpin through the DNA duplex, resulting in the two damaged base pairs to be flipped out of the helix. Interestingly it is only the nucleotides of the undamaged strand that are recognized by Rad4, whereas the two complementary CPD-linked nucleotides become disordered. This data indicated that lesion recognition and binding of Rad4 leads to thermodynamic destabilization of the Watson–Crick bonds and as a result local helix bending that facilitates flipping-out of the nucleotides. RAD23B and CETN2 components play facilitating roles increasing affinity of XPC for the lesion. It has been shown that the more helix-distorting lesion, the higher affinity of XPC binding (Sugasawa, 2008; Sugasawa et al., 1998). Interestingly, UV-induced cyclobutane pyrimidine dimers (CPDs) that are important target lesions for NER only minorly distort the helix and XPC has little affinity towards them. This is why some additional proteins are required to facilitate NER repair of CPDs. UV-damaged DNA binding complex (UV-DDB) recognizes the CPDs and bends the DNA introducing a bigger distortion that is then bound by XPC with high affinity (Sugasawa, 2010).TC-NER is activated when RNA Polymerase II (RNAPII) stalls at damaged lesions (Fousteri and Mullenders, 2008). Upon activation Cockayne Syndrome Protein A (CSA) and B (CSB) are recruited and are believed to dislocate the RNAPII complex from DNA to facilitate downstream repair (Hanawalt and Spivak, 2008). Despite mechanistic differences in the recognition step between GG-NER and TC-NER, downstream repair events follow the same scenario. First, TFIIH – a multi-subunit transcription factor is recruited by XPF in GG-NER and CSA/CSB in TC-NER. Subsequently, XPB and XPD – TFIIH associated helicases unwind 25-35 nucleotides of DNA adjacent to the lesion. This facilitates XPA protein recruitment, which provides an additional damage verification mechanism (Marteijn et al., 2014). RPA

binding stabilizes single stranded DNA and the protein complex. The excision step is executed by two structure specific endonucleases: XPG and XPF/ERCC1 (Godon et al., 2012). The lesion containing DNA fragment of about 30 nucleotides is cleaved out and the gap is filled by Polymerase δ or ε using the undamaged strand as a template (Mocquet et al., 2008; Overmeer et al., 2011). The process is finished when DNA ligase seals the nicked DNA on the newly synthetized strand. It is worth mentioning that involvement of particular sets of proteins in the final steps of NER is dependent on the proliferative status of the cell and a number of different complexes might be recruited (Ogi et al., 2010).

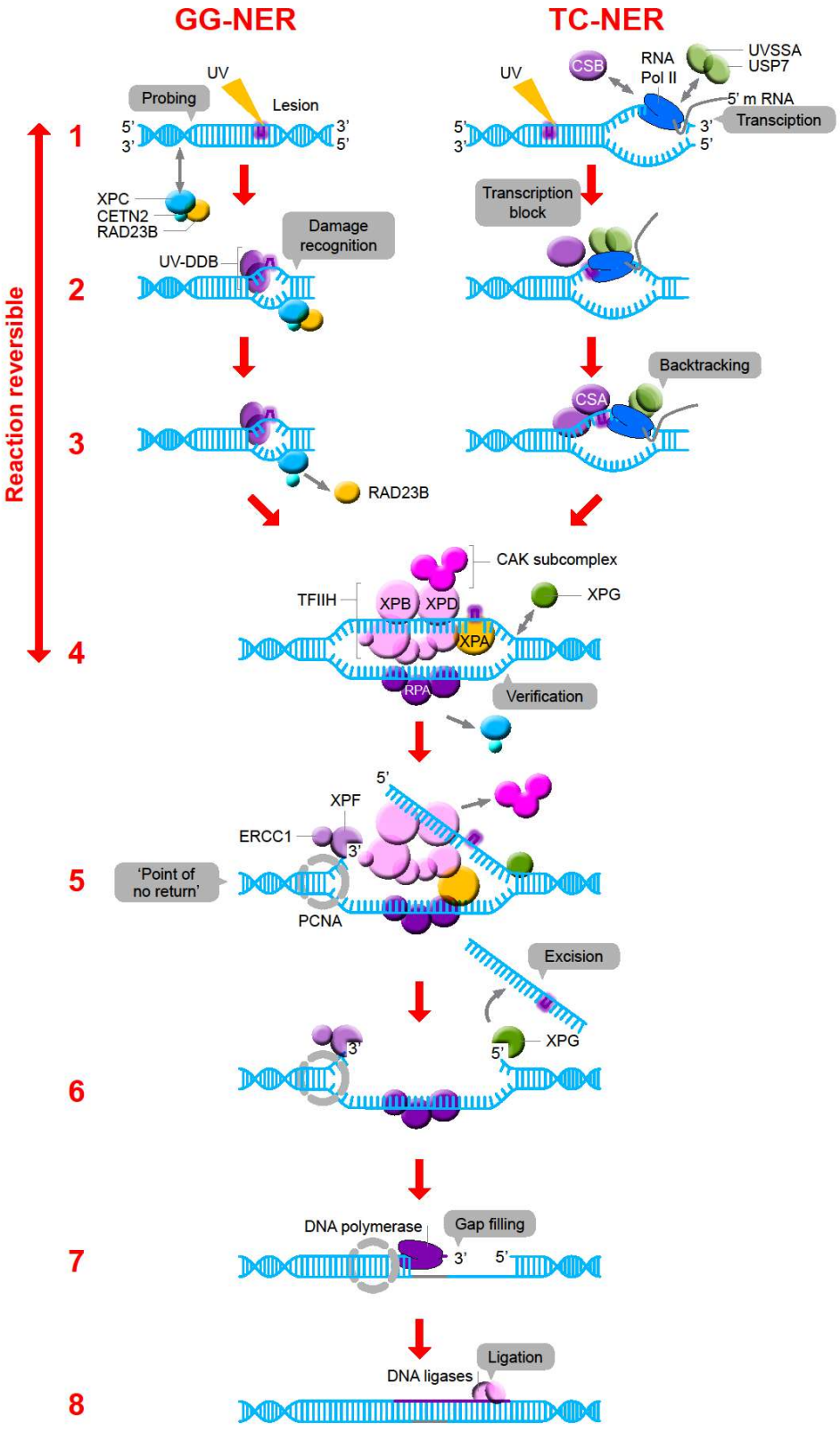

GG-NER
TC-NER
Reaction reversible
1
2
3
4
5
6
7
8
UV
Probing
Lesion
5'
3'
3'
5'
XPC
CETN2
RAD23B
Damage recognition
UV-DDB
RAD23B
UV
CSB
RNA Pol II
UVSSA
USP7
5' m RNA
Transciption
Transcription block
Backtracking
CSA
CAK subcomplex
TFIIH
XPB
XPD
XPG
XPA
RPA
Verification
XPF
ERCC1
'Point of no return'
PCNA
Excision
XPG
DNA polymerase
Gap filling
DNA ligases
Ligation

Figure 1-5 Schematic representation of GG-NER and TC-NER.

Left panel: GG-NER. **Step 1**: DNA is constantly scanned for helix-distortion by a complex formed by XPC, RAD23B and CETN2. **Step 2**: damage recognition by XPC-RAD23B-CETN2 is supported by UV–DDB complex. **Step 3**: Upon binding of the XPC complex to the damage, RAD23B dissociates from the complex. **Right panel:** TC-NER. **Step 1**: During unperturbed transcription SSA, USP7 and CSB proteins transiently interact with RNA Polymerase. Step 2: UV damage is sensed indirectly by RNA Polymerase II collision with the UV-induced lesion. Upon transcription machinery stalling affinity of CSB to polymerase increases and CSA—CSB complex is formed**. Step 3:** CSA-CSB complex formation leads to backtracking (reverse translocation) of polymerase which makes the damaged region more accessible for repair machinery. After damage recognition, further repair events are the same for both GG- and TC-NER. **Step 4:** TFIIH is recruited to the lesion and XPG nuclease binds to the pre-incision NER complex which leads to release of the CAK subcomplex (CDK-activating kinase) from the core TFIIH complex . TFIIH unwinds DNA around the damage which allows XPD to further verify presence of the lesion with the help of ATPase activity of XPB and altered-nucleotide binding activity of XPA. RPA proteins coat the undamaged strand providing stabilisation. **Step 5:** XPA recruits XPF-ERCC1 endonuclease that is directed to the damage-affected strand by RPA and performs incision 5' of the lesion. Following the incision event PCNA is loaded. Incision by XPF-ERCC1 marks step beyond which the reaction is irreversible ("point of no return"). **Step 6**: activated XPG cleaves the affected strand 3′ to the damage, which releases 22–30 nucleotide-long strand containing the lesion. **Step 7:** PCNA promotes gap filling synthesis by recruiting DNA Pol δ, DNA Pol κ or DNA Pol ε. **Step 8**: NER cycle is completed when DNA ligase 1 or DNA ligase 2 seal the remaining nick.

1.2 Formation and repair of DNA Interstrand Cross-Links (ICLs).

ICLs are lesions that covalently link nucleotides on both strands of the DNA double helix. They are introduced by chemical agents with two active groups that attack bases on the complementary strands. ICLs efficiently prevent strand separation thus inhibiting the key functions of DNA such as transcription and replication. The ability to affect both strands of the double helix translates into profound toxicity of ICL-inducing agents and makes them efficient anti-cancer therapeutics (Lawley and Phillips, 1996).

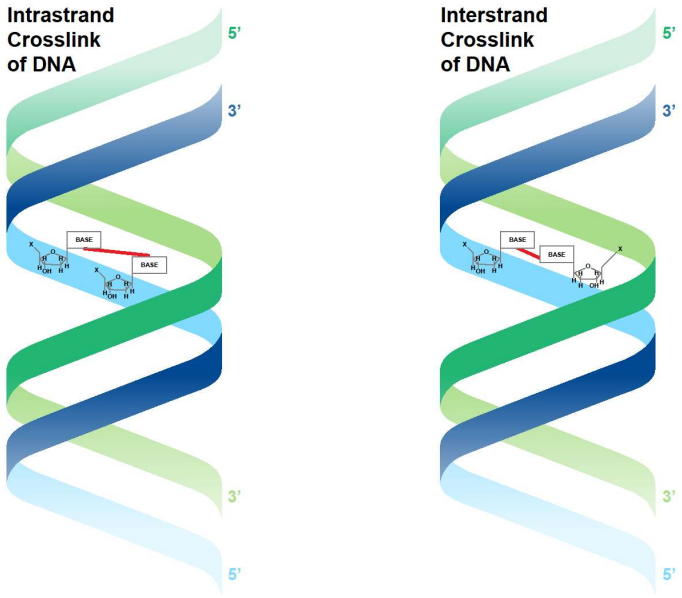


Figure 1-6 Schematic representation of Intrastrand Crosslink (left) and Interstrand Crosslink of DNA (right)

In contrast to Intrastrand Crosslinks that are covalent bonds formed between nucleotides in the same strand, Interstrand Crosslinks affect nucleotides in both strands of double helix.

It has been shown that for a mammalian cell with defects in relevant repair mechanisms, as few as 20 ICLs can be lethal (Murnane and Byfield, 1981).

Unlike previously described lesions, ICLs affect both strands of the double helix their repair and therefore cannot follow the usual scenario where undamaged strand is

used as a template for repair process. ICLs repair can be either erroneous or error-free and the mechanism depends on the type of the cell affected as well as cell cycle stage. Studies focusing on how ICLs are formed, recognized and repaired are of great clinical relevance as agents that induce ICLs were one of the earliest, and still are the most widely used forms of chemotherapeutic drugs.

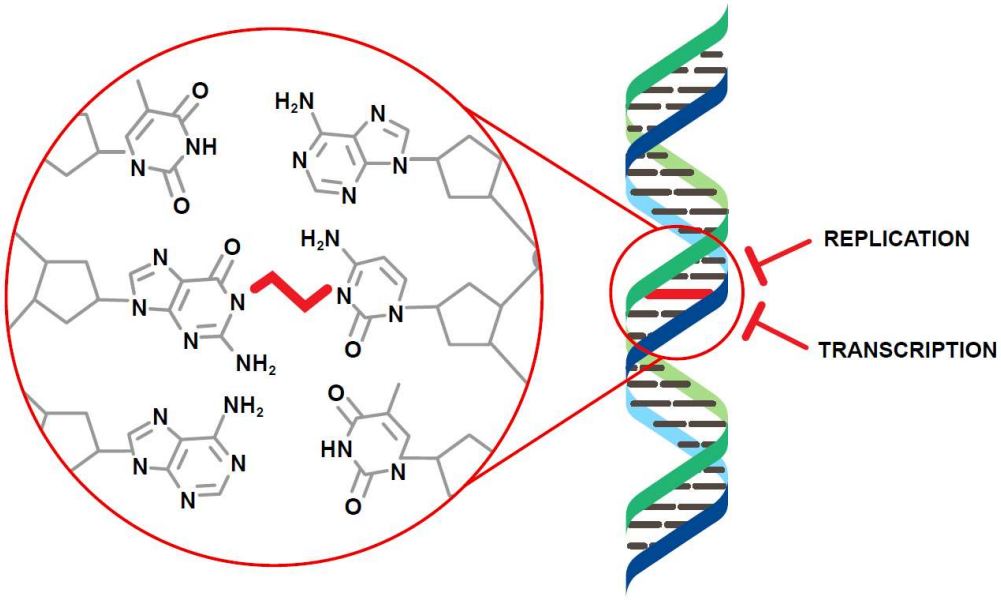


Figure 1-7 Schematic representation of interstrand crosslink (ICL).

ICL forms between bases in antiparallel strands of the double helix therefore preventing their separation and blocking crucial DNA processes such as transcription and replication.

1.2.1 ICL-inducing agents.

Formation of DNA Interstrand crosslinks is a two-step reaction and requires geometric alignment of the two electrophilic groups of the reacting chemical with nucleotides in the DNA double helix (Noll et al., 2006). Both, endogenous and exogenous sources of crosslinking agents are known, however, exogenous chemicals are much better characterized, mostly due to successful application as anticancer agents.

1.2.1.1 Endogenous sources of ICL-inducing agents

Until recently little has been known about internal sources of crosslinking chemicals. In the past ten years a considerable research effort has been put in this area and

several crosslinkers have been identified. Interestingly these molecules are frequently by-products of normal cellular metabolism. For example, alcohol and lipid metabolism leads to formation of aldehydes that may introduce crosslinks (Guainazzi and Schärer, 2010; Stone et al., 2008). Work from the Patel lab focusing on metabolism of reactive aldehydes revealed that ethanol can be a potent source of ICLs and that Fanconi anemia pathway is involved in counteracting its genotoxic effect (Langevin et al., 2011). In cells, ethanol is oxidized to a highly reactive acetaldehyde, which is then further metabolized by the ALDH2 enzyme. Effective catabolism of acetaldehyde is crucial for development of Fancd2$^{-/-}$ mouse embryos. Interestingly, acetaldehyde competent mothers can support development of Fancd2$^{-/-}$ Aldhd2$^{-/-}$ knockout embryos. However, double mutant mice are severely sensitive to alcohol and develop acute leukemia.

Another example of endogenous ICL sources are nitrates – a group of chemicals used in meat preservation. Their metabolism gives rise to nitrous acid that can be converted into nitric oxide – an important signaling molecule but also a potential crosslinking agent – it has been shown to covalently link the opposite guanines in DNA (Kirchner et al., 1992). Interestingly, these examples suggest that dietary habits affect the burden of endogenous sources of crosslinking agents. Recently a novel, potentially significant source of ICLs have been identified. As described in paragraph 1.1.2.2 spontaneous hydrolysis of the glycosidic bond in DNA leads to formation of abasic sites – a very abundant type of lesions that oscillates between two isoforms – hemiacetal and ring open aldehyde. Hemiacetal form can readily react with exocyclic amino group of adenine or guanine of the opposite strand thus creating an ICL. Due to the high frequency of abasic sites formation in human cells this mechanism may be an important contributor of endogenous ICLs, however further research is needed to confirm this conclusion (Dutta et al., 2007; Guan and Greenberg, 2009; Johnson et al., 2013).

1.2.1.2 Exogenous ICL-inducing agents

A number of chemicals produced by plants and fungi have been shown to contain two reactive groups that are prone to introduce covalent links between opposite

nucleotides in the DNA double helix. Additionally, in the second part of twentieth century many more ICLs agents were obtained synthetically to optimize cancer treatment.
It was during the Second World War when nitrogen sulphur compounds were used to produce bombs and poisonous gases. Clinical observation of the affected soldiers led to first clinical use of methyl-bis(β-chloroethyl)amine hydrochloride and tris(β-chloroethyl)amine hydrochloride for Hodgkin's lymphoma, lymphosarcoma & leukaemia in 1946 (GOODMAN & WINTROBE, 1946)

Although the therapeutic potential of ICL inducing agents has a long history, until recently little has been known about their biological mechanism of action and cellular response to DNA damage induced by these compounds. Research on ICL inducing agents is complicated by the fact that this class of chemicals induces a spectrum of lesions, ICLs being usually less than 10%. (Gargiulo et al., 1995). Some of the most commonly used ICL inducing agents include platinum compounds, nitrogen mustards, Mitomycin C and psoralens.

1.2.1.2.1 Platinum compounds.

Platinum (II)-based compounds consist of a generic planar, squared shape central region that defines the molecule's geometry and a number of *cis*- ligands that determines crosslinking activities. Currently there are three major platinum-based derivatives used as anticancer drugs – cisplatin, carboplatin and oxaliplatin (Yam, 2010).

Cisplatin (c*is*-diamminedichloroplatinum(II)) is the most broadly used compound. It enters the cell as inactive form and undergoes water-induced chloride displacement to form a biologically active species *cis*-$[Pt(NH_3)_2(H2O)_2]^{2+}$. Carboplatin (cis-diammine(1,1-cyclobutanedicarboxylato)platinum(II)) structurally differs from cisplatin in that instead of chloride the ligand is cyclobutane-1,1-dicarboxylic acid. Although carboplatin's reaction with DNA leads to formation of the same spectrum of products, the kinetics of DNA binding is slower, which may explain less severe side effects of carboplatin-based chemotherapy regimens. (Ali et al., 2013; Noll et

al., 2006). Oxaliplatin (oxalato(trans-l-1,2-diaminocyclohexane)platinum) contains two diaminocyclohexene (DACH) groups instead of the amine groups of cisplatin and oxalate leaving ligands instead of chloride atoms. In contact with water oxaliplatin as with other platinum compound undergoes aquation that leads to formation of its biologically active form – dichloro (DACH) platinum (Alcindor and Beauger, 2011; Woynarowski et al., 2000).

Platinum based compounds are known to react with N7-nitrogen of purines and the main products include intrastrand cross-links between guanine residues in 5′ -GG-3′, 5′ -GNG-3′, and between adenine and guanine in 5′ -AG-3′ sequences (Reishus and Martin, 1961). Although the exact structure of the ICL introduced by various platinum compounds may differ it is known that they all cause a significant deformation of the DNA helix, that locally mimics Z-form DNA (Noll et al., 2006; Takahara et al., 1995). Detailed structural studies involving NMR, X-rays and EMSAs showed that at the site of damage, the helix is unwound by 79° and bent by 45° (Coste et al., 1999; Huang et al., 1995; Malinge et al., 1994; Paquet et al., 1996).

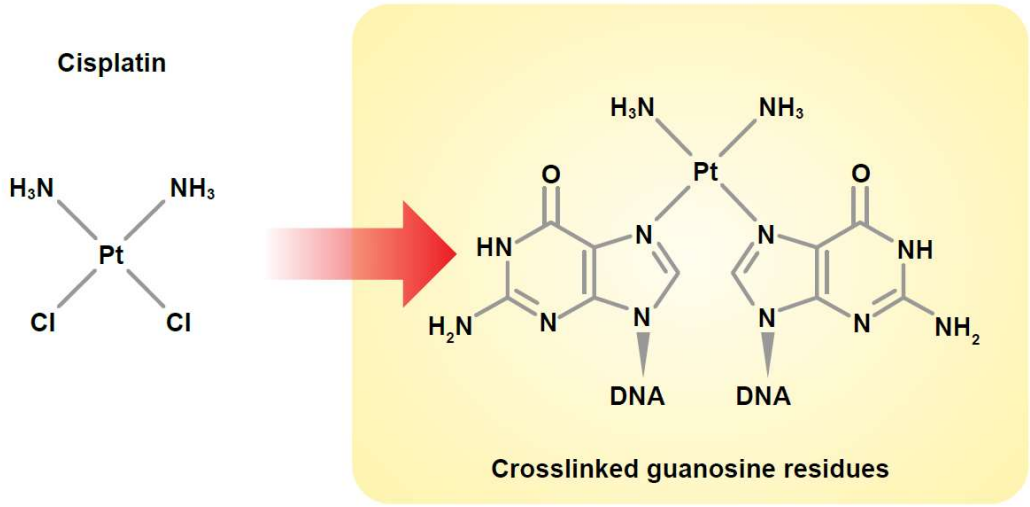


Figure 1-8 Cisplatin-induced ICL.

Schematic representation of cisplatin compound (left) and cisplatin linked guanosine residues of antiparallel DNA strands.

1.2.1.2.2 Mitomycin C

Mitomycin C (MMC) is a bifunctional alkylating agent produced by *Streptomyces caespitosis* that reacts with guanine residues of 5′-CG-3′ sequences through the minor groove of DNA. In its native form MMC is relatively inert, in cells it undergoes activation that requires reduction of its quinone ring (Gargiulo et al., 1995). Formation of ICLs requires a series of reactions and rearrangement involving production of highly reactive intermediates: hydroquinone that after tautomerization reacts with N2-amino group of guanine to produce monoadduct and vinylogous hydroquinone methide, which alkylates the guanine on the opposite strand of DNA therefore forming ICL (Tokuda and Bodell, 1987; Wijen et al., 2000; Yaghi et al., 1998). Although molecular modeling of MMC-induced ICL suggested only minimal perturbance introduced to the DNA helix, structural studies involving EMSA and NMR revealed that the minor grove undergoes limited widening (Norman et al., 1990; Rajski and Williams, 1998; Rink et al., 1996).

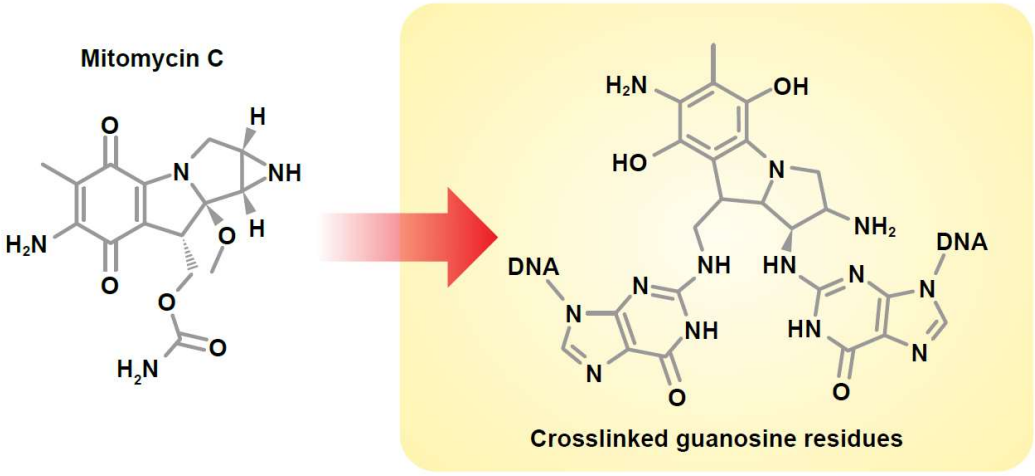


Figure 1-9 Mitomycin C induced ICL.

Schematic representation of MMC compound (left) and MMC linked guanosine residues of antiparallel DNA strands.

1.2.1.2.3 Nitrogen mustards

Nitrogen mustards is a group of bifunctional alkylating agents that contain a reactive N,N bis-(2-chloroethyl)amine group. Reaction of nitrogen mustards with guanine, typically in the sequence 5’CG involves formation of an aziridinium intermediate that

is then transformed into an alkylated guanine derivative (Falnes et al., 2002; Trewick et al., 2002). Newly formed guanine monoadducts can then engage into formation of another aziridinium intermediate that in turn can react, among other possibilities with other guanine residue to form the ICL (Falnes and Rognes, 2003; Moore et al., 1992; Rink and Hopkins, 1995).

Inter-strand crosslinks produced by nitrogen mustards are thought to introduce distortion into the DNA double helix. However, more data is needed to fully understand their structure. Structural studies conducted on mechlorethamine-induced ICL showed a bend in DNA of around 14˚(Rink and Hopkins, 1995).

1.2.1.2.4 Psoralens

Psoralen and its derivatives belong to a family furocumarins- of tricyclic, planar chemicals naturally occurring in plants (McHugh et al., 2001). Psoralens intercalate into DNA and upon activation with a long wave ultraviolet light can react with adjacent pyrimidines to form Interstrand cross-links (Hearst, 1981). Although psoralens stochastically bind to DNA, regardless of the actual base sequence, due to molecular structure they preferentially form adducts with thymines in 5′ -TA-3′ and 5′ -AT-3′ sequences. This is because when intercalated into B-Form helix the 5,6-double bonds of the thymines and 3,4-pyrone ring of the psoralens are aligned allowing for a light-induced cycloaddition reaction between pyrone/furan of the psoralen and thymine. This leads to formation of a cyclobutane-containing monoadduct. Upon absorption of another photon, the monoadduct can initiate a second cycloaddition reaction with the thymine located in the opposite strand. Interestingly it is only the furan-side (as opposed to pyrone-side) monoadduct that can absorb another quantum of light and form ICL (Edelson et al., 1987; Gasparro, 1996).

NMR studies carried out on a number of psoralen-induced interstrand cross-links suggest that although a considerable distortion is introduced at the site of damage, it is limited to a small number of nucleotides and the helix retains its native conformation 3 nucleotides up- and downstream of the ICL (Hwang et al., 1996;

Spielmann et al., 1995a; Spielmann et al., 1995b). No significant bending of the helix has been observed (Haran and Crothers, 1988).

1.2.2 ICL Repair mechanisms

Considering the molecular and mechanistic challenges of repairing lesions affecting both strands of the double helix, involvement of numerous factors representing diverse repair pathways is not surprising. Despite extensive research advances in the field of ICL repair, some aspects of this complex processes remain elusive. This is mostly due to the fact that ICL-inducing agents used in the laboratory introduce also other type of DNA damage. In fact, ICLs usually represent less than 8% of all induced lesions. It is extremely complicated to study how cells respond to ICL if many different types of lesions are being recognized and repaired at the same time. Additionally, as discussed below, ICL induced by various agents may differ in terms of structure and extent of distortion imposed on the DNA helix. This diversity may lead to some variations in the recognition and repair processes. Importantly eukaryotic cells employ different repair mechanisms depending on the cell cycle stage, further complicating interpretation of results.

Despite many differences between bacteria and eukaryotes, ICL repair generally follows two possible mechanisms in most organisms (Dronkert and Kanaar, 2001; Legerski, 2010; McHugh et al., 2001). A central event in both ways is initial ICL incision that leads to unhooking of the lesion and the main difference is involvement of HR. In the recombination-independent repair pathway the lesion is firstly excised from one strand and then the gap is filled with translesion bypass polymerases (Berardini et al., 1999; Berardini et al., 1997; Grossmann et al., 2001b). In *E.coli* the remaining monoadduct linked with the other strand of the helix is then removed by bacterial NER components and the gap is filled in by Polymerase II. In mammals this process is less understood, however NER also plays a crucial role (Wu et al., 2001). Homologous Recombination-independent pathway is the predominant repair mechanism in non-replicating cells.

The second pathway also begins with excision step, but it is followed by HR-mediated repair of the gap. Although in *E.coli* and yeast NER provides the nucleolytic activity necessary for excision, mammalian cells deficient for other NER components are only mildly sensitive to ICL-inducing agents (Andersson et al., 1996; Cole, 1973; De Silva et al., 2000b; Miller and Heflich, 1982).

Homologous Recombination-dependent ICL removal is the main repair mechanism in cycling mammalian cells. It takes place in S-phase of the cell cycle and is initiated by replication forks colliding with the ICL. Due to the key role bifunctional alkylating agents play in cancer therapy, the majority of ICL damage research has been focused on S-phase repair mechanisms. However, most cells in vertebrates are differentiated and non-dividing. It has been suggested that long-term cancer survivors who have been exposed to ICL-inducing therapeutic compounds suffer from Acquired Premature Progeroid Syndrome (APPS) characterized by decline of cognitive functions visual deterioration, musculoskeletal decline and osteoporosis (Grillari et al., 2007; Meinardi et al., 2000). Consistently similar symptoms of premature aging are observed in individuals diagnosed with loss of function mutations in genes implicated in ICL-repair such as XPF or Fanconi anemia (Grillari et al., 2007; Mitchell et al., 2003). Given that ICLs can also be formed by physiological metabolic processes (as discussed in chapter 1.2.1.2) replication-independent repair is the predominant ICL removal mechanism in non-replicating cells. Taken together, thorough understanding of ICL repair processes in both S- and G_1 phase may help to design therapies characterized by high toxicity to diving cells and reduced toxicity to benign, non-diving cells.

The differences between ICL repair during and outside of S phase

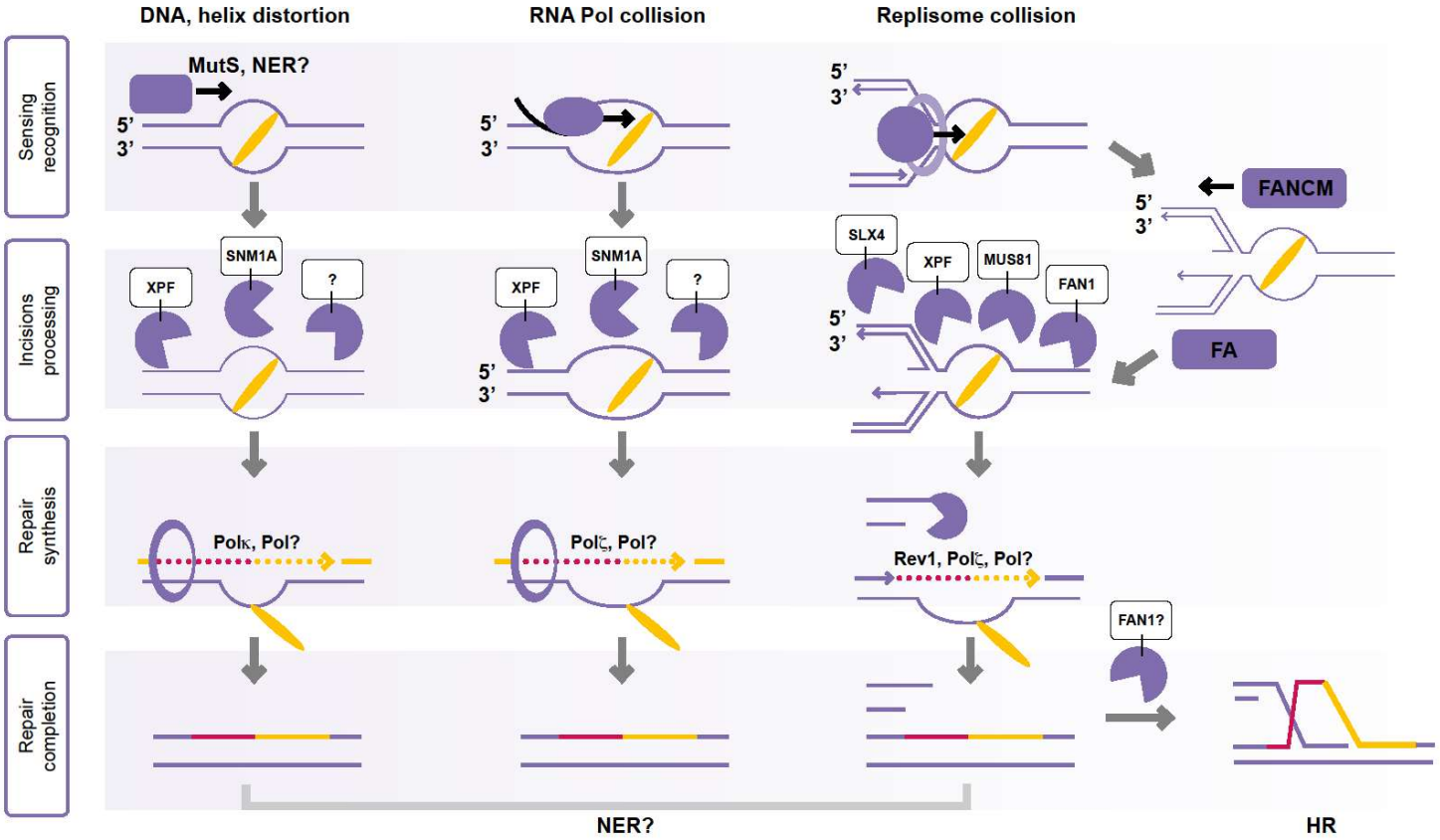


Figure 1-10 Schematic representation of differences in ICL repair during and outside of S phase

Although majority of ICLs is sensed and repaired during S phase when replisome **collides with the crosslink (right panel)**, other replication-independent mechanisms also contribute, in particular outside of S phase: **DNA helix distortion sensing (left panel)** and **RNA Polymerase collision with ICL (middle panel)**. Regardless of the cell cycle stage and nature of the relevant ICL repair mechanism, four universal steps can be distinguished: **ICL sensing and recognition, incisions and processing, repair synthesis and completion of the repair** (left). Key repair factors are depicted in simplified manner, detailed description of the repair pathways can be found in chapter 1.2.

1.2.3 Recognition of ICL

In order for the repair process to begin the sites where ICL has been formed must be recognized. In contrast with NER and BER pathways where adduct recognition mechanisms are well understood, ICL sensing remains elusive. Structural diversity of the ICL and differences in the degree of helix alterations led to a number of experimental discrepancies and contradicting observations. Although many alternative recognition mechanisms have been proposed in eukaryotes, in mammalian cells there are three ways for ICL recognition. Firstly, crosslinks can be sensed by proteins detecting local DNA bending imposed by the lesion. Alternatively, in non-cycling cells, ICL may be also recognized by the transcription machinery stalling at the site of damage. Finally, in cells undergoing division, replication forks colliding with ICLs trigger activation of repair mechanisms, which is the main recognition mechanism in S phase of the cell cycle.

1.2.3.1 Replication-independent recognition of ICL

As mentioned above in the absence of replication, cells sense ICL thanks to transcription polymerase collisions or by utilizing proteins that recognize local helical distortion at the site of cross-link. Despite substantial research efforts, ICL recognition processes are relatively well understood only in lower organisms. Surprisingly, NER appears to play a crucial role in those early stages of ICL repair. Given that cross-links covalently bind both strands of the double helix, no base-pairs can be flipped-out, which is a crucial step for NER factors binding to the unaffected strand at the site of lesion (as described in chapter 1.1.2.3). To explain this phenomenon, it has been proposed that ICL may locally induce some degree of unwinding that would create a binding site for proteins such as XPC.

1.2.3.1.1 The role of NER in ICL recognition

The NER pathway is essential for the ICL repair in both bacteria and yeast. Work in *Saccharomyces cerevisiae* implicated Rad4 (XPC homolog) involvement in ICL

recognition in G_1 phase (Lehoczký et al., 2007; Sarkar et al., 2006b; Wu et al., 2004). Given the convincing data from bacteria and yeast, it has been postulated that NER factors are also involved in ICL recognition in higher eukaryotes but studies on vertebrates led to some confusing conclusions. In experiments involving repair of cisplatin-crosslinked oligonucleotide duplex in cell-free systems with reconstituted NER factors, the initial incision was observed only in the case of intrastrand cross links, no effect was observed for ICLs (Zamble et al., 1996b). Work of Furuta and McKay showed on the other hand that although XPC-mutated mammalian cells are no more sensitive to cisplatin than wild type cells inactivation of post-recognition NER factors such as XPA, XPD, XPF and XPG leads to pronounced sensitivity (Furuta et al., 2002; McKay et al., 2001).
Taken together it appears that despite the strong structural distortion imposed by cisplatin-induced ICL, XPC plays a minor if any role in recognition of this type of damage.

The role of NER in psoralen-induced ICL recognition has also been investigated. In an *in vitro* study using triplex forming technology, XPC-RAD23B complex has been shown to recognize psoralen-induced cross-links. Additionally XPC-RAD23B and XPA-RPA protein complexes were also observed to bind psoralen ICLs simultaneously, demonstrating not only that psoralen-induced ICLs can be sensed and bound by XPC-RAD23B alone, but also that XPA-RPA protein complexes may act cooperatively with XPC-RAD23B forming a multimeric NER complex at the site of damage (Thoma, 2005). The main caveat of this study is the fact that we do not know how these protein complexes would respond to psoralen-induced crosslinks without conjugation to the triplex nucleotide, which obviously are not a physiological type of damage.

In an interesting study, Muniandy and colleagues showed that XPC is indispensable for ICL repair and that the protein is rapidly recruited to the stripes of laser-induced TMP crosslinks in G_1 phase of wild type mammalian cells (Muniandy et al., 2009a). On the contrary, experiments with plasmids containing a single psoralen cross-link located in the promoter of a reporter gene showed that although NER proteins were required for efficient expression of the reporter, XPC was not as important as other factors such as XPA-ERCC (Muniandy et al., 2010).

The mechanism of mitomycin C - induced ICL recognition remains unclear. In plasmid-repair based experiments analogous to the one described above, XPC deficiency led to pronounced decrease in repair efficiency, comparable with the effect observed for other NER factors (Zheng et al., 2003). Given what is known about XPC mechanism of action it is a surprising observation as mitomycin C-induced ICL are the least helix-distorting lesions. At the same time XPC deficient cells were only mildly sensitive to cisplatin, although crosslinks introduced by this compound introduce a significant helical distortion. Taken together the role of XPC in ICL recognition is not in line with what could be expected based on current state of knowledge. It has been postulated that ICL recognition by XPC may require an additional factor that facilitates DNA binding. One possible candidate could be HMGB1 protein that has been shown to be involved in XPC-driven recognition of psoralen crosslinks (Lange et al., 2009).

The role of MMR proteins in replication-independent ICL repair remains unclear. In yeast MutS complexes are supposed to be involved during lesion unhooking rather than in lesion recognition (Barber et al., 2005a). Furthermore, MSH2 is dispensable for repair of laser-induced psoralen ICL in G_1 phase of mammalian (Muniandy et al., 2009a). In the same study it was shown that although MSH3 is recruited to the psoralen ICL it only comes into play after XPC, which is thought to be the recognition factor in this process. In contrast, MSH2 has been shown to play a crucial role in mitomycin C and psoralen crosslinks repair in mammalian cells (Fiumicino et al., 2000a). In vitro experiments showed MutSβ binding to oligonucleotide duplexes with cisplatin engineered ICL (Zhu and Lippard, 2009). In another study MutSβ has been shown to bind with high affinity to oligonucleotides containing TMP induced crosslinks in PCNA-dependent manner (Zhang et al., 2002b).

Work from the Lambert group identified an unexpected player in ICL recognition – in a series of experiments they showed that α-Spectrin II binds psoralen cross-links. Furthermore siRNA-mediated silencing of α-Spectrin II leads to cross-linking agent sensitivity and chromosome aberrations (McMahon et al., 2001; Mcmahon et al., 2009). Interestingly in the same study a number of FA proteins (FANCA, FANCC

and FANCG) also showed affinity to DNA with psoralen-induced ICL (McMahon et al., 2001).

1.2.3.2 Transcription-coupled ICL repair

In contrast to monoadduct, ICL completely block progression of transcriptional machinery on DNA thus in actively transcribed regions of the genome this can be a potent mechanism in which ICL are recognized (Islas et al., 1991; Larminat et al., 1993). In general, there are two ways by which transcription can trigger repair. In the first mechanism, transcription polymerases stall on the lesion, which activates repair pathways. Additionally, chromatin opening, and relaxation allows access of repair factors to DNA. Naturally, the first mechanism is CSA- and CSB-dependent, and would preferentially repair lesions on the actively transcribed strand. The second mechanism is strand-independent (1.1.2.3).

The majority of transcription-coupled ICL repair studies have been conducted with help of Host Cell Reactivation assay. In such experiments a plasmid with a single, site-specific ICL located between a strong promoter and a reporter gene is introduced to mammalian cells and expression of the reporter (usually luciferase) is used as surrogate of repair efficiency. Unfortunately, expression plasmids are inherently biased towards transcription, which complicates and blurs interpretations of such experiments. On the contrary, work by Muniandy and colleagues on randomly introduced psoralen crosslinks in G1 phase cells showed that there was a dependence on XPC - global NER factor (Muniandy, Thapa, Thazhathveetil, Liu, & Seidman, 2009). This result clearly show that global repair mechanisms may also contribute to ICL repair in the absence of replication.

The first studies looking at the role of transcription-coupled repair of ICL were carried out by Hanawalt's group in the early 1990s and showed that in intensively transcribed regions of the genome psoralen ICL formation and repair processes are more rapid (Islas et al., 1991). Early studies using reporter plasmids suggested NER factors play an important role in transcription-coupled recognition of ICLs (Islas et al., 1994; Vos and Hanawalt, 1987). These results have been reinforced by more recent work by Enoiu and colleagues in which repair of cisplatin-induced ICLs was reported to

depend on transcription-coupled NER protein, CSB as well as the general nucleotide excision repair proteins XPA, XPF and XPG. XPC (GG-NER factor) was not involved in this process (Enoiu et al., 2012b). Transcription-coupled repair is influenced by factors such as size and shape of the lesion as well as local DNA sequence (Scicchitano et al., 2004). Although it is not entirely known whether the distortion introduced to the double helix impacts where the polymerase stalls and repair is initiated, work of Smeaton and colleagues show that the structure of the lesions determines the repair process (Smeaton et al., 2009).

1.2.3.3 Post recognition in the absence of replication

As discussed above, the crucial step of the entire ICL repair process is dual incision on both sides of the lesion that leads to unhooking of the strands. An excellent tool to study dynamics of the unhooking event in living cells is the alkaline comet assay. In cells, DNA is condensed and packed in the nucleus in the form of super-coiled loops. Introduction of DNA breaks relaxes the strands and facilitates unwinding in alkaline pH. When an electric field is applied, relaxed DNA fragments migrate towards the anode. In physiological conditions enough spontaneous DNA breaks is formed to reveal the characteristic comet pattern when single cells are subjected to electrophoresis. Compact, super-coiled DNA forms the comet's head whereas the nicked DNA fragments migrate forming the tail-like pattern.

Cells treated with crosslinking agents present with a different pattern where all DNA is bound by ICLs and no tail is observed. As unhooking releases the DNA, this restores the comet pattern, so can be used as a method to monitor efficient ICL uncoupling. Importantly DNA monoadducts introduced by crosslinking agents do not interfere with the read out in the comet assay allowing dissecting of ICL repair pathways. Furthermore, the assay can be performed on synchronized cells. As such, dynamics of ICL uncoupling in specific cell cycle stages can be monitored. What should be considered when working with comet assays is that agents causing single or double strand break formation can restore the tail hence blurring the results. Consequently, apoptotic cells should not be analysed as associated DNA fragmentation profoundly contributes to tail size.

In a recent study by Iyama and colleagues, the comet assay was successfully used to show that CSB (Cockayne Syndrome B protein) depletion leads to perturbation in replication-independent unhooking of ICLs in G1 phase of mammalian cells. The authors postulated that CSB can stimulate the SNM1A exonuclease activity on single- and double-stranded oligonucleotides by enhancing substrate affinity. Additionally, as SNM1A recruitment to ICLs was decreased in CSB-deficient cells, the authors suggested that CSB is key to coordinate efficient SNM1A assembly at damage sites (Iyama et al., 2015). SNM1A protein has been shown to load at a single-strand nick and digest past the ICL to facilitate the TLS step (Wang et al., 2011a). Earlier studies of replication independent nucleolytic processing of psoralen and nitrogen mustard ICLs revealed that unhooking in G1-enriched cells was considerably more efficient when compared with asynchronous population and measured by the comet assay (Rothfuss and Grompe, 2004). Similar studies by Richards and colleagues showed tail recovery is severely affected in cell deficient of XPD, which suggests NER factor involvement (Richards et al., 2005)

Given the requirement of NER factors in ICL repair in G1 phase, the role of incision nucleases XPG (3' incision) and ERCC1-XPF (5' incision) is particularly interesting in the context of the unhooking mechanism. Although XPF and ERCC1 deficient cells are extremely sensitive to crosslinking agents, XPG deficiency itself leads to only a minor sensitivity (De Silva et al., 2000a; Hoy et al., 1985). As described in chapter 1.1.2.3 in canonical NER the two DNA strands are unwound around the lesion. However, with ICL, unwinding is unlikely without incision, so it is unclear how NER factors would be able to perform incision on both sides of the damaged lesion. Indeed, in already cited work of Zamble and colleagues, synthetic ICL were not a substrate for NER-mediated incision in cell-free extracts (Zamble et al., 1996a). Psoralen and alkyl ICL on the other hand underwent dual incision 5' of the lesion (Bessho et al., 1997; Mu et al., 2000) leading to a futile repair cycle. Whether such intermediates are substrate for further nucleolytic processing to facilitate crosslink removal *in vivo* remains to be uncovered.

In a very insightful study, Williams and colleagues used *Xenopus laevis* cell-free extract to identify key components of the replication independent ICL repair pathway

(Williams et al., 2012). The authors showed that Pol κ (member of the Y family of polymerases) is recruited to a site-specific ICL and that its catalytic activity is essential for repair. This observation implies that Pol κ performs non-redundant DNA synthesis in ICL repair. It has also been shown that although Pol κ UBZ domains are essential for ICL repair, mutations of the PIP box sequence (which is a potential PCNA binding motif within Pol κ) do not affect the repair efficiency. In the same work, the authors provided evidence that in the absence of replication, PCNA is recruited to ICL-containing DNA and that this requires lysine 164 mono ubiquitination. Although most of the experiments were carried out in Xenopus membrane-free cytosol (HSS extract), which does not support plasmid replication, the authors recapitulated their findings in MEFs using a modified cell reactivation assay.

1.2.3.4 Replication dependant ICL repair

It generally accepted that the main ICL repair pathway in vertebrates occurs in the S phase of the cell cycle (Zhang and Walter, 2014). This complex mechanism requires concerted actions of several repair pathways including the Fanconi anemia (FA) pathway, structure-specific endonucleases, translesion DNA polymerases and recombinases (Clauson et al., 2013b). According to the current model, repair mechanisms are activated as DNA replication forks collide with the ICL. This event leads to formation of a DNA substrate recognized by structure-specific endonucleases, immediately recruited to the lesion. Structure specific nucleases process the fork, which leads to unhooking of the daughter duplex and generation of a double-stranded DNA break. The unhooked ICL is then bypassed thanks to activity of translesion DNA polymerases. Finally, the fork is restored via homologous recombination process.

Replication-dependent ICL repair is therefore a unique situation in which stalled replication forks undergo programmed and strictly controlled collapse, a feature that directly reflects the complexity of ICL repair and distinguishes these processes from other repair pathways. A large body of evidence show that these complex processes are synchronized and orchestrated by the Fanconi Anemia pathway (Knipscheer et al., 2009a).

1.2.3.4.1 ICL repair and the Fanconi Anemia pathway

Fanconi anemia (FA) is a rare genetic disorder inherited as a Mendelian recessive trait caused by biallelic mutations in any of 19 FANC genes (complementation groups). While mutations in most complementation groups manifest clinically with the full spectrum of FA-associated phenotypes (congenital abnormalities, early onset bone marrow failure, increased risk of malignancies such as acute myeloid leukemia and solid tumors), some complementation groups exhibit only subsets of these characteristics (FANCD1, N, O, S and R) (Auerbach, 2009; Clauson et al., 2013a; Duxin and Walter, 2015)

In brief, Fanconi anemia proteins coordinate various steps of the ICL repair and can be divided into three groups based on their role in the process. Group I includes FANCA, B, C, E, F, G, L, and M, as well as three so called Fanconi anemia associated proteins (FAAP20, FAAP24 and FAAP100). Members of this group assemble into a large multi-subunit FA core complex that serves as an E3 ubiquitin ligase. The main role of the FA core complex is to bind to chromatin and monoubiquitylate the group II members: FANCD2 and FANCI heterodimer (also called the ID complex) (Alpi et al., 2008; Garcia-Higuera et al., 2001; Joo et al., 2011; Smogorzewska et al., 2007)

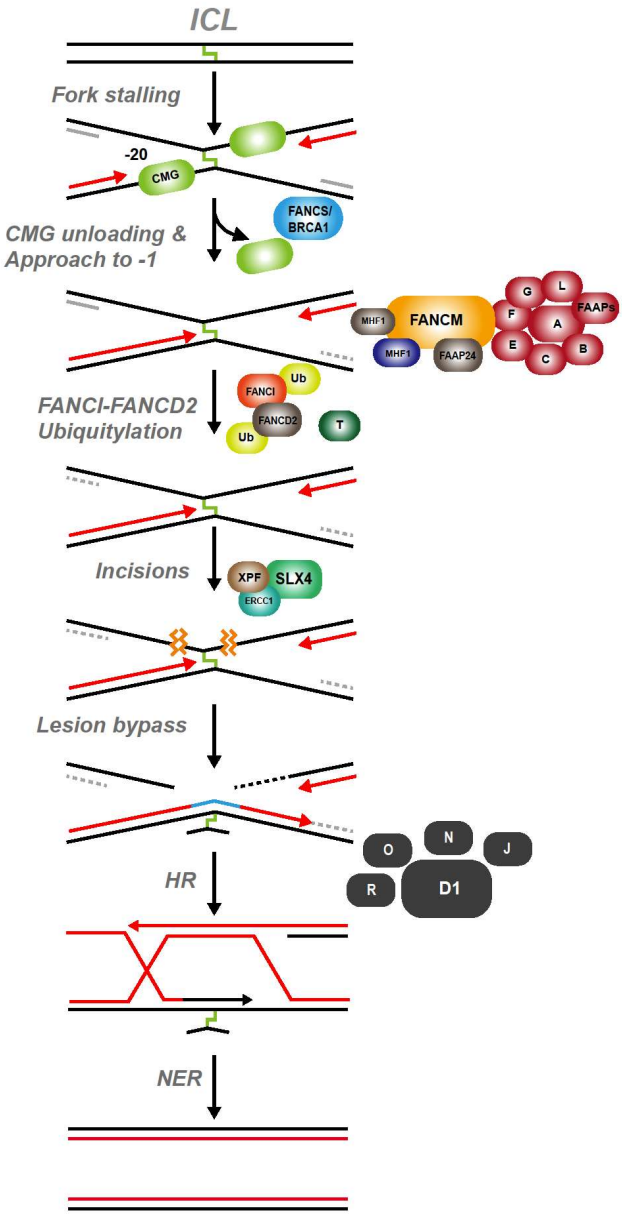


Figure 1-11 Schematic representation of FA role in ICL repair.

ICL repair is initiated when replication forks converge at a single ICL creating x-shape structure (**Fork stalling**). Due to the footprint of the CMG helicase complex the leading strands initially stall 20–40 nt in front of the lesion. CMG complexes then undergo BRCA1-dependent ubiquitylation, and p97 segregase evicts ubiquitylated CMG from chromatin, allowing approach of the leading strand to 1 nt of the lesion (**CMG unloading & Approach to -1)**. At the same time FA core complex assembles and ubiquitylates FANCI-FANCD2 heterodimer, which binds near the lesion and recruits structure specific endonucleases XPF(FANCQ)-ERCC1 and SLX4 (FANCP) (**FANCI-FANCD2 Ubiquitylation).** Double incisions by the recruited endonucleases unhooks the ICL (**Incisions**) allowing a complex of the translesion DNA synthesis polymerases to bypass the lesion (**Lesion bypass**). Finally, the double-stranded break generated by incisions

is repaired by homologous recombination **(HR)** and two double strands are recreated with the help of NER **(NER)**.

Monoubiquitylated ID complex is essential for downstream repair that is executed by the group III FA proteins. Group III includes the following members: the scaffolding protein SLX4 (FANCP) (Kim et al., 2011; Stoepker et al., 2011), XPF nuclease (FANCQ) (Bogliolo et al., 2013), as well as a number of the homologous recombination factors PALB2 (FANCN) (Reid et al., 2007), BRCA2 (FANCD1) (Howlett et al., 2002), RAD51, RAD51 C (FANCO) (Vaz et al., 2010), BRCA1 (FANCS) (Sawyer et al., 2015) and FANCJ (BRIP1) (Levran et al., 2005). To complete the repair cycle and deactivate the FA pathway USP1 protein (ubiquitin specific peptidase 1) together with its activating factor UAF1 deubiquitinates the ID heterodimer (Cohn et al., 2009; Nijman et al., 2005).

1.2.3.4.2 ICL Recognition in the context of replication

As ICLs covalently link the Watson and Crick strands of DNA duplex they prevent strand separation and form an impenetrable obstacle for the replicative helicases. As the replication fork (or possibly two replication forks) approaches the lesion, the complex slows down and eventually stalls triggering replication-depended ICL repair processes. A number of factors can contribute to replication forks stalling; hence the cell needs to distinguish the molecular context of ICL-dependent event from other reasons and to direct the repair to an ICL pathway.

It is thought that FANCM together with its associated proteins play the role of molecular sensor that detects ICL-induced fork stalling, bind to unwound DNA and recruit the downstream repair factors. FANCM is a highly conserved member of the XPF-heterodimeric 3'-fap endonuclease family. The subcomplex formed by FANCM, FAAP24 as well as MHF1 and MHF2 (also known as FAAP16 and FAAP10) has been shown to bind branched DNA structures *in vitro* (Ciccia et al., 2007; Singh et al., 2010; Yan et al., 2010). Although due to changes in key residues FANCM has lost its nuclease activity (Niedernhofer, 2007) the protein is an ATP-depended translocase that promotes migration of both Holiday junction and replication fork branch points (Gari et al., 2008a; Gari et al., 2008b; Rosado et al., 2009).

FANCM-dependent translocation leads to exposition of single stranded DNA that is immediately bound by the RPA protein (Huang et al., 2010; Vare et al., 2012). RPA accumulation at the side of ICL is a prerequisite for the key event in DNA repair checkpoint activation: ATR kinase recruitment and activation (Ben-Yehoyada et al., 2009; Zou and Elledge, 2003). Once activated ATR triggers a number of downstream processes, one of the more important being phosphorylation and thus activation of the CHK1 kinase that then blocks cell cycle entry into mitosis (Cui et al., 2009). Interestingly although FANCM has been shown to promote model fork reversal it has not yet been established whether it directly interacts with the stalled fork *in vivo*.

Although in cells depleted of FANCM we do not observe the key ICL repair events such as FA core complex recruitment or FANCD2-FANCI monoubiquitylation (Kim et al., 2008) this has to be seen in the broader experimental context. A great deal of work on FANCM function has been done in EUFA867 cell line derived from the only *FANCM* biallelic mutated patient identified to date. Strikingly, given the statistical odds of such coincidence, detailed genetic analysis revealed the cells also harbour a biallelic mutation in another FA factor – FANCA, which blurs interpretation of the results. Although complementation of the FANCA defect in EUFA867 cells did not change their hypersensitivity to MMC, in contrast to cells defective in other FA core complex members, FANCA-corrected EUFA867 cells were proficient in FANCD2 monoubiquitination (Singh et al., 2009). Additionally FANCM deletion in chicken DT40 cells as well as isogenic human cell lines did not completely abrogate FANCD2 monoubiquitination (Rosado et al., 2009). It is therefore very likely that FANCM is not the exclusive FA activator and other factors play redundant role in this process.

An obvious group of candidate proteins are human MutS homologs that work in the MMR mechanism and recognize mismatched bases in DNA duplex (Huang et al., 2011). Huang and colleagues showed that cells with MSH2, MSH3 and MSH6 siRNA-mediated depletion treated with MMC fail to effectively recruit FA core complex and to monoubiquitinate FANCD2-FANCI heterodimer. Importantly double depletion of FANCM and MSH2 resulted in a more severe phenotype than depletion of any of the proteins alone. Although caution is needed when interpreting effects of double siRNA depletion, these experiments indicate MSH2 and FANCM may play a

similar role in ICL-dependent FA activation. This observation is consistent with the fact that MMR factors have been shown to physically interact with the FA proteins (Williams et al., 2011). Moreover, studies in *in vitro* mammalian cell extracts have shown that MutSβ is required for ICL-induced DNA synthesis (Zhang et al., 2002a). Genetic studies in buddying yeast by Barber and colleagues further support a role for MutS factors in ICL repair as MSH2 was found to be involved in the processing of ICLs, a step required prior to the repair of ICL-induced DNA double-strand breaks that form during replication (Barber et al., 2005b). In a more recent study FANCM ortholog Mph1 was found to physically and functionally interact with the MutSα mismatch repair factor (Ward et al., 2012). However, the role of MSH2 still remains contentious as contradicting data concerning sensitivity of mutant cells is available. Although in a number of experiments lack of MSH2 conferred sensitivity to crosslinking agents (Aquilina et al., 1998; Fiumicino et al., 2000b; Wu et al., 2005) more recent studies showed no such effect (Enoiu et al., 2012a; Hlavin et al., 2010; Muniandy et al., 2009b). This discrepancy can be explained by the fact that as shown with other repair factors MSH-2 may play a redundant role in response to certain type of ICL-induced structure.

The process of ICL recognition is still a field of intensive research and it is very plausible that new players will be identified. For instance, in a very recent study UHRF1 protein has been characterized as a novel ICL recognition factor (Tian et al., 2015). UHRF1 is ubiquitin-like with PHD and RING finger domain 1 protein that was shown to be recruited to DNA lesions *in vivo* and to bind directly to ICL-containing DNA. Cells deficient in UHRF1 were hypersensitive to crosslinking agents. Additionally, in the absence of UHRF1 ICL lesion processing was delayed and ineffective most likely due to reduced recruitment of repair nucleases to the damaged site. UHRF1 was also shown to interact with ICL repair nucleases: MUS81 and ERCC1. Interestingly depletion of UHRF1 and components of the Fanconi anemia pathway led to increased crosslinking sensitivity as compared to the defect observed in single depletion. To summarize, the finding so far implicate UHRF1 in ICL recognition processes where it promotes recruitment of ICL-processing factors as it works as a nuclease recruitment scaffold. Most likely UHRF1 protein activity is parallel to the FA pathway.

It is conceivable that it is both: the ICL itself (and the DNA structural aberration it induces) and the stalled replication fork act independently in S phase to signal presence of the lesion that activates repair processes. Helix distortion induced by the ICL would attract MutSβ (and possibly other factors) that would lead to recruitment of lesion processing complexes and downstream repair factors. Stalled replication fork on the other hand would activate Fanconi anemia pathways that would stabilize the fork and locally remodel the chromosome and then recruit translesion synthesis and homologous recombination factors.

As mentioned above upon replication fork stalling group I Fanconi anemia proteins assemble into a large nuclear complex called the core complex. Correct core complex assembly is necessary for the key Fanconi anemia pathway event – ID complex monoubiquitination that triggers the downstream signalling and repair. Interestingly out of the 11 core complex proteins possess a recognized enzymatic activity. It has been postulated that the role of majority of the core complex members is to stabilize the complex via protein-protein interaction. Interestingly it has been observed by several groups that members of Fanconi anemia core complex (or group I) can engage in formation of smaller sub-complexes. For instance, FANCA, FANCG and FAAP20 form a subcomplex, where TPR motifs of FANCG are thought to function as a possible scaffold and the UBZ domain of FAAP20 provides affinity to ubiquitinated histones (Ali et al., 2012; Leung et al., 2012; Leveille et al., 2004; Yan et al., 2012). Another subcomplex is formed by the FANCB, FANCL and FAAP100 proteins (B-L-100) (Medhurst et al., 2006). It has the E3 ligase activity of FANCL but the function of FANCB and FAAP100 remains unclear (Ling et al., 2007). Third of the identified sub-complexes is made up by FANCC, FANCE, and FANCF proteins(C-E-F). FANCF was also postulated to perform an adaptor role (Leveille et al., 2004).

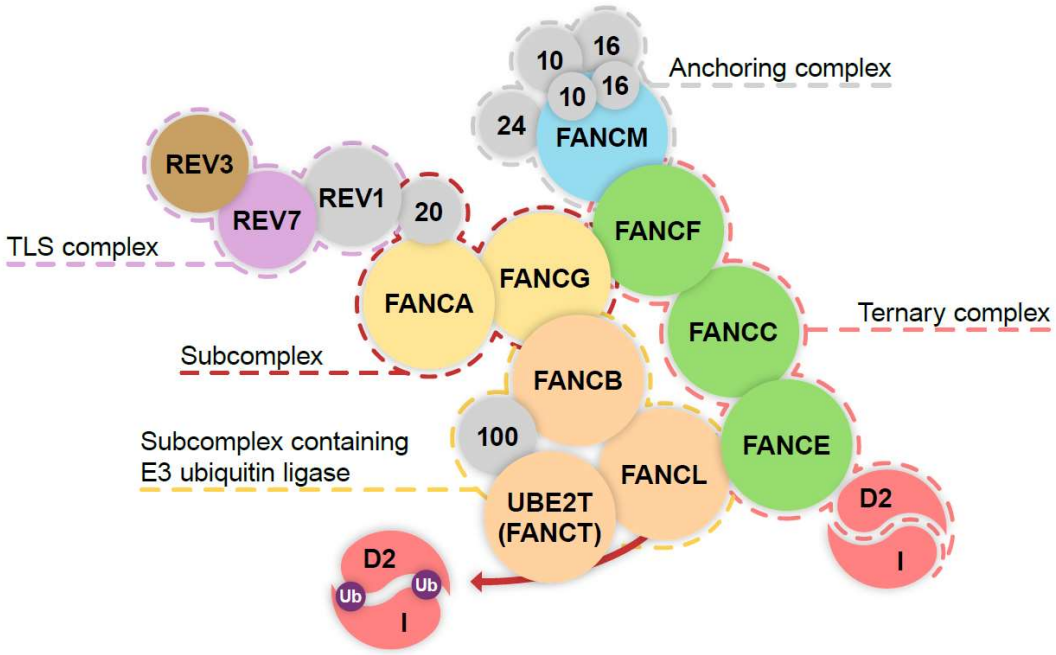


Figure 1-12 Architecture of the FA complexes including FAAPs (Fanconi anemia-associated proteins).

Fanconi anemia proteins form a spectrum of subcomplexes interconnected by structural and functional links. For example, the-FANCG-FANCA- FAAP20 subcomplex links translesion synthesis complex and the FA pathway as FAAP20 interacts directly with REV1 component of TLS. FANCA stability increases upon binding to FAAP20 due to reduced SUMOylation and RNF4-mediated degradation. The ternary complex formed of FANCF-FANCC-FANCE bridges FANCD2-FANCI heterodimer with FANCM and FAAPs complex that recognizes ICL and binds to DNA. FANCD2 becomes a substrate for monoubiquitination performed by FANCB-FANCL-FANC100-UBE2T complex.

In a very insightful recent study by Huang and colleagues, the FA core complex was shown to be a highly dynamic structure with extensive interplay between the subcomplexes (Huang et al., 2014). As not all components of the classical FA pathway are conserved in the organisms traditionally used to study interactions between genes such as *S. cerevisiae* or *D. melanogaster* the authors generated a set of human and chicken FA genes isogenic single and double mutant cell lines and performed comprehensive epistasis studies. In their work they showed that the core complex is assembled in a modularized manner. Consistently with previous studies, their findings confirmed the existence of the three subcomplexes. However, they clearly showed that not all of the subcomplexes contribute equally to the function of the core complex and resistance to DNA damage. In brief, the authors postulated

that although L-B-100 subcomplex is indispensable for E3 ligase activity, the A-G-20 as well as C-E-F subunits are dispensable as their main role is to increase DNA binding of the core complex. Given that until recently our understanding of the FA core complex dynamics was that disruption of any of its components completely abrogates ID ubiquitination and downstream repair, the work of Huang and colleagues sheds a new light on early steps of the FA pathway activation.

The crystallographic studies of FANCL revealed 3 functional domains within the full length protein: E2-like fold domain in the N-terminus, a central double RWD domain, and a C-terminal RING domain (Cole et al., 2010). Although the E2-like fold domain is found in all FANCL homologs its function is not entirely clear. Miles and colleagues postulated that it facilitates a non-covalent interaction between FANCL and ubiquitin moieties. The RING domain was shown to interact with UBE2T, the E2 conjugating enzyme of the FA pathway (Hodson et al., 2011) and the double RWD domain provides the substrate binding site (Cole et al., 2010; Hodson et al., 2011).

Monoubiquitination of the ID complex is clearly the key event in FA pathway activation and has a crucial role for downstream repair, recruitment of DNA processing endonucleases as well as transcriptional activation of tumour suppressor genes (Collis et al., 2006; Deans and West, 2009; Park et al., 2014; Sugahara et al., 2012). It seems that maintenance of the ubiquitin groups on FANCD2 and FANCI depends on both proteins being modified (Smogorzewska et al., 2007). Upon monoubiquitination FANCD2 translocates into chromatin foci that are considered sites of DNA repair (Garcia-Higuera et al., 2001). FANCD2 and FANCI consist of 1451 and 1328 amino acids, respectively. Sequence analysis revealed that although neither of the proteins harbours any recognizable enzymatic or functional motifs although they share a homologous stretch of around 150 amino acid residues flanking the monoubiquitination sites. In fact, FAND2 and FANCI have been proposed to be paralogs (Smogorzewska et al., 2007). In 2011 a high resolution crystal structure of the ID complex became available revealing some interesting insights into DNA binding and activity of the complex (Joo et al., 2011). Due to similar shapes and electrostatic charge of DNA binding groves of both FANCI and FANCD2 it has been postulated that both proteins have comparable affinity for DNA. Crystal structure of the FAND2-FANCI complex suggests there are two functional sets of

DNA binding motifs arranged in a pseudo-twofold symmetric manner. It therefore has been suggested by the authors that the ID complex binds to two dsDNA/ssDNA structures cross-linked by an ICL. When replication fork stalls at the ICL, a splayed Y DNA structure is initially formed and then when leading strands advance to direct proximity of the ICL it generates a cross-linked 5' flap structure (Knipscheer et al., 2009a; Räschle et al., 2008). If a two replication forks collision scenario is considered (as described in section 1.2.3.5) the ICL would be flanked from both sides by either two 5' flaps or one 5' flap and one Y structure. Interestingly, both of these structures would spatially align with the symmetric pseudo-twofold arrangement of the DNA binding site of FANCD2-FANCI complex.

Crystallographic analysis revealed that monoubiquitination and phosphorylation sites are located on the heterodimerization interface, which implies these modifications are applied on monomeric proteins or an opened-up conformation of the complex and may stabilize the dimer. In 2016 first cryo-EM structure of FANCD2-FANCI was obtained adding more resolution to our understanding of domain organisation of the complex (Liang et al., 2016). It revealed that the heterodimer has a pocket‐like main body, a thumb‐like structure consisting of the C‐terminus of FANCI and a distinctive fork‐like tower consisting of the C‐terminus of FANCD2 component. Pathogenic, mutations in the Tower domain are found in several FA patients. The study also revealed that recruitment of the complex to stalled replication fork triggers monoubiquitination event.

Taken together, DNA binding activity of the FANCD2-FANCI complex allows it to recognize a range of possible DNA structures that occur when replication fork stalls. The complex is thought to bind to the lesion not only to stabilize and protect it but also to provide specificity for initial incision by recruiting endonucleases. FANCD2 monoubiquitination does not depend on nucleolytic processing of the ICL (Bhagwat et al., 2009; McCabe et al., 2008) but is required for this step to occur (Knipscheer et al., 2009b).

1.2.3.4.3 Nucleolytic processing by Structure-Specific endonucleases

As mentioned before, due to the complex nature of ICLs a minimum of two incision steps on both sides of the lesion are required in order to promote DNA synthesis past

the lesion. A range of nucleases have been shown to play a role in ICL repair, but we still do not have enough insight into mechanistic aspects of the process. To date three endo/exo nucleases (FAN1, SNM1A, SNM1B) and a number of structure-specific endonucleases (XPF-ERCC1, MUS81-EME1, SLX4-SLX1) have been shown to contribute to ICL repair. Little is known though about exact nuclease being required for each of the multiple incision and resection events in the repair process, as well as the extent of redundancy between the enzymes XPF-ERCC1.

XPF-ERCC1 endonuclease was shown to cut double-stranded DNA adjacent to 3' single strand region and is best known for its crucial role in the NER pathway. Its role in ICL repair outside of the S-phase is discussed in chapter 1.2.3.1. Both ERCC1 and XPF mutants are considerably more sensitive to ICL inducing agents than other NER factors. Moreover mutation of ERCC1 residues involved in interaction with XPA affects NER but has no effect on ICL repair. Cumulatively this data suggest distinct roles of the ERCC1-XPF complex in various DNA repair pathways regulated by protein-protein interactions (Orelli et al., 2010). In 2013 XPF was identified as a bona fide Fanconi anemia member – FANCQ (Bogliolo et al., 2013; Kashiyama et al., 2013). In replicating cells treated with ICL agents absence of XPF-ERCC1 activity leads to formation of DSBs. This observation suggests that XPF-ERCC1 is not responsible for the initial, unhooking incision. Seemingly contradictory observations were made by the Knipscheer's group – in XPF-depleted *Xenopus* egg extract ICL-stalled replication forks were not at all incised (Klein Douwel et al., 2014). This discrepancy can potentially be explained provided XPF is indeed responsible for the initial incision in both experimental system but in replicating cells, the persisting fork becomes a substrate for MUS81 nuclease (A. T. Wang et al., 2011b) (Wang et al., 2015).

MUS81-EME1

Cells with defect in MUS81-EME1 are hypersensitive to ICL inducing agents but this phenotype is not as severe as for XPF-ERCC1-mutated lines (Crossan et al., 2011; A. T. Wang et al., 2011a)

This observation indirectly suggests that MUS81-EME1 plays a secondary role in the ICL processing. Similarly, as with XPF-ERCC1 the exact step where MUS81-EME1 is involved is not entirely clear. Initial work by Hanada and colleagues showed the

heterodimer is required for formation of ICL-induced DSBs in normally cycling mammalian cells. More recent findings however, suggest MUS81 is necessary for DSBs formation only in situations when normal fork processing is impaired (Bell et al., 2011; Hanada et al., 2006; Sengerova et al., 2011). This observation is in line with experiments using the *Xenopus* model where no effect on replication fork was observed in MUS81-depleted extracts (Klein Douwel et al., 2014).

Given MUS81 favourable substrate structure is 3' flap containing a 5' end it seems possible that the complex only cuts a fraction of possible fork intermediates. It has been proposed that the main role of MUS81-EME1 in ICL repair is to cut on the 3' side of X-shaped structures in which a 5' end is adjacent to the ssDNA-dsDNA junction. Such structures are likely to form in a subset of fork traverse events.

SLX1-SLX4

SLX1-SLX4 is another structure-specific endonuclease that was identified to play a role in ICL repair. The SLX4 component of the complex was found to be another Fanconi anemia complementation group (FANCP) (Kim et al., 2011; Stoepker et al., 2011). Although the preferred substrate for SLX1-SLX4 is 5′ flap DNA, the complex exhibits affinity for a range of DNA structures such as splayed arms, 3′ flaps, and holiday junctions (Fricke and Brill, 2003; Wyatt et al., 2013). When considering the molecular context of ICL-stalled replication fork, 5'-like structure are generated when the leading strand advances up to -1 position. Similarly to MUS81-EME1 the endonuclease activity of SLX1 is important for Holliday junction resolution (Svendsen et al., 2009).

SLX4 is the master scaffolding factor for nucleases involved in various incision steps during ICL repair. It co-precipitates with XPF-ERCC1, SLX1, and MUS81-EME (Andersen et al., 2009; Fekairi et al., 2009; Muñoz et al., 2009; Svendsen et al., 2009)

Deletion of SLX4 results in a more severe hypersensitivity to crosslinking agents than deletion of SLX1 or MUS81 nucleases (Castor et al., 2013). Similarly, to XPF-ERCC1, SLX4 is also required for ICL-stalled fork processing in *Xenopus* egg extract (Klein Douwel et al., 2014). Interaction with SLX4 is critical for SLX1 and XPF-ERCC1 function in ICL repair (Castor et al., 2013; Crossan and Patel, 2012) but it is dispensable for MUS81 activity. Furthermore, SLX4 appears to dramatically

stimulate nucleolytic activity of XPF with ICL-containing structures (Hodskinson et al., 2014). Conclusively, XPF-ERCC1-SLX4-SLX1 is likely the main complex that cuts ICL-associated X-shaped DNA structures and the role of MUS81-EME1 is limited to scenarios described above.

FAN1

Establishing of FAN1 as a Fanconi anemia-associated nuclease in 2010 stirred a heated debate in the field of DNA repair (Liu et al., 2010; MacKay et al., 2010; Smogorzewska et al., 2010; Yoshikiyo et al., 2010). FAN1 amino acid sequence analysis identified a N-terminal UBZ4-type ubiquitin binding domain that was originally found to interact with monoubiquitylated FANCD2, a SAP DNA binding domain and a C-terminal PD-D/E(X)K nuclease motif, which classifies FAN1 in the same group as MUS81 or XPF (Kratz et al., 2010). Similar to SLX1, the preferred substrate for FAN1 are 5'flap structures but in contrast to SLX1 that cuts exactly at the junction, FAN1 cleaves 4 nucleotides 3' of the branch point. Cells depleted of FAN1 activity are hypersensitive to crosslinking agents but this phenotype is less severe than for *bona fide* Fanconi anemia members depletion (MacKay et al., 2010; Smogorzewska et al., 2010; Yoshikiyo et al., 2010). Interestingly mutations in FAN1 and FA genes are not epistatic, clearly suggesting that FAN1 has a distinct function in ICL repair (Yoshikiyo et al., 2010). Furthermore, in contrast to FA genes, FAN1 mutations in humans do not manifest in a typical FA phenotype but lead to karyomegalic interstitial nephritis, a kidney disease unseen in FA patients (Trujillo et al., 2012; Zhou et al., 2012).

In a very insightful study from the Rouse lab, FAN1 was shown to play an important role in stalled replication fork processing that is essential to maintain genomic stability. The authors showed that FAN1 recruitment mediated by monoubiquitinated FANCD2 is dispensable for ICL repair. Moreover, in the event of ICL-independent replication fork stalling, FAN1 activity is necessary to halt replication forks progression thus preventing chromosomal aberrations. Interestingly, the authors also showed that the genetic variant of FAN1 associated with elevated risk of pancreatic cancers abolishes recruitment by monoubiquitinated FANCD2 (Lachaud et al., 2016)

SNM1 family

Human homologs of the yeast Snm1/Pso2 protein include SNM1A, SNM1B and SNM1C with the latter playing a role in end joining of double strand breaks (Cattell et al., 2010). Cells lacking SNM1A and SNM1 B are hypersensitive to ICL inducing agents and double mutants exhibit even more severe sensitivity, which suggest the proteins may play some non-redundant roles in the ICL repair. There are contradictory data about the relationship of SNM1A/B protein with the FA pathway. Genetic experiments in chicken DT40 cells showed no epistasis with neither FA nor HR pathways. However, work conducted in mammalian cell lines suggests SNM1A collaborates with XPF-ERCC1 (Ishiai et al., 2004; Wang et al., 2011b).

SNM1A functions as a 5'→3' exonuclease with no obvious role in replication fork incision (Hazrati et al., 2008; Hejna et al., 2007). Using its 5'→3' exonuclease activity SNM1A is able to digest one DNA strand several nucleotides past the lesion. Thus, as long as an incision is introduced on the 5'side of the crosslink (possibly thanks to activity of SLX1 or FAN1 as described above), SNM1A might be able to complete the unhooking reaction without the need for 3' endonucleolytic cleavage performed by either XPF-ERCC1 or MUS81-EME1 (Wang et al., 2011b). This is an interesting scenario as it shows SLX1 could potentially substitute for the 3' incision when the 3' side of the lesion does not form a preferred incision substrate for either XPF-ERCC1 or MUS81-EME1.

Orchestrating ICL processing

Although DNA structures arising at ICL-stalled replication forks could be cut by nucleases in a number of combinations there are only two incision scenarios that lead to effective unhooking. The first one involves cutting the parental DNA strand on both sides of the ICL. Alternatively, incision introduced 5' of the lesion followed by exonuclease-driven degradation past the ICL would also result in effective unhooking. Considering the complexity of ICL nucleolytic processing, the mechanism on all of its stages is coordinated by a number of the FA factors, FANCD2-FANCI heterodimer seemingly being the main regulator. As monitored by chromatin immunoprecipitation in *Xenopus* egg extracts, monoubiquitylated FANCD2-FANCI is recruited to the lesion directly before nucleolytic incision takes place (Klein Douwel et al., 2014). Given the crystallographic data and functional studies the current model encompasses monoubiquitylated FANCD2-FANCI directly binding to the crosslinked

DNA where it exerts control over the unhooking process by regulating recruitment of other proteins to the site of damage. It is therefore plausible that FANCD2-FANCI performs its role by recruiting the XPF-ERCC1-SLX4-SLX1 complex that is able to cleave the parental strand on both sides of the ICL on one hand and physically blocking access to the lesion for other nucleases on the other.
In 2017 the alternative scenario was examined. The authors showed that while purified XPF-ERCC1 introduces incision 5' of ICL in the structures mimicking replication forks, presence of a nascent leading strand that models the molecular context of replication arrest was able to inhibit nucleolytic event. Interestingly addition of RPA to the assay was able to restore XPF-ERCC1 endonuclease activity on the replication arrest structure. The authors postulated a model in which SNM1A loads from XPF-ERCC1 induced incision and digest past the ICL leading to completeion of the unhooking event (Abdullah et al., 2017).

Translesion Polymerases (TLS)

After unhooking the nascent DNA strand needs to be extended past the lesion in order to generate a valid substrate for HR-mediated DSBs repair. As the replisome is still blocked by the ICL, normal replicative polymerases cannot catalyse this bypass and a switch to TLS enzymes is necessary. Synthesis of DNA past the damage site requires replacement of high-fidelity replicative polymerases with low-fidelity, low-speed TLS polymerases.

There are several human TLS polymerases, such as Pol ι, Pol η, Pol κ, REV1, and Pol ν, that are able to insert a nucleotide opposite, and bypass a range of structurally diverse ICLs. Chemistry of the polymerase reaction catalysed by these enzymes and in consequence the efficiency of ICL bypass depends on the actual structure of the ICL as well as the number of DSBs in close proximity of the lesion (Ho et al., 2011; Klug et al., 2012; Minko et al., 2008; Yamanaka et al., 2010). Indications for TLS involvement in ICL repair came from genetic studies in yeasts. Mutations in Rev1, Rev3 and Rev7 genes result in strong hypersensitivity to crosslinking agents (Larimer et al., 1989; McHugh et al., 2000; Sarkar et al., 2006a). Interestingly, PCNA monoubiquitination substantially increases its affinity for TLS polymerases facilitating polymerase switching to TLS at the site of damage (Stelter and Ulrich, 2003). This affinity increase is possible thanks to ubiquitin-binding domains of TLS polymerases

that bind directly to the ubiquitinated PCNA. Although this observation was first made in *S.cerevisiae* it was then shown to also be truth for polymerases found in other organisms (Bi et al., 2006; Plosky et al., 2006).

In mammalian cells Polζ consisting of the REV3, REV7 subunits together with REV1 play crucial role in ICL repair processes as shown by a pronounced hypersensitivity to crosslinking agents of the mutant cell lines (Gan et al., 2008; Nojima et al., 2005). REV1 has been shown to not only to play a scaffolding role for other TLSs that facilitates polymerase class switch when replication fork stalls but also has a deoxycytidyl transferase activity that may be involved in inserting a dCMP residue opposite an ICL (Minko et al., 2008; Sharma et al., 2013). According to a recent model, translesion synthesis is a multi-step, multi-protein process that begins with recruitment of one of the Y-family polymerases (Pol κ, Pol η, Pol τ, or Rev1) that is facilitated by the above mentioned monoubiquitination of PCNA. The main purpose of this step is to incorporate a nucleotide opposite of the damage site. Subsequently, replicative polymerases are replaced by extension TLS polymerases that adds a patch of around 18 nucleotides, which allows synthesis without detection by proofreading activity of replicative polymerases. It is believed that this step is facilitated by the Pol ζ heterodimer. Finally, the TLS polymerase is switched back to the standard replicative polymerase.

Role of Homologous Recombination in ICL repair

The importance of HR in ICL repair is highlighted by the severe sensitivity of HR-defective cells to ICL-inducing agents (Takata et al., 2001). For example, in the well-characterized Chinese hamster ovary (CHO) model system, Rad51 paralog-defective mutants (xrcc2, rad51d, and xrcc3) are up to 100-fold more sensitive to crosslinking agents when compared with wild-type cells (Cartwright et al., 1998; Liu et al., 1998).

The fact that both strands of the double helix participate in the ICL lesion implicates the need for DSBs formation and repair. There are several potential sources of DSB formation during ICL repair in the S phase of the cell cycle. In the first scenario nucleolytic processing of the lesion (unhooking) takes place before the replication fork encounters the ICL leading to fork collapse and generation of a one-ended DSB

(Räschle et al., 2008). In an alternative scenario, the replication fork regresses to form a “chicken foot” structure at the ICL thus generating a one-ended DSB. As mentioned above FANCM has been shown to promote fork regression *in vitro* (Gari et al., 2008c).

Eukaryotic cells utilize two mechanisms to deal with DSBs: the error-free, homologous recombination-based pathway and error-prone non-homologous-end-joining. A number of genetic studies revealed that whereas depletion of homologous recombination genes sensitizes cells to ICL-inducing agents, depletion of NHEJ-factor does not, indicating that homology-directed repair might be a preferred pathway in this molecular context (Bhattacharyya et al., 2000; Grossmann et al., 2001a; Pace et al., 2010). In Homology-directed repair identical (or almost identical) sequences are used as a template for repair of the DSB containing region. The template sequence preferentially comes from a sister chromatid available in late S and G2 phases of the cell cycle. Homologous recombination repair plays a crucial role in a number of processes such as replication fork restart, repair of two-ended, direct DSBs and antibodies maturation (Helleday et al., 2007). The significance of HR in ICL repair is underlined by clinical observations of ovarian and breast cancer patients. Women with mutations in *BRCA1* and/or *BRCA2* genes whose cancers are characterized by deficiency of HR repair benefit most from therapies based on ICL-inducing agents such as cis-platin (ovarian cancer) and carboplatin (breast cancer) (Alsop et al., 2012; Koster et al., 2007).

As described in the previous sections FANCM-FAAP24-MHF complex is recruited to ICL-stalled replication fork where it unwinds DNA and leads to ssDNA exposure. RPA protein is then recruited to coat the ssDNA fragments leading in a consequence to activation of ATR kinase signalling. Subsequently FA pathway is activated, and the homologous recombination machinery is loaded onto DNA. The DSB intermediate is recognized and resected by MRN complex (formed of MRE11, RAD50, NBS1) in a CtIP dependent manner. Resection generates a 3’ ssDNA overhang which is amenable to HR (Sartori et al., 2007) and is required for resistance to cross-link damage (Duquette et al., 2012). CtIP is ubiquitinated by BRCA1 and has been shown to accumulate at sites of locally induced ICLs

downstream of FANCM (Duquette et al., 2012). It then promotes accumulation of RPA, ATR, and FANCD2 at damage sites.

BRCA2 has been shown to load RAD51 proteins on stalled replication fork independently of activation of the Fanconi anemia pathway (Kitao et al., 2006; Long et al., 2011). Work in the *Xenopus* system clearly demonstrate that RAD51 loading onto stalled fork takes place ahead of nucleolytic incision event (Long et al., 2011). Upon incision, RAD51 filaments facilitate strand invasion of the damaged end onto the intact sister chromatid and promotes Homologous Recombination-dependent repair of the lesion. RAD51 and FANCD2 colocalize, however their recruitment to the ICL is not interdependent (Kitao et al., 2006). A number of additional factors regulate and tightly coordinate Homologous Recombination-dependent repair at its various steps of the process. One particularly interesting group of HR-mediating proteins are paralogs of RAD51. Five (RAD51B, RAD51 C, RAD51D, XRCC2, XRCC3) paralogs sharing extensive sequence similarity have been identified in mammalian cells (Lin et al., 2006). The proteins have been shown to form various subcomplexes the most important ones being BCDX2 (RAD51B, RAD51C, RAD51D, XRCC2) and CX3 (RAD51C, XRCC3) (Masson et al., 2001; Yonetani et al., 2005). Although crucial, the exact role of RAD51 paralogs in regulating Homologous Recombination has long been a subject of debate. RAD51C has been identified as a *bona-fide* Fanconi anemia factor – FANCO (Vaz et al., 2010). RAD51C functions downstream of FANCD2 monoubiquitination and formation of the DSB but it is essential for repair of ICL-induced DSBs in HR-dependant manner during the S-phase of the cell cycle (Somyajit et al., 2012). Although the RAD51 paralogs have been shown to play distinct roles on various stages of the HR their most important mechanism of action is to stimulate HR by Rad51 filament remodelling and thus facilitating strand exchange with the undamaged template (Taylor et al., 2015).

Another group of HR mediators that has recently been in focus are MCM8 and MCM9, which are the last discovered members of the MCM family of proteins (Lutzmann et al., 2005). MCM8 and MCM9 form a complex that is required for resistance to ICL-inducing agents and has been shown to promote RAD51 recruitment to facilitate HR-dependant DSB repair (Lutzmann et al., 2012; Nishimura et al., 2012; Park et al., 2013). In a more recent report MCM8/9 have been shown to be required for DNA

resection by MRN complex to generate ssDNA ends (Lee et al., 2015). Interestingly mutations in MCM8 and MCM9 have been associated with primary ovarian insufficiency and ovarian carcinogenesis (Fauchereau et al., 2016).

A number of key homologous recombination players have been identified as FA factors including, FANCD1/BRCA2, FANCN/PALB2, FANCJ/BRIP1/BACH1, FANCO/RAD51C. Interestingly activation of Fanconi anemia network does not appear to be required for the loading of HRR machinery upon ICL-inducing agents treatment based on ability to Rad51 foci formation. RAD51 focus induction is not affected in cells from any Fanconi anemia complementation groups other than BRCA2 and PALB2 (Godthelp et al., 2006). Studies in *Xenopus laevis* cell-free system revealed that RAD51 filaments are formed with single-stranded regions of the stalled fork even before nucleolytic incisions occur (Long et al., 2011). Upon incision RAD51 nucleation facilitates strand invasion of the broken end into the intact sister chromatid for HR-mediated repair of the DSB. As mentioned before upon cross-link induction RAD51 colocalizes with FANCD2 after cross-link induction however, recruitment of either protein to the broken DNA is not dependent on the presence of the other (Kitao et al., 2006).

FANC gene	Alternative name	Description or molecular function of protein
Bona fide Fanconi anaemia genes		
FANCA		Fanconi anaemia core complex
FANCB		Fanconi anaemia core complex
FANCC		Fanconi anaemia core complex
*FANCD1**	*BRCA2*	Homologous recombination and fork stabilization
FANCD2		Binds to FANCI; has multiple functions
FANCE		Fanconi anaemia core complex
FANCF		Fanconi anaemia core complex
FANCG	*XRCC9*	Fanconi anaemia core complex
FANCI		Binds FANCD2; has multiple functions
FANCJ	*BRIP1*	Homologous recombination and translesion synthesis
FANCL		Fanconi anaemia core complex; E3 ubiquitin ligase for FANCD2-I ubiqitylation
*FANCN**	*PALB2*	Homologous recombination
FANCP	*SLX4*	Nuclease scaffold
FANCQ	*ERCC4, XPF*	DNA incision, functions in nucleotide excision repair
FANCT	*UBE2T*	Fanconi anaemia core complex; E2 ubiquitin ligase for FANCD2-I ubiquitylation
Fanconi anaemia-like genes		
FANCM		Fanconi anaemia core complex; lesion recognition
*FANCO**	*RAD51C*	Homologous recombination
FANCR	*RAD51*	Homologous recombination and fork stabilization
*FANCS**	*BRCA1*	Homologous recombination and fork stabilization
FAAP100		Fanconi anaemia core complex
FAAP24		Fanconi anaemia core complex
FAAP20		Fanconi anaemia core complex
FAAP16	*MHF1*	Fanconi anaemia core complex; histone fold-containing protein
FAAP10	*MHF2*	Fanconi anaemia core complex; histone fold-containing protein
BOD1L		RAD51 nucleofilament stabilization and fork stabilization
UHRF1		Lesion recognition
USP1		FANCD2-I deubiquitylation
UAF1		FANCD2-I deubiquitylation
FAN1		Nuclease

Table 1-1 Summary of FA and FA-like genes.

1.2.3.5 One or two fork converge?

Replication-dependent ICL repair was successfully recapitulated in *Xenopus* egg cell-free extracts, allowing detailed studies of repair intermediates. In the system developed by the Walter lab a 6kb plasmid harbouring a single, site-specific ICL is incubated in the egg extract where the lesion is recognized and repaired in a replication dependent manner (Räschle et al., 2008). Bidirectional replication of the plasmid implies that two replication forks arriving from both sides converge on the ICL triggering the repair processes. Repair in this system can be analysed as progression of the replication forks measured by incorporation of radiolabelled-nucleotides into the newly synthetized strands. Although some aspects of the dual fork collapse mechanism likely apply to the single fork model, it is currently debated which model best recapitulates the physiological situation in cells.

The dual fork convergence inevitably observed in the case of a plasmid is a much less likely event for genomic DNA replication where origins are spaced 100 kb apart and forks move at a rate of 1.5 kb/minute (Duderstadt et al., 2014). Unless the lesion is located halfway between two contemporaneously firing origins, one fork will always arrive to ICL earlier than the other. On the other hand, before stalled forks can be processed, MCM2-7 helicases need to dissociate, the leading strand has to approach the ICL as well as a number of structure specific nucleases need to be recruited to the damage site to promote opening of a 30-40 minutes temporal window for the other fork to arrive (Fu et al., 2011; Klein Douwel et al., 2014).

The major criticism of the single fork model is built on fact that in metazoans, in contrast to bacteria and yeast there are no known mechanisms for replicative helicases re-loading in the S-phase. This practically means that the restored replication fork would need to await another fork arrival before resuming synthesis therefore posing a risk of genomic rearrangements (Petermann et al., 2010).
Recently an interesting data from the Seidman group shed new light on the single vs dual fork model. Using DNA combing and a fluorescently tagged ICL, the group showed that both single and dual fork collisions constitute around 15–20% of ICL encounters. Strikingly, in circa 60% of observed cases, replication forks bypassed ICLs in a FANCM-dependent manner without unhooking (Huang and Li, 2013).

Taken together, it seems that both single and dual fork convergence events can be observed *in vivo* and pathways to resolve both situation exist in cells. An interesting new phenomenon of replication fork traverse has been observed but more research is required to fully understand it.

1.3 The role of HELQ helicase in ICL repair

Mammalian HELQ also known as Hel308a or Hjm (in *archaea*), Mus301 or Spn-C (in *D. melanogaster*), and HEL-308/HELQ-1 (in *C. elegans*) is a 3'-5' superfamily 2 helicase with ssDNA dependent ATPase activity (Woodman and Bolt, 2009). In *in vitro* studies, HELQ has been shown to unwind a range of diverse structures with 3' ssDNA overhangs and to remove the lagging strand from synthetic replication fork

substrates (Fujikane, Komori, Shinagawa, & Ishino, 2005; Guy & Bolt, 2005a; Marini & Wood, 2002). Crystal structure analysis of archaeal Hel308 combined with biochemical studies indicate that the DNA strand separation activity occurs via a ratchet-like manner (Buttner et al., 2007). From a mechanistic point of view, DNA unwinding is coupled with ATP hydrolysis thanks to a C-terminal ssDNA binding domain that has also been shown to auto-inhibit the helicase activity (Richards et al., 2008a; Woodman and Bolt, 2011).

Analysis of *HelQ*-deficient mutants in metazoans provided evidence that it plays an important role as DNA repair factor: in *D. melanogaster Mus301* is required for resistance to nitrogen mustard – a potent ICL inducing agent as well as methyl methanesulfonate, a compound known for its alkylating activity, which efficiently blocks replication forks (Boyd et al., 1981). Mus301 (Spn-C) was also shown to be required for repair of meiotic DSBs and proper segregation of meiotic chromosomes in flies (Laurencon et al., 2004; McCaffrey et al., 2006). *C. elegans helq-1* mutants are comparably sensitive to a variety of ICL-inducing agents but do not show any problems with meiotic DSB repair (Muzzini et al., 2008). Furthermore, studies in mammalian models confirmed that HELQ is required in the response to mutagens such as mitomycin C (MMC) and camptothecin (CPT) that block or impede the replication fork progression (Moldovan et al., 2010; Tafel et al., 2011). The actual role of HELQ in the ICL repair processes was unknown.

Taken together, existing data clearly positioned HELQ as DNA damage resistance factor with important role in the repair of ICLs. I decided to undertake research aiming to better understand how HELQ contributes to maintaining genome stability. In particular I was interested to characterise the stage of ICL repair process in which ICL performs its function as well as to disseminate the molecular mechanism in which HELQ counteracts genotoxic stress. Obtaining such insights would contribute to better understanding of the ICL detection and repair – processes of clinical significance. As mentioned before, formation of ICL is a basis for therapeutic use of many anti-cancer agents. Improving our understanding of the mechanisms in which cells repair ICLs together with knowledge of intermediates of the process and molecular players involved can potentially help to address important clinical problems such as development of chemotherapy resistance as well as to identify new therapeutic agents.

To gain insight into the role of HELQ in the DNA damage response and to extend the scope of our understanding of how cells counteract genotoxic stress arising during DNA replication, we have generated a mouse model of HELQ deficiency (C. Adelman) and have performed proteomic and molecular analyses of HELQ. Interestingly, we found that *HelQ* mutant mice are highly reminiscent of the Fanconi Anemia mouse models manifesting with subfertility, germ cell attrition, predisposition to tumors and hypersensitivity to DNA ICL-inducing agents. Furthermore, *HelQ* mutant females are prone to ovarian tumors.

We have created a number of HELQ-expressing mammalian cell lines to perform proteomic studies that revealed HELQ it is physically and functionally involved in the replication stress response network as defined by interactions with the ssDNA binding protein RPA70, the checkpoint kinase ATR, the Fanconi Anemia factors FANCD2 and FANCI, and the RAD51 paralog complex: RAD51B, RAD51C, RAD51D and XRCC2. Interaction with these proteins supports the *in vivo* observations of HELQ deficiency; we establish that HELQ cooperates with the FA machinery and downstream RAD51 paralog BCDX2 HR factors to promote repair. Furthermore, the ovarian cancer phenotype of *HelQ*-deficient female mice supports recent findings demonstrating a link between germline RAD51 paralog mutations and increased ovarian cancer risk (Loveday et al., 2011; Meindl et al., 2010; Pelttari et al., 2011; Vuorela et al., 2011; Wickramanyake et al., 2012)

These results clearly position HELQ as an integral component of the cellular response to DNA damage encountered during replication and its interaction with the RAD51 paralog BCDX2 complex underscores its importance to HR and tumor suppressor.

Chapter 2. Materials & Methods

2.1 Enzymes and reagents

2.1.1 Enzymes

All the restriction enzymes used in this study were purchased from New England Biolabs Inc. (NEB). Proteinase K, recombinant PCR grade, was purchased from Roche. RNase A was purchased from Qiagen.

2.1.2 General reagents

All chemical reagents were purchased from Sigma unless stated otherwise.
Amersham: Hyperfilm™ ECL High Performance Chemiluminescence film
Bio-Rad: 30% Acrylamide/Bis Solution, 37.5:1; 40% Acrylamide/Bis Solution 19:1; bromophenol blue and xylene cyanol; All Blue Precision Plus Protein™ Standards.
Chromotek: GFP-TRAP®_A beads
EMD Millipore: KOD Hot Start DNA polymerase
Expedeon: InstantBlue™
Fermentas: GeneRuler 1kb DNA ladder
GIBCO: Tetracycline-free Fetal Bovine Serum (FBS)
Life technologies™: SYPRO® Ruby Protein gel stain; NuPAGE® 10% Bis-Tris gels and NuPAGE® 7% Tris-Acetate gels; NuPAGE® MOPS and Tris-Acetate SDS running buffer (20x); UltraPure™ agarose; NuPAGE® LDS sample buffer; ProLong® Gold Antifade Reagent with DAPI; Lipofectamine RNAiMAX; zeocin; geneticin; SOC medium; Gateway® LR and BP reactions
Millipore: Immobilon®-P transfer membrane
NBS Biologicals: SafeView
Promega: Fugene HD transfection reagent
Qiagen: Effectene® transfection reagent
Sigma: Anti-FLAG M2 Affinity Gel resin; Hoechst 33342; 5-bromo-2'-deoxyuridine (BrdU); phosphatase inhibitor cocktail 2 and 3; propidium iodide (PI); formamide, >99.5%; Fetal Bovine Serum (FBS); pepsin from porcine gastric mucosa; dimethyl sulfoxyde (DMSO); thymidine.

2.1.3 Antibodies

2.1.3.1 Primary antibodies

Antibodies used for western blot analysis: Flag (Sigma F1804, 1:2,000), HELQ (Santa Cruz 81095, 1:200), His (Clontech 631212, 1:2,000), PARP1 (Trevigen 4338-ML-50, 1:1,000), CHK1 (Sigma C9358, 1:500), S345-P-CHK1 (Cell Signalling 2348, 1:500), CHK2/p-CHK2 (Upstate 05-649, 1:400), ATM (Sigma A1106, 1:2,000), S1981-p-ATM (Cell signalling 4526, 1:1,000), histone H3 (Abcam 10799, 1:2,000), α-tubulin (Sigma T6199, 1:2,000), RAD51 (Santa Cruz 8349, 1:200), FANCD2 (Epitomics 2986-1, 1:1,000), γH2AX (Cell Signalling 2577, 1:1,000), RPA32 (Abcam 12F3.3, 1:1,000), BRCA2 (Santa Cruz 8326, 1:200), TFIIH p89 (Santa Cruz 293, 1:200). All RAD51 paralogue antibodies were a kind gift from S. West's laboratory, RAD51B (IH3 mouse monoclonal antibody, 1:500), RAD51C (2H11 mouse monoclonal antibody, 1:500), RAD51D (5B3 mouse monoclonal antibody, 1:400), XRCC2 (7B7 mouse monoclonal antibody, 1:400), XRCC3 (10F1 mouse monoclonal antibody, 1:400).

Antibodies used for immunoprecipitation: RAD51C (R68 rabbit antibody), XRCC3 (10F1 mouse monoclonal antibody), mouse IgG (Abcam 18413) and rabbit IgG (Abcam 46540) were used where appropriate as negative control immunoprecipitates.

2.1.3.2 Secondary antibodies

Polyclonal Mouse and rabbit horseradish peroxidase-conjugated secondary antibodies were from ThermoFisher, and signals were visualized with ECL western blotting detection reagent (Amersham) or SuperSignal West Femto reagent (Thermoscientific).

2.2 Buffers and solutions

2.2.1 Media and protein buffers

Western blot blocking buffer: 3-5% (w/v) skimmed milk powder in PBS + 0.1% (v/v) Tween-20
PBS: 140 mM NaCl, 3.4 mM KCl, 10 mM Na_2HPO_4, 18 mM KH_2PO_4
Tris-buffered saline (TBS): 30 mM Tris-HCl (pH 7.4), 150 mM NaCl
Standard Lysis buffer: TBS supplemented with 1x phosphatase inhibitor cocktail 2, 1x phosphatase inhibitor cocktail 3, 1x protease inhibitor cocktail, 0.5% (v/v) NP-40
Benzonase lysis buffer: 20 mM Tris-Cl, pH 7.5, 75 mM NaCl, 10% glycerol, 2 mM $MgCl_2$, 0.5% NP40, 30 U/ml benzonase, protease inhibitors
CSK buffer 10 mM PIPES, pH 6.8, 100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 1 mM EGTA, pH 7, 0.5% Triton X-100
FLAG Beads wash buffer: 20 mM Tris-Cl, pH 7.5, 150 mM NaCl, 3 mM EDTA, 0.5% NP40
FLAG elution buffer: TBS supplemented with 200 µg/ml 3xFLAG peptide (SIGMA)
Western blot transfer buffer: 25 mM Tris base, 190 mM glycine, 20% (v/v) methanol
Protein loading buffer (2x): NuPAGE® LDS loading buffer supplemented with 10% (v/v) β-mercaptoethanol
Storage buffer: 50 mM Tris-HCl (pH 8.0), 10% (v/v) glycerol, 100 mM NaCl, 1 mM DTT
Protein dilution buffer: 50 mM Tris-HCl (pH 8.0), 10% (v/v) glycerol, 100 mM NaCl, 1 mM DTT, 0.1 mg/ml purified BSA

2.2.2 DNA buffers

TBE: 89 mM Tris base, 89 mM boric acid, 2 mM EDTA
TE: 10 mM Tris-HCl (pH 8.0), 1 mM EDTA
DNA loading buffer (6x): 0.35% (w/v) bromophenol blue, 0.35% (w/v) xylene cyanol, 30% (v/v) glycerol

2.3 Bacterial strains

2.3.1 Bacterial strains

E. coli DH5α competent cells were used for all plasmid transformations.
E. coli One Shot® TOP10 chemically competent cells (Life Technologies™) were used for the transformation of DNA ligation reactions.
E.coli DH10BAC™ cells were used to generate bacmids for *Xenopus laevis* HelQ expression.

2.3.2 Transformation of chemically competent cells

E. coli cells were grown to an OD~0.5, centrifuged, resuspended in 500 µl of ice-cold TSB buffer and incubated on ice for 10 min or aliquoted at -80°C.
For transformation with 1 µg of plasmid DNA, 50 µl of competent cells were incubated with DNA, 1x KCM and water (up to 50 µl) on ice for 20 min and then at RT for 20 min. Cells were incubated with 500 µl of SOC media for 40-60 min in a shaking 37°C incubator and plated on LB agar plates.

2.4 Cell lines and cell culture methods

Table 2-1 List of cell lines used in this study

CELL LINE	DESCRIPTION
NIH-3T3	*Mus musculus* embryonic fibroblast cell line
U2OS	*Homo sapiens* osteosarcoma cell line
HEK293	*Homo sapiens* embryonic kidney cell line
SF9	*Spodoptera frugiperda* – insect cell line

All cells were cultured in Dulbecco's Modified Eagle Media (DMEM) supplemented with 10% (v/v) FBS. The cultures were incubated in a 10% CO_2 humidified incubator at a temperature of 37°C.

2.4.1 siRNA transfection

siRNA oligonucleotides:

Targeted gene:	Oligonucleotide sequence 5'→3'
RAD51	-AAGGGAAUUAGUGAAGCCAAA-
BRCA2	-AACAACAAUUACGAACCAAAC-
HELQ_1	-GUUUGAAGAUUGCAACGAA-
HELQ_3	-AAUGUGAGGUGAUUAAGAA-
HELQ_4	-GGUAGAAGAGUUACUAAGA-
HELQ_17	-GUUUGAAGAUUGCAACGAA-
XRCC2	-CAGGGTACTACGCAAGCCT-
XRCC3	-CAGAATTATTGCTGCAATT-
RAD51C	-AAGAGAATGTCTCACAAAT-
RAD51D	-CTGGGTGGAAATAAGCTTA-

The short-interfering RNAs (siRNAs) used in this study were purchased from Thermofisher (siGENOME). Control Luciferase siRNA and control scramble siRNA were purchased from Thermofisher.

U2OS cells were subjected to two or three rounds of reverse transfections depending on the targeted gene (as indicated in respective experiments) using siGENOME siRNA and Dharmafect1 (Thermofisher) according to the manufacturer's protocol. For U2OS cells 3 x 10^5 or 6 x 10^5 cells were seeded in 6 cm or 10 cm cell culture plates respectively in each round of reverse transfections.

Mass spectrometry and proteomics

HELQ–Flag and Flag control cells (~ 100 mln of each) were collected and lysed in benzonase lysis buffer (20 mM Tris-Cl, pH 7.5, 75 mM NaCl, 10% glycerol, 2 mM $MgCl_2$, 0.5% NP40, 30 U ml−1 benzonase, protease inhibitors). NaCl concentration was adjusted to 150 mM, EDTA to 3 mM and lysates were cleared by centrifugation. Supernatants were pre-cleared with Protein G agarose beads for 30 min at 4 °C. Pre-cleared lysates were incubated with anti-Flag affinity agarose resin (Sigma) for 4 h at 4 °C. Beads were washed five times with wash buffer (20 mM Tris-Cl, pH 7.5, 150 mM NaCl, 3 mM EDTA, 0.5% NP40) and once with PBS. Bound proteins were

eluted by boiling in SDS–PAGE sample buffer and eluates were resolved on NuPAGE Bis-Tris gels (Invitrogen) and stained with Sypro Ruby (Invitrogen). Gel slices were excised and processed for mass spectrometry using the Janus automated liquid handling system (PerkinElmer). Peptides were analysed by nanoscale capillary liquid chromatography–electrospray ionization/multi-stage mass spectrometry (LC–ESI MS/MS), data were processed using Mascot Distiller (Matrix Science) and exported to Scaffold for viewing (Proteome Software).

The Biological General Repository for Interaction Data sets (BioGRID, http://thebiogrid.org/), the Molecular INTeraction database (MINT, http://mint.bio.uniroma2.it/mint), and Search Tool for the Retrieval of Interacting Genes/Proteins database (STRING, http://string-db.org/) were used to compile the protein interaction network.

Cell lysates, in vitro binding assay and fractionation for western blot analyses
All cell lines used in this study were short tandem repeat-profiled and tested for mycoplasma infection before use. All lysis buffers were supplemented with protease inhibitor cocktail (Roche) and phosphatase inhibitors (Sigma).

For validation of mass spectrometry data, HELQ–Flag- and Flag-expressing cells were used. (This was due to our inability to validate these interactions using endogenous HELQ, stemming from the fact that it is expressed at very low levels in most human cell lines, and no antibodies were found to reliably immunoprecipitate the human version. Validation using endogenous mouse HELQ was similarly hindered by a lack of reagents available for detection of the mouse RAD51 paralogues.) Cells were lysed in the presence of benzonase and 2 mg of total protein were immunoprecipitated with anti-Flag affinity resin as above. Beads were washed, bound proteins eluted with 1 × NuPAGE LDS sample buffer and analysed by western blot. Similar methods were employed using lysates prepared from 293T cells to examine endogenous HELQ or RAD51C coimmunoprecipitates.

For in vitro binding assays, HELQ–Flag, ALC1–Flag and Flag cells were lysed in the presence of benzonase and pre-cleared lysate was used for Flag immunoprecipitate as described above. Flag-immunocomplexed beads were then washed eight times

with a modified wash buffer containing from 250mM to 1 M NaCl to remove bound co-precipitates, and then with dropping NaCl concentration from 1 M to 250mM and finally once with in vitro binding buffer (20 mM Tris-Cl, pH 7.5, 280 mM NaCl, 3 mM EDTA, 0.5% NP40). Washed beads were incubated with recombinant RAD51 paralogue BCDX2 complex (gift of S. West's laboratory) in binding buffer for 4 h at 4 °C and washed four times with the same buffer. Eluates were analysed by western blot.

GFP-tagged HELQ was stably expressed in NIH3T3 cells using a BAC recombineering method to C-terminally Flag/GFP-tag the BAC-containing full-length genomic Helq, which included the endogenous promoter. This allowed HELQ–GFP to be expressed at physiological levels.

For chromatin fractionation of embryonic fibroblasts and siRNA-treated U2OS cells, cells were treated with or without 3 μM aphidicolin for 6 h or with or without 1 mM MMC for 24 h, collected and fractionated: pellets were re-suspended in CSK buffer (10 mM PIPES, pH 6.8, 100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 1 mM EGTA, pH 7, 0.5% Triton X-100), incubated for 10 min on ice (a small fraction of this was removed and SDS–PAGE sample buffer was added to obtain WCEs), pelleted at low speed and supernatants reserved as soluble fraction. Pellets were washed in CSK buffer, and re-pelleted. Pellets were re-suspended in benzonase CSK buffer (10 mM PIPES, pH 7.5, 100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 0.5% Triton X-100, 0.1 U μl−1 benzonase), lysates were incubated for 30 min at 37 °C, pelleted and supernatants reserved for chromatin fraction. Pellets were re-suspended in high-salt CSK (recipe as above except NaCl was added to 500 mM), lysates were incubated for 10 min on ice, cleared at high speed and supernatants pooled with benzonase CSK lysates to yield chromatin extracts25. In total, 25 μg of soluble and 10 μg of chromatin proteins were analysed by western blotting.

Induction of DNA-damage and ATR inhibition

Thirty-six hours after the last reversed transfection, cells were treated for 14 h with 2 μM aphidicolin. For ATR inhibition experiment, 3 μM ATR inhibitor (gift from Oscar Fernandez Capetillo) was added to cultures 30 min before aphidicolin treatment.

MMC induced RAD51 chromatin recruitment in HELQ depleted cells

U2OS cells were subjected to three rounds of reversed siRNA transfection and treated with 1μM MMC for 24 hours. Subsequently the cells were harvested, chromatin fraction was isolated and probed for RAD51, histone H3 and Tubulin.

Checkpoint activation status in HELQ-deficient cells

Mouse HelQ$^{\Delta C/\Delta C}$ cells

5 mln mouse HelQ$^{\Delta C/\Delta C}$ and HelQ wt (HelQ$^{+/+}$) cells were treated with or without 500ng/ml MMC or 50nM CPT, lysed with lysis buffer (20 mM Tris-Cl, pH 7.5, 150 mM NaCl, 10% glycerol, 2 mM $MgCl_2$, 0.5% NP40, 30 U/ml benzonase, protease inhibitors). Lysates were centrifugated and supernatants boiled with Sample Buffer and probed for appropriate checkpoint indices by wb.

U2OS si-RNA depleted cells

U2OS cells were subjected to three rounds of reverse HELQ siRNA transfection with two independent sequences (HELQ_1 and HELQ_3) according to the manufacturers (Thermofisher Dharmafect) protocol. HELQ depleted U2OS cells were treated with MMC (1μM for 24h), lysed as above and probed for checkpoint activation indices.

RAD51 and γH2AX chromatin dynamics in the absence of HELQ

Mouse HelQ$^{\Delta C/\Delta C}$ and wt (HelQ$^{+/+}$) cells were grown on 6-well plates and exposed to +/- 100ng/ml MMC for 24 hours. After washing off the drug with PBS (2 separate washes), fresh media were added, and cells were grown for the next 4 days. Every 24 hours cells samples were derived, counted and pellets frozen down. Once all samples have been collected, counted and frozen down, insoluble (chromatin) fraction was isolated according to the above described protocol and probed for RAD51, H3 and Tubulin by western blot.

U2OS cells were subjected to three rounds of reversed transfection, grown on 6-wells plate and exposed to 1μM MMC for 24 hours in order to induce DNA crosslinks. Subsequently the MMC was washed twice with PBS and fresh media was added to

allow cells to repair broken DNA for 4 consecutive days. Cell samples were taken every 24 hours during the experiment, counted and frozen down. After all cells samples have been taken and frozen down, chromatin fraction was isolated according to the above described protocol and probed by western blot for RAD51, H3 and Tubulin.

Homologous recombination reporter assays

DSB repair efficiency by homologous recombination was measured in DR–GFP U2OS cells. In brief, 48 h after the first round of siRNA transfection (40 nM) using Lipofectamine RNAiMAX (Invitrogen), cells were either mock-transfected (pcDNA3.1) or transfected with 0.6 μg of an I-SceI expression plasmid (pCBASce) together with siRNA (20 nM) using 3.6 μl of Lipofectamine 2000 (Invitrogen). The media was replaced 3 h after I-SceI transfection and cells were analysed for GFP expression by flow cytometry on a Cyan ADP (Dako) 72 h after I-SceI transfection. To confirm siRNA efficiency, western blotting was carried out on 50 μg of NP-40 lysates plus sonication run on 4–15% Precast SDS–PAGE gels (Bio-Rad).

2.4.2 Cell line derivation (work done with or by C. Adelman)

Ear fibroblasts for primary and SV40 immortalized cultures were generated as follows: mice were euthanized, and ear tissue was collected using sterile scissors, ear fragments were rinsed twice in 70% ethanol followed by two rinses in PBS supplemented with 100 μg ml−1 kanamycin. Tissue was transferred into 0.3 ml of protease solution (4 mg ml−1 each of collagenase D and dispersed in DMEM; filter sterilized), and incubated at 37 °C for 45 min. In total, 1.5 ml DMEM containing 10% FBS, 1× glutamine and 5× antibiotic–antimycotic solution were added to protease solution containing ear fragments, and samples were incubated at 37 °C overnight. Cells were dissociated by pipetting, passed through a 40-μm mesh cell strainer, and plated in DMEM as above except using 1× antibiotic–antimycotic solution. Cells were passaged upon reaching confluence to five dishes, and upon reaching confluence, cells were frozen at passage 1 or used immediately for immortalization or experiments.

Fibroblasts were immortalized via transfection with a vector expressing SV40 large T antigen. Constitutively expressed HELQ–Flag, Flag, and ALC1–Flag cell lines were generated using the 293 Flp-In system according to the manufacturer's protocol (Invitrogen). NIH3T3 cells stably expressing GFP-tagged mouse HELQ (consisting of a bacterial artificial chromosome (BAC) containing the entire Helq promoter and genomic locus) were generated according to the BAC recombineering method described previously9. These and all other cell lines used in this study (293T, HeLa, NIH3T3 and U2OS) were grown in DMEM supplemented with 10% FBS and L-glutamine. Cells were grown in 5% CO2 incubators at atmospheric O2 concentrations (~21%) with noted exceptions where samples were cultured at physiological O2 concentrations (~5%).

For HELQ–GFP transient transfections in HeLa cells, human HELQ was cloned into the pcDNA6.2/C-EmGFP-DEST vector using Gateway technology (Invitrogen). The vector was transfected into HeLa cells with Lipofectamine 2000 using the manufacturer's protocol (Invitrogen). Live or paraformaldehyde (PFA)-fixed cells (fixed cells were counterstained with DAPI (4′ ,6-diamidino-2-phenylindole)) were visualized 48–96 h after transfection under epifluorescence using a Zeiss Axio Imager M1 microscope with an ORCA-ER camera (Hamamatsu), and images were acquired using the Volocity software (Improvision, Perkin Elmer).

2.4.3 Clonogenic survival assay

For all experiments, fibroblast lines established from littermates or siblings were used wherever possible. Experiments involving primary cells were conducted in physiological O2 using cell lines of similar passage number.

For clonogenic survival assays, SV40 immortalized mouse ear fibroblasts and siRNA-treated U2OS cells were plated in triplicate on 10-cm dishes at clonal density, allowed to adhere for 8–16 h, and damage treatments administered (CPT medium was changed after 24 h). For Olaparib treatment cells were exposed to 0μM, 0,5 μM and 1 μM concentrations of Olaparib in DMSO for 24h. Media were then changed, and cells were grown for 12 days. For other damaging agents after 8–10 days of growth, plates were rinsed, fixed/stained in 20% ethanol/4% crystal violet (w/v),

rinsed in distilled water and colonies tabulated. All results were normalized to untreated to adjust for plating efficiency and determine percentage survival. Survival experiments were carried out on at least two independent sets of mutant and control mouse cell lines, and in most cases cell lines were tested in at least two independent experiments. Similar results were obtained across all experiments and sets of cell lines.

Immunofluorescence – RAD51 foci dynamics in absence of HELQ

Cells were first washed in PBS and then fixed with 2% PFA at room temperature (18 °C) for 15 min, and then washed three times in PBS. The fixed cells were further permeabilized with 3% BSA in PBS plus 0.1% Triton X-100 for 30 min at room temperature. Primary antibodies (RAD51; 1:500, RPA; 1:1,000, γH2AX) were added and incubated at 37 °C for 1 h. After washing with PBS plus 0.1% Triton X-100, secondary antibodies (provided by Jackson ImmunoResearch) were applied and incubated for 1 h in the dark. The stained coverslips were mounted with prolong Gold Antifade reagent (Invitrogen). Imaging was carried out using Axio Imager (Zeiss) or Axioplan 2 Imaging (Zeiss) microscope and analysed by Axiovision software (Zeiss).

HELQ and BCDX2 stability in U2OS cells

U2OS cells were grown on 6-wells plates and were subjected to three rounds of siRNA transfection targeting indicated proteins (HELQ and RAD51 paralogs). 48h post final transfection equal numbers of cells were harvested, counted and lysed using the lysis buffer (20 mM Tris-Cl, pH 7.5, 150 mM NaCl, 10% glycerol, 2 mM $MgCl_2$, 0.5% NP40, 30 U ml−1 benzonase, protease inhibitors). Whole cell extracts protein concentrations were adjusted samples were probed by western blot to determine relevant protein levels.

2.5 Gel electrophoresis

2.5.1 SDS-polyacrylamide gel electrophoresis (PAGE)

SDS-PAGE was performed as described (Laemmli, 1970). Briefly, protein samples were prepared by adding an equal volume of protein loading buffer (2x) and boiling for 5 min prior to electrophoresis. Gel electrophoresis was performed with NuPAGE®

10% Bis-Tris gels or NuPAGE® 7% Tris-Acetate gels in an XCell SureLock™ Mini-Cell Electrophoresis System (Life Technologies™) using NuPAGE® MOPS or Tris-Acetate SDS running buffer (20x). Gels were run at RT for 90 min at 180 V or 150 V. Proteins were visualised by western blotting or by staining with InstantBlue or SYPRO® Ruby.

2.5.2 Agarose gel electrophoresis

Gels were prepared in Sub-Cell® GT Agarose Gel Electrophoresis System (Bio-Rad) and contained 1% (w/v) Agarose in TBE buffer supplemented with a 1:10 dilution of SafeView Samples were loaded in 1x DNA loading buffer and run at 120 V for 20-40 minutes at RT. Gels were imaged and photographed using the Molecular Imager® GelDoc™ XR+ (Bio-Rad).

2.5.3 InstantBlue staining

After SDS-PAGE, gels were incubated with InstantBlue® for up to 12 hr and destained in water.

2.5.4 SYPRO® Ruby staining

After SDS-PAGE, protein gels were incubated in a solution containing 50% (v/v) methanol and 7% (v/v) acetic acid for 15 min at RT. SYPRO® Ruby Protein gel stain was added to cover the gels, microwaved for 30 s, agitated for 30 s at RT and then microwaved for an additional 30 s. The gel was covered with aluminium foil and incubated in the SYPRO® Ruby Protein gel stain for 30 min at RT before washing with a solution containing 10% (v/v) methanol and 7% (v/v) acetic acid (3x for 15 min each). Proteins were visualised using the Molecular Imager® GelDoc™ XR+ (Bio-Rad).

2.6 General methods of DNA and protein manipulation

2.6.1 DNA concentration determination

DNA concentrations were determined by measuring the sample absorbance at 260 nm using a NanoDrop® ND-1000 Spectrophotometer.

2.6.2 PCR for cloning

PCR reactions for cloning purposes were carried out with KOD DNA polymerase (Novagen). PCR reaction were set up as follows (unless stated otherwise):

29 μl dH_2O

5 μl 10X buffer

5 μl dNTPs (stock, 2 mM each)

2 μl $MgSO_4$ (stock, 25 mM)

1 μl template DNA (100-300 ng)

1 μl 5' primer (stock , 5 μM)

1 μl 3' primer (stock, 5 μM)

1 μl DMSO

1 μl KOD hot start polymerase

2.6.3 Gateway recombinational cloning

Cloning using the Gateway technology (Invitrogen) was performed according to the manufacturer's instructions.

2.6.4 Ethanol precipitation

DNA samples were mixed with 0.1 vol of 3 M sodium acetate (pH 5.2) and two volumes of 100% ice-cold ethanol. After 20 min incubation in dry ice, samples were centrifuged at 14000 rpm for 30 min. The precipitated DNA was washed in 70% (v/v) ethanol, air-dried and resuspended in TE buffer

2.6.5 Protein concentration determination

Protein concentrations were quantified using Bio-Rad D_c protein assay kit. The absorbance at 750 nm was determined using μQuant (BIO-TEK instruments Inc.) and compared to a standard curve obtained with known concentrations of BSA.

2.6.6 Western blotting

After SDS-PAGE electrophoresis, gels were transferred onto Immobilon®-P transfer membrane (pre-activated in 100% methanol for 5 min at RT) in western blot transfer buffer at 100 V for 90 min at 4°C. Membranes were incubated in blocking buffer for 1 hr at RT. Primary antibodies were diluted in blocking buffer and added to the membranes. All incubations with primary antibodies were performed overnight at 4°C. The next day, membranes were washed 3x in TBS supplemented with 0.05% (v/v) Tween-20 for 10 min each. Secondary antibodies were diluted in blocking buffer and added to the membranes for 1 hr at RT. The membranes were then washed 3x with TBS supplemented with 0.05% (v/v) Tween-20 for 10 min (RT). Following the application of ECL western blotting detection reagents, membranes were exposed to Amersham Hyperfilm™ ECL High Performance Chemiluminescence film for 10 s to 10 min.

2.7 Baculovirus and insect cells

2.7.1 Production of the bacmid DNA

Bacmids were generated by repeated transformation of codon-optimized *X.laevis* HelQ plasmid DNA into *E. coli* DH10BAC™ chemically competent cells according to the manufacturer's instructions. Transformants were selected on LB plates containing 10 μg/ml tetracycline, 50 μg/ml kanamycin, 7 μg/ml gentamicin, 40 μg/ml IPTG and 100 μg/ml X-gal. As the mini Tn7 sites were placed in the middle of the *lacZ* gene, white colonies were generated by successful transposition. White colonies were picked from plates and re-streaked to verify recombination. Two colonies were used for viral production and amplified to P3.

Isolated bacmid DNA (1 µg) was transfected into two wells of a 6-well plate containing 2 x 10^6 *Spodoptera frugiperda* Sf9 cells using Fugene HD transfection reagent and FBS-free Grace's media. Transfection was performed according to Fugene HD's manufacturer's protocol and cells were cultured at 27°C. After 5 to 6 hr, FBS was added to a final concentration of 10% (v/v). The next day, growth media was replaced with fresh Grace's media supplemented with 10% (v/v) FBS. The supernatant of transfected Sf9 cells was collected 72 hr after transfection (P1 virus), centrifuged 5 min at 3000 rpm and transferred to a fresh tube covered with aluminium foil. Cells were pelleted by centrifugation, lysed in 300 µl of protein loading buffer and screened for optimal protein expression by western blotting. To test for expression of xHelQ, P3 virus was used to infect High5 cells. The culture was allowed to grow for 3 days after which the cells were lysed, and the soluble fraction incubated with Ni-NTA resin. The insoluble fraction was washed and dissolved in buffer containing 6 M urea and 5 mM DTT. An aliquot was taken for analysis by SDS-PAGE and the remaining solution incubated with Ni-NTA resin for 30 mins at room temperature. The resin was then washed extensively, and any bound protein eluted by the addition of SDS-PAGE

2.8 Peptide arrays – BCDX2-HELQ interaction site mapping

Peptide arrays were designed in such a way that each of the proteins of interest was divided into 20-mer peptide sequences immobilized on a membrane and each spot on the membrane had 19 residues in common with the adjacent spot (each position was shifted by one residue relative to its neighbour spots). The membranes were produced by the LRI Peptide Synthesis Facility.

2.8.1 Peptide arrays-based mapping of HELQ-FLAG binding site within BCDX2 complex

To identify HELQ interaction site within BCDX2 components, the RAD51 paralogs sequences were immobilized on membranes in a manner described above.

Recombinant HELQ was purified from HEK293 cell line expressing FLAG-tagged HELQ used previously in Mass Spectrometry and in vitro binding studies .

In brief, 20 mln of cells were harvested and lysed in Lysis Buffer (20 mM Tris-Cl, pH 7.5, 150 mM NaCl, 10% glycerol, 2 mM $MgCl_2$, 0.5% NP40, 30 U/ml benzonase, protease inhibitors). Lysates were cleared by centrifugation 3500 rpm, 10 minutes at 4C. Supernatants were then pre-cleared with Protein G agarose beads for 30 min at 4 °C on a rotating wheel. Pre-cleared lysates were incubated with anti-Flag beads (Sigma) for 4 h at 4 °C. To strip bound HELQ of its interacting partners the beads were washed five times with wash buffer with growing concentration of salt (20 mM Tris-Cl, pH 7.5, from 150 mM to 1M NaCl, 3 mM EDTA, 0.5% NP40) and then 4 more times with dropping concentration of salt 1M NaCl to 150 mM NaCl and once with FLAG Beads wash buffer (20 mM Tris-Cl, pH 7.5, 150 mM NaCl, 3 mM EDTA, 0.5% NP40). HELQ was eluted with 1ml of FLAG peptide elution buffer (TBS supplemented with 200 µg/ml 3xFLAG peptide) on a rotating wheel for 60 minutes at 4C.

The membranes containing arrays were activated by soaking in methanol for two minutes, washed in TBST (TBS+0,1% Tween-20), and then blocked for five minutes at RT in TBST-M. Membranes were incubated overnight at 4C in 20ml TBST-M buffer containing 100 µl of purified HELQ-FLAG. The following day, the TBST-M was removed, and the membranes were washed 4 times with 20 ml of TBST to wash off unbound proteins. To minimize the amount of primary antibodies used the membranes were transferred onto parafilm sheets and incubated with 7ml of TBST-M with anti-FLAG-antibody (1:1000) for an hour at RT. The incubation was carried out in a sealed container to minimize evaporation. The arrays were washed 4 times for 10 minutes in each with TBST, then incubated with 20 ml of TBST-M with rabbit and mouse-HRP conjugated secondary antibody (1:3000) for one hour at RT. Three 15-minutes washes with TBST were performed and membranes were subjected to visualisation with ECL substrate. After developing, the membranes were Ponceau stained for 45 minutes and then destained in successive washes of 5% acetic acid (v/v in dH_2O) until peptides spots were visible.

In an alternative approach, an attempt has been made to visualise HELQ-FLAG with anti HELQ antibodies. To do so the previously used arrays were stripped of bound proteins and antibodies using the protocol described below and incubated in similar

way with an exception of antibodies used: primary - mouse anti-HELQ (1:800) and rabbit anti-mouse HR-HRP conjugated.
To control for binding specificity a duplicate array were used treated according to the same protocol, with the exception that instead of HELQ-FLAG, similarly to in-vitro binding experiment, ALC-FLAG expressing cells were used.

2.8.2 BCDX2 complex binding site within HELQ studies by peptide arrays

In reciprocal approach HELQ sequence was immobilized in a form of peptide arrays and incubated with recombinant BCDX2. Recombinant HIS-tagged BCDX2 complex from Sf9 cells, similarly as in *in vitro* binding experiment was used. The arrays were probed in two different ways – primary mouse antibodies against HIS-tag (1:800) with secondary rabbit-HRP anti mouse antisera (1:3000) or, after stripping, rabbit anti RAD51C (1:400) with secondary mouse-HRP conjugated anti rabbit.
As a binding specificity control recombinant HIS-tagged SFRS1 protein was used.
To control for the specificity of antigen recognition identical arrays handled the same way but without primary antibodies in the detection step were used.

Chapter 3. HELQ is a DNA response protein that promotes RAD51 paralogue-dependent repair of DSBs

3.1 HELQ interacts with known DNA repair factors.

To gain insights into HELQ function in vertebrates and to elucidate the molecular context that it operates within, I decided to investigate the HELQ interaction network, reasoning that insights into its function could be made based on the identification of HELQ interacting partners (Figure 3-1). I first set out to purify HELQ and its binding partners from mammalian cells, resolve the immune-purified complexes on polyacrylamide gels and identify the co-precipitating proteins using Mass Spectrometry (MS).

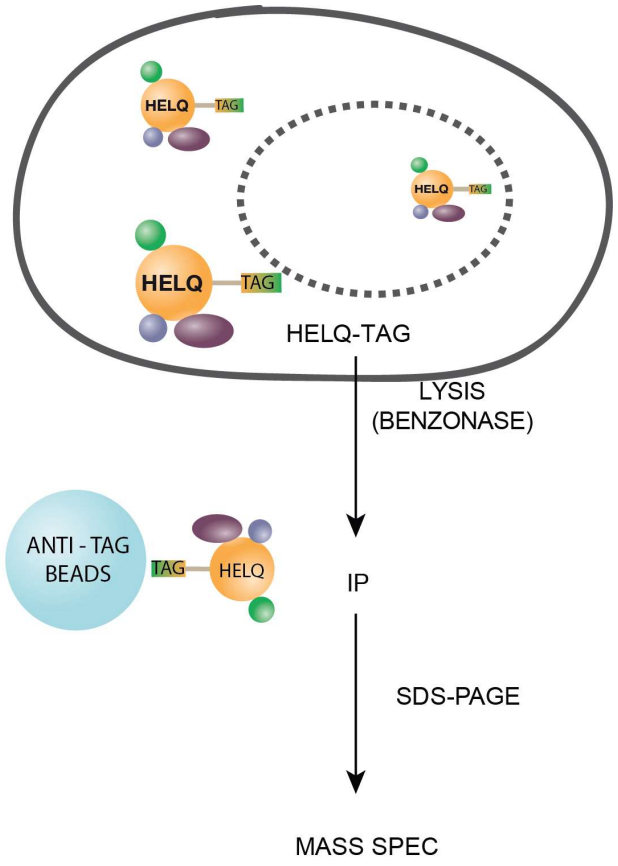


Figure 3-1 HELQ Purification Scheme.

A set of human and mouse HELQ-tag expressing cell lines were developed. In order to avoid non-specific interactions via DNA bridging between proteins, cells were lysed in presence of Benzonase. To identify interacting partners HELQ was affinity purified using

antibodies specific to the tag and immunoprecipitates were resolved on polyacrylamide gels. Proteins were then eluted from sliced gels, digested and peptides subjected to Mass Spectrometry analysis.

3.1.1 HELQ expression systems

In order to purify mammalian HELQ I tested two cell lines stably expressing tagged versions of HELQ – NIH-3T3 (mouse) and human HEK293.

3.1.1.1 Mouse NIH-3T3 HelQ-GFP

An NIH-3T3 cell line expressing GFP tagged HelQ was developed using the Bacterial Artificial Chromosome (BAC) recombineering method (Figure 3-2) (Poser et al., 2008). Due to its capacity, BAC recombineering allows insertion of transgene together with its native promoter and genomic locus sequence. Since the transgenes is expressed from its own promotor expression at near physiological levels is achieved, which reduces the likelihood of toxicity or non-specific interactions. This has proven to be particularly relevant in HELQ studies as the gene is expressed at extremely low levels and over-expression of HELQ leads to protein aggregates in cells. An additional advantage of this method is the introduction of two tags (EGFP, S-peptide) as well as two cleavage sites (TEV, PreSicission), which allows high specificity tandem purification and therefore obtaining pure protein prep. As a control, parental NIH-3T3 cell line was used.

3.1.1.2 Human HEK293 HELQ-FLAG

A Human HEK293 cell line expressing N-terminal FLAG tagged HELQ was also derived using Flp-In™ technology developed by Invitrogen (Figure 3-2). This system uses *Saccharomyces cerevisiae* flippase recombinase to stably integrate the gene of interest into a previously introduced FRT site in the genome of HEK293 cells (O'Gorman, Fox, & Wahl, 1991). N-FLAG tagged HELQ transgene was cloned under the control of the CMV promoter, which drives relatively high levels of expression. As a control, 293 FLAG Flp-In cells were used (293 cell line with engineered empty FLAG tag expression).

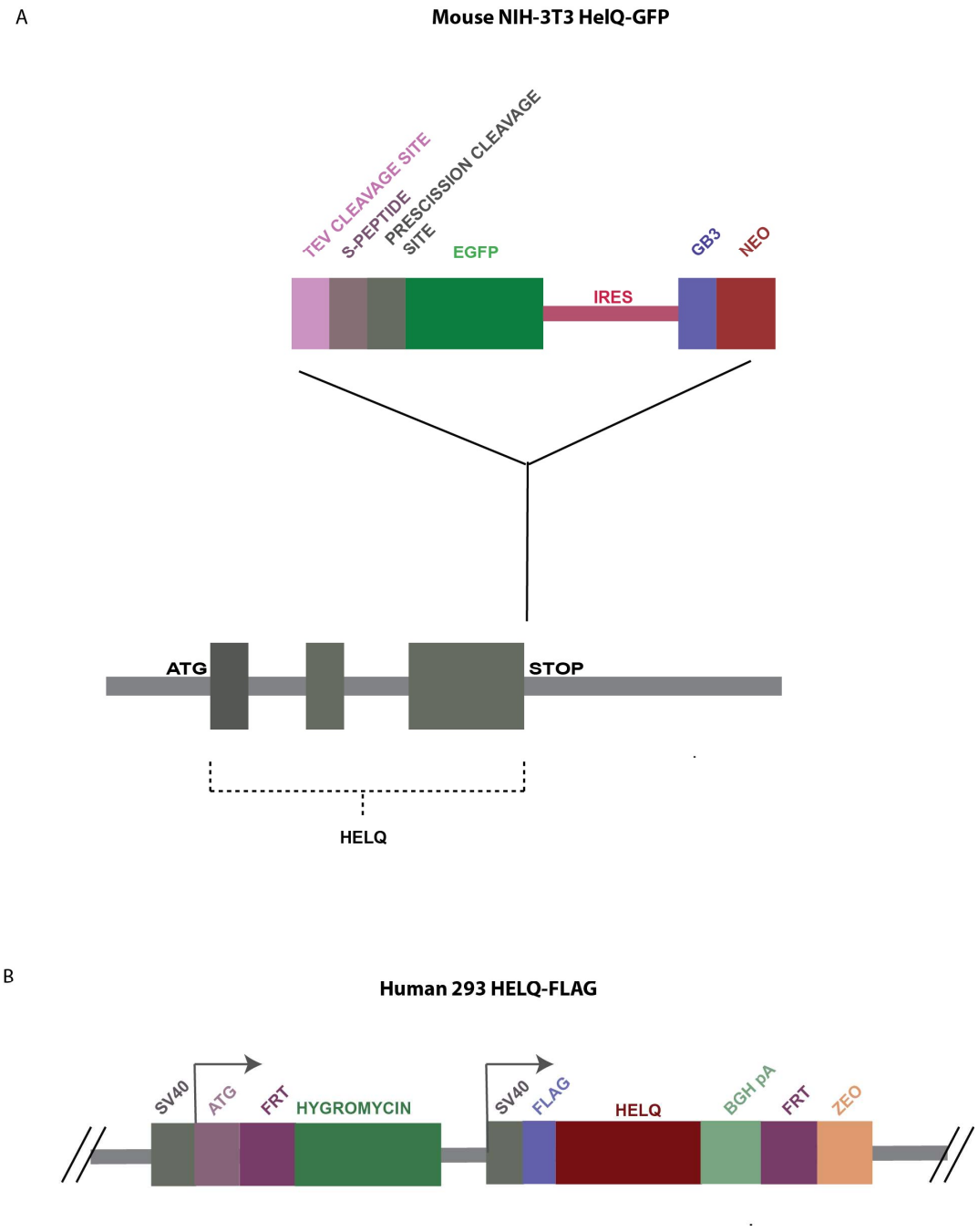


Figure 3-2 Schematic representation of HELQ expression constructs.

(A) C-terminal tag expression cassette used to develop NIH 3T3 HelQ-GFP cell line. EGFP, Enhanced Green Fluorescent Protein; GB3, bacterial promoter; IRES, Internal Ribosome Entry Site; NEO, Neomycin resistance gene. **(B)** N-terminal tag expression cassette used to develop 293 HELQ-FLAG cell line. BGH pA, Bovine Growth Hormone polyadenylation site; FRT, Flp Recombination Target; SV40, Simian vacuolating virus 40 promoter sequence; ZEO, Zeocin resistance gene.

3.1.2 HELQ expression validation

Before proceeding with the HELQ complex purification I first validated HELQ expression levels in both cell lines. Western blot analysis with antibodies against the respective tags were used, and clear bands of the appropriate size (around 120kDa for HELQ-FLAG and 150kDa for mouse HelQ-GFP) were present in both cell lines (Figure 3-3). Importantly, no equivalent signal was detected in the control cells. It is noteworthy that the western blot signal from mouse NIH-3T3 HelQ-GFP cells was more specific than from the human cells despite lower expression levels. This can be attributed to higher specificity of anti GFP antibodies. Additionally, it is notable that a smear visible just below 97kDa in 293 HELQ-FLAG cells most likely indicates minor degradation or premature translation termination, which is often reported for proteins expressed from strong promoters.

I also used anti HELQ antibodies to validate expression levels. Interestingly, I was only able to detect endogenous HELQ in HEK293 cells, with significantly stronger signal from HELQ-FLAG cells. I was not able to detect endogenous HelQ in the mouse NIH-3T3 cells with any of the available antibodies, suggesting that the levels of HELQ expression in these cells was low and below the detection of western blotting.

3.1.3 HELQ purification procedure – small scale

Once I had confirmed the validity of the HELQ expressing cell lines, I conducted a number of small scale test purifications (1-10 million cells per sample). Because of the differences in expression levels and tags used, optimization involved testing conditions such as cell number, lysis protocol and length of the immunoprecipitation step. The goal was to find conditions that resulted in satisfactory ratios between quantities and purity of the immunoprecipitate. Table 3-1 shows the most optimal conditions found for the small-scale purifications and Figure 3-4 presents examples of optimization steps for both cell lines.

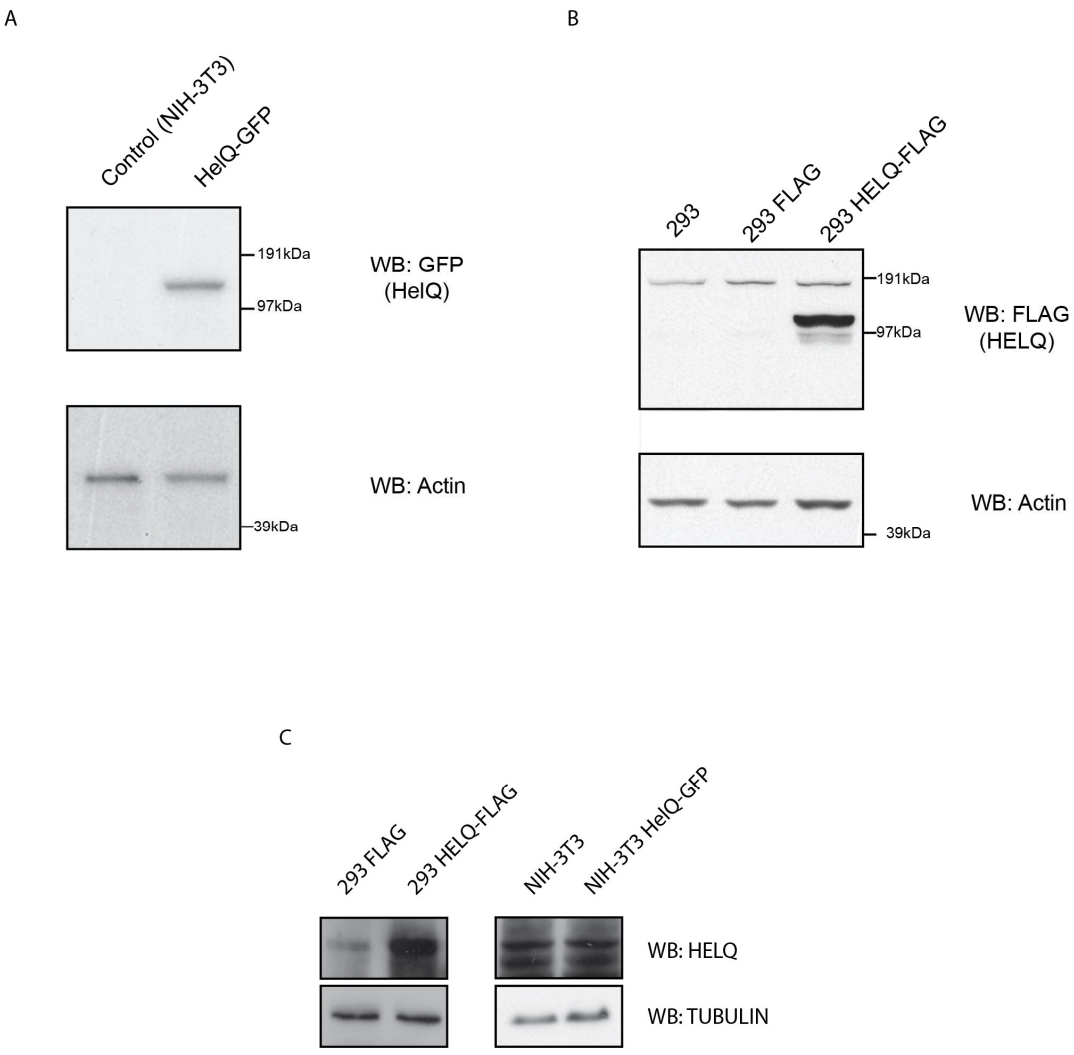


Figure 3-3 HELQ expression validation.

(A) Western blot analysis of NIH-3T3 HelQ-GFP cell line. Anti GFP antibodies detected only one band of the appropriate size. **(B)** Western blot analysis of HEK293 HELQ-FLAG cell line. Anti-FLAG antibody detected a strong band of the expected size (± 120kDa) and a number of smaller bands, which are most likely degradation products. **(C)** Western blot analysis of both cell lines using anti HELQ antibody. In control HEK293 FLAG cells endogenous HELQ was detected after long exposure times. Consistently with overexpression in HEK293 HELQ-FLAG, much larger band was detected. None of the available antibodies detected endogenous HelQ in mouse NIH-3T3 cells.

	293 HELQ-FLAG	NIH 3T3
Cell number	2 mln	10^6
Lysis method	Buffer 2 (no scraping)	Buffer 2 (no scraping)
IP time	2 hours	16 hours
Elution	FLAG peptide competition	Boiling

Table 3-1 Optimized HELQ IP conditions - small scale.

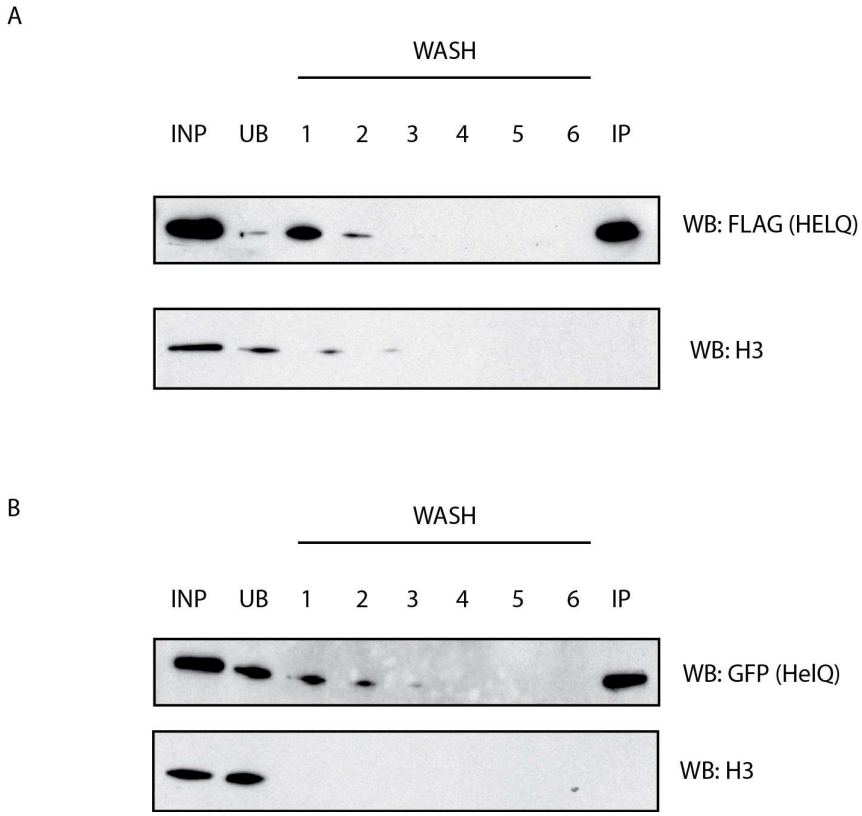


Figure 3-4 Exemplary HELQ IP optimization - small scale.

Cells (2 mln of 293 HELQ-FLAG and 10mln of NIH-3T3 HelQ GFP) were lysed, incubated with IP beads for 2 hours and subjected to 6 rounds of washing. Immunoprecipitates were released by boiling the beads. **(A)** 293 HELQ-FLAG cell line: 2 hours incubation with M2 resin was enough to IP almost entire HELQ-FLAG pool as shown by unbound fraction (UB). 2-3 washes were sufficient to remove detectable histone H3 levels IP line - eluate. **(B)** HELQ-GFP IP from NIH-3T3 cells. 2 hours incubation with anti GFP resin was not enough to remove majority of HELQ pool from cell lysate. Additionally, some of the bound fraction was released in the first 2 – 3 washing steps IP line – eluate.

As HELQ is a DNA processing enzyme it is noteworthy that the cell lysis step had to be optimized for use with Benzonase – a potent nuclease from *Serratia mercescens* that efficiently digests nucleic acid (DNA and RNA) and prevents formation of DNA fibre-mediated bridges that could potentially lead to false-positive hits in the MS analysis.

Expression of the epitope-tagged proteins allow for affinity purification using highly specific antibodies, which makes it a robust method to study recombinant proteins and their interacting partners. Mass Spectrometry is a powerful tool to interrogate the composition of protein complexes, however, confident identification of molecules is a function of quality (purity) and quantity of the sample. To design an optimal expression and purification system coupled with MS analysis a number of factors should be considered and optimized for individual protein. From previous work in the laboratory we knew that HELQ is present in cells at extremely low levels therefore we decided to try two alternative expression strategies.

EGFP is an excellent tag that increases solubility of the fusion proteins as well as allows very clean purification as there are highly specific antibodies available. At the same time, we decided to test the Flip-in system with human 293 cells that yields high level of protein expression and is a frequent choice in proteomic studies. Due to its small size and reasonably high hydrophilic properties FLAG peptide is less likely to affect the structure of the fusion proteins or cause degradation problems.

Between these two approaches in which one allows physiological level of expression of GFP-tagged HelQ and the other obtaining large quantities of HELQ-FLAG, I was able to find optimal conditions for efficient HELQ purification and identification of immunocomplexes.

3.1.4 HELQ purification procedure – large scale.

Once optimal small-scale immunoprecipitation conditions were established, the protocols were expanded to obtain sufficient material for MS analysis. After a number of minor adjustments in the protocols I was able to perform large-scale HELQ purification from both cell lines. Based on the differences in expression levels

between cell lines, I adjusted the number of cells accordingly (HelQ-GFP IP around 600 million of cells; HEK 293 HELQ-FLAG 100 million cells). Eluted immunocomplexes were resolved on polyacrylamide gels and were then sliced and subjected to MS analysis (Figure 3-5)

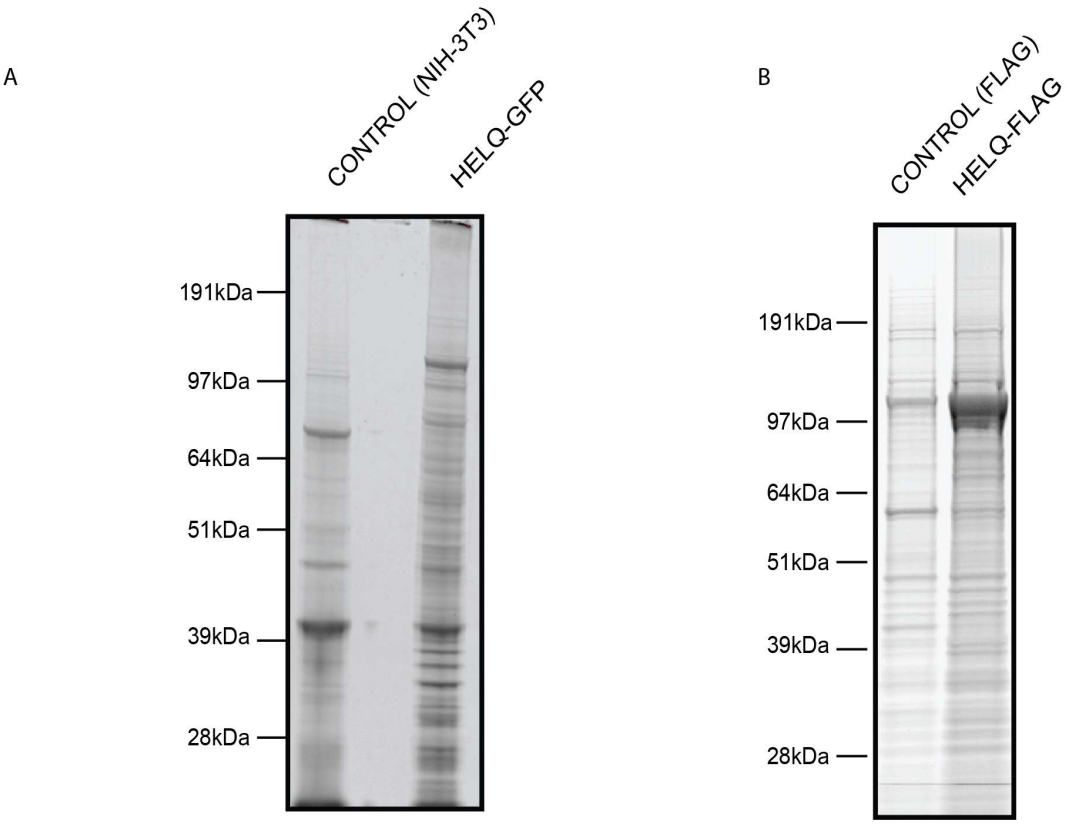


Figure 3-5 SYPRO Ruby-stained SDS-PAGE gels with large scale HELQ immunoprecipitation and copurified proteins.

(A) 600x10^6 NIH-3T3 cells per sample was lysed in the presence of Benzonase nuclease. Cell lysates were used for GFP immunoprecipitation with GFP-Trap resin. Upon extensive washing the beads were boiled to elute bound complexes and resolved on SDS-PAGE gels. **(B)** 100x10^6 human HEK293 cells per sample was lysed in the presence of Benzonase and the lysates were used for FLAG immunoprecipitation using M2 resin. Washed beads were boiled to elute complexes that were then resolved on SDS-PAGE gel. Proteins were visualized with SYPRO Ruby stain.

3.1.5 NIH-3T3 HelQ-GFP Mass Spectrometry analysis

The list of the most abundant peptides/proteins identified by MS from NIH 3T3 cell line is presented in (Table 3-2). The high numbers of peptides corresponding to cell skeletal proteins found in the control cells suggest contamination of the samples

most likely due to insufficient washing. As mentioned before obtaining immunoprecipitates of high purity is crucial for successful MS analysis as proteins non-specifically binding to the resin may mask positive hits.

	Protein	Number of total peptides	
		HelQ-GFP IP	Control IP
1	Myosin-9	890	68
2	Vimentin	812	197
3	Myosin-10	638	10
4	Myosin-XVIIIa	240	0
5	Myosin Va	89	0
6	Lamin A	71	8
7	Tubulin beta	58	20
8	Tubulin alpha	46	15
9	Filamin-B	44	17
10	**HelQ**	**24**	**0**

Table 3-2 List of most abundant proteins copurified with HelQ-GFP (mouse NIH - 3T3 cell line). The list lacks respective numbers of unique peptides.

Detailed analysis of the full list of identified peptides revealed several potentially interesting hits with lower peptide numbers. Most promising candidates included four of the Structural Maintenance of Chromosome (Smc) proteins – a group of ATPases involved in chromosomes dynamic and organization. Additionally, an uncharacterized ring finger protein, Rnf213 was also found (Table 3-3).

	Protein	Number of total peptides	
		HelQ-GFP	Control
1	Rnf213	26	0
2	Smc1A (Structural maintenance of chromosomes 1A)	9	0
3	Smc4 (Structural maintenance of chromosomes 4)	5	0

4	Smc3 (Structural maintenance of chromosomes 3)	2	0
5	Smc2 (Structural maintenance of chromosomes) 2	2	0
6	Mcmbp (Mini-chromosome maintenance complex-binding protein	3	0
7	Smchd Structural maintenance of chromosomes flexible hinge domain-containing protein 1	3	0

Table 3-3 List of potential binding partners of mouse HelQ-GFP. The list lacks respective numbers of unique peptides.

3.1.6 Human 293 HELQ-FLAG Mass Spectrometry analysis.

In contrast to NIH-3T3, samples derived from human HEK293 cells were of good purity and MS analysis provided high quality data. Analysis of the list of proteins co-precipitated with HELQ-FLAG revealed several potentially very interesting candidates with good peptide coverage (Figure 3-6).

I identified a number of intra-S-phase checkpoint and DNA damage response proteins including ATR, the replication checkpoint master kinase as well as ssDNA binding protein RPA70. Additionally, four components of the RAD51 paralogs BCDX2 complex were found – RAD51B, RAD51C, RAD51D and XRRC2. Lastly,

A

	Uniq / Tot Peptides	
ATR	24/32	ATR signalling
RPA70	21/25	
RAD51B	15/20	BCDX2 complex
RAD51C	8/13	
RAD51D	13/21	
XRCC2	9/15	
FANCD2	5/5	D2/I heterodimer
FANCI	2/2	

B

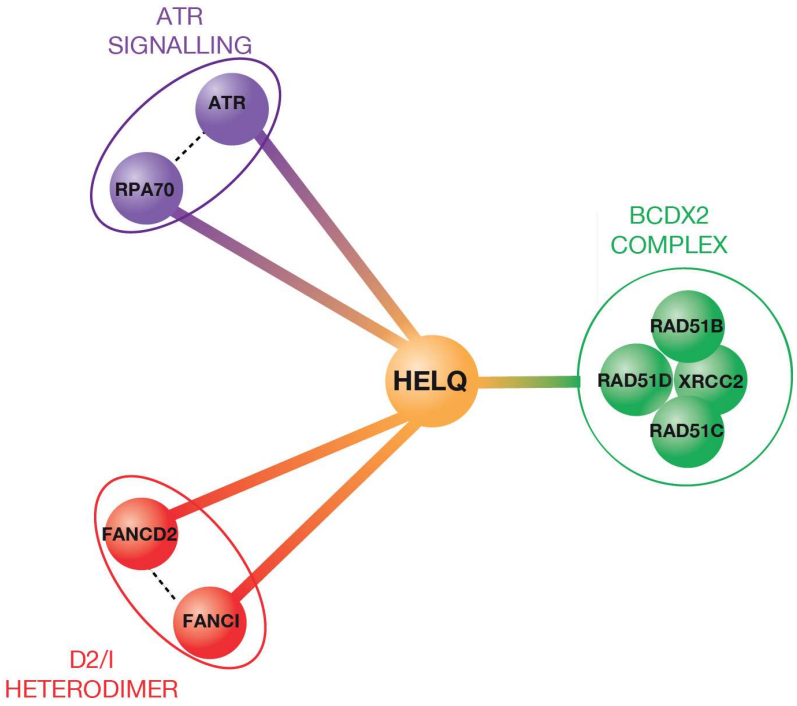


Figure 3-6 Mass Spectrometry results for HELQ-FLAG IP from human 293 FlpIn cells.

(A) List of most interesting candidate proteins showing unique and total number of peptides identified. **(B)** Schematic representation of HELQ binding partners identified in MS analysis.

I also found a small number of peptides corresponding to the FANCD2 and FANCI heterodimer, which are key components of the Fanconi anemia pathway.

3.1.7 Validation of the interactions identified by Mass Spectrometry

Mass Spectrometry analysis presented in this work would benefit from calculating Mascot probability index. For technical and software licensing reasons this calculation was not performed retrospectively at the time the thesis was revised.

candidate hits. To validate the HELQ-FLAG candidate hits I repeated FLAG IPs from the same cell line and probed the eluates with antibodies against proteins identified by MS. Although this confirmed the presence of FANCD2, ATR and RPA70, finding the optimal western blot conditions for RAD51 paralogs proved to be highly problematic and required extensive optimization efforts, almost certainly because of the very low expression levels of these proteins. In short, in order to find optimal protocol for western blot analysis of the BCDX2 complex components, I tested a number of different conditions. Apart from using various antibodies (and concentrations of thereof) I also assessed several types of polyacrylamide gels, transfer membranes as well as dilution buffers of various stringency. The most convincing results were obtained using high antibody concentrations (between 1:150 to 1:300), which caused strong unspecific background signal. Re-using antibodies solutions helped to deplete the pool of antibodies responsible for unspecific signal (Figure 3-8). Once optimal western blot conditions had been set up, I was able to detect all of the proteins previously identified in the HELQ-FLAG eluates (Figure 3-8)

To further validate the HELQ MS data, I took a reciprocal approach in which I used RAD51C as a bait and looked for HELQ-FLAG in the eluate (Figure 3-9). Extremely low expression levels of both HELQ and RAD51 paralogs together with lack of well-established antibodies proved to be a major obstacle to thoroughly analyse the interaction using immunoaffinity techniques such as WB, IP or IF.

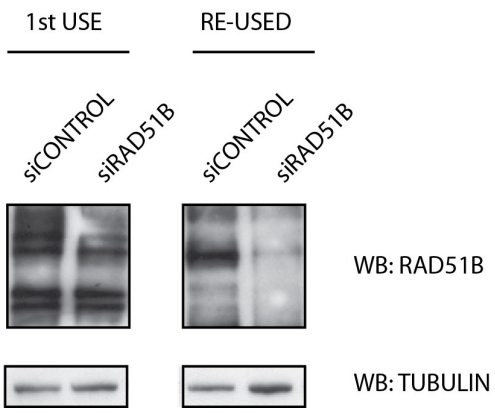


Figure 3-7 RAD51B western blot signal improvement over time.

U2OS cells were treated with control and anti RAD51B siRNA. Re-using antibody solution improves the specificity of the signal as pool of unspecific antibodies is depleted.

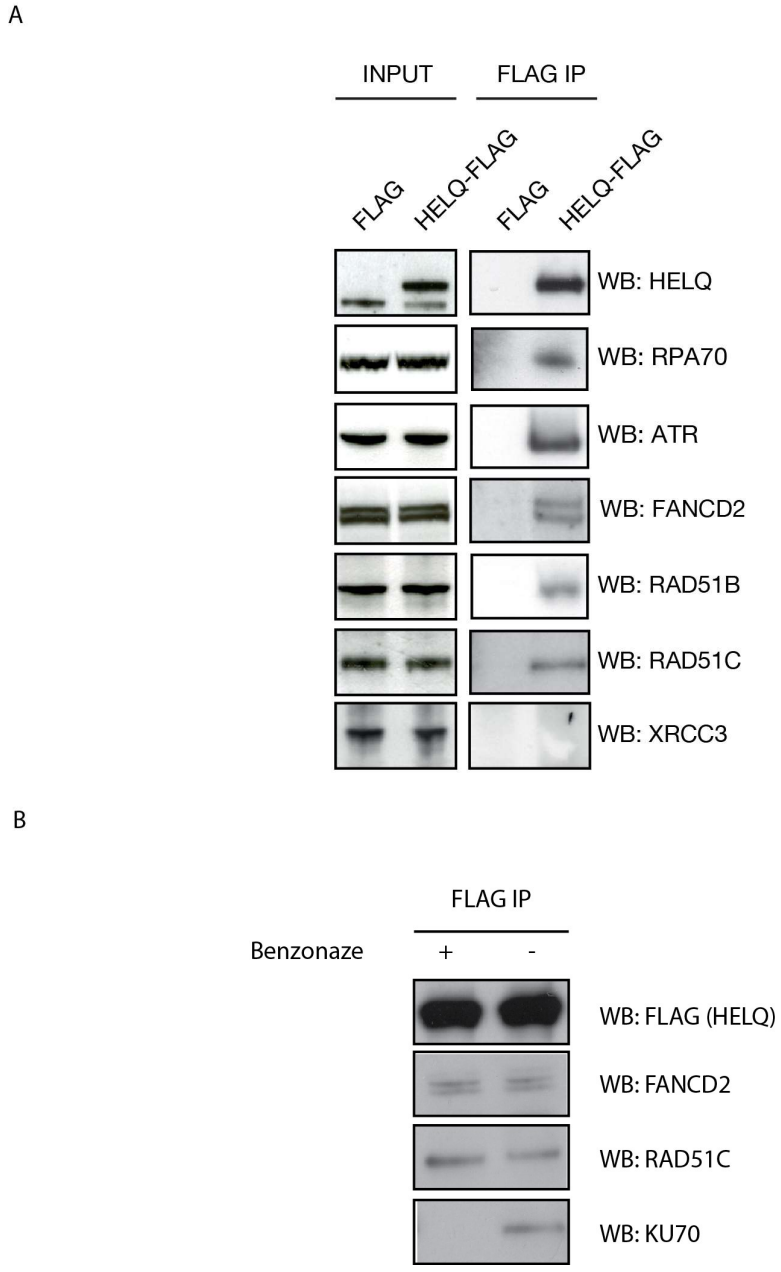


Figure 3-8 Western blot validation of HELQ interactions identified by Mass Spectrometry.

(A) Several rounds of FLAG IP was performed using the 293 HELQ-FLAG cell line. All proteins identified by MS were reproducibly found in the eluates. XRCC3 - another member of the RAD51 paralogs family was not found in the MS analysis and it did not coprecipitate with HELQ. **(B)** Control for benzonase treatment. FANCD2 and RAD51C interaction with HELQ was benzonase resistant. KU70 – an abundant DNA binding protein was found only in IP from benzonase untreated lysate.

Despite many attempts, the reciprocal IP of endogenous RAD51 paralogs in other cell lines turned out to be technically challenging. Reproducible results were obtained for RAD51C pull downs from HEK293 cells where endogenous HELQ was found to co-precipitate with the BCDX2 complex (Figure 3-9). The HEK293 cell line has the highest level of endogenous HELQ of the cell lines I have tested, which may explain why it was possible to detect the interaction between two weakly expressed proteins at their endogenous levels.

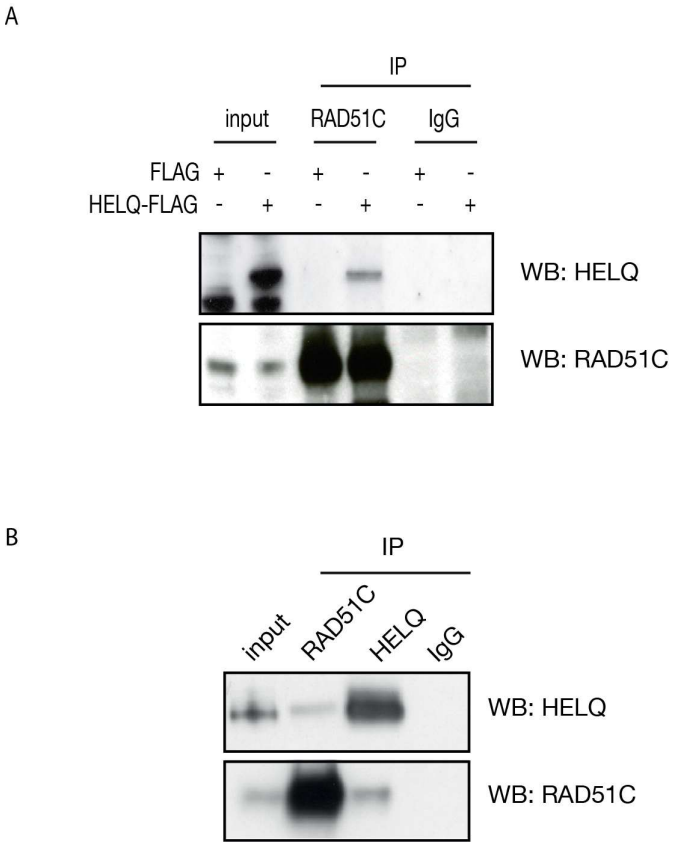


Figure 3-9 Western blot analysis of RAD51C IP

(A) RAD51C IP from human 293 HELQ-FLAG cell line. FLAG-tagged HELQ copurified with RAD51C. **(B)** Endogenous RAD51C IP from 293T cells. HELQ was detected in the eluates confirming the interaction on endogenous level.

As both HELQ and RAD51 paralogs are implicated in DNA repair I was curious to know whether the interaction is affected by DNA damage. To examine this possibility, I treated HEK293 cells with MMC, a potent DNA crosslinking agent and performed immunoprecipitation of endogenous RAD51C and HELQ.

As shown in Figure 3-10 the strength of the interaction between these two proteins does not seem to be influenced by DNA damage, as similar amounts of RAD51C were detected in HELQ eluates from cells treated with MMC or the untreated control sample.

Taken together, these results suggest that only a small fraction of cellular HELQ is bound to RAD51C (hence the BCDX2 complex) and the interaction is independent of DNA damage status.

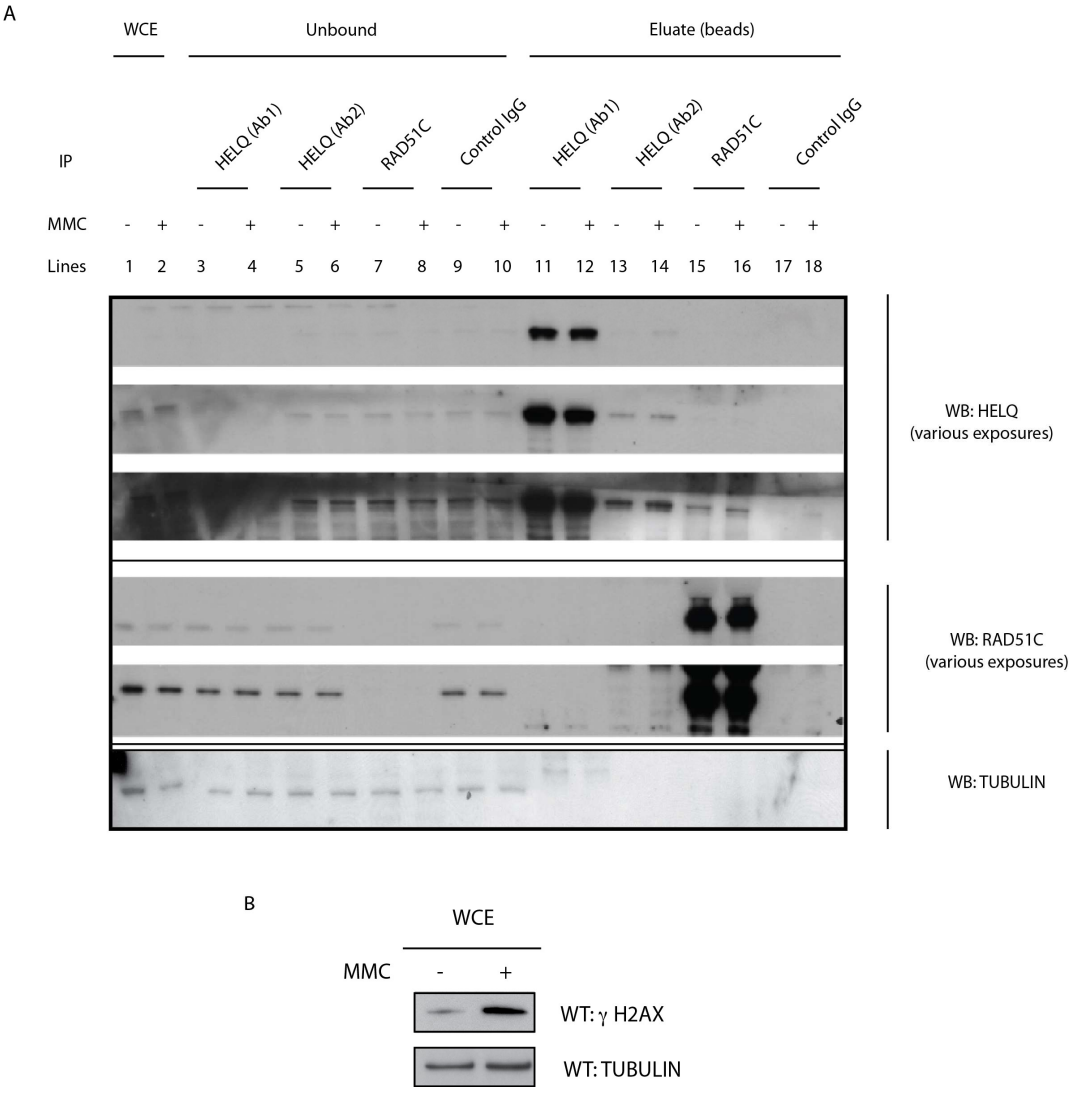


Figure 3-10 IP of endogenous HELQ and RAD51C from 293 cells.

(A) Endogenous HELQ coprecipitated with RAD51C (lines13 and 14) and the interaction was independent of MMC treatment (2μM). Although RAD51C IP depleted all detectable levels of the protein from cell lysate as shown by the unbound fraction, it did not have a significant effect on HELQ levels (lines 7,8) indicating that only a small fraction of cellular RAD51C is bound to HELQ. **(B)** MMC treatment control - histone γ–H2AX wb.

3.2 HELQ specifically interacts with BCDX2 but not CX3 of the RAD51 paralogs complexes.

In total there are five known RAD51 paralogs, which have been shown to assemble into several sub-complexes in cells, the two major ones including the BCDX2 and CX3 complexes (composed of RAD51C and XRCC3). It was notable that only the components of the BCDX2 complex were found in the MS analysis raising the possibility that HELQ interacts specifically with the BCDX2 complex but not the CX3 complex.

3.2.1 Endogenous CX3 complex IP

To confirm that HELQ does not interact with the CX3 complex, I probed HELQ-FLAG IP eluates for the XRCC3 protein. In repeated experiments I was unable to detect XRCC3 in HELQ IPs. (Figure 3-8). Additionally, I took a reciprocal approach using anti XRCC3 antibodies to IP endogenous CX3 complex from HEK293 HELQ-FLAG cells (Figure 3-11). As expected, RAD51C was detected in the XRCC3 IPs whereas HELQ was absent from the eluates. These data provide further evidence that HELQ interacts specifically with the BCDX2 but not with the CX3 complex of RAD51 paralogs. Given the separation of function of both complexes in homologous recombination, this specificity of interaction could be biologically relevant.

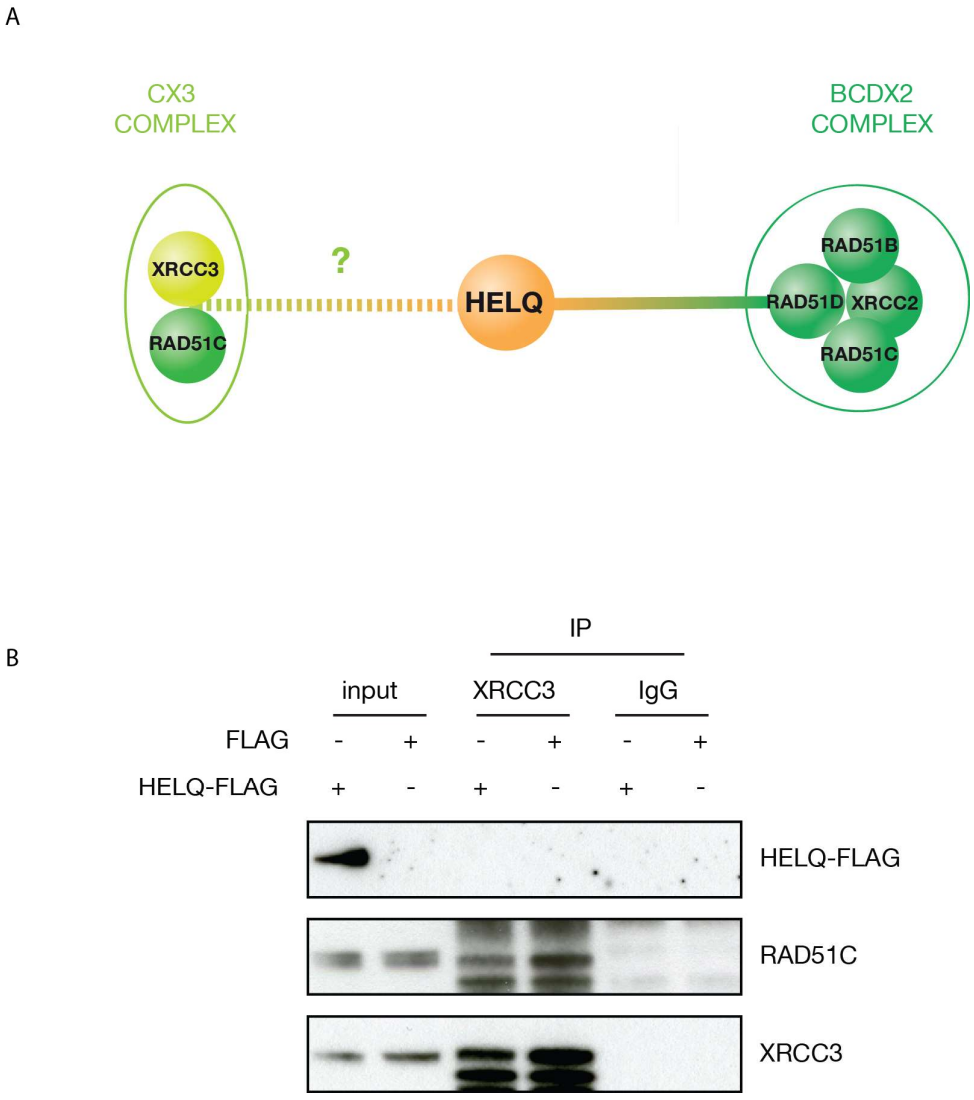


Figure 3-11 Endogenous XRCC3 IP.

(A) Schematic representation of confirmed and potential interaction between HELQ and RAD51 paralogs complexes. **(B)** Western blot analysis of endogenous XRCC3 IP from 293 HELQ-FLAG cell line. RAD51C was found in XRCC3 eluate, which indicates that the IP was performed in conditions preserving interaction within the CX3 complex. HELQ-FLAG did not co-purify with XRCC3.

3.2.2 HELQ – BCDX2 *in vitro* binding assay

The functional link between HELQ and BCDX2 complex was potentially very interesting so I decided to focus my efforts to study this interaction. The first question I asked was whether HELQ binds directly to the BCDX2 complex or whether this interaction is conferred by an intermediate factor(s). To investigate this problem I established an *in vitro* binding assay. First, HELQ-FLAG was affinity purified to homogeneity from HEK293 cells using M2 beads with a modified high salt protocol. In order to strip HELQ of all its endogenous interacting partners that could blur the binding assay, the beads were extensively washed with buffers of increasing salt concentration (up to 1M NaCl) (Figure 3-13 C). Following the wash step, beads were incubated with recombinant BCDX2 complex purified from either insect Sf9 cells or from *E.coli* (Figure 3-13). After the incubation beads were washed extensively and boiled. Eluates were then analysed by western blotting. As components of the recombinant BCDX2 complex were HIS-tagged I was able to detect the complex using anti HIS antibodies, which ruled out the risk of detecting any residual endogenous interacting partners. To control for the specificity of binding I also conducted two negative controls: HEK293 FlpIn engineered cell lines expressing either FLAG only or ALC1-FLAG proteins. ALC1 (Amplified in Liver Cancer 1) is a protein similar in size (100 kDa) to HELQ with functions in chromatin biology (Ahel et al., 2009).

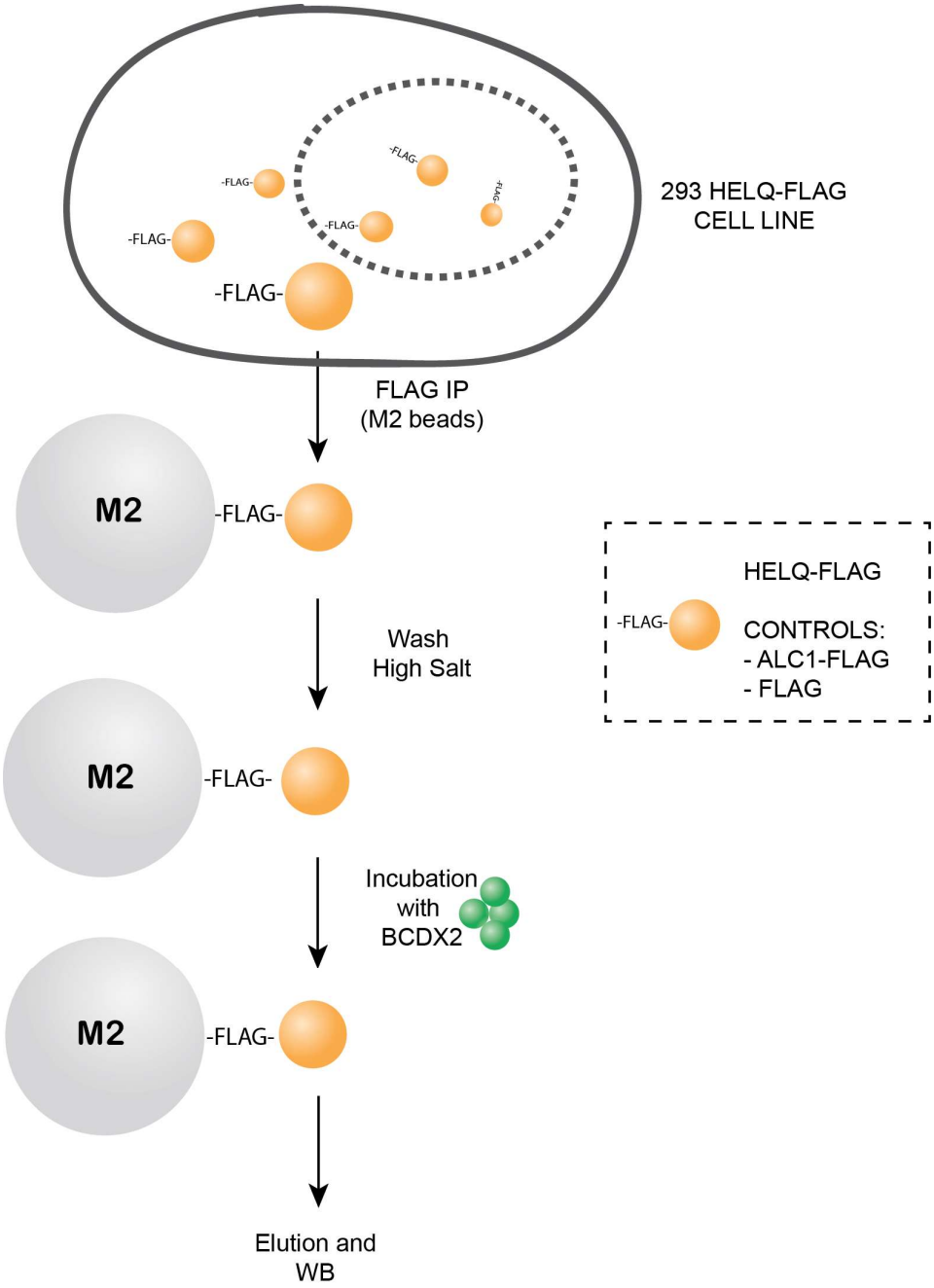


Figure 3-12 Schematic representation of HELQ-FLAG and BCDX2 complex in vitro binding assay.

Human 293 HELQ-FLAG cells were lysed as described before (control: 293 FLAG and 293 ALC1-FLAG). Upon FLAG IP, the M2 beads were washed with increasing concentration of NaCl to disrupt any protein-protein interactions. Following wash step, the beads were incubated with recombinant BCDX2 complex and washed again. Eluates were analysed by western blot.

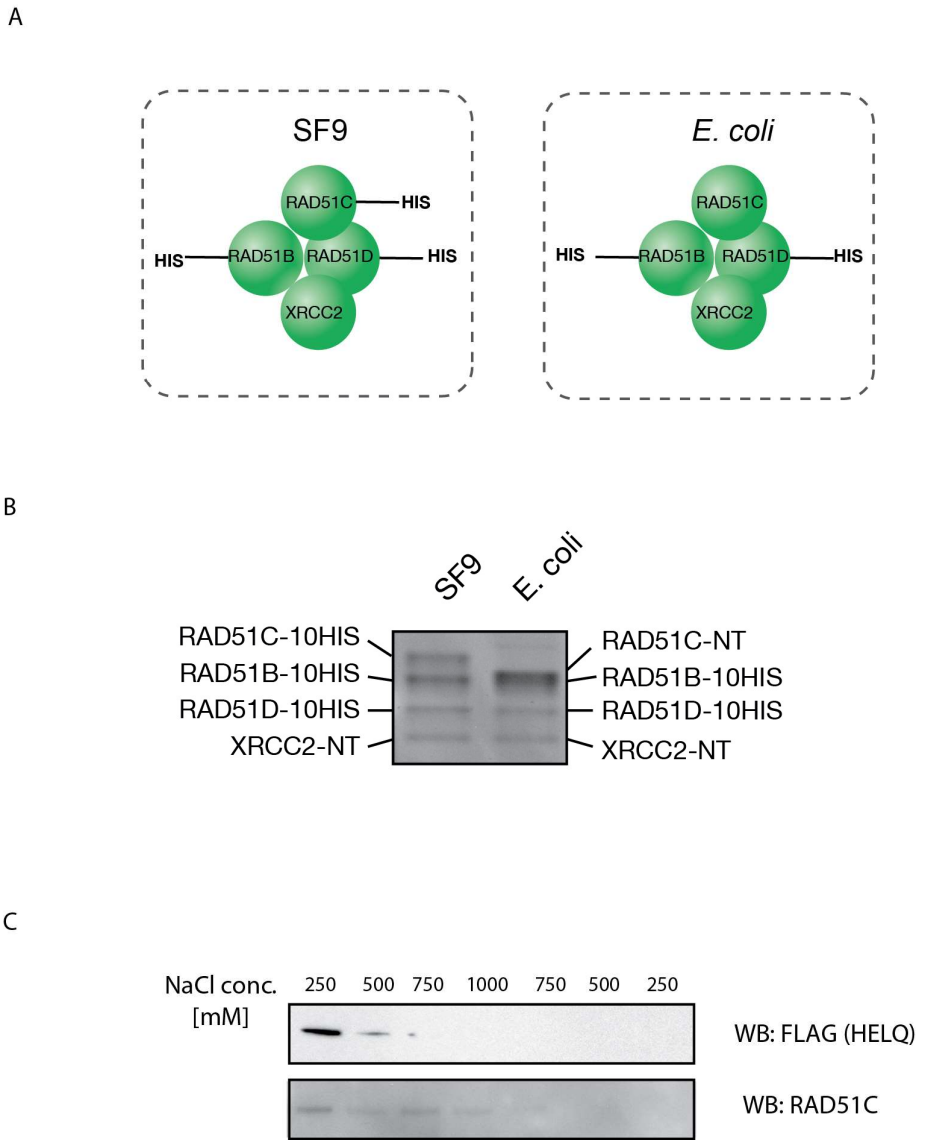


Figure 3-13 Recombinant BCDX2 complexes.

(A) Schematic representation of BCDX2 complexes from SF9 and *E.coli* cells. Three of the RAD51 paralogs (RAD51B, RAD51C and RAD51D) of insect origin had been HIS tagged whereas only two (RAD51B, RAD51D) from bacteria had HIS tag. **(B)** Sypro Ruby-stained SDS-PAGE with recombinant BCDX2 complexes from both sources. Because of different tagging strategy RAD51B and RAD51C from E.coli migrate as one band. **(C)** To remove endogenous binding partners, beads were washed with increasing concentrations of NaCl (up to 1000mM) and then gradually equilibrated to 250mM to ensure stringency of the subsequent binding assay. As indicated by RAD51C, BCDX2 complex was removed in the first 4 wash steps. Recombinant RAD51 paralogs were a kind gift from Stephen West.

As shown in (Figure 3-14) recombinant BCDX2 complex from both sources (bacteria and insect cells) bound to HELQ-FLAG but not to beads with just FLAG tag. The difference in the number of detected bands on western blot bands comes from differences in the tags used in bacteria and insects cells; three proteins (RAD51B, RAD51C and RAD51D) of the BCDX2 complex purified from insect cells carry a 6xHIS-tag whereas only two proteins (RAD51B and RAD51D) carry a 6xHIS-tag from the complex purified from bacteria (Figure 3-13). The BCDX2 complex was found to specifically bind to HELQ, RAD51C is found with both – HELQ-FLAG and the control FLAG only beads (for the later the interaction is notably weaker, though). One of the major on-going technical problems in the field of RAD51 paralogs has been very poor solubility of the components of both complexes when expressed separately. However, recent advances with insect cell expression of codon optimized cDNAs, together with the inclusion of an MBP tag and ATP in the purification buffers, has led to significantly improved purity and yields of the human CX3 complex.

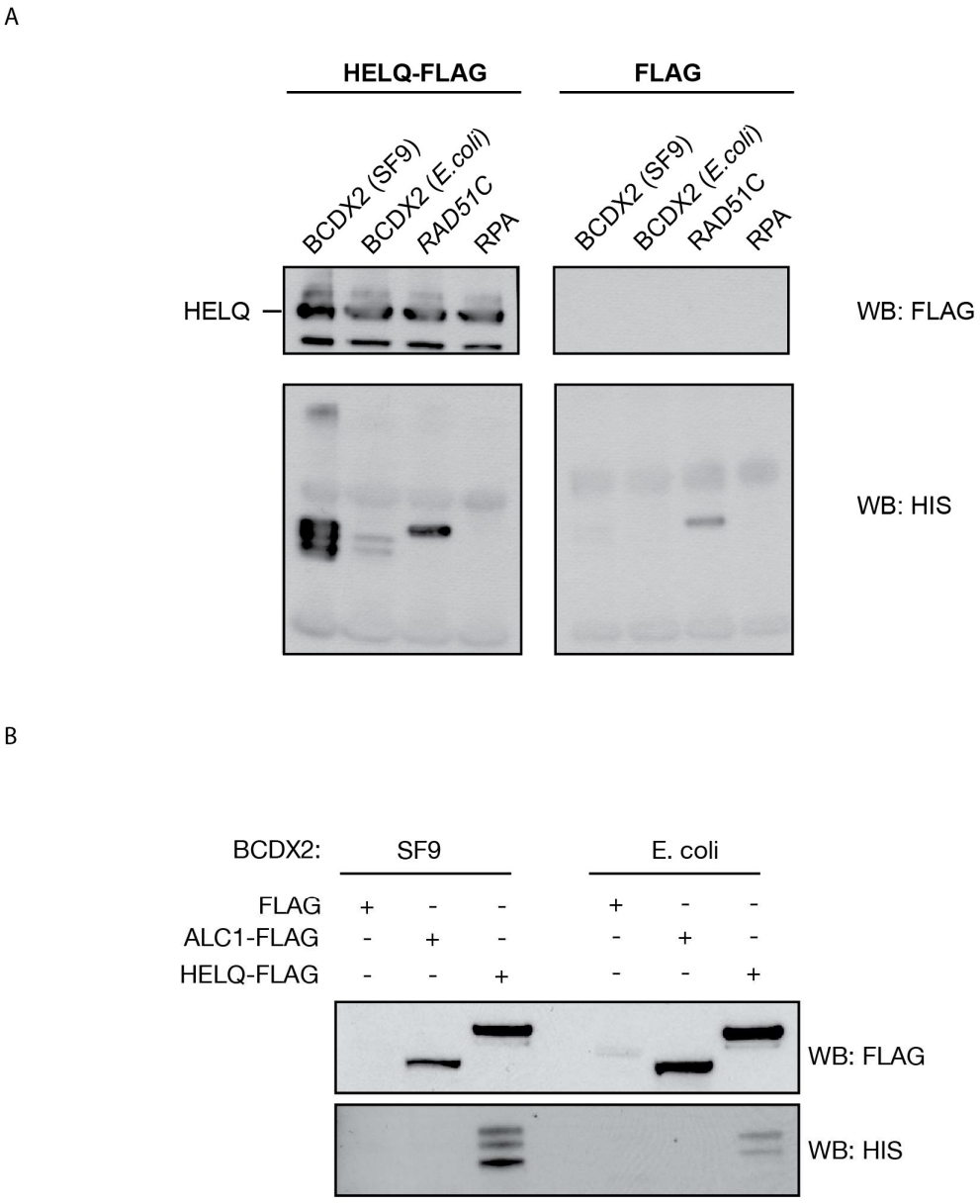


Figure 3-14 HELQ-FLAG - BCDX2 binding in vitro assay.

(A) Wb analysis of HELQ-FLAG and control FLAG-only binding to recombinant BCDX2-HIS complexes as well as RAD51C-HIS and RPA-HIS proteins. RPA did not bind to HELQ-FLAG (as RPA changes conformation upon DNA binding to draw more definite conclusions, binding experiment should include additional sample with RPA protein pre-bound to ssDNA). Small quantities of RAD51C bound to both HELQ-FL:AG and control FLAG only beads. **(B)** Wb analysis of HELQ-FLAG and BCDX2-HIS. As additional negative control ALC1-FLAG IP was introduced. For anti-HIS wb, the difference in bands number is caused by tagging strategies – three components of the BCDX2 (SF9 origin) are HIS-tagged and only two components of the BCDX2 from bacteria. Recombinant RAD51 paralogs were a kind gift from Stephen West.

It seems that forming complexes increases RAD51 paralogs stability in cells. It is very likely that RAD51C alone is misfolded and some hydrophobic residues are exposed to external environment, which may lead to unspecific interaction with M2 resin and/or FLAG tag. In the light of this data and what was known about RAD51 paralogs solubility, at the time, I decided that RAD51C was not a good candidate for *in vitro* binding experiment as it may give false positive results. I therefore decided to repeat the experiment with BCDX2 complexes under more stringent conditions and include additional controls. Figure 3-14 shows results of the modified experiment set up – apart from introducing ALC1-FLAG as additional control of unspecific (hydrophobic) interaction, incubation with BCDX2 complex was performed in 300mM NaCl (rather than 150mM NaCl). Under these more stringent conditions, HELQ-FLAG was found to specifically interact with the purified BCDX2 complexes from both insect and bacterial origins indicating that the interaction is almost certainly direct and does not require any mediating factors.

3.3 Discussion

Based on the existing data on HELQ's role in maintaining genome stability in lower organisms, I set out to gain insights into its function in vertebrates. I started by elucidating the molecular context it operates within by studying its interaction profile. In summary, I created a number of expression systems allowing for flexibility with tagging, purification and promoter strength. After optimizing expression and purification conditions I was able to obtain HELQ immunocomplexes from human HEK293 cells stably expressing FLAP-tagged HELQ that subsequently were analysed by MS.

I identified a number of intra-S-phase checkpoint and DNA damage response proteins including ATR, the replication checkpoint master kinase as well as ssDNA binding protein RPA70. Additionally, four components of the RAD51 paralogs BCDX2 complex were found – RAD51B, RAD51C, RAD51D and XRRC2. Lastly, FANCD2 and FANCI proteins were also found to coprecipitate with HELQ. This interaction was then validated by a number of approaches including western blotting of the HELQ immunoprecipitates and reciprocal IP with identified interactors as a bait.

I showed that HELQ binds specifically to BCDX2 complex but not CX3 of the RAD51 paralogs and that this interaction is most likely direct given recombinant HELQ and BCDX2 form a complex in vitro.

There are several ways in which the experimental setup or additional data analysis would reinforce obtained results. Calculation of Mascot score for each of the Mass Spectrometry-identified interacting partners would provide another line of validation ‘which would be particularly valuable for proteins with small /peptide numbers such as FANCD2 and FANCI. Although both proteins were repeatedly found in HELQ pull downs (as detected by Mass Spectrometry and Western blot) small peptide numbers together with functional data presented in Chapter 4 are reasons for caution in drawing definite conclusions regarding HELQ-ID complex interaction.
Additionally, the way original Western blot films were scanned made in impossible at the time of writing the thesis to properly position molecular weight marker on final figures. This problem affects all Western blot containing figures in the thesis.
As mentioned before, *in-vitro* binding studies were challenging due to difficulties with obtaining soluble recombinant HELQ. If more time became available, it would be interesting to extend the analysis of recombinant FANCD2- FANCI and also to see if addition of various DNA repair intermediate affects HELQ binding to its partners.

Identified interaction profile positions HELQ at the centre of repair events triggered by replication forks stalling at ICLs. In next steps it will be interesting to better understand the mechanism is which HELQ is recruited to damaged chromatin and what is its role in promoting repair mechanisms.

Chapter 4. Examining HELQ role in replication stress.

Considering the fact that BCDX2 complex as well as FANCD2/FANCI heterodimer and ATR together with RPA70 play an important role in response to DNA replication fork stalling, I was intrigued to examine HELQ subcellular localisation upon treatment with DNA damaging agents.

4.1.1 HELQ subcellular localization studies (Immunofluorescence)

My first approach was to perform IF studies but despite many attempts and testing different cellular systems my efforts failed due to lack of specific antibodies coupled with the extremely low expression levels of HELQ and the RAD51 paralogs.

Figure 4-1 shows results of some of my optimization efforts for endogenous HELQ IF studies performed in U2OS and HeLa cells. Apart from testing several antibodies and cell lines I also tried various fixation and pre-extraction methods. Unfortunately, I was unable to develop a reliable IF protocol in any of the cell lines tested. As attempts to visualize endogenous HELQ failed, Carrie Adelman, a postdoc in the lab, developed several cell lines expressing (either stably or transiently) tagged HELQ and tried to develop optimal IF procedures.

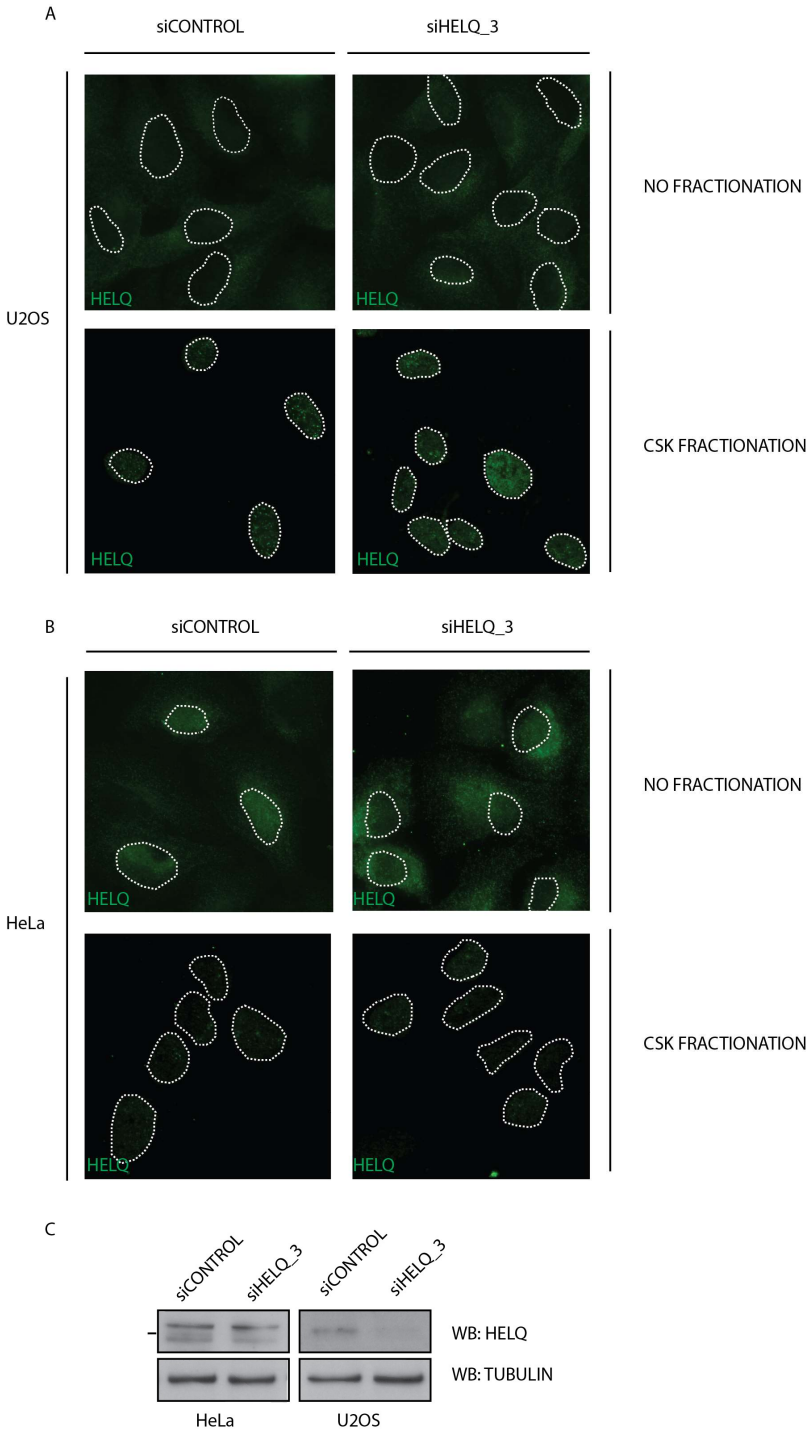


Figure 4-1 Endogenous HELQ IF (C. Adelman)

U2OS **(A)** and HeLa **(B)** had been depleted for HELQ with siHELQ_3 siRNA. In half of the samples soluble fraction had been extracted with CSK buffer. Cells were then fixed with formaldehyde and incubated with primary anti HELQ antibodies. **(C)** Wb confirmation of HELQ knock down in both cell lines.

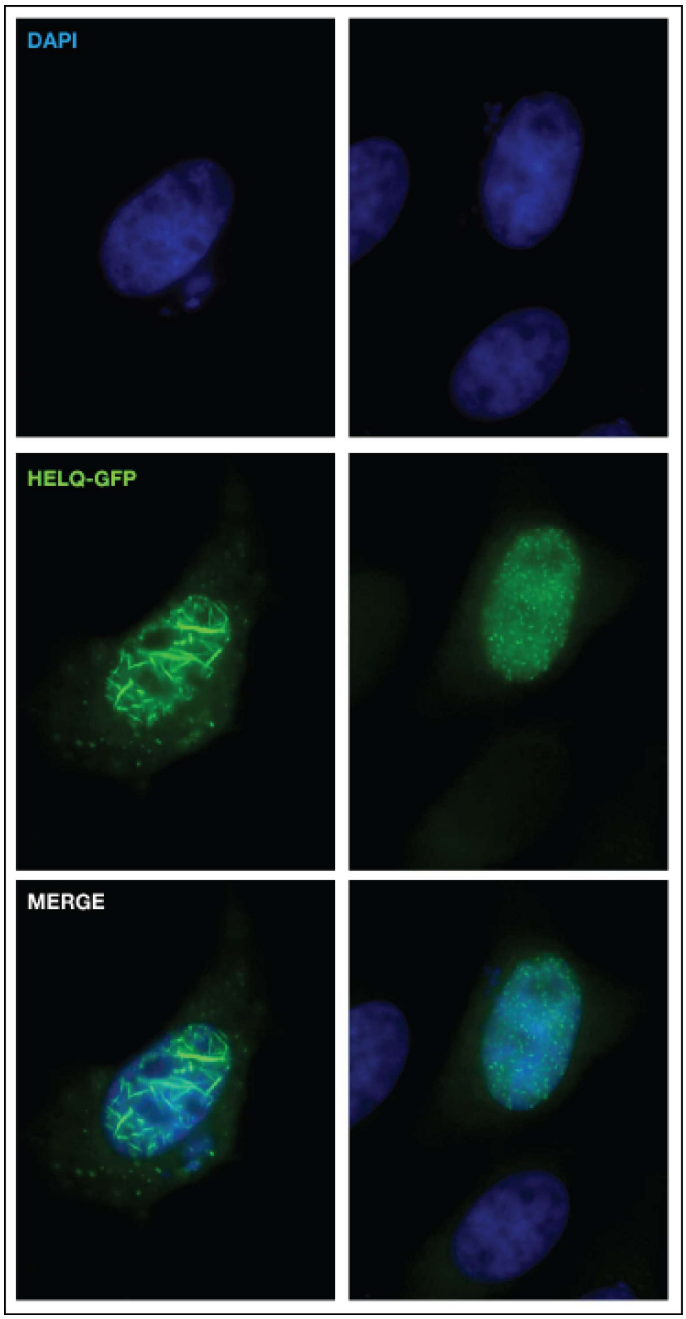


Figure 4-2 IF of HeLa cells transiently expressing HELQ-GFP (C. Adelman)

Cells were fixed and stained with DAPI. DAPI: blue; HELQ-GFP: green. Spontaneous nuclear aggregation is visible: filamentous (left panels), focal (right panel).

4.1.2 HELQ subcellular localization studies (Fractionation)

As our attempts to study the intracellular dynamics of HELQ by IF proved unsuccessful, I decided to take an alternative approach. Using subcellular fractionation techniques, I was able to isolate the cytoplasm, nuclear and chromatin compartments and analyse them separately by immunoblotting. The purity of the fractions was controlled by monitoring compartment-specific markers such as histone H3 that is only present in the chromatin fraction, tubulin that predominantly localizes to the cytoplasm and should not be found on chromatin.

I first established conditions for fractionation and then performed western blotting for endogenous HELQ and RAD51 paralogs in the separate fractions. Figure 4-3 shows fractionation results of the U2OS cell line. I was able to repeatedly obtain clear separation of the cellular compartments without major cross-fraction contamination. Unfortunately, despite extensive efforts, it was impossible to visualize endogenous HELQ or RAD51 paralogs on chromatin. To see a western blot signal for these proteins in the whole cell extract required high concentrations of the antibodies (1:200 for HELQ and 1:350 for RAD51C), which as mentioned before resulted in strong, unspecific background. It can be expected that only a part of total cellular HELQ pool is present on chromatin at a given time, therefore obtaining strong and specific western blot signal can be even more difficult than for the whole cell lysate.

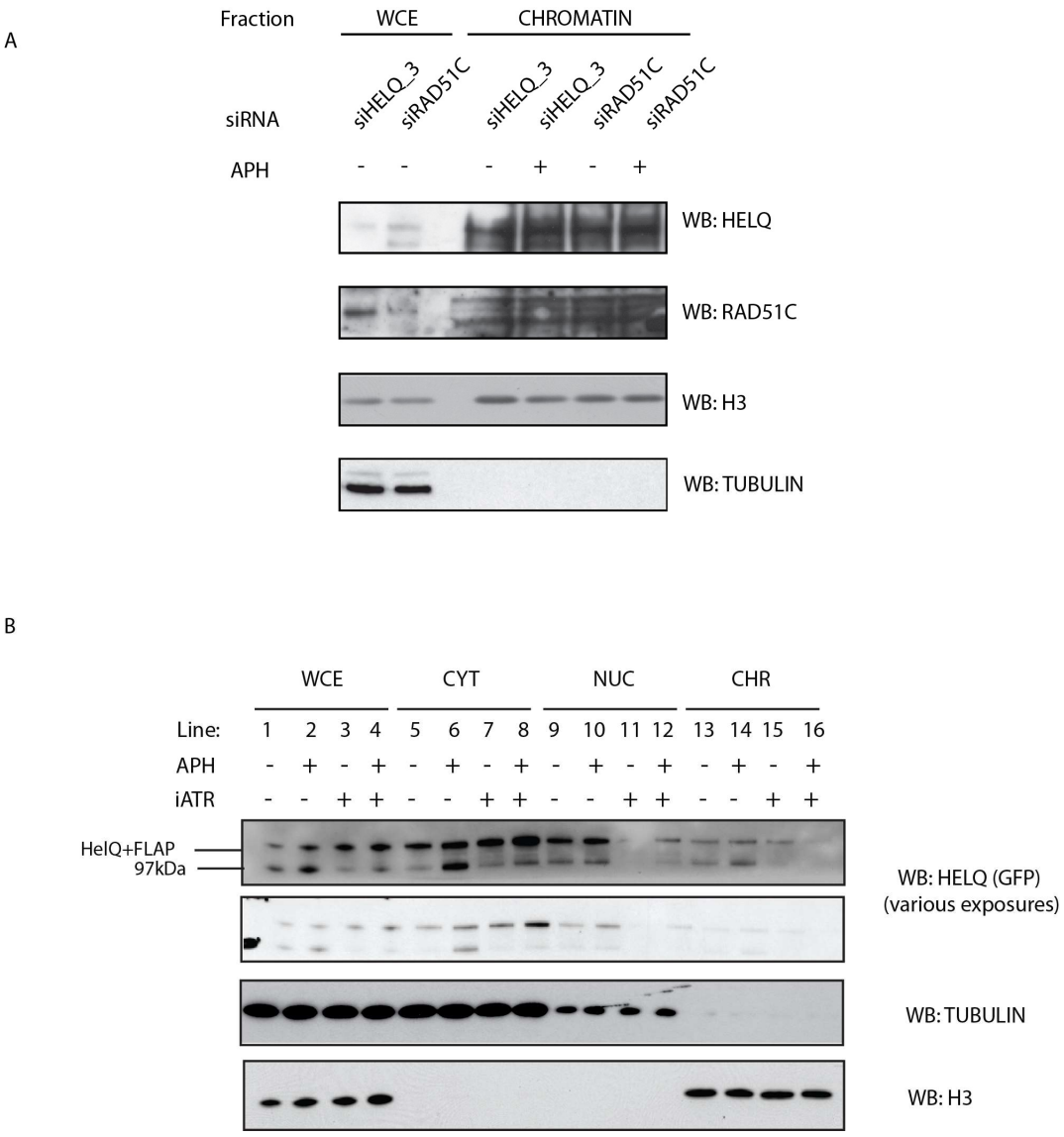


Figure 4-3 U2OS and NIH-3T3 HelQ-GFP cell lines fractionation results.

(A) Exemplary results of U2OS cells fractionation. HELQ and RAD51C siRNAs depleted U2OS cells were treated with APH to stall replication forks, Chromatin bound fraction was isolated and analysed by wb. As shown in wce HELQ and RAD51C knock downs were efficient indicating that bands observed in the chromatin fraction were unspecific. HELQ and RAD51C wb – Tris Acetate gel, Tubulin and H3 – BIS TRIS gel **(B**) Exemplary results of NIH-3T3 HelQ-GFP cell line fractionation. Cells were treated with APH and ATR kinase inhibitor to assess whether HelQ chromatin recruitment is DNA damage and ATR dependent.

As attempts to detect endogenous HELQ in the chromatin fraction failed I decided to test cell lines expressing tagged HELQ. My first choice was the mouse NIH-3T3 in

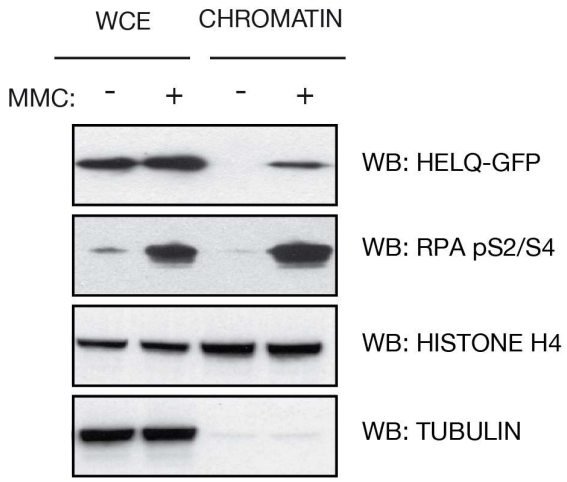
which GFP tagged HelQ is expressed from its native regulatory elements ensuring near physiological protein levels. Given our experience with HELQ overexpression leading to nuclear aggregation, the mouse NIH-3T3 cells seemed to be the best tool to study intracellular localization. Figure 4-3 shows exemplary fractionation optimization steps. In addition to localization studies, I also wanted to test whether aphidicolin treatment (leading to replication fork stalling) affects HelQ chromatin recruitment status and whether ATR plays a role in this effect. The latter was prompted by the fact that ATR functions in the replication stress response and was detected as a HELQ interacting partners.

As equal amounts of each fraction were loaded on gel (Figure 4-3) and there was no major cross-fraction contamination (as indicated by H3 and Tubulin markers) it is possible to compare how much of cellular HelQ localizes to different compartments. It seems that majority of HelQ pool is in cytoplasm and nucleus, and only a very small fraction is bound to chromatin. Interestingly, in cells treated with aphidicolin I observed a significant increase in the levels of HELQ in the chromatin fraction and this effect was abrogated by ATR inhibition.

Taking into consideration the fact that vast majority of cellular HELQ pool is in the cytoplasm and nucleus I assumed that even very small contamination of the chromatin fraction might contribute to dramatic changes in the HELQ levels. To minimize this possibility, I decided to change the fractionation procedure to a more stringent one. With the new method I was able to isolate two fractions – a soluble one (cytoplasm and nucleus) and an insoluble fraction containing DNA with bound proteins. After optimization this new method produced extremely clean chromatin fraction in a highly reproducible manner and was used in all subsequent fractionation experiments.

As preliminary results were promising, I decided to further study the effect of DNA damage on HELQ chromatin recruitment. As shown in Figure 4-4, MMC treatment result in the enrichment of HelQ-GFP on chromatin. Importantly, the increased HelQ chromatin status was not caused by elevated expression levels (as indicated by similar HELQ levels in whole cell lysate). Importantly, HelQ-GFP chromatin enrichment was also compromised upon ATR inhibition.

A



B

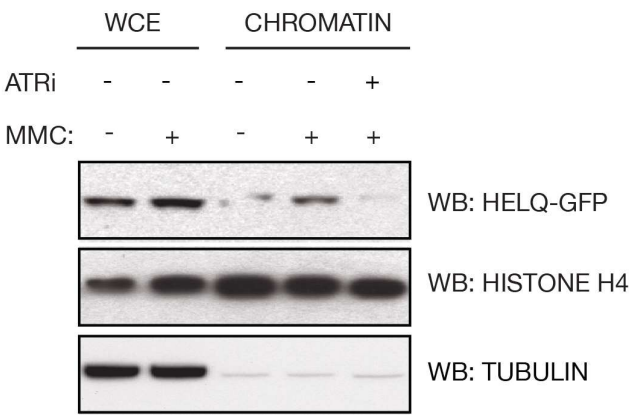


Figure 4-4 Western blot analysis of NIH-3T3 HelQ-GFP fractionation.

(A) Whole cell extract (WCE) and chromatin bound fraction of NIH-3T3 HelQ-GFP cells treated with or without 100ng/ml MMC. Phospo RPA – positive control of MMC-induced DNA damage. **(B)** Whole cell extract (WCE) and chromatin bound fraction of NIH-3T3 HelQ-GFP cells treated with or without MMC 100ng/ml and ATR inhibitor 3mM.

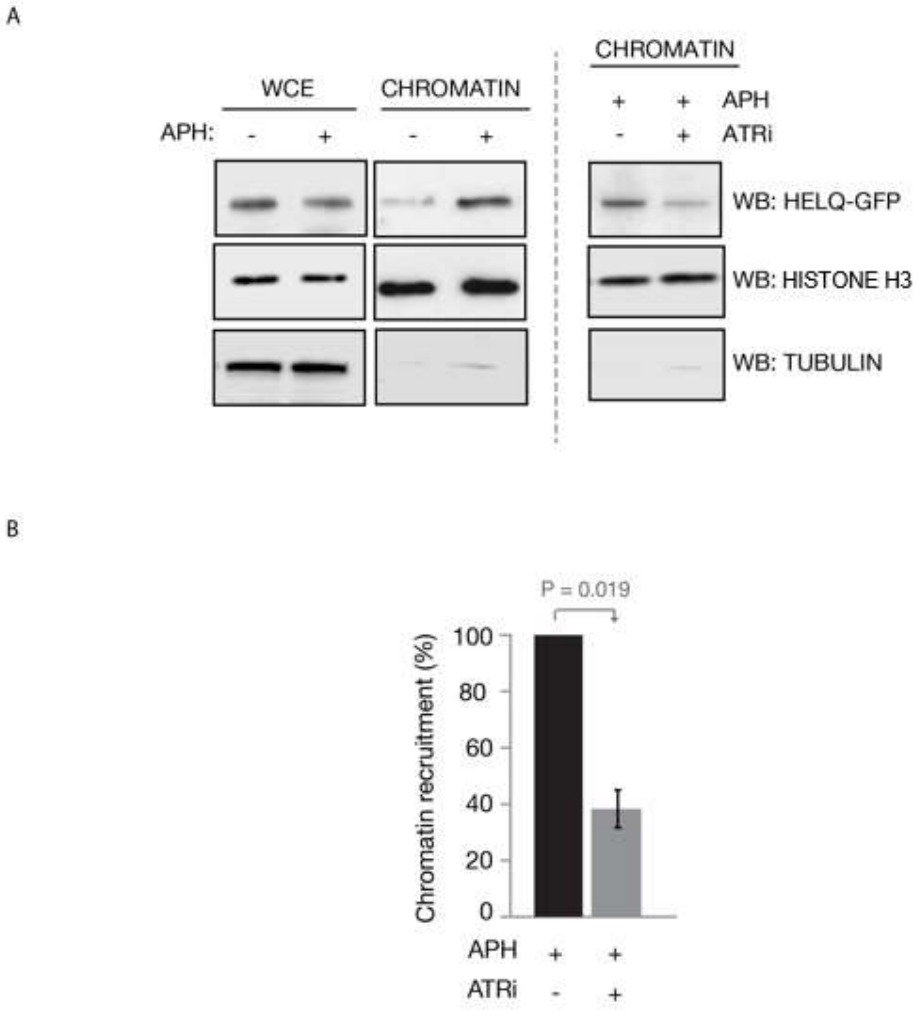


Figure 4-5 Western blot and densitometry analysis of DNA damage induced HelQ-GFP chromatin recruitment.

(A) Whole cell extract and chromatin fraction of NIH-3T3 HelQ-GFP cells treated with or without 2μM APH and 3mM ATR kinase inhibitor for 24h. **(B)** Densitometric comparison of HelQ abundance on chromatin in NIH-3T3 HelQ-GFP cells treated as in A. HelQ-GFP signal was normalized to histone H3 levels.

Densitometric analysis of APH treated HelQ-GFP chromatin levels from three independent experiments showed that upon ATR inhibition HelQ-GFP levels drop to around 40% of the values noted for APH only samples (Figure 4-5).

Taken together, localization studies showed that a vast majority of cellular HelQ is localised in the cytoplasm and nucleus and in undamaged cells, with only a very

small fraction of the protein bound to the chromatin. Upon replication fork stalling, HelQ is recruited to chromatin where it presumably performs its function in the repair process. Importantly, HelQ chromatin enrichment appears to be regulated, at least in part, by ATR kinase activity.

Although subcellular fractionation provided some important insights into HelQ dynamics upon DNA damage, poor resolution of this method only allowed monitoring significant shifts of the protein pools between cellular compartments. Employing more direct tools such as IF would allow studying those processes in a more detailed way. There is still a number of important questions that remain unanswered. It would be interesting to know the mechanism of HelQ chromatin enrichment – whether it is an active process in which the HelQ is transported to DNA or perhaps ATR (or its downstream targets) modify HelQ hence increasing its affinity to DNA (or DNA repair machinery). Additionally, IF studies would allow to assess whether HelQ colocalizes to DNA repair machinery markers such γH2AX or ubiquitinated FANCD2.

4.1.3 iPOND-based studies of HELQ dynamics at replication forks.

Although I showed that HELQ is recruited to damaged DNA in an ATR-dependent manner I decided to gain further insights into this phenomenon. I was curious to validate whether HELQ is recruited to broken replication forks where it may perform its function to promote HR-dependent repair. IF is usually a method of choice with regards to monitoring protein dynamics at replication fork. However, given the fact that HELQ is expressed at extremely low levels and our attempts with immunofluorescence of overexpressed HELQ failed due to aggregation problems, I decided to try isolation of proteins on nascent DNA (iPOND) (Sirbu et al., 2012; Sirbu et al., 2011; Sirbu et al., 2013)

iPOND allows the study of protein recruitment to active, stalled and collapsed replication forks and offers a high spatio-temporal resolution of the dynamics. In brief, this technique is based on incorporation of 5-ethynyl-2'-deoxyuridine (EdU) – a thymidine analog into a nascent strand of newly synthetized DNA and click chemistry with the aim to pull-down the EdU-labeled strand together with associated proteins.

Adding EdU to cell culture medium in short pulses causes rapid incorporation and labeling of newly-synthetized DNA dependent on the time of the pulses. The cells are then treated with formaldehyde to snap-stop replication and crosslink proteins to DNA. Subsequently EdU is conjugated to biotin in the copper-ion dependent click manner. Cells are then lysed, and DNA is fragmented by sonication. Using streptavidin-coated beads the nascent EdU-labelled DNA fragments are purified. Eluted DNA-protein complexes can be subjected to Western blot or Mass Spectrometry analysis to identify precipitated proteins.

Spatio-temporal resolution of this method is therefore determined by replication fork velocity, length of the labeling pulses and the size of DNA fragments after sonication. Apart from identification of proteins associated with active replisome and studying chromatin changes upon replication fork passage, iPOND is a very useful method to detect proteins or their posttranslational modifications associated with damaged replication fork. This requires a different experimental set-up where cells are shortly pulsed with EdU and then a damaging agent is added. A chase with thymidine step following EdU pulse can be included into the experiment to provide additional control – only proteins present at active replisome should be present in the pulse fraction. Proteins such as histones should be found in both pulse and chase samples. Although many fork-stalling agents are known, for this application hydroxyurea (HU) is particularly useful as in high concentration it efficiently stalls all replication forks (Petermann et al., 2010).

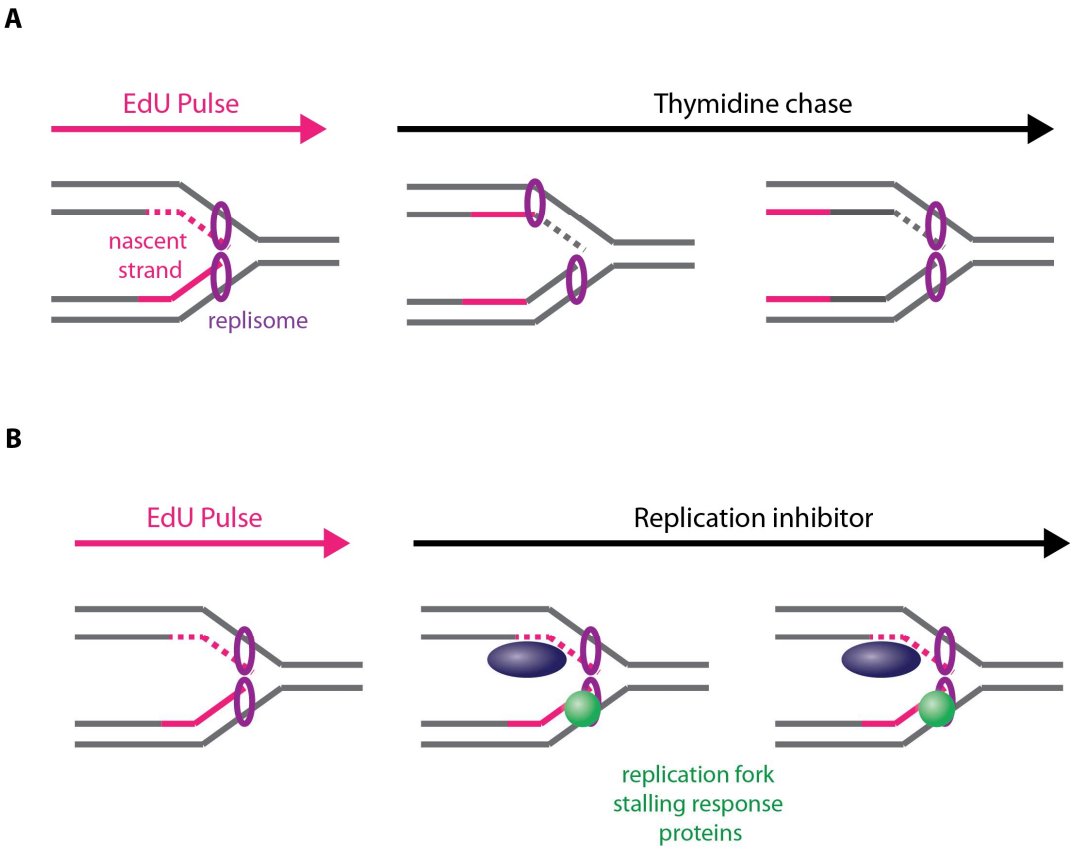


Figure 4-6 Identification of Proteins On Nascent DNA (iPOND) Methodology

Schematic of pulse and chase experiment **(A).** Cells are pulsed with EdU, washed and Thymidine is added. Only proteins found in the chase sample are associated with replisome. **(B)** Identification of proteins associated with stalled replication fork is obtained by EdU pulse immediately followed by addition of fork stalling agent such as HU.

To verify whether HELQ is recruited to stalled replication forks I used HEK293 stably expressing HELQ-FLAG (Figure 4.7). In my experiment I included a pull-down control – a sample pulsed with EdU and treated exactly the same way but without addition of biotin. To control for proteins generally present on chromatin a sample pulsed with EdU and chased with thymidine was included into the experiment set up.

Our DNA fibers experiments showing the replication fork dynamics is affected in MEFs derived from HelQ deficient mice (data not shown – experiments conducted by Carrie Adelman) prompted me to check whether HELQ is recruited to active replication forks. iPOND is a great tool to study proteins associated with active replisome – this is possible by analyzing the sample only pulsed with EdU in which the replication forks are unaffected.

And finally, to detect proteins present at stalled replication fork I included sample where cells had been pulsed with EdU and treated with HU for two hours.

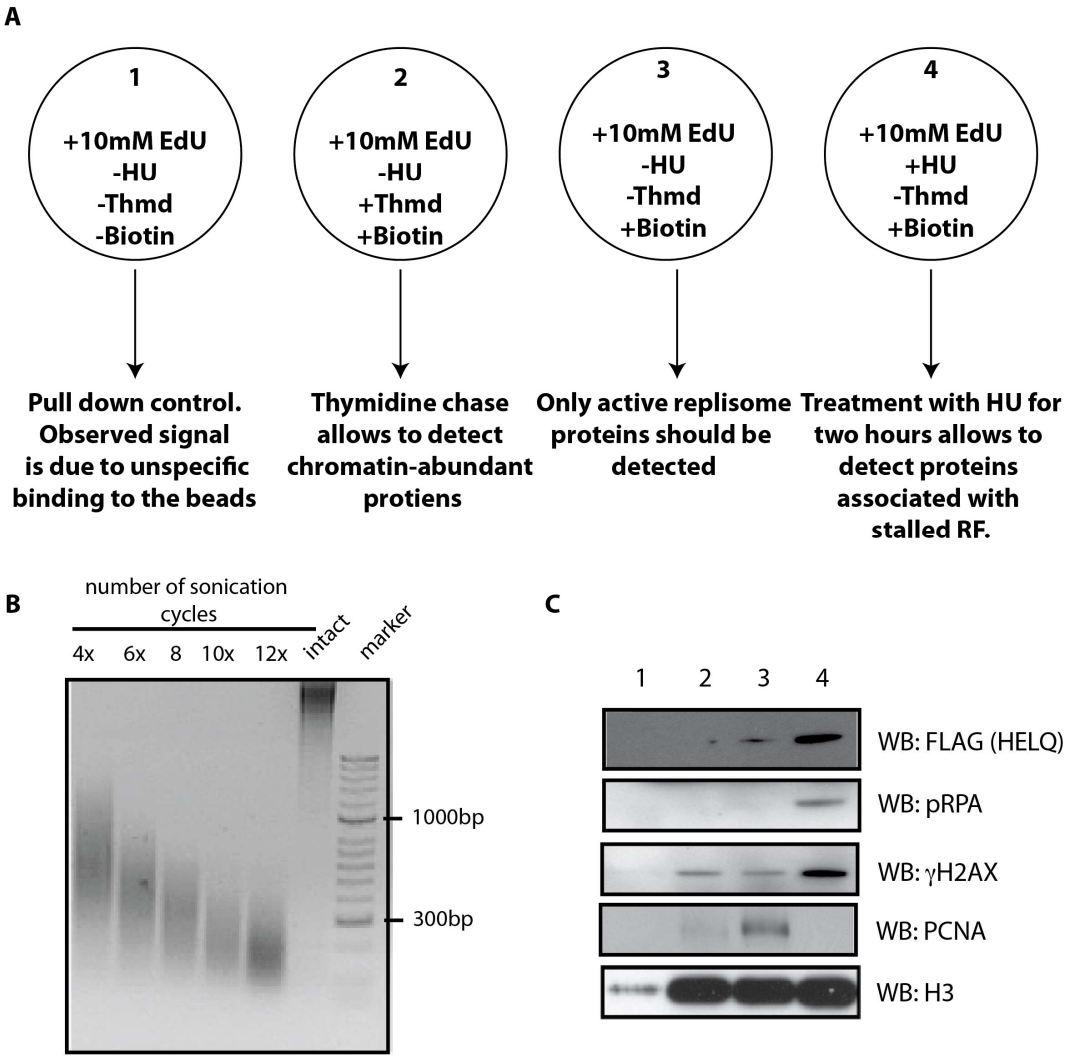


Figure 4-7 HELQ is recruited to stalled replication forks as shown by iPOND.

(A) iPOND - Experimental set up. **(B)** DNA sonication optimization. To obtain optimal DNA fragments (100-300bp) 12 cycles of sonication was needed. **(C)**. Eluate from streptavidin-coated beads was resolved on SDS-Page gel and subjected to Western blot. Sample numbers corresponds with numbering in panel A.

The procedure required optimization consisting mainly on increasing the number of cells per sample as well obtaining the right size of DNA fragments upon sonication.

It is crucial for maintaining the right resolution as well as reproducibility of the method to work with DNA 100 to 300 nucleotides long. As sonication efficiency depends on the cell type, buffer in which the procedure is performed and a number of other factors I put a considerable effort to optimize the sonication protocol (Figure 4.7 B). A major improvement in reproducibility of sonication was achieved by using sonicating water-bath and replacing ice with ice-cold water.

Once the crucial steps of the protocol were optimized I proceeded to perform the full experiment. As shown in Figure 4.7 C, the pull-down control is relatively clean as indicated by H3 levels in the sample without biotin. PCNA is a component of active replication forks and can be found in sample pulsed with EdU (line 3). Only very small levels of PCNA are detected in sample chased with thymidine (line 2). Although traditionally phosphorylation of histone H2AX is an indicator of double strand break (Dickey et al., 2009), there have been reports showing that γH2AX is present at chromatin at much earlier points than formation of DSB – just upon HU triggered replication fork stalling (Sirbu et al., 2012). Accordingly, γH2AX is detected in sample treated with HU (line 4). Weak γH2AX signal is also present in lines 2 and 3.

Interestingly, DNA-PK has been shown to phosphorylate RPA at persistently stalled forks, a phenomenon that occurs at around 2 hours of HU treatment (Anantha et al., 2007). S4/S8 phosphorylation of RPA was detected in sample treated with HU for 2 hours indicating prolonged fork stalling.

Intriguingly, HELQ-FLAG was detected mainly in the HU-treated sample (line 4) but also, small quantities could be found in the unaffected, active replisome (line 3). Together these results suggest that HELQ may play some discreet role on active replication fork and upon treatment with fork stalling agents it is enriched to perform its repair function.

4.2 HELQ functions in the DNA Interstrand Cross Links repair in parrarel to the FANCD2-FANCI.

To study the effects of HELQ deficiency in vertebrates Carrie Adelman, a postdoctoral researcher in Simon Boulton's laboratory generated *Helq*$^{\Delta C/\Delta C}$ mouse strain carrying deletion of the C-terminal region of the protein.

4.2.1 HelQ mouse model generation (Carrie Adelman)

The ΔC/ΔC mutation was designed to remove part of the helicase domain as well as the entire ratched region that is required for the separation of DNA strands. The resulting N-terminus together with β–Gal fusion product was most likely degraded as neither HelQ nor β–Gal epitopes were undetactable in the cells derived from the mutant strain (Figure 4-8).

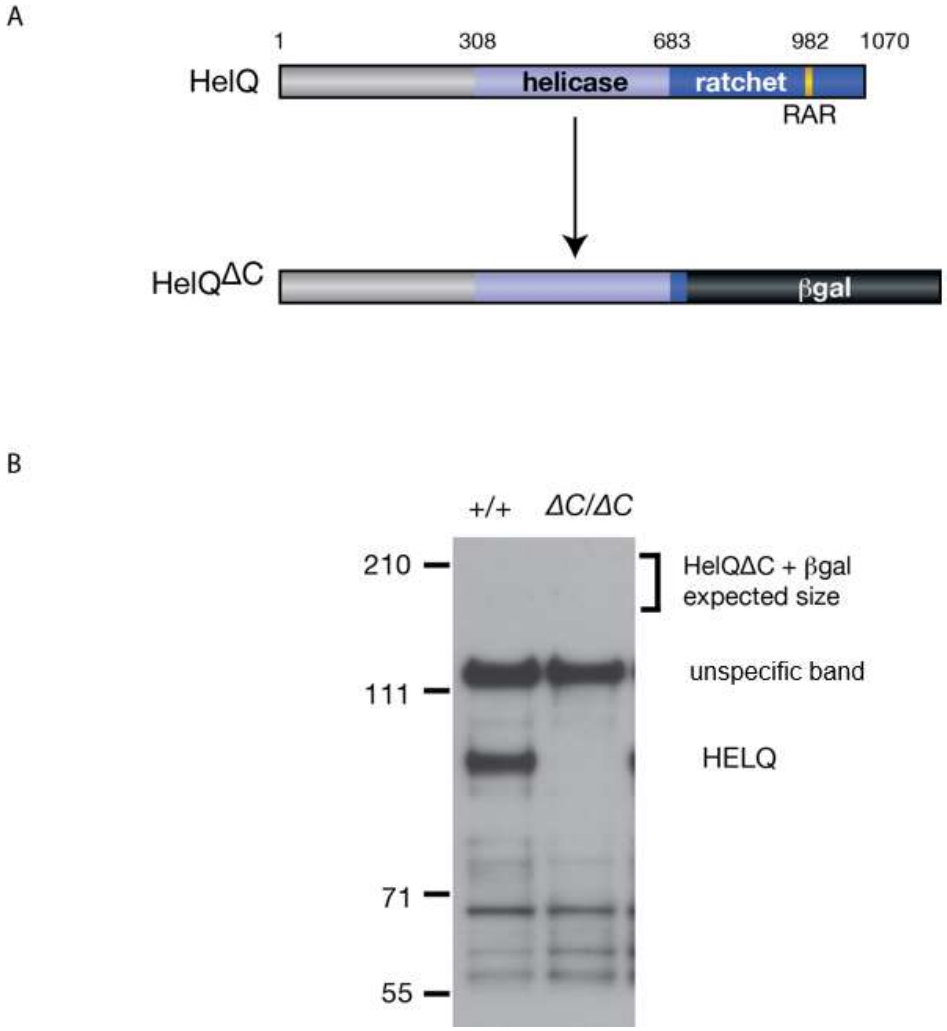


Figure 4-8 Schematic representation of HelQ$^{\Delta C/\Delta C}$ deletion introduced in the mouse model.

(A) Cartoon representation of insertional mutation in *Helq* gene introduced using gene trapping. N-terminal region including the entire rachet domain as well as a small fragment of the helicase domain were replaced by β-gal gene. **(B)** Wb analysis of wild type (+/+) and HelQ mutant MEFs. HelQ$^{\Delta C}$+βgal fusion protein was not detected.

HelQ deficient mice phenotype was characterized mostly by Carrie Adelman. I am providing a short summary of Carrie's findings as it is absolutely crucial to fully understand and appreciate the role of HelQ in DNA repair in vertebrates. A short summary of the phenotype can be found in Table 4-1. Figures containing C. Adelman's data are labelled accordingly.

Mouse Embryonic Fibroblasts (MEFs) derived from HelQ$^{\Delta C/\Delta C}$ mice together with accompanying cells from the control strain were an excellent tool to study effects of HelQ deficiency and were extensively used by both C. Adelman and myself.

	HelQ$^{\Delta C/\Delta C}$	**FancD2$^{-/-}$**
Viability	**+**	**+**
Subfertility	**+**	**+**
Hypogonadism	**+**	**+**
Cancer prone	**+**	**+**
Radial chromosomes upon ICLs	**+**	**+**
Radio sensitivity	**-**	**-**

Table 4-1 Comparison of HELQ and FANCD2 model mouse phenotypes.

4.2.2 HelQ confers crosslinking agents resistance.

As shown in Figure 4-9 HelQ deficient cells exhibit hypersensitivity to replication fork stalling agents such as MMC or CPT yet we observed no obvious sensitivity to either ionizing radiation or ultraviolet light. Interestingly, cells lacking HelQ activity seemed to be relatively more sensitive to the type of damage introduced by MMC than CPT. This phenotype, together with the lack of UV sensitivity and my interaction data, point towards a role for HelQ in the DNA interstrand cross links repair pathway, which would be consistent with our previous observations in *C. elegans* (Ward et al., 2010a).

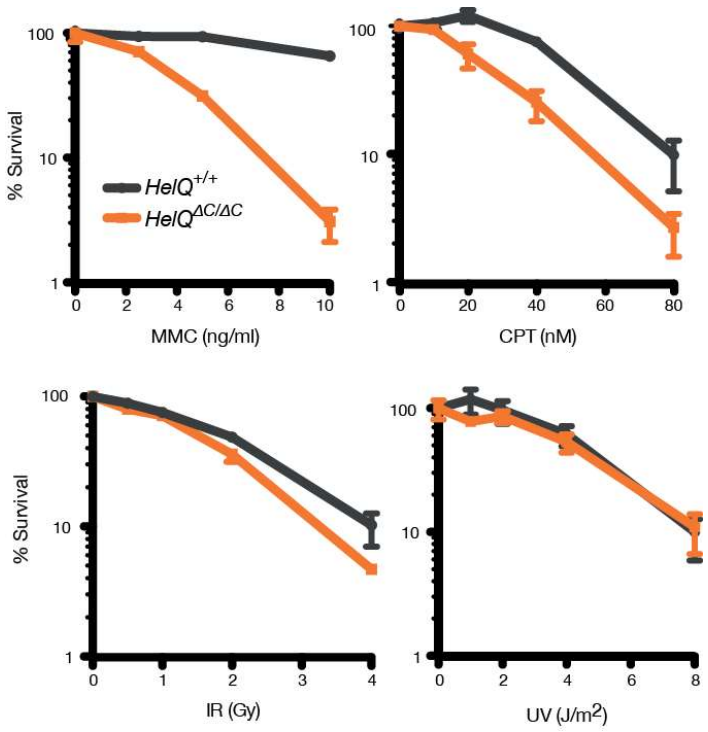


Figure 4-9 HelQ$^{\Delta C/\Delta C}$ cells sensitivity towards DNA damage. (C. Adelman)

Clonogenic survival assays of immortalized HelQ$^{\Delta C/\Delta C}$ cells treated with various damaging factors.

4.2.3 HelQ is dispensable for Intra-S-phase checkpoint activation.

To examine the underlying cause of HelQ$^{\Delta C/\Delta C}$ mutant cells sensitivity to DNA crosslinking agents, we examined integrity of checkpoint activation and recruitment of downstream repair proteins to damaged DNA. Figure 4-10 shows results of the checkpoint indices analysis. It is clear that HelQ is dispensable for checkpoint activation as measured by phosphorylation of ATM, CHK1, CHK2, and γH2AX proteins (phosphorylation of these factors was induced to comparable levels in response to MMC/CPT in both *HelQ*$^{\Delta C/\Delta C}$ and the control *HelQ*$^{+/+}$ cells). Similar results were obtained for human U2OS cells with siRNA driven HELQ knock down Figure 4-11.

Having determined that HELQ is dispensable for intra-S phase checkpoint activation we wanted to examine the status of downstream DNA repair processes such as the Fanconi anemia pathway and homologous recombination.

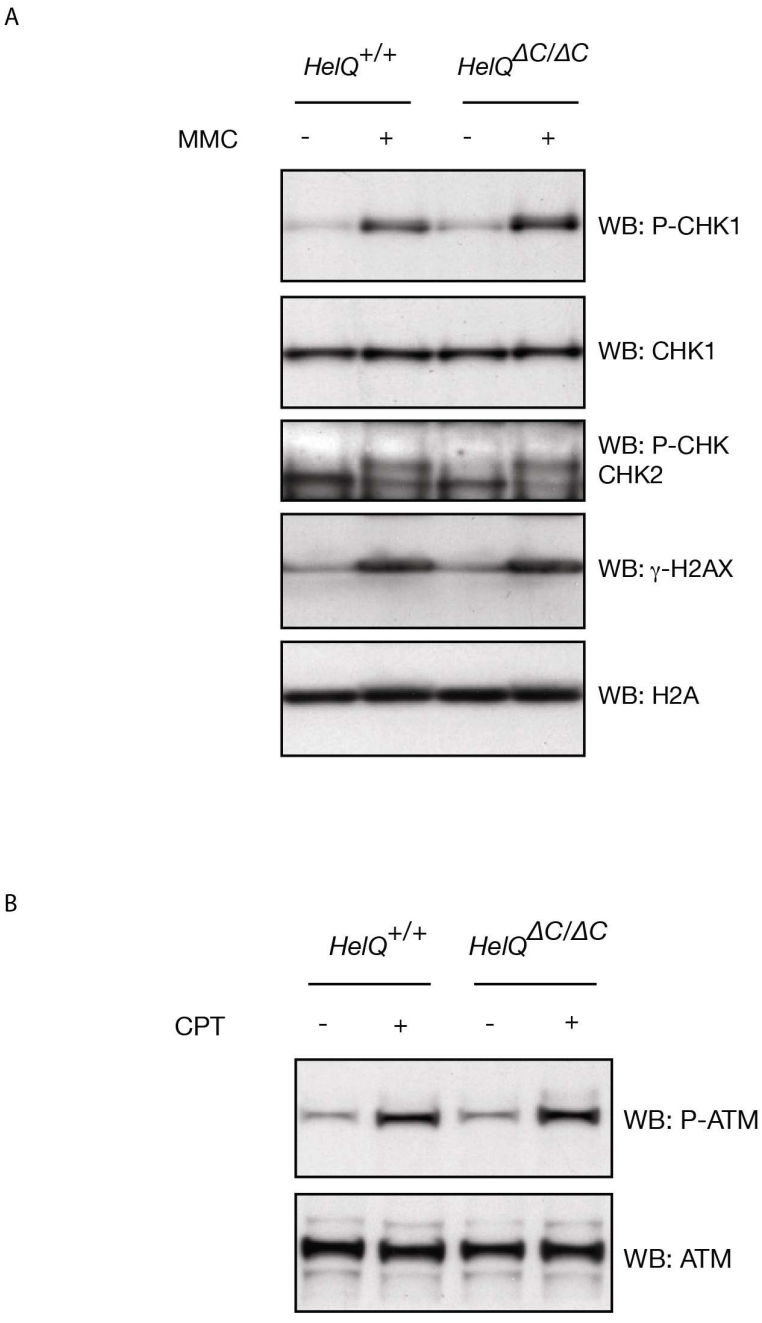


Figure 4-10 Checkpoint activation status in mouse HelQ$^{\Delta C/\Delta C}$ cells.

HelQ$^{\Delta C/\Delta C}$ and HelQ wt (HelQ$^{+/+}$) cells were treated with or without 500ng/ml MMC **(A)** or 50nM CPT **(B)**, lysed and probed for checkpoint activation markers. HelQ status did not affect checkpoint activation in cells treated with both compounds (C. Adelman)

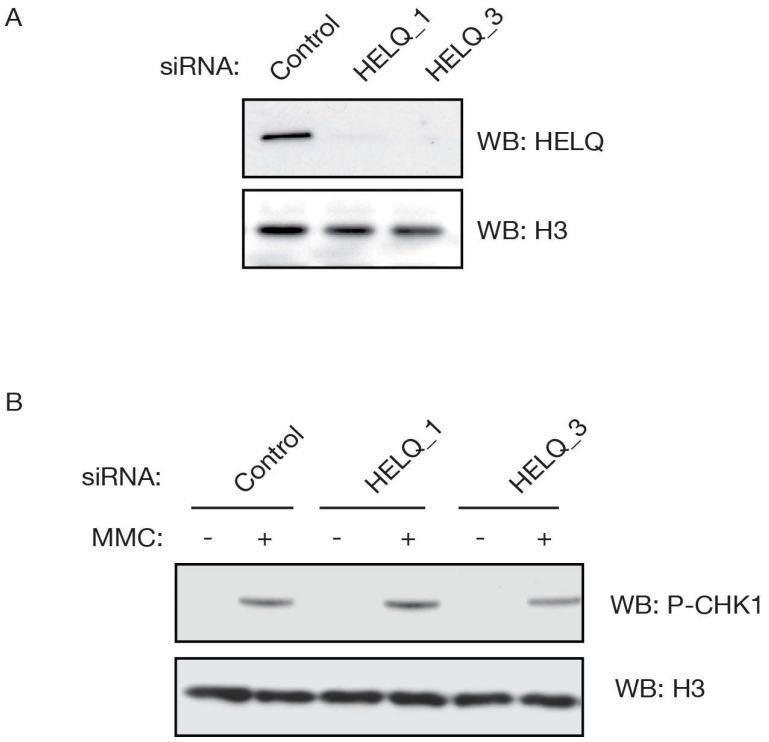


Figure 4-11 Checkpoint activation status in USOS cells depleted of HELQ.

(A) U2OS cells were subjected to three rounds of reverse HELQ siRNA transfection with two independent sequences (HELQ_1 and HELQ_3). **(B**) HELQ depleted U2OS cells were treated with MMC (1μM for 24h), lysed and probed for checkpoint activation indices. HELQ is dispensable for CHK1 phosphorylation

4.2.4 HelQ is dispensable for FANCD2-FANCI ubiquitination.

One of the most frequently used readouts of the Fanconi pathway activation is mono-ubiquitination of the FANCD2 protein in response to DNA lesions that block replication fork progression. We were interested to know whether HelQ is required for this modification process to take place. To assess this possibility, we treated cells with aphidicolin a potent DNA polymerase alpha inhibitor that blocks eukaryotic replication and activates the Fanconi Anemia pathway response. As shown in Figure 4-12, aphidicolin treatment triggered FANCD2 monoubiquitination in both HelQ deficient and wild type cells to comparable levels. This data suggests that HelQ is dispensable for the Fanconi anemia pathway activation and most likely plays a role in the repair process downstream of or in parallel with this step.

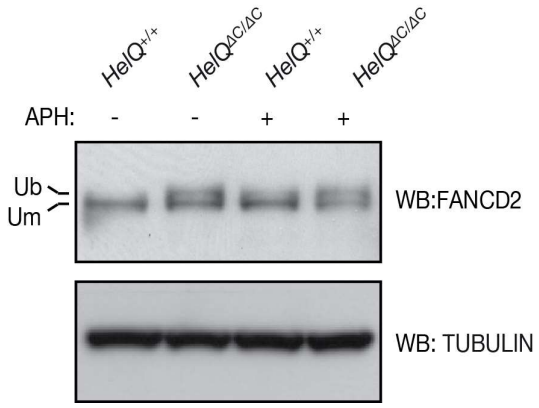


Figure 4-12 Fanconi anemia pathway activation in the absence of HelQ.

HelQ deficient and wt cells were grown under physiological oxygen conditions (5% O_2) and subjected to 3mM aphidicolin (APH) treatment for 6 hours. Cells were then lysed and probed for FANCD2 (Um – unmodified; Ub – ubiquitinated)

Interestingly, we also observed similar levels of ubiquitinated FANCD2 in untreated HelQ$^{\Delta C/\Delta C}$ cells indicating that the Fanconi anemia pathway is constantly active in the absence of HelQ. It seems that in cells lacking HelQ activity, endogenous DNA damage increased or not repaired efficiently, causing constitutive FA pathway activation.

4.3 HELQ is involved in the Homologous Recombination.

Activation of the FANCD2-FANCI heterodimer promotes downstream repair and chromatin recruitment of the Homologous Recombination machinery. A key step of the HR process is BRCA2 dependent RAD51 loading on the ssDNA followed by the homology search. Since I have shown that HELQ directly interact with the BCDX2 complex of the RAD51 paralogs and given the constitutive activation of the FA pathway in HelQ$^{\Delta C/\Delta C}$ cells, we suspected that HELQ may function at the HR stage of repair, downstream of the FANCD2 mono-ubiquitination step.

To test that idea, we decided to assess RAD51 recombinase chromatin loading upon MMC treatment in the HelQ deficient background. In human U2OS cells depleted for HELQ with two different siRNAs we observed equal levels of chromatin RAD51 upon MMC treatment as in control cells (Figure 4.13).

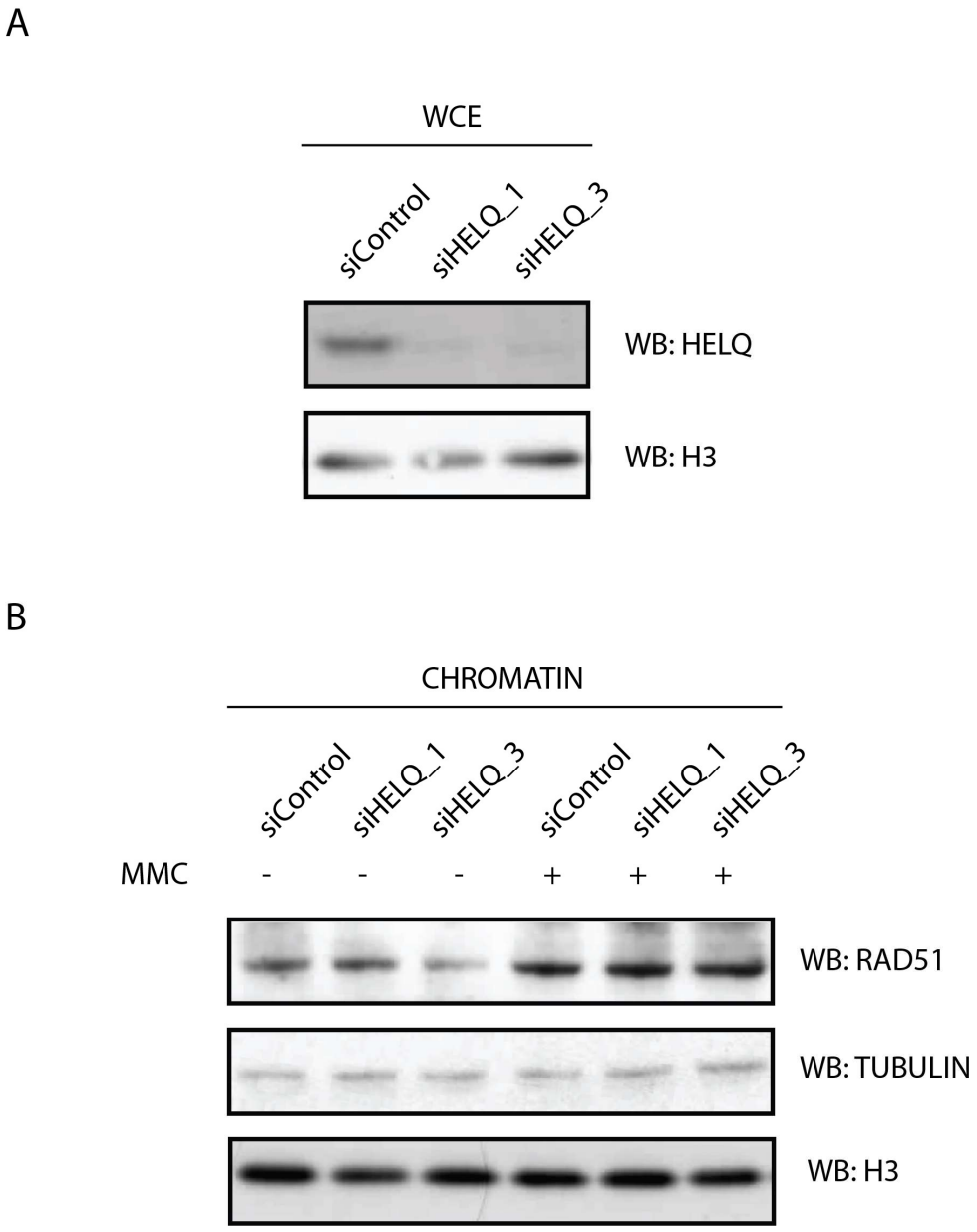


Figure 4-13 MMC induced RAD51 chromatin recruitment in HELQ depleted cells.
U2OS cells were subjected to three rounds of reversed siRNA transfection **(A)** and treated with 1μM MMC for 24 hours. Cell were harvested, chromatin fraction was isolated and probed for RAD51, histone H3 (loading control) and Tubulin (cytoplasm fraction contamination marker) **(B)**.

In addition to cell fractionation, we also assessed damage induced RAD51 and RPA32 chromatin loading in HELQ depleted cells by IF studies. Due to technical

difficulties and limited access to raw IF data at the time of revision, presented IF data lacks scale bars. Consistently with fractionation results, we showed that HELQ depletion in human U2OS cells did not affect formation of MMC induced RAD51 and RPA32 foci (Figure 4-14 and Figure 4-15 respectively) providing further evidence indicating that HELQ is dispensable for early steps of HR.

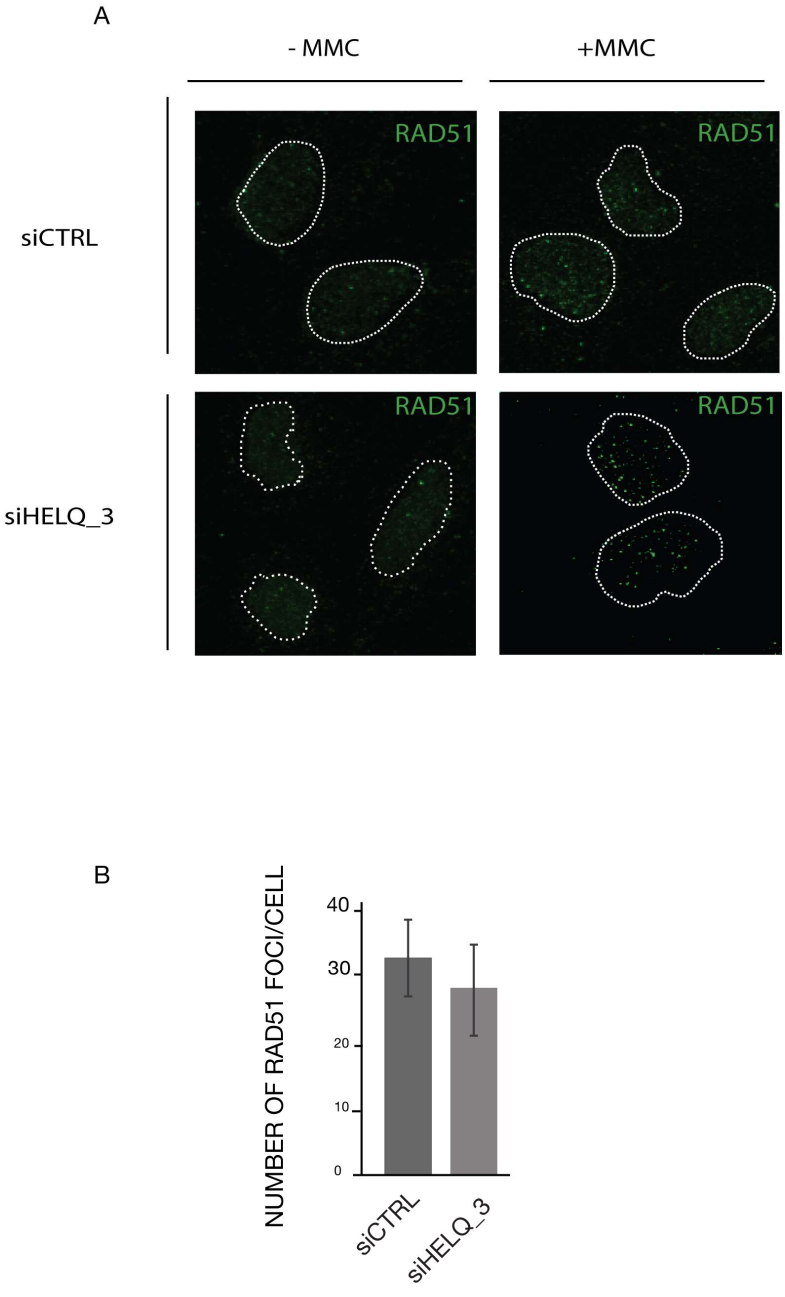


Figure 4-14 MMC induced RAD51 foci formation in HELQ depleted U2OS cells.

(A) U2OS cells depleted for HELQ with HELQ_3 siRNA from Figure 4-13 were treated with +/- 1μM MMC for 24 hours, fixed and stained with RAD51 antibodies. (B) Quantification of RAD51 foci numbers. Error bars represent s.e.m.

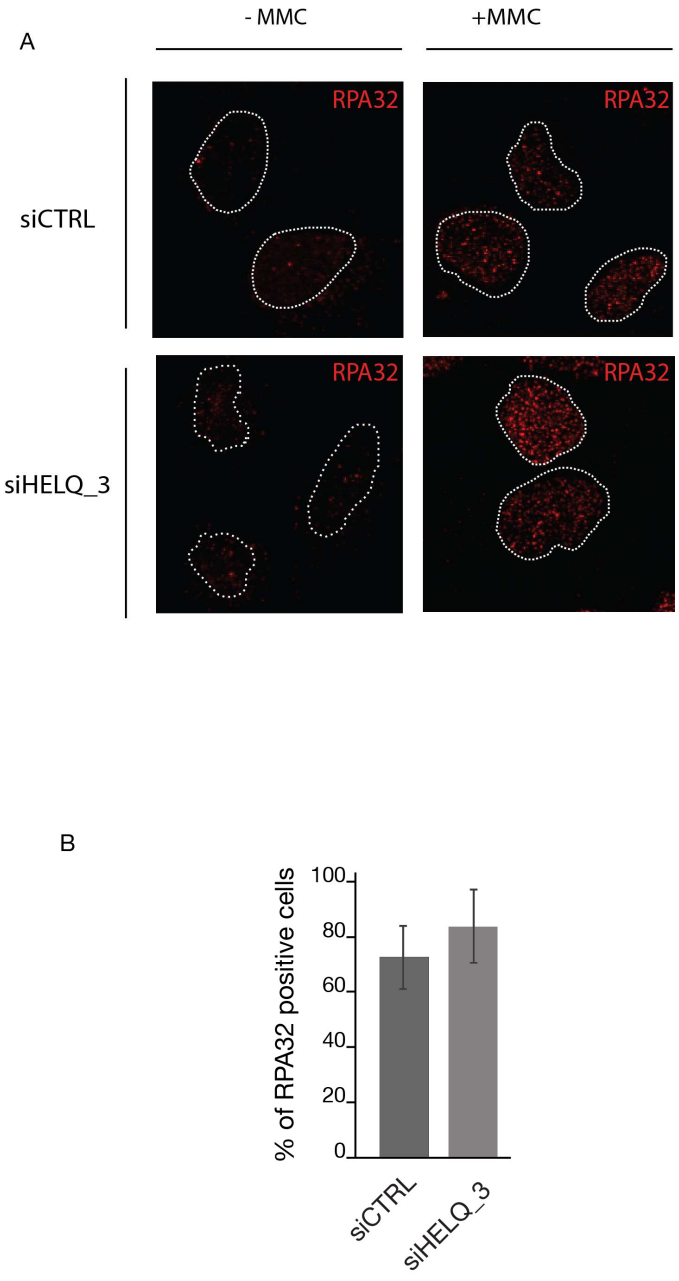


Figure 4-15 MMC induced RPA32 foci formation in HELQ depleted U2OS cells.

(A) Cells were treated as in Figure 4-13 fixed and stained for RPA32 protein. **(B)** quantification of RPA32 foci. Error bars represent s.e.m.

As our results indicate that HELQ depletion does not affect chromatin RAD51 loading and HR repair is initiated as normally, we were interested to know whether it may play a role in the later steps of the process. To assess HR progression in the HELQ deficient background we developed two independent approaches.

First, we performed time point studies of MMC induced DNA damage repair by HR as measured by RAD51 and γH2AX chromatin dynamics. Additionally, we used U2OS cell line-based DR-GFP assay that allows comparing HR efficiency in different siRNA depleted backgrounds. As discussed in Chapter 1, in normal conditions when DNA repair mechanisms are intact, MMC induced ICLs are predominantly sensed and repaired during S phase of the cell cycle when replication forks collide with the lesions. Replication fork stalling triggers the repair machinery loading followed by processing of the damaged DNA. Progression of the HR repair can be monitored by γH2AX and RAD51 chromatin recruitment over time.

To study HELQ role in HR repair I developed a procedure in which cells were exposed to MMC for 24 hours to induce DNA crosslinks. Subsequently, MMC was washed out and fresh media was added to allow cells to repair broken DNA for four consecutive days. I retrieved cell samples every 24 hours during the experiment and then isolated chromatin fraction to compare RAD51 levels by western blot. Complementary to fractionation and western blot analysis I also assessed RAD51 and γH2AX chromatin foci formation by IF.

As shown in Figure 4-16 undamaged, replicating cells had very low basal levels of chromatin RAD51 but 24 hours upon MMC treatment I observed recruitment of the recombinase to damaged DNA, independently of HELQ status. Washing off MMC allowed the wild type cells to repair the damage and RAD51 dropped to physiological levels as soon as 24 hours after drug removal. In contrast, for $HelQ^{\Delta C/\Delta C}$ cells, RAD51 was retained on chromatin as long as 48 hours after washing MMC off and it was only after three days that the protein was no longer enriched on chromatin.

A similar trend was observed in human U2OS cells depleted for HELQ with siRNA. Again, MMC treatment led to RAD51 chromatin recruitment in both samples and

whereas the control cells were able to repair damaged DNA within 24 hours upon drug removal, HELQ depleted U2OS cells retained RAD51 on chromatin until the last time point (48h).

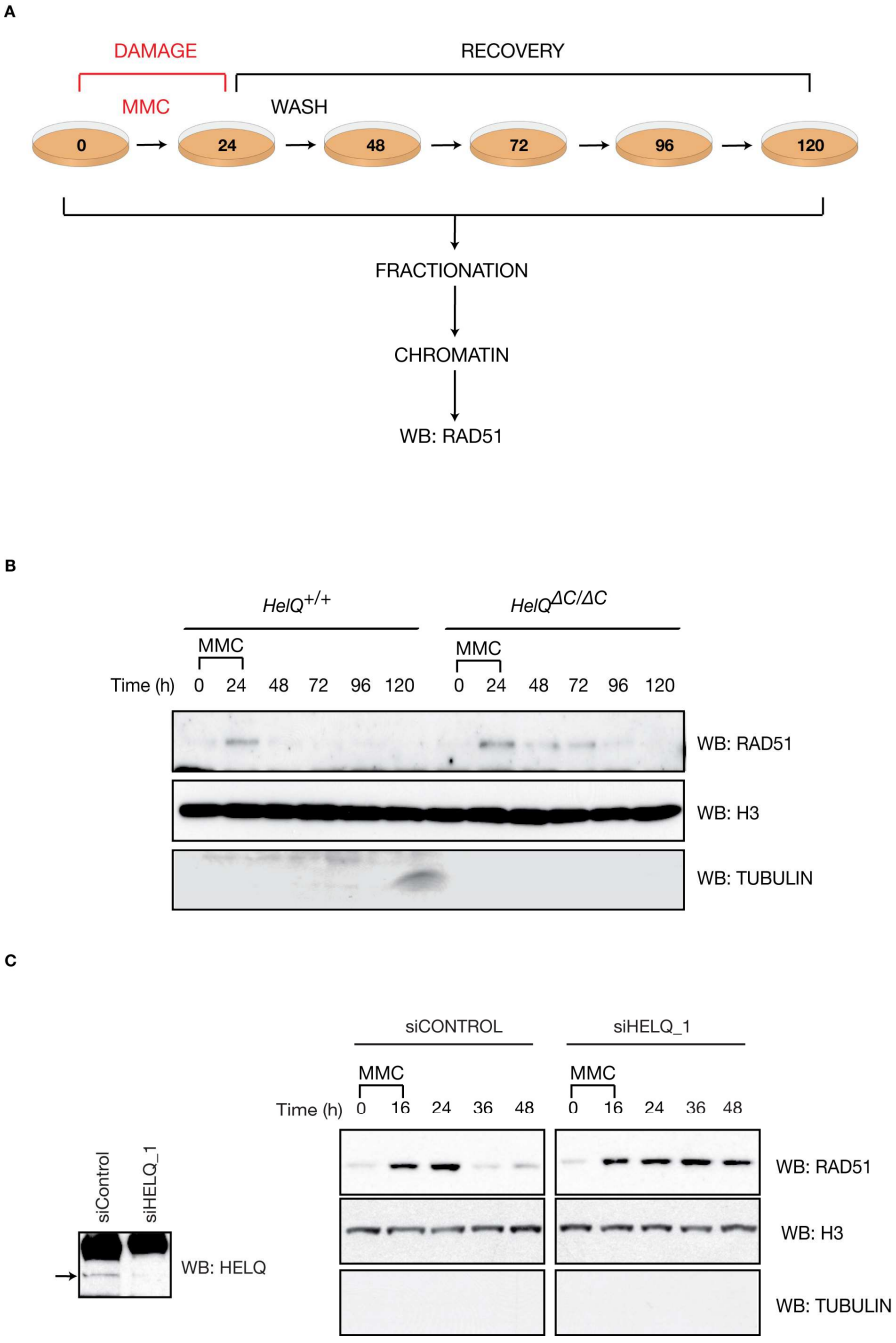


Figure 4-16 HELQ deficiency affects MMC induced RAD51 chromatin dynamics.

(A) RAD51 chromatin dynamics experimental procedure. For siRNA depleted U2OS cells, three rounds of reversed transfection had been performed prior to MMC treatment.

(B) Mouse HelQ$^{\Delta C/\Delta C}$ and wt (HelQ$^{+/+}$) cells were exposed to +/- 100ng ml^{-1} MMC for 24 hours. After washing off the drug, fresh media were added, and cells were grown for 4 days. Every 24 hours cells samples were derived, chromatin fraction was isolated and probed for RAD51, H3 and Tubulin. **(C)** Chromatin fraction of HELQ depleted U2OS cells treated as in B.

Complementary to fractionation and western blot I also assessed formation of chromatin repair foci by IF which allowed more quantitative analysis. As shown in Figure 4-17 and Figure 4-18 for mouse and Figure 4-19 and Figure 4-20 for human cells both RAD51 and γH2AX accumulated in chromatin foci in response to MMC treatment independently of HELQ status. Consistently with the fractionation results I observed dissociation of the RAD51 repair foci in control cells whereas the foci persisted up to 72 hours in the HELQ deficient cells. The same pattern was observed for γH2AX foci.

Collectively these data suggest that in the absence of HELQ, ICL-induced DSB repair is impaired and the defect occurs at a step downstream of RAD51 loading onto broken replication forks.

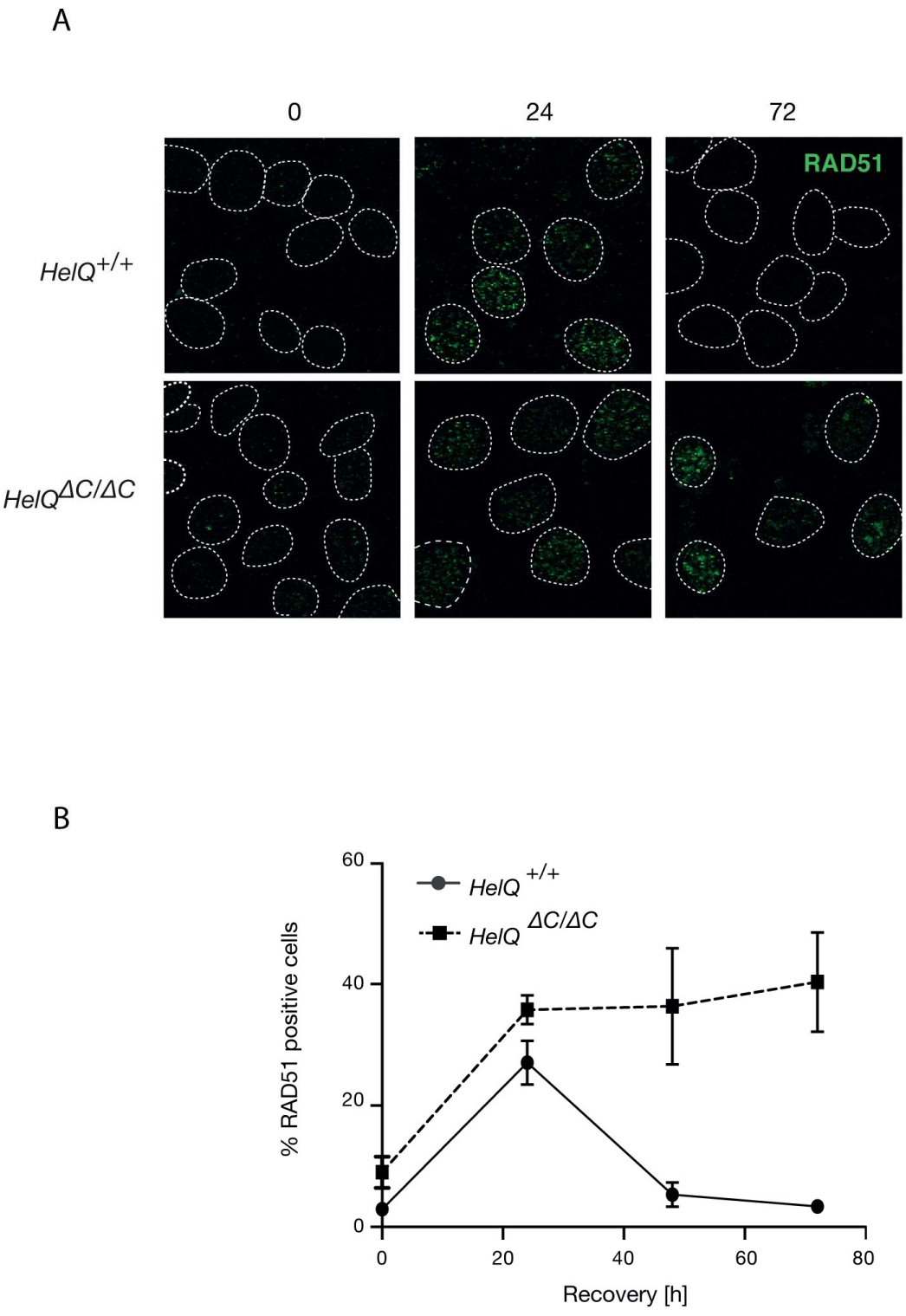


Figure 4-17 MMC induced RAD51 nuclear foci dynamics in mouse HelQ$^{\Delta C/\Delta C}$ cells

(A) HelQ$^{\Delta C/\Delta C}$ cells were treated as in Figure 4-16, instead of fractionation, cells were fixed and stained for RAD51. **(B)** Quantification of MMC induced RAD51 foci over time. Error bars represent s.e.m.

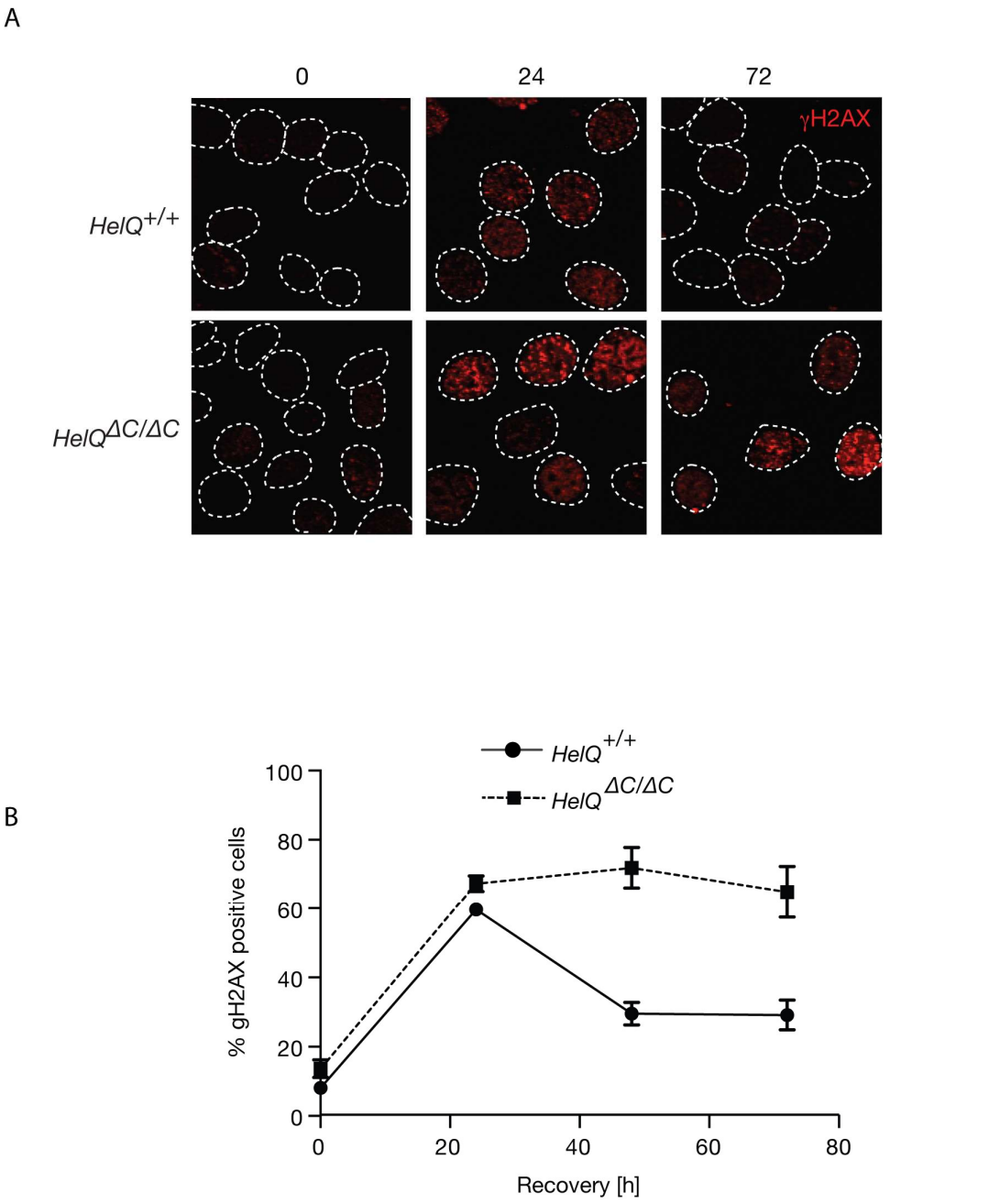


Figure 4-18 MMC induced γH2AX nuclear foci dynamics in mouse HelQ$^{\Delta C/\Delta C}$ cells

(A) HelQ$^{\Delta C/\Delta C}$ cells were treated as in Figure 4-16 instead of fractionation, cells were fixed and stained for γH2AX. **(B)** Quantification of MMC induced γH2AX foci over time. Error bars represent s.e.m.

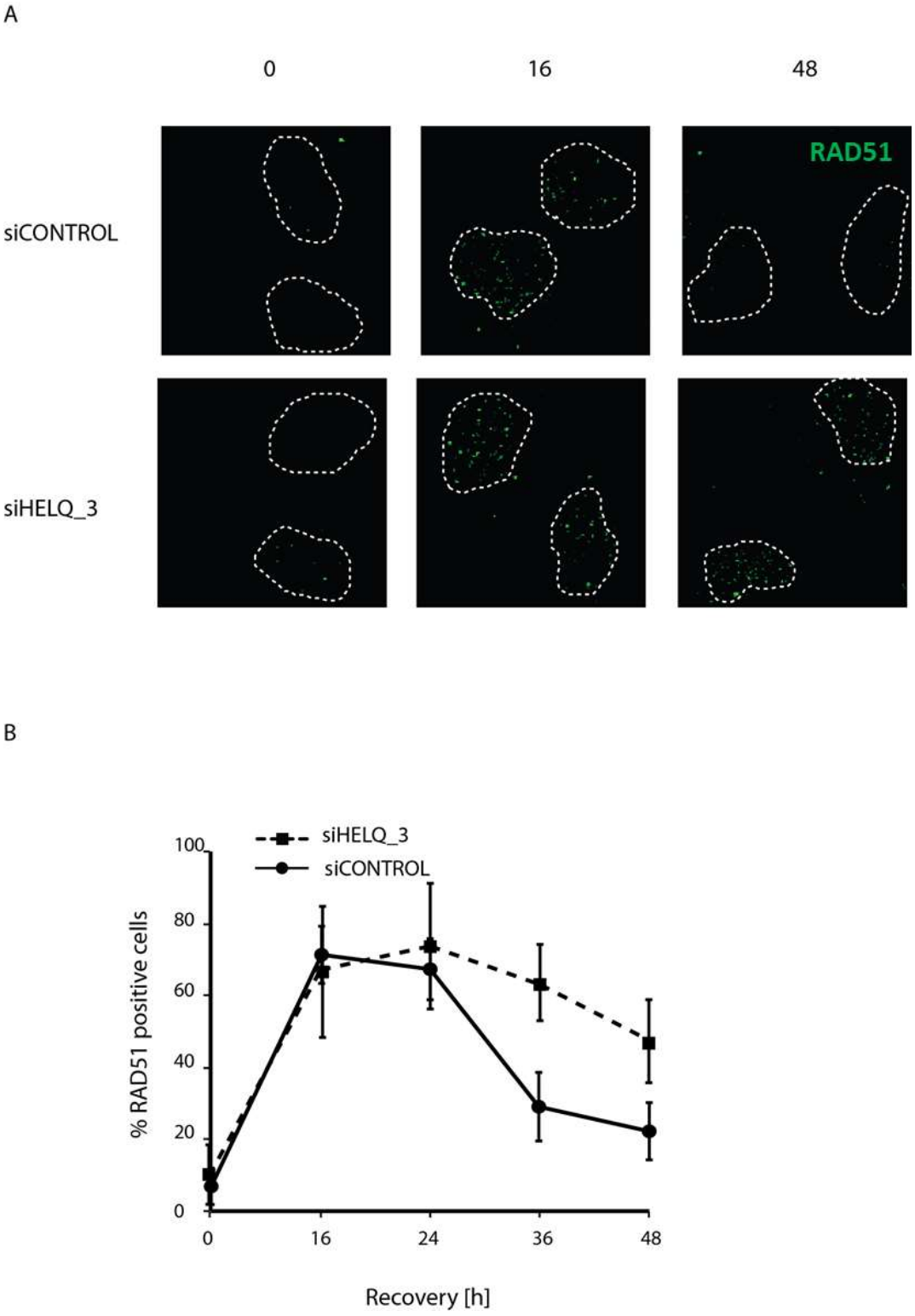


Figure 4-19 MMC induced RAD51 nuclear foci dynamics in HELQ depleted human U2OS cells.(A) Cells were subjected to three rounds of reversed siRNA transfection and treated as in Figure 4-16. **(B)** Quantification of MMC induced RAD51 foci over time. Error bars represent s.e.m.

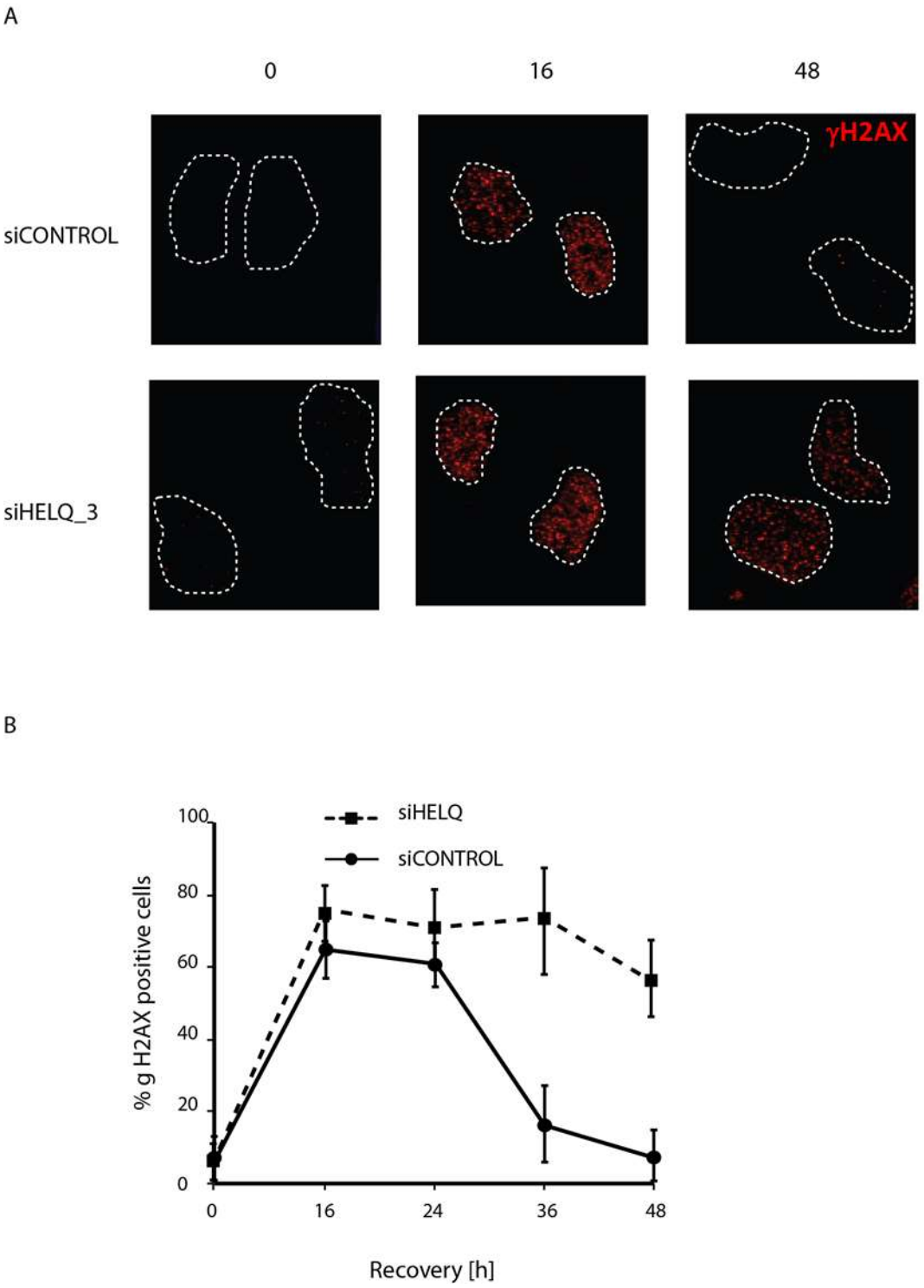


Figure 4-20 MMC induced γH2AX nuclear foci dynamics in HELQ depleted human U2OS cells.

(A) Cells were subjected to three rounds of reversed siRNA transfection and treated as in Figure 4-16. Instead of chromatin fractionation cells were fixed and stained for γH2AX. **(B)** Quantification of MMC induced γH2AX foci over time. Error bars represent s.e.m.

To gain further insights into the role of HELQ in HR we decided to assess efficiency of DSBs repair via HR using fluorescence-based DR-GFP (Direct Repeat – GFP) reporter essay developed in Maria Jasin's group.
In short, this system employs engineered U2OS cell line with integrated one copy of DR-GFP construct allowing formation of a single, localised DSB upon expression of I-SceI endonuclease. HR-based repair of the lesion is then monitored by formation of functional GFP gene. Knocking down genes in the DR-GFP U2OS cell line using specific siRNAs allows assessing whether the gene of interest is involved in DSBs repair as its knock down leads to decrease of fluorescent signal.

To efficiently knock down HELQ we used previously established protocol with three consecutive round of reversed siRNA transfections. 24h after the third round, cells were transformed with plasmid coding I-SceI endonuclease and cultivated for the following 3 days. Subsequently, cells were harvested, and green fluorescence was measured by FACS.

Although pilot experiments looked promising – we observed a drop in GFP-positive cell number in HELQ-depleted background, we realized that this effect was difficult to reproduce (Figure 4-21).
As a positive control we included cells depleted for RAD51 in which we would expect to observe a dramatic decrease of HR efficiency, but this effect fluctuated as well.

There are two crucial steps in the DR-GFP assay - the gene of interest needs to be efficiently knocked down in all cells and transfection with I-SceI should be equally efficient across samples and experiments. If these conditions are suboptimal it can result in considerable variations of the DR-GFP readout. WB analysis showed that we were able to obtain a consistently good knock down of HELQ gene, we therefor decided to focus on DNA transfection step optimization.
To measure efficiency of plasmid delivery across different samples we included a construct coding RFP protein (in 1:10 molar ration) to I-SceI transfection mixture.

This allowed us to assess the population of cells transformed with the DNA and then normalize FACS results to transfection efficiency. As showed in Figure 4-22 our efforts did not resolve the reproducibility problems.

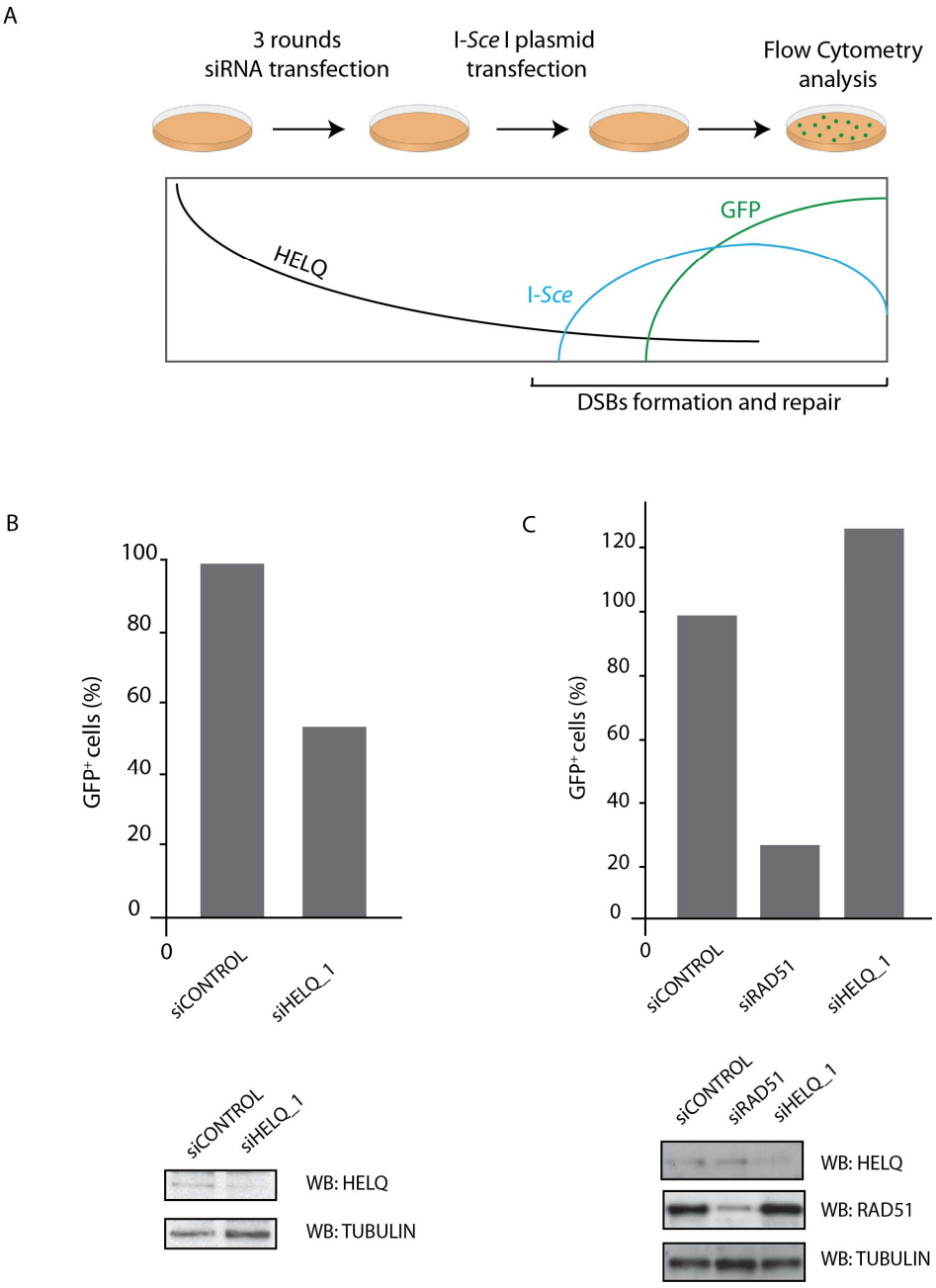


Figure 4-21 Homologous Recombination frequencies in DR-GFP reporter cells depleted for HELQ.

(A) Schematic representation of the DR-GFP reporter assay experiment design: following siRNA HELQ depletion, cells were transfected with I-SceI endonuclease coding

plasmid to induce DSBs formation. Efficient HR repair leads to generation of functional GFP gene that is used as a readout. **(B)** and **(C)** Two independent DR-GFP experiments showing effect of HELQ depletion. In C, a RAD51 depletion as a positive control was introduced.

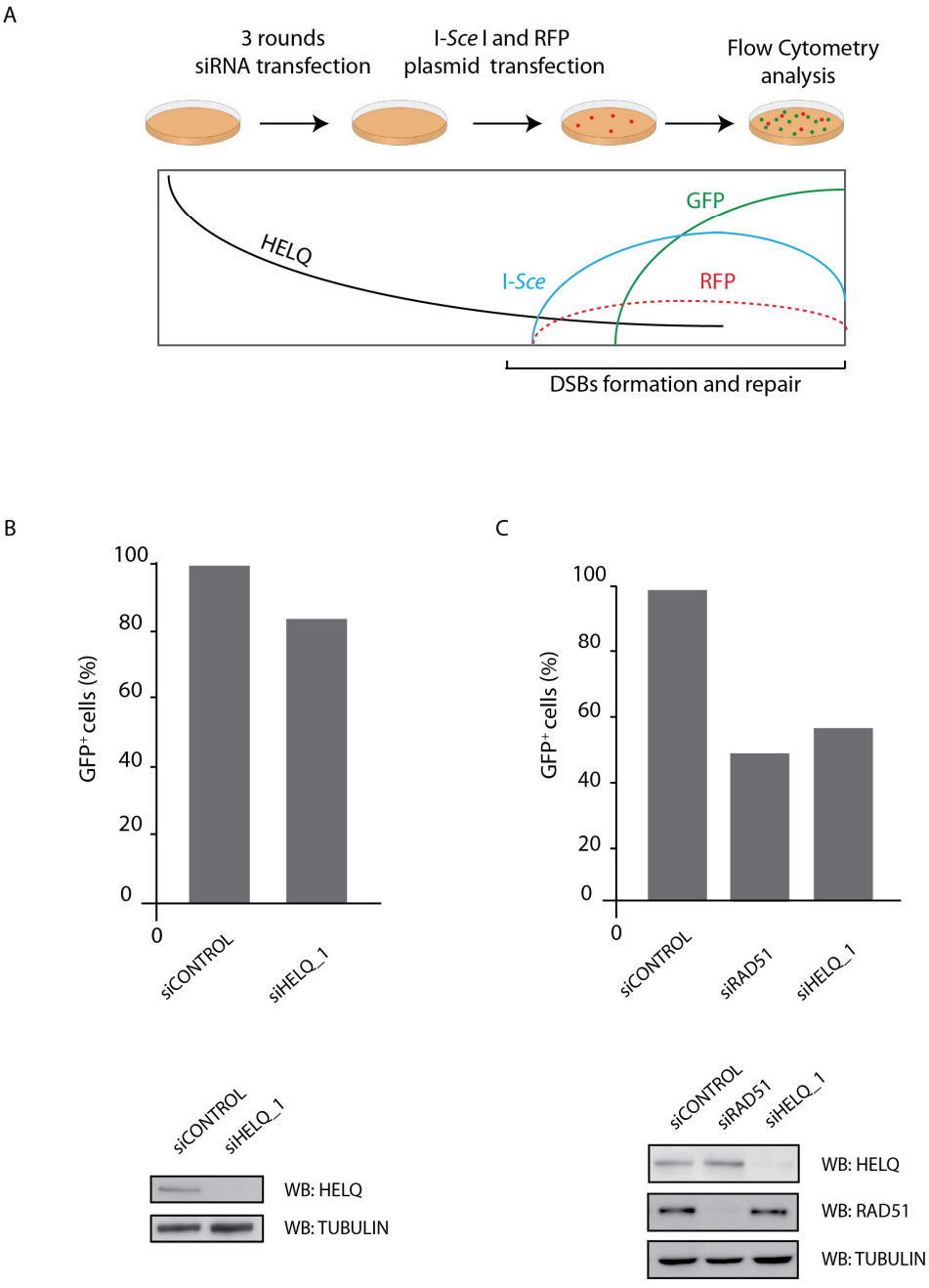


Figure 4-22 Homologous Recombination frequencies in DR-GFP reporter cells depleted for HELQ – transfection efficiency normalization.

(A) Modified DR-GFP assay with transfection efficiency control (RFP plasmid). **(B)** Homologous recombination frequencies in cells treated with indicated siRNAs (as in Figure 4-21) GFP signal was normalized for transfection efficiency (RFP).

To be able to coordinate publication of our research with another group we decided to request this experiment from a laboratory of dr. Alessandro Sartori who specializes in this type of assays.

As shown in Figure 4-23, HELQ depletion with two different siRNAs led to two to three-fold drop in DSBs repair by HR as compared to control cells (siRNA targeting luciferase). Although the decrease was not as pronounced as for *BRCA2* or *RAD51* knock down, these data clearly indicate that HELQ plays a role in promoting Homologous Recombination repair of DSBs.

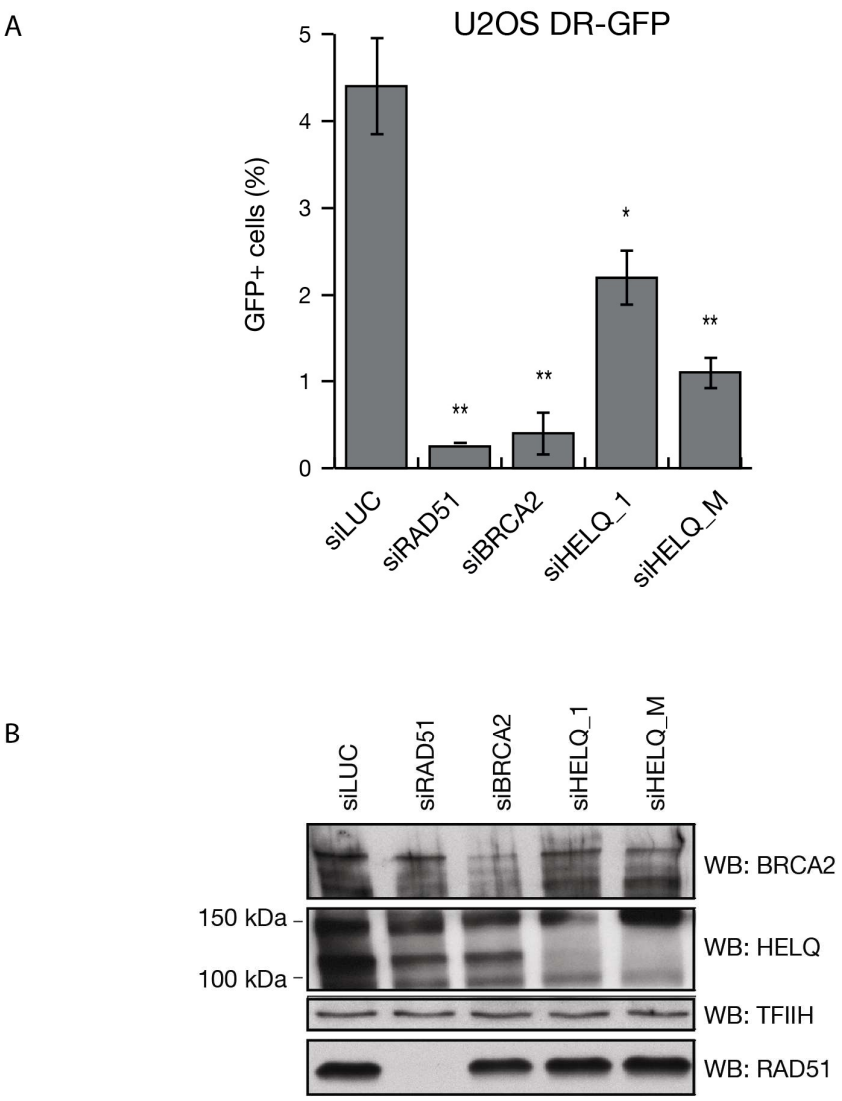


Figure 4-23 Homologous Recombination frequencies in DR-GFP reporter cells depleted for HELQ (Alessandro Sartori).

(A) DR-GFP procedure optimized by A. Sartori's group was used. After initial siRNA transfection, cells were co-transfected with siRNA and I-SceI plasmid. Green signal was analysed 72 hours later, siLUC – luciferase siRNA. **(B)** Wb confirmation of siRNA driven protein depletion.

Given that sensitivity to PARP inhibitors is one of the main characteristics of HR-deficient cells and that some of HELQ interacting partners are sensitive to such agents we decided to test HELQ-deficient cells to Olaparib – a potent inhibitor of PARP1 and PARP2 activity (Bryant et al., 2005; Huehls, Wagner, Huntoon, & Karnitz, 2012; Loveday et al., 2011b, 2012).

Clonogenic assay performed in siRNA depleted U2OS indicated that depletion of HELQ with two different siRNAs led to pronounced Olaparib sensitivity. In fact the survival rate was decreased by more than four-fold, when compared to the control siRNA. Similar results were obtained for HELQ deficient MEFs. This result further confirms HELQ role in promoting Homologous Recombination repair.

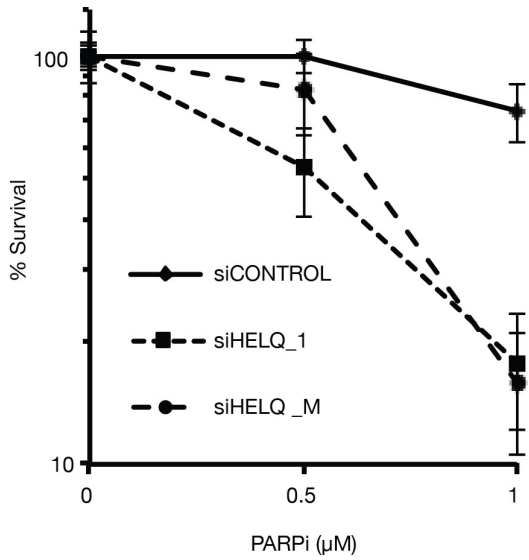


Figure 4-24 HELQ depleted U2OS cells are sensitive to PARP inhibitor, Olaparib.

Cells were subjected to three rounds of reversed siRNA transfection and exposed to indicated concentrations of PARP inhibitor (Olaparib).

4.4 Discussion

Considering the fact that BCDX2 complex as well as FANCD2/FANCI heterodimer and ATR together with RPA70 play an important role in response to DNA replication fork stalling, I decided to characterize HELQ subcellular localization and verify whether it is recruited to chromatin upon DNA damage. Low endogenous expression of HELQ coupled with lack of specific enough antibodies hindered IF studies of HELQ dynamics that I started with. To bypass this problem, I decided to use cell fractionation, a method that allows separation of cytoplasmic, nuclear and chromatin bound proteins hence providing similar insights to IF studies. Apart from localization I also wanted to test whether aphidicolin treatment (leading to replication fork stalling) affects HELQ chromatin recruitment status and whether ATR plays a role in this effect. Although it was impossible to detect endogenous HELQ in chromatin-bound fraction, I successfully used cell line expressing HelQ-GFP at endogenous level. Taken together, localization studies showed that a vast majority of cellular HelQ is localized in the cytoplasm and nucleus and in undamaged cells only a very small fraction of the protein is bound to the chromatin. Upon aphidicolin-induced replication fork stalling, HelQ is recruited to chromatin where it presumably performs its function in the repair process. Importantly, HelQ chromatin enrichment appears to be regulated, at least in part, by ATR kinase activity as this effect is abrogated by inhibition of the ATR kinase signaling.

Complimentary to my previous work showing HELQ is recruited to damaged DNA in ATM-dependent manner I was curious to check whether HELQ can be found on broken replication forks. To examine this possibility, I set up the iPOND system in our laboratory that allows the study of protein recruitment to active, stalled and collapsed replication fork with a high spatio-temporal resolution. In my experiments, I showed that low levels of HELQ can be found on unaffected replication forks and treatment with HU leads to accumulation of the protein. This suggests that HELQ may be engaged with active replication forks and upon treatment with fork stalling agents it is enriched to perform its repair function and promote HR.

I also showed that HelQ is dispensable for Intra-S-Phase activation as indicated by unaffected MMC-induced phosphorylation of Chk1, Chk2 and gH2AX in HelQ depleted cells human and mouse cells. Furthermore, depletion of HelQ does not affect Fanconi Anemia pathway activation as aphidicolin-induced replication fork stalling leads to wild-type levels of monoubiquitinated FANDC2. This data suggests that HELQ performs its repair function downstream of or in parallel to FANCD2 monoubiquitination.

Assessment of RAD51 recruitment to damaged replication forks revealed that RAD51 is enriched on chromatin in response to MMC treatment in HELQ-deficient mouse and human cells. However, RAD51 and γH2AX persisted on chromatin and remained present in repair foci at later time points in HELQ-deficient mouse and human cells, pointing towards a defect in the absence of HELQ occurring at a step downstream of RAD51 recruitment to damaged replication forks. This was shown by both, cell fractionation and western blotting of chromatin bound-fraction and IF studies.

To gain further insights into HELQ's role in HR we assessed efficiency of DSBs repair via HR using fluorescence-based DR-GFP reporter system. We showed that HELQ depletion with two different siRNAs leads to two-to-three-fold drop in DSBs repair by HR as compared to control cells. Although the decrease was not as pronounced as for BRCA2 or RAD51 knock down, these data clearly indicates that HELQ plays a role in promoting Homologous Recombination repair of DSBs.

Lastly to confirm HELQ role is HR we showed that cells lacking HELQ are hypersensitive to Olaparib - potent PARP1 and PARP2 inhibitor. Sensitivity to PARP inhibitors is a hallmark of HR-directed DSBs repair problems (Farmer et al., 2005)

Chapter 5. HELQ-BCDX2 interaction mapping.

As presented in Chapter 3, HELQ binds to the BCDX2 complex of RAD51 paralogs. The interaction is most likely direct and independent of DNA damage status. Despite their pivotal role in HR, RAD51 paralogs remain largely elusive group of proteins, I therefore decided to further elucidate the mutual interplay between HELQ and components of the BCDX2 complex. It would be interesting to identify amino acid sequence within HELQ and/or BCDX2 necessary for binding and in the next step to check whether mutation of the identified sequence gives rise to any particular phenotype in cell lines and potentially mouse models.

5.1 HELQ and BCDX2 stability studies.

Given the constitutive character of HELQ-BCDX2 interaction I wanted to check whether removing any of the binding partners affects stability of the other components of the complex. Additionally, this would be a particularly important information when interpreting effects of HELQ deficiency in cells.

To deplete the proteins of interest in U2OS cells I used four different siRNAs targeting HELQ and one siRNA sequence targeting each of the BCDX2 components: RAD51C, RAD51D, XRCC2 as well as XRCC3 (part of CX3 complex).

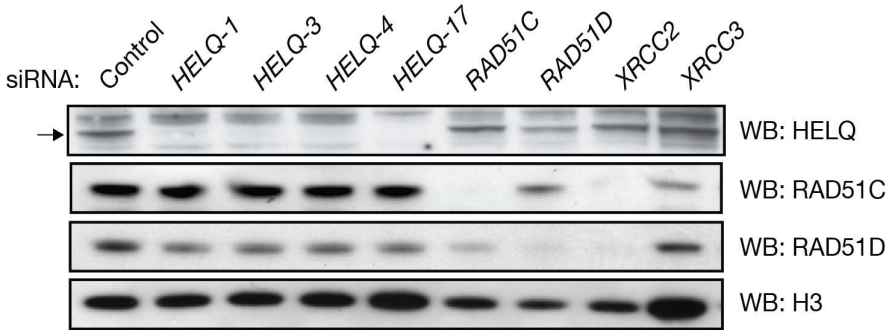


Figure 5-1 HELQ and BCDX2 stability in U2OS cells.

U2OS cells were subjected to three rounds of siRNA transfection targeting indicated proteins. 48h post final transfection equal numbers of cells were lysed and whole cell extracts were assessed by western blot to determine relevant protein levels.

Figure 5-1 presents comparison of relevant proteins levels in siRNA depleted U2OS cells. HELQ and RAD51 paralogs were efficiently depleted using the previously established protocols. Although loading was not perfect in some of the lines (as indicated by histone H3 levels), a few interesting observations can be made based on this experiment.

Consistent with previous observations (Chun, Buechelmaier, & Powell, 2013; Lio, Schild, Brenneman, Redpath, & Chen, 2004), knocking down any of the BCDX2 complex proteins negatively affects stability of the other components (for example, removing XRCC2 protein causes RAD51C and RAD51D protein to drop to undetectable levels). Although I observed no obvious effect on RAD51 paralogs stability in cells depleted for HELQ, one might argue that RAD51D levels seemed to be mildly lower in cells transfected with all four siRNAs targeted for HELQ. Consistently with the *in-vitro* interaction studies, we observed no effect of XRCC3 depletion on HELQ levels.

5.2 HELQ and BCDX2 interaction site analysis

Having established that HELQ-BCDX2 complex interaction is most likely direct, it was important to determine the binding site involved in this interaction. In other words, I wanted to find a region of HELQ that mediates binding to BCDX2 and *vice-versa* - to determine which of the four RAD51 paralogs is/are responsible for HELQ binding. The most straightforward approach to determine which of the BCDX2 complex components is responsible for interaction with HELQ would be to perform a series of siRNA mediated knock downs and assess HELQ binding by IP and western blot. Unfortunately, as shown in chapter 5.1 depletion of any of the BCDX2 components destabilizes the whole complex and leads to increased degradation, making the interaction studies in such system impossible.

5.2.1 Interaction mapping - peptide arrays.

To circumvent technical difficulties encountered in mapping the interaction sites, I decided to utilize peptide array technology. In this method a series of partially

overlapping peptides covering the entire protein sequence is synthetized and covalently immobilized on a membrane. Incubation with a known binding partner followed by immuno- or radio-detection allows to determine potential interaction sites within the peptide sequence. This method has been successfully used in a number of recent interaction studies in our laboratory.

The arrays were constructed in such a way that each of the proteins of interest was divided into 20-mer peptide sequences immobilized on a membrane and each spot on the membrane had 19 residues in common with the adjacent spot (each position was shifted by one residue relative to its neighbour spots). The membranes were made by the LRI Peptide Synthesis Facility.

The strategy to identify BCDX2 complex component(s) that interact with HELQ was to purify HELQ-FLAG from HEK293 (M2 beads IP followed by FLAG peptide elution) and to incubate the eluate with RAD51 paralogs arrays (as schematically show in in Figure 5-2). In reciprocal approach I decided to make use of the recombinant HIS-tagged BCDX2 complex from Sf9 cells, similarly as in *in vitro* binding experiment (paragraph 3.2.2.) This protein prep was shown to reproducibly bind to recombinant HELQ in previous assays, given the binding motif that facilitates this interaction does not require complicated folding, this was a promising starting point for peptide array experiment. To control for the specificity of antigen recognition I used identical arrays handled the same way but without primary antibodies in the detection step.

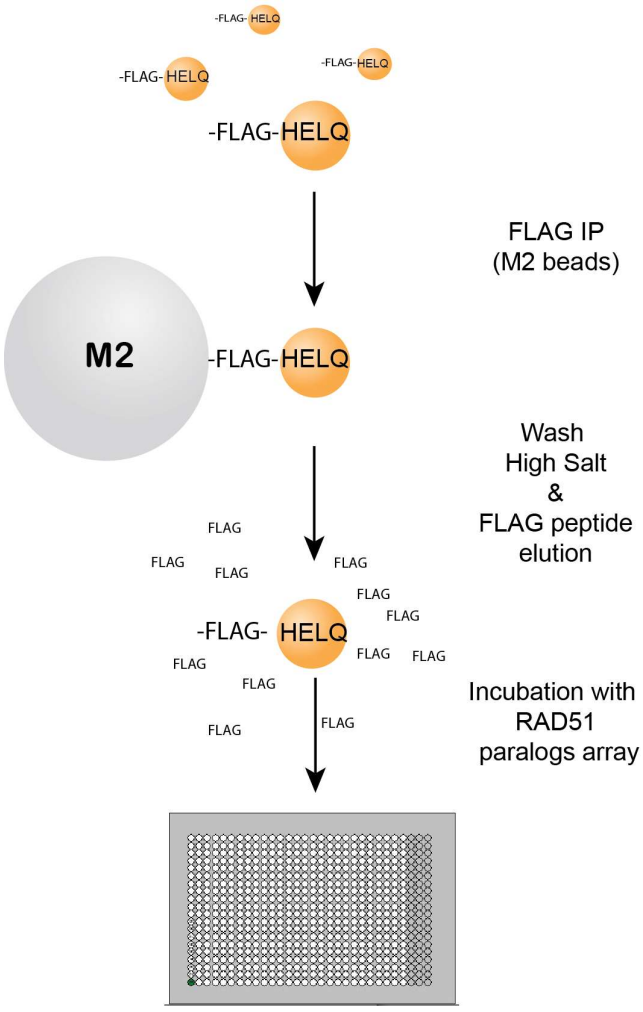


Figure 5-2 HELQ-BCDX2 interaction site mapping by peptide array - schematic.

HELQ-FLAG was purified as described before from HEK 293 cells. To strip HELQ from its interacting partners, beads were thoroughly washed with high salt buffer. Arrays containing of RAD51 paralogs peptide sequence were incubated with eluate containing HELQ-FLAG together with free FLAG peptide used for elution. Binding was assessed by immunoblotting with anti-FLAG antibodies.

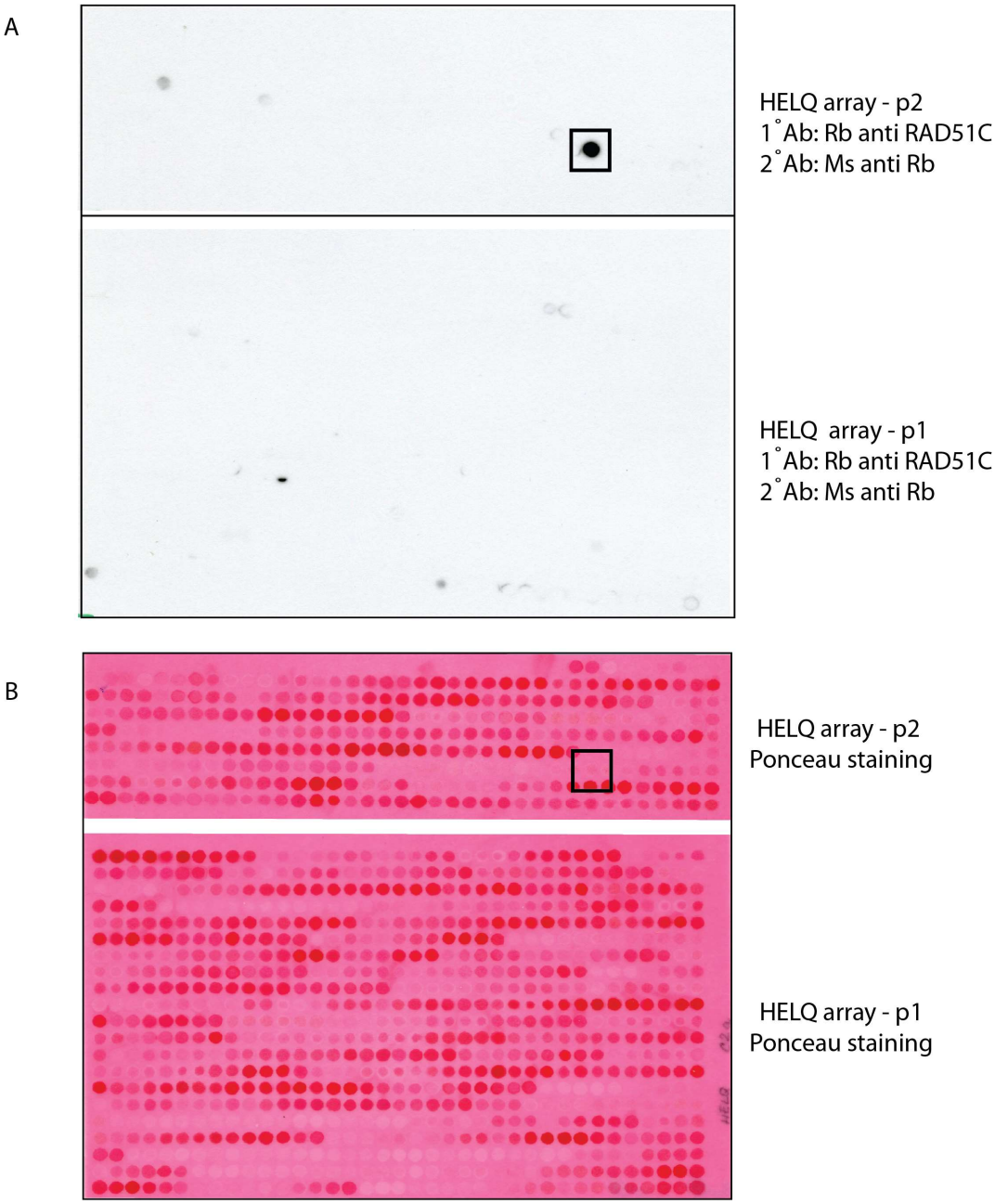


Figure 5-3 Peptide array-based mapping of BCDX2-HIS binding site within HELQ protein – detection with anti RAD51C antibodies.

An array consisting of 20-mer peptides that correspond to full length HELQ protein was created and spread across two membranes (p1, p2). The membranes were then incubated with recombinant, HIS-tagged BCDX2 complex from Sf9 cells. **(A)** Membranes were then thoroughly washed and probed with anti RAD51C antibodies raised in rabbit. To visualise western blot signal, secondary mouse anti rabbit antibodies conjugated to HRP were used. Single spot corresponding to C-terminal stretch: **E-G-L-V-L-E-S-L-L-H-L-I-Y-L-T-T-P-Y-D-L. (B)** The membranes were Ponceau stained to identify peptide position within the array.

5.2.1.1 BCDX2 complex binding site within HELQ studies by peptide arrays.

Because of size of the protein, HELQ peptide array was split into two separate membranes (Figure 5-3). After incubation with the BCDX2 complex and thorough washing steps, the membranes were incubated with anti RAD51C antibodies. As shown in Figure 5-3 a single peptide was able to interact with RAD51C (BCDX2). The identified amino acid stretch (**EGLVLESLLHLIYLTTPYDL)** was located away from any functional domains in the C-terminal region of HELQ protein. What is interesting about this region is the fact that despite relatively low evolutionary conservation of this region, three amino acids: HLI are identical from bacteria to human Figure 5-4 and the fact that it harbours the minimal consensus sequence (T-P) for CDK-dependent phosphorylation (Moreno & Nurse, 1990).
The control membranes showed only a minor background signal spread equally across the entire sequence of the protein.

As the BCDX2 complex components were tagged with HIS-tag, I decided to repeat the experiment but instead of using RAD51C antibodies as a detection tool, I probed the membranes for the HIS-tag. It would be reasonable to expect a similar signal pattern provided the binding is antigen-specific. As a negative control I used a HIS-tagged SFRS1 splicing factor. Both membranes were treated in the same manner.
As shown in Figure 5-5 using antibodies specific to HIS-tag changes the binding signal dramatically. Instead of a single peptide, there are now few peptides located in different regions of the protein with the most promising one mapping to C terminal region, downstream of the peptide identified in the previous experiment. Of note, the single peptide identified with anti RAD51C antibodies is not lit up. The negative control array incubated with SFRS1-HIS did not show any significant binding (Figure 5-6).

The simplest explanation for the differences between these two experiments is that the binding observed is unspecific. Since. two of the four BCDX2 complex components are HIS-tagged (RAD51B and RAD51D), one cannot rule out a situation where there is a subset of RAD51 paralogs molecules in the solution that are unengaged in a complex and flow freely in the incubation buffer. Those particles

could potentially bind to various regions of the membrane, either via site specific interaction or simply driven by hydrophobic forces.

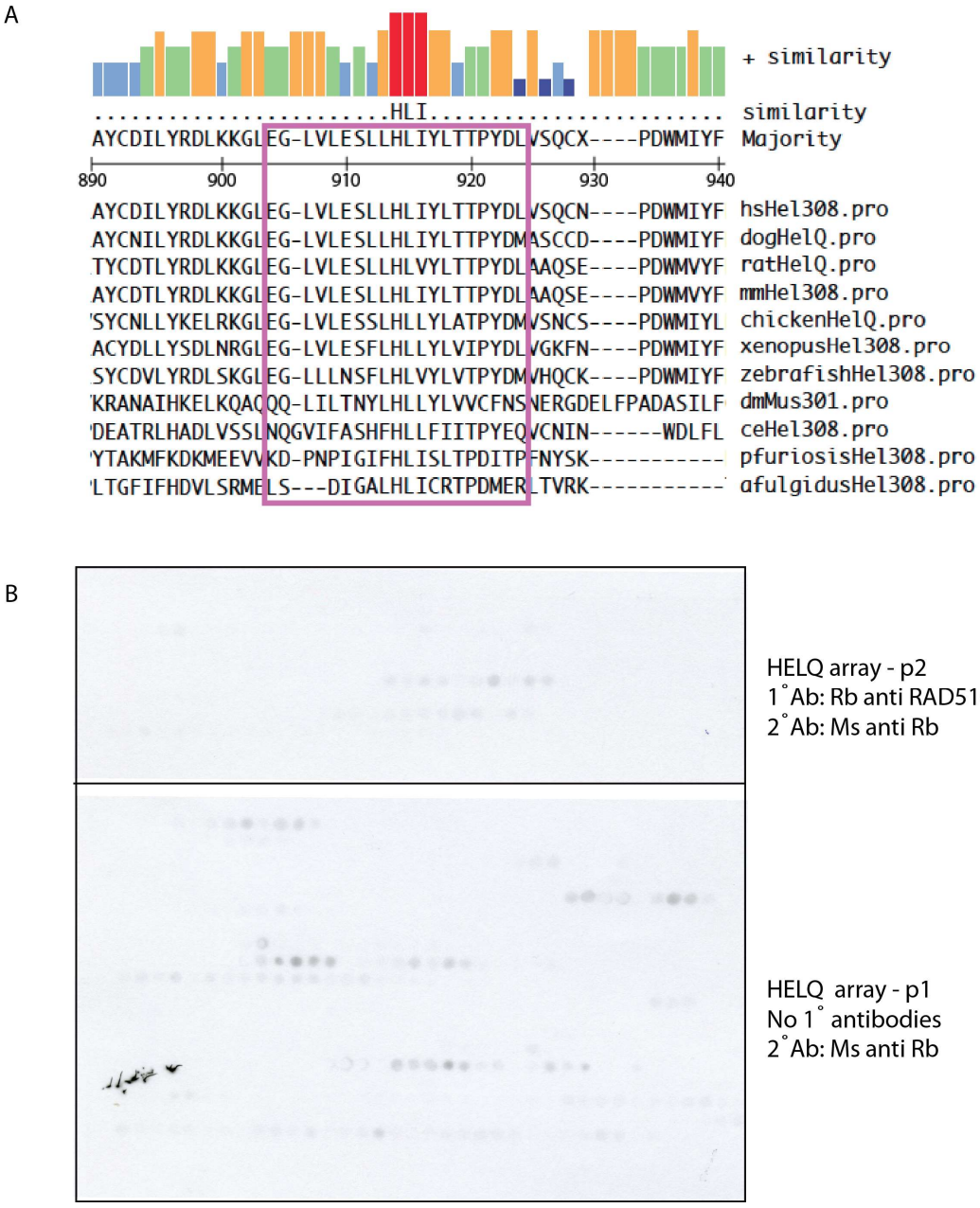


Figure 5-4 Peptide array - based mapping of BCDX2 complex binding site within HELQ.

(A) Identified **E-G-L-V-L-E-S-L-L-H-L-I-Y-L-T-T-P-Y-D-L** peptide stretch localizes in a relatively poorly conserved C-terminal region of HELQ protein. Within the candidate binding site only three consecutive amino acids: HLI are highly conserved across

species. **(B)** Mock membrane controlling specificity of antibody detection was treated as the experimental membrane but without primary antibodies incubation step.

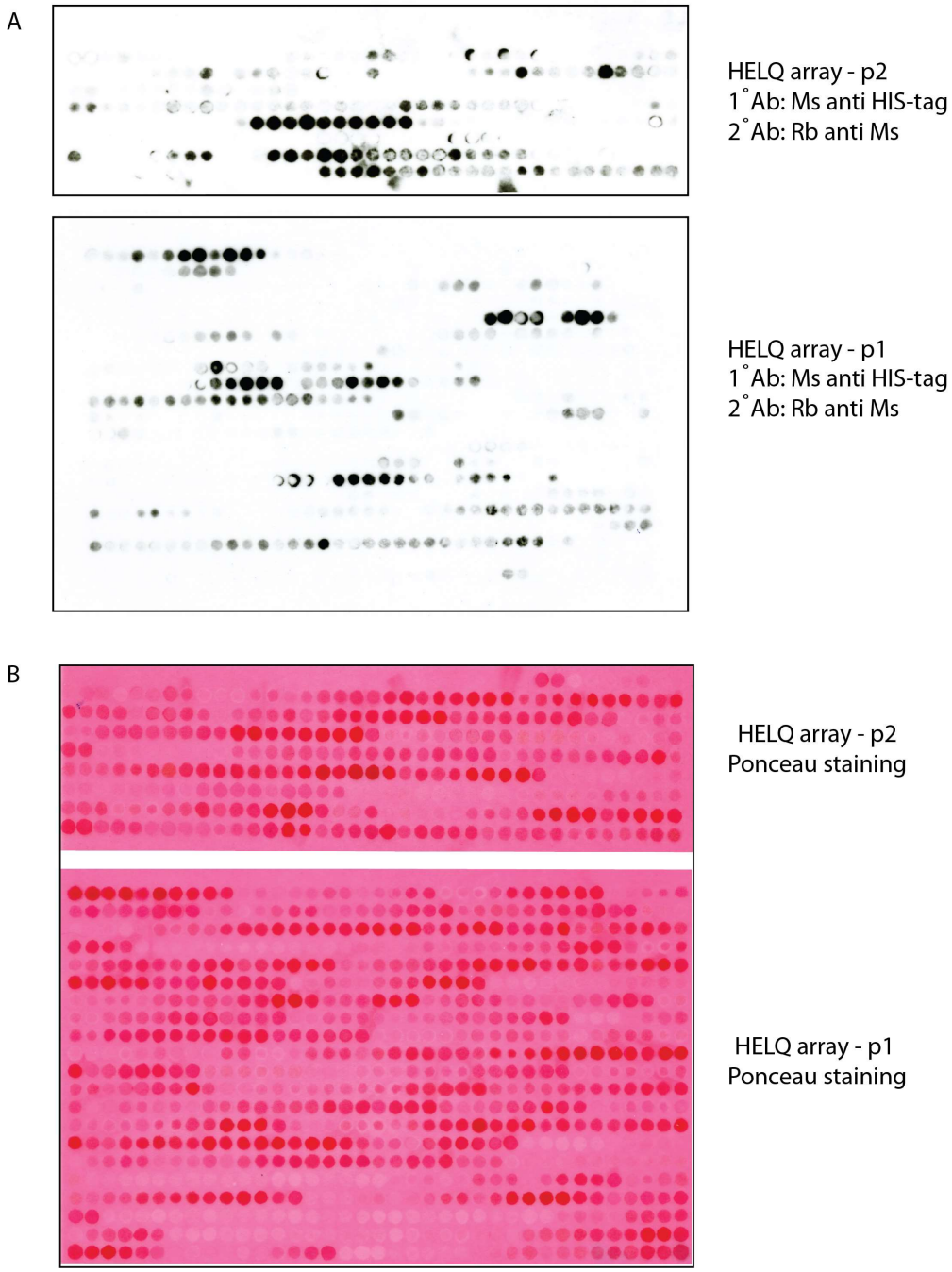


Figure 5-5 Peptide array-based mapping of BCDX2-HIS binding site within HELQ protein – detection with anti HIS antibodies.

(A) Two membranes with peptides corresponding to full length HELQ protein were incubated with recombinant, HIS-tagged BCDX2 complex from Sf9 cells. After a wash step the membranes were probed with anti HIS antibodies raised in mouse. To visualise

western blot signal, secondary mouse anti rabbit antibodies conjugated to HRP were used. **(B)** The membranes were Ponceau stained to determine peptides positions within the array.

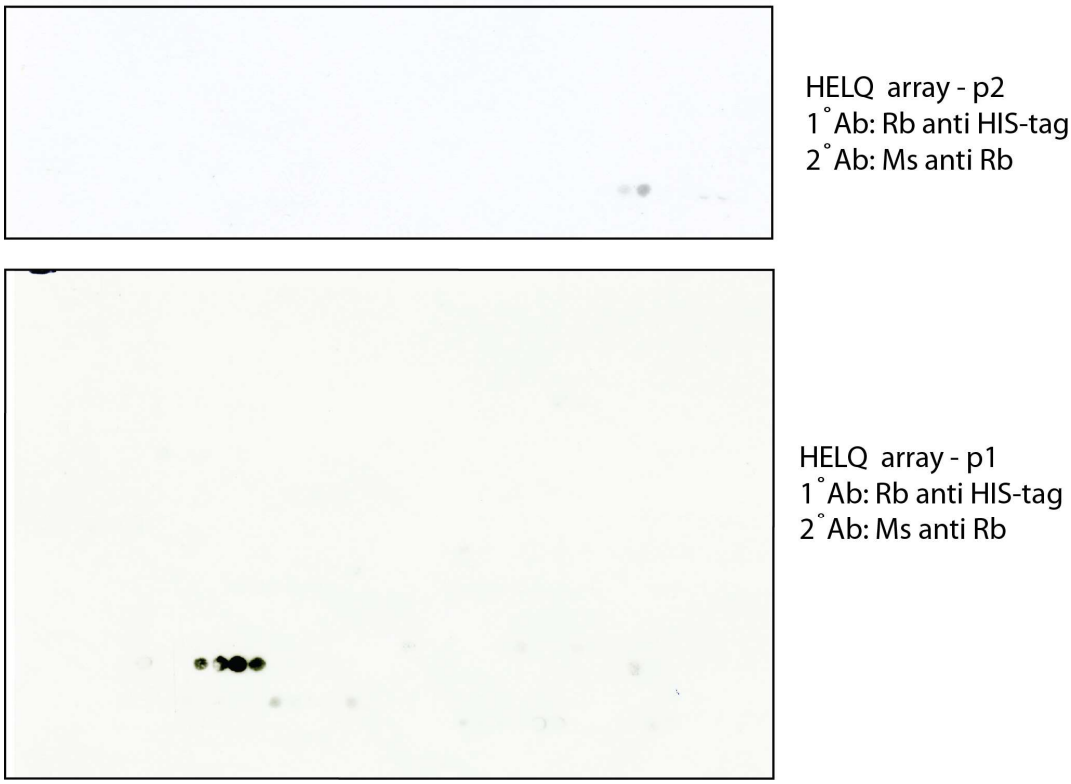


Figure 5-6 Control HLEQ membrane.

The control membrane was treated the same way with one modification – instead of incubation with BCDX2-HIS, a HIS-tagged SFRS1 protein was used as a negative binding control.

5.2.1.2 HELQ binding site within BCDX2 complex studies by peptide arrays

I also tried a reciprocal approach in which RAD51 paralogs peptide libraries were immobilized on membranes and incubated with FLAG-tagged HELQ purified from HEK293 cells (Figure 5-7, Figure 5-8, Figure 5-9).

Although no obvious signal was detected for RAD51B, RAD51C (Figure 5-7 A and B) and XRCC2 (Figure 5-8 B) there was a strong binding stretch in RAD51D (Figure 5-8 A). The spots correspond with G-L-T-A-S-R-L-L-Q-L-L-Q-A-K-T-Q-D-E-E-E peptide sequence located in the middle of RAD51-like ATPase fold between Walker A and Walker B motives. Noteworthy, QT is a potential ATM/ATR phosphorylation site.

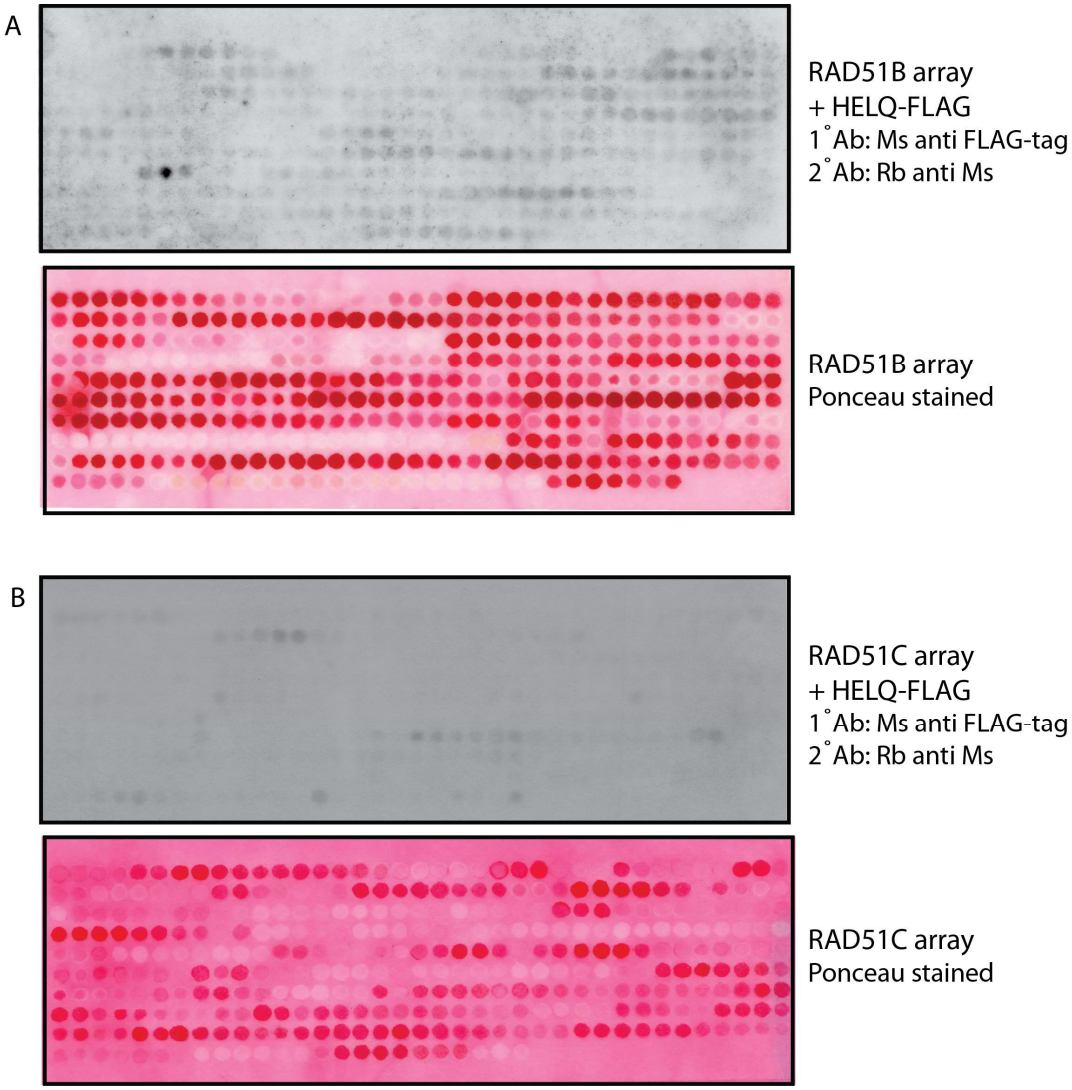


Figure 5-7 Peptide array-based mapping of HELQ-FLAG binding site within BCDX2 complex (RAD51B, RAD51C).

Membranes corresponding to RAD51B **(A)** and RAD51C **(B)** were incubated with HELQ-FLAG purified from HEK293. Antibodies specific for FLAG tag were used for detection. No obvious signal was detected in neither array. To visualize immobilized peptides spots, membranes were stained with Ponceau.

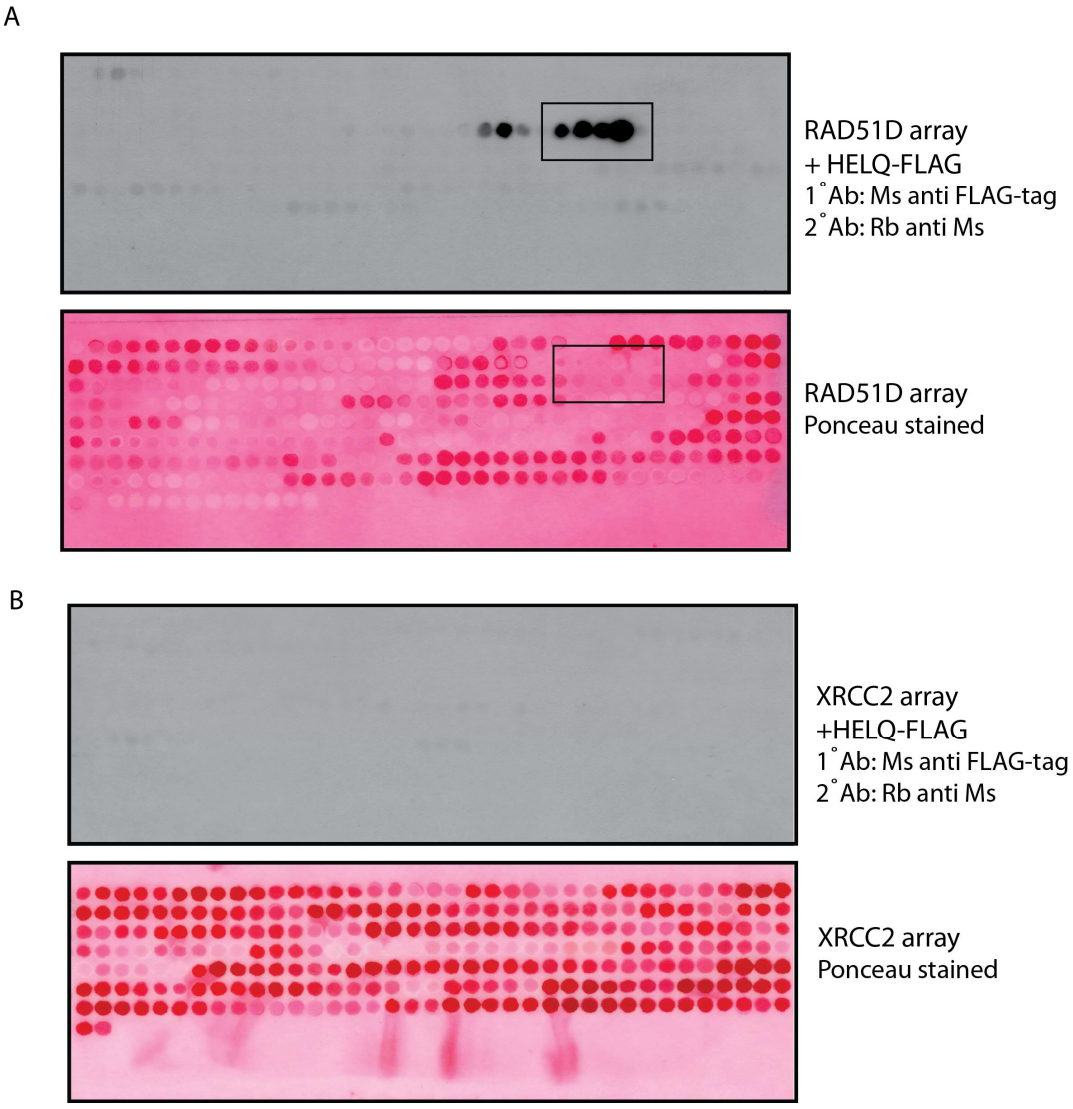


Figure 5-8 Peptide array-based mapping of HELQ-FLAG binding site within BCDX2 complex (RAD51D, XRCC2)

Membranes with peptides corresponding to RAD51D **(A)** and XRCC2 **(B)** were incubated with HELQ-FLAG purified from HEK293. Antibodies specific to FLAG-tag were used for detection. Strong signal in RAD51D array was detected.
The binding spot corresponds to **G-L-T-A-S-R-L-L-Q-L-L-Q-A-K-T-Q-D-E-E-E** peptide. To visualize the spots, membranes were stained with Ponceau.

As a negative control, I also used an XRCC3 peptide array – the only RAD51 paralog that does not interact with HELQ. As expected there was no FLAG-tag signal detected from the XRCC3 array.

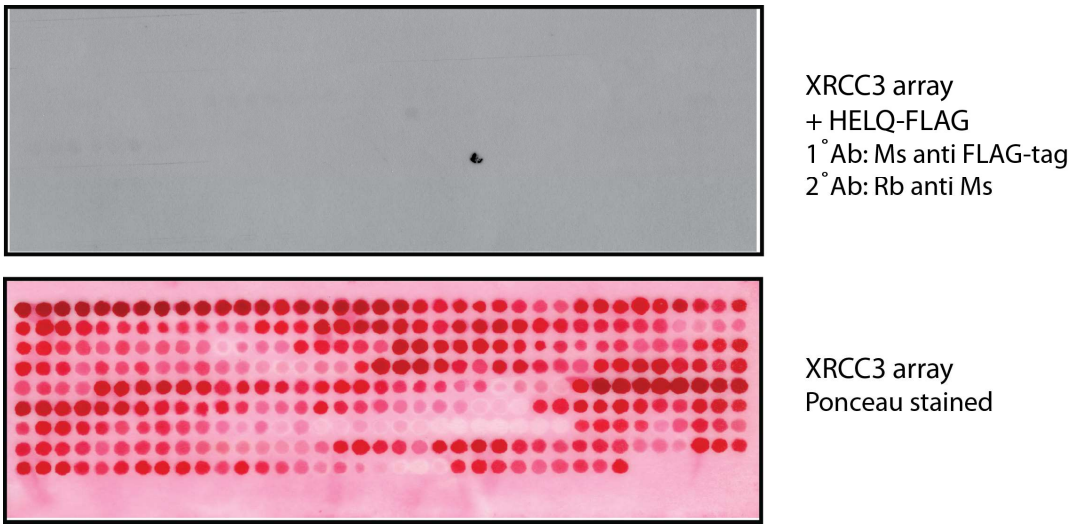


Figure 5-9 Negative control - XRCC3 peptide array.

A membrane containing peptides corresponding to XRCC3 protein was incubated with HELQ-FLAG purified from HEK293 cells. Antibodies specific to the FLAG tag were used for detection. No binding was observed. In order to visualize peptide spots, membrane was stained with Ponceau.

To further confirm these results, I repeated the experiment with a different detection tool. Additionally, to the anti-FLAG tag antibodies, I also used antibodies against human HELQ that we previously optimized in the lab.

As shown in Figure 5-10, the detected signal was dramatically different from the results obtained in the previous experiments. FLAG-tag antibodies allowed for a very low background signal whereas using anti HELQ antibodies led to detection of many spots and stretches across all paralogs, including XRCC3. Interestingly, despite high level of noise, the region in RAD51D that was identified to bind HELQ in the previous experiment is also lit up using this detection method.

Taken together, it seems that the most promising HELQ binding candidate is the G-L-T-A-S-R-L-L-Q-L-L-Q-A-K-T-Q-D-E-E-E peptide located in central region of

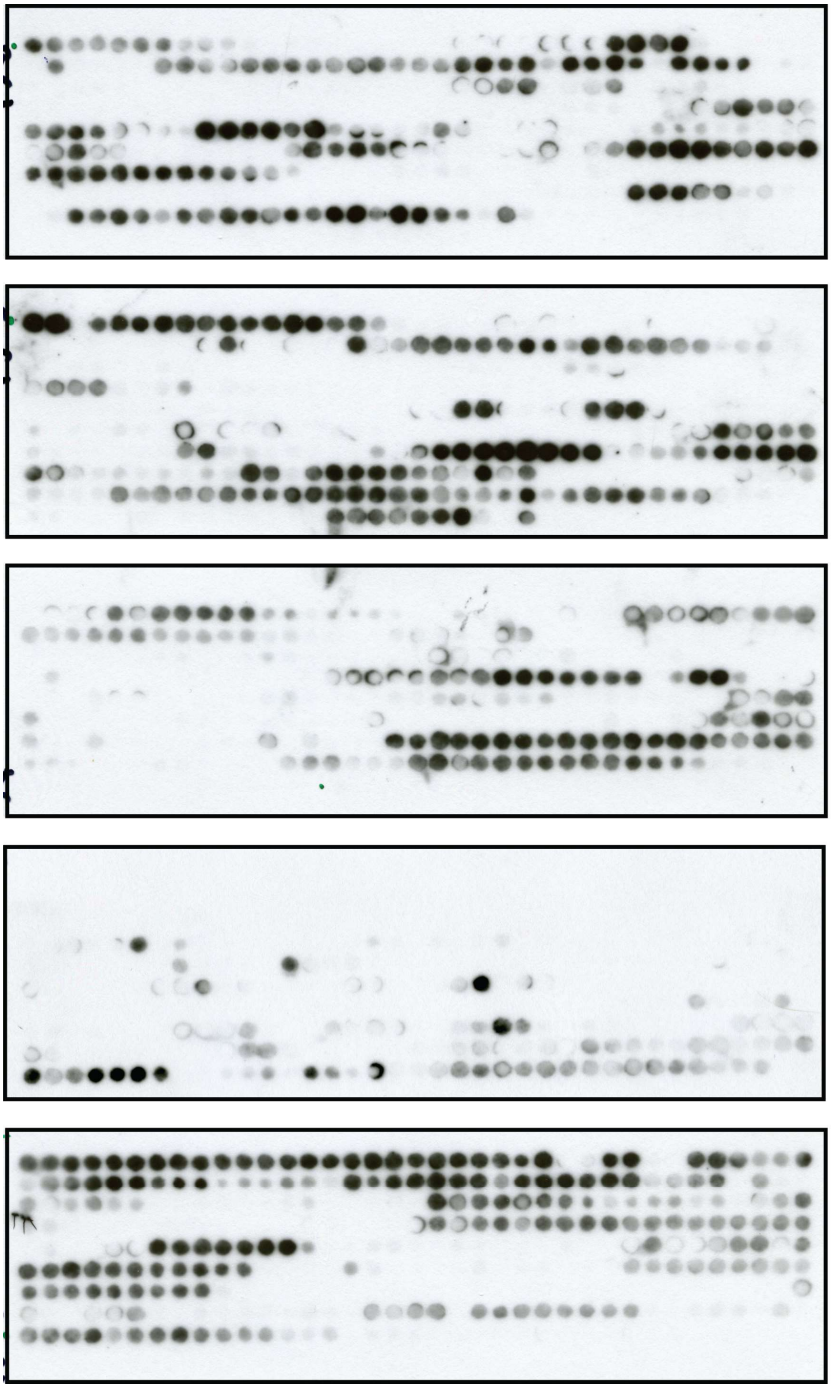


Figure 5-10 Peptide array-based mapping of HELQ-FLAG binding site within BCDX2 complex using anti HELQ antibodies for detection.

Membranes with peptide arrays corresponding to RAD51 paralogs were incubated with HELQ-FLAG. Antibodies specific for human HELQ were used for signal detection. Observed binding pattern is most likely a result of unspecific interaction driven by very low specificity of the HELQ antibodies.

RAD51D (identified by detection with anti-FLAG antibodies). To explore this possibility further, I decided to take two independent approaches. The first method was to determine whether it was possible to pull-down HELQ-FLAG with the identified peptide. Secondly, I used a substitution array to identify amino acids crucial for the interaction with HELQ protein.

Using synthetic, tagged peptides as a bait is a common way to map protein-protein interactions. The amino acids stretch identified in RAD51D was synthetized and biotin-tagged to facilitate pull-down (LRI Peptide Synthesis Facility). As a negative control a scrambled peptide sequence of the same length was prepared and handled in the same way.

As shown in Figure 5-11, it was possible to pull down HELQ-FLAG from HEK293 cells using biotinylated RAD51D peptide but not with the scrambled-sequence control peptide. Of note, as much as 40% of the eluate volume had to be loaded on gel in order to obtain a detectable signal. This may suggest that the observed interaction is weak and only a fraction of molecules is engaged in complex formation. This observation is consistent with my previous experiment where I immunoprecipitated endogenous RAD51C. Although I was able to clear all detectable levels of RAD51C with the pull down and HELQ was detectable in the eluate, the levels of HELQ in the remaining cell lysate where unaffected. Taken together these observations may imply that only a small fraction of cellular HELQ is engaged in the interaction with BCDX2 complex.

In the next step I was interested to determine the key residues that drive the interaction by a substitution peptide array. The mutagenized arrays were created in a way that each of the amino acids in the identified stretch were sequentially mutated to each of the 20 possible amino acids. The substitution array was handled in the same way as described above and binding signal was detected with anti-FLAG antibodies.

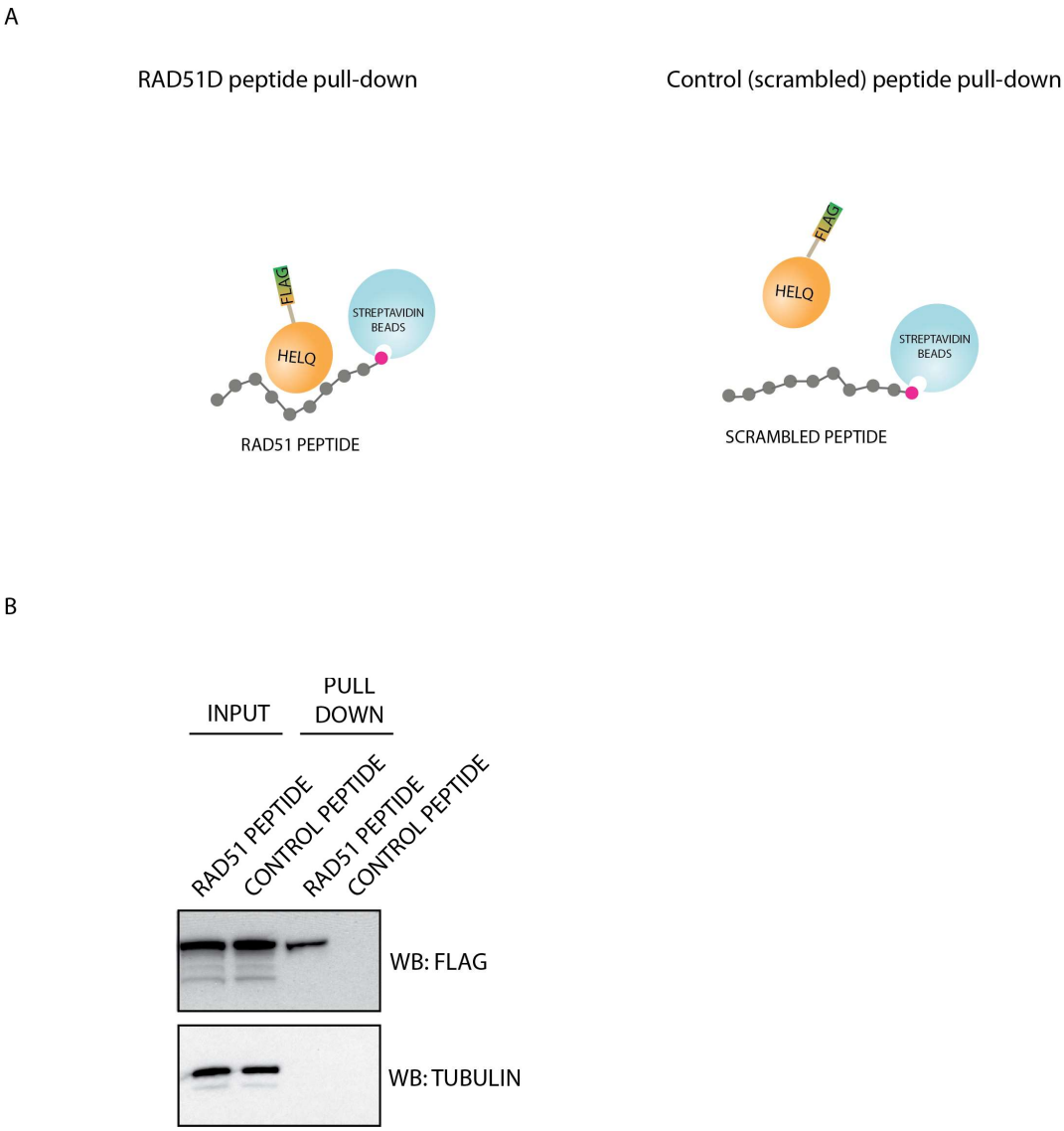


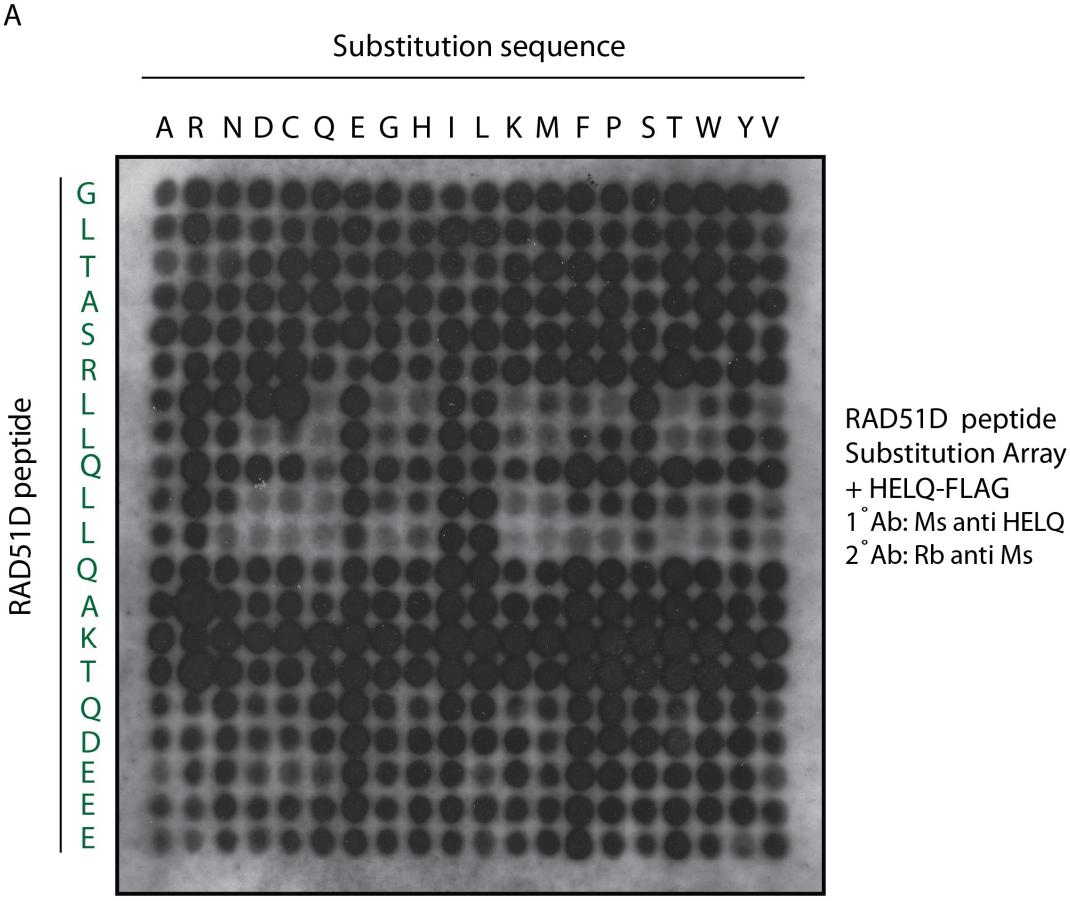
Figure 5-11 RAD51D peptide pull down from HEK293 HELQ-FLAG

(A) Schematic representation of the peptide pull-down experiment. **(B)** HEK293 cells expressing HELQ-FLAG were lysed. Whole cell lysate was cleared by centrifugation and incubation with agarose beads. Pre-treated lysate was incubated with biotinylated RAD51 and control peptides. After streptavidin-coated agarose beads pull-down eluates were subjected to anti-FLAG western blot to detect bound HELQ protein. 3% of the input sample and 40% of the pull-down sample was loaded on a gel.

As shown in Figure 5-12 the key binding motif within the identified RAD51D peptide seems to be LLQLL. Alteration of these residues, abrogated HELQ-FLAG binding (with an interesting exception - substitution of the leucines to isoleucines or glutamic acids did not affect the binding ability). Although analysis of the sequence of interest shows a relatively good evolutionary conservation it should be noted that RAD51D is in general a well conserved protein. Another interesting observation is that analysis of the sequence conservation across several species revealed that variations in the LLQLL sequence are usually substitutions to amino acids that according to the array maintain HELQ-FLAG binding ability.

Taken together, my experiments point towards RAD51D as a likely binding partner of HELQ, although there are clear ambiguities with the peptide array experiments so additional experiments are needed to confirm the nature of this interaction. Given the peptide pull-down result it is tempting to speculate that the BCDX2 complex – HELQ interaction we discovered is mediated by the RAD51D component. To further support this hypothesis it should be pointed out that depletion of HELQ from U2OS cells with four different siRNA sequences led to a minor decrease of RAD51D levels – an effect not observed for other RAD51 paralogs (Figure 5-1).

A



B

G L T A S R L L Q L L Q A K T Q D E E E

```
sp|O75771|RA51D_HUMAN      SRLLQLLQAKTQDEEE
sp|O55230|RA51D_MOUSE      SRLLQLLQARTQDEEK
sp|Q2HJ51|RA51D_BOVIN      SRILQLLQARTPDEEE
tr|M3XB01|M3XB01_FELCA     SRLLQLLQARTPDEEE
tr|E1BVK2|E1BVK2_CHICK     SRLYQMLQARVEDKEE
tr|Q6DJL8|Q6DJL8_XENLA     SRLLQLVQSRTENEDE
tr|F1R474|F1R474_DANRE     NRLLQMLQTKTSNEQE
sp|Q9LQQ2|RA51D_ARATH      RRIAQFICS-SSDATL
```

Figure 5-12 Substitution analysis of HELQ-FLAG binding to RAD51D

(A) The HELQ-FLAG interacting peptide identified by RAD51D peptide array was subjected to substitution analysis. Each amino acid in the peptide was mutagenized to each of the 20 amino acids in a 400-spot array. Membrane was incubated with HELQ-FLAG and after washing subjected to western blotting detection with anti-FLAG-tag antibodies. **(B)** The RAD51D peptide identified to bind to HELQ-FLAG. Residues supposed to be crucial for the interaction are highlighted in red. Alignment of the relevant sequence shows relative conservation of the key residues.

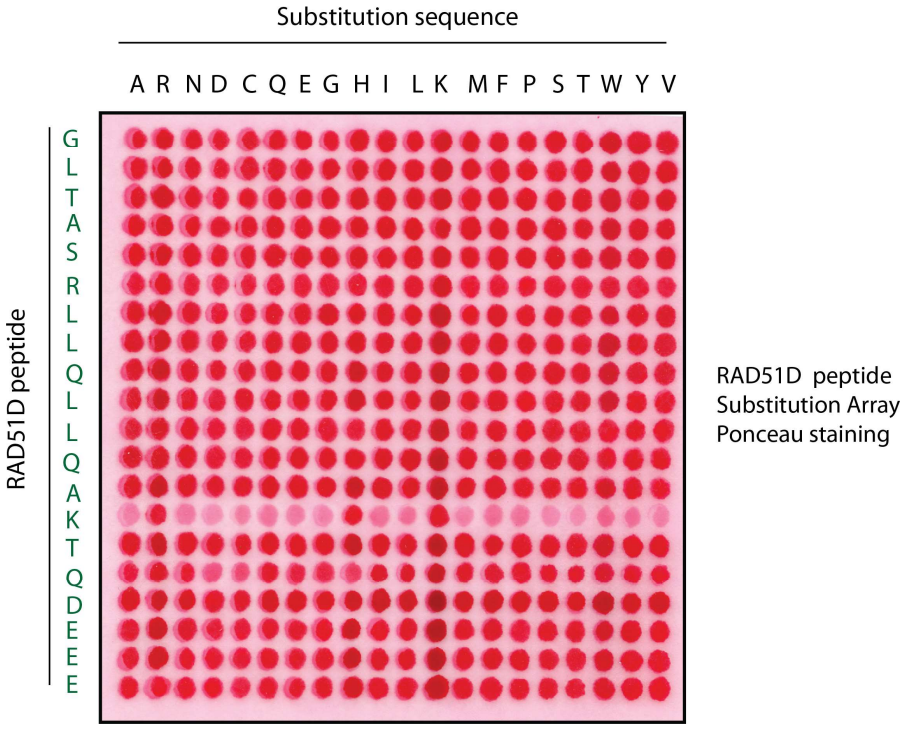


Figure 5-13 Substitution analysis of HELQ-FLAG binding to RAD51D Ponceau staining.

To determine localization of the relevant spots, membrane was stained with Ponceau solution.

5.3 Discussion

Based on the data presented in 0 I wanted to gain a more detailed insight into the nature of HELQ-BCDX2 interaction. Interaction studies proved to be very challenging,

which I contribute to the fact that stability of the entire complex is interdependent on each of its components but also to low expression levels of both RAD51 paralogs and HELQ.

In line with previously published data, my experiments revealed that depletion of any of the BCDX2 complex components affects stability of all of its components, as seen by reduced protein levels in western blots. siRNA-mediated depletion of HELQ had no obvious effect on stability of the BCDX2 with the possible exception of RAD51D levels, that seemed to be mildly lowered in cells transfected with all four siRNAs targeted against HELQ. This observation however requires confirmation with further experiments.

Using peptide arrays to attempt to circumvent the stability problems in cells did not lead to conclusive observations. Incubation of HELQ peptide arrays with recombinant HIS-tagged BCDX2 complex from Sf9 cells (previously used in *in-vitro* binding experiments) led to two different set of binding patterns depending on the visualisation method (using anti-RAD51C antibodies or anti-HIS antibodies). A possible explanation for this discrepancy across these two experiments is that the observed binding was driven by unspecific interactions. Only two out of the four BCDX2 complex components were HIS-tagged (RAD51B and RAD51D), and it is possible that there was a pool of free of RAD51 paralogs molecules in the solution (disengaged with BCDX2 complex). These paralogs molecules could potentially bind to various regions of the HELQ sequence either via site specific interaction or simply sticking via hydrophobic regions. In such a situation probing for HIS-tag would identify a wider range of peptides rather than only detecting RAD51C as it was the case with using anti-RAD51C antibodies.

In a reciprocal approach in which RAD51 paralogs peptide libraries were immobilized on membranes and incubated with HELQ-FLAG purified from HEK293 cells the most promising results were obtained for RAD51D protein and anti-FLAG peptide as a detection tool. The identified spots corresponded with G-L-T-A-S-R-L-L-Q-L-L-Q-A-K-T-Q-D-E-E-E peptides located in the middle of RAD51-like ATPase fold between walker A and Walker B motives and contained a potential ATM/ATR phosphorylation substrate (TQ). Unfortunately, the identified binding pattern was not confirmed when the experiments were repeated but anti-HELQ antibody was used instead of anti-FLAG. Given our experience with the anti-HELQ antibodies I proceeded to validate RAD51D binding peptide. I was able to pull down HELQ-FLAG from HEK 293 using

the identified peptide and by using substitution array I identified the residues critical for maintaining binding. The minimal binding motif seems to be LLQLL. Alteration of these residues, especially leucines abrogated HELQ-FLAG binding. Interestingly, substitution of the leucines to isoleucines or glutamic acids did not affect binding.

To gain more clarity and further validate the nature of this interaction I could also try to use alternative approach, taking advantage of MS techniques, such as chemical cross-linking coupled with MS (XL-MS) or Hydrogen Deuterium Exchange with MS (HDX-MS). Combining cross-linking with MS and bioinformatics allows identification of amino acid pairs that are positioned in close proximity to each other and therefore modelling and mapping structural details of functional complexes in solution.

I could potentially express all components of the BCDX2 complex together with HELQ, purify the complexes and use chemical cross linking to convert non-covalent interactions between BCDX2 and HELQ into covalent bonds. Such artificially fused HELQ-BCDX2 complex would withstand denaturating conditions and thus could facilitate analysis with methods that normally dissociate protein complexes. Alternatively, an interesting approach would be to use HDX-MS, a technique that allows to quantitatively monitor the rates in which protons of a particular protein exchange with a deuterated solvent. This exchange rate is, among other factors, a function of solvent's accessibility. The intramolecular differences in exchange rate are caused by the fact that surface amide protons unengaged in the internal interactions of secondary and tertiary structural elements exchange rapidly, while those that are involved in hydrogen bonding or buried in the protein interior exchange at a much lower rate. To take advantage of HDX-MS I would need to express HELQ and components of the BCDX2 complex and optimize conditions for the exchange reaction, replace aqueous (H_2O-based) solvent with D_2O-based solvent and quench the exchange reaction at various time points. The protein complexes would be then subjected to MS and by exploiting distinctive spectroscopic properties of deuterium compared to hydrogen the exchange rates could be converted into structural data allowing to identify proximal aa moieties of HELQ and BCDX2 complex. Given the solubility problems I encountered when expressing HELQ and BCDX2 complex, this technology could be particularly useful as it only requires small quantities of sample (500–1,000 picomoles) and allows to analyse complexes that can only be handled at low concentrations (as low as 0.1 μM).

Ultimately, the identified aa residues responsible for HELQ-BCDX2 interactions should be validated in cells. To do so, I would need to mutate the identified motif, introduce such construct to cells lacking given gene (or cells that have been subjected to siRNA silencing) and validate if HELQ-BCDX2 interaction is compromised (for instance by IP).

Chapter 6. HELQ functional studies using *Xenopus laevis* egg extract immunodepletion system.

6.1 *X. laevis* egg extract immunodepletion experimental system to study the role of HELQ in DNA interstrand cross links.

Removing a factor from the system within which it performs its function and monitoring the consequences of such depletion is a long-standing approach to study biological function of proteins. Extensive research has been done to develop tools and experimental systems facilitating such manipulations. Gene knock-outs or RNA interference are the most widely used methods that contributed to many break-through findings. It is, however much more challenging to remove already translated proteins from a cell. To address the latter, several cell-free systems have been developed, one of the most successful being the South African clawed frog (*Xenopus laevis*) egg extract. As in most vertebrates, frogs' eggs are arrested in metaphase of meiosis II. Fertilization (addition of sperm DNA) triggers a chain of events – in just 7 hours the eggs undergo as many as 11 rounds of cell divisions with only minor signs of transcription and translation (Gillespie, Gambus, & Blow, 2012). This protein-rich, soluble extract prepared from unfertilized female gametes is therefore a perfect tool to study essential cell cycle and replication related processes. It also has been extensively utilised in the field of replication coupled-DNA repair and many modifications of the system have been developed. One particular variation aiming to study proteins' function in the DNA ICL repair was developed by Johannes Walter's laboratory. The system's design allows to monitor progression of replication-dependent repair of a single, well-defined ICL introduced site-specifically in a plasmid, in an egg extract depleted for a certain protein with specific antibodies. Recently this system has been successfully utilised to study the function of several Fanconi anemia factors (Knipscheer et al., 2009; Räschle et al., 2008). Noteworthy, the group successfully used this system to study the mechanism of HR-dependent ICL repair – the very step of the ICL-repair process where we think HelQ may perform its function (Long, Räschle, Joukov, & Walter, 2011). Taken together, HelQ functional

studies in this system seem to be a perfect complimentary approach to the mouse and molecular biology work performed so far.

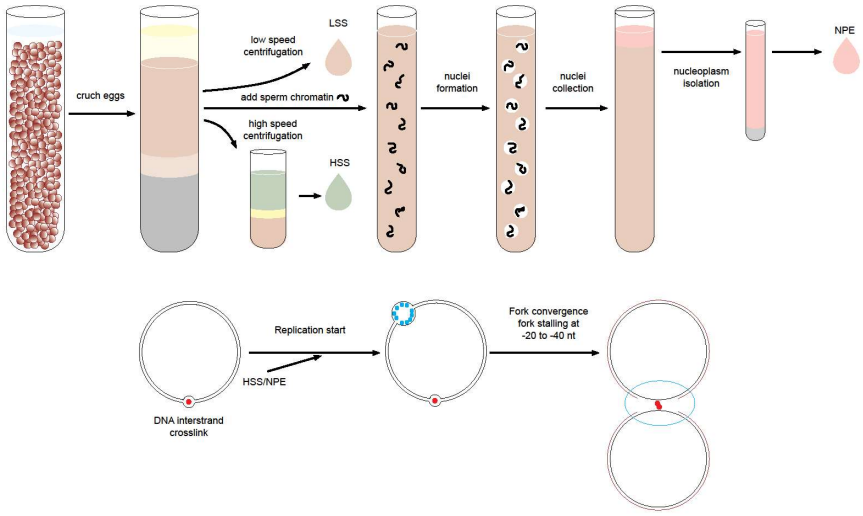


Figure 6-1 Schematic representation of *Xenopus* egg extract preparation and replication dependent ICL repair plasmid model.

Top panel: unfertilized *Xenopus laevis* eggs are crushed and the crude cytoplasmic fraction is collected. Separation by centrifugation allows to obtain cytoplasmic extract including the membranes LSS (low speed centrifugation) and cytoplasmic extract without membranes, HSS (high speed centrifugation). When the crude cytoplasmic extract is incubated with sperm chromatin and ATP nuclei form. Nucleoplasm, NPE can be isolated by centrifugation of the nuclei. **Bottom panel:** model for DNA replication-dependent ICL repair in *Xenopus* egg extract. Plasmid containing a single, defined ICL is incubated in membrane-free HSS extract resulting in the assembly of pre-replication complexes. Addition of a highly concentrated NPE triggers replication initiation and allows a single, complete round of DNA replication. Circular nature of plasmid DNA and bi-directional replication result in two replication forks converging on the ICL.

6.2 Generation of *X. laevis* HelQ expression constructs.

There are two critical conditions to be met in order to successfully employ the *X. laevis* egg extract experimental system. In order to specifically and efficiently deplete the protein of interest, it is necessary do develop high quality antibodies. Additionally, to prove that the observed effect is caused by depletion of the protein of interest (as opposed to an off-target effect) it is critical to perform an add-back experiment in which a recombinant, functional protein is added to the extract to correct the defect observed following protein depletion.

In order to learn the principles of preparing and handling *X. laevis* egg extract as well as to familiarize with the experimental set up, we established a collaboration with Johannes Walter as well as securing an UCL travel grant to visit his laboratory in Boston.

The strategy was to express and purify full length *X. laevis* HelQ and use it to immunize rabbits to raise polyclonal antibodies. As the cDNA sequence for *X. laevis* HelQ was not present in databases at the time (xgc.nci.nih.gov, xenbase.org) I first had to use bioinformatic tools to obtain the sequence of interest.

6.2.1 Determining the *X.laevis* HelQ cDNA sequence

X. laevis is one of the model organisms frequently used in molecular biology. However, its genome sequencing project has proved challenging, mostly due to the fact that it is tetraploid. To bypass these problems, research groups working with *X. laevis* use *X. tropicalis (a close* relative of *X. laevis*), to find *X. laevis* genes transcript sequences of interest. I utilized a similar strategy in which I blasted *X. tropicalis* HelQ cDNA sequence against a *X. laevis* E.S.Ts. (Expressed Sequence Tag) database. I identified a number of candidate sequences but none of them spanned across the entire HelQ gene.

The top two candidates were: E.S.T. 7764769 covering the 1-678 amino acid region and E.S.T 4674388 consisting of 318 amino acid sequence of the C-terminus. According to alignment of the obtained sequence with human and *X. tropicalis* HelQ cDNA the identified E.S.Ts were lacking a region of about 60 amino acids. Based on the determined sequence I then designed PCR primers flanking the missing region and made a cDNA library from mRNA extracted from egg extracts. After sequencing the PCR product I was able to determine full length *X. laevis* HelQ cDNA sequence.

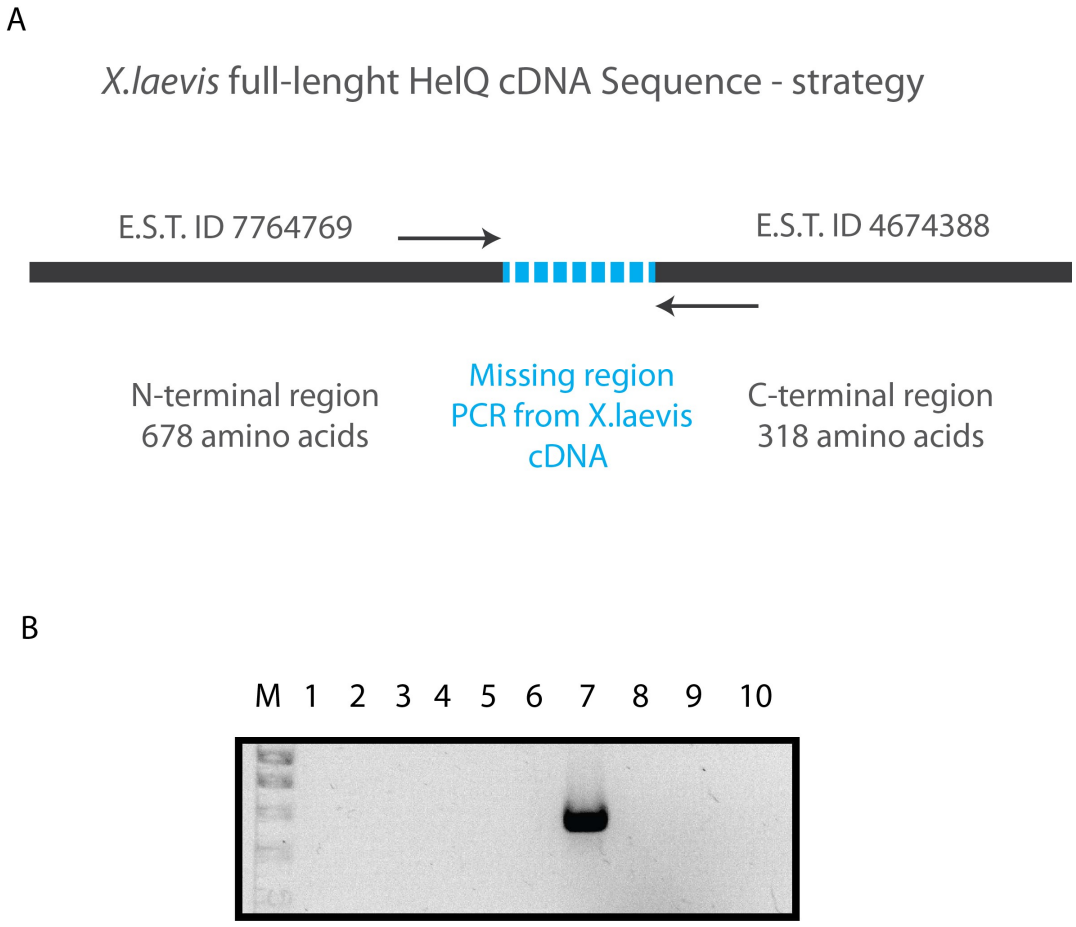


Figure 6-2 Obtaining *Xenopus laevis* HelQ cDNA sequence.

(A) *X. laevis* HelQ cDNA sequence was determined from two independent Expressed Sequence Tags. The remaining region not covered by the ESTs was obtained from a cDNA library using PCR reaction with primers flanking the missing sequence. **(B)** Total RNA was purified from *X. laevis* egg extract reverse-transcribed to obtain cDNA. cDNA was then used as a template in a PCR reaction with primers flanking the missing region of the gene. Several different PCR conditions were used to optimize the reaction. Only condition in line number 7 facilitated PCR reaction.

6.2.2 *De novo* synthesis of codon optimized *X.laevis* HelQ cDNA and expression of the full-length protein in insect High5 cell line.

To be able to produce large quantities of properly folded and soluble HelQ, I decided to use an insect cell-based expression system. Using Life Technologies online sequence analysis tool I codon optimized *X. laevis* HelQ for expression in insect cells.

An optimized cDNA was then provided to the CRUK Protein Expression Unit where it was cloned into appropriate vectors to prepare viral genomes in the form of a bacmid.

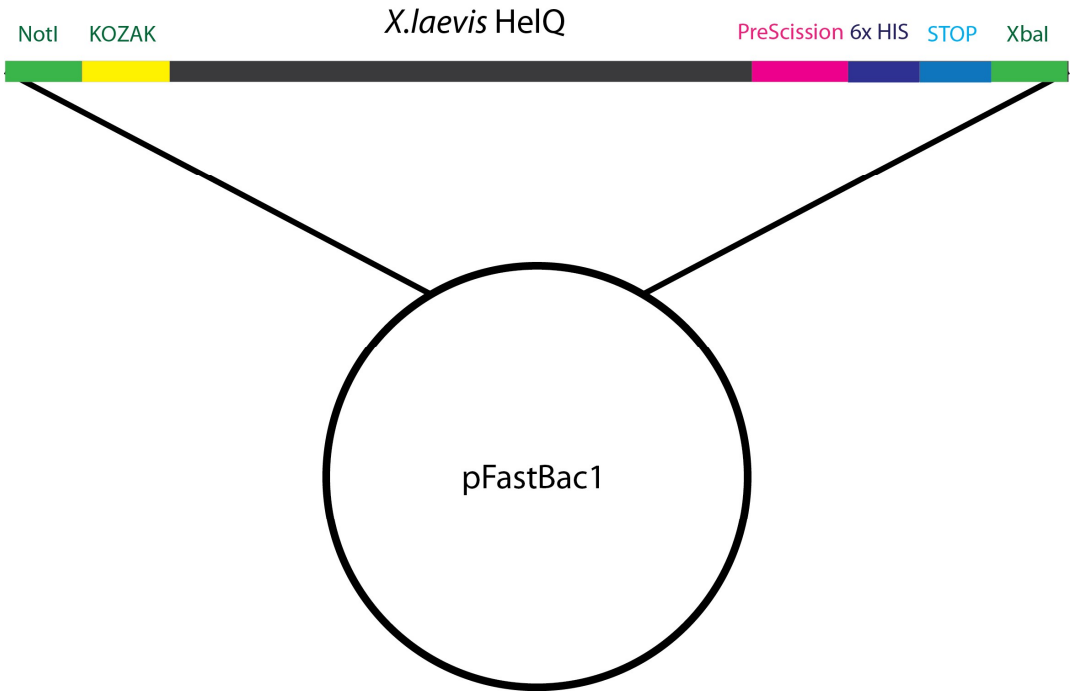
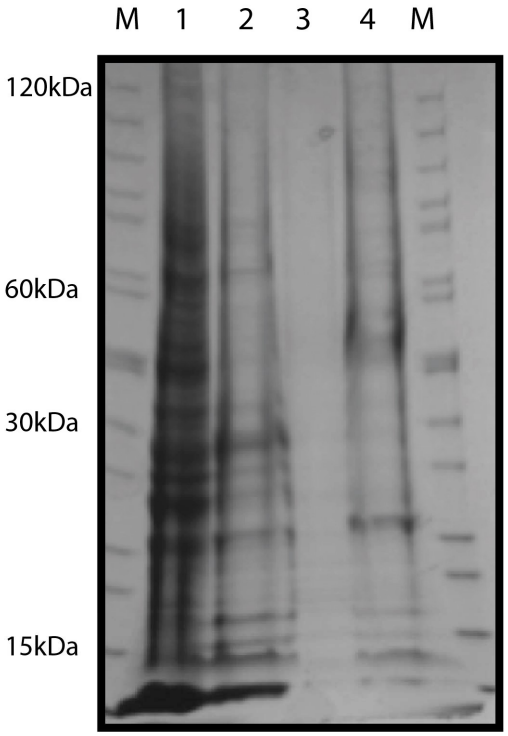


Figure 6-3 *X. laevis* HelQ expression construct - schematic.

X. laevis HelQ cDNA was synthesised de-novo by Life Technologies. To facilitate downstream cloning and expression a NotI restriction enzyme recognition site as well as Kozak sequence were added 5' of the cDNA. Downstream of HelQ sequence, Prescission nuclease, six repetitions of HIS-tags sequence, STOP codon and XbaI restriction site were introduced. The construct was inserted into pFastBac plasmid.

A


M 1 2 3 4 M
120kDa
60kDa
30kDa
15kDa

B

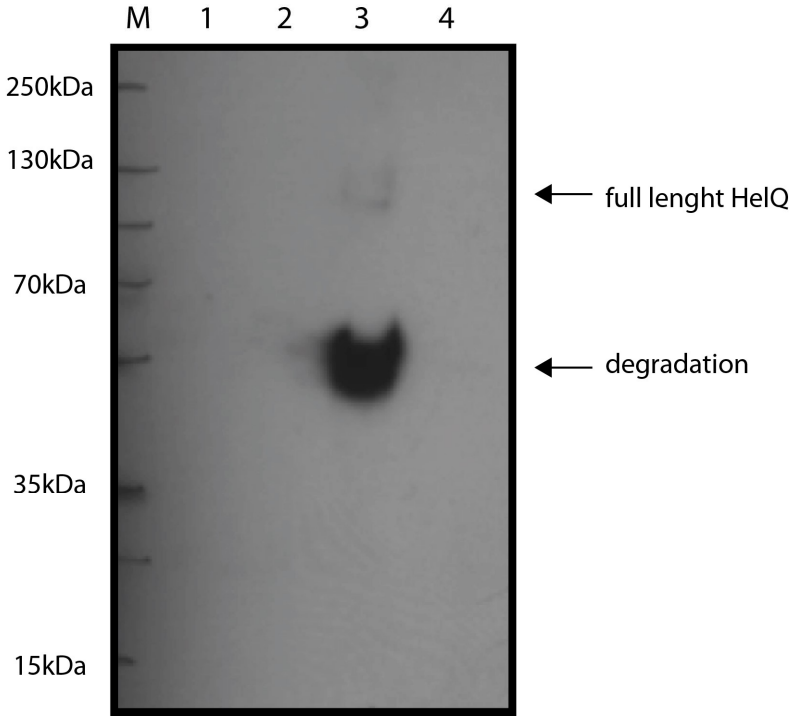

M 1 2 3 4
250kDa
130kDa
70kDa
35kDa
15kDa
full lenght HelQ
degradation

Figure 6-4 Analysis of *X. laevis* expression in the High5 cell line by SDS-Page (A) and Western blot (B)

Bacmids were generated by repeated transformation of *X. laevis* plasmid DNA into DH10 *E. coli*. White colonies (indicating recombination of exogenous protein) were picked from plates and re-streaked to verify recombination. Two colonies were used for viral production and amplified to P3. To test for expression of xHelQ, P3 virus was used to infect High5 cells. The culture was allowed to grow for 3 days after which the cells were lysed, and the soluble fraction incubated with Ni-NTA resin. The insoluble fraction was washed and dissolved in buffer containing 6 M urea and 5 mM DTT. An aliquot was taken for analysis by SDS-PAGE and the remaining solution incubated with Ni-NTA resin for 30 mins at room temperature. The resin was then washed extensively, and any bound protein eluted by the addition of SDS-PAGE loading buffer. Samples were analysed by SDS-PAGE followed by Coomassie staining and Western blot. The SDS-PAGE gel **A** does not show a distinct band indicating expression of *X.laevis* HelQ in the soluble or insoluble fractions of the lysate. The Western blot **B** using anti-His antibody for detection detects a protein at 70 kDa in the insoluble fraction of the cell lysate after incubation with Ni-NTA. Two bands of lower intensity are also detected and are resolved between 130 and 100 kDa. M-molecular marker; 1-soluable fraction; 2-insoluable fraction; 3- elution of insoluble fraction bound to NiNTA; 4- elution of soluble fraction bound to NiNTA resin. The experiment was repeated with a new set of viruses to verify this result.

6.2.3 Expression of the fragments of *X. laevis* HelQ in *E.coli*

Due to the solubility problems encountered in the insect cell line I decided to test several alternatives. A commonly used technique in raising antibodies for immunodepletion is expression of fragments of the protein of relatively well-defined tertiary structure and size ranging from 25kDa to 30kDa. Due to the time limitations I switched from insect cells to bacterial expression.

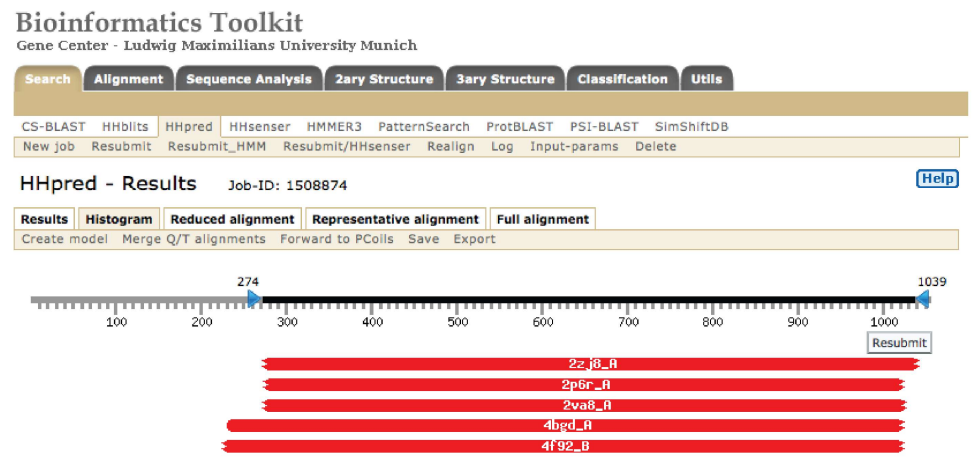
6.2.4 *X.laevis* HelQ tertiary structure prediction

To determine the three-dimensional properties of *X. laevis* HelQ, the amino acid sequence was analysed by HHPred software developed by Gene Center Munich and available at the url: http://toolkit.lmb.uni-muenchen.de/hhpred/.

X. laevis HelQ amino acid sequence was introduced and analysed with the following settings: HMM database to be searched: pdb; MSA generation method: HHblist; maximum MSA generation iterations: 3; score secondary structure: on; alignment mode: local. In brief, the HHpred software is a tool that compares the query sequence with the databases of proteins for which crystal structures are available. It returns the list of best alignments with parameters describing the quality of alignment. For

X.laevis HelQ the best hit was 2zj8_A - *archaeal* DNA helicase Hjm apo state in form 2 from *Pyrococcus furiosus* (Oyama et al., 2009). The alignment was described by the following parameters: probability 100%; E-value: $2.1E^{-82}$; P-value: $3.1E^{-85}$; score:777.1; SS:67.1; cols:707; query HMM: 244-1039; template HMM:6-719; HMM: 720.

A



B

	No	Hit		Prob	E-value	P-value	Score	SS	Cols	Query HMM	Template HMM
☐	1	2zj8_A	DNA helicase, putative	100.0	2.1E-82	6.3E-87	797.1	67.1	705	274-1039	6-719 (720)
☐	2	2p6r_A	Afuhel308 helicase; pro	100.0	1E-80	3.1E-85	779.3	54.5	669	275-1022	9-685 (702)
☐	3	2va8_A	SSO2462, SKI2-type heli	100.0	2.7E-78	8.2E-83	759.3	71.5	673	274-1025	13-713 (715)
☐	4	4bgd_A	Protein SNU246, PRE-mRN	100.0	4.9E-63	1.5E-67	672.3	62.4	712	231-1022	1-758 (1722)
☐	5	4f92_B	U5 small nuclear ribonu	100.0	4.7E-64	1.4E-68	681.7	50.2	719	226-1022	18-780 (1724)

Figure 6-5 *X.laevis* HelQ sequence alignment with known proteins from PDB database: 5 hits with highest scores.

Bar diagram **(A)** with top five hits for *X.laevis* HelQ amino acids alignment. As indicated by the bars, the first 200-280 amino acids of the query do not align with any of the hits. The summary hits list **(B)** presents the top hits together with parameters describing the quality of a particular alignment.

Alignment with the highest quality parameters (2zj8_A) was then used to model the three-dimensional structure of X.laevis HelQ with the MODELLER software developed by Andrej Sali's group and maintained by Ben Webb available on: https://salilab.org/modeller/ (Eswar et al., 2006; Martí-Renom et al., 2000)

MODELLER is an on-line tool used for comparative or homology-based modelling of proteins' three-dimensional structure. It requires an alignment of the sequence of interest with a related sequence with determined structure. The software returns a calculated model with non-hydrogen atoms' coordinates.

Using the described algorithm I chose three fragments to be expressed in E.coli and used as immunogens (Figure 5-5).

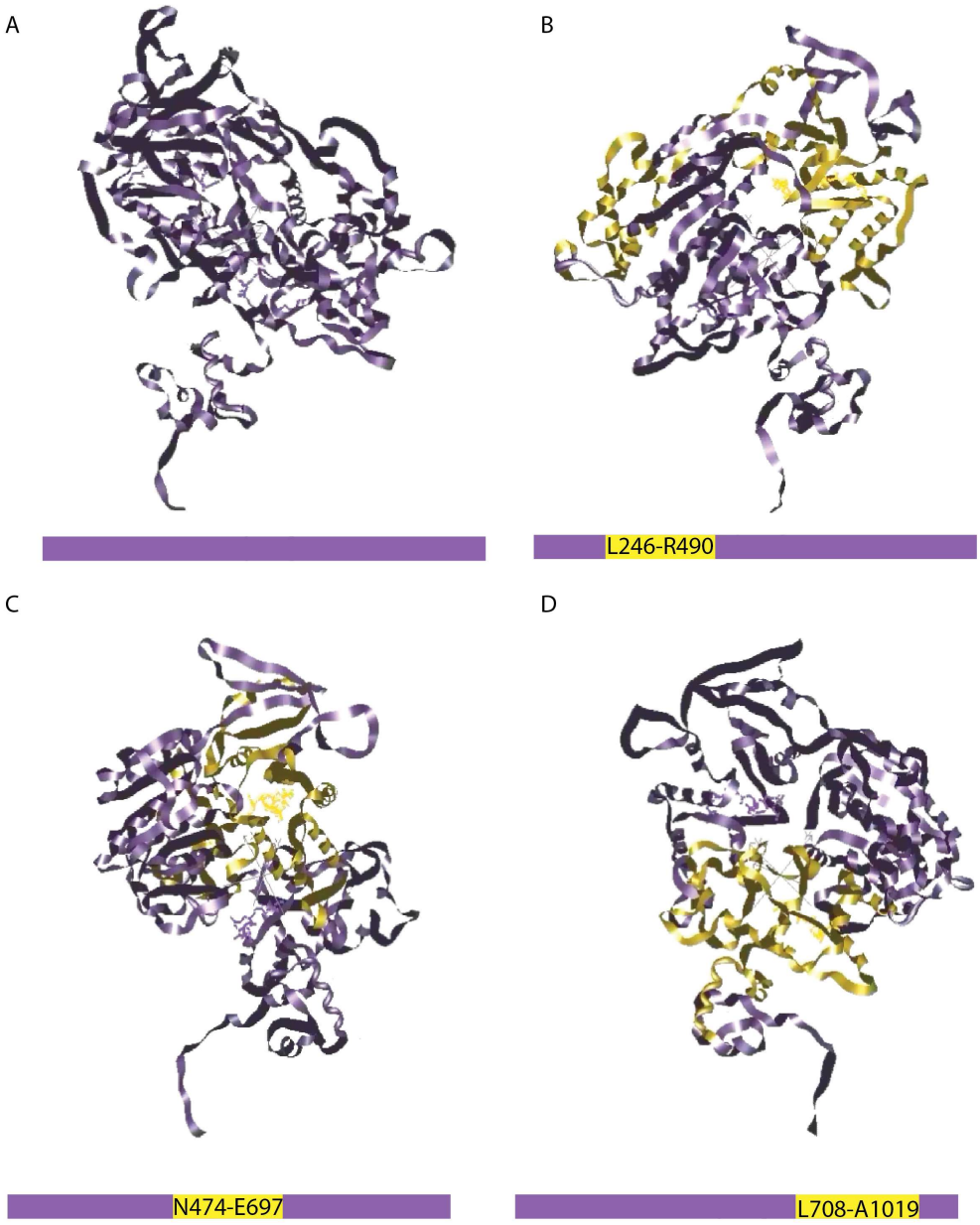


Figure 6-6 *X.laevis HelQ* three-dimensional structure prediction.

X.laevis tertiary structure was modelled using MODELLER software **(A)**. Three fragments of approximately 25kDa were chosen to be cloned and expressed in *E.coli* as antigen for antibodies production **(B)**,**(C)**, **(D)**.

Fragment	Location	Size
Fragment 1	L240-R490	29kDa
Fragment 2	N474-E697	27kDa
Fragment 3	L708-A1019	25kDa

Table 6-1 Characterization of the *X.laevis* HelQ fragments.

6.2.5 Cloning, expression and purification of *X. laevis* HelQ immunogenic fragments.

Once the selected domains were amplified with PCR I introduced the fragments into the Gateway system to generate appropriate constructs and tested their expression in *E. coli*.

Gateway is a powerful cloning system that facilitates fast switching of your gene of interests across various expression and tagging systems. Expression of foreign proteins or their fragments in *E. coli* frequently affects protein folding, thus impacting solubility, which is why I decided to express the domains using two different tagging systems. I choose GST as it is a large soluble tag that should positively affect the solubility of the fragments. Additionally, I also tried expression with HIS-tag as it allows affinity purification under denaturing conditions – a particularly helpful feature when the expressed protein turns out to be insoluble.

6.2.5.1 Expression and purification of X. laevis HelQ Fragment 3-HIS

To test expression, overnight start cultures were added to fresh LB media and cultivated at 37C on a rotating shaker. When OD reached 0,7 IPTG was added to induce expression of the construct and cultivated for another 3 hours. Samples were taken every hour from the moment of induction (first being taken before IPTG was added). Samples were immediately spun down, and supernatant was taken away. Pellets were snap-frozen in liquid hydrogen. Once all the samples were collected the pellets were thawed and lysed. To check whether expressed fragments were soluble, the lysates were centrifugated to separate fractions. SDS sample buffer was added to pellets (insoluble fractions) and supernatant (soluble fractions) and all the samples were run on SDS-Page gel that was then stained (Figure 6-7)

Unfortunately, none of the created constructs expressed well. There were no bands of expected size in IPTG-induced lines of soluble fractions and only Fragment 3-HIS and Fragment 1-GST and Fragment 2-GST seemed to be found in the insoluble fractions (data shown for the Fragment 3-HIS expression). As HIS-tag allows for purification from denaturating conditions I decided to carry on with Fragment 3-HIS expression and purification on larger scale. Of note, the band of appropriate size (26kDa) is also present in the sample before IPTG addition however, it is only a small fraction of the samples taken upon induction). This can be explained by Lac-operon leaking which is not uncommonly observed.

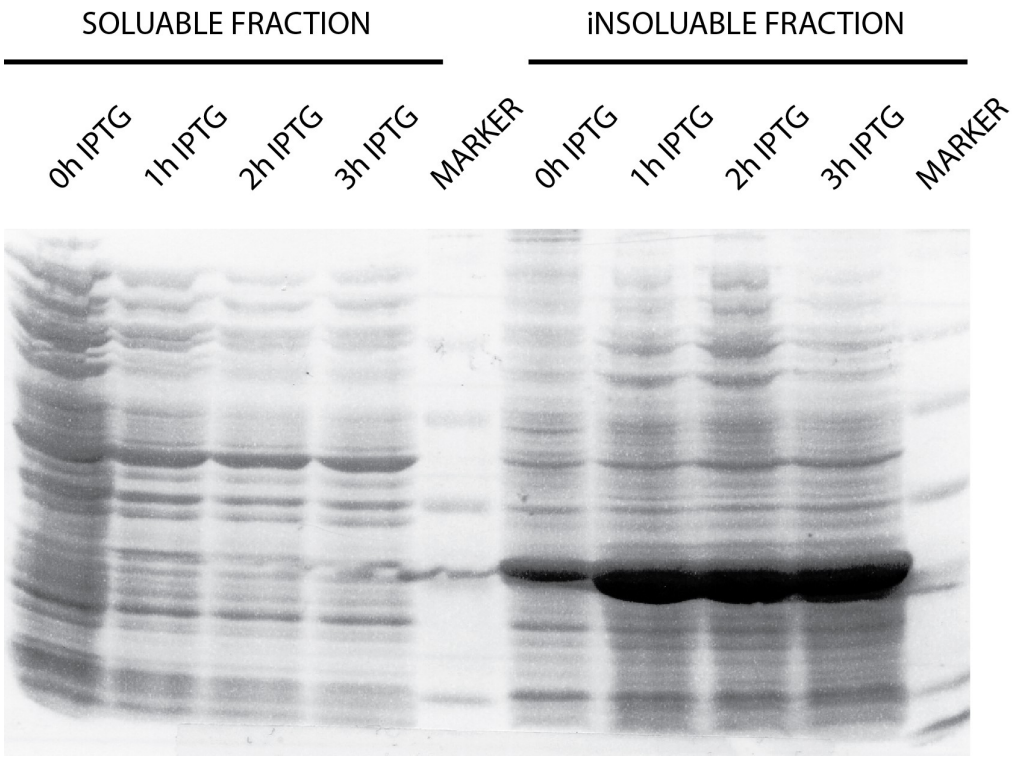


Figure 6-7 *X. laevis* HelQ Fragment 3-HIS expression in *E.coli*

Overnight culture of BL21 bacteria transformed with appropriate plasmid was used to inoculate fresh LB and cultivated in a shaker at 37C. When OD reached 0,7 IPTG was added to induce expression and the culture was left to grow for 3 hours. Samples of the culture were taken every hour. When all the samples were collected, bacteria were spin down and the pellets were lysed. Lysates were centrifugated to separate soluble and insoluble fractions. SDS-sample buffer was added to the pellets (insoluble fraction) and to the supernatant (soluble) and run on SDS Page gels that were then subjected to staining.

As the fragment was found only in the insoluble fraction, large=scale purification was performed under denaturing condition using Talon beads that contain cobalt-based

resin rather than nickel. In brief, the lysis was performed with help of a French press, extraction was performed in a buffer containing 6M guanidine to ensure denaturing conditions and elution was done in an imidazole-based buffer.

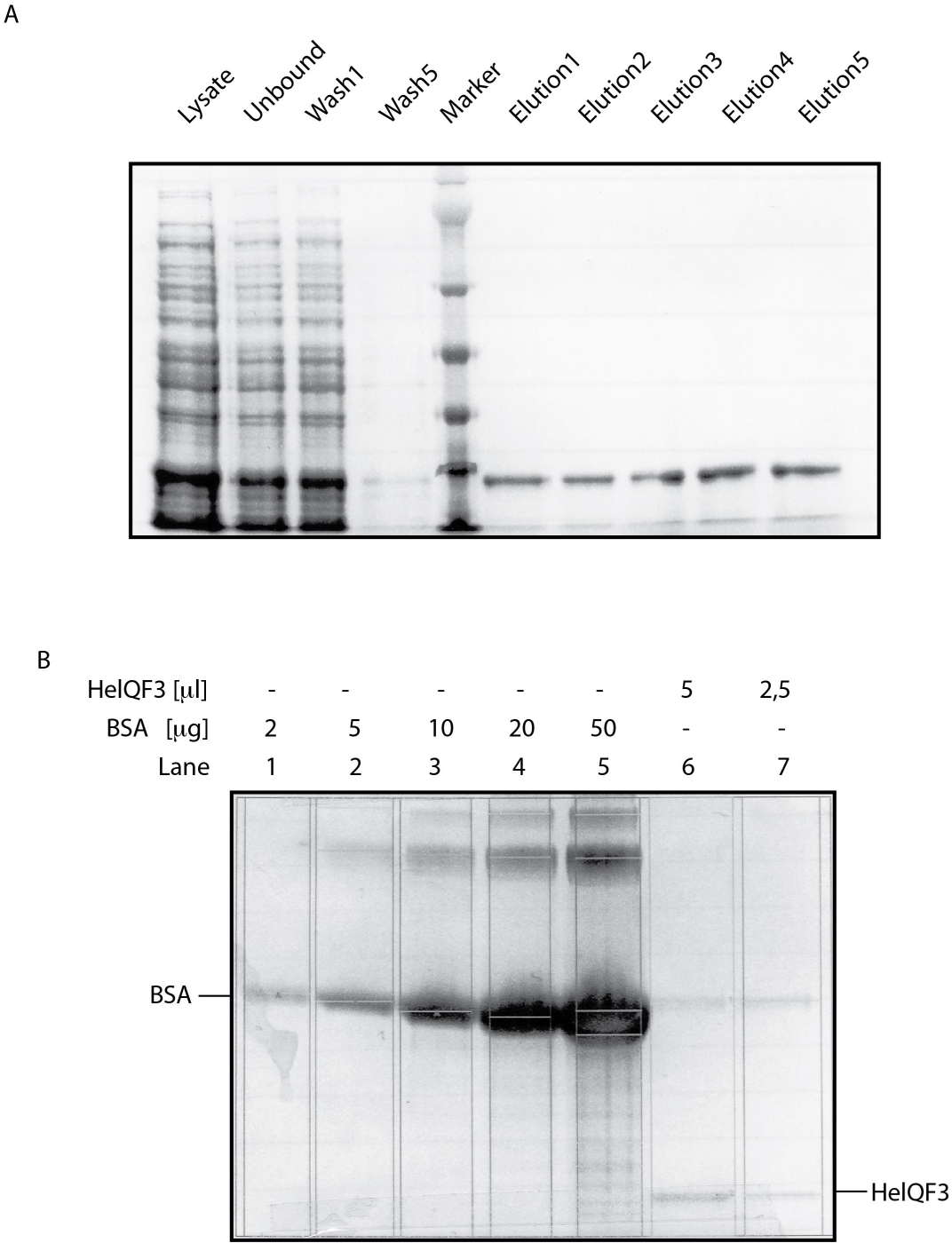


Figure 6-8 *X.laevis* HelQ fragment 3 purification

(A) 2 litres of bacterial culture were harvested and lysed with French press. The lysate incubated with Talon beads. Unbound fraction was collected to monitor efficiency of

binding. Beads were thoroughly washed and subjected to elution. Eluate fractions were collected. All samples were resolved on SDS-Page gel and stained. **(B)** BSA standard (2-50μg) together with two concentrations of pooled eluate were resolved on SDS-Page gel to assess concentration of eluted protein. The concentration was calculated as 240ng/μg

To assess the efficiency of HelQ-HIS binding to the cobalt resin, a sample of the lysate after incubation with beads was taken. The beads were subjected to 5 rounds of thorough washes. Subsequently the beads were incubated with a gravity-flow of elution buffer (Figure 6-8 A).

By looking at fragment 3 levels in unbound and eluate fractions it seems that the binding was suboptimal as the protein was still present after incubation with beads. Additionally, some of the already bound protein was lost during the wash step.

Although stringent conditions during binding and wash decreased the amount of purified protein, the eluate was extremely clean, and the only contaminating band was very weak and coming from a small molecular mass protein.

As all the elution fractions were equally pure I pooled them to facilitate subsequent dialysis. To inject into rabbit guanidine had to be exchange for urea so the eluate was dialysed against 5M urea.

After dialysis, it was important to approximate the total amount of purified protein. To do so I used BSA standards (a series of dilutions from 2 to 50 μg) and two concentrations of the purified protein (Figure 6-8B). The samples were run on SDS-Page gel that was then stained with Coomassie Brilliant Blue. The concentration of HelQ was estimated by gel imager software as 240ng/μl. As the total volume of the purified protein was 20 ml, I obtained around 4800μg of *X.laevis* HelQ fragment 3.

The goal of this experiment was to purify a small amount of very clean protein to be used as antigen. If large quantities of the protein were needed the procedure could be improved in terms of efficiency, however for this particular application, stringent conditions allowed me to obtain enough protein of the required purity.

6.2.6 Antibodies directed against *X.laevis HelQ*- validation

Antigenic HelQ fragment 3 was sent to Eurogentec company in Belgium who immunised two rabbits to raise polyclonal antibodies.

Materials that Eurogentec provided us with were the following: pre-immunisation sera (control of specificity of the raised antibodies) and sera from rabbits drew upon immunization and antibodies maturation.

6.2.6.1 Western blot validation

To evaluate the sera I performed a western blot on *X. laevis* egg extracts. It was meant to be a quick pilot test and therefore it lacks some aspects of a proper experimental set-up. Namely, to quickly test various concentrations of the sera I cut the nitrocellulose membrane into slices corresponding to samples' migration lines in the gel and incubated those with a range of post-immunisation sera concentrations (1:200 – 1:100). Before visualisation the slices were assembled back, and the cuts are marked with dashed-lines (Figure 6-9). Western blot results were encouraging – sera from both rabbits not only recognized the recombinant Fragment 3 of HelQ used as immunogens (Figure 6-9 lines 1 and 9) but there was also a strong band of appropriate size of full length *X. laevis* HELQ in the egg extract suggesting the antibodies recognize the endogenous protein as well. Admittedly, there were unspecific bands present as well but given our experience with human and mouse HELQ antibodies, the results with crude sera were very encouraging. Comparing western blot signals from both rabbits it seems that there are no major differences apart from the fact that serum Rb472 gives less unspecific signal at higher dilutions (1:2000 and 1:800).

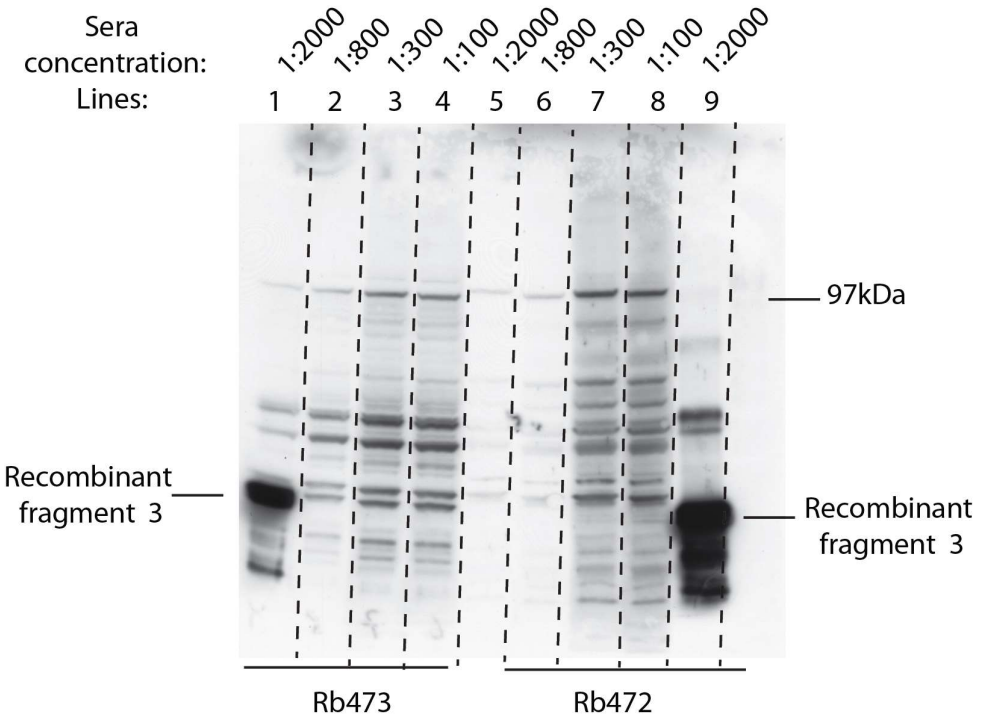


Figure 6-9 *X. laevis* HelQ antibodies testing.

Sera obtained from two rabbits (Rb473 and Rb472) were tested. 1µl of egg extract was run on SDS-Page gel and transferred onto nitrocellulose membrane. The membrane was stained with Ponceau S to distinguish lines. In order to test several concentrations of the sera, the membrane was cut into slices separating samples lines (dashed lines). The slices were incubated with a range of different sera concentrations. Before visualisation the slices were assembled back together. Note: to line 1 and 9 recombinant fragment 3 of the *X.laevis* HelQ had been added as a control.

6.2.6.2 HelQ immunodepletion from egg extract.

As described in section 6.2.5 the antibodies were raised based against denatured antigens. The fact that the sera recognize endogenous *X. laevis* HelQ in Western blot is promising but does not imply the antibodies will work for immunodepletion where recognition of the native antigen structure is necessary.

I therefore decided to test sera from two rabbits simultaneously in four rounds of immunodepletion. As a specificity control I performed a mock immunodepletion with pre-immunization sera. To monitor decreasing levels of a protein being depleted as well as to enhance depletion rate it is common practice to perform several rounds of incubation with antibody coated beads. After each round a sample of the extract is taken away and run on a gel alongside with pre-depletion extract (input). Beads after

depletion can be subjected to elution in SDS-Sample buffer and heat to release bound proteins that can be analysed by Western blot. Results are shown in Figure 6-10.

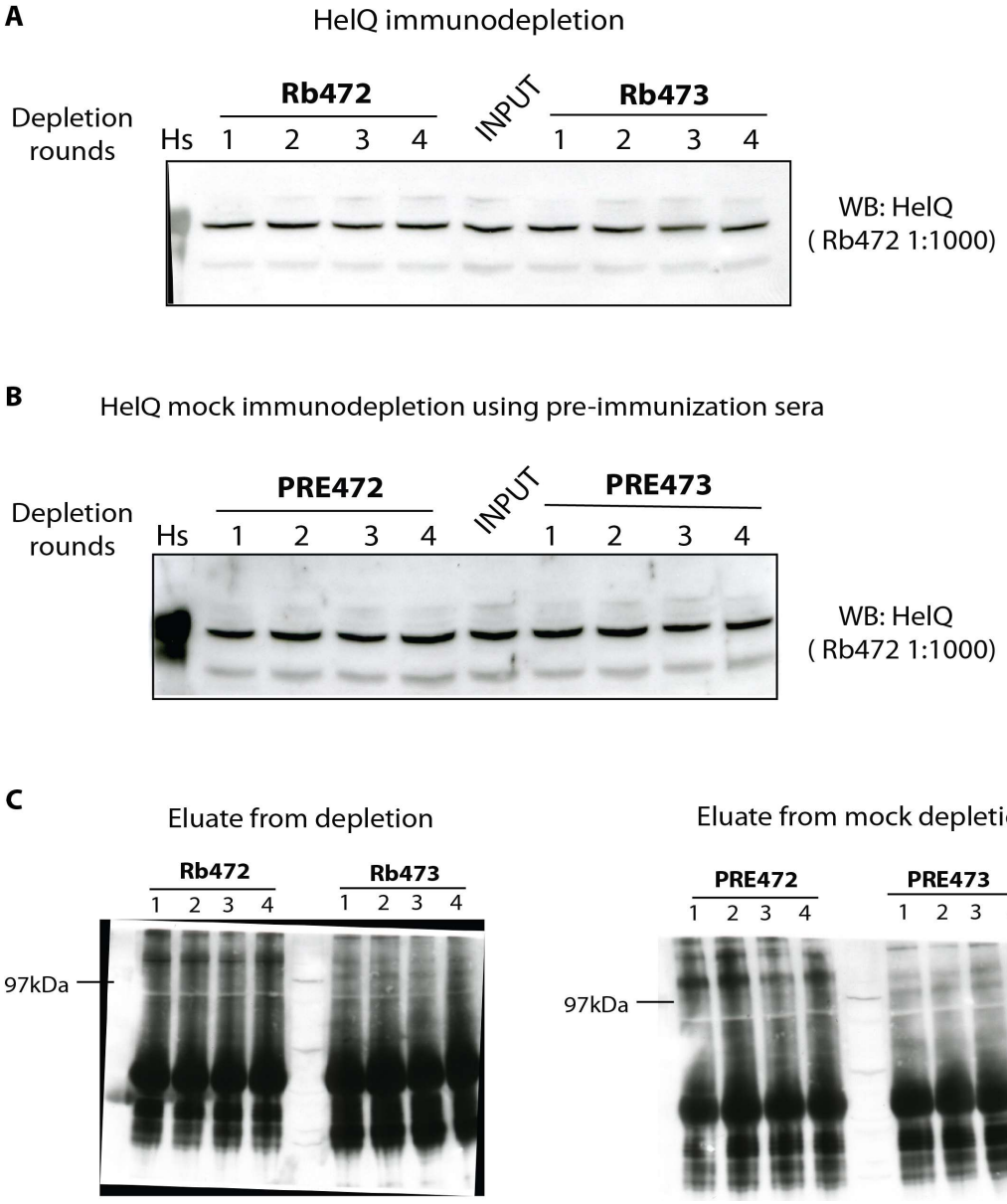


Figure 6-10 *X. laevis* HelQ immunodepletion testing.

(A) 80 μl of sepharose beads was incubated with 100μl of serum overnight at 4C. Four 30-minutes' rounds of immunodepletion with 15μl of antibodies-coated beads were performed at 4C. After each round of depletion a sample of egg extract was taken and subjected to western blot to analyse levels of HelQ in comparison with input. There is no changes in intensity of the bands corresponding to HelQ and migrating at the same size as human HelQ-FLAG from 293 cells (Hs) described in earlier chapters. **(B)** Control,

mock immunodepletion was performed similarly with pre-immunization sera. **(C)** Eluates from beads after each round of depletion show no band corresponding to HelQ size indicating that HelQ was not bound by the antibodies. Due to high background signal caused by the presence of released antibodies conclusions from this control are limited.

Unfortunately, the results of the immunodepletion experiment were negative. Neither sera bound native HelQ as indicated by unchanged levels of the protein after each round of depletion and as compared with the input (Figure 6-10 A). The main control in this experiment - mock immunodepletion - gave exactly the same results further indicating lack of binding. Finally, no bands corresponding to HelQ were detected in the eluate from beads.

6.2.7 Generation of *X. laevis* HelQ antibodies with antigenic peptides.

In parallel to expression of *X.laevis* HelQ domains, I was also taking an alternative approach by production of short peptides corresponding to *X. laevis HelQ* regions.
For the peptides to be a good antigenic material they need to meet several criteria. Best results are usually obtained with peptides 12-18 amino acids long. They should be based within regions relatively well exposed to the external environment so that antibodies can interact and bind to the sequence. Additionally, to avoid cross reactivity with other antigens they should be unique to the proteins of interest.
Similarly, as with domain expression, I started by looking at the simulation of *X. laevis* HelQ tertiary structure. I determined and excluded the regions that were likely to be hidden within the folded protein, and from the remaining fragments I chose several sequences that were then subjected to BLAST analysis against all *X. tropicalis* proteins. This approach is commonly taken by labs working with the *Xenopus* model as despite certain nucleotide sequence diversity between the two species, protein sequences are very well conserved.
As previously, we requested services of antibody production from Eurogentec. The sequences were run through their in-house developed immunogenicity prediction algorithms and only two were chosen for injection into rabbits (Table 6-2), Figure 6-11.

Peptide	Peptide location	Peptide sequence
XHELQ488-503	488-503	CDRIYEVDSNAENNLT
XHELQ684-699	684-699	DKELVHELISKPIENC

Table 6-2 *X.laevis* immunogenic peptides used for raising antibodies

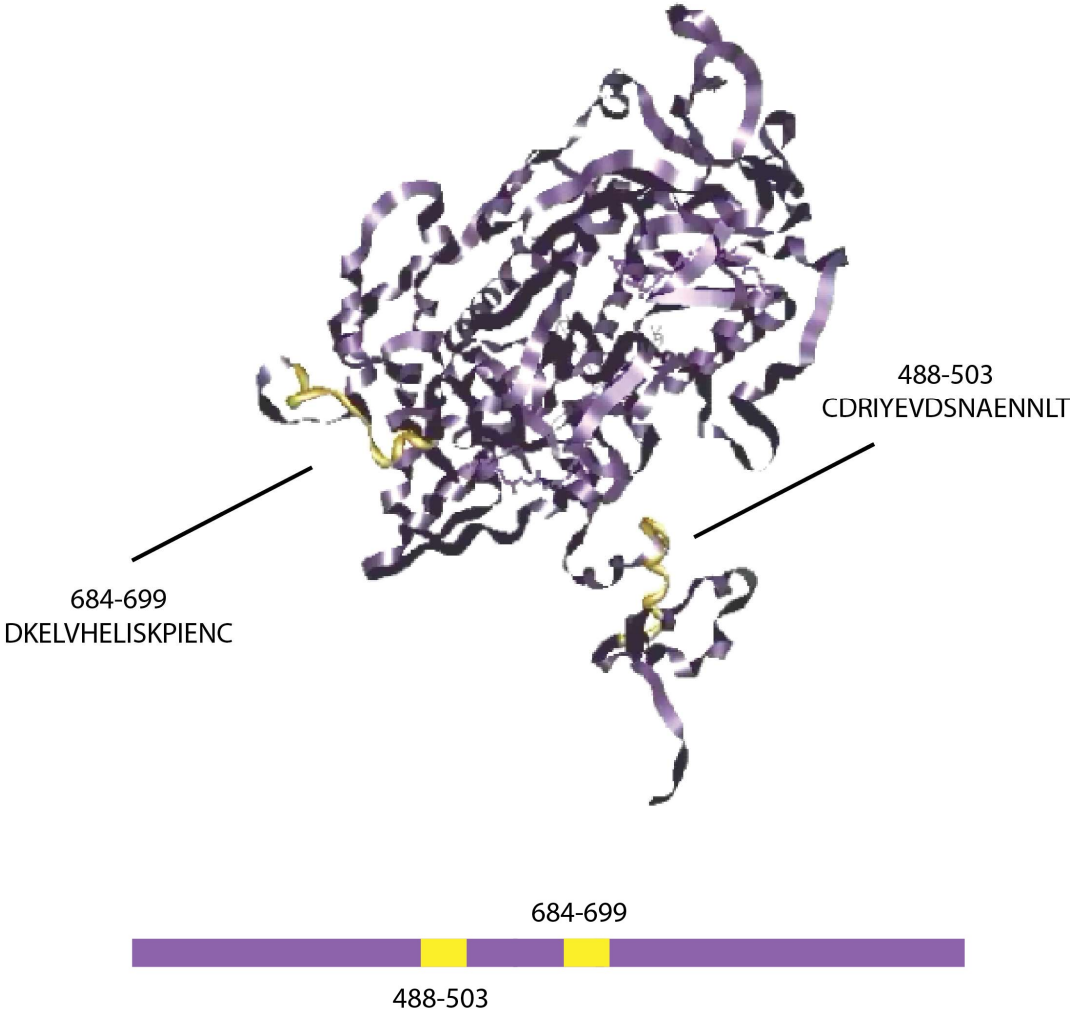
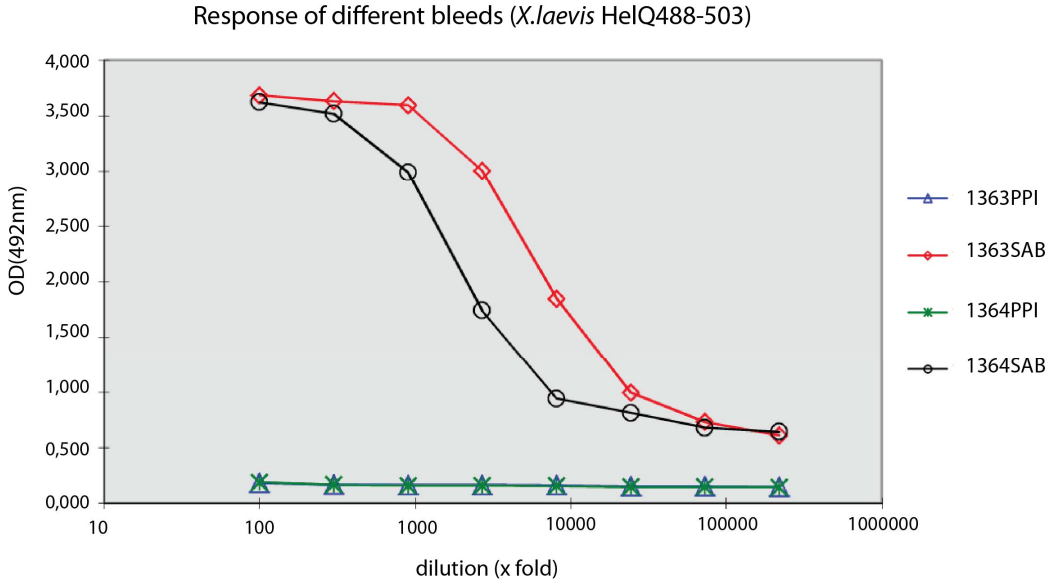


Figure 6-11 *X.laevis* HelQ structure prediction with two antigenic peptides mapped.
X.laevis HelQ tertiary structure was predicted as described in chapter 6.2.4. Two peptide sequences were chosen to be synthetized and injected into rabbits: **Peptide1**:488-503 (CDRIYEVDSNAENNLT) and **Peptide2**:684-699 (DKELVHELISKPIENC).

Peptides were synthetized by CRUK peptide synthesis facility. At Eurogentec, a mixture of both peptides was injected into two rabbits (Rb1363 and Rb1364). The antibodies were raised using a speed protocol where rabbits were injected with antigens on day 0, 7, 10,18 and bled on days 0, 21 and 28. Reactivity of pre-immunization and mature sera was assessed against the two antigenic peptides Figure 6-12).

A



B

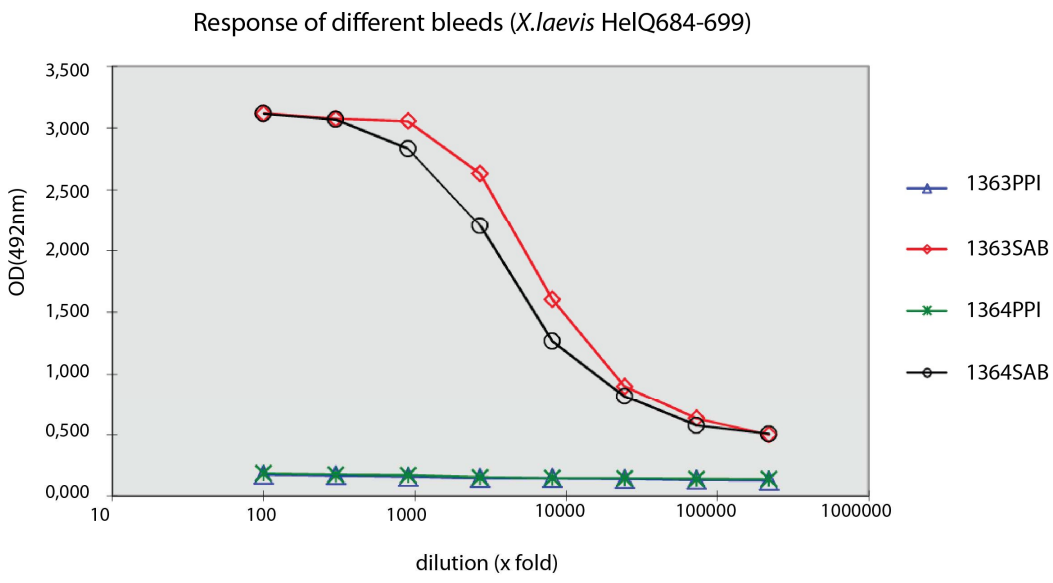


Figure 6-12 Response of different bleeds to the antigenic peptides as measured by ELISA

Pre-immune sera (PPI) and mature bleeding sera (SAB) of rabbits 1363 and 1364 were tested for reactivity against the peptides used as antigens: *X. laevis Hel*Q488:503 **(A)** and *X. laevis Hel*Q684:689 **(B)**.

For both peptides the curves fall into a sigmoidal shape from a saturation plateau at a dilution of 1:1000. The background is reached above dilution of 1:100000. The results suggest that both rabbits produced a mixed population of antibodies with high and low affinity. The fact that a subpopulation of the antibodies still recognize antigen at dilution beyond 1:100000 is a promising predictor. Overall it seems that antibodies raised by rabbit 1363 were more specific than those of rabbit 1364.

Serum of rabbit 1363 was chosen for affinity purification conducted by Eurogentec using peptide HelQ488:503 and we were provided with the remaining sera of rabbit 1363 and serum from rabbit 1364 as well as purified antibodies from rabbit 1363.

6.2.8 Anti-*X.laevis* HelQ peptide antibodies testing.

The purified antibodies (antibodies purified from serum of rabbit 1363 against peptide *X.laevis* HelQ488:503) were tested with egg extracts. Protein A Sepharose beads were coated with the purified antibodies and incubated with $200\mu l$ of egg extract for two hours. The beads were then thoroughly washed, and the bound fraction was eluted by boiling in SDS-Page sample buffer. Eluates were resolved on polyacrylamide gels and subjected to Western blot analysis using the same antibodies as for IP (Figure 6-13). Although no bands were detected in egg extracts and in the unbound fraction, a band of the expected size was present in the immunoprecipitates. As an additional reference I decided to run HEK293 HELQ-FLAG lysate and blot for HELQ to see the location of the band in relation to *X. laevis* HelQ. Human HELQ is a protein of around 124kDa whereas *X.laevis* HelQ is slightly smaller and should migrate around 118kDa. Comparing the bands on Figure 6-13 it seems very likely that the protein immunoprecipitated from egg extract is HelQ (it migrates at around 120kDa). What is interesting is the fact that in neither the egg extract nor unbound fraction was the band observed. One possible explanation is that although the egg extract is extremely protein-rich the amount loaded on gel ($3\mu l$)

is not enough for the antibodies to detect it. HELQ proved to be present at very low levels in human and mouse cells and one can expect similar expression level in *Xenopus*. It might be that the IP enriched the pool of HelQ and in 200 μl of the extract used for incubation with beads there was enough protein for detection.

Although the peptide-raised antibodies showed promise additional validation and optimization steps are required before they can be used for immunodepletion.

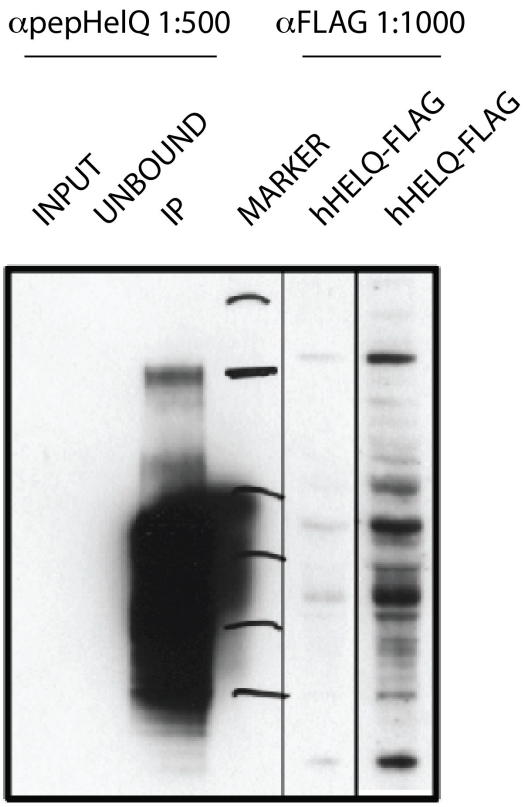


Figure 6-13 *X.laevis* HelQ IP from egg extract with antibodies raised with immunogenic peptides.

Studies on *Xenopus* HelQ were halted at this stage due to time limitations. In order to synchronize our publication with another group I prioritized finalising experiments required to present a consistent story of the role of HELQ in DNA repair and could not finish the *Xenopus* studies in the scope of PhD course (Adelman et al., 2013; Luebben et al., 2013; Takata et al., 2013).

6.3 Discussion

In order to compliment molecular biology and mouse studies in determining the role of HELQ in replication-dependant ICL repair, I decided to set up a *X. laevis* egg extract protein depletion system. I started with generating crucial tools – expression of *X. laevis* HelQ and raising specific antibodies able to efficiently deplete the protein from egg extract. As the sequence of *X.laevis* HelQ was unknown at the time I performed a number of bioinformatic analysis using *X.tropicalis* sequences as a template. I successfully identified two E.S.T. covering the 1-678 amino acid region as well as 318 amino acid sequence of the C-terminus. By analysing alignment of the obtained sequence with human and *X.tropicalis* HelQ cDNA I was able to determine a region of about 60 amino acids not covered by the ESTs. To obtain a full length HelQ sequence I then designed PCR primers flanking the region of interest, prepared a cDNA library from mRNA isolated from egg extract and amplified the missing sequence that I then engineered into the full gene.

To be able to obtain large quantities of properly folded *X.laevis* HelQ with posttranslational modifications I decided to use insect cells expression system. I codon optimized the sequence and requested a *de-novo* synthesis of the gene from an external provider. Optimized cDNA was then provided to the CRUK Protein Expression Unit where it was cloned into relevant vectors to produce bacmids.

Production of full length HelQ in High5 cells was unsuccessful due to solubility and degradation problems. Anti HIS-tag western blot analysis of the cell lysate and Ni-NTA-beads eluate showed that HelQ is exclusively found in the insoluble fraction and that it is degraded to a product of around 70kDa. Band of the full-length protein size was almost invisible on Western blot.

To circumvent solubility and degradation problems of full length HelQ I performed bioinformatic analysis of tertiary structure and selected three fragments to be expressed in *E.coli*. From a number of attempts with different tagging and purification protocols I obtained sufficient quantity of pure HelQ fragment corresponding to L708-A1019 amino acids. Antigenic HelQ fragment 3 was sent to Eurogentec that injected it into two rabbits to raise polyclonal antibodies and provided me with sera from the immunised animals. I then showed by western blot that sera from both rabbits recognize a protein of appropriate size in *X.laevis* egg extract as well as the recombinant HelQ fragment 3. Unfortunately, the sera did not recognize native HelQ

as I was not able to immunodeplete the protein from the extract (as shown by unchanged levels of the protein after each round of depletion and as compared with the input levels)

In parallel I raised additional antibodies based on 2 short peptide sequences chosen in a bioinformatic analysis of the tertiary structure of HelQ. In a preliminary immunodepletion experiment the peptides-based antibodies gave promising results, but further work needs to be done in order to properly verify their ability to immunodeplete *X.laevis* HelQ from egg extract.

In summary, I successfully used bioinformatic tools to determine *X.laevis* HelQ sequence. I then expressed recombinant fragments of the frog protein that allowed me to generate preliminary tools for immunodepletion experiments. Both sets of antibodies I raised seemed to specifically recognize the antigen, but more validation and optimization is needed. Unfortunately, due to the time constraints and prioritization experiments included in the publication I did not manage to finish my work with the *X. laevis* immunodepletion system. However, the tools I generated show promise and warrant further investigation.

Chapter 7. Discussion

Maintenance of genome stability is a critical condition that ensures that daughter cells acquire an accurate copy of the genetic information from the parental cell. DNA replication stress that arises from blocked replication forks can be a major challenge to genome integrity. Cells have therefore developed complex mechanisms to detect and deal with the replication-associated DNA damage. Intra-S-phase ATR checkpoint, FA pathway and RAD51 paralog BCDX2 complex together constitute key components of the replication stress response system that is essential to sense, repair and restart damaged forks. Here, I establish that HELQ plays a pivotal role in the replication stress response, interfacing between the central FANCD2/FANCI heterodimer and the downstream RAD51 paralog BCDX2 complex, which is critical for HR repair.

HELQ and the Fanconi Anemia pathway

Analysis of an HELQ deficient mouse model developed by Carrie Adelman in the lab showed extensive phenotypic similarities to FA mice, including hypogonadism, subfertility as well as tumour predisposition, hypersensitivity and formation of radial chromosomes upon treatment with ICL-inducing agents (Parmar et al., 2009). Some interesting differences with FA mouse models were also observed though; *HelQ* mutants were born in normal Mendelian ratio and showed no growth retardation or microphthalmia – characteristic developmental defects observed in some subtypes of FA mice. Strikingly, bone marrow hematopoietic stem and progenitor cell proliferation defects, a hallmark of Fanconi Anemia and FA mouse models, were also not detected in HELQ deficient mice. Cells derived from *HelQ*$^{\Delta C/\Delta C}$ mice displayed only moderate sensitivity to MMC compared with *FancD2*$^{-/-}$ cells, and conversely, HELQ deficient cells were more sensitive to CPT than FANCD2 deficient cells.

Given the phenotypic similarities between HELQ and FANCD2 deficient mice and the fact that I found that FANCD2 and FANCI were HELQ's major interacting partner it was expected that these proteins perform a cooperative role in ICL repair. Surprisingly, we found that HelQ$^{\Delta C/\Delta C}$*;FancD2*$^{-/-}$ double mutant mice and cell lines derived from such animals exhibit additive phenotypes in some assays, including hypogonadism, germ cell attrition, and chromosomal aberrations and cellular

sensitivity to ICL-inducing agents. This observation suggests that HELQ acts independently from FANCD2 in a parallel ICL repair pathway. Considering the complexity and mutual interdependence of DNA repair pathways, however, it cannot be completely ruled out that HELQ and FANCD2 perform a cooperative role in repair of certain kind of replication-dependant damage but act independently in other processes.

Phenotypic additivity of *HelQ*$^{\Delta C/\Delta C}$ and *FancD2*$^{-/-}$ double mutants is seemingly counterintuitive to the observation that FAND2 and FANCI coprecipitate with HELQ, which would suggest cooperative or antagonistic roles of both proteins. It should be noted however, that there are well documented examples of mutation additivity between genes clearly cooperating within the same pathway. A good example of such phenomenon are Sgs1 and Exo1, which are involved in DSB end resection, yet *sgs1;exo1* double mutants manifest much more severe resection defects than either of the single mutants (Mimitou and Symington, 2008; Zhu et al., 2008). Furthermore, mutated members of the Spn family of Drosophila (HELQ and RAD51 paralogs also belong to this group) show additivity in terms of meiotic DSB repair defects (Abdu et al., 2003; McCaffrey et al., 2006)

We believe that a number of observations support the concept that HELQ functions at a late step of ICL repair to influence downstream HR events. These include the following: physical interaction between HELQ and the BCDX2 complex of RAD51 paralogs, which function to promote RAD51 loading and HR and has been implicated in FA; secondly, the impaired unloading of the ssDNA binding protein RPA32 and HR strand exchange protein RAD51 in *HelQ* mutants; and the sensitivity of *HelQ*$^{\Delta C/\Delta C}$ cells to PARP inhibition, which has been shown to be synthetic lethal in combination with HR mutants.

Further evidence supporting a downstream role for HELQ in HR is the observation that HELQ deficiency shares similarities with BRCA2 (FANCD1) and PALB2 (FANCN) deficiency- two factors which function at late steps in the FA pathway during HR. No obvious hematopoietic stem or progenitor cell defects in *HelQ* mutants is consistent with the lack of bone marrow dysfunction reported for FA patients with *PALB2* mutations (Reid et al., 2007), and the reduced frequency with which this phenotype is observed in FA patients with *BRCA2* mutations (Hirsch et al., 2004; Offit et al., 2003). Furthermore, *BRCA2* and *PALB2* mutations are associated with familial ovarian cancer risk and we have observed a similar link between HELQ

deficiency and ovarian cancer in mice (Dansonka-Mieszkowska et al., 2010; Petrucelli et al., 2010). Additionally, unaffected monoubiquitination of FANCD2 indicates that early steps of FA-pathway activation are intact in *HelQ* mutants. Finally, the observation that RAD51 and γH2AX persisted on chromatin and remained present in repair foci at later time points in HELQ-deficient mouse and human cells clearly indicates that the defect in the absence of HELQ occurs at a late step - downstream of RAD51 recruitment to damaged replication forks. All of these findings together are consistent with HELQ functioning at a late stage of ICL repair during HR.

It is worth noting that many of the phenotypes observed in HELQ deficient mice also bear a striking resemblance to the *Mcm8* and *Mcm9* mutant mice and cell lines. MCM8 and MCM9 are two factors implicated in the DNA damage response to DSBs and replication fork blocks (Lutzmann et al., 2012; Nishimura et al., 2012). Similar to *HelQ*$^{\Delta C/\Delta C}$ mice, *Mcm8*$^{-/-}$ and *Mcm9*$^{-/-}$ mice exhibit subfertility, gonadal atrophy and ovarian tumours, including the rare sex cord stromal tumours observed in HELQ and FANCF deficient mice (Bakker et al., 2012). *Mcm8* and *Mcm9* deficient cell lines display indices of genome instability, replication fork instability, hypersensitivity to DNA replication stress, and a defect in recruitment and retention of HR factors such as MRE11, RPA32 and RAD51 in response to replication stress. It appears that the MCM8-MCM9 complex promotes RAD51 recruitment at DNA damage sites to facilitate homologous recombination (Park et al., 2013). Despite the striking similarities it would appear that MCM8/MCM9 may act independently of HELQ as these proteins were not detected in our mass spectrometry analysis.

HELQ and the RAD51 Paralog BCDX2 complex

The RAD51 paralogs are essential for execution of HR at damaged replication forks, but molecular and biochemical details of how the RAD51 paralogs operate in HR had long been relatively unclear. This is in part attributable to the fact that the purified proteins from vertebrates have been intractable for many *in vitro* biochemical studies. In a recent work from our laboratory, Martin Taylor discovered that a RAD51 paralog complex from nematode *C. elegans*, RFS-1/RIP-1, plays a crucial role downstream of filament assembly by binding to- and remodelling RAD-51-ssDNA filaments to a conformation more competent to undergo strand exchange. RFS-1/RIP-1 acts by abolishing RAD-51 dissociation from ssDNA by capping the 5' end of RAD-51-ssDNA

filaments and that filament capping and stabilization are dependent on nucleotide binding, but not hydrolysis by RFS-1/RIP-1. This insightful work provided a definite data describing the mechanism of RAD51 filament remodelling by RAD51 paralogs. Apart from the RAD51 paralogs' well established relationship with the RAD51 recombinase (Suwaki et al., 2011), and a number of candidate interacting partners identified in a previous MS screening study, the isolation of RAD51 paralog BCDX2 complex in immune precipitates with HELQ is one of the only three bona fide RAD51 paralog interacting partner with functional data (Amendola et al., 2017; Braybrooke et al., 2003; Rajesh, Gruver, Basrur, & Pittman, 2009). Discovery of this interaction offers a plausible explanation for the phenotypic similarities observed for HELQ (Mus301/Spn-C) and RAD51 paralog fly-deficient mutants (Abdu, González-Reyes, Ghabrial, & Schüpbach, 2003; Ghabrial & Schüpbach, 1999; Laurencon et al., 2004; McCaffrey, St Johnston, & González-Reyes, 2006).

In my studies I established that HELQ specifically interacts with BCDX2 complex of the RAD51 paralogs but not with the CX3 complex, I was not able however to identify the specific site or complex's component that mediate this interaction. I believe this interaction is crucial as underlined by my finding that loss of HELQ leads to reduced ability to promote chromatin loading and retention of the key HR proteins such as RAD51 and RPA32 in replication stress response. This is reminiscent of the moderate reduction in formation of RAD-51 foci in response to ICLs in helQ-1 mutant worms (Muzzini, Plevani, Boulton, Cassata, & Marini, 2008). Loss of RAD51 paralogs in various organisms has been previously shown to compromise RAD51 loading at damaged replication forks albeit to a greater extent than observed in the absence of HELQ (Ward, Barber, Petalcorin, Yanowitz, & Boulton, 2007; Yonetani et al., 2005). Although I showed that the interaction between the two complexes is a direct one, a question that remains largely unanswered is which components are engaged in the recognition and binding. Attempts to decipher HELQ – BCDX2 interactions proved to be challenging due to complexity of the studied model and expression/solubility problems. Although RAD51D seems the most promising candidate to confer this interaction, more experiments are needed to confirm that observation.

An important question that should be asked is whether HELQ may play multiply roles during ICL repair. It is recruited to chromatin at early stage of the repair process in

an ATR-dependent manner – one can speculate it might also be involved in earlier processes of ICL repair.

Biochemical and bioinformatic studies of HELQ have implicated it in unwinding duplex DNA from a 3' ssDNA overhang (Fujikane et al., 2005; Guy & Bolt, 2005b; Li et al., 2008), removing lagging strands from model replication forks (Li et al., 2008; Tafel, Wu, & McHugh, 2011) migration of Holliday junctions (Fujikane et al., 2005) , replication fork regression (Li et al., 2008), and translocation (Richards et al., 2008; Ward et al., 2010). Such a broad spectrum of potential enzymatic activity could be employed at various stages of repair – from assisting in unwinding lagging strands at blocked forks, permitting the timely and appropriate incision of the obstructing lesion; post lesion incision, to promoting formation of a 3' ssDNA end substrate for loading of RAD51 assisted by the BCDX2 paralogs; and/or, removing post-strand exchange RAD51 filaments, allowing repair synthesis to ensue.

Additionally, it cannot be ruled out that apart from a purely enzymatic role, HELQ may also act as structural platform, physically linking the FA core components with downstream HR factors in order to temporally coordinate these events. It might therefore be possible that HELQ deficiency reflects defects at several distinct steps in the repair and restart of blocked replication forks, which has also been observed for the RAD51 paralogs (Suwaki et al., 2011).

HELQ in ovarian cancer and human disease

One of the most important and interesting results of our studies was the finding that HelQ deficient mice are predisposed to develop ovarian tumours. This observation may also help to explain the prevalence of non-synonymous variants in *HELQ*, which are associated with upper aerodigestive tract carcinomas, as well as the fact that HELQ variants have been shown to be associated with early onset of menopause (McKay et al., 2011; Stolk et al., 2012).

The link between HR-defect (or HR-insufficiency) and carcinomas of ovarian and breast origins is well documented, however the pathomechanism remains unclear.

Considering our findings describing functional interplay between HELQ and FA/HR factors and the fact that HELQ-deficiency leads to increased risk of cancer development is not at all unexpected. In fact, mutations in almost all of HELQ interacting partners found in my MS analysis have well documented links with elevated cancer incidence (Durocher et al., 2006; Heikkinen, Mansikka, Karppinen,

Rapakko, & Winqvist, 2005; Hilbers et al., 2012; Houghtaling et al., 2003; Johnson, Healey, Khanna, & Chenevix-Trench, 2011; Loveday et al., 2011a; Parmar, D'Andrea, & Niedernhofer, 2009; Somyajit, Subramanya, & Nagaraju, 2010; Vachon et al., 2012). Interestingly, *FancF*$^{-/-}$ mice develop granulosa cell tumors, which are very reminiscent of those we found in *HelQ*$^{\Delta C/\Delta C}$ females.

As previously mentioned *HELQ* is located in region 4q which has been linked with development of various type of cancers. Genetic alterations of this region have been documented in hepatocellular cancer, ductal pancreatic adenocarcinoma oral squamous cell cancer as well as gastric cancer (Fan et al., 2012; Homayounfar et al., 2013; Luebke et al., 2012; Ribeiro et al., 2014). Interestingly deletions within 4q21 region is one of the most common type of genomic alterations in ovarian cancer, found in as many as 54% of ovarian cancer cases (Kim et al., 2003). The ovarian cancer phenotype in *HelQ* mutant mice together with HELQ interaction with the BCDX2 complex and location in region implicated in ovarian carcinogenesis therefore make *HELQ* a compelling candidate gene for screening in familial cases of ovarian and breast cancer.

To date, several studies have looked for *HELQ* aberrations in the context of elevated risk of ovarian and breast cancer. In a study from Chen's lab *HELQ* coding sequence was analysed by Next Generation Sequencing in a group of 192 women with secondary amenorrhoea – an indicator of premature ovarian failure. Six known single-nucleotide polymorphisms were identified in both control and assessed group, but no mutation was found. The group used in the study was selected based on the premature ovarian failure criterion which is highly reminiscent of the HelQ mutant mice (subfertility, dysgenesis/atrophy of ovaries). Analysis of only 192 individuals however only allows to identify factors of high penetrance and was insufficient to draw strong conclusions (W. Wang et al., 2015).

In another study on the Finish population, the authors screened *HELQ* coding region for germline variation in 185 breast or ovarian cancer families and ran analysis of haplotypes in 1517 breast cancer cases, 308 ovarian cancer cases, and 1234 population controls. Only one likely pathogenic missense variant (substitution 1309A>G) was found in one case despite subsequent genotyping of additional 332 breast and ovarian cancer patients with familial history. Additionally, there was no

clear differences in the haplotype distribution between ovarian and/or breast cancer group and controls. Similarly, as with the Chinese study, larger groups are needed to assess variants of lower penetrance, but this study provided further evidence that *HELQ* is not a major breast or ovarian cancer susceptibility gene, at least in the studied population (the Finish population is thought to be relatively homogenous).

Given the link between HELQ and RAD51 paralogs with hereditary breast and ovarian cancer risk another interesting question is whether mutations and copy number variation of those genes are linked with sporadic cancers in human. Our analysis of available databases (mainly TCGA) led to the observation that *HELQ* copy number loss affects a significant subset of sporadic ovarian cancer samples. Additionally, the analyses of germline *RAD51C* and *RAD51D* mutations in familial ovarian cancer showed that loss of heterozygosity (LOH) affects most but not all *RAD51*-mutated tumours and just over half of tumour bearing RAD51D mutations (Loveday et al., 2011c; Meindl et al., 2010; Osher et al., 2012; Pelttari et al., 2011; Wickramanyake et al., 2012). This observation suggests that tumour formation is not fully dependent on losing the remaining wildtype allele. Consistently, our studies on human ovarian cancer samples suggest that hemizygous loss of *HELQ* may be sufficient to drive tumour formation. Although the concept in which loss of just one copy of an onco-suppressor is sufficient to propel cancer development is controversial, a growing body of evidence supports this view. A study from the Elledge lab from 2012 revealed recurrent loss of regions harbouring high numbers of tumour suppressor genes and low numbers of essential genes across many tumour samples and types. The authors proposed that hemizygous loss of such clusters of genes could lead to cumulative selective advantage to cells that would eventually drive tumorigenesis (Solimini et al., 2012). The problem of hemizygous loss of HR-promoting genes (such as *HELQ*) in the context of cancer development has become particularly important subject of discussion in 2014 when the first PARP inhibitor (PARPi) became clinically available (Lee, Ledermann, & Kohn, 2014). Accurate patient selection strategies for therapies based on PARPi and full apprehension of their long-term clinical benefits and side effects requires a thorough understanding of the role of hemizygous loss of HR-promoting genes.

PARP inhibition is a relatively new approach to cancer therapy that emerged as a result of successful translational research in the DNA repair field. Olaparib (Lynparza) was the first-in-class molecule registered in 2014 for maintenance treatment in patients with high grade serous ovarian cancer positive for deleterious mutation in *BRCA1* or *BRCA2* genes. Since then new molecules and broader indication have become available and the whole DNA response inhibitors class is gaining momentum. Therapeutic effect of PARP inhibition in BRCA-deficient tumor is based on synthetic lethality, a phenomenon in which otherwise viable cells harbouring defect in DSBs repair (caused by mutation in BRCA1 or BRCA2) become lethal upon blocking SSBs repair by PARPi (Ashworth, 2008). This concept have been validated in-vivo (Bryant et al., 2005; Evers, Helleday, & Jonkers, 2010; Farmer et al., 2005; Farmer H, McCabe N, Lord CJ, Tutt AN, Johnson DA, Richardson TB, Santarosa M, Dillon KJ, Hickson I, Knights C, Martin NM, Jackson SP, Smith GC, 2055) and in clinics (Fong, Boss, Yap, & Tutt, 2009).

Our results revealed that HELQ-deficient cells are sensitive to olaparib, potentially opening a route to therapeutic approach for patients with *HELQ*-mutated tumours. Interestingly, despite intensive sequencing efforts only a few *HELQ*-mutated patients have been identified suggesting a very low penetrance. On a poster presented at AACR Ovarian Meeting in 2015 in Orlando, Elizabeth Swisher from University of Washington School of Medicine showed interesting results of rucaparib (potent PARPi) treatment of ovarian cancer patients with mutations in various DNA-repair genes including *HELQ*. They identified a patient with deleterious germline mutation truncating the protein product. Interestingly however, the patient had no clinical response to rucaparib. Clearly, to draw conclusions on potential clinical benefits of PARPi treatment of HELQ-mutated tumours assessment of a much bigger dataset is needed. Low penetrance of the *HELQ* mutations can hinder these efforts but with growing availability of Next Generation Sequencing in clinics, even low penetrance mutations gain clinical significance. A great example is RECQL that has been identified as breast cancer susceptibility gene independently in two populations with family history, with very low incidence (7 mutation carriers in a group of 1013 and 30 in a group of 13136) (Cybulski et al., 2015).

In an alternative PARPi patients selection strategy, rather than evaluating genes promoting HR, long regions of genome are sequenced in search of the characteristic pattern left by deficiency of HR, so called genomic scar (Abkevich et al., 2012). Most commonly loss of heterozygosity at SNP is looked at and analysed to assess the level of HR-deficiency. In the AACR rucaparib poster, Swisher et al. presented results of such analysis in patients with germline mutations in various DNA response genes, including *HELQ* and its interacting partners. Interestingly, the level of LOH (used as surrogate of HR defect) was qualified as “high” for *RAD51B*, *RAD51C*, *RAD51D* and *HELQ*.

Taken together, our study established *HELQ* helicase as an important DNA response player anchored within the molecular network of replication dependent DBSs repair. To perform its function *HELQ* operates in parallel to the Fanconi Anemia pathway by promoting HR-dependant repair. Additionally, our data positions HELQ as an ovarian cancer susceptibility gene with low penetrance effect. This potentially opens an exciting therapeutic approach based on PARP inhibition. To verify clinical implication of our data, more studies are needed, in particular focusing on germline or somatic *HELQ*-mutations cancer patients

Reference List

Abdu, U., Gonzalez-Reyes, A., Ghabrial, A., and Schupbach, T. (2003). The Drosophila spn-D gene encodes a RAD51C-like protein that is required exclusively during meiosis. Genetics *165*, 197-204.
Adelman, C.A., Lolo, R.L., Birkbak, N.J., Murina, O., Matsuzaki, K., Horejsi, Z., Parmar, K., Borel, V., Skehel, J.M., Stamp, G.*, et al.* (2013). HELQ promotes RAD51 paralogue-dependent repair to avert germ cell loss and tumorigenesis. Nature *502*, 381-384.
Alcindor, T., and Beauger, N. (2011). Oxaliplatin: a review in the era of molecularly targeted therapy. In Curr Oncol, pp. 18-25.
Ali, A.M., Pradhan, A., Singh, T.R., Du, C., Li, J., Wahengbam, K., Grassman, E., Auerbach, A.D., Pang, Q., and Meetei, A.R. (2012). FAAP20: a novel ubiquitin-binding FA nuclear core-complex protein required for functional integrity of the FA-BRCA DNA repair pathway. Blood *119*, 3285-3294.
Ali, I., Wani, W.A., Saleem, K., and Haque, A. (2013). Platinum compounds: a hope for future cancer chemotherapy. In Anticancer Agents Med Chem, pp. 296-306.
Alpi, A.F., Pace, P.E., Babu, M.M., and Patel, K.J. (2008). Mechanistic insight into site-restricted monoubiquitination of FANCD2 by Ube2t, FANCL, and FANCI. Mol Cell *32*, 767-777.
Alsop, K., Fereday, S., Meldrum, C., deFazio, A., Emmanuel, C., George, J., Dobrovic, A., Birrer, M.J., Webb, P.M., Stewart, C.*, et al.* (2012). BRCA mutation frequency and patterns of treatment response in BRCA mutation-positive women with ovarian cancer: a report from the Australian Ovarian Cancer Study Group. J Clin Oncol *30*, 2654-2663.
Anantha, R.W., Vassin, V.M., and Borowiec, J.A. (2007). Sequential and Synergistic Modification of Human RPA Stimulates Chromosomal DNA Repair. In Journal of Biological Chemistry, pp. 35910-35923.
Andersen, S.L., Bergstralh, D.T., Kohl, K.P., LaRocque, J.R., Moore, C.B., and Sekelsky, J. (2009). Drosophila MUS312 and the vertebrate ortholog BTBD12 interact with DNA structure-specific endonucleases in DNA repair and recombination. Mol Cell *35*, 128-135.
Andersson, B.S., Sadeghi, T., Siciliano, M.J., Legerski, R., and Murray, D. (1996). Nucleotide excision repair genes as determinants of cellular sensitivity to cyclophosphamide analogs. In Cancer Chemother Pharmacol, pp. 406-416.
Aquilina, G., Ceccotti, S., Martinelli, S., Hampson, R., and Bignami, M. (1998). N-(2-chloroethyl)-N'-cyclohexyl-N-nitrosourea sensitivity in mismatch repair-defective human cells. Cancer Res *58*, 135-141.
Auerbach, A.D. (2009). Fanconi anemia and its diagnosis. Mutat Res *668*, 4-10.
Bakker, S.T., van de Vrugt, H.J., Visser, J.A., Delzenne-Goette, E., van der Wal, A., Berns, M.A., van de Ven, M., Oostra, A.B., de Vries, S., Kramer, P.*, et al.* (2012). Fancf-deficient mice are prone to develop ovarian tumours. J Pathol *226*, 28-39.
Barber, L.J., Ward, T.A., Hartley, J.A., and McHugh, P.J. (2005a). DNA interstrand cross-link repair in the Saccharomyces cerevisiae cell cycle: overlapping roles for PSO2 (SNM1) with MutS factors and EXO1 during S phase. In Molecular and Cellular Biology, pp. 2297-2309.
Barber, L.J., Ward, T.A., Hartley, J.A., and McHugh, P.J. (2005b). DNA interstrand cross-link repair in the Saccharomyces cerevisiae cell cycle: overlapping roles for PSO2 (SNM1) with MutS factors and EXO1 during S phase. Mol Cell Biol *25*, 2297-2309.

Bell, D., Berchuck, A., Birrer, M., Chien, J., Cramer, D.W., Dao, F., Dhir, R., Disaia, P., Gabra, H., Glenn, P.*, et al.* (2011). Integrated genomic analyses of ovarian carcinoma. In Nature, pp. 609-615.
Ben-Yehoyada, M., Wang, L.C., Kozekov, I.D., Rizzo, C.J., Gottesman, M.E., and Gautier, J. (2009). Checkpoint signaling from a single DNA interstrand crosslink. Mol Cell *35*, 704-715.
Berardini, M., Foster, P.L., and Loechler, E.L. (1999). DNA polymerase II (polB) is involved in a new DNA repair pathway for DNA interstrand cross-links in Escherichia coli. In J Bacteriol, pp. 2878-2882.
Berardini, M., Mackay, W., and Loechler, E.L. (1997). Evidence for a recombination-independent pathway for the repair of DNA interstrand cross-links based on a site-specific study with nitrogen mustard. In Biochemistry, pp. 3506-3513.
Bessho, T., Mu, D., and Sancar, A. (1997). Initiation of DNA interstrand cross-link repair in humans: the nucleotide excision repair system makes dual incisions 5' to the cross-linked base and removes a 22- to 28-nucleotide-long damage-free strand. Mol Cell Biol *17*, 6822-6830.
Bhagwat, N.R., Roginskaya, V.Y., Acquafondata, M.B., Dhir, R., Wood, R.D., and Niedernhofer, L.J. (2009). Immunodetection of DNA repair endonuclease ERCC1-XPF in human tissue. Cancer Res *69*, 6831-6838.
Bhattacharyya, A., Ear, U.S., Koller, B.H., Weichselbaum, R.R., and Bishop, D.K. (2000). The breast cancer susceptibility gene BRCA1 is required for subnuclear assembly of Rad51 and survival following treatment with the DNA cross-linking agent cisplatin. J Biol Chem *275*, 23899-23903.
Bi, X., Barkley, L.R., Slater, D.M., Tateishi, S., Yamaizumi, M., Ohmori, H., and Vaziri, C. (2006). Rad18 regulates DNA polymerase kappa and is required for recovery from S-phase checkpoint-mediated arrest. Mol Cell Biol *26*, 3527-3540.
Bogliolo, M., Schuster, B., Stoepker, C., Derkunt, B., Su, Y., Raams, A., Trujillo, J.P., Minguillon, J., Ramirez, M.J., Pujol, R.*, et al.* (2013). Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. Am J Hum Genet *92*, 800-806.
Boyd, J.B., Golino, M.D., Shaw, K.E., Osgood, C.J., and Green, M.M. (1981). Third-chromosome mutagen-sensitive mutants of Drosophila melanogaster. Genetics *97*, 607-623.
Buttner, K., Nehring, S., and Hopfner, K.P. (2007). Structural basis for DNA duplex separation by a superfamily-2 helicase. Nat Struct Mol Biol *14*, 647-652.
Cartwright, R., Tambini, C.E., Simpson, P.J., and Thacker, J. (1998). The XRCC2 DNA repair gene from human and mouse encodes a novel member of the recA/RAD51 family. In Nucleic Acids Res, pp. 3084-3089.
Castor, D., Nair, N., Declais, A.C., Lachaud, C., Toth, R., Macartney, T.J., Lilley, D.M., Arthur, J.S., and Rouse, J. (2013). Cooperative control of holliday junction resolution and DNA repair by the SLX1 and MUS81-EME1 nucleases. Mol Cell *52*, 221-233.
Cattell, E., Sengerova, B., and McHugh, P.J. (2010). The SNM1/Pso2 family of ICL repair nucleases: from yeast to man. Environ Mol Mutagen *51*, 635-645.
Ciccia, A., Ling, C., Coulthard, R., Yan, Z., Xue, Y., Meetei, A.R., Laghmani el, H., Joenje, H., McDonald, N., de Winter, J.P.*, et al.* (2007). Identification of FAAP24, a Fanconi anemia core complex protein that interacts with FANCM. Mol Cell *25*, 331-343.
Clauson, C., Scharer, O.D., and Niedernhofer, L. (2013a). Advances in Understanding the Complex Mechanisms of DNA Interstrand Cross-Link Repair. In Cold Spring Harbor Perspectives in Biology, pp. a012732-a012732.
Clauson, C., Schärer, O.D., and Niedernhofer, L. (2013b). Advances in understanding the complex mechanisms of DNA interstrand cross-link repair. In Cold Spring Harbor Perspectives in Biology, pp. a012732.

Cleaver, J.E., Lam, E.T., and Revet, I. (2009). Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. In Nat Rev Genet, pp. 756-768.
Cohn, M.A., Kee, Y., Haas, W., Gygi, S.P., and D'Andrea, A.D. (2009). UAF1 is a subunit of multiple deubiquitinating enzyme complexes. J Biol Chem *284*, 5343-5351.
Cole, A.R., Lewis, L.P., and Walden, H. (2010). The structure of the catalytic subunit FANCL of the Fanconi anemia core complex. Nat Struct Mol Biol *17*, 294-298.
Cole, R.S. (1973). Repair of DNA containing interstrand crosslinks in Escherichia coli: sequential excision and recombination. In Proceedings of the National Academy of Sciences of the United States of America, pp. 1064-1068.
Collado, M., Gil, J., Efeyan, A., Guerra, C., Schuhmacher, A.J., Barradas, M., Benguría, A., Zaballos, A., Flores, J.M., Barbacid, M.*, et al.* (2005). Tumour biology: senescence in premalignant tumours. In Nature, pp. 642.
Collis, S.J., Barber, L.J., Ward, J.D., Martin, J.S., and Boulton, S.J. (2006). C. elegans FANCD2 responds to replication stress and functions in interstrand cross-link repair. DNA Repair (Amst) *5*, 1398-1406.
Constantin, N., Dzantiev, L., Kadyrov, F.A., and Modrich, P. (2005). Human mismatch repair: reconstitution of a nick-directed bidirectional reaction. In J Biol Chem, pp. 39752-39761.
Coste, F., Malinge, J.M., Serre, L., Shepard, W., Roth, M., Leng, M., and Zelwer, C. (1999). Crystal structure of a double-stranded DNA containing a cisplatin interstrand cross-link at 1.63 A resolution: hydration at the platinated site. In Nucleic Acids Research, pp. 1837-1846.
Crossan, G.P., and Patel, K.J. (2012). The Fanconi anaemia pathway orchestrates incisions at sites of crosslinked DNA. J Pathol *226*, 326-337.
Crossan, G.P., van der Weyden, L., Rosado, I.V., Langevin, F., Gaillard, P.H., McIntyre, R.E., Sanger Mouse Genetics, P., Gallagher, F., Kettunen, M.I., Lewis, D.Y.*, et al.* (2011). Disruption of mouse Slx4, a regulator of structure-specific nucleases, phenocopies Fanconi anemia. Nat Genet *43*, 147-152.
Cui, B., Johnson, S.P., Bullock, N., Ali-Osman, F., Bigner, D.D., and Friedman, H.S. (2009). Bifunctional DNA alkylator 1,3-bis(2-chloroethyl)-1-nitrosourea activates the ATR-Chk1 pathway independently of the mismatch repair pathway. Mol Pharmacol *75*, 1356-1363.
Dansonka-Mieszkowska, A., Kluska, A., Moes, J., Dabrowska, M., Nowakowska, D., Niwinska, A., Derlatka, P., Cendrowski, K., and Kupryjanczyk, J. (2010). A novel germline PALB2 deletion in Polish breast and ovarian cancer patients. BMC Med Genet *11*, 20.
De Silva, I.U., McHugh, P.J., Clingen, P.H., and Hartley, J.A. (2000a). Defining the roles of nucleotide excision repair and recombination in the repair of DNA interstrand cross-links in mammalian cells. Mol Cell Biol *20*, 7980-7990.
De Silva, I.U., McHugh, P.J., Clingen, P.H., and Hartley, J.A. (2000b). Defining the roles of nucleotide excision repair and recombination in the repair of DNA interstrand cross-links in mammalian cells. In Molecular and Cellular Biology, pp. 7980-7990.
Deans, A.J., and West, S.C. (2009). FANCM connects the genome instability disorders Bloom's Syndrome and Fanconi Anemia. Mol Cell *36*, 943-953.
Dickey, J.S., Redon, C.E., Nakamura, A.J., Baird, B.J., Sedelnikova, O.A., and Bonner, W.M. (2009). H2AX: functional roles and potential applications. In Chromosoma, pp. 683-692.
Dronkert, M.L., and Kanaar, R. (2001). Repair of DNA interstrand cross-links. In Mutat Res, pp. 217-247.
Duderstadt, K.E., Reyes-Lamothe, R., van Oijen, A.M., and Sherratt, D.J. (2014). Replication-fork dynamics. Cold Spring Harb Perspect Biol *6*.

Duquette, M.L., Zhu, Q., Taylor, E.R., Tsay, A.J., Shi, L.Z., Berns, M.W., and McGowan, C.H. (2012). CtIP is required to initiate replication-dependent interstrand crosslink repair. PLoS Genet *8*, e1003050.
Dutta, S., Chowdhury, G., and Gates, K.S. (2007). Interstrand cross-links generated by abasic sites in duplex DNA. In J Am Chem Soc, pp. 1852-1853.
Duxin, J.P., and Walter, J.C. (2015). What is the DNA repair defect underlying Fanconi anemia? Curr Opin Cell Biol *37*, 49-60.
Echols, H., and Goodman, M.F. (1991). Fidelity mechanisms in DNA replication. In Annu Rev Biochem, pp. 477-511.
Edelson, R., Berger, C., Gasparro, F., Jegasothy, B., Heald, P., Wintroub, B., Vonderheid, E., Knobler, R., Wolff, K., and Plewig, G. (1987). Treatment of cutaneous T-cell lymphoma by extracorporeal photochemotherapy. Preliminary results. In N Engl J Med, pp. 297-303.
Enoiu, M., Jiricny, J., and Scharer, O.D. (2012a). Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. Nucleic Acids Res *40*, 8953-8964.
Enoiu, M., Jiricny, J., and Scharer, O.D. (2012b). Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. In Nucleic Acids Research, pp. 8953-8964.
Falnes, P.Ø., Johansen, R.F., and Seeberg, E. (2002). AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli. In Nature, pp. 178-182.
Falnes, P.Ø., and Rognes, T. (2003). DNA repair by bacterial AlkB proteins. In Res Microbiol, pp. 531-538.
Fauchereau, F., Shalev, S., Chervinsky, E., Beck-Fruchter, R., Legois, B., Fellous, M., Caburet, S., and Veitia, R.A. (2016). A non-sense MCM9 mutation in a familial case of primary ovarian insufficiency. Clin Genet *89*, 603-607.
Fekairi, S., Scaglione, S., Chahwan, C., Taylor, E.R., Tissier, A., Coulon, S., Dong, M.Q., Ruse, C., Yates, J.R., 3rd, Russell, P.*, et al.* (2009). Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. Cell *138*, 78-89.
Fiumicino, S., Martinelli, S., Colussi, C., Aquilina, G., Leonetti, C., Crescenzi, M., and Bignami, M. (2000a). Sensitivity to DNA cross-linking chemotherapeutic agents in mismatch repair-defective cells in vitro and in xenografts. In Int J Cancer, pp. 590-596.
Fiumicino, S., Martinelli, S., Colussi, C., Aquilina, G., Leonetti, C., Crescenzi, M., and Bignami, M. (2000b). Sensitivity to DNA cross-linking chemotherapeutic agents in mismatch repair-defective cells in vitro and in xenografts. Int J Cancer *85*, 590-596.
Fortini, P., and Dogliotti, E. (2007). Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways. In DNA Repair (Amst), pp. 398-409.
Fousteri, M., and Mullenders, L.H.F. (2008). Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects. In Cell Res, pp. 73-84.
Fricke, W.M., and Brill, S.J. (2003). Slx1-Slx4 is a second structure-specific endonuclease functionally redundant with Sgs1-Top3. Genes Dev *17*, 1768-1778.
Fu, Y.V., Yardimci, H., Long, D.T., Ho, T.V., Guainazzi, A., Bermudez, V.P., Hurwitz, J., van Oijen, A., Scharer, O.D., and Walter, J.C. (2011). Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. Cell *146*, 931-941.
Fukui, K. (2010). DNA mismatch repair in eukaryotes and bacteria. In J Nucleic Acids.
Furuta, T., Ueda, T., Aune, G., Sarasin, A., Kraemer, K.H., and Pommier, Y. (2002). Transcription-coupled nucleotide excision repair as a determinant of cisplatin sensitivity of human cells. In Cancer Res, pp. 4899-4902.

Galio, L., Bouquet, C., and Brooks, P. (1999). ATP hydrolysis-dependent formation of a dynamic ternary nucleoprotein complex with MutS and MutL. In Nucleic Acids Res, pp. 2325-2331.
Gan, G.N., Wittschieben, J.P., Wittschieben, B.O., and Wood, R.D. (2008). DNA polymerase zeta (pol zeta) in higher eukaryotes. Cell Res *18*, 174-183.
Garcia-Higuera, I., Taniguchi, T., Ganesan, S., Meyn, M.S., Timmers, C., Hejna, J., Grompe, M., and D'Andrea, A.D. (2001). Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. Mol Cell *7*, 249-262.
Gargiulo, D., Kumar, G.S., Musser, S.S., and Tomasz, M. (1995). Structural and function modification of DNA by mitomycin C. Mechanism of the DNA sequence specificity of mitomycins. In Nucleic Acids Symp Ser, pp. 169-170.
Gari, K., Decaillet, C., Delannoy, M., Wu, L., and Constantinou, A. (2008a). Remodeling of DNA replication structures by the branch point translocase FANCM. Proc Natl Acad Sci U S A *105*, 16107-16112.
Gari, K., Decaillet, C., Stasiak, A.Z., Stasiak, A., and Constantinou, A. (2008b). The Fanconi anemia protein FANCM can promote branch migration of Holliday junctions and replication forks. Mol Cell *29*, 141-148.
Gari, K., Décaillet, C., Stasiak, A.Z., Stasiak, A., and Constantinou, A. (2008c). The Fanconi anemia protein FANCM can promote branch migration of Holliday junctions and replication forks. In Mol Cell, pp. 141-148.
Gasparro, F.P. (1996). Psoralen photobiology: recent advances. In Photochem Photobiol, pp. 553-557.
Godon, C., Mourgues, S., Nonnekens, J., Mourcet, A., Coin, F., Vermeulen, W., Mari, P.-O., and Giglia-Mari, G. (2012). Generation of DNA single-strand displacement by compromised nucleotide excision repair. In EMBO J, pp. 3550-3563.
Godthelp, B.C., Wiegant, W.W., Waisfisz, Q., Medhurst, A.L., Arwert, F., Joenje, H., and Zdzienicka, M.Z. (2006). Inducibility of nuclear Rad51 foci after DNA damage distinguishes all Fanconi anemia complementation groups from D1/BRCA2. Mutat Res *594*, 39-48.
GOODMAN, L.S., and WINTROBE, M.M. (1946). Nitrogen mustard therapy; use of methyl-bis (beta-chloroethyl) amine hydrochloride and tris (beta-chloroethyl) amine hydrochloride for Hodgkin's disease, lymphosarcoma, leukemia and certain allied and miscellaneous disorders. In J Am Med Assoc, pp. 126-132.
Grillari, J., Katinger, H., and Voglauer, R. (2007). Contributions of DNA interstrand cross-links to aging of cells and organisms. In Nucleic Acids Research, pp. 7566-7576.
Grossmann, K.F., Ward, A.M., Matkovic, M.E., Folias, A.E., and Moses, R.E. (2001a). S. cerevisiae has three pathways for DNA interstrand crosslink repair. Mutat Res *487*, 73-83.
Grossmann, K.F., Ward, A.M., Matkovic, M.E., Folias, A.E., and Moses, R.E. (2001b). S. cerevisiae has three pathways for DNA interstrand crosslink repair. In Mutat Res, pp. 73-83.
Guainazzi, A., and Schärer, O.D. (2010). Using synthetic DNA interstrand crosslinks to elucidate repair pathways and identify new therapeutic targets for cancer chemotherapy. In Cell Mol Life Sci, pp. 3683-3697.
Guan, L., and Greenberg, M.M. (2009). DNA interstrand cross-link formation by the 1,4-dioxobutane abasic lesion. In J Am Chem Soc, pp. 15225-15231.
Hanada, K., Budzowska, M., Modesti, M., Maas, A., Wyman, C., Essers, J., and Kanaar, R. (2006). The structure-specific endonuclease Mus81-Eme1 promotes conversion of interstrand DNA crosslinks into double-strands breaks. In EMBO J, pp. 4921-4932.
Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. In Cell, pp. 646-674.

Hanawalt, P.C., and Spivak, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. In Nat Rev Mol Cell Biol, pp. 958-970.
Haran, T.E., and Crothers, D.M. (1988). Phased psoralen cross-links do not bend the DNA double helix. In Biochemistry, pp. 6967-6971.
Hazrati, A., Ramis-Castelltort, M., Sarkar, S., Barber, L.J., Schofield, C.J., Hartley, J.A., and McHugh, P.J. (2008). Human SNM1A suppresses the DNA repair defects of yeast pso2 mutants. DNA Repair (Amst) *7*, 230-238.
Hearst, J.E. (1981). Psoralen photochemistry. In Annu Rev Biophys Bioeng, pp. 69-86.
Hejna, J., Philip, S., Ott, J., Faulkner, C., and Moses, R. (2007). The hSNM1 protein is a DNA 5'-exonuclease. Nucleic Acids Res *35*, 6115-6123.
Helleday, T., Lo, J., van Gent, D.C., and Engelward, B.P. (2007). DNA double-strand break repair: from mechanistic understanding to cancer treatment. DNA Repair (Amst) *6*, 923-935.
Hewish, M., Lord, C.J., Martin, S.A., Cunningham, D., and Ashworth, A. (2010). Mismatch repair deficient colorectal cancer in the era of personalized treatment. In Nat Rev Clin Oncol, pp. 197-208.
Hirsch, B., Shimamura, A., Moreau, L., Baldinger, S., Hag-alshiekh, M., Bostrom, B., Sencer, S., and D'Andrea, A.D. (2004). Association of biallelic BRCA2/FANCD1 mutations with spontaneous chromosomal instability and solid tumors of childhood. Blood *103*, 2554-2559.
Hlavin, E.M., Smeaton, M.B., and Miller, P.S. (2010). Initiation of DNA interstrand cross-link repair in mammalian cells. In Environ Mol Mutagen, pp. NA-NA.
Ho, T.V., Guainazzi, A., Derkunt, S.B., Enoiu, M., and Scharer, O.D. (2011). Structure-dependent bypass of DNA interstrand crosslinks by translesion synthesis polymerases. Nucleic Acids Res *39*, 7455-7464.
Hodskinson, M.R., Silhan, J., Crossan, G.P., Garaycoechea, J.I., Mukherjee, S., Johnson, C.M., Scharer, O.D., and Patel, K.J. (2014). Mouse SLX4 is a tumor suppressor that stimulates the activity of the nuclease XPF-ERCC1 in DNA crosslink repair. Mol Cell *54*, 472-484.
Hodson, C., Cole, A.R., Lewis, L.P., Miles, J.A., Purkiss, A., and Walden, H. (2011). Structural analysis of human FANCL, the E3 ligase in the Fanconi anemia pathway. J Biol Chem *286*, 32628-32637.
Howlett, N.G., Taniguchi, T., Olson, S., Cox, B., Waisfisz, Q., De Die-Smulders, C., Persky, N., Grompe, M., Joenje, H., Pals, G.*, et al.* (2002). Biallelic inactivation of BRCA2 in Fanconi anemia. Science *297*, 606-609.
Hoy, C.A., Thompson, L.H., Mooney, C.L., and Salazar, E.P. (1985). Defective DNA cross-link removal in Chinese hamster cell mutants hypersensitive to bifunctional alkylating agents. Cancer Res *45*, 1737-1743.
Huang, H., Zhu, L., Reid, B.R., Drobny, G.P., and Hopkins, P.B. (1995). Solution structure of a cisplatin-induced DNA interstrand cross-link. In Science, pp. 1842-1845.
Huang, M., Kennedy, R., Ali, A.M., Moreau, L.A., Meetei, A.R., D'Andrea, A.D., and Chen, C.C. (2011). Human MutS and FANCM complexes function as redundant DNA damage sensors in the Fanconi Anemia pathway. DNA Repair (Amst) *10*, 1203-1212.
Huang, M., Kim, J.M., Shiotani, B., Yang, K., Zou, L., and D'Andrea, A.D. (2010). The FANCM/FAAP24 complex is required for the DNA interstrand crosslink-induced checkpoint response. Mol Cell *39*, 259-268.
Huang, Y., Leung, J.W., Lowery, M., Matsushita, N., Wang, Y., Shen, X., Huong, D., Takata, M., Chen, J., and Li, L. (2014). Modularized functions of the Fanconi anemia core complex. Cell Rep *7*, 1849-1857.
Huang, Y., and Li, L. (2013). DNA crosslinking damage and cancer - a tale of friend and foe. In Transl Cancer Res, pp. 144-154.

Hwang, G.S., Kim, J.K., and Choi, B.S. (1996). The solution structure of a psoralen cross-linked DNA duplex by NMR and relaxation matrix refinement. In Biochem Biophys Res Commun, pp. 191-197.
Ishiai, M., Kimura, M., Namikoshi, K., Yamazoe, M., Yamamoto, K., Arakawa, H., Agematsu, K., Matsushita, N., Takeda, S., Buerstedde, J.M.*, et al.* (2004). DNA cross-link repair protein SNM1A interacts with PIAS1 in nuclear focus formation. Mol Cell Biol *24*, 10733-10741.
Islas, A.L., Baker, F.J., and Hanawalt, P.C. (1994). Transcription-coupled repair of psoralen cross-links but not monoadducts in Chinese hamster ovary cells. In Biochemistry, pp. 10794-10799.
Islas, A.L., Vos, J.M., and Hanawalt, P.C. (1991). Differential introduction and repair of psoralen photoadducts to DNA in specific human genes. In Cancer Res, pp. 2867-2873.
Iyama, T., Lee, S.Y., Berquist, B.R., Gileadi, O., Bohr, V.A., Seidman, M.M., McHugh, P.J., and Wilson, D.M., 3rd (2015). CSB interacts with SNM1A and promotes DNA interstrand crosslink processing. Nucleic Acids Res *43*, 247-258.
Jackson, A.L., and Loeb, L.A. (2001). The contribution of endogenous sources of DNA damage to the multiple mutations in cancer. In Mutat Res, pp. 7-21.
Johnson, K.M., Price, N.E., Wang, J., Fekry, M.I., Dutta, S., Seiner, D.R., Wang, Y., and Gates, K.S. (2013). On the formation and properties of interstrand DNA-DNA cross-links forged by reaction of an abasic site with the opposing guanine residue of 5'-CAp sequences in duplex DNA. In J Am Chem Soc, pp. 1015-1025.
Joo, W., Xu, G., Persky, N.S., Smogorzewska, A., Rudge, D.G., Buzovetsky, O., Elledge, S.J., and Pavletich, N.P. (2011). Structure of the FANCI-FANCD2 complex: insights into the Fanconi anemia DNA repair pathway. In Science, pp. 312-316.
Kadyrov, F.A., Dzantiev, L., Constantin, N., and Modrich, P. (2006). Endonucleolytic function of MutLalpha in human mismatch repair. In Cell, pp. 297-308.
Kashiyama, K., Nakazawa, Y., Pilz, D.T., Guo, C., Shimada, M., Sasaki, K., Fawcett, H., Wing, J.F., Lewin, S.O., Carr, L.*, et al.* (2013). Malfunction of nuclease ERCC1-XPF results in diverse clinical manifestations and causes Cockayne syndrome, xeroderma pigmentosum, and Fanconi anemia. Am J Hum Genet *92*, 807-819.
Kim, J.M., Kee, Y., Gurtan, A., and D'Andrea, A.D. (2008). Cell cycle-dependent chromatin loading of the Fanconi anemia core complex by FANCM/FAAP24. Blood *111*, 5215-5222.
Kim, Y., Lach, F.P., Desetty, R., Hanenberg, H., Auerbach, A.D., and Smogorzewska, A. (2011). Mutations of the SLX4 gene in Fanconi anemia. Nat Genet *43*, 142-146.
Kirchner, J., Sigurdsson, S., and Hopkins, P. (1992). Interstrand cross-linking of duplex DNA by nitrous acid: covalent structure of the dG-to-dG cross-link at the sequence 5'-CG. In J Am Chem Soc, pp. 4021-4027.
Kitao, H., Yamamoto, K., Matsushita, N., Ohzeki, M., Ishiai, M., and Takata, M. (2006). Functional interplay between BRCA2/FancD1 and FancC in DNA repair. J Biol Chem *281*, 21312-21320.
Klein Douwel, D., Boonen, R.A., Long, D.T., Szypowska, A.A., Raschle, M., Walter, J.C., and Knipscheer, P. (2014). XPF-ERCC1 acts in Unhooking DNA interstrand crosslinks in cooperation with FANCD2 and FANCP/SLX4. Mol Cell *54*, 460-471.
Klug, A.R., Harbut, M.B., Lloyd, R.S., and Minko, I.G. (2012). Replication bypass of N2-deoxyguanosine interstrand cross-links by human DNA polymerases eta and iota. Chem Res Toxicol *25*, 755-762.
Knipscheer, P., Raschle, M., Smogorzewska, A., Enoiu, M., Ho, T.V., Scharer, O.D., Elledge, S.J., and Walter, J.C. (2009a). The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. Science *326*, 1698-1701.

Knipscheer, P., Räschle, M., Smogorzewska, A., Enoiu, M., Ho, T.V., Schärer, O.D., Elledge, S.J., and Walter, J.C. (2009b). The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. In Science, pp. 1698-1701.
Koster, D.A., Palle, K., Bot, E.S., Bjornsti, M.A., and Dekker, N.H. (2007). Antitumour drugs impede DNA uncoiling by topoisomerase I. Nature *448*, 213-217.
Kratz, K., Schöpf, B., Kaden, S., Sendoel, A., Eberhard, R., Lademann, C., Cannavó, E., Sartori, A.A., Hengartner, M.O., and Jiricny, J. (2010). Deficiency of FANCD2-associated nuclease KIAA1018/FAN1 sensitizes cells to interstrand crosslinking agents. In Cell, pp. 77-88.
Lachaud, C., Moreno, A., Marchesi, F., Toth, R., Blow, J.J., and Rouse, J. (2016). Ubiquitinated Fancd2 recruits Fan1 to stalled replication forks to prevent genome instability. Science *351*, 846-849.
Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680-685.
Lange, S.S., Reddy, M.C., and Vasquez, K.M. (2009). Human HMGB1 directly facilitates interactions between nucleotide excision repair proteins on triplex-directed psoralen interstrand crosslinks. In DNA Repair (Amst), pp. 865-872.
Langevin, F., Crossan, G.P., Rosado, I.V., Arends, M.J., and Patel, K.J. (2011). Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. In Nature, pp. 53-58.
Larimer, F.W., Perry, J.R., and Hardigree, A.A. (1989). The REV1 gene of Saccharomyces cerevisiae: isolation, sequence, and functional analysis. J Bacteriol *171*, 230-237.
Larminat, F., Zhen, W., and Bohr, V.A. (1993). Gene-specific DNA repair of interstrand cross-links induced by chemotherapeutic agents can be preferential. In J Biol Chem, pp. 2649-2654.
Laurencon, A., Orme, C.M., Peters, H.K., Boulton, C.L., Vladar, E.K., Langley, S.A., Bakis, E.P., Harris, D.T., Harris, N.J., Wayson, S.M.*, et al.* (2004). A large-scale screen for mutagen-sensitive loci in Drosophila. Genetics *167*, 217-231.
Lawley, P.D., and Phillips, D.H. (1996). DNA adducts from chemotherapeutic agents. In Mutat Res, pp. 13-40.
Lee, K.Y., Im, J.S., Shibata, E., Park, J., Handa, N., Kowalczykowski, S.C., and Dutta, A. (2015). MCM8-9 complex promotes resection of double-strand break ends by MRE11-RAD50-NBS1 complex. Nat Commun *6*, 7744.
Legerski, R.J. (2010). Repair of DNA interstrand cross-links during S phase of the mammalian cell cycle. In Environ Mol Mutagen, pp. 540-551.
Lehoczký, P., McHugh, P.J., and Chovanec, M. (2007). DNA interstrand cross-link repair in Saccharomyces cerevisiae. In FEMS Microbiol Rev, pp. 109-133.
Leung, J.W., Wang, Y., Fong, K.W., Huen, M.S., Li, L., and Chen, J. (2012). Fanconi anemia (FA) binding protein FAAP20 stabilizes FA complementation group A (FANCA) and participates in interstrand cross-link repair. Proc Natl Acad Sci U S A *109*, 4491-4496.
Leveille, F., Blom, E., Medhurst, A.L., Bier, P., Laghmani el, H., Johnson, M., Rooimans, M.A., Sobeck, A., Waisfisz, Q., Arwert, F.*, et al.* (2004). The Fanconi anemia gene product FANCF is a flexible adaptor protein. J Biol Chem *279*, 39421-39430.
Levran, O., Attwooll, C., Henry, R.T., Milton, K.L., Neveling, K., Rio, P., Batish, S.D., Kalb, R., Velleuer, E., Barral, S.*, et al.* (2005). The BRCA1-interacting helicase BRIP1 is deficient in Fanconi anemia. Nat Genet *37*, 931-933.
Li, G.-M. (2008). Mechanisms and functions of DNA mismatch repair. In Cell Res, pp. 85-98.

Lin, Z., Kong, H., Nei, M., and Ma, H. (2006). Origins and evolution of the recA/RAD51 gene family: evidence for ancient gene duplication and endosymbiotic gene transfer. Proc Natl Acad Sci U S A *103*, 10328-10333.
Lindahl, T. (1993). Instability and decay of the primary structure of DNA. In Nature, pp. 709-715.
Ling, C., Ishiai, M., Ali, A.M., Medhurst, A.L., Neveling, K., Kalb, R., Yan, Z., Xue, Y., Oostra, A.B., Auerbach, A.D*., et al.* (2007). FAAP100 is essential for activation of the Fanconi anemia-associated DNA damage response pathway. EMBO J *26*, 2104-2114.
Liu, N., Lamerdin, J.E., Tebbs, R.S., Schild, D., Tucker, J.D., Shen, M.R., Brookman, K.W., Siciliano, M.J., Walter, C.A., Fan, W*., et al.* (1998). XRCC2 and XRCC3, new human Rad51-family members, promote chromosome stability and protect against DNA cross-links and other damages. In Mol Cell, pp. 783-793.
Liu, T., Ghosal, G., Yuan, J., Chen, J., and Huang, J. (2010). FAN1 acts with FANCI-FANCD2 to promote DNA interstrand cross-link repair. In Science, pp. 693-696.
Long, D.T., Räschle, M., Joukov, V., and Walter, J.C. (2011). Mechanism of RAD51-dependent DNA interstrand cross-link repair. In Science, pp. 84-87.
Longley, M.J., Pierce, A.J., and Modrich, P. (1997). DNA polymerase delta is required for human mismatch repair in vitro. In J Biol Chem, pp. 10917-10921.
Loveday, C., Turnbull, C., Ramsay, E., Hughes, D., Ruark, E., Frankum, J.R., Bowden, G., Kalmyrzaev, B., Warren-Perry, M., Snape, K*., et al.* (2011). Germline mutations in RAD51D confer susceptibility to ovarian cancer. In Nature Publishing Group, pp. 879-882.
Luebben, S.W., Kawabata, T., Akre, M.K., Lee, W.L., Johnson, C.S., O'Sullivan, M.G., and Shima, N. (2013). Helq acts in parallel to Fancc to suppress replication-associated genome instability. In Nucleic Acids Res, pp. 10283-10297.
Lutzmann, M., Grey, C., Traver, S., Ganier, O., Maya-Mendoza, A., Ranisavljevic, N., Bernex, F., Nishiyama, A., Montel, N., Gavois, E*., et al.* (2012). MCM8- and MCM9-deficient mice reveal gametogenesis defects and genome instability due to impaired homologous recombination. Mol Cell *47*, 523-534.
Lutzmann, M., Maiorano, D., and Mechali, M. (2005). Identification of full genes and proteins of MCM9, a novel, vertebrate-specific member of the MCM2-8 protein family. Gene *362*, 51-56.
MacKay, C., Déclais, A.-C., Lundin, C., Agostinho, A., Deans, A.J., MacArtney, T.J., Hofmann, K., Gartner, A., West, S.C., Helleday, T*., et al.* (2010). Identification of KIAA1018/FAN1, a DNA repair nuclease recruited to DNA damage by monoubiquitinated FANCD2. In Cell, pp. 65-76.
Malinge, J.M., Pérez, C., and Leng, M. (1994). Base sequence-independent distorsions induced by interstrand cross-links in cis-diamminedichloroplatinum (II)-modified DNA. In Nucleic Acids Research, pp. 3834-3839.
Marteijn, J.A., Lans, H., Vermeulen, W., and Hoeijmakers, J.H.J. (2014). Understanding nucleotide excision repair and its roles in cancer and ageing. In Nature Publishing Group, pp. 465-481.
Masson, J.Y., Tarsounas, M.C., Stasiak, A.Z., Stasiak, A., Shah, R., McIlwraith, M.J., Benson, F.E., and West, S.C. (2001). Identification and purification of two distinct complexes containing the five RAD51 paralogs. In Genes Dev, pp. 3296-3307.
McCabe, K.M., Hemphill, A., Akkari, Y., Jakobs, P.M., Pauw, D., Olson, S.B., Moses, R.E., and Grompe, M. (2008). ERCC1 is required for FANCD2 focus formation. Mol Genet Metab *95*, 66-73.
McCaffrey, R., St Johnston, D., and Gonzalez-Reyes, A. (2006). Drosophila mus301/spindle-C encodes a helicase with an essential role in double-strand DNA break repair and meiotic progression. Genetics *174*, 1273-1285.
McCulloch, S.D., and Kunkel, T.A. (2008). The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. In Cell Res, pp. 148-161.

McHugh, P.J., Sones, W.R., and Hartley, J.A. (2000). Repair of intermediate structures produced at DNA interstrand cross-links in Saccharomyces cerevisiae. Mol Cell Biol *20*, 3425-3433.
McHugh, P.J., Spanswick, V.J., and Hartley, J.A. (2001). Repair of DNA interstrand crosslinks: molecular mechanisms and clinical relevance. In Lancet Oncol, pp. 483-490.
McKay, B.C., Becerril, C., and Ljungman, M. (2001). P53 plays a protective role against UV- and cisplatin-induced apoptosis in transcription-coupled repair proficient fibroblasts. In Oncogene, pp. 6805-6808.
McMahon, L.W., Sangerman, J., Goodman, S.R., Kumaresan, K., and Lambert, M.W. (2001). Human alpha spectrin II and the FANCA, FANCC, and FANCG proteins bind to DNA containing psoralen interstrand cross-links. In Biochemistry, pp. 7025-7034.
Mcmahon, L.W., Zhang, P., Sridharan, D.M., Lefferts, J.A., and Lambert, M.W. (2009). Knockdown of αII spectrin in normal human cells by siRNA leads to chromosomal instability and decreased DNA interstrand cross-link repair. In Biochem Biophys Res Commun, pp. 288-293.
Medhurst, A.L., Laghmani el, H., Steltenpool, J., Ferrer, M., Fontaine, C., de Groot, J., Rooimans, M.A., Scheper, R.J., Meetei, A.R., Wang, W.*, et al.* (2006). Evidence for subcomplexes in the Fanconi anemia pathway. Blood *108*, 2072-2080.
Meinardi, M.T., Gietema, J.A., van Veldhuisen, D.J., van der Graaf, W.T., de Vries, E.G., and Sleijfer, D.T. (2000). Long-term chemotherapy-related cardiovascular morbidity. In Cancer Treat Rev, pp. 429-447.
Meindl, A., Hellebrand, H., Wiek, C., Erven, V., Wappenschmidt, B., Niederacher, D., Freund, M., Lichtner, P., Hartmann, L., Schaal, H.*, et al.* (2010). Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. Nat Genet *42*, 410-414.
Miller, J.H., and Heflich, R.H. (1982). A model for the effect of excision repair on the survival of human cells exposed to chemical carcinogens. In Chem Biol Interact, pp. 45-55.
Mimitou, E.P., and Symington, L.S. (2008). Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature *455*, 770-774.
Minko, I.G., Harbut, M.B., Kozekov, I.D., Kozekova, A., Jakobs, P.M., Olson, S.B., Moses, R.E., Harris, T.M., Rizzo, C.J., and Lloyd, R.S. (2008). Role for DNA polymerase kappa in the processing of N2-N2-guanine interstrand cross-links. J Biol Chem *283*, 17075-17082.
Mitchell, J.R., Hoeijmakers, J.H.J., and Niedernhofer, L.J. (2003). Divide and conquer: nucleotide excision repair battles cancer and ageing. In Curr Opin Cell Biol, pp. 232-240.
Mocquet, V., Lainé, J.P., Riedl, T., Yajin, Z., Lee, M.Y., and Egly, J.M. (2008). Sequential recruitment of the repair factors during NER: the role of XPG in initiating the resynthesis step. In EMBO J, pp. 155-167.
Moldovan, G.L., Madhavan, M.V., Mirchandani, K.D., McCaffrey, R.M., Vinciguerra, P., and D'Andrea, A.D. (2010). DNA polymerase POLN participates in cross-link repair and homologous recombination. Mol Cell Biol *30*, 1088-1096.
Moore, D.H., Mendelsohn, M.L., and Lohman, P.H. (1992). International Commission for Protection Against Environmental Mutagens and Carcinogens. A method for comparing and combining short-term genotoxicity test data: the optimal use of dose information. In Mutat Res, pp. 27-42.
Mu, D., Bessho, T., Nechev, L.V., Chen, D.J., Harris, T.M., Hearst, J.E., and Sancar, A. (2000). DNA interstrand cross-links induce futile repair synthesis in mammalian cell extracts. Mol Cell Biol *20*, 2446-2454.

Muniandy, P.A., Liu, J., Majumdar, A., Liu, S.-T., and Seidman, M.M. (2010). DNA interstrand crosslink repair in mammalian cells: step by step. In Crit Rev Biochem Mol Biol, pp. 23-49.
Muniandy, P.A., Thapa, D., Thazhathveetil, A.K., Liu, S.-T., and Seidman, M.M. (2009a). Repair of laser-localized DNA interstrand cross-links in G1 phase mammalian cells. In J Biol Chem, pp. 27908-27917.
Muniandy, P.A., Thapa, D., Thazhathveetil, A.K., Liu, S.T., and Seidman, M.M. (2009b). Repair of laser-localized DNA interstrand cross-links in G1 phase mammalian cells. J Biol Chem *284*, 27908-27917.
Munoz, I.M., Hain, K., Declais, A.C., Gardiner, M., Toh, G.W., Sanchez-Pulido, L., Heuckmann, J.M., Toth, R., Macartney, T., Eppink, B.*, et al.* (2009). Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair. Mol Cell *35*, 116-127.
Murnane, J.P., and Byfield, J.E. (1981). Irreparable DNA cross-links and mammalian cell lethality with bifunctional alkylating agents. In Chem Biol Interact, pp. 75-86.
Muzzini, D.M., Plevani, P., Boulton, S.J., Cassata, G., and Marini, F. (2008). Caenorhabditis elegans POLQ-1 and HEL-308 function in two distinct DNA interstrand cross-link repair pathways. In DNA Repair (Amst), pp. 941-950.
Niedernhofer, L.J. (2007). The Fanconi anemia signalosome anchor. Mol Cell *25*, 487-490.
Nijman, S.M., Huang, T.T., Dirac, A.M., Brummelkamp, T.R., Kerkhoven, R.M., D'Andrea, A.D., and Bernards, R. (2005). The deubiquitinating enzyme USP1 regulates the Fanconi anemia pathway. Mol Cell *17*, 331-339.
Nishimura, K., Ishiai, M., Horikawa, K., Fukagawa, T., Takata, M., Takisawa, H., and Kanemaki, M.T. (2012). Mcm8 and Mcm9 form a complex that functions in homologous recombination repair induced by DNA interstrand crosslinks. Mol Cell *47*, 511-522.
Nojima, K., Hochegger, H., Saberi, A., Fukushima, T., Kikuchi, K., Yoshimura, M., Orelli, B.J., Bishop, D.K., Hirano, S., Ohzeki, M.*, et al.* (2005). Multiple repair pathways mediate tolerance to chemotherapeutic cross-linking agents in vertebrate cells. Cancer Res *65*, 11704-11711.
Noll, D.M., Mason, T.M., and Miller, P.S. (2006). Formation and repair of interstrand cross-links in DNA. In Chem Rev, pp. 277-301.
Norman, D., Live, D., Sastry, M., Lipman, R., Hingerty, B.E., Tomasz, M., Broyde, S., and Patel, D.J. (1990). NMR and computational characterization of mitomycin cross-linked to adjacent deoxyguanosines in the minor groove of the d(T-A-C-G-T-A).d(T-A-C-G-T-A) duplex. In Biochemistry, pp. 2861-2875.
Offit, K., Levran, O., Mullaney, B., Mah, K., Nafa, K., Batish, S.D., Diotti, R., Schneider, H., Deffenbaugh, A., Scholl, T.*, et al.* (2003). Shared genetic susceptibility to breast cancer, brain tumors, and Fanconi anemia. J Natl Cancer Inst *95*, 1548-1551.
Ogi, T., Limsirichaikul, S., Overmeer, R.M., Volker, M., Takenaka, K., Cloney, R., Nakazawa, Y., Niimi, A., Miki, Y., Jaspers, N.G.*, et al.* (2010). Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. In Molecular Cell, pp. 714-727.
Orelli, B., McClendon, T.B., Tsodikov, O.V., Ellenberger, T., Niedernhofer, L.J., and Scharer, O.D. (2010). The XPA-binding domain of ERCC1 is required for nucleotide excision repair but not other DNA repair pathways. J Biol Chem *285*, 3705-3712.
Overmeer, R.M., Moser, J., Volker, M., Kool, H., Tomkinson, A.E., van Zeeland, A.A., Mullenders, L.H.F., and Fousteri, M. (2011). Replication protein A safeguards genome integrity by controlling NER incision events. In J Cell Biol, pp. 401-415.
Oyama, T., Oka, H., Mayanagi, K., Shirai, T., Matoba, K., Fujikane, R., Ishino, Y., and Morikawa, K. (2009). Atomic structures and functional implications of the archaeal RecQ-like helicase Hjm. In BMC Struct Biol, pp. 2.

Pace, P., Mosedale, G., Hodskinson, M.R., Rosado, I.V., Sivasubramaniam, M., and Patel, K.J. (2010). Ku70 corrupts DNA repair in the absence of the Fanconi anemia pathway. Science *329*, 219-223.
Paquet, F., Pérez, C., Leng, M., Lancelot, G., and Malinge, J.M. (1996). NMR solution structure of a DNA decamer containing an interstrand cross-link of the antitumor drug cis-diamminedichloroplatinum (II). In J Biomol Struct Dyn, pp. 67-77.
Park, J., Long, D.T., Lee, K.Y., Abbas, T., Shibata, E., Negishi, M., Luo, Y., Schimenti, J.C., Gambus, A., Walter, J.C.*, et al.* (2013). The MCM8-MCM9 complex promotes RAD51 recruitment at DNA damage sites to facilitate homologous recombination. Mol Cell Biol *33*, 1632-1644.
Park, J.Y., Singh, T.R., Nassar, N., Zhang, F., Freund, M., Hanenberg, H., Meetei, A.R., and Andreassen, P.R. (2014). Breast cancer-associated missense mutants of the PALB2 WD40 domain, which directly binds RAD51C, RAD51 and BRCA2, disrupt DNA repair. Oncogene *33*, 4803-4812.
Parmar, K., D'Andrea, A., and Niedernhofer, L.J. (2009). Mouse models of Fanconi anemia. Mutat Res *668*, 133-140.
Pavlov, Y.I., Minnick, D.T., Izuta, S., and Kunkel, T.A. (1994). DNA replication fidelity with 8-oxodeoxyguanosine triphosphate. In Biochemistry, pp. 4695-4701.
Pelttari, L.M., Heikkinen, T., Thompson, D., Kallioniemi, A., Schleutker, J., Holli, K., Blomqvist, C., Aittomaki, K., Butzow, R., and Nevanlinna, H. (2011). RAD51C is a susceptibility gene for ovarian cancer. Hum Mol Genet *20*, 3278-3288.
Petermann, E., Orta, M.L., Issaeva, N., Schultz, N., and Helleday, T. (2010). Hydroxyurea-Stalled Replication Forks Become Progressively Inactivated and Require Two Different RAD51-Mediated Pathways for Restart and Repair. In Molecular Cell, pp. 492-502.
Petrucelli, N., Daly, M.B., and Feldman, G.L. (2010). Hereditary breast and ovarian cancer due to mutations in BRCA1 and BRCA2. Genet Med *12*, 245-259.
Plosky, B.S., Vidal, A.E., Fernandez de Henestrosa, A.R., McLenigan, M.P., McDonald, J.P., Mead, S., and Woodgate, R. (2006). Controlling the subcellular localization of DNA polymerases iota and eta via interactions with ubiquitin. EMBO J *25*, 2847-2855.
Rajski, S.R., and Williams, R.M. (1998). DNA Cross-Linking Agents as Antitumor Drugs. In Chem Rev, pp. 2723-2796.
Räschle, M., Knipscheer, P., Knipsheer, P., Enoiu, M., Angelov, T., Sun, J., Griffith, J.D., Ellenberger, T.E., Schärer, O.D., and Walter, J.C. (2008). Mechanism of replication-coupled DNA interstrand crosslink repair. In Cell, pp. 969-980.
Reid, S., Schindler, D., Hanenberg, H., Barker, K., Hanks, S., Kalb, R., Neveling, K., Kelly, P., Seal, S., Freund, M.*, et al.* (2007). Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. Nat Genet *39*, 162-164.
Reishus, J., and Martin, D. (1961). cis-Dichlorodiammineplatinum(II). Acid Hydrolysis and Isotopic Exchange of the Chloride Ligands1. In J Am Chem Soc, pp. 2457-2462.
Richards, J.D., Johnson, K.A., Liu, H., McRobbie, A.-M., McMahon, S., Oke, M., Carter, L., Naismith, J.H., and White, M.F. (2008a). Structure of the DNA repair helicase hel308 reveals DNA binding and autoinhibitory domains. In J Biol Chem, pp. 5118-5126.
Richards, J.D., Johnson, K.A., Liu, H., McRobbie, A.M., McMahon, S., Oke, M., Carter, L., Naismith, J.H., and White, M.F. (2008b). Structure of the DNA repair helicase hel308 reveals DNA binding and autoinhibitory domains. J Biol Chem *283*, 5118-5126.
Richards, S., Liu, S.T., Majumdar, A., Liu, J.L., Nairn, R.S., Bernier, M., Maher, V., and Seidman, M.M. (2005). Triplex targeted genomic crosslinks enter separable deletion and base substitution pathways. Nucleic Acids Res *33*, 5382-5393.

Rink, S.M., and Hopkins, P.B. (1995). A mechlorethamine-induced DNA interstrand cross-link bends duplex DNA. In Biochemistry, pp. 1439-1445.
Rink, S.M., Lipman, R., Alley, S.C., Hopkins, P.B., and Tomasz, M. (1996). Bending of DNA by the mitomycin C-induced, GpG intrastrand cross-link. In Chem Res Toxicol, pp. 382-389.
Rodriguez-Bigas, M.A., Boland, C.R., Hamilton, S.R., Henson, D.E., Jass, J.R., Khan, P.M., Lynch, H., Perucho, M., Smyrk, T., Sobin, L.*, et al.* (1997). A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. In J Natl Cancer Inst, pp. 1758-1762.
Rosado, I.V., Niedzwiedz, W., Alpi, A.F., and Patel, K.J. (2009). The Walker B motif in avian FANCM is required to limit sister chromatid exchanges but is dispensable for DNA crosslink repair. Nucleic Acids Res *37*, 4360-4370.
Rothfuss, A., and Grompe, M. (2004). Repair kinetics of genomic interstrand DNA cross-links: evidence for DNA double-strand break-dependent activation of the Fanconi anemia/BRCA pathway. Mol Cell Biol *24*, 123-134.
Sarkar, S., Davies, A.A., Ulrich, H.D., and McHugh, P.J. (2006a). DNA interstrand crosslink repair during G1 involves nucleotide excision repair and DNA polymerase zeta. EMBO J *25*, 1285-1294.
Sarkar, S., Davies, A.A., Ulrich, H.D., and McHugh, P.J. (2006b). DNA interstrand crosslink repair during G1 involves nucleotide excision repair and DNA polymerase ζ. In EMBO J, pp. 1285-1294.
Sartori, A.A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., Baer, R., Lukas, J., and Jackson, S.P. (2007). Human CtIP promotes DNA end resection. Nature *450*, 509-514.
Sattler, U., Frit, P., Salles, B., and Calsou, P. (2003). Long-patch DNA repair synthesis during base excision repair in mammalian cells. In EMBO Rep, pp. 363-367.
Sawyer, S.L., Tian, L., Kahkonen, M., Schwartzentruber, J., Kircher, M., University of Washington Centre for Mendelian, G., Consortium, F.C., Majewski, J., Dyment, D.A., Innes, A.M.*, et al.* (2015). Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype. Cancer Discov *5*, 135-142.
Scicchitano, D.A., Olesnicky, E.C., and Dimitri, A. (2004). Transcription and DNA adducts: what happens when the message gets cut off? DNA Repair (Amst) *3*, 1537-1548.
Sengerova, B., Wang, A.T., and McHugh, P.J. (2011). Orchestrating the nucleases involved in DNA interstrand cross-link (ICL) repair. Cell Cycle *10*, 3999-4008.
Sharma, S., Helchowski, C.M., and Canman, C.E. (2013). The roles of DNA polymerase zeta and the Y family DNA polymerases in promoting or preventing genome instability. Mutat Res *743-744*, 97-110.
Singh, T.R., Bakker, S.T., Agarwal, S., Jansen, M., Grassman, E., Godthelp, B.C., Ali, A.M., Du, C.H., Rooimans, M.A., Fan, Q.*, et al.* (2009). Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M. Blood *114*, 174-180.
Singh, T.R., Saro, D., Ali, A.M., Zheng, X.F., Du, C.H., Killen, M.W., Sachpatzidis, A., Wahengbam, K., Pierce, A.J., Xiong, Y.*, et al.* (2010). MHF1-MHF2, a histone-fold-containing protein complex, participates in the Fanconi anemia pathway via FANCM. Mol Cell *37*, 879-886.
Sirbu, B.M., Couch, F.B., and Cortez, D. (2012). Monitoring the spatiotemporal dynamics of proteins at replication forks and in assembled chromatin using isolation of proteins on nascent DNA. In Nat Protoc, pp. 594-605.
Sirbu, B.M., Couch, F.B., Feigerle, J.T., Bhaskara, S., Hiebert, S.W., and Cortez, D. (2011). Analysis of protein dynamics at active, stalled, and collapsed replication forks. In Genes & Development, pp. 1320-1327.

Sirbu, B.M., McDonald, W.H., Dungrawala, H., Badu-Nkansah, A., Kavanaugh, G.M., Chen, Y., Tabb, D.L., and Cortez, D. (2013). Identification of proteins at active, stalled, and collapsed replication forks using isolation of proteins on nascent DNA (iPOND) coupled with mass spectrometry. In J Biol Chem, pp. 31458-31467.
Smeaton, M.B., Hlavin, E.M., Noronha, A.M., Murphy, S.P., Wilds, C.J., and Miller, P.S. (2009). Effect of cross-link structure on DNA interstrand cross-link repair synthesis. Chem Res Toxicol *22*, 1285-1297.
Smogorzewska, A., Desetty, R., Saito, T.T., Schlabach, M., Lach, F.P., Sowa, M.E., Clark, A.B., Kunkel, T.A., Harper, J.W., Colaiácovo, M.P*., et al.* (2010). A genetic screen identifies FAN1, a Fanconi anemia-associated nuclease necessary for DNA interstrand crosslink repair. In Mol Cell, pp. 36-47.
Smogorzewska, A., Matsuoka, S., Vinciguerra, P., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Ballif, B.A., Gygi, S.P., Hofmann, K., D'Andrea, A.D*., et al.* (2007). Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. Cell *129*, 289-301.
Somyajit, K., Subramanya, S., and Nagaraju, G. (2012). Distinct roles of FANCO/RAD51C protein in DNA damage signaling and repair: implications for Fanconi anemia and breast cancer susceptibility. J Biol Chem *287*, 3366-3380.
Spielmann, H.P., Dwyer, T.J., Hearst, J.E., and Wemmer, D.E. (1995a). Solution structures of psoralen monoadducted and cross-linked DNA oligomers by NMR spectroscopy and restrained molecular dynamics. In Biochemistry, pp. 12937-12953.
Spielmann, H.P., Dwyer, T.J., Sastry, S.S., Hearst, J.E., and Wemmer, D.E. (1995b). DNA structural reorganization upon conversion of a psoralen furan-side monoadduct to an interstrand cross-link: implications for DNA repair. In Proceedings of the National Academy of Sciences of the United States of America, pp. 2345-2349.
Stelter, P., and Ulrich, H.D. (2003). Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. Nature *425*, 188-191.
Stoepker, C., Hain, K., Schuster, B., Hilhorst-Hofstee, Y., Rooimans, M.A., Steltenpool, J., Oostra, A.B., Eirich, K., Korthof, E.T., Nieuwint, A.W*., et al.* (2011). SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. Nat Genet *43*, 138-141.
Stone, M.P., Cho, Y.-J., Huang, H., Kim, H.-Y., Kozekov, I.D., Kozekova, A., Wang, H., Minko, I.G., Lloyd, R.S., Harris, T.M*., et al.* (2008). Interstrand DNA cross-links induced by alpha,beta-unsaturated aldehydes derived from lipid peroxidation and environmental sources. In Acc Chem Res, pp. 793-804.
Sugahara, R., Mon, H., Lee, J.M., and Kusakabe, T. (2012). Monoubiquitination-dependent chromatin loading of FancD2 in silkworms, a species lacking the FA core complex. Gene *501*, 180-187.
Sugasawa, K. (2008). XPC: its product and biological roles. In Adv Exp Med Biol, pp. 47-56.
Sugasawa, K. (2010). Regulation of damage recognition in mammalian global genomic nucleotide excision repair. In Mutat Res, pp. 29-37.
Sugasawa, K., Ng, J.M., Masutani, C., Iwai, S., van der Spek, P.J., Eker, A.P., Hanaoka, F., Bootsma, D., and Hoeijmakers, J.H. (1998). Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. In Molecular Cell, pp. 223-232.
Suwaki, N., Klare, K., and Tarsounas, M. (2011). RAD51 paralogs: Roles in DNA damage signalling, recombinational repair and tumorigenesis. In Seminars in Cell and Developmental Biology, pp. 898-905.
Svendsen, J.M., Smogorzewska, A., Sowa, M.E., O'Connell, B.C., Gygi, S.P., Elledge, S.J., and Harper, J.W. (2009). Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. Cell *138*, 63-77.

Tafel, A.A., Wu, L., and McHugh, P.J. (2011). Human HEL308 localizes to damaged replication forks and unwinds lagging strand structures. J Biol Chem *286*, 15832-15840.
Takahara, P., Rosenzweig, A., Frederick, C., and Lippard, S. (1995). Crystal structure of double-stranded DNA containing the major adduct of the anticancer drug cisplatin. In Nature, pp. 649-652.
Takata, K., Reh, S., Tomida, J., Person, M.D., and Wood, R.D. (2013). Human DNA helicase HELQ participates in DNA interstrand crosslink tolerance with ATR and RAD51 paralogs. Nat Commun *4*, 2338.
Takata, M., Sasaki, M.S., Tachiiri, S., Fukushima, T., Sonoda, E., Schild, D., Thompson, L.H., and Takeda, S. (2001). Chromosome instability and defective recombinational repair in knockout mutants of the five Rad51 paralogs. Mol Cell Biol *21*, 2858-2866.
Taylor, M.R.G., Spirek, M., Chaurasiya, K.R., Ward, J.D., Carzaniga, R., Yu, X., Egelman, E.H., Collinson, L.M., Rueda, D., Krejci, L.*, et al.* (2015). Rad51 Paralogs Remodel Pre-synaptic Rad51 Filaments to Stimulate Homologous Recombination. Cell *162*, 271-286.
Thoma, B.S. (2005). Human XPC-hHR23B interacts with XPA-RPA in the recognition of triplex-directed psoralen DNA interstrand crosslinks. In Nucleic Acids Research, pp. 2993-3001.
Tian, Y., Paramasivam, M., Ghosal, G., Chen, D., Shen, X., Huang, Y., Akhter, S., Legerski, R., Chen, J., Seidman, M.M.*, et al.* (2015). UHRF1 contributes to DNA damage repair as a lesion recognition factor and nuclease scaffold. Cell Rep *10*, 1957-1966.
Tokuda, K., and Bodell, W.J. (1987). Cytotoxicity and sister chromatid exchanges in 9L cells treated with monofunctional and bifunctional nitrogen mustards. In Carcinogenesis, pp. 1697-1701.
Tran, P.T., Erdeniz, N., Symington, L.S., and Liskay, R.M. (2004). EXO1-A multi-tasking eukaryotic nuclease. In DNA Repair (Amst), pp. 1549-1559.
Trewick, S.C., Henshaw, T.F., Hausinger, R.P., Lindahl, T., and Sedgwick, B. (2002). Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. In Nature, pp. 174-178.
Trujillo, J.P., Mina, L.B., Pujol, R., Bogliolo, M., Andrieux, J., Holder, M., Schuster, B., Schindler, D., and Surralles, J. (2012). On the role of FAN1 in Fanconi anemia. Blood *120*, 86-89.
Vare, D., Groth, P., Carlsson, R., Johansson, F., Erixon, K., and Jenssen, D. (2012). DNA interstrand crosslinks induce a potent replication block followed by formation and repair of double strand breaks in intact mammalian cells. DNA Repair (Amst) *11*, 976-985.
Vaz, F., Hanenberg, H., Schuster, B., Barker, K., Wiek, C., Erven, V., Neveling, K., Endt, D., Kesterton, I., Autore, F.*, et al.* (2010). Mutation of the RAD51C gene in a Fanconi anemia-like disorder. Nat Genet *42*, 406-409.
Vos, J.M., and Hanawalt, P.C. (1987). Processing of psoralen adducts in an active human gene: repair and replication of DNA containing monoadducts and interstrand cross-links. In Cell, pp. 789-799.
Vuorela, M., Pylkas, K., Hartikainen, J.M., Sundfeldt, K., Lindblom, A., von Wachenfeldt Wappling, A., Haanpaa, M., Puistola, U., Rosengren, A., Anttila, M.*, et al.* (2011). Further evidence for the contribution of the RAD51C gene in hereditary breast and ovarian cancer susceptibility. Breast Cancer Res Treat *130*, 1003-1010.
Wang, A.T., Kim, T., Wagner, J.E., Conti, B.A., Lach, F.P., Huang, A.L., Molina, H., Sanborn, E.M., Zierhut, H., Cornes, B.K.*, et al.* (2015). A Dominant Mutation in Human RAD51 Reveals Its Function in DNA Interstrand Crosslink Repair Independent of Homologous Recombination. Mol Cell *59*, 478-490.

Wang, A.T., Sengerova, B., Cattell, E., Inagawa, T., Hartley, J.M., Kiakos, K., Burgess-Brown, N.A., Swift, L.P., Enzlin, J.H., Schofield, C.J.*, et al.* (2011a). Human SNM1A and XPF-ERCC1 collaborate to initiate DNA interstrand cross-link repair. Genes Dev *25*, 1859-1870.
Wang, A.T., Sengerová, B., Cattell, E., Inagawa, T., Hartley, J.M., Kiakos, K., Burgess-Brown, N.A., Swift, L.P., Enzlin, J.H., Schofield, C.J.*, et al.* (2011b). Human SNM1A and XPF-ERCC1 collaborate to initiate DNA interstrand cross-link repair. In Genes Dev, pp. 1859-1870.
Ward, J.D., Muzzini, D.M., Petalcorin, M.I., Martinez-Perez, E., Martin, J.S., Plevani, P., Cassata, G., Marini, F., and Boulton, S.J. (2010). Overlapping mechanisms promote postsynaptic RAD-51 filament disassembly during meiotic double-strand break repair. Mol Cell *37*, 259-272.
Ward, T.A., Dudasova, Z., Sarkar, S., Bhide, M.R., Vlasakova, D., Chovanec, M., and McHugh, P.J. (2012). Components of a Fanconi-like pathway control Pso2-independent DNA interstrand crosslink repair in yeast. PLoS Genet *8*, e1002884.
Wickramanyake, A., Bernier, G., Pennil, C., Casadei, S., Agnew, K.J., Stray, S.M., Mandell, J., Garcia, R.L., Walsh, T., King, M.C.*, et al.* (2012). Loss of function germline mutations in RAD51D in women with ovarian carcinoma. Gynecol Oncol *127*, 552-555.
Wijen, J.P., Nivard, M.J., and Vogel, E.W. (2000). The in vivo genetic activity profile of the monofunctional nitrogen mustard 2-chloroethylamine differs drastically from its bifunctional counterpart mechlorethamine. In Carcinogenesis, pp. 1859-1867.
Williams, H.L., Gottesman, M.E., and Gautier, J. (2012). Replication-independent repair of DNA interstrand crosslinks. Mol Cell *47*, 140-147.
Williams, S.A., Longerich, S., Sung, P., Vaziri, C., and Kupfer, G.M. (2011). The E3 ubiquitin ligase RAD18 regulates ubiquitylation and chromatin loading of FANCD2 and FANCI. In Blood, pp. 5078-5087.
Woodman, I.L., and Bolt, E.L. (2009). Molecular biology of Hel308 helicase in archaea. Biochem Soc Trans *37*, 74-78.
Woodman, I.L., and Bolt, E.L. (2011). Winged helix domains with unknown function in Hel308 and related helicases. In Biochem Soc Trans, pp. 140-144.
Woynarowski, J.M., Faivre, S., Herzig, M.C., Arnett, B., Chapman, W.G., Trevino, A.V., Raymond, E., Chaney, S.G., Vaisman, A., Varchenko, M.*, et al.* (2000). Oxaliplatin-induced damage of cellular DNA. In Molecular Pharmacology, pp. 920-927.
Wu, H.I., Brown, J.A., Dorie, M.J., Lazzeroni, L., and Brown, J.M. (2004). Genome-wide identification of genes conferring resistance to the anticancer agents cisplatin, oxaliplatin, and mitomycin C. In Cancer Res, pp. 3940-3948.
Wu, Q., Christensen, L.A., Legerski, R.J., and Vasquez, K.M. (2005). Mismatch repair participates in error-free processing of DNA interstrand crosslinks in human cells. EMBO Rep *6*, 551-557.
Wu, X., Guo, D., Yuan, F., and Wang, Z. (2001). Accessibility of DNA polymerases to repair synthesis during nucleotide excision repair in yeast cell-free extracts. In Nucleic Acids Research, pp. 3123-3130.
Wyatt, H.D., Sarbajna, S., Matos, J., and West, S.C. (2013). Coordinated actions of SLX1-SLX4 and MUS81-EME1 for Holliday junction resolution in human cells. Mol Cell *52*, 234-247.
Yaghi, B.M., Turner, P.M., Denny, W.A., Turner, P.R., O'Connor, C.J., and Ferguson, L.R. (1998). Comparative mutational spectra of the nitrogen mustard chlorambucil and its half-mustard analogue in Chinese hamster AS52 cells. In Mutat Res, pp. 153-164.
Yam, V.W.W. (2010). Behind platinum's sparkle. In Nature Publishing Group, pp. 790-790.
Yamanaka, K., Minko, I.G., Takata, K., Kolbanovskiy, A., Kozekov, I.D., Wood, R.D., Rizzo, C.J., and Lloyd, R.S. (2010). Novel enzymatic function of DNA polymerase nu in

translesion DNA synthesis past major groove DNA-peptide and DNA-DNA cross-links. Chem Res Toxicol *23*, 689-695.
Yan, Z., Delannoy, M., Ling, C., Daee, D., Osman, F., Muniandy, P.A., Shen, X., Oostra, A.B., Du, H., Steltenpool, J.*, et al.* (2010). A histone-fold complex and FANCM form a conserved DNA-remodeling complex to maintain genome stability. Mol Cell *37*, 865-878.
Yan, Z., Guo, R., Paramasivam, M., Shen, W., Ling, C., Fox, D., 3rd, Wang, Y., Oostra, A.B., Kuehl, J., Lee, D.Y.*, et al.* (2012). A ubiquitin-binding protein, FAAP20, links RNF8-mediated ubiquitination to the Fanconi anemia DNA repair network. Mol Cell *47*, 61-75.
Yonetani, Y., Hochegger, H., Sonoda, E., Shinya, S., Yoshikawa, H., Takeda, S., and Yamazoe, M. (2005). Differential and collaborative actions of Rad51 paralog proteins in cellular response to DNA damage. In Nucleic Acids Res, pp. 4544-4552.
Yoshikiyo, K., Kratz, K., Hirota, K., Nishihara, K., Takata, M., Kurumizaka, H., Horimoto, S., Takeda, S., and Jiricny, J. (2010). KIAA1018/FAN1 nuclease protects cells against genomic instability induced by interstrand cross-linking agents. In Proceedings of the National Academy of Sciences of the United States of America, pp. 21553-21557.
Zamble, D.B., Mu, D., Reardon, J.T., Sancar, A., and Lippard, S.J. (1996a). Repair of cisplatin--DNA adducts by the mammalian excision nuclease. Biochemistry *35*, 10004-10013.
Zamble, D.B., Mu, D., Reardon, J.T., Sancar, A., and Lippard, S.J. (1996b). Repair of cisplatin--DNA adducts by the mammalian excision nuclease. In Biochemistry, pp. 10004-10013.
Zhang, J., and Walter, J.C. (2014). Mechanism and regulation of incisions during DNA interstrand cross-link repair. DNA Repair (Amst) *19*, 135-142.
Zhang, N., Lu, X., Zhang, X., Peterson, C.A., and Legerski, R.J. (2002a). hMutSbeta is required for the recognition and uncoupling of psoralen interstrand cross-links in vitro. Mol Cell Biol *22*, 2388-2397.
Zhang, N., Lu, X., Zhang, X., Peterson, C.A., and Legerski, R.J. (2002b). hMutSbeta is required for the recognition and uncoupling of psoralen interstrand cross-links in vitro. In Molecular and Cellular Biology, pp. 2388-2397.
Zhang, Y., Yuan, F., Presnell, S.R., Tian, K., Gao, Y., Tomkinson, A.E., Gu, L., and Li, G.-M. (2005). Reconstitution of 5'-directed human mismatch repair in a purified system. In Cell, pp. 693-705.
Zheng, H., Wang, X., Warren, A.J., Legerski, R.J., Nairn, R.S., Hamilton, J.W., and Li, L. (2003). Nucleotide excision repair- and polymerase eta-mediated error-prone removal of mitomycin C interstrand cross-links. In Molecular and Cellular Biology, pp. 754-761.
Zhou, W., Otto, E.A., Cluckey, A., Airik, R., Hurd, T.W., Chaki, M., Diaz, K., Lach, F.P., Bennett, G.R., Gee, H.Y.*, et al.* (2012). FAN1 mutations cause karyomegalic interstitial nephritis, linking chronic kidney failure to defective DNA damage repair. Nat Genet *44*, 910-915.
Zhu, G., and Lippard, S.J. (2009). Photoaffinity labeling reveals nuclear proteins that uniquely recognize cisplatin-DNA interstrand cross-links. In Biochemistry, pp. 4916-4925.
Zhu, Z., Chung, W.H., Shim, E.Y., Lee, S.E., and Ira, G. (2008). Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. Cell *134*, 981-994.
Zou, L., and Elledge, S.J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. Science *300*, 1542-1548.