Characterization of UHRF2 (RNF107) as a novel sensor for DNA ICLs in the Fanconi Anemia Pathway

Anna Motnenko

Department of Biochemistry

Wolfson College

2017

Abstract

The Fanconi Anemia (FA) pathway is important for repairing interstrand crosslinks (ICLs) between the Watson-Crick strands of the DNA helix. An initial and essential stage in the repair process is the detection of the ICL. In a previous study, UHRF1 was identified as an ICL sensor in humans, it was shown that downstream repair factors were not recruited normally to the ICL in the absence of UHRF1. However, the mechanism underlying this process was unknown. Here, we report the identification of UHRF2, a paralogue of UHRF1, as an ICL sensor protein. We show that UHRF2 cooperates with UHRF1, to ensure recruitment of FANCD2 to the ICL. We show that a direct protein-protein interaction with FANCD2 is formed via the SRA domain of UHRF1, and a direct interaction is formed between UHRF2 and UHRF1. Moreover, we demonstrate that the essential monoubiquitination of FANCD2 is stimulated by UHRF1/UHRF2, by mediating its retention on chromatin. Taken together, we uncover the mechanism of FANCD2 activation by monoubiquitination via recruitment and retention at ICLs dependent on an interaction with UHRF1/UHRF2.

Acknowledgments

First I would like to extend my gratitude to my supervisor Martin Cohn for giving me the opportunity to undertake this project. I would like to thank him for his patients with me, and all of his guidance along the way.

I would like to thank all my lab members. Especially, I would like to thank Eric, who was there from the start, for all his help and advice. To my lab besty Anna, thanks for being there to exchange all the gossip and conspiracy theories, we got each other through all of this. Thank you to David and Kate for all your help and just being great people to work with, and putting a smile on my face every day. Thanks to my Russian posse Alina and Nadiya for teaching me the Russian ways and offering all the support. Last, thank you to everyone not mentioned by name from the Cohn, Pears, Lakin, and Watts groups, you are all great. Honorable mentions to Florien for all the philosophical discussions, and to Javier for always making me laugh.

I would like to extend my deepest gratitude to my mom, my dad, my grandmother, and the rest of my family for all of their emotional support. Thank you for always worrying about me and pushing me to keep doing my best, for always believing in me. And of course thank you for investing so heavily into me to give me the opportunity to live my dream of attending a top university, I will never forget this.

Lastly, a big thank you for my wonderful boyfriend Justin. You supported me even when I put an ocean between us, you supported me when I was at my lowest, you were there for me every step of the way. You were always trying to make me see the upside, you always reassured me everything will be ok and never stopped believing in me. You went out of your way to spend time with me in person and online, I never felt alone. I couldn't have done this without you. You mean the world to me, I love you.

Dedication

I would like to dedicate this PhD. thesis to the memory of my late grandfather Peter Hombach. You were my guide and inspiration to always keep aiming higher. You encouraged my curiosity every day with great discussions on all topics, that were over my head most of the time. You set the example of a true scientist; you never stopped learning new things in new fields. You taught me so much over my life, and I will always carry that with me. I hope that someday I will be as knowledgeable about so many things as you were.

In loving memory of Dedushka

Table of Contents

Table of Figures

List of Abbreviations

5hmC	5' hydroxymethyl cytosine
5mC	5' methyl cytosine
A	Adenosine
AML	Acute myelogenous leukemia
AP	Apurinic/apyrimidinic
ATM	Ataxia-telangiectasia mutated (ATM)

ATR	Ataxia telangiectasia and Rad3-related protein
ATRIP	Ataxia telangiectasia and Rad3-related protein interacting protein
BER	Base excision repair
bp	Base pairs
BRCA1	Breast cancer type 1 susceptibility protein
BRCA2	Breast cancer type 2 susceptibility protein
C	Cytosine
CGI	CpG islands
CoIP	Co-immunoprecipitation
CPD	Cyclobutane pyrimidine dimer
CtBP	Ct binding protein
CtIP	CtBP interacting protein
DDR	DNA damage response
DNA	Deoxyribonucleic acid
DNA-PKc	DNA dependent protein kinase catalytic subunit
DNMT1	DNA methyltransferase 1
DNMT3	DNA methyltransferase 3
DSB	Double stranded break
EMSA	Electrophoretic mobility shift assay
FANC	Fanconi anemia complementation
G	Guanidine
GG-NER	Global genomic nucleotide excision repair
HAT	Histone acetyl transferase

HMT	Histone methyl trasnferase
HR	Homologous recombination
ICL	Interstrand crosslink
IP	Immunoprecipitation
K	lysine
MBP	Maltose binding protein
mCh	mCherry
MMC	Mitomycin C
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
p53	Tumor protein p53
PAGE	Polyacrylamide gel electrophoresis
PCNA	Proliferating cell nuclear antigen
PCNP	PEST-containing- nuclear protein
PHD	Plant Homeodomain
PNK	Polynucleotide kinase
Pol	Polymerase
pRB	Retinoblastoma 1 protein
RING	Really Interesting New Gene
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RPA	Replication protein A
SET	Su(var)3–9, Enhancer of zeste (E(z)), and trithorax (trx)

shRNA	Short Hairpin RNA
SRA	SET and RING Associated
SSB	Single stranded break
T	Thymidine
TC-NER	Transcription coupled nucleotide excision repair
TET	Ten-eleven translocase
TLS	Translesion synthesis
TMP	4,5′,8-Trimethylpsoralen, Trioxsalen
TTD	Tandem Tudor Domain
UBL	Ubiquitin-like
UHRF1	Ubiquitin-like, containing PHD and RING finger domains, 1
UHRF2	Ubiquitin-like, containing PHD and RING finger domains, 2
UVA	Ultraviolet A

Chapter 1: Literature Review

1.1 Chromatin Modifications

Everything that makes every individual unique is dictated by their deoxyribonucleic acid (DNA) sequences, a blue print of incredible complexity. The human genome project estimates that human genome is 3 billion base pairs (bp) long spread across 23 pairs of chromosomes (Venter et al. 2001). This staggering quantity of information must be somehow packaged, bookmarked for efficient reading, and maintained so that it can be passed down to the next generation.

1.1.1 Histones

The cell uses an organisation mechanism which wraps DNA around specific proteins like "beads on a string", these beads are called nucleosomes. Nucleosomes consist of two copies for each of four histone proteins: H3, H4, H2A, and H2B with 146 bp of the DNA double helix wrapping around them (Luger et al. 1997). Together, histones and DNA pack into chromatin which can be categorized into two types: euchromatin and heterochromatin. Euchromatin is considered to be transcriptionally active DNA that is loosely packed, while heterochromatin is usually gene poor, full of repetitive sequences and transcriptionally repressed or silent and tightly packed (Tartof et al. 1989). Each histone has a polypeptide tail extending out past the DNA wrapped around the core, this tail is accessible to modifying enzymes which can chemically change the tail and influence how the nucleosome packs into chromatin. The three type of histone tail modifications which are most heavily studied are: phosphorylation, acetylation, and methylation (Strahl & Allis 2000). Acetylation of histone tails promotes chromatin de-compaction and is associated with euchromatin while methylation of histone tails is more in line with compaction and therefore heterochromatin. Phosphorylation of histone tails is not as black

and white, with phosphorylation promoting transcription, either directly or by preceding acetylation (Nowak & Corces 2004). Histone tails are also be poly(ADP-Ribosyl)ated, which affects chromatin dynamics (Thomas et al. 2014); they may also be ubiquitinated during DNA damage response (DDR) (Uckelmann & Sixma 2017).

1.1.2 Methylation

Methylation is one of the chromatin modifying markers found throughout the genome. Methylation can occur directly on DNA, specifically the DNA base cytosine, or on the tails of histones.

1.1.2.1 DNA methylation

Direct DNA methylation has been historically considered as a long term memory for DNA modification, as protein-DNA interactions are far more fragile than a chemical modification of the DNA bases (Razin & Riggs 1980). The mammalian genome is full of regions that are high in CG content, termed CpG islands (CGIs). The best studied CGIs are in transcriptional start sites, where the CGI density is quite high, however CGIs can be found in gene exon regions as well as other locations in the genome (Takai & Jones 2002). CGI methylation is thought to be associated with transcriptional repression, gene silencing, X-chromosome inactivation, and imprinting and is thought to be hereditary and tissue specific (Riggs 1975; Holliday & Pugh 1975). Paradoxically, CGIs at transcriptional start sites are mostly non-methylated with only a few exceptions, which are usually associated with the aforementioned long term silencing.

Most CpG methylation occurs elsewhere in the genome and conspicuously not on CGIs (Bird 1986). As the bulk of the mammalian genome is methylated regardless of the genetic element the true nature of genomic methylation is still under debate. It is hypothesized that CpG

methylation may serve to modulate gene expression through protein interactions with the methylated regions; in fact as the methylation profiles of different tissue types differ, it is thought that methylation contributes to the differential gene expression profiles, however this remains a hypothesis as no conclusive data had been found (Jones 2001).

It is well documented that mammalian cells require methylation when differentiating past the embryonic stem cell stage, therefore methylation is a vital aspect of the genomic landscape. The current model of maintaining the methylation markers on DNA relies on two paths: *de novo* methylation and replication methylation. The *de novo* methylation is handled by a pair of proteins known as DNMT3A and DNMT3B, which establish methylation patterns in the germ line cells (Okano et al. 1999). In the developed organism methylation is then maintained through the actions of DNMT1 and UHRF1 which faithfully copy the methylation mark on the parent DNA strand to the newly synthesized strand during DNA replication (Song et al. 2012; Berkyurek et al. 2014).

1.1.2.2 Histone methylation

Histone tails are a key factor in DNA packaging and chromatin dynamics. The tails of each histone can carry multiple methylation marks at specific lysines. Most methylation marks are on histone H3, specifically at lysines (K)4, K9, and K27 (Strahl et al. 1999). Proteins that could carry out acetylation of histone tails, known as histone acetyl transferases (HATs) were already well studied, while histone methyl transferases (HMTs) were only predicted to exist. Using screens in *S. pombe* and *D. melanogaster*, proteins with SET domains were identified as being part of repressive and enhancing chromatin modulating loci; as such they were labelled as "promiscuous", and laid the foundation for the search for HMTs (Jenuwein et al. 1998). Early work in the exciting new field of repressive epigenetics had found the protein SUV39H1 which

had a conserved SET domain that was found to act as an HMT targeting H3K9 (Melcher et al. 2000; Rea et al. 2000). Later, the protein G9a was discovered to have HMT activity towards H3K9 and K27 (Tachibana et al. 2001). The protein MLL, which is a functional homologue of the *D. melanogaster* trithorax (a SET containing protein) was found to be HMT targeting H3K4 (Yokoyama et al. 2004). The location of the methyl mark alone is not everything in histone methylation, the quantity of methyl groups plays an important role as well. Histones can be mono, di, and tri methylated on each lysine which confer different effects on the genomic landscape around them. Generally, H3K9me3 (tri methylated) is a repressive epigenetic mark frequently found in transcriptionally silenced heterochromatin, while H3K4me3 is found to be associate with active transcription elements, and H3K27me3 is usually present in developmentally controlled repressed genes (Boyer et al. 2006)

Both DNA methylation and histone methylation serve to sculpt the genetic landscape and in fact may be interdependent with each other as proteins that can bridge the two systems do exist.

1.1.3 DNA damage

DNA is constantly under attack from both endogenous and exogenous sources with an estimated 70,000 lesions per day in a cell (Tubbs & Nussenzweig 2017). Such damage may manifest in various forms such as: single and double strand breaks (SSB and DSB respectively), base dimers, base deletions, deamination of bases, abasic sites, base adducts, intrastrand crosslinks, and interstrand crosslinks (ICLs) (Lindahl 1993). This damage can come from within the cell as a by product of cellular metabolism such as aldehydes (Voulgaridou et al. 2011) or reactive oxygen species (ROS). The source of damage can also be external in the form of radiation, both UV and ionizing (Cadet & Wagner 2013). Lastly, genotoxic drugs or chemicals

are also a source of DNA damage. All types of DNA damage are deleterious to the cell, causing mutations, which can lead to cell death or cancer. ICLs are particularly dangerous as they can interfere with transcription, and DNA replication effectively halting them and can lead to chromosomal aberrations and mitotic catastrophe if not properly repaired.

1.1.4 Origins of ICLs

ICLs are incredibly toxic, with 40 ICLs being sufficient to kill a human cell (Dronkert & Kanaar 2001). The best studied ICLs are the ones produced by DNA crosslinking agents used in chemotherapy. Mitomycin C (MMC) is frequently used as a chemotherapy drug and, it forms covalent bonds between guanines on opposite DNA strands, it can also form a mono adduct to one guanine not fully crosslinking the strands. MMC is not chemically active until it enters the cellular environment where it is activated by enzymatic reactions opening its aziridine ring, which is unstable and reactive. When examining the structure of MMC complexed with DNA, it is observed that it resides in the minor groove of the helix creating minimal distortion (Tomasz 1995).

Cisplatin is another common drug which also links two opposing guanine bases, the frequency of single adducts with cisplatin is lower than that my MMC. Cisplatin also sits in the minor groove of the DNA helix; however, the distortion caused by the ICL is very large causing the two cytosines flanking the crosslinked guanines to flip out and the helix to unwind 110° distorting the major groove (Malinge et al. 1999).

Lastly, psoralen is a class of photosensitive ICL agents, more commonly used for skin conditions such as psoriasis. Psoralens are planar molecules that can intercalate DNA but cause no harm unless they are photo activated by UV light. To form a crosslink the psoralen should be intercalated at a TpA site and undergo two photo activation events linking the opposing

thymidines together, as such mono adducts are common with psoralens. Psoralen crosslinks cause a very minor distortion of the DNA helix (Hwang et al. 1996).

ICLs can also occur endogenously in the body as by products of metabolism. Metabolism of alcohol and fatty acids yield aldehydes which are the current candidates for endogenously formed ICLs (Pang & Andreassen 2009; Langevin et al. 2011).

1.2 DNA Repair

All forms of damage to DNA must be repaired as quickly and accurately as possible to ensure no long term damage, in the event that repair is impossible the cells will either die via necrosis or undergo programmed cell death known as apoptosis. As the frequency of DNA lesions is quite high and occurs every day, multiple mechanism have evolved, which carry out repaire and are collectively called the DNA damage response (DDR). These pathways carry out independent repair, as well as collaborate in repairing more complicated types of lesions.

1.2.1 BER

Base excision repair (BER) is a system that is used by the cell to repair chemically modified bases and abasic sites (AP), additionally BER can also repair single stranded breaks (SSBs). While modified bases do not pose a direct threat to DNA replication as translesion polymerases (TLS) can replicate past the damaged site, the TLS pol is not accurate and point mutations can be introduced.

The BER pathway relies on four core proteins: a DNA glycosylase, an AP endonuclease, a DNA polymerase, and a DNA ligase. There are several types of DNA glycosylases as they are lesion specific. Their function is to remove the base from the ribose sugar by cleave the glycosidic bond, this leaves an AP site (Krokan et al. 2000). The AP site is more toxic than the modified base and should be further processed. To that end, the AP endonuclease cleaves the

DNA backbone 5' to the AP site to generate a 3' OH. Some glycosylases (like NEIL1/2) leave a phosphate on the 3' end, which requires processing by a polynucleotide kinase (PNK) to remove the phosphate group and again leave a 3'OH (Hegde et al. 2008). The resulting gap can be filled in either by a short or long patch repair. Short patch repair is carried out by polymerase β (Horton et al. 2008). In the case of long patch repair where the DNA ends are more difficult to process, the polymerase used is pol δ/ε which adds extra nucleotides peeling off bases as it goes and the nuclease FEN1/XRCC1 trims off the leftover flap. Finally the nick in the backbone is ligated together by DNA ligase I, thus repairing the damage (Parsons & Dianov 2013).

1.2.2 NER

Nucleotide excision repair (NER) is used by the cell to repair bulky DNA lesions. This pathway is the means by which cyclobutane pyrimidine dimers (CPDs) caused by UV radiation are repaired. People with a genetic defect in early NER proteins have a condition called Xeroderma Pigmentosum (XP) or Cockayne syndrome, which renders them extremely sensitive to UV light among other anomalies (Cline & Hanawalt 2003).

The NER pathway has two modes of action, either through transcription coupled repair (TC-NER) or global genomic repair (GG-NER). In GG-NER the DNA lesion is recognized by DDB2-DDB1 which recruit the XPC-Rad23B-Centrin2 complex to the opposite strand of the lesion creating a kink, which explains why many lesion types can be recognized by the same complex (Lee et al. 2014). Once recognized the local DNA is melted by the XPC-Rad23B-Centrin2 complex and the TFIIH complex is recruited. In the case of TC-NER, the RNA polymerase II (RNA pol II) stalls when it encounters a CPD, this acts as the recognition step recruiting TFIIH. TFIIH is a multi protein complex, it includes XPB and XPD which unwind the local DNA creating a 10-20 nucleotide bubble. XPG, RPA and XPA are then recruited into the

TFIIH complex. RPA stabilizes the ssDNA and XPG acts as a support. XPA then further recruits ERCC1-XPF which together with XPG act as nucleases cleaves on either side of the DNA lesion. Then PCNA and pol δ/κ (for non-replicating cells) or ε (for replicating cells) synthesize new DNA from the other strand and displace the cut out lesion. DNA ligase I then ligates the DNA together (Spivak 2015).

1.2.3 NHEJ

Among the most catastrophic types of DNA damage is the double stranded DNA breaks (DSB). If a DSB occurs and is not quickly repaired, whole genes or even chromosomes can be lost leading to massive mutagenesis or cell death. DSBs occur when a replication fork collapses or ionising radiation cleaves both the phosphodiester backbones of the DNA helix. Non-homologous end joining is a repair mechanism used by cells to quickly fix DSBs. This mechanism unfortunately is error prone as the repair machinery just ligates two DNA ends together regardless of if they were joined to begin with. As the ends of the DNA are in no way processed, basses are frequently lost or even added in this process, which is highly mutagenic. This pathway however is useful as it functions in all stages of the cell cycle and does not require homologous sequences for repair.

NHEJ mediated repair is initiated with the recruitment of a heterodimer called Ku, consisting of a pair of proteins called Ku70 and Ku80. Ku is very dynamic within the nucleus associating with DSB ends within seconds of their formation (Mari et al. 2006). Once Ku is bound to the DNA ends it protects them from non-specific processing, and recruits DNA-PKc-Artemis to the site. The binding of DNA-PKc to the DNA activates the enzyme's kinase activity, which likely autophosphorylates itself. This causes the Ku complex to move away from the break into the DNA exposing the ends to further processing (Uematsu et al. 2007). Artemis is an

endonuclease that can do some end processing. Ku also directly interacts with XRCC4/DNA ligase IV where XRCC4 stabilizes the DNA ligase (Nick McElhinny et al. 2000) also part of the complex are PAXX and XLF (Zhou et al. 2016). The ligase then sticks the two DNA ends together (Lieber 2008; Davis & Chen 2013). There is also another NHEJ based pathway which uses more end processing and micro homology called alternative NHEJ, reviewed in (Deriano & Roth 2013).

1.2.4 HR

The most error free way to repair DSBs is through homologous recombination (HR). Unfortunately HR requires an intact copy of the DNA to be present for the repair process and as such only functions in late S-phase and G2.

To initiate HR the MRE11-Rad50-NSB1 complex (MRN) binds to the DNA ends. MRE11 has nuclease activity which is stimulated by NSB1 (Paull & Gellert 1998). The binding of MRN recruits and activates ataxia-telangiectasia mutated (ATM) through NBS1. ATM is an important checkpoint kinase in DDR, which arrests replication and promotes DSB repair (Difilippantonio et al. 2005). Once MRN is bound to the DSB, CtBP interacting protein (CtIP) is also recruited and together they carry out limited 5' to 3' resection at the DSB site (Sartori et al. 2007). The resection is then further extended by Bloom syndrome protein (BLM), DNA2, and Exo1 producing a long stretch of ssDNA (Gravel et al. 2008; Mimitou & Symington 2009). This ssDNA is rapidly coated by replication protein A (RPA) creating a DNA-protein filament to stabilize the DNA and protect it. The RPA coated DNA stimulates another important kinase, ataxia telangiectasia and Rad3-related protein (ATR), through its interacting partner ataxia telangiectasia and Rad3-related protein interacting protein (ATRIP) (Zou & Elledge 2003). ATR signaling has many functions some of which are still poorly understood. ATR is known to

phosphorylate Chk1 and RPA as well as maintain integrity of PCNA and pol ε at the replication fork (Cimprich & Cortez 2008). The RPA is then displaced by Rad51 with the assistance of breast cancer type 2 susceptibility protein (BRCA2) and breast cancer type 1 susceptibility protein (BRCA1). BRCA2 and its interacting partner PALB2 load Rad51 into ssDNA and prevent loading on dsDNA using the conserved BRC domains on BRCA2 which are thought to act like an assembly line for Rad51 polymerization (Powell & Kachnic 2003). The Rad51 coated ssDNA can then perform a homology search to find the complimentary DNA in the sister chromatid. Once homology is found the ssDNA stand invades the intact duplex and crease a D-loop structure priming the site of DNA synthesis (Heyer et al. 2010).The D-loop creates a double stranded DNA duplex which allows for DNA pol to extend the resected end using the sister chromatid as a template, and DNA ligase fills in the nick. The intertwined DNA duplexes create two junctions called holiday junctions which can migrate back and forth on the DNA. In the case of the two junctions migrating towards each other they can be "dissolved" by the activity of the BLM/TOPIIIα/RMI1 complex. The junctions can also be resolved by MUS81/EME1(Mms4), GEN1, and SLX1/SLX4 activity. Regardless of the method the result is restoration of the two separate DNA duplexes, thus creating two intact error free DNA sequences (Jasin & Rothstein 2013).

1.2.5 Fanconi Anemia Pathway

Fanconi Anemia (FA) is a rare genetic disorder affecting approximately 1 in 200 000–400 000 live births and first described by Guido Fanconi in 1927 (Mathew 2006). It is characterized by distinctive phenotypes such as: short stature, congenital abnormalities, bone marrow failure and a predisposition to cancers (Auerbach 2009). These patients will develop aplastic anemia requiring bone marrow transplantation, they are also susceptible to acute

myelogenous leukemia (AML) and solid tumor formations, especially of the anogenital regions and head and neck (Dong et al. 2015). The cancer predisposition of FA patients is due to their inability to repair ICL damage which may arise from exogenous or endogenous sources. This deficiency in damage repair stems from a defect in one of twenty two genes responsible for the FA pathway proteins (see Table 1). Typically, for a protein to be characterized as an FA gene it must be found to be mutated in an FA patient and complementation of this gene must restore normal tolerance to ICLs. As such it is worth noting that FANCM may not be an FA gene as it was found in only a single patient which also had with a FANCA mutation (Meetei et al. 2005; Singh et al. 2009).

Name	Alternate name	Function	Reference of identification
FANCA		Core complex subunit	(Lo Ten Foe et al. 1996)
FANCB		Core complex subunit	(Meetei et al. 2004)
FANCC		Core complex subunit	(Strathee et al. 1992)
FANCE		Core complex subunit	(De Winter et al. 2000)
FANCF		Core complex subunit	(de Winter et al. 2000)
FANCG	XRCC9	Core complex subunit	(de Winter et al. 1998)
FANCL		Core complex subunit: E3 ligase	(Meetei et al. 2003)
FANCM		Core complex subunit	(Meetei et al. 2005)
FANCT	UBE2T	Core complex subunit: ubiquitin E2	(Rickman et al. 2015; Hira et al. 2015)
FANCD2		Target of core complex Central pathway regulator	(Timmers et al. 2001)
FANCI		Target of core complex Central pathway regulator	(Dorsman et al. 2007; Sims et al. 2007)
FANCD1	BRCA2	Part of HR	(Howlett 2002)
FANCJ	BRP1	Part of HR	(Litman et al. 2005)
FANCN	PALB2	Part of HR	(Xia et al. 2007)
FANCO	RAD51C	Part of HR	(Vaz et al. 2010)
FANCP	SLX4	Nuclease scaffold protein Interacts with SLX1, MUS81/EME1, XPF/ERCC1	(Yamamoto et al. 2011; Stoepker et al. 2011; Kim et al. 2011)
FANCQ	XPF	Endonuclease	(Bhagwat et al. 2009; Bogliolo et al. 2013)
FANCR	RAD51	Part of HR	(Ameziane et al. 2015; Wang et al. 2015)
FANCS	BRCA1	Part of HR	(Sawyer et al. 2015)
FANCU	XRCC2	Part of HR	(Park et al. 2016)
FANCV	REV7	Subunit of TLS pol ζ (in conjunction with REV3 and REV1)	(Bluteau et al. 2017)
FANCW	RFWD3	Part of HR	(Knies et al. 2017)

Table 1. FA proteins and their functions

The FA pathway consists of four basic steps (Figure 1) which can be broken down as: recognition, activation, repair, and dissociation. The FA pathway is considered to be only active in S-phase and G2 as it uses elements from the HR pathway as part of the repair process. It is the predominant belief in the field that the ICL is recognised and subsequently repaired when it is approached from either one or both sides by a replication fork causing stalling (Zhang et al. 2015). In the recognition step UHRF1, which is predominantly interacting with chromatin

encounters an ICL and through preferential binding affinity is retained at the damage site (Liang et al. 2015), this can occur in any stage of the cell cycle. It has also been suggested that FANCM can act as a sensor for ICLs when a replication fork collides with the ICL. FANCM can then traverse the lesion leaving behind the ICL to be repaired by the canonical ICL repair pathway in early S-phase (Huang et al. 2013). Once the ICL is recognized the central proteins of the pathway, FANCD2 and FANCI are recruited to the ICL site by UHRF1. FANCD2/FANCI are at this point also phosphorylated by ATR/ATM. In the activation stage FANCD2/FANCI which are now on DNA become ubiquitinated. This is carried out by the FA core complex, which was recruited to the ICL by FANCM. The ubiquitination of FANCD2 and FANCI now locks these proteins onto DNA and sets them into their active conformation (Liang et al. 2016).

The now active FANCD2/FANCI signals repair proteins to initiate the repair process. ICL repair borrows aspects from several DNA repair pathways, primarily NER and HR. The initial step in the repair is termed ICL unhooking which is carried out by XPF-ERCC1 in conjunction with the scaffold protein SLX4. The unhooking step cleaves the DNA on either side of the ICL on one strand, creating a ssDNA break (Klein Douwel et al. 2014). With the ICL still attached to one strand the polymerase ζ-REV1 is recruited to the site and carries out TLS restoring one intact DNA duplex (Budzowska et al. 2015). The DNA strand that was cleaved before is now repaired through HR. Lastly the duplex bearing the ICL adduct is repaired by NER. The final step is dissociation, where USP1/UAF, the deubiquitinating enzyme, remove the ubiquitins on FANCD2/FANCI and the complex dissociates and repair is complete (Cohn et al. 2009).

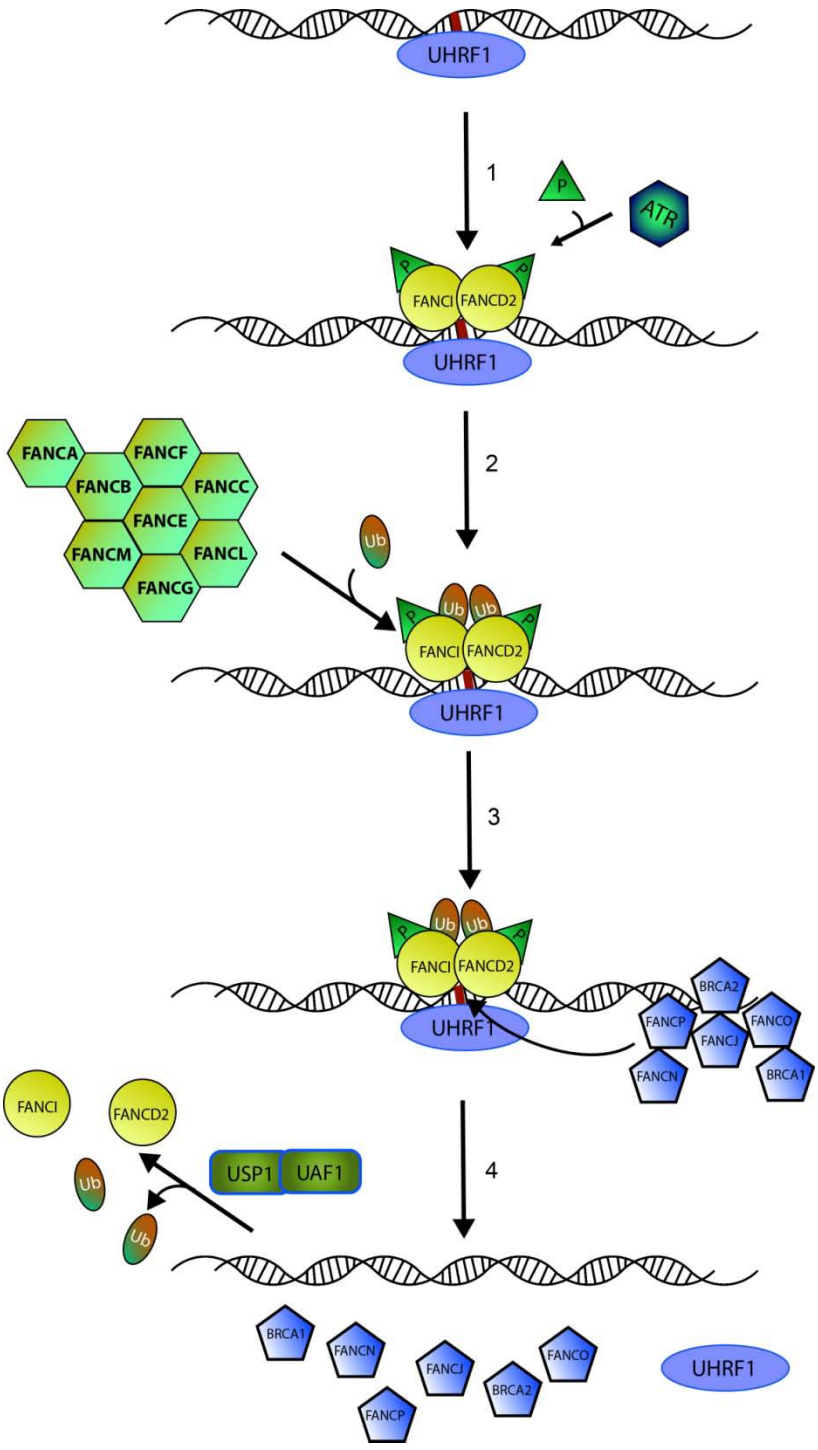
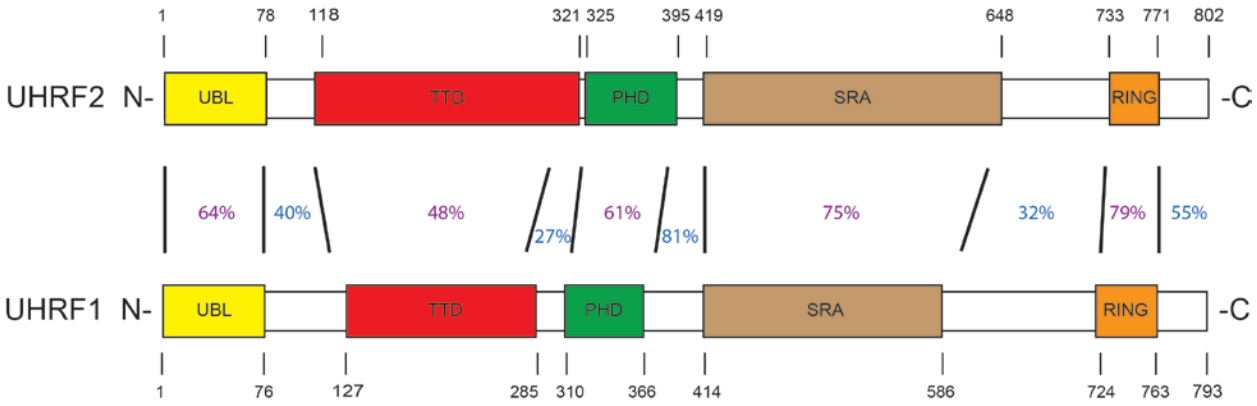


Figure 1. Simplified FA pathway

A simplified schematic depicting the key steps in the FA pathway. 1) UHRF1 detects ICL and recruits FANCD2/FANCI which are phosphorylated by ATR (on or off DNA). 2) FA core complex ubiquitinated FANCD2/FANCI locking them onto the DNA. 3) Now active FANCD2/FANCI signal downstream effecter proteins to carry out repair. 4) FANCD2/FANCI are deubiquitinated and all components diffuse away from the now repaired ICL.

1.3 The UHRF family

The Ubiquitin-like containing PHD and RING finger domains (UHRF) family consists of two proteins, UHRF1 and UHRF2. These two proteins share identical domain organizations however, are only 54% identical on an amino acid sequence level. UHRF1 is evolutionarily conserved to *Danio rerio* while UHRF2 appears to be a more recent evolutionary change, appearing in *Xenopus laevis*. UHRF2 was likely the result of a gene duplication event and has since evolved to have its own function in the cell. Both proteins have: a ubiquitin-like domain (UBL), which structurally resembles ubiquitin, Tandem-Tudor domain (TTD) and Plant Homeodomain (PHD) which interact with histone methylation markers, SET and RING associated domain (SRA) which is a DNA binding domain, and Really interesting new gene domain (RING) which is an E3 ligase (Figure 2). Both of these proteins have been implicated in cancer progression (Bronner et al. 2007; Mousli et al. 2003; Lu et al. 2016; Wang et al. 2012). However, UHRF1 and UHRF2 each appear to serve different roles in the cell and are in fact not redundant.



	UHRF1		UHRF2	
	Function	**Interaction**	**Function**	**Interaction**
UBL domain	Unknown	-	Unknown	-
TTD domain	Unknown	H3K9me2 binding Ligase 1 + G9a	Unknown	H3K9me2/3 binding
PHD domain	Unknown	Histone H3 binding	Unknown	Histone H3 binding
SRA domain	DNA methylation maintenance	Hemi 5mC binding DNMT1 ICLs	Unknown	Full 5hmC binding
RING domain	E3 ligase ubiquitinates substrates	DNMT1 DNMT3A TIP60 UHRF1 (auto ub)	E3 ligase ubiquitinates substrates	DNMT3A TIP60 PCNP Cyclin D1 E1

Figure 2. UHRF1 and UHRF2 homology

Schematic representation of UHRF1 and UHRF2 depicting the amino acid sequence identity obtained from NCBI BLAST. Overall sequence identity is 54% with individual domain identity shown in purple and inter domain regions shown in blue. The interactions of each domain are listed below the diagram.

1.3.1 UHRF1

UHRF1 has been heavily studied for its role in methylation maintenance in the cell. UHRF1 interacts with DNMT1 and carries out methylation copying after DNA replication to copy over the methylation epigenome (Sharif et al. 2007; Bostick et al. 2007). The SRA domain of UHRF1 has a high binding affinity for hemi 5' methyl cytosine (5mC), which then binds the DNA after replication through a base flipping mechanism exposing the DNA to DNMT1 which then places a methyl group on the opposing cytosine thereby creating a fully methylated DNA, which UHRF1 has a low affinity for and the complex dissociates (Hashimoto et al. 2008; Avvakumov et al. 2008). Later research has shown that the TTD domain of UHRF1 also plays a role in the correct localisation of UHRF1 to 5mC sites as the UHRF1 TTD domain has an affinity for H3K9me3 (Rothbart et al. 2012; Liu et al. 2013). The interaction of UHRF1 with DNMT1 is also further stabilized by the deubiquitinating enzyme USP7 (Felle et al. 2011). Loss of UHRF1 is lethal in differentiated cells, it is speculated that this is due to loss of methylation. Only non differentiated embryonic stem cells are able to tolerate the loss of UHRF1, simply due to the fact that they can tolerate methylation loss (Bostick et al. 2007).

The role of UHRF1 in DNA repair has only recently been confirmed, demonstrating the involvement of UHRF1 in the FA pathway as a sensor for ICLs (Liang et al. 2015; Tian et al. 2015). This role warrants further examination.

1.3.2 UHRF2

UHRF2 had been considered for a long time to be redundant with UHRF1, with its true functions being poorly understood. Initial research into UHRF2 function found that it interacted with PEST-containing nuclear proteins (eg. c-Myc, cyclin D1, and cyclin E1) and was capable of ubiquitinating them to target them for degradation, giving UHRF2 a role in cell cycle regulation (Mori et al. 2004). It was also later found that UHRF2 was implicated in the intranuclear

degradation of polyglutamine aggregates, again through ubiqutination, implicating UHRF2 in neural cell health (Iwata et al. 2009). Further research into UHRF2 interaction unearthed that UHRF2 interacts with cyclins (A2, B1, D1, and E1), cyclin-dependent kinases (CDK2, CDK4, and CDK6), retinoblastoma 1 protein (pRB), tumor protein p53 (p53), and proliferating cell nuclear antigen (PCNA), firmly implicating UHRF2 in cell cycle control (Mori et al. 2012).

More recently UHRF2 was found to be a specific reader of 5' hydroxymethyl cytosine (5hmC), which has been reported to be a stable epigenetic mark (Branco et al. 2012; Spruijt et al. 2013). 5hmC is a stable degradation intermediate of 5mC trough the ten-eleven translocase (TET) enzyme pathway (Tahiliani et al. 2009). The domain responsible for this 5hmC recognition is the SRA domain of UHRF2, which has a high binding affinity for both hemi and fully 5hmC DNA, a function that the SRA domain of UHRF1 cannot match (Zhou et al. 2014). This 5hmC recognition has drawn interest in the role of UHRF2 in neurons as those are the type of cells with a higher proportion of this epigenetic mark (Kriaucionis & Heintz 2009). One study found that loss of UHRF2 in mice while viable renders the mice susceptible to seizures (Liu et al. 2017). The interest in UHRF2 in differentiated cells such as neurons is congruent with its distribution as UHRF2 is expressed in all cell types but expression level is low in non differentiated cells and far higher in terminally differentiated cells, suggesting that UHRF2 is important in cell differentiation (Mori et al. 2012).

There are some aspects of UHRF2 that are still poorly understood. UHRF2 has an increased affinity for H3K9me2/me3 histones through its TTD and PHD domain which causes UHRF2 to be predominantly bound and interacting with pericentric heterochromatin (Pichler et al. 2011). The purpose of this interaction is presently unknown. UHRF2 has also been shown to interact with DNMT1 much like UHRF1 during S-phase, however unlike UHRF1 it is unable to

carry out methylation maintenance and so the purpose of this interaction is still unknown (Zhang et al. 2011). However, UHRF2 has been shown to interact with DNMT3A in regulating *de novo* methylation (Jia et al. 2016). Lastly, the role of UHRF2 in DNA damage repair has not really been addressed. There was one study that demonstrated the importance of UHRF2 in DSB repair in smooth muscle cells (Luo et al. 2013). As research this aspect of UHRF2 is lacking, it is therefore important to further examine the role of UHRF2 in DNA repair.

Chapter 2: Materials and Methods

2.1 Cell Lines, plasmids and antibodies

HeLa and HEK293T cells were grown in DMEM (D5796, Sigma) supplemented with 2.5-10% FBS. Antibodies used were as follows: anti-UHRF1 antibody (sc-373, Santa Cruz Biotechnology); anti-FANCD2 (sc-20022, Santa Cruz Biotechnology); anti-α-Tubulin (5829, Millipore); anti-UHRF2 (sc-54252, Santa Cruz Biotechnology T18; sc-398953, Santa Cruz Biotechnology C-10); anti-p300 (sc-584, Santa Cruz Biotechnology) and anti-HA (mouse monoclonal antibody clone 12CA5). EGFP-fused FANCD2 and mCherry-fused UHRF1 cDNA were expressed using a derivative of the pOZ-N plasmid (Nakatani & Ogryzko 2003). EGFP-fused UHRF2 cDNA was expressed using a derivative of the pOZ-C plasmid. shRNA-mediated knockdown of the UHRF1 and UHRF2 genes was achieved by expressing the target sequences 5’-AGATATAACGTTAGGGTTT-3’ and 5'-TGTAAAGGCTGGTGAAAGA-3', in the pSuper.retro vector (Clontech). Transfections of plasmid DNA were carried out using FuGENE6 (Promega) according to the manufacturer’s instructions. The UHRF2 domain deletion plasmids were generated as above. The deleted regions of UHRF2 are amino acids 1-78 for the UBL domain, 118-321 for the TTD domain, 325-395 for the PHD domain, 419-648 for the SRA domain and 733-802 for the RING domain.

2.2 Cell Lysates

Cells were scraped off the dishes, and centrifuged at 1,000 rpm for 5 minutes. Cell pellets were resuspended and incubated in equal volume of Benzonase buffer (2mM $MgCl_2$, 20mM Tris pH 8.0, 10% glycerol, 1% Triton X-100 and 12.5 units/ml Benzonase (E1014, Sigma) on ice for 10 minutes. The cells were then lysed by the addition of an equal volume of 2% SDS to reach a final concentration of 1%. Samples were heated at 70°C for 2 minutes. The protein concentration was determined by Bradford assay (Bio-Rad Life Science).

2.3 CRISP-Cas9

HeLa UHRF2 -/- cells were generated using plasmid pX459 (Addgene #48139)(Ran et al. 2013) . The targeting sequence used in the sgRNA was: 5'-GTGCCCGTCTTATTGATCC-3'. Primers, 5'- caccgGTGCCCGTCTTATTGATCC -3' and 5'-aaacGGATCAATAAGACGGGCACc -3', were annealed and introduced into the pX459 plasmid through its BbsI site. HeLa cells were transfected with 2 μg of the resulting pX459 plasmid and selected with 4 μg/ml puromycin after 24h. After another 24h cells were plated at low density and clones were picked after 2 weeks. Clones were analyzed using immunoblot analysis and genotyped by sequencing the gRNA target site of the gene locus.

2.4 Clonogenic Survival Assay

Cells (250–4,000) were plated in 6-well plates and treated with different dosages of mitomycin C (MMC). For TMP/UVA treatment, the cells were treated with 50ng/ml 4,5′,8-trimethylpsoralen (TMP) for 30 minutes, and irradiated with the UVA dosages indicated. Colony formation was scored after 10-14 days after fixing the colonies and staining them using 1% (w/v) crystal violet in methanol.

2.5 Microscopy

EGFP-fused FANCD2, EGFP-fused UHRF2, and mCherry-fused UHRF1 cDNA were inserted into the pOZ vector as described above. Live cell imaging were carried out with an OLYMPUS IX81 microscope connected to PerkinElmer UltraView Vox spinning disk system equipped with a Plan-Apochromat 60x/1.4 oil objective using Volocity software 6.3 for image capturing. EGFP and mCherry were excited with 488 nm and 561 nm laser lines, respectively. Throughout the experiment, these cells were maintained at 5% CO_2, and 37°C using a live cell environmental chamber (Tokai hit). Confocal image series were typically recorded with a frame size of 512x512 pixels and a pixel size of 139 nm. For localized DNA damage induction, cells were seeded in glass bottom dish (MatTek) and sensitized by incubation in DMEM supplemented with 2.5% FBS and 20 μg/ml 4,5′,8-trimethylpsoralen (TMP) for 30 min at 37°C. Microirradiation was performed using the FRAP preview mode of the Volocity software by scanning (each irradiation time was 100 ms) a preselected area (50x3 pixels) within the nucleus 20-75 times with a 405nm laser set to 100% laser power. The mCherry and EGFP intensities at microirradiated sites were quantified using ImageJ with Fiji, and normalized by their intensities before microirradiation.

2.6 IP

2.6.1 In vivo

FLAG-HA tagged recombinant proteins were inserted into pOZ vectors with an IL-2 receptor marker. The proteins were stably expressed using a retroviral system. The positive cells were selected using magnetic Dynabeads™ (ThermoFisher) beads coupled to IL-2 antibody. The cells were lysed in a hypotonic buffer (10 mM KCL, 10 mM Tris pH 7.3, and 1.5 mM $MgCl_2$) using a dounce homogenizer (25 strokes). The lysate was spun for 15 min at 4k RPM to separate the nuclear pellet and cytoplasmic fraction. The nuclear pellet was resuspended in low salt buffer

(0.02 M KCl, 20 mM Tris pH 7.3, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 25% Glycerol, supplemented with 1µl/ml 25 unit/µl Benzonase, 0.1% Tween 20). Post incubation a high salt buffer (1.2 M KCl, 20 mM Tris pH 7.3, 1.5 mM $MgCl_2$, 0.2 mM EDTA and 25% Glycerol) was added dropwise in a 1:1 ratio with the low salt buffer. The mixture was incubated for 30 min, centrifuged at 13.3k RPM, then dialyzed against BC-100 buffer (100 mM KCl, 20 mM Tris pH 7.3, 0.2 mM EDTA, and 20% glycerol). The sample was then collected and centrifuged again at 13.3k RPM, the supernatant was mixed with Protein A sepharose coupled with anti-HA IgG and incubated for 1 hour at 4℃ with slow rotation. The beads were then transferred to Micro Bio-Spin Chromatography Columns (Bio-Rad), washed with BC-100 buffer. The proteins were eluted in 0.5 mg/ml HA peptide (Sigma) in BC-100 for 2 hours.

2.6.2 In vitro

Recombinant UHRF1, UHRF2 and FANCD2 were purified from Sf9 insect cells. MBP was purified from *E.coli*. 1µg of each protein was mixed in the reaction buffer containing 10µg BSA (NEB), 20 mM Tris-HCl pH 8.0, 150 mM KCl, 10% Glycerol, 0.05% Tween-20, 2 mM β-ME and 0.2 mM PMSF in 10µl. The mixture was first incubated at 37°C for 1h for protein complex formation. Protein A sepharose coupled with anti-HA IgG were added subsequently, and the mixture was incubated at 4°C with gentle mixing for 30 minutes. The mixture was then transferred to Micro Bio-Spin Chromatography Columns (Bio-Rad), and washed with the reaction buffer supplemented with 0.1% Tween-20. The proteins were eluted in a buffer contained 100 mM Tris-HCl pH 6.8, 100 mM KCl, 0.1% Tween 20, 0.2mM EDTA, 10% glycerol, 0.1% Tween-20 and 0.5 mg/ml HA peptide (Sigma). In the case of His-tag purification the proteins were washed with the reaction buffer supplemented with 0.1% Tween-20 and 20 mM Imidazole. The proteins were eluted in the wash buffer containing 200 mM Imidazole.

2.7 DNA Probes

The DNA oligos were annealed in the buffer containing 10 mM Tris-HCl pH7.5, 100 mM NaCl and 1mM EDTA. 4,5′,8-trimethylpsoralen (TMP, Sigma, T6137)/UVA (365nm) crosslinking induction was described previously (Espositos et al. 1988). Interstrand crosslink was confirmed by 8 M urea 20% denaturing polyacrylmide gel electrophoresis.

2.8 EMSA

EMSA was performed as previously described (Cohn et al. 2001) with the following modifications: The binding reaction that contained 1 μg of UHRF1 or UHRF2 and 1 nM radiolabeled DNA, was performed in 10μl containing 25 mM Tris-HCl pH 8.0, 100 mM NaCl, 6% glycerol, 1 mM dithiothreitol (DTT), 5 ng poly(dI·dC)-poly(dI·dC)and 1μg bovine serum albumin (BSA, New England Biolabs). For super-shift 2μg anti-HA antibody was added.

2.9 Recombinant Protein Purification

Proteins purified from Sf9 cells were expressed using the pFastBac1 vector (Life Technologies) with an engineered N-terminal Flag-HA tag. Cell pellets were resuspended in lysis buffer (20 mM Tris- HCl pH 8.0, 0.1M KCl, 10% glycerol, 0.1% Tween-20, 2mM β-ME and 0.2mM PMSF). Lysates were clarified by centrifugation, and the supernatants were incubated with M2 anti-FLAG agarose resin for 2 hr. The resin was washed extensively, and the protein was eluted in the same buffer containing 0.5 mg/ml FLAG peptide, but excluding Tween-20. 6xHis-MBP was purified from BL21(DE3) cells using the pET28a vector (plasmid kindly provided by Dr. Mark Howarth) following standard purification methods.

2.10 FANCD2 *in vitro* ubiquitination

Reaction volumes of 25 μl contained 17nM UBA1, 0.64 μM UBE2T or UBCH5c, 1.86 μM FLAG-HA-FANCL or FLAG-HA-UHRF1 or His-UHRF2, 4.2 μM His-Ub, 0.25 μM

FANCD2–FANCI complex, 20 μM pBlueScript SKII (+) when indicated, in the following reaction buffer: 50 mM Tris (pH 7.5), 100 mM KCl_2, 2 mM $MgCl_2$, 0.5 mM dithiothreitol and 2 mM ATP. Reactions were incubated at room temperature for the indicated time. In all, 5μl of 6x SDS loading buffer containing DTT was used to terminate reactions. Samples were loaded onto an SDS-PAGE gel and subjected to Coomassie blue staining or immunoblotting. This protocol was adapted from (Liang et al. 2016). Recombinant FANCDL, UBA1, UBE2T and Ubiquitin were purified from Sf9 insect cells by lab members. Recombinant UHRF1, UHRF2 and UBCH5c were purified from Sf9 insect cells by me.

2.11 DNA nucleoside analysis by mass spectrometry (HPLC–QQQ)

1 μg of DNA in 200 μl of water was added to 200 μl of hydrolysis solution (100 mM NaCl, 20 mM MgCl2, 20 mM Tris pH 7.9, 1000 U/ml Benzonase, 600 mU/ml Phosphodiesterase I, 80 U/ml Alkaline phosphatase, 36 μg/ml EHNA hydrochloride, 2.7 mM deferoxamine). The mixture was incubated for two hours and then lyophylised by SpeedVac.

The lyophylisate was resuspended in 1000 μl of buffer A and 300 μl was transferred into an LC-MS vial for analysis. A sample 100 times more dilute was prepared by dilution 5 μl of the original sample into 495 μl of Buffer A.

For the analysis by HPLC– QQQ mass spectrometry, a 1290 Infinity UHPLC was fitted with a Zorbax Eclipse plus C18 column, (1.8 μm, 2.1 mm 150 mm; Agilent) and coupled to a 6495a Triple Quadrupole mass spectrometer (Agilent Technologies) equipped with a Jetstream ESI-AJS source. The data were acquired in dMRM mode using positive electrospray ionisation (ESI1).

The AJS ESI settings were as follows: drying gas temperature 230 °C, the drying gas flow 14 $lmin^{-1}$, nebulizer 20 psi, sheath gas temperature 400 °C, sheath gas flow 11 $lmin^{-1}$,Vcap 2,000 V and nozzle voltage 0 V. The iFunnel parameters were as follows: high pressure RF 110 V, low pressure RF 80 V. The fragmentor of the QQQ mass spectrometer was set to 380 V and the delta EMV set to +200. The gradient used to elute the nucleosides started by a 5-min isocratic gradient composed with 100% bufferA (10 mM ammonium acetate, pH 6) and 0% buffer B (composed of 40% CH3CN) with a flow rate of 0.400 ml min^{-1} and was followed by the subsequent steps: 5-8 min, 94.4% A; 8–9 min, 94.4% A; 9–16min 86.3% A; 16–17 min 0% A; 17– 21 min 0% A; 21–24.3 min 100% A; 24.3–25min 100%A. The gradient was followed by a 5min post time to re-equilibrate the column.

The raw mass spectrometry data was analysed using the MassHunter Quant Software package (Agilent Technologies, version B.07.01). The transitions and retention times used for the characterization of nucleosides and their adducts are summarized in Supplementary table below. For the identification of compounds, raw mass spectrometry data was processed using the dMRM extraction function in the MassHunter software. For each nucleoside, precursor ions corresponding to the $M\text{-}H^+$ and $M\text{-}Na^+$ species were extracted, and the average of the signal observed from each target ion weighted by response was used for quantification. To utilise the linear range of for each nucleoside, the quantifications of dC, dG and dA were carried out with the diluted samples and quantification of dT and mdC was carried out with the concentrated sample.

Compound	Adduct	Precursor Ion (*m/z*)	MS1 Resolution	Product Ion (*m/z*)	MS2 Resolution	Retention Time (min)	Dwell Time (ms)	CE (V)	CA (V)	Polarity
dC	dC-H^+	228	Unit	112	Unit	3.5	500	10	4	Positive
	dC-Na^+	250	Unit	134	Unit	3.5	500	10	4	Positive
5hmdC	hmdC-H^+	258	Unit	142	Unit	4.1	500	12	4	Positive
	hmdC-Na^+	280	Unit	164	Unit	4.1	500	12	4	Positive
5mdC	mdC-H^+	242	Unit	126	Unit	8.7	500	10	4	Positive
	mdC-Na^+	264	Unit	148	Unit	8.7	500	10	4	Positive
dG	dG-H^+	268	Unit	152	Unit	9.4	500	10	4	Positive
	dG-Na^+	290	Unit	174	Unit	9.4	500	10	4	Positive
T	dT-H^+	243	Unit	127	Unit	10.9	500	10	4	Positive
	dT-Na^+	265	Unit	149	Unit	10.9	500	10	4	Positive
dA	dA-H^+	252	Unit	136	Unit	13.4	500	10	4	Positive
	dA-Na^+	274	Unit	158	Unit	13.4	500	10	4	Positive

Method written by Paolo Spingardi

Chapter 3: Results

3.1 UHRF2 interacts with ICLs *in vivo* and *in vitro*

3.1.1 Introduction

When the cell encounters a DNA lesion, it must be swiftly repaired. In the case of ICLs, the cell relies on the FA pathway. Before any repair takes place, the damage must first be recognized. UHRF1 was recently published as a sensor for ICLs (Liang et al. 2015), however, the protein UHRF2 was initially identified from the same IP by our research group. UHRF2 contains five domains, three if which have the potential to interact with DNA, the TTD, PHD, and SRA domains.

The SRA domain has been characterized as a specific DNA binding domain. In *A. thaliana* the SRA domain of the histone methyl transferase SUVH5 is needed for maintenance of methylation status and histone methylation (Rajakumara et al. 2016). In mammals the SRA domain of UHRF1 is required for maintenance of the cell's methylation status via its interaction with hemi-methylated DNA and the methyltransferase DNMT1 (Zhang et al. 2011). UHRF2 is present in mouse and human, the UHRF2 SRA domain is similar but non-identical to the SRA of its paralogue UHRF1, although they are the only two proteins in humans that contain this domain. UHRF2 is not required for maintenance of cell methylation status, although it too has been shown to interact with DNMT1 (Zhang et al. 2011). The SRA domain has a saddle shape structure consisting of seven β-sheets and four α-helices with the concave inner surface having a strong positive charge. The SRA of UHRF2 contains a disordered NKR loop, this flexible structure allows for recognition of dual flipped out bases, thereby allowing the UHRF2 SRA to recognise fully modified DNA. This domain has been shown to preferential bind 5-hydroxymethylcytosine and has lower binding affinity to just methylcytosine, this is likely due to

reduction in steric hindrance in its binding pocket which accommodates the bulkier hydroxyl group (Zhou et al. 2014).

The TTD and PHD domains have been identified as binding methylation markers on histones. The Tandem-Tudor Domain does not appear to have high affinity for methylated histones on its own; however, in combination with the PHD domain, this region appears to recognize and bind H2K9me2 and H3K9me3 up to 4 fold more strongly (Pichler et al. 2011).

3.1.2 Results

3.1.2.1 UHRF2 binds crosslinked DNA in *vitro*

UHRF2 was previously identified as a potential ICL binding protein by mass spectrometry (Liang et al. 2015). As UHRF2 has similar domain architecture to UHRF1, which was previously characterized as a human protein preferentially binding DNA ICLs (Liang et al. 2015); UHRF1 was used as a reference to compare with the DNA binding affinity of UHRF2. Recombinant human N-terminal tagged FLAG-HA-UHRF2 was purified to homogeneity from *Sf9* insect cells (Figure 3A lane 2). DNA probes were designed with various modifications and radiolabelled with ^{32}P (Figure 3B). To generate a site specific ICL, the DNA probe containing a unique TA site was incubated with Trioxsalen (TMP) and irradiated with UVA, followed by biochemical purification. However, TMP is able to react with non TA sites (requiring only T) at low frequencies (Espositos et al. 1988), additionally; TMP requires two photo-activation events to fully cross-link the DNA strands (Hwang et al. 1996). Due to the nature of the TMP reaction a specific probe was engineered utilizing a more complex crosslinking and purification method (Figure 3C).

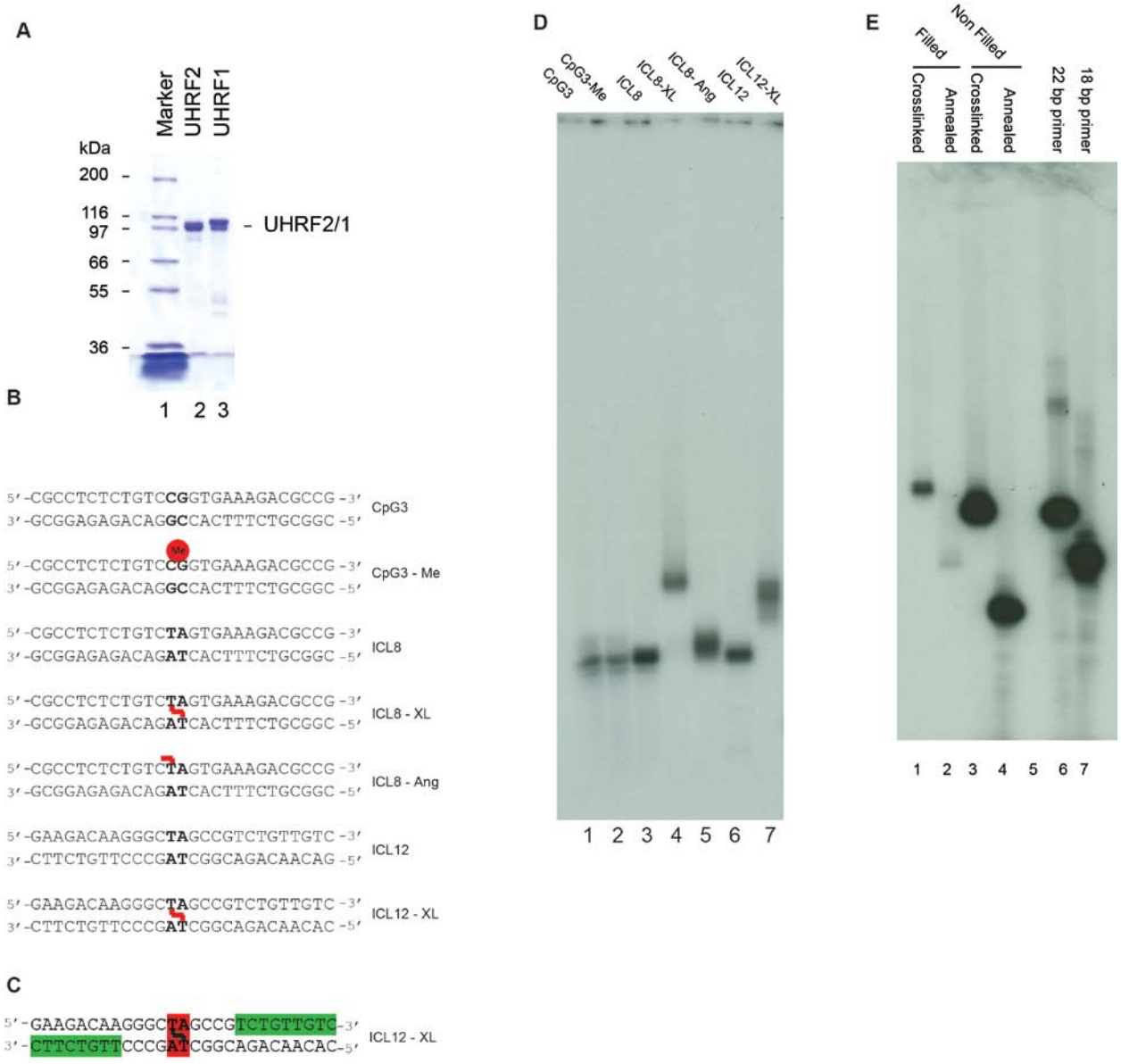


Figure 3. Reagents used in EMSA.

A) Coomassie staind gel of FLAG-HA tagged UHRF2 and UHRF1 from *Sf*9 insect cells. B) Schematic representations of DNA baits used in EMSA experiments. ME indicates methylation. Red line indicates TMP linkage (either cross linked or mono adducted). C). Representation of processing used in ICL12. Green sequence was introduced post crosslinking to generate full length probe. D) 8M Urea PAGE depicting probes used in EMSA experiments. Probes indicated as crosslinked are pure with no ssDNA present. E) 8M urea PAGE of intermediates and final generation of probe ICL12-XL.

Specifically, the core of the probe (10 annealed, 27 total mer) was subjected to TMP crosslinking, as the core contains only one T, that is where the ICL would form (Figure 3E lanes 3 and 4). The now crosslinked core was treated with Klenow fragment to fill in the ssDNA arms (3' to 5' -exo) to produce a dsDNA 27 mer probe with a unique crosslink (Figure 3E lanes 1 and 2). Probe ICL12-XL was therefore verified to be a defined ICL (Zhang et al. 2015). To ensure the purity of each final probe, they were radiolabeled and visualised on an 8M urea denaturing gel. As expected, each probe produced a single band and the crosslinked probes appeared to run slower than their non crosslinked counterparts (Figure 3D).

Using an electrophoretic mobility shift assay (EMSA) the binding affinities of purified UHRF1 and UHRF2 (Figure 4) were tested against ICL8 and ICL8-XL DNA probes (Figure 3B). Both proteins formed a weak complex with the non-ICL8 (lanes 1 and 3), and as expected UHRF1 formed a noticeably stronger complex with ICL8-XL than with ICL8 (lanes 3 and 4). UHRF2 appeared to also form a stronger complex with ICL8-XL than ICL8 (lanes 1 and 2). The complex could be super shifted with the addition of anti HA antibody, as both proteins are HA tagged, indicating it is a specific interaction (lanes 5 and 6). These data suggests that UHRF2 has a higher binding affinity to ICLs *in vitro*.

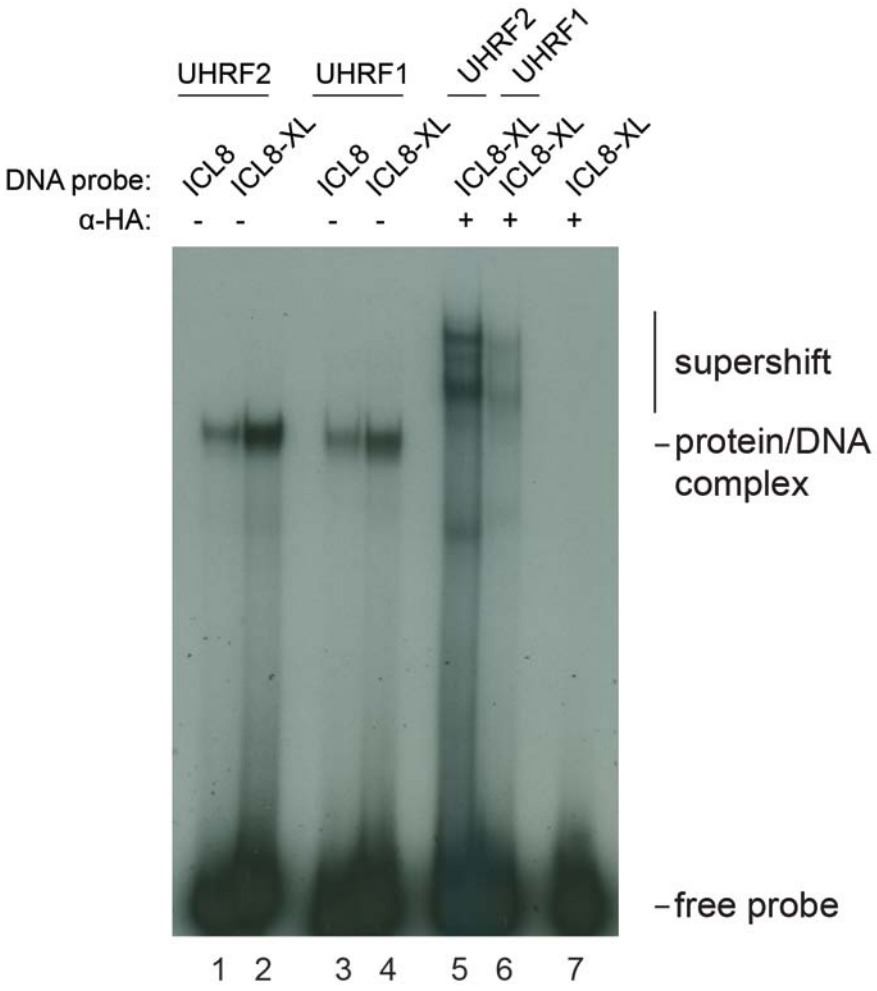


Figure 4. UHRF2 and UHRF1 DNA binding.

EMSA indicating DNA binding affinities of UHRF1 and UHRF2 using non-crosslinked and crosslinked DNA probes. UHRF1 and UHRF2 are FLAG-HA tagged.

To further investigate the DNA binding capability of UHRF2, an EMSA was carried out with additional types DNA probes (Figure 5). It was again observed that the binding capability of UHRF2 towards the crosslinked (ICL8) probe was higher than that of the non crosslinked (ICL8-XL) (lanes 1 and 2). Additionally, the comparative binding of the pure single crosslinked probe (ICL12-XL) was even higher compared to the unmodified form (lanes 5 and 6). Lastly, in good agreement with literature, it does not appear that UHRF2 has differential binding preferences to unmodified (CpG) or hemi methylated DNA (CpG3-Me) (lanes 3 and 4) (Zhou et al. 2014; Pichler et al. 2011). It is apparent that UHRF2 has a preferential binding affinity for crosslinked DNA.

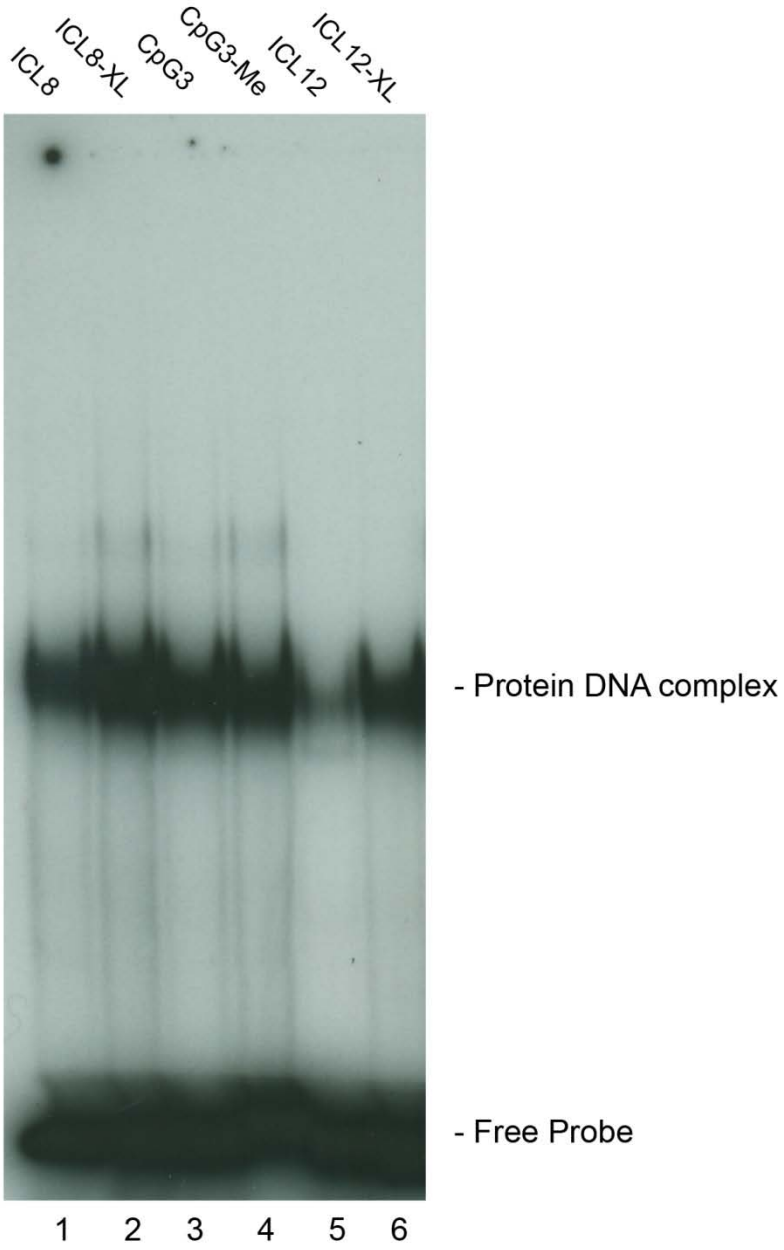


Figure 5. EMSA of UHRF2 DNA binding to various substrates.

EMSA depicting the binding preferences of UHRF2 to various types of DNA probe described in Figure 3A,B.

3.1.2.2 UHRF2 is Recruited to ICLs *in vivo*

UHRF2 was shown to preferentially bind ICLs *in vitro*, which may be indicative of its function *in vivo*. To test whether UHRF2 will bind ICLs in living cells, UHRF2 was C-terminally tagged with EGFP and stably expressed in HeLa cells. Using a micro-irradiation protocol established in our research group the cells expressing EGFP tagged UHRF2 were subjected to a precise micro irradiation stripe by a 405 nm laser in the presence and absence of TMP (Figure 7). 405 nm is a sufficient wavelength to photoactivate the now DNA intercalated TMP, causing ICLs specifically at sites where the laser was passed. Cells that were pretreated with TMP showed rapid accumulation of EGFP-tagged UHRF2 at the irradiation site (Figure 7 bottom panels). A stripe became visible 30 seconds after irradiation and accumulation leveled off at approximately 1.4 times intensity compared to the cell intensity within 3 minutes. Cells that were not pre-treated with TMP (Figure 7 top panels) showed no accumulation of EGFP UHRF2 for the entire duration of the acquisition, indicating that the recruitment visible in the TMP treated cells was due to formation of ICLs in response to the laser rather than the laser itself. This suggests that UHRF2 is rapidly recruited to ICL sites in living cells.

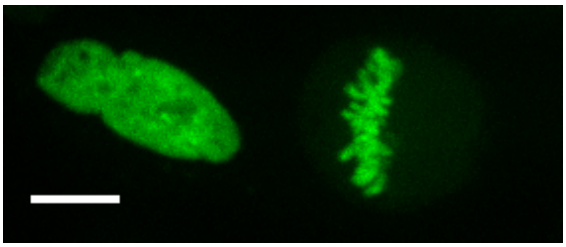


Figure 6. UHRF2 is chromatin bound.

EGFP-UHRF2 was expressed in HeLa cells. The cell on the left is in G1/S/G2 phase with the nucleus glowing green, as UHRF2 is nuclear only. The nucleosome is visible as a region with less green signal towards the lower right. The cell on the right has entered M phase with UHRF2 being bound to the condensed chromosomes. Scale bar 10 µm.

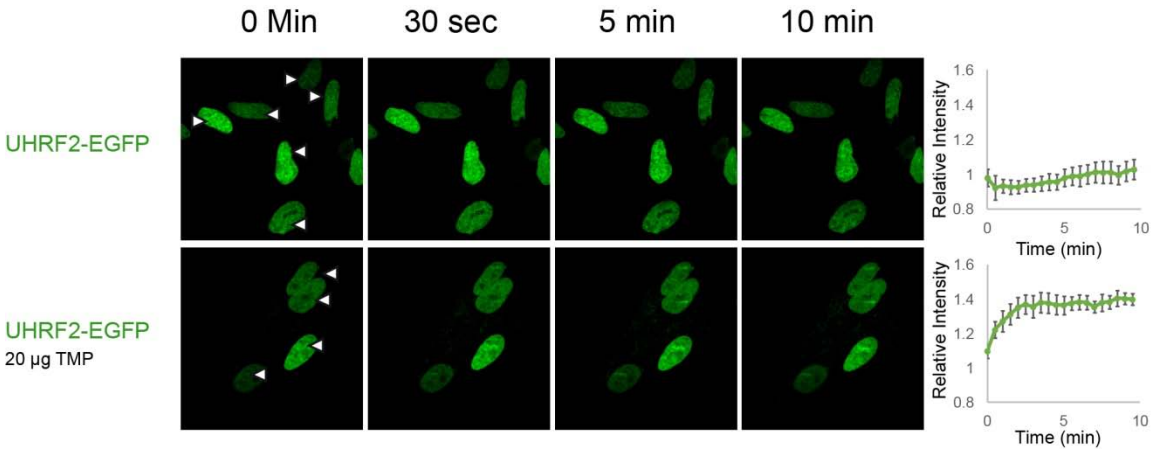


Figure 7. UHRF2 EGFP Recruitment to ICLs.

Microscopy time course showing the recruitment of UHRF2 EGFP to sites of ICL damaged induced by irradiation of the cells, with 405 nm laser at the indicated sites in a stripe (white arrow). The relative stripe to cell intensity is plotted on the right. n=4 Scale bar 10 µm.

3.1.3 Discussion

UHRF2 interacts with crosslinked DNA. It was shown that UHRF2 can specifically recognize and bind TMP induced ICLs *in vitro*. This binding is likely occurring through the SRA domain, that has a 75% sequence identity to UHRF1 which is known to be the DNA binding domain in UHRF1 (Liang et al. 2015). It has been previously shown that the UHRF1 and UHRF2 SRA domain, while being very similar in sequence, possessed different binding recognition. The SRA domain of UHRF1 has preferential binding affinity to hemi-methylated DNA while the SRA of UHRF2 was shown to prefer fully hydroxy-methylated DNA and shows no preference for non-methylated or hemi-methylated DNA (Zhou et al. 2014), which is in line with our results. The DNA methylation status preference of our proteins may act as an indicator that our proteins are potentially active and correctly folded, giving some credence to the binding affinities observed with the ICLs in the EMSA experiments. The affinity of just the UHRF2 protein towards naked ICLs makes it a likely candidate for recognition of this damage.

It was also shown that UHRF2 is rapidly recruited to sites of TMP induced ICLs *in vivo*. An important challenge in studying ICL repair is observing the repair directly and in real time. The irradiation protocol established in our research group provides an important mechanistic insight into the repair process as it allows us to study the same cell pre and post damage live. After the initial irradiation the local tagged UHRF2 proteins were bleached, however within seconds new non bleached proteins diffused to the location, after this diffusion an accumulation was observed above background indicating the recruitment of more UHRF2. UHRF2 is normally chromatin bound at all stages of the cell cycle. As the initial bleach of the local UHRF2 protein is briefly visible (not shown), it is likely that UHRF2 exists in a bound and unbound equilibrium allowing it to diffuse about the nucleus only remaining strongly bound when it encounters an ICL.

These data suggest that UHRF2 may be scanning along the DNA and when an ICL is encountered the binding affinity increases, immobilizing the protein at the ICL site. This is likely accomplished through the SRA domain as seen in EMSA but may be further assisted by the TTD and PHD domains.

3.2 UHRF2 is Involved in the DNA Damage Response

3.2.1 Introduction

When searching for a function for a protein it is logical to begin by examining the known/recognizable domains within it. UHRF2 consist of 5 known domains with varying functionality. It is easy to group them into two main categories: ubiquitin related and methylation related.

The UBL domain bears a resemblance to ubiquitin in its folded form, however the sequence similarity is not high (60% positives). At present the function of this domain is unknown. The RING domain is an E3 ligase with several function attributed to it. It has been shown play a role in histone acetylation through interaction with TIP60 (Zeng et al. 2016). In addition, UHRF2 RING domain has been shown to ubiquitinate PCNP targeting it for degradation (Mori et al. 2002) as well as ubiquitinating polyglutamine aggregates targeting them for degradation too (Iwata et al. 2009). The E3 activity of the RING domain is also tied to cell cycle regulation though its ubiquitination of G_1 cyclins (Mori et al. 2011) Lastly, the UHRF2 RING domain is able to auto-ubiquitinate (Mori et al. 2004).

The methylation related domains of UHRF2 are TTD, PHD, and SRA. The Tandem-Tudor Domain of UHRF2 binds H3K9me2/3 histone markers, this binding is enhanced when the PHD domain is also present, this enhancement was attributed to the conserved linker region between the two domains (Pichler et al. 2011). The PHD domain of UHRF2 has been shown to interact and affect levels of H3K9ac (Zhang et al. 2016). These two domains work together to

facilitate the interaction of UHRF2 with histone markers. The SRA domain is a specific DNA binding domain, interacting with DNA methylation marks. The UHRF2 SRA domain binds more strongly to 5' hydroxymethylated DNA as well as fully methylated DNA (Zhou et al. 2014). It has also been shown to interact with DNMT1 in S-phase, however UHRF2 does not contribute to methylation maintenance like UHRF1 (Zhang et al. 2011).

The structure of UHRF2 is rather unique with several domains with different functions. Its domain architecture is identical UHRF1, but they do appear to serve different roles in the cell. As UHRF1 is indispensable for methylation maintenance in the cell through its interaction with DNMT1, it can only be knocked out in ES cells. Ectopic expression of UHRF2 does not rescue cell methylation in its absence (Pichler et al. 2011; Zhang et al. 2011). The two proteins also display differential expression patterns with UHRF2 up regulating during differentiation and remaining high in terminally differentiated cells and UHRF1 being more uniformly expressed (Mori et al. 2012). As UHRF2 interacts with cell cycle regulators, histones, and DNA it is likely the protein acts as a messenger hub between different cell networks.

3.2.2 Results

3.2.2.1 Generation of CRISPR/Cas9 Knockout of UHRF2

To assess the importance of UHRF2 in ICL repair it was important to observe the cell response to ICL agents in the absence of UHRF2. To that end the endogenous *UHRF2* was deleted using the new method of CRISPR/Cas9 directed knockout. To achieve this a guide RNA was designed that would match and anneal to the start of the second exon of the *UHRF2* coding sequence. Once cut by the Cas9 the resulting protein was truncated just a few amino acids into the TTD domain (Figure 8A). To ensure all possible copies of *UHRF2* were damaged, the genomic DNA of the clone was extracted. Primers specific to sites up and down stream of the Cas9 target site were used to amplify a genomic fragment, which was subsequently sub-cloned

into pBluescript (pBS) and transformed into *E.coli*. The individual plasmids were isolated sequenced, individual sequences were aligned to the wild type genomic sequence. From the alignments four distinct types of variations or "indel" (deletions of genomic DNA) were identified (Figure 8B), no wild type sequence was detected leading to the conclusion that there are four copies of the *UHRF2* gene in HeLa cells and all four were destroyed by CRISPR. Lastly, as the UHRF2 antibody was dirty, the western blot appearance of the knock out was established. To that end, varying concentrations of HeLa lysate were loaded against 10 µg (standard concentration used in western blots) of UHRF2 knockout HeLa lysate (Figure 8C). There is a non-specific band that appears just above the UHRF2 band. The lower band is absent from the UHRF2 knockout cell line leaving only this non-specific band visible. More recently a superior antibody became available for UHRF2, this new antibody depicts the loss of UHRF2 in a far more unambiguous way (Figure 8D).

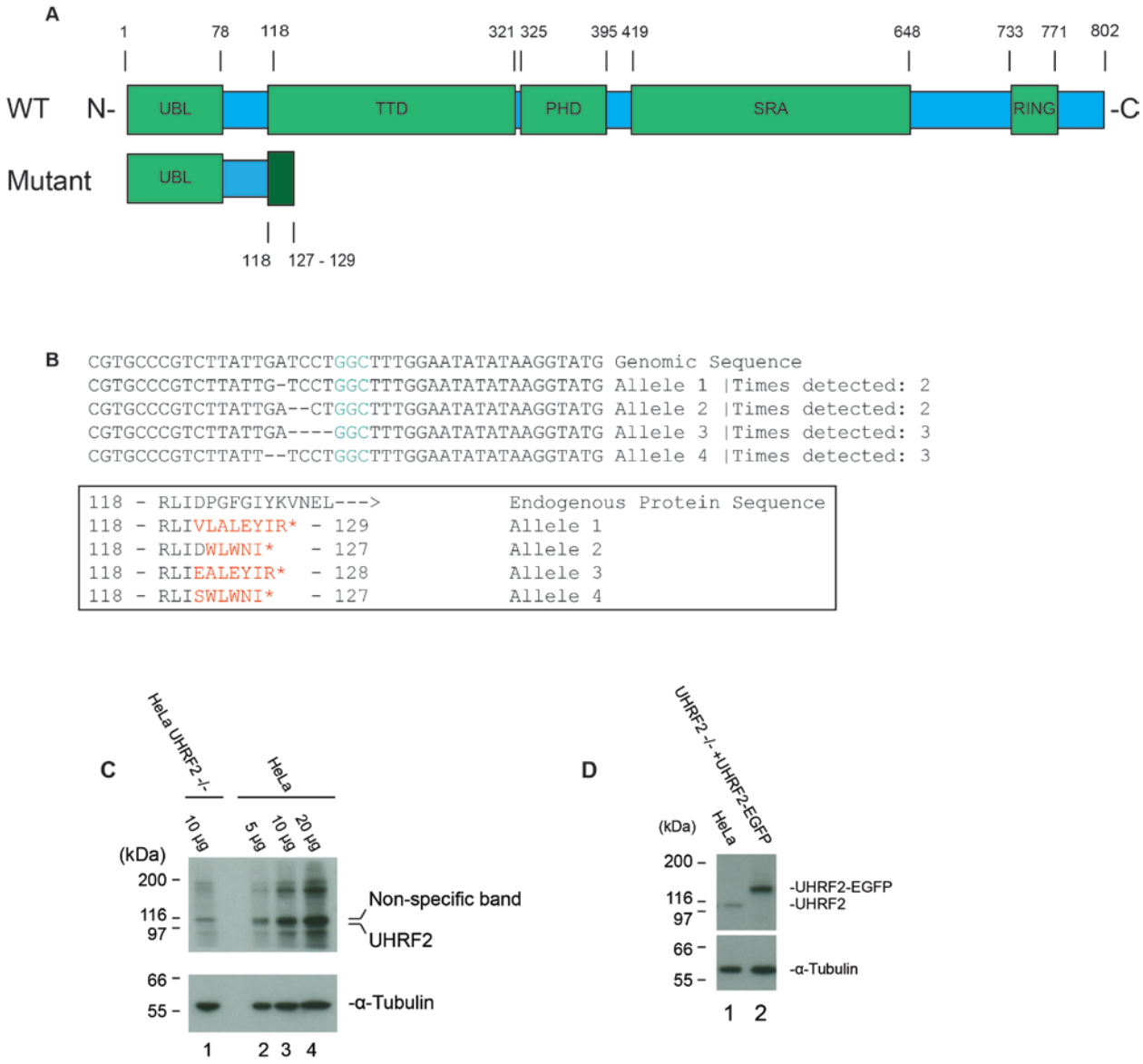


Figure 8. Characterization of UHRF2 CRISPR/Cas9 Knockout in HeLa.

A) Schematic representation of the UHRF2 protein and the truncated version of the protein that results from CRISPR/Cas9 mediated disruption of the gene. The numbers represent amino acids. B) Sequences derived from genomic DNA of the knockout clone (Top set). Primers flanking the Cas9 target region were used. Mutant sequences were aligned to the wild type sequence to demonstrate where NHEJ repair caused a deletion. Marked in green is the PAM sequence. The sequences in the box are the amino-acid sequences (based on DNA above) that result from the improper repair of the Cas9 damaged DNA. The subsequent frameshift mutations and resulting truncations are marked in red. C) Western blot of HeLa vs. UHRF2 knockout HeLa depicting the non-specific band and the disappearance of the specific band in the knockout (anti-UHRF2 T-18) D) Western blot of HeLa vs. UHRF2 knockout HeLa complemented with EGFP-tagged UHRF2 (anti-UHRF2 C-10).

3.2.2.2 UHRF2 is needed for response to ICL inducing agents

Using the UHRF2 knockout cell line it was now possible to access the cell lines viability when challenged with ICL inducing agents. Clonogenic survival assays were plated in triplicate using both non-modified HeLa cells and UHRF2 knockout cells. The plates were then treated with MMC (Figure 9A) or TMP and UVA (Figure 9B) and allowed to grow for approximately 2 weeks. The cells were then fixed and stained with crystal violet and the number of colonies were counted. With both agents there was a significant reduction in the colony count of the UHRF2-/- cells, suggesting the loss of UHRF2 rendered the cells more sensitive to ICLs, likely due to defects in ICL repair. As the UHRF2-/- cell line is clonal, another survival assay was carried out using shRNA mediated UHRF2 knockdowns compared to the knockout and non-modified HeLa (Figure 9C). The knockdown was achieved with an shRNA targeting a region between the PHD and SRA domains in the *UHRF2* mRNA, a target that is vastly different to the CRISPR/Cas9 target region which is at the start of the TTD domain. These two separate targets (difference sequences) with two separate mechanisms of action reduces the likelihood of the observed phenotype being due to clonal effect or off target effects, however the shRNA cells are not 100% null of UHRF2. The resulting graph shows that the shRNA knockdown of UHRF2 also renders the cells sensitive to MMC but not to the same extend as UHRF2-/-.

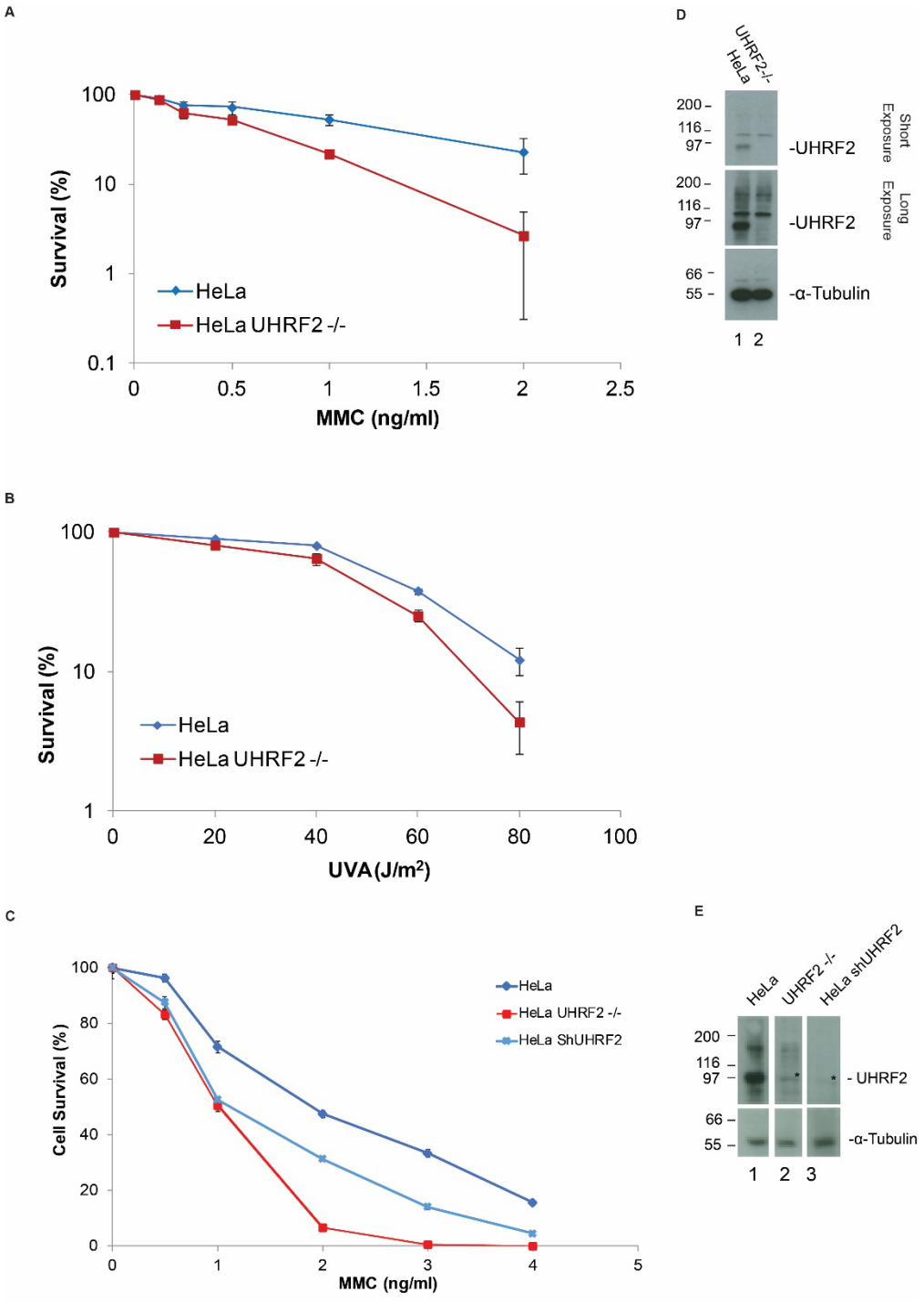


Figure 9. Depletion of UHRF2 reduces cell survival in response to ICL inducing agents.

Clonogenic survival assays of HeLa UHRF2 -/- CRISPR/Cas9 mediated knockout cells compared to HeLa. A) The cells were treated with varying concentrations of MMC. B) The cells were treated with 50 ng/ml TMP and irradiated at various intensities of UVA. C) UHRF2 was stably knocked down in HeLa cells using shRNA and treated as in A. Error represented as SEM. All experiments carried out in technical triplicate. A is an average of 3 repeats, B is an average of 2 repeats. D) Western blot of cell lines used in A and B. E) Western blot of cells used in C. Figure also shown in full in Figure 23A (lanes 1,4, and 7). * non specific bamd.

3.2.2.3 Characterization of the UHRF2 Domains

As it was shown that UHRF2 is important for the cells response to ICLs, the next question is which part of UHRF2 is most important for carrying out the repair function, especially considering the multi domain architecture of the protein. A series of constructs were created where each domain of UHRF2 was sequentially deleted (Figure 10). The constructs were EGFP-tagged and stably expressed in UHRF2-/- cells (Figure 11B).

The various deletion cell lines were pre-treated with 20 μg TMP for 30 minutes and allowed to equilibrate in a temperature and CO_2 controlled chamber in the microscope. The cells were then subjected to microirradiation at the sites indicated by the white arrows (Figure 11A). The wild type EGFP-tagged UHRF2 was recruited to the ICL site within 30 seconds. The UHRF2-ΔUBL, UHRF2-ΔTTD, UHRF2-ΔPHD, and UHRF2-ΔRING mutants also recruited at the same rate as the wild type. However, the UHRF2-ΔTTD, UHRF2-ΔPHD mutants did not reach the same intensity as the wild type. The UHRF2-ΔSRA mutant was not recruited.

To further assess the impact of each domain on UHRF2 function in ICL repair the complemented UHRF2 knockout cells were subjected to survival assays. When examining the survival of the domain deletions in response to MMC (Figure 12) it was found that the ΔSRA and ΔTTD expressing cells displayed the lowest survival much like the knockout. Conversely the WT, ΔUBL, ΔPHD, and ΔRING all showed an intermediate survival. Full complementation back to HeLa was not achieved.

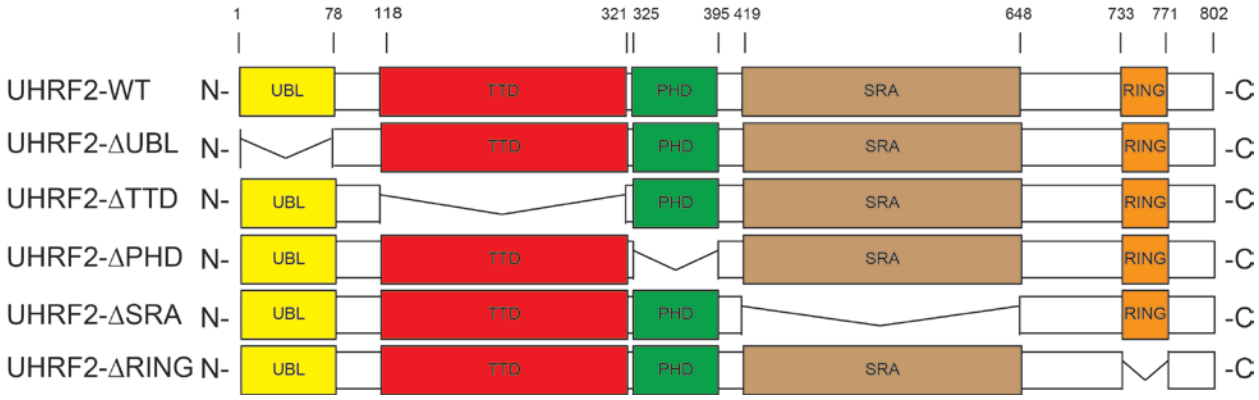


Figure 10. Diagram of UHRF2 domain deletions.

Schematic representation of UHRF2 domain deletion constructs. All constructs were also EGFP tagged at the C-terminus.

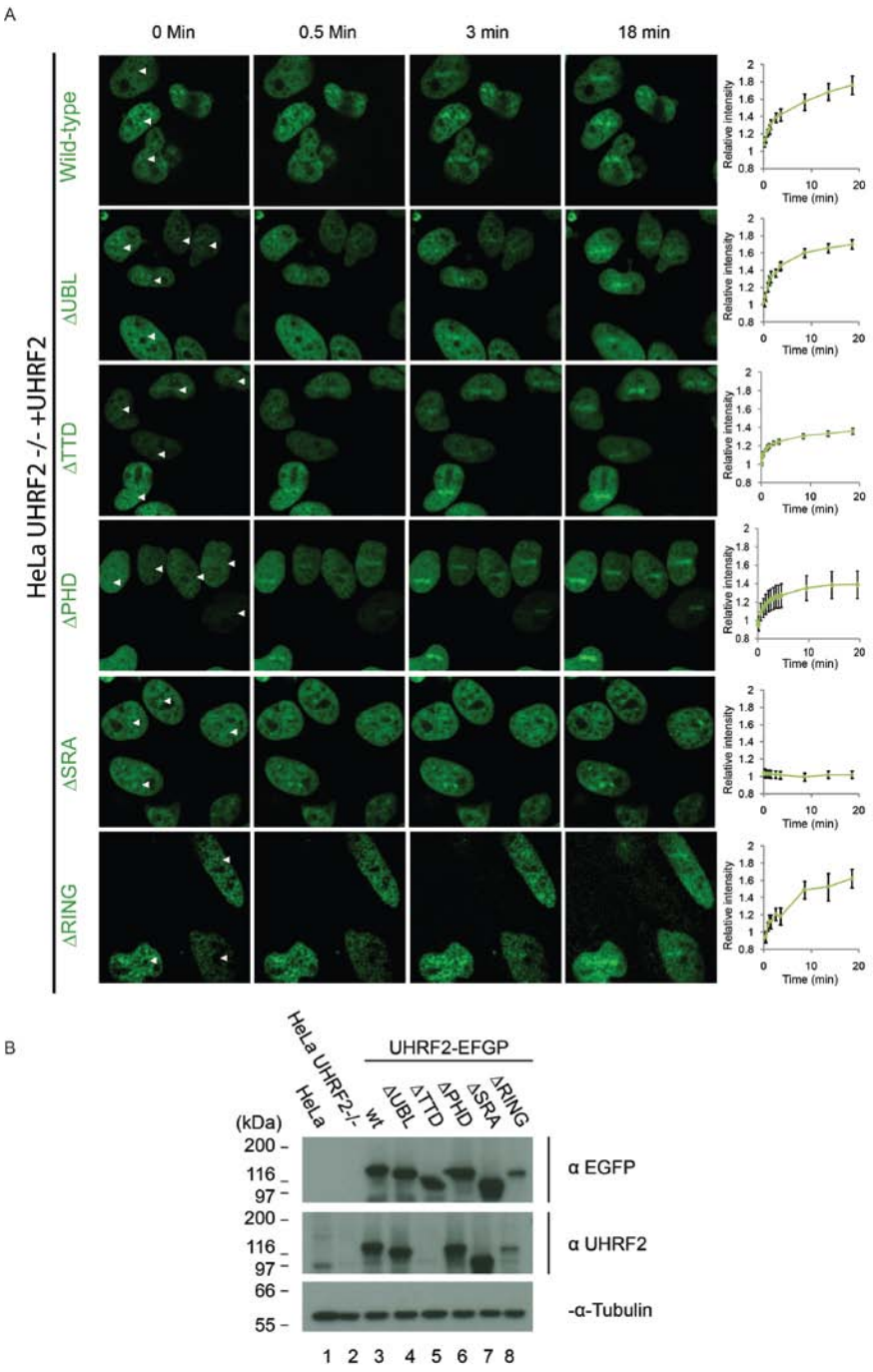


Figure 11. Recruitment of UHRF2 domain deletions.

A) Live-cell imaging of HeLa UHRF2 -/- cells (CRIPR/Cas9-mediated knockout) complemented with EGFP-tagged UHRF2 containing deletions as indicated in (Figure 10). Cells were pre-treated with TMP and microirradiated at the sites indicated with white arrows. Scale bar indicates 10 μm. Charts indicate quantification of relative intensity of signal at the irradiated sites. Error bars show SEM n = 8/treatment. B) Expression of EGFP-tagged UHRF2 and derivatives in UHRF2-/- HeLa cells. UHRF2 -/- HeLa cells were stably transfected with EGFP-tagged wild-type UHRF2 and the various UHRF2 domain deletion mutants as indicated. The anti-UHRF2 antibody recognizes an epitope in the TTD region, as such deletion of this domain causes the disappearance of the protein from the anti-UHRF2 blot.

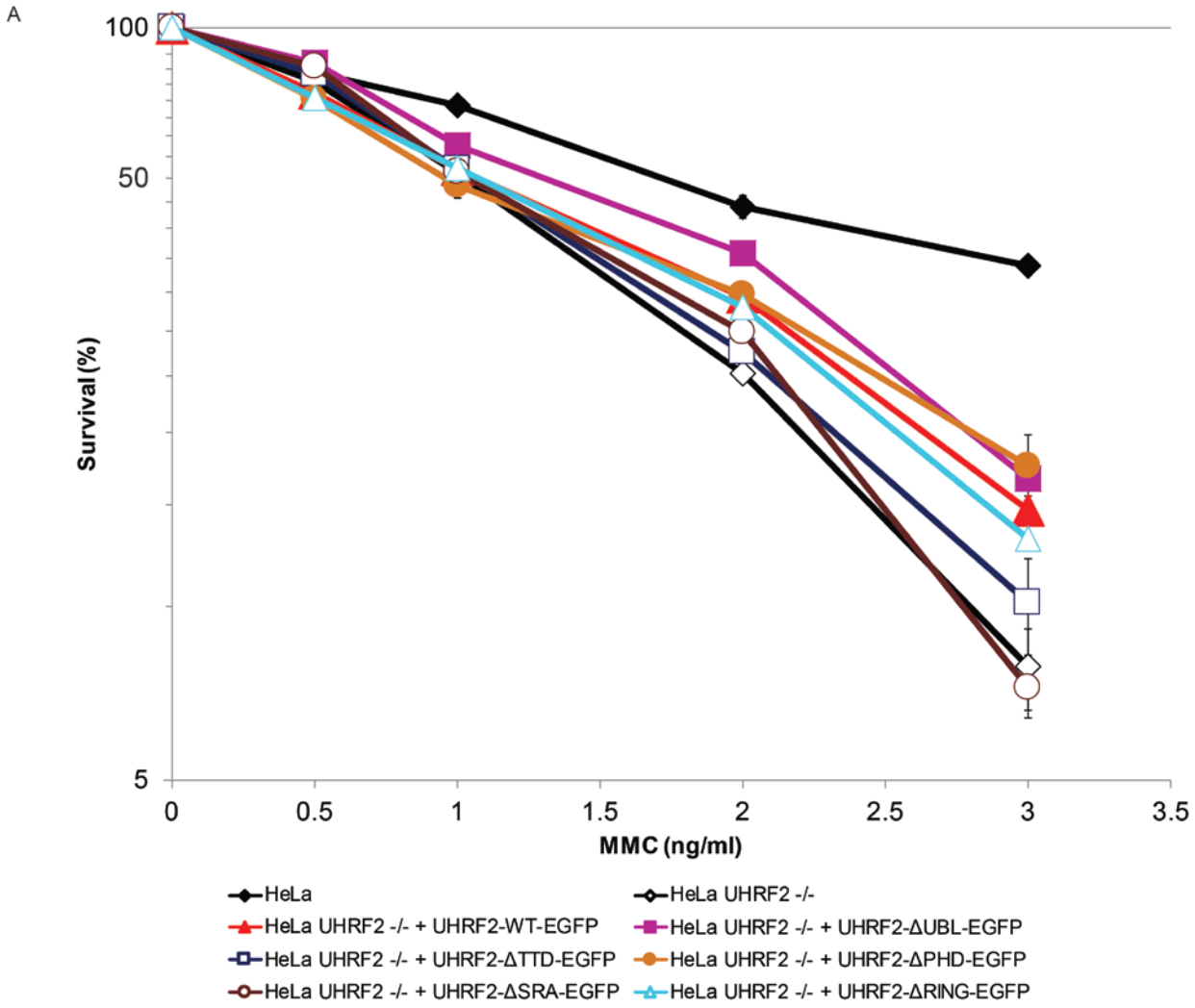


Figure 12. Survival assays of UHRF2 deletion mutants.

A) Clonogenic survival assay of HeLa and HeLa UHRF2-/- cells complemented with UHRF2 domain deletion mutants in (Figure 10) in response to MMC. Depletion of UHRF2 reduces cell survival. $p<0.05$ between HeLa and UHRF2 -/-. The error is represented as SEM. This is an average between 3 experiments.

3.2.2.4 UHRF2 is recruited to ICLs independent of DNA methylation

As UHRF2 and UHRF1 can recognize methylation marks in DNA, it was important to test whether the methylation status of the DNA is important for recruitment to ICLs. To that end, a cytosine analogue 5-aza-2'-deoxycytidine (Aza) was used to reduce global methylation levels in HeLa cells. The cells were treated with 5 μM Aza for 7 days, replenishing the media every day. On the 7th day the cells were subjected to microirradiation in the presence of TMP and the recruitments of UHRF1 and UHRF2 were quantified and compared to just TMP treated cells (Figure 13A). The stripe intensity was measured for each condition and plotted on the right of the time course. It was found that the recruitment of UHRF1 and UHRF2 was unaffected by depletion of DNA methylation. To ensure the Aza treated cells were truly methylation deficient, genomic DNA was harvest from the same cells used for microscopy. The genomic DNA was broken down into nucleosides and subjected to HPLC analysis. The peaks pertaining to cytosine and methylcytosine were quantified and plotted (Figure 13B). The untreated cells had a relative percent of 3.5% methycytosine (mdC) to cytosine (dC), while the Aza treated cells had only 0.5% mdC to dC, a 7 fold decrease. This result demonstrates the Aza treatment worked, suggesting that indeed UHRF1 and UHRF2 recruitment was not affected by the methylation status of the cells.

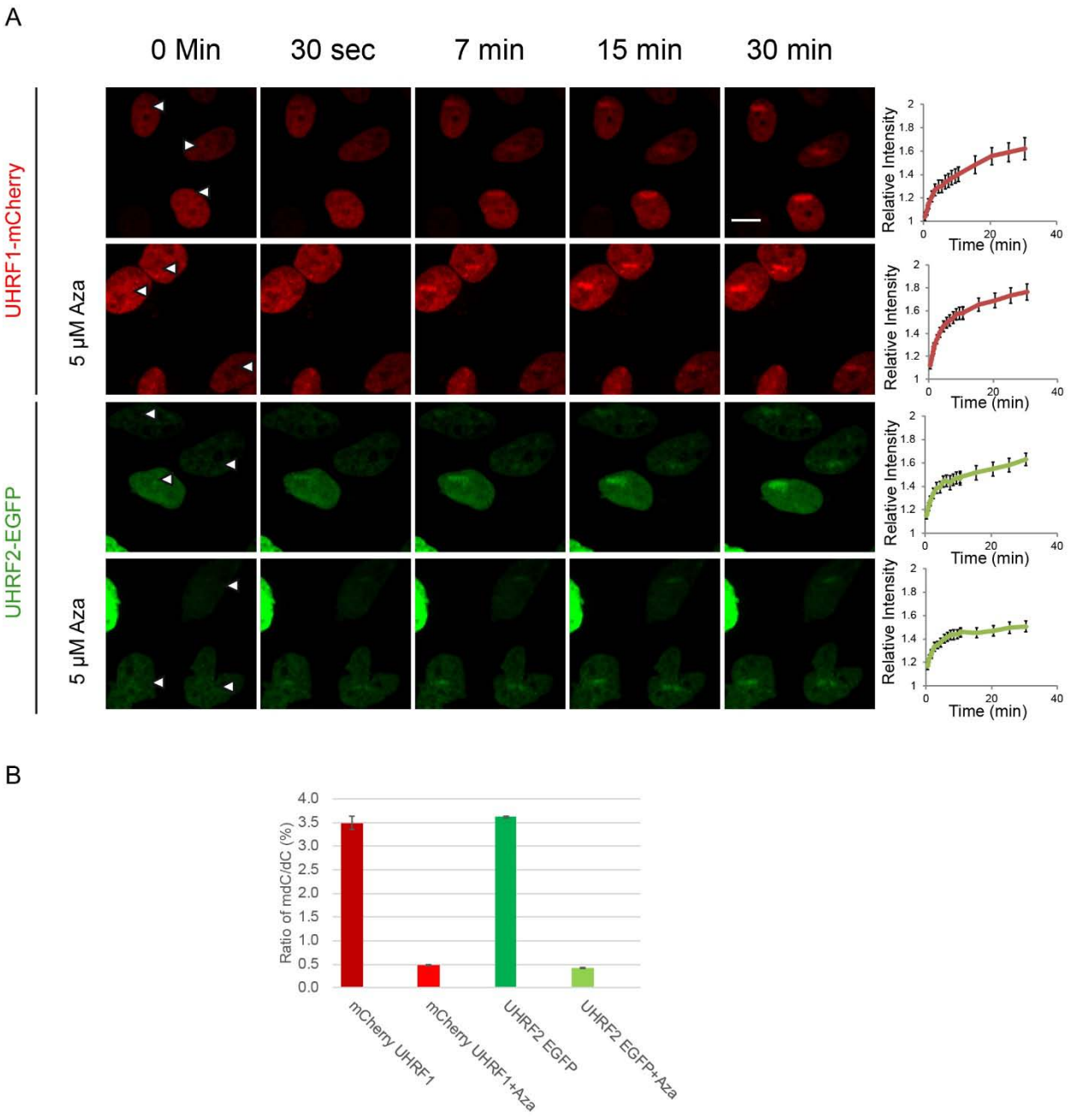


Figure 13. Recruitment of UHRF1 and UHRF2 in the absence of methylation.

A) Live cell imaging of HeLa cells expressing mCherry-tagged UHRF1, and UHRF2 -/- cells complemented with EGFP-UHRF2 treated with 5 µM Aza for 7 days or untreated cells. The cells were pre-treated with 20 µg TMP and subjected to microirradiation at indicated sites (white arrows). Stripe quantification represented as graphs. n = 8 error is shown as SEM, scale 10 µm. B) Quantification of methylated cytosine nucleosides vs. cytosine nucleosides in the genomic DNA of the cells used in A. Work done in collaboration with Paolo Spingardi and Skirmantas Kriaucionis.

3.2.3 Discussion

UHRF2 is important in the ICL repair process in the cell. The initial generation of the UHRF2 CRISPR knockout cell line was difficult with low frequency of full knockout. Most frequently at least one intact allele would be found in the clones screened. This could either be due to the poor guide RNA target region, or perhaps the detrimental effects on the cell if all copies of UHRF2 are lost. When the knockout cells were subjected to ICL inducing agents the cells displayed a marked reduction in survival. To ensure this was not due simply to a defect in this particular clonal cell line, the same survival assay was carried out with a pool of cells where UHRF2 was stably knocked down using shRNA. Those cells, while not having the same extent of survival impairment as the knockout, likely due to residual UHRF2, were also sensitized. This rules against the phenotype observed being simply due to clonal effects or even off target effects, as the CRISPR guide RNA and the shRNA targets are in separate locations within the gene.

It was also shown that specific domains within UHRF2 are of more importance to the ICL repair process than others. When examining the recruitment of UHRF2 to ICLs it appeared that the loss of three domains, the TTD, PHD, and SRA domains, appeared to affect the recruitment the most. The deletion of the SRA domain abrogated the recruitment of UHRF2 to ICLs completely. This result is quite logical as the SRA domain is known as a DNA binding domain. The other two domains that displayed a defect in recruitment were the TTD and PHD domains, this too makes sense as the TTD and PHD domains have been shown to bind methylated histones. It is interesting to note that the localization in all three of these mutants was also visibly affected. The ΔSRA mutant appeared to be more strongly excluded from heterochromatin rich regions in the nucleus, while the ΔTTD and ΔPHD mutants conversely seemed to localize far less stringently with UHRF2-ΔTTD and UHRF2-ΔPHD appearing to be more uniformly present throughout the nucleus. This observation falls well in line with the

known functions of these domains. The SRA domain is chiefly responsible for binding DNA, specifically fully methylated and hydroxy methylated DNA (Zhou et al. 2014), thus the loss of this domain would exclude UHRF2 from regions marked by these modifications. The TTD PHD domains of UHRF2 has been found to bind H3K9me2/3 histones (Pichler et al. 2011) and loss of these domains would reduce the association of UHRF2 with di/tri methylated histone tails.

To further dissect the importance of each domain in the ICL response process the deletion mutants were subjected to survival assays. The UHRF2-/- cells again demonstrated a diminished survival compared to the wild type. The deletion mutants however failed to give a black or white response, falling in between HeLa and the knockout. When using MMC the wild type protein failed to complement the knockout all the way back to HeLa survival levels, this is likely due to changes that happened within the cell due to prolonged lack of UHRF2 during the knockout generation process before complementation with UHRF2 was possible. Recent development in the field have linked UHRF2 to 5' hydroxymethylation in the cell (Zhou et al. 2014), which is a possible epigenetic marker regulating gene expression (Ficz et al. 2011). As UHRF2 is frequently deregulated in cancers (Zhang et al. 2016), the loss of the protein could prompt irreversible changes in the cell. Taken as such, it is still worth comparing the deletion mutants to the wild type. The SRA deletion and the TTD deletion mutants were comparable to the knockout while the others were in line with the wild type, suggesting the SRA domain and the TTD domain are important in the response to MMC damage.

Lastly, as the UHRF2 SRA domain consistently appears to be important it was worth investigating whether the methylation status of the cell plays a role in recruitment to ICLs. When cells were reduced from 3.5 to 0.5% genomic methylation it was found that the recruitment of UHRF1 and UHRF2 to ICLs were not impeded. This reduction is in line with pervious work

done with UHRF1, where knockout of UHRF1 in ES cells saw a drop from 4% genome methylation to 1.5% (Liu et al. 2013). This is an interesting result as the SRA domain of UHRF1 is directly implicated in binding of hemi methylated DNA and the SRA domain of UHRF2 has been shown to bind fully methylated DNA as well as hydroxymethyated DNA (which is a degradation derivative of methylation). This could perhaps underscore a multi-functionality of the SRA domain, as its deletion abrogates recruitment, but when the SRA domain is present the recruitment it is not methylation dependent.

Taken together these data demonstrate that UHRF2 plays a role in ICL repair and that it is likely dependent on the SRA domain of the protein, but is functionally distinct from epigenetic regulation.

3.3 UHRF1 and UHRF2 Interact with each other *in vivo* and *in vitro*

3.3.1 Introduction

The protein UHRF1 has been much more extensively studied than UHRF2, mostly due to its importance in methylation maintenance in the cell (Bostick et al. 2007). UHRF2 on the other hand has not been studied in much depth, until recently, when it was discovered as a specific reader of hydroxymethyl cytosine (Spruijt et al. 2013). UHRF1 and UHRF2 appear to serve different functions in the cell based on this research, however; UHRF1 and UHRF2 are similar with 54% over-all sequence identity, or 68% similarity when allowing similar amino acid substitutions. Both proteins interact with DNMT1(Zhang et al. 2011), which could indicate a possible overlap of function.

The question of redundancy between these two proteins has been considered in the past with respect to methylation maintenance, where it was seen that over expression of UHRF2 does not rescue the UHRF1 depletion phenotype in ES cells (Pichler et al. 2011). Conversely, it has been reported that UHRF1 and UHRF2 may share a conserved function in regulating *de novo*

methylation through their interaction with DNMT3A (Jia et al. 2016). What has not been addressed is the possible redundancy of UHRF1 and UHRF2 with respect to DNA damage. UHRF1 has been implicated in ICL repair previously, and UHRF2 has been reported to function in DSB repair (Luo et al. 2013). Both UHRF1 and UHRF2 have a function in ICL repair, but the question if their functions are redundant or not in ICL repair has yet to be addressed.

3.3.2 Results

3.3.2.1 UHRF1 and UHRF2 cooperate in ICL repair

It was previously shown herein that UHRF2 depletion results in reduced tolerance to ICL inducing agents, such as MMC. It was also previously published that the UHRF2 paralogue UHRF1, is also needed for survival in response to MMC (Liang et al. 2015). It was therefore interesting to test if these two proteins are epistatic with each other or not, to deduce whether they are redundant. As it is not possible to knockout UHRF1 in HeLa cells, (due to lethality) shRNA was used to stably knockdown UHRF1 (<1%) in both HeLa cells and the previously characterized UHRF2-/- cell line. The cells were then used for a clonogenic survival assay and treated with MMC (Figure 14A). As expected, the UHRF1 knockdown reduced cell survival in response to MMC as did the UHRF2 knockout. The double reduction of UHRF1 and UHRF2 together further reduced cell survival indicating a relationship between the two proteins, they appear to function in two separate pathway and are not epistatic with each other.

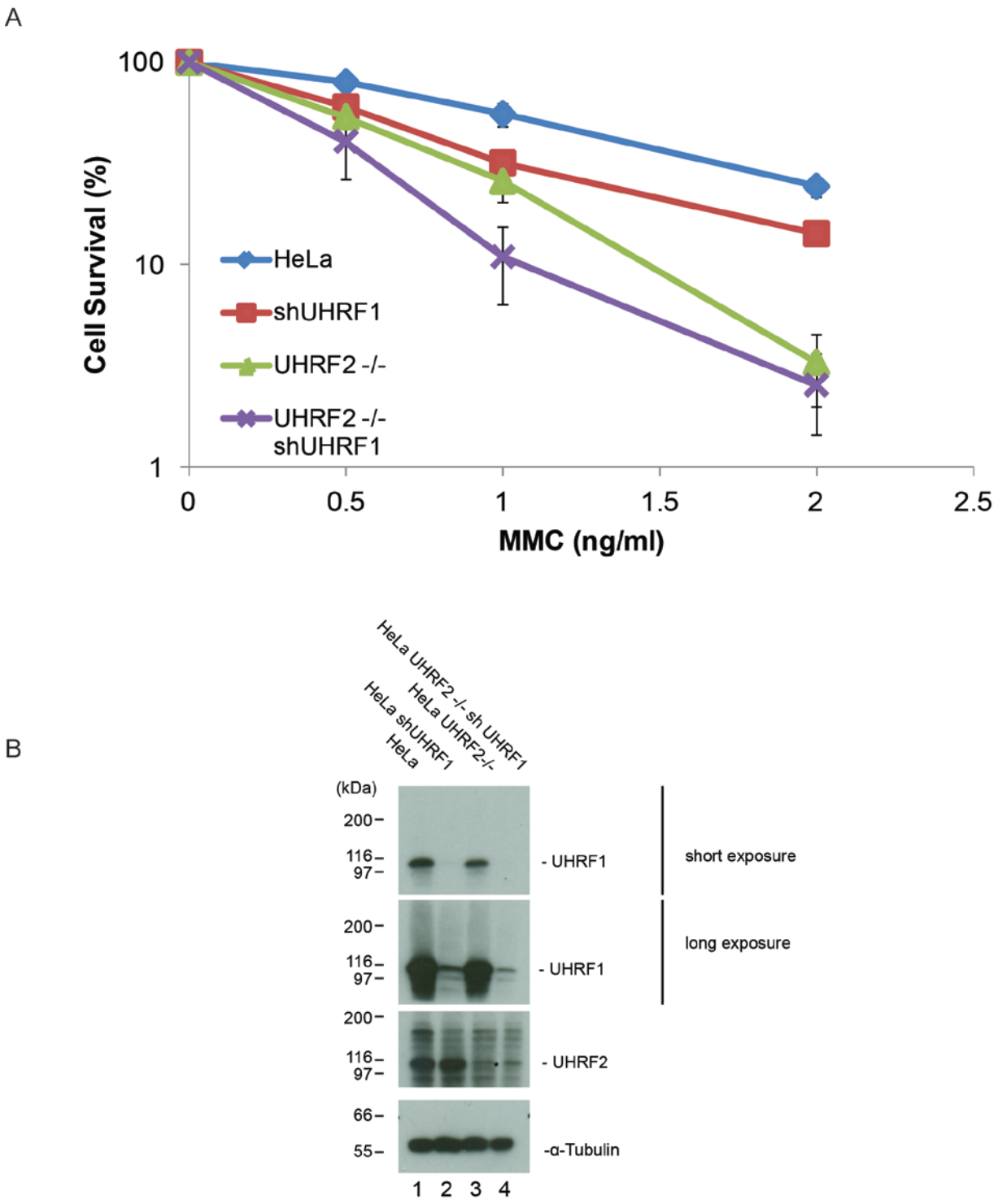


Figure 14. UHRF2 and UHRF1 are needed for survival in response to MMC.

A) Clonogenic survival assay of HeLa and HeLa UHRF2-/- cells treated with shUHRF1 shRNA in response to MMC. Depletion of UHRF2 and reduction of UHRF1 reduces cell survival. The effect is enhanced with a double reduction. The error is represented as SEM. This is an average between 2 experiments, each done in triplicate. B) Representative western blot of UHRF1 and UHRF2 expression in cells used for survival assay in A. *non-specific band on anti-UHRF2 blot.

3.3.2.2 UHRF1 and UHRF2 directly interact with each other

As UHRF2 and UHRF1 appear to both be important for survival in response to MMC (Figure 14A) it is important to address whether they can interact with each other or if they act completely independently. In order to elucidate a possible interaction two immunoprecipitation approaches were used. S3 HeLa cells expressing either Flag-HA-UHRF1 or Flag-HA-UHRF2 were harvested after treatment with TMP+UVA. The cells were then fractionated into a nuclear pellet and a cytoplasmic supernatant. The nuclear pellets were lysed and subjected to immunoprecipitation with anti-Flag antibody. The immunoprecipitated proteins were analyzed by western blotting. It was found that when UHRF2 was subjected to IP, endogenous UHRF1 was CoIPed with it; in the reverse scenario where UHRF1 was pulled down, endogenous UHRF2 was also IPed (Figure 15A), providing evidence that these two proteins are interacting, either directly or indirectly.

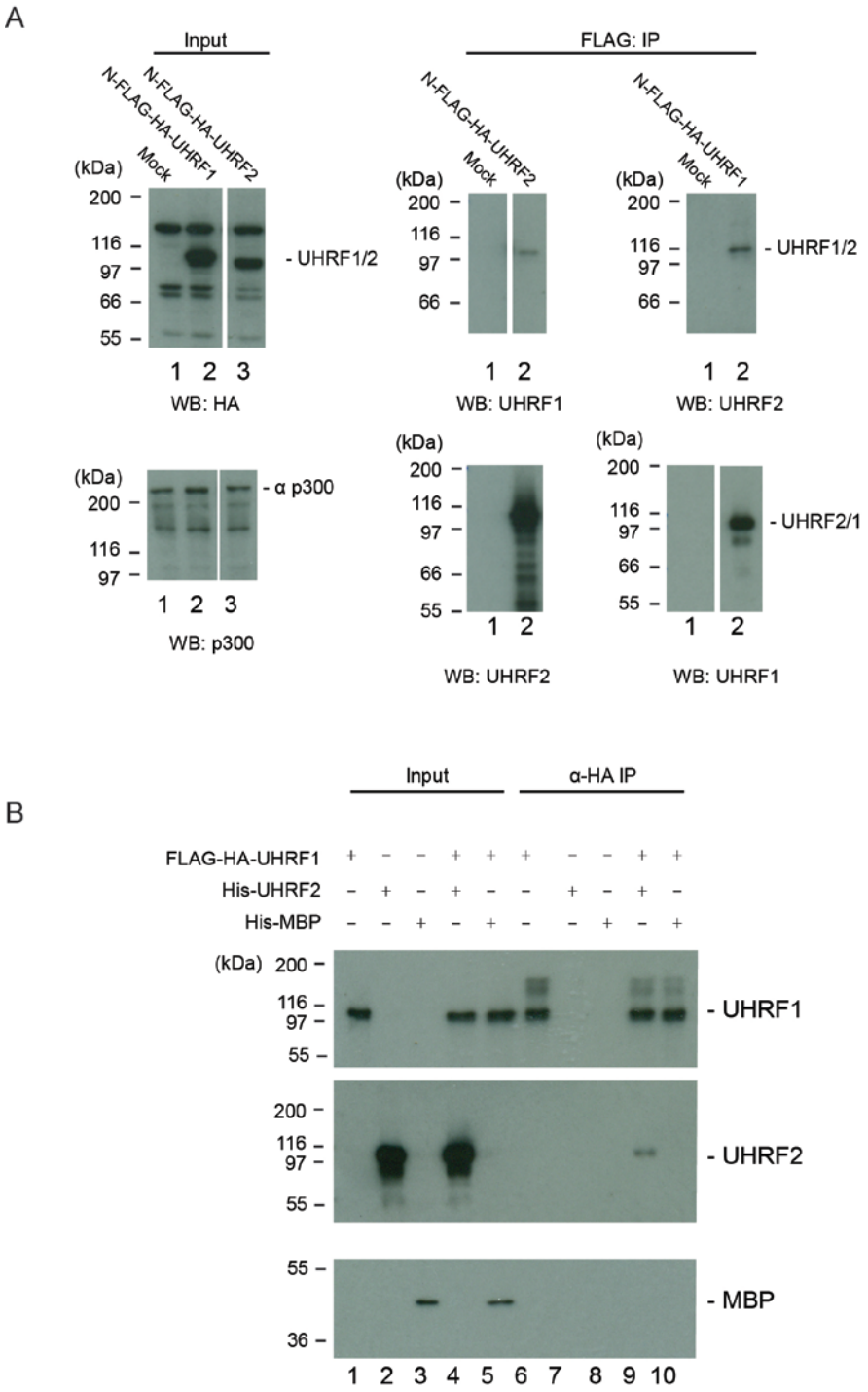


Figure 15. UHRF1 and UHRF2 interact physically.

A) Flag-HA-tagged UHRF2 or Flag-HA-tagged UHRF1 were expressed in HeLa cells and immunoprecipitated from nuclear extracts using anti-Flag antibodies. Immunoprecipitates were analyzed by immunoblotting using antibodies, as indicated. UHRF2 is co-immunoprecipitated with UHRF1 and vice versa. p300 is used as a loading control. B) Flag-HA-UHRF1 and His-UHRF2 were expressed and purified from *Sf9* insect cells, His-MBP was expressed and purified from *E.coli*. The proteins were mixed as indicated and immunoprecipitated with anti-HA antibody. UHRF2 directly interacts with UHRF1. MBP was used as a negative control.

To define whether the interaction observed was direct protein-protein interaction or whether it was mediated via another protein, Flag-HA-UHRF1 and His-UHRF2 were expressed in *Sf9* cells and purified. The proteins were mixed and incubated to allow an interaction to form. The mixtures were then immunoprecipitated with anti-HA antibody and tested by western blotting (Figure 15B). It was found that UHRF1 and UHRF2 did co-immunoprecipitate, proving a direct protein-protein interaction between them. MBP was used as a negative binding control.

Lastly, to further characterize the interaction between UHRF1 and UHRF2 a series of domain deletions of UHRF1 were used. Flag-UHRF1 and Flag-UHRF1 domain deletion variants as well as His-UHRF2 were expressed and purified from *Sf9* insect cells. UHRF2 was mixed with each variant of UHRF1 in turn and all mixtures were allowed to form an interaction. The protein mixtures were then subjected to His IP using Ni-NTA agarose. The resulting eluates were then evaluated using western blotting. The wild-type UHRF2 and UHRF1 were found to once again interact. The deletion of UBL,TTD,PHD, and the RING domains did not interfere with the ability of UHRF1 to interact with UHRF2. However, the deletion of the SRA domain of UHRF1 abrogated the interaction of UHRF1 with UHRF2, limiting the interacting region between these two proteins to the confined of the UHRF1 SRA domain.

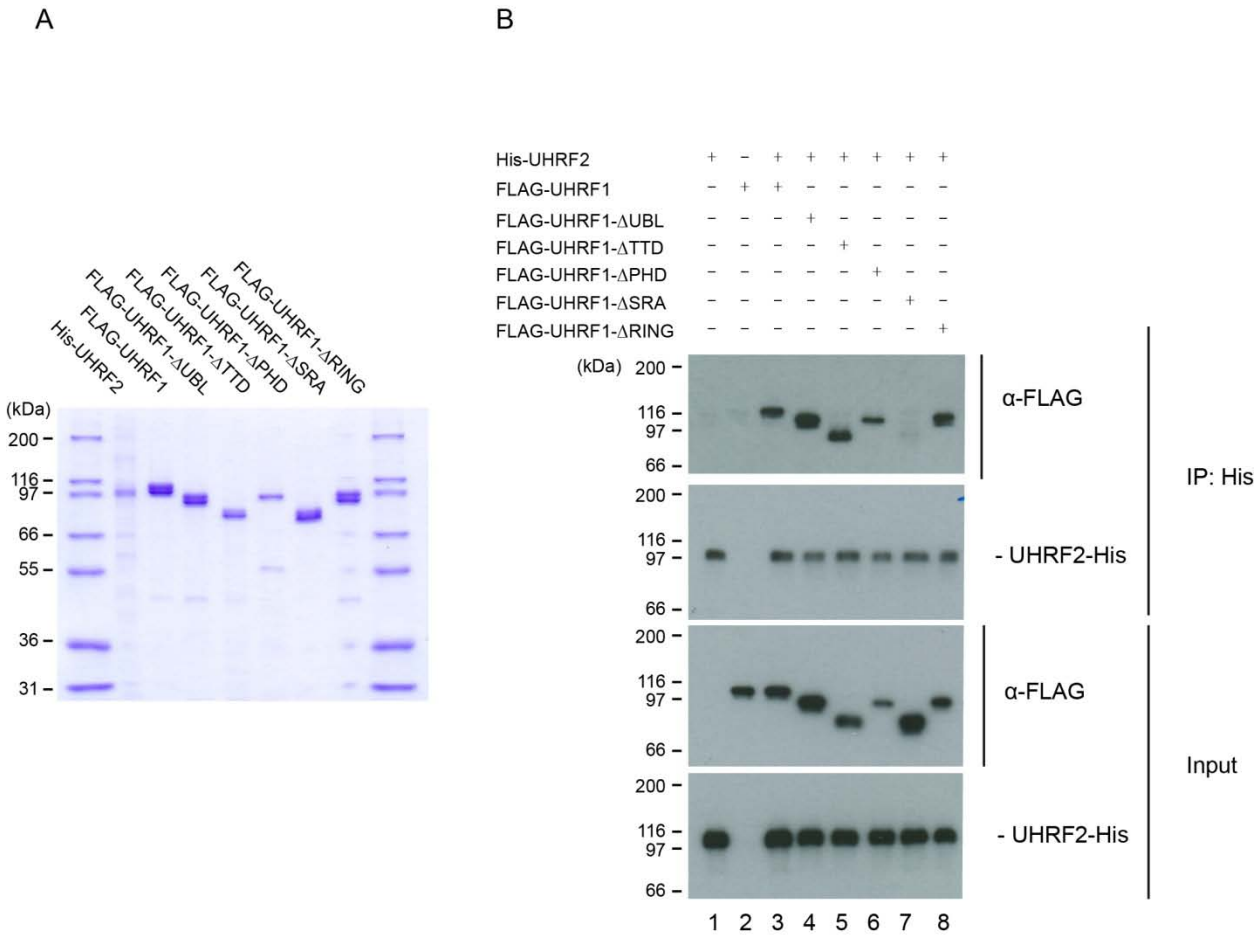


Figure 16. UHRF1 UHRF2 interacting domain mapping

A) Protein purification control. His-UHRF2 and Flag-UHRF1 as well as Flag-UHRF1 domain deletion variants were expressed and purified form *Sf9* insect cells. B) Proteins purified in A were mixed as indicated and subjected to Ni-NTA-His pull down.

3.3.2.3 UHRF1 and UHRF2 are recruited to ICLs independently

UHRF2 and UHRF1 are both recruited to ICLs at similar time frames and have now been shown to directly interact. This raises the question whether their recruitment is dependent on each other or if those are independent events. In order to test this cells were engineered to express fluorophore tagged UHRF1 or UHRF2 while the other protein was knocked down or knocked out. The UHRF2-/- cell line was complemented with UHRF2-EGFP or mCherry-UHRF1, HeLa cells were also stably transfected with mCherry UHRF1. The UHRF2-EGFP expressing cells were then subjected to shRNA mediated depletion of UHRF1. The recruitment of UHRF1/2 in the absence of the other protein was tested using laser microirradiation microscopy (Figure 17). UHRF2 was recruited at the same rate in the presence or absence of UHRF1. UHRF1 was slightly more enriched at the ICLs in the absence of UHRF2. These results suggest that UHRF1 and UHRF2 are recruited independently of each other to sites of ICL damage.

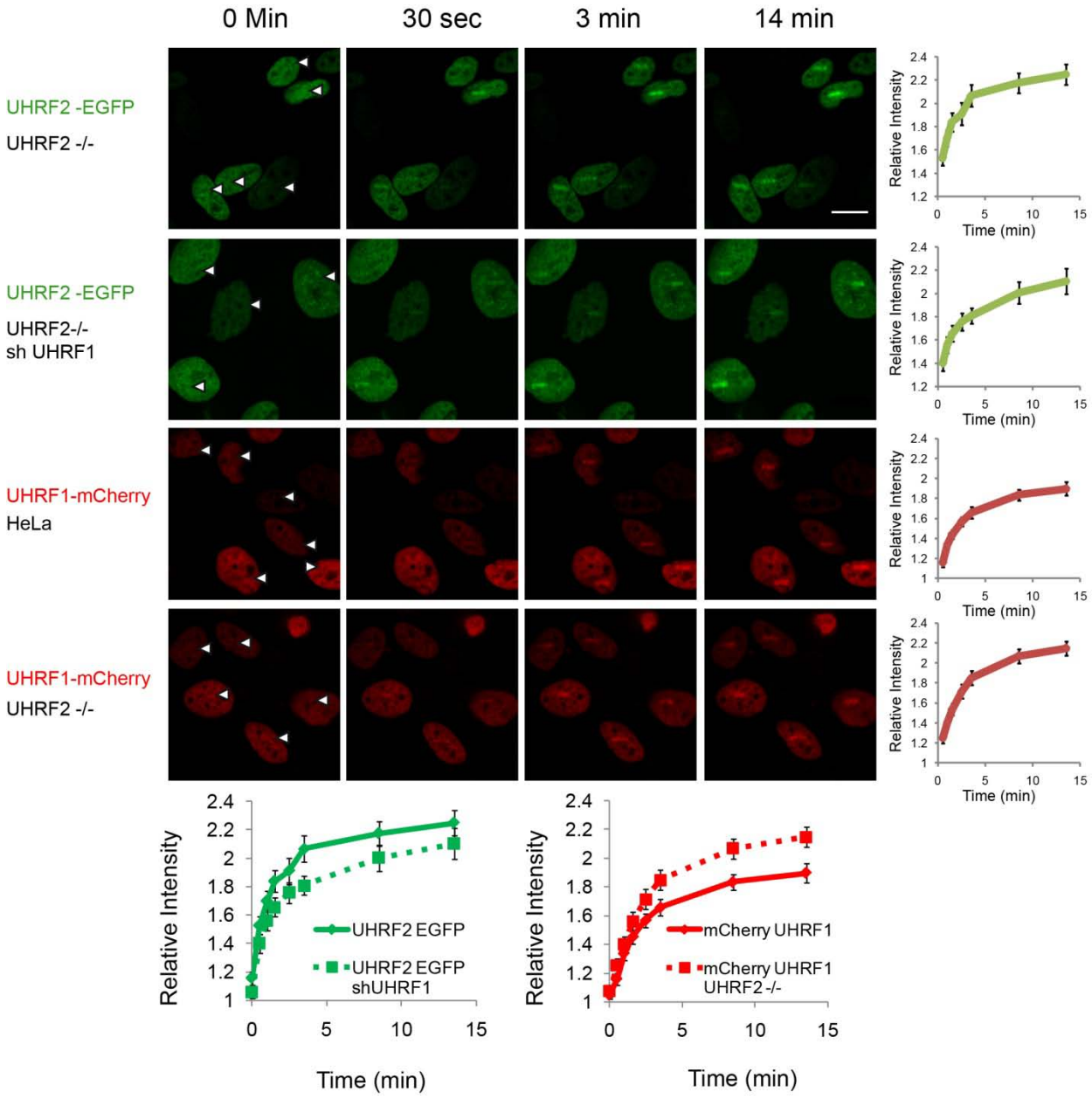


Figure 17. UHRF1 and UHRF2 are recruited to ICLs independently

HeLa cells expressing either EGFP-tagged UHRF2 or mCherry-tagged UHRF1, were depleted of endogenous UHRF2 (CRISPR/Cas-9) and/or UHRF1 (shRNA). The cells were pre-treated with TMP and then microirradiated at the indicated sites. Charts indicate quantification of relative intensity of signal at the irradiated sites. Scale bar indicates 10 μm. Error bars show SEM n = 8/treatment.

3.3.3 Discussion

The paralogue of UHRF2, UHRF1, has been previously shown to be important for sensing ICLs and promoting repair through the FA pathway (Liang et al. 2015). Here it was shown that UHRF2 may either be cooperating with UHRF1 or acting as a backup to facilitate ICL repair.

Firstly, when assessing the impact of co-depletion of UHRF1 and UHRF2 on cell survival when challenged with MMC it was apparent that the knockdown of UHRF1 and the knockout of UHRF2 had a similar level of impact on the cells ability to respond to MMC. This could suggest that UHRF1 and UHRF2 each serve a similar role and that they are not entirely redundant. This is further demonstrated when UHRF2 and UHRF1 are both depleted, the overall survival of the cells is reduced beyond the levels seen in a single depletion, that is the effect is additive. However, UHRF1 and UHRF2 cannot be fully redundant with each other as cell survival is still impacted in the absence of each. Such a situation hints at the possibility that UHRF1 and UHRF2 are each able to carry out a function in ICL repair, but they may need to cooperate for their full function to take place, which argues against redundancy.

In order to assess if UHRF1 and UHRF2 could cooperate it is necessary to examine whether they are able to interact. It was found that each protein was co-immunoprecipitated with the other from HeLa cells. To deduce if this interaction is direct or mediated via a third mediator protein, UHRF1 and UHRF2 were purified and allowed to interact *in vitro*. An interaction between UHRF1 and UHRF2 was again observed, supporting the idea that UHRF1 and UHRF2 interact directly, this however does contradict previously published research stating there is no interaction (Zhang et al. 2011).

As a direct interaction between UHRF1 and UHRF2 was established, it was important to attempt to map a possible region of interaction between them. UHRF1 and UHRF2 both have

five distinct domains, making it an logical step to narrow to narrow the interaction to a domain if possible. Using UHRF1 domain deletion mutants it was found that the interaction between UHRF1 and UHRF2 was lost when the SRA domain of UHRF1 was deleted. This is an interesting result as the SRA domain of UHRF1 is found to interact with FANCD2 (Liang et al. 2015), it has been reported that the SRA of UHRF1 also interacts directly with DNMT1 (Achour et al. 2008; Berkyurek et al. 2014; Bostick et al. 2007). The domain of UHRF2 that interacts with UHRF1 has not been identified yet, it is possible that due to the similarity of the proteins it is also the SRA domain. The SRA domain in general has a saddle like shape (Arita et al. 2008; Zhou et al. 2014) with the DNA fitting inside the hydrophobic core's positively charged patch. It is possible that UHRF1 and UHRF2 could form a clamp on DNA by interacting with each other on the edges of the saddle shape with the DNA fitting inside with FANCD2 interacting on another exposed patch of UHRF1. This hypothesis would fit well with the survival data as UHRF1 and UHRF2 would effectively form half a clamp, allowing some degree of repair to take place, but in both their absence there would be no clamp and severely hampered repair.

Lastly, the recruitment kinetics of both UHRF1 and UHRF2 were examined in the absence of the other protein. Both UHRF1 and UHRF2 recruit to DNA within 30 seconds of microirradiation induced ICLs, but are they recruiting as a dimmer or are they independent? To answer this question cells expressing either EGFP-tagged UHRF2 in the absence of UHRF1, or mCherry-tagged UHRF1 in the absence of UHRF2 were subjected to laser microirradiation in the presence of TMP to induce ICLs. UHRF2 was recruited to ICLs in the same way with or without UHRF1, UHRF1 was recruited slightly more strongly in the absence of UHRF2. These data suggest that UHRF1 and UHRF2 are recruited independently and if an interaction between them occurs, it is only after they are bound on DNA. The enhanced recruitment of UHRF1

however poses an interesting option, as the kinetics of UHRF1 recruitment appeared to look the same with only the final intensity being higher, in the absence of UHRF2, UHRF1 could be forming a dimmer at the ICL site.

Taken together these data support the idea that UHRF1 and UHRF2 are both important for ICL repair and that perhaps this function is facilitated via a direct protein-protein interaction through the SRA domain of UHRF1. This interaction likely takes place on the DNA, as the two proteins are recruited independently.

3.4 UHRF2 is Needed for FANCD2 Mediated Repair

3.4.1 Introduction

At the core of the FA pathway are the two proteins FANCD2 and FANCI. FANCD2 is a key activator for initiating repair and without it the cell is incapable of repairing ICLs in the FA pathway. There are instances of FA patients that harbor either defective FANCD2 or extremely low amounts of it (~2% (Wang 2007)); as an example, the patient derived cell line PD20 which has been used as a FANCD2 null cell line prior to CRISPR/Cas9 (Timmers et al. 2001). It has recently been demonstrated that FANCD2 and FANCI are ubiquitinated after they are successfully recruited to chromatin thereby increasing the affinity of FANCD2 to DNA (Liang et al. 2016). However FANCD2 and FANCI must first be recruited to the ICL site. Previously it was proposed that FANCM, which contains DNA helicase motifs, acts as a landing platform for the FANCD2/FANCD complex (Yan et al. 2010), it was also later proposed that UHRF1 serves the role of sensing the ICL and recruiting FANCD2 to the damage site (Liang et al. 2015). Upon the recruitment of FANCD2 and FANCI to the ICL, the core complex which consists of 9 proteins, including an E3 ligase termed FANCL, ubiquitinates FANCD2 and FANCI initiating the repair via recruitment of downstream effecter proteins.

It is therefore important to understand the mechanism by which FANCD2 is recruited to DNA and subsequently monoubiquitinated.

3.4.2 Results

3.4.2.1 UHRF2 is needed for full FANCD2 recruitment

In order to initiate the FA repair pathway, the key activator, FANCD2 must be present at the ICL site. As UHRF2 and UHRF1 both appear at the ICL site very quickly after damage (detected within 30 sec), they could be preceding the recruitment of FANCD2 and perhaps influencing it. To test whether UHRF2 and UHRF1 play a role in FANCD2 recruitment to ICLs, stably expressed mCherry-tagged FANCD2 was expressed in HeLa and UHRF2 -/- cells, additionally the cells had shRNA mediated UHRF1 knockdown generating four cell lines: HeLa mCh-FANCD2, UHRF2-/- mCh-FANCD2, shUHRF1 mCh-FANCD2, and UHRF2-/- shUHRF1 mCh-FANCD2. The cells were then pre-treated with TMP and subjected to laser microirradiation at the indicated sites (Figure 18A). FANCD2 in wild type cells was found to recruit to ICL sites, becoming visible at approximately 10 min post irradiation. When UHRF1 was depleted from the cells the recruitment of FANCD2 was greatly diminished both delaying its appearance at the ICL sites and reducing overall signal intensity. In cells lacking UHRF2 the recruitment of FANCD2 was severely reduced. When both UHRF2 and UHRF1 are depleted from cells again FANCD2 recruitment was nearly abrogated.

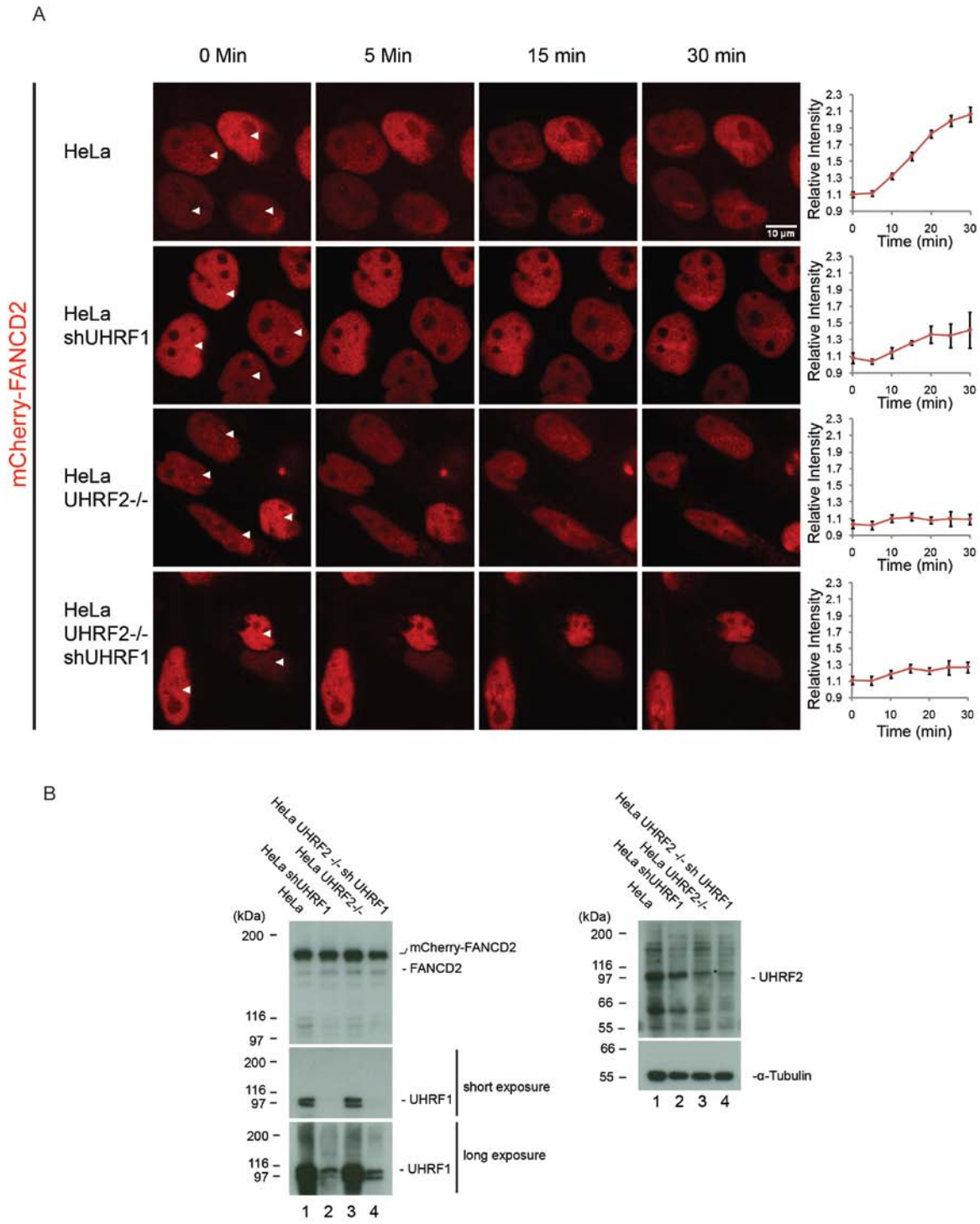


Figure 18. FANCD2 recruitment to ICLs is dependent on UHRF2 and UHRF1

A) HeLa cells expressing mCherry-tagged FANCD2 where subjected to depletion of UHRF1 by shRNA and/or depletion of UHRF2 by CRISPR/Cas9-mediated knockout, pre-treated with TMP, and then microirradiatedat the sites indicated with white arrows. Charts indicate quantification of relative intensity of signal at the irradiated sites. Depletion of UHRF1 and UHRF2 impairs FANCD2 recruitment. Scale bar indicates 10 µm n = 8/treatment. B) Western blot of the cells used in A, demonstrating expression level of all relevant proteins. Asterisk is non-specific band.

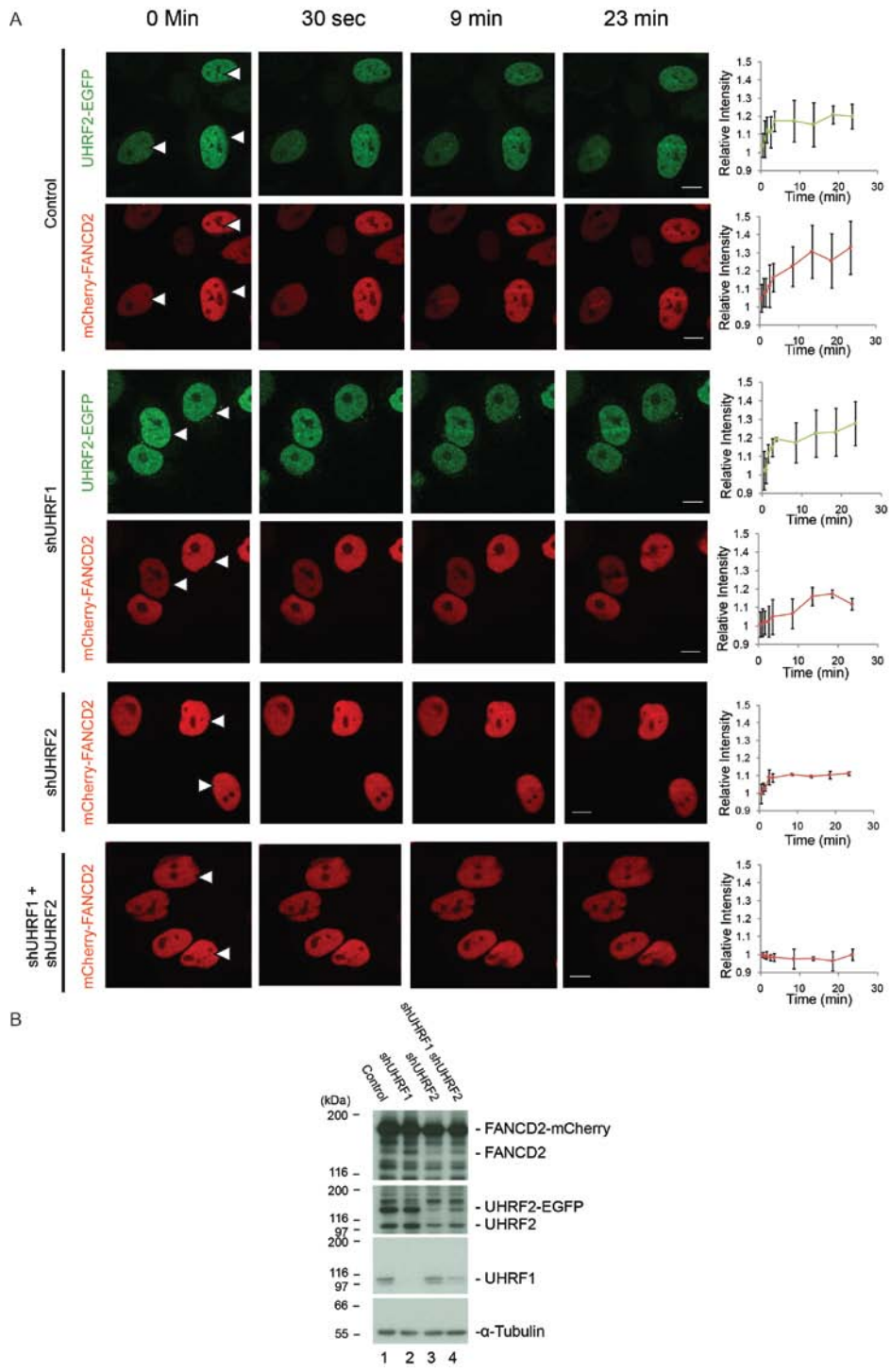


Figure 19. FANCD2 recruitment to ICLs is impeded with UHRF1 and UHRF2 knockdown

A) HeLa cells expressing mCherry-tagged FANCD2, and UHRF2-EGFP when indicated, where subjected to depletion of UHRF1 and/or UHRF2 by shRNA, pre-treated with TMP, and then microirradiatedat the sites indicated with white arrows. Charts indicate quantification of relative intensity of signal at the irradiated sites. Depletion of UHRF1 and UHRF2 impairs FANCD2 recruitment. Scale bar indicates 10 μm. Error bars show SEM, n = 5/treatment. B) Western blot of the cells used in A, demonstrating expression level of all relevant proteins.

As the UHRF2-/- cells are a clone, to compensate for possible clonal effects another experiment was carried out where HeLa cells expressing mCherry-tagged FANCD2 were depleted of UHRF1 and UHRF2 using shRNA. The cells were pre-treated with TMP and subjected to laser microirradiation at the sites indicated (Figure 19A). It was again found that wild-type FANCD2 was recruited to ICLs becoming visible at about 9 minutes. When depleted of UHRF1 the recruitment of FANCD2 to the ICL sites was reduced, the same was observed in cells depleted of UHRF2. When UHRF1 and UHRF2 were co-depleted the recruitment of FANCD2 was nearly abolished.

3.4.2.2 UHRF2 directly interacts with FANCD2

UHRF1 and UHRF2 preceded the recruitment of FANCD2 to ICLs and in their absence FANCD2 recruitment is severely diminished. It is therefore possible that UHRF1 and UHRF2 serve to recruit FANCD2 to the ICLs site, possibly through a direct protein-protein interaction. To access whether that is the case purified recombinant Flag-HA-FANCD2, Strep-His-UHRF1, and His-UHRF2 were used in a co-immunoprecipitation assay. Using anti-HA beads FANCD2 was immunoprecipitated in the presence of UHRF1, UHRF2, UHRF1 and UHRF2 together, and MBP (Figure 20). It was found that both UHRF1 and UHRF2 co-immunoprecipitated with FANCD2 *in vitro* (Figure 20 lanes 12 and 13), while MBP, which was used as a control, did not. These results demonstrate a direct protein-protein interaction between UHRF1/UHRF2 and FANCD2.

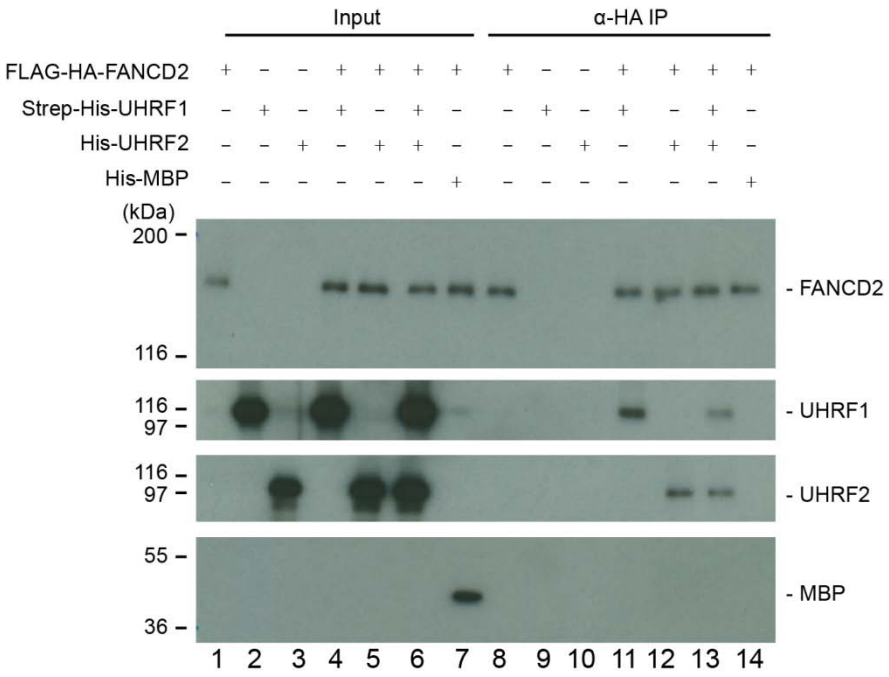


Figure 20. UHRF1 and UHRF2 interact directly with FANCD2 *in vitro*

In vitro binding assay using Flag-HA-FANCD2, Strep-His-UHRF1 , His-UHRF2 (purified from *Sf9* insect cells) and His-MBP (purified from *E.coli*). HA-FANCD2 was immunoprecipitated and immunoblotting demonstrated co-immunoprecipitation of UHRF1 and/or UHRF2 but not of MBP (used as a negative control).

3.4.2.3 UHRF2 plays a role in FANCD2 ubiquitination

Monoubiquitination of FANCD2 on lysine 561 is absolutely required for its activation and accumulation on ICLs *in vivo*. Therefore, it was speculated that UHRF1 and UHRF2 might affect the monoubiquitination of FANCD2 as they precede the recruitment of FANCD2 to ICLs and are able to directly interact with FANCD2. To check this, UHRF1, UHRF2 or both UHRF1 and UHRF2 were depleted in HeLa cells. To determine the ability of the modified cells to monoubiquitinate FANCD2, the cells were treated with TMP and UVA and harvested as time points post irradiation, the proportion of ubiquitinated FANCD2 was assessed with western blotting. It was observed that there was a clear increase in the amount of monoubiquitinated FANCD2 in control cells after 3 and 6 hours (Figure 21A lanes 2 and 3). In contrast, monoubiquitination was markedly reduced upon depletion of UHRF1 and/or UHRF2 (Figure 21A lanes 5-12). The ratio of ubiquitinated FANCD2 to non-ubiquitinated FANCD2 was calculated from three independent experiments (Figure 21B). The overall change from non treated to 6 hours post treatment in wild type HeLa is represented by the blue arrow on the right with an increase of 0.4. The overall change from non-treated to 6 hours post treatment of the double depleted cells is represented by the purple arrow on the right with an increase of 0.2, so half as much as the control. As UHRF1 is involved in methylation (Bostick et al. 2007) and UHRF2 is known to be involved in the cell cycle (Mori et al. 2011) a FACS analysis was carried out on the cells used in (Figure 21A) without damage (Figure 22A). It was determined that there were no significant changes in the populations of S-phase cells upon reduction of UHRF1 and UHRF2, ruling out that the observed phenotype could be caused by altered cell cycle progression (Figure 22B). It is worth noting that traditionally FA cells will exhibit G2 peak accumulation in the presence of crosslinking agents (Reid et al. 2007). While the cells were not treated, a minor G2 peak accumulation is visible in the double knockdown cells.

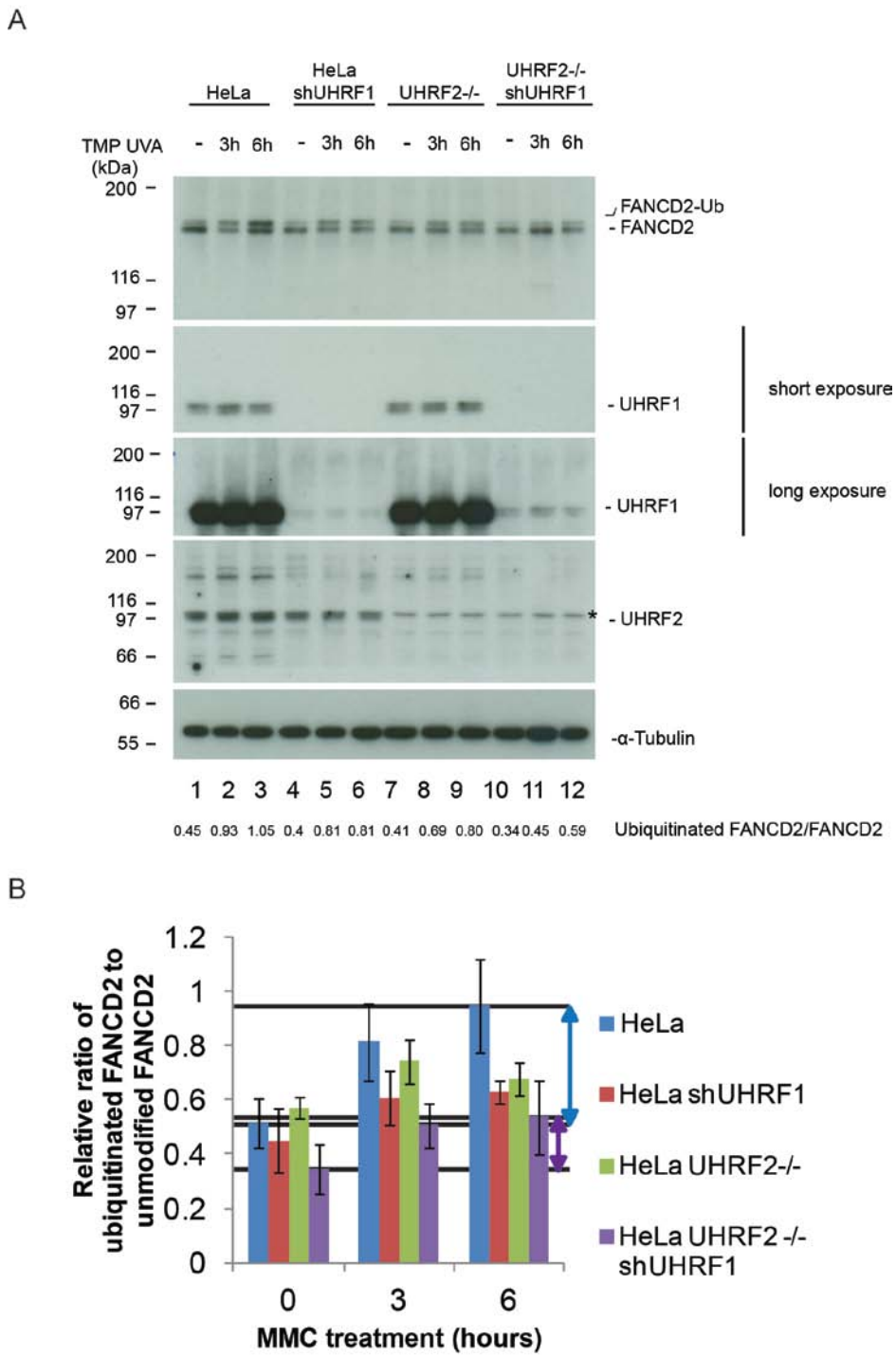


Figure 21. UHRF2 and UHRF1 affect FANCD2 moniubiquitination

A) A representative western blot analysis of lysates from HeLa cells or HeLa cells where UHRF1 and/or UHRF2 were depleted using shRNA-mediated knockdown or CRISPR/Cas9-mediated knockout following treatment with TMP/UVA and harvested at 3 and 6 hours. Strong accumulation of monoubiquitinated FANCD2 (FANCD2-Ub) occurs in HeLa cells but is significantly reduced when UHRF1 and UHRF2 are depleted. Asterisk indicate unspecific band. B) Chart represents data from three independent experiments.

To ensure that the observed phenotype is real a series of controls were carried out. The monoubiquitination in response to TMP UVA was carried out in UHRF2-/- cells as well as cells depleted of UHRF2 using shRNA (Figure 23A) in an effort to rule out clonal effect. It was found that in both cases of UHRF2 depletion, the ubiquitination of FANCD2 was affected in a similar way as seen by the quantification of the ratios between ubiquitinated FANCD2 and unmodified FANCD2 (Figure 23B). Historically UHRF1 and UHRF2 were thought to be redundant so to assess if they are, in regard to FANCD2 monoubiquitination, UHRF1 was over expressed in UHRF2 knockout cells. These cells were also treated with TMP+UVA and it was observed that over expression of UHRF1 in UHRF2 deficient cells did not provide a significant rescue of the phenotype (Figure 23 lanes 10-12).

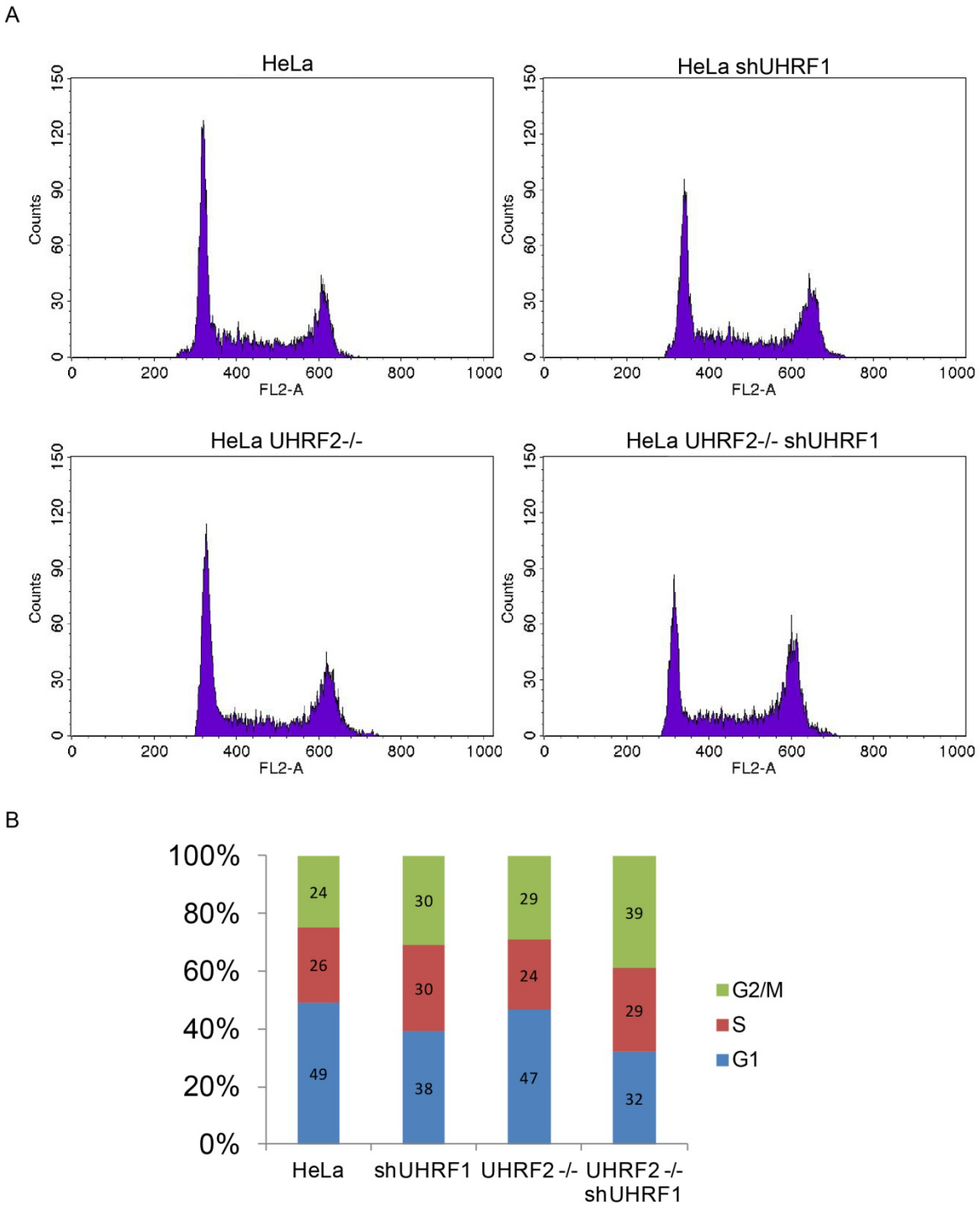


Figure 22. Depletion of UHRF2 and UHRF1 does not significantly affect the cell cycle distribution

A) FACS analysis of cell lines used in Figure 21A. B) Graph showing percent of total cells in the G1 and G2/M peaks and S-phase region.

It was demonstrated herein that UHRF1 and UHRF2 are needed for FANCD2 localization to ICLs and its monoubiquitination in response to damage, it was important to investigate whether UHRF1 and UHRF2 are able to directly ubiquitinate FANCD2 via their RING domains (are E3 ligases). To test this, an *in vitro* ubiquitination assay was used. The purified components of the ubiquitination reaction: UBA1 (E1), UBE2T (E2), FANCL (E3), Ubiquitin, FANCI, and FANCD2, are used in conjunction with ATP and DNA (pBluescript plasmid) to carry out a robust ubiquitination reaction of FANCD2 in a test tube (Figure 24 lanes 2 and 3). When switching out FANCL for UHRF1 or UHRF2 there was little to no visible ubiquitination of FANCD2 (Figure 24 lanes 5 and 7). Suspecting that perhaps the E2 ligase UBE2T was not compatible with UHRF1 and UHRF2, it was switched to UBCH5c (UBE2D3), UBCH5c is an E2 with low selectivity for E3 ligases and would therefore be more likely to function with UHRF1 or UHRF2 (Gatti et al. 2015). Again no ubiquitination of FANCD2 occurred (Figure 24 lanes 9 and 11). Taken together these experiments show that UHRF1 and UHRF2 are unable to directly ubiquitinate FANCD2.

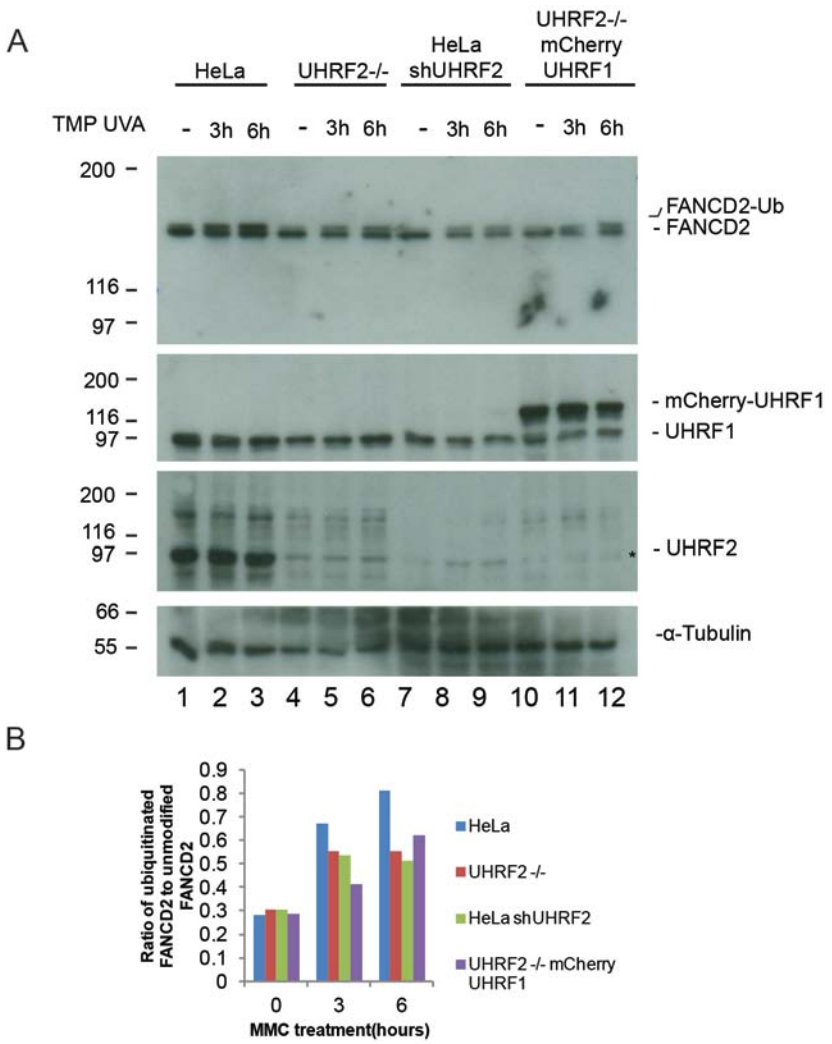


Figure 23. UHRF2 is needed for FANCD2 monoubiquitination

A) Western blot analysis of lysates from HeLa cells or HeLa cells where UHRF2 were depleted using shRNA-mediated knockdown or CRISPR/Cas9-mediated knockout, in addition UHRF2 -/- cells were over expressing mCherry-tagged UHRF1 following treatment with TMP/UVA and harvested at 3 and 6 hours. B) Quantification of the ratio of ubiquitinated FANCD2 to unmodified FANCD2 in the western blot in A. Strong accumulation of monoubiquitinated FANCD2 (FANCD2-Ub) occurs in HeLa cells but is significantly reduced when UHRF2 is depleted, over expression of UHRF1 does not rescue phenotype. Asterisk is non-specific band. Lanes 1, 4, and 7 are also used in Figure 9E

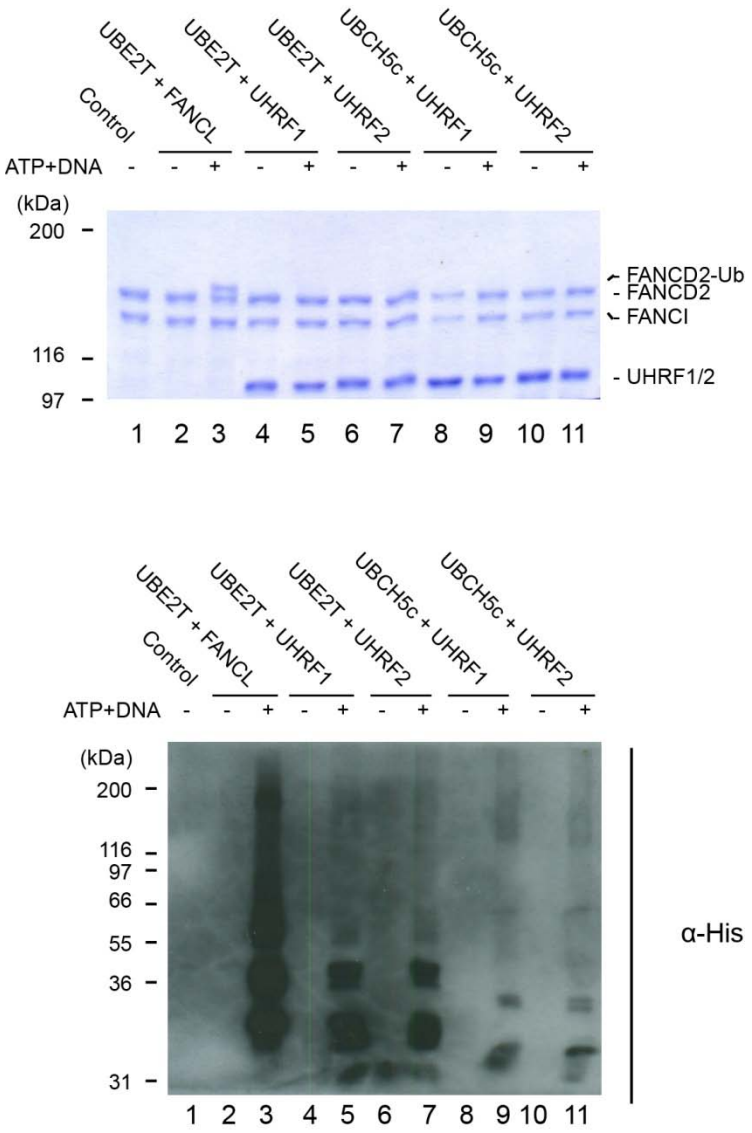


Figure 24. UHRF1 and UHRF2 cannot ubiquitinate FANCD2

In vitro ubiquitination assay of FANCD2. FANCL (E3 ligase) supports robust ubiquitinatination of FANCD2. Replacement of FANCL by UHRF1 or UHRF2 does not support FANCD2 monoubiquitination. Switching the E2 ligase UBE2T with UBCH5c (C5) does not allow ubiquitination of FANCD2. Work was done in collaboration with David Lopez-Martinez.

3.4.3 Discussion

The recruitment and activation, via ubiquitination at lysine 561, of FANCD2 is vital for the activation of the FA pathway and successful repair of the ICL (Garcia-Higuera et al. 2001). Until fairly recently there was no known factor responsible for the localization of FANCD2 to ICLs, UHRF1 was published as the protein responsible for FANCD2 recruitment to ICLs (Liang et al. 2015). Here it was demonstrated that UHRF2 also plays an important role in FANCD2 localization and subsequent activation.

To answer the question of whether UHRF2 and UHRF1 are required for FANCD2 localization to ICLs, cells expressing mCherry fluorophore-tagged FANCD2 was depleted of either UHRF1 or UHRF2 or both at the same time. The cells were then exposed to TMP and a 405 nm laser, to create localised ICLs and the recruitment of FANCD2 was observed. When UHRF1 was depleted the recruitment of FANCD2 was diminished but not abrogated, on the other had when the UHRF2 protein was deleted, FANCD2 recruitment was abolished. While this may appear to suggest that FANCD2 recruitment is mostly depended on UHRF2 it would be an unfair comparison as *UHRF2* was deleted using CRISPR/Cas9 while UHRF1 was only reduced in the cell using shRNA and residual UHRF1 was still present, all be it in minute amounts (<1%). As such, the possible additive effect UHRF1 and UHRF2 may have on FANCD2 recruitment may not be fully visible, as when both UHRF1 and UHRF2 were depleted the recruitment of FANCD2 was also severely reduced. Using only shRNA to reduce the levels of UHRF1 and UHRF2 provides a much fairer comparison. When each protein was depleted, the recruitment of FANCD2 was diminished but this time not abrogated. It was only when both proteins were depleted that the recruitment was again abolished. These results point to UHRF1 and UHRF2 cooperating in the recruitment of FANCD2 to ICLs and either one is not sufficient on its own for this purpose.

It was established that UHRF2 and UHRF1 recruitment precedes the appearance of FANCD2 at ICLs. In fact the presence of UHRF2 and UHRF1 support the subsequent localization of FANCD2 to the site, it was important to address if this cascade is mediated via a direct interaction or if there is another unknown protein mediating the interaction. To answer this question, purified recombinant FANCD2, UHRF1 and UHRF2 were incubated together and FANCD2 was immunoprecipitated. It was found that both UHRF2 and UHRF1 were co-precipitated with FANCD2, demonstrating a direct protein-protein interaction. This would suggest that UHRF2 and UHRF1 are able to directly bind FANCD2, which could be the mechanism by which FANCD2 is recruited to the ICL in *vivo*.

A key step in the activation of the FA pathway is the monoubiquitination for FANCD2. The FA core complex is responsible for carrying out this ubiquitination, with the catalytic E3 ligase being FANCL. This ubiquitination event takes place while the FANCD2/FANCI dimer is already present on DNA (Liang et al. 2016). As UHRF2 and UHRF1 are E3 ligases that directly interact with FANCD2, it was speculated that they may ubiquitinate or regulate the ubiquitination of FANCD2. To test if that is the case the *in vivo* ubiquitination of FANCD2 in response to ICL induction was measured in a wild type scenario as well as when UHRF1, UHRF2, and UHRF1 and UHRF2 together were depleted. The results demonstrate that loss of either UHRF1 or UHRF2 reduces the proportion of FANCD2 that is ubiquitinated in response to damage. However, double depletion does not abrogate ubiquitination completely, but does further reduce it. As UHRF1 is still present in small quantity in the double depleted cells perhaps it could have still been contributing to FANCD2 ubiquitination. In order to control for all factors in the FANCD2 ubiquitination, an *in vitro* ubiquitination experiment was carried out using purified recombinant proteins. The assay uses the absolute minimum combination of proteins to

carry out robust ubiquitination of FANCD2. When the E3 ligase FANCL was substituted for UHRF1 or UHRF2, FANCD2 was not ubiquitinated to any appreciable degree (there was a very light shadow of a band visible). In an effort to rule out incompatibility of the E2 ligase, a more broad spectrum E2 was chosen. Substituting UBE2T for UBCH5c did not allow FANCD2 to be ubiquitinated. This collectively rules out FANCD2 as a substrate for UHRF1 and UHRF2 E3 ligase activity.

Taken together these data support UHRF1 and UHRF2 being recruited to ICLs and directly interacting with FANCD2 at the ICL site, allowing it to be ubiquitinated by the core complex and initiate the repair.

Chapter 4: General discussion

The proper recognition and marking of ICLs is indispensable to their repair. In the past it was suggested that FANCM recruits FANCD2 and FANCI to the ICL site and facilitates core complex recruitment (Kim et al. 2008). More recently UHRF1 was identified as an ICL sensor, recruiting FANCD2/FANCI to ICLs (Liang et al. 2015). In this study the role of UHRF2, the paralogue of UHRF1, was added into the overall pathway of ICL repair. This was achieved through several approaches including *in vivo* and *in vitro* studies.

Using EMSA it was established that UHRF2 can bind DNA and has a higher affinity for ICLs compared to non-crosslinked DNA, it was also confirmed that UHRF2 does not differentiate between non-methylated and hemi methylated DNA, as stated in literature (Zhou et al. 2014). However, this result only factors in one DNA interacting domain in UHRF2, the SRA domain. UHRF2 is able to interact with DNA through its SRA domain, which recognizes DNA directly as well as DNA modifications, such as methylation, hydroxymethylation, and ICLs. Additionally, UHRF2 can interact with DNA in a more indirect fashion though its TTD and PHD

domains. TTD and PHD recognize histone tails and histone tail methylation marks, specifically histone H3 and histone H3K9 di and tri methylation. The histone methylation recognition aspects of UHRF2 have been explored in the past but there have been no direct claims as to the function these domains have in the cell other than UHRF2 localization, which is predominantly to pericentric hetrochromatin (Pichler et al. 2011; Zhang et al. 2011). Together these domains work to localize UHRF2 to DNA for various purposes (Figure 25). When considering that the SRA domain binds naked DNA, and the TTD-PHD domains bind histones it is logical to conclude that using naked DNA in EMSA only accounts for a portion of the UHRF2-DNA interaction. To address the interaction of UHRF2 with DNA in a more biological context *in vivo* analysis was the next step; where the DNA interaction properties of UHRF2 were examined using microirradiation and live-cell microscopy. Using laser microirradiation to induce ICLs, it was shown that UHRF2 was recruited to the ICLs in live cells very quickly (within 30 seconds). This recruitment was abrogated when the SRA domain of UHRF2 was deleted, which supports that the ICL recognition function of UHRF2 is linked to the SRA domain. However, the recruitment intensity was reduced when the TTD or PHD domains were deleted. This could suggest that while the primary ICL recognition of UHRF2 relies on the SRA domain, the TTD and PHD domains may serve to strengthen the UHRF2-DNA interaction through the SRA.

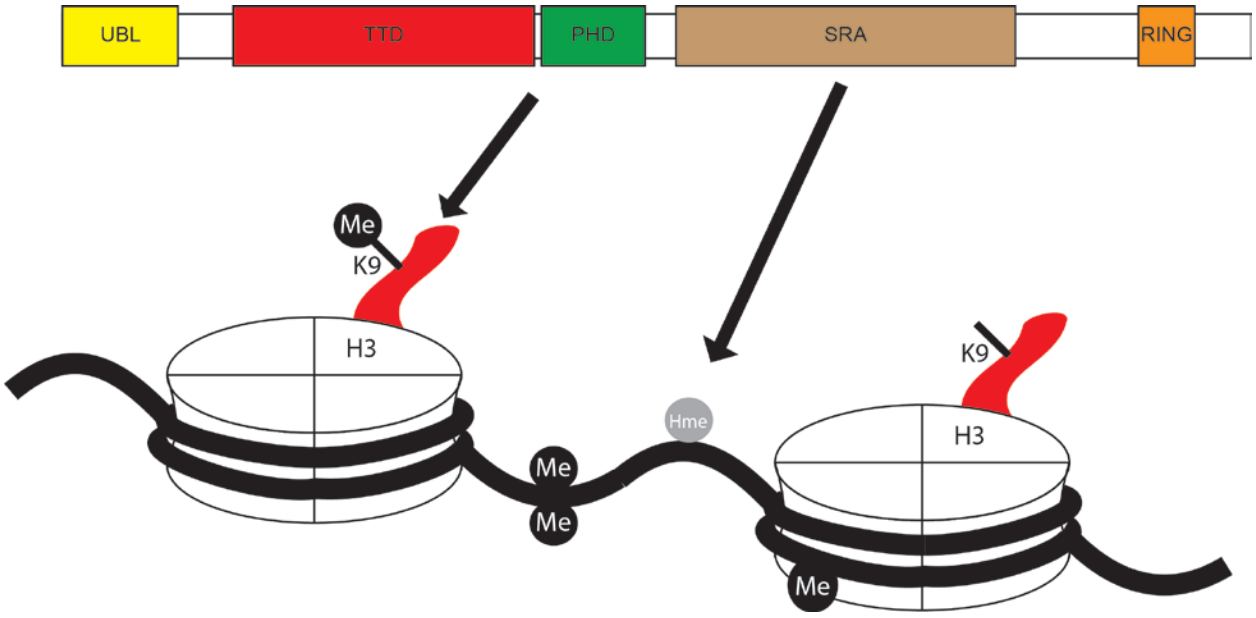


Figure 25. Schematic of DNA recognising domains of UHRF2

The SRA domain of UHRF2 can recognise hydroxy-methylation or methylation markers on cytosine on DNA. The TTD and PHD domains of UHRF2 cooperatively interact and bind histone H3 methylation markers.

As the SRA domain in UHRF1 has been shown to be important for hemi methylated DNA maintenance and the SRA domain of UHRF2 has been tied to hydroxymethyl cytosine recognition, it was important to consider those methylation markers in the context of ICLs. Ergo, when cells were treated with a cytosine analogue which blocks propagation of methylation markers in cells during division, thereby causing methylation loss, it was found that binding of UHRF2 and UHRF1 was independent of the DNA methylation status of the cells as UHRF1 and UHRF2 bound to microirradiation induced ICLs in an identical fashion to the non treated cells. This further demonstrated the versatility of the SRA domain, as it was recognizing ICLs independently of methylation. UHRF1 and UHRF2 are the only human proteins to contain the SRA domain so this versatility is quite interesting because it gives UHRF2 and UHRF1 a second role that is unrelated to epigenetics, opening up the possibility to multiple functions of the SRA domain.

Having established that UHRF2 also recognizes ICLs much like UHRF1 it was important to answer the question of redundancy between these two proteins. Early literature had noted that UHRF2 ectopic expression failed to rescue methylation defects caused by UHRF1 knockout, this was an early argument for non-redundancy between UHRF1 and UHRF2 relating to mathylation maintenance (Pichler et al. 2011). Later it was shown that UHRF2 preferentially recognizes hydroxymethylated DNA, a function distinct from UHRF1 (Zhou et al. 2014; Spruijt et al. 2013). These functions set UHRF1 and UHRF2 apart in term of function in the cell; however, when this question was addressed in the context of ICL repair, a convergent relationship was found. UHRF1 or UHRF2 depletion had similar effects on cells' ability to survive MMC treatment, recruitment and foci formation of FANCD2, and the level of FANCD2 mono ubiquitination post DNA damage. When UHRF1 and UHRF2 were depleted together the effects were additive. A

much-reduced survival in the presence of MMC, less foci and less recruitment of FANCD2 post damage, and a strong reduction of FANCD2 mono ubiquitination after DNA damage were observed. These results argue for some degree of redundancy between UHRF1 and UHRF2, but both have a function in ICL response, one is not a complete replacement for the other.

To follow from the conclusion that UHRF1 and UHRF2 are not fully redundant, the question becomes what role does each play in the ICL response. It was shown that UHRF1 and UHRF2 both precede and are needed for subsequent FANCD2 recruitment. Those data were obtained in cells where one of the two proteins was still present, therefore it was important to address if the diminished FANCD2 recruitment was due to the absence of one of the proteins directly or if the present protein was poorly recruited because the other was missing, and therefore so was FANCD2. It was demonstrated that UHRF1 and UHRF2 are each recruited independently of each other and therefore the recruitment phenotype of FANCD2 was a direct results of the absence of UHRF1 or UHRF2, suggesting these two proteins may cooperate in some way. This was further strengthened by the observed direct protein-protein interaction between UHRF1 and UHRF2.

The protein-protein interaction between UHRF1 and UHRF2 is surprising considering the differential expression pattern of these proteins. UHRF1 is predominantly present in undifferentiated cells and its levels decrease as the cells undergo specialization, although its expression is never turned off as this protein is essential for methylation (Bostick et al. 2007; Miura et al. 2001). UHRF2 on the other hand, while also being ubiquitously expressed, is predominantly present in differentiated cells (Pichler et al. 2011). The expression patterns of these proteins demonstrate that their cellular function is related to processes other than ICL

repair; however, the results described herein argue that there is a subset population within each cell that responds to ICL damage.

Chapter 5: Conclusion

Taken all together these data can be used in the construction of a possible updated model of the early steps in the ICL repair pathway (Figure 26). When an ICL forms UHRF1 and UHRF2 independently detect the damage and are recruited to the damage site. Once at the site they are able to interact directly through their SRA domains. FANCD2 and FANCI come to the damage site, either through a yet unknown signal or simply through stochastic interaction. UHRF1 and UHRF2 together interact with FANCD2 and hold it in place, increasing the retention time long enough for the FA core complex to ubiquitinate the FANCD2/FANCI dimer and lock it onto the DNA. The now active dimer is able to carry out additional signalling and initiate downstream repair.

Here we propose that UHRF2 is a potential new ICL sensor protein working in the FA pathway to repair DNA interstand crosslinks.

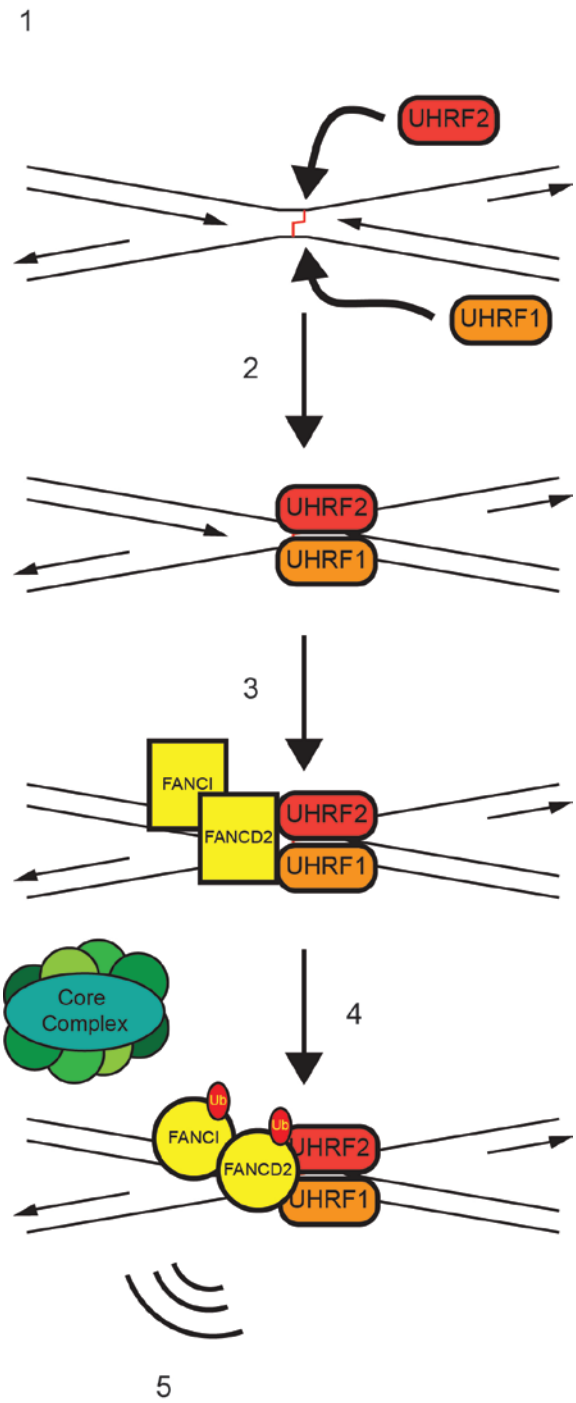


Figure 26. Proposed model of ICL damage response

A basic schematic of ICL recognition and FA pathway activation proposed on the date collected. An ICL is formed in the cell and is detected by UHRF1 and UHRF2 (1). Both UHRF1 and UHRF2 are independently recruited to the ICL site (2) marking it. It is possible that UHRF1 and UHRF2 are interacting with each other through their SRA domain, acting like a clamp. FANCD2 and FANCI are recruited to the ICL site and FANCD2 directly interacts with UHRF1 and UHRF2 (3). While FANCD2 is held in place by UHRF1 and UHRF2 the FA core complex is recruited and ubiquitinates FANCD2 and FANCI (4) thereby changing their confirmation and locking them onto the DNA. The now active heterodimer activates the downstream signaling cascade and recruits the necessary exo nucleases and recombinases to repair the ICL (5).

Chapter 6: References

Achour, M. et al., 2008. The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of VEGF gene expression. *Oncogene*, 27(May 2007), pp.2187–2197.

Ameziane, N. et al., 2015. A novel Fanconi anaemia subtype associated with a dominant-negative mutation in RAD51. *Nature Communications*, 6(May), pp.1–11.

Arita, K. et al., 2008. Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature*, 455(7214), pp.818–821.

Auerbach, A.D., 2009. Fanconi anemia and its diagnosis. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, 668(1–2), pp.4–10.

Avvakumov, G. V et al., 2008. Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature*, 455(October), pp.822–826.

Berkyurek, A.C. et al., 2014. The DNA methyltransferase Dnmt1 directly interacts with the SET and RING finger-associated (SRA) domain of the multifunctional protein Uhrf1 to facilitate accession of the catalytic center to hemi-methylated DNA. *Journal of Biological Chemistry*, 289(1), pp.379–386.

Bhagwat, N. et al., 2009. XPF-ERCC1 participates in the Fanconi anemia pathway of cross-link repair. *Molecular and cellular biology*, 29(24), pp.6427–37.

Bird, a P., 1986. CpG-rich islands and the function of DNA methylation. *Nature*, 321(6067), pp.209–213.

Bluteau, D. et al., 2017. Biallelic inactivation of *REV7* is associated with Fanconi anemia. *Journal of Clinical Investigation*, 127(3), pp.1117–1117.

Bogliolo, M. et al., 2013. Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *American Journal of Human Genetics*, 92(5), pp.800–806.

Bostick, M. et al., 2007. UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells. *Science*, (September), pp.1760–1765.

Boyer, L.A. et al., 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature*, 441(7091), pp.349–353.

Branco, M.R., Ficz, G. & Reik, W., 2012. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nature reviews. Genetics*, 13(1), pp.7–13.

Bronner, C. et al., 2007. The UHRF family: Oncogenes that are drugable targets for cancer therapy in the near future? *Pharmacology and Therapeutics*, 115(3), pp.419–434.

Budzowska, M. et al., 2015. Regulation of the Rev 1 – pol ζ complex during bypass of a DNA interstrand cross-link. *EMBO Journal*, 34(14), pp.1971–1985.

Cadet, J. & Wagner, R.J., 2013. DNA Base Damage by Reactive Oxygen Species , Oxidizing Agents, and UV Radiation. *Cold Spring Harb Perspect Biol*, 5, pp.1–18.

Cimprich, K.A. & Cortez, D., 2008. ATR: an essential regulator of genome integrity. *Nature Reviews Molecular Cell Biology*, 9(8), pp.616–627.

Cline, S.D. & Hanawalt, P.C., 2003. Who's on first in the cellular response to DNA damage? *Nature Reviews Molecular Cell Biology*, 4(5), pp.361–373.

Cohn, M.A. et al., 2009. UAF1 is a subunit of multiple deubiquitinating enzyme complexes. *Journal of Biological Chemistry*, 284(8), pp.5343–5351.

Cohn, M. a et al., 2001. Characterization of Sp1, AP-1, CBF and KRC binding sites and minisatellite DNA as functional elements of the metastasis-associated mts1/S100A4 gene intronic enhancer. *Nucleic acids research*, 29(16), pp.3335–3346.

Davis, A. & Chen, D., 2013. DNA double strand break repair via non-homologous end-joining. *Translational Cancer Research*, 2(3), pp.130–43.

Deriano, L. & Roth, D.B., 2013. Modernizing the Nonhomologous End-Joining Repertoire: Alternative and Classical NHEJ Share the Stage. *Annual Review of Genetics*, 47(1), pp.433–455.

Difilippantonio, S. et al., 2005. Role of Nbs1 in the activation of the Atm kinase revealed in humanized mouse models. *Nature cell biology*, 7(7), pp.675–685.

Dong, H. et al., 2015. Update of the human and mouse Fanconi anemia genes. *Human Genomics*, 9(1), p.32.

Dorsman, J.C. et al., 2007. Identification of the Fanconi anemia complementation group I gene, FANCI. *Cellular oncology : the official journal of the International Society for Cellular Oncology*, 29(3), pp.211–218.

Dronkert, M.L.. & Kanaar, R., 2001. *Repair of DNA interstrand cross-links*,

Espositos, F., Brankamp, R.G. & Sinden, R.R., 1988. DNA Sequence Specificity of 4,5 ',8-Trimethylpsoralen Cross-linking. *The Journal of Biological Chemistry*, 263(23), pp.11466–11472.

Felle, M. et al., 2011. The USP7/Dnmt1 complex stimulates the DNA methylation activity of Dnmt1 and regulates the stability of UHRF1. *Nucleic Acids Research*, 39(19), pp.8355–8365.

Ficz, G. et al., 2011. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature*, 473(7347), pp.398–402.

Garcia-Higuera, I. et al., 2001. Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Molecular Cell*, 7(2), pp.249–262.

Gatti, M. et al., 2015. RNF168 promotes noncanonical K27ubiquitination to signal DNA damage. *Cell Reports*, 10(2), pp.226–238.

Gravel, S. et al., 2008. DNA helicases Sgs1 and BLM promote DNA double-strand break resection. *Genes and Development*, 22(20), pp.2767–2772.

Hashimoto, H. et al., 2008. The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature*, 455(7214), pp.826–829.

Hegde, M.L., Hazra, T.K. & Mitra, S., 2008. Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. *Cell Research*, 18(1), pp.27–47.

Heyer, W.-D., Ehmsen, K.T. & Liu, J., 2010. Regulation of Homologous Recombination in Eukaryotes. *Annual Review of Genetics*, 44(1), pp.113–139.

Hira, A. et al., 2015. Mutations in the gene encoding the E2 conjugating enzyme UBE2T cause fanconi anemia. *American Journal of Human Genetics*, 96(6), pp.1001–1007.

Holliday, A.R. & Pugh, J.E., 1975. DNA Modification Mechanisms and Gene Activity during Development. *Science*, 187(4173), pp.226–232.

Horton, J.K. et al., 2008. XRCC1 and DNA polymerase beta in cellular protection against cytotoxic DNA single-strand breaks. *Cell Research*, 18(1), pp.48–63.

Howlett, N.G., 2002. Biallelic Inactivation of BRCA2 in Fanconi Anemia. *Science*, 297(5581), pp.606–609.

Huang, J. et al., 2013. The DNA Translocase FANCM/MHF Promotes Replication Traverse of DNA Interstrand Crosslinks. *Molecular Cell*, 52(3), pp.434–446.

Hwang, G.S., Kim, J.K. & Choi, B.S., 1996. The solution structure of a psoralen cross-linked DNA duplex by NMR and relaxation matrix refinement. *Biochemical and biophysical research communications*, 219(1), pp.191–7.

Iwata, A. et al., 2009. Intranuclear degradation of polyglutamine aggregates by the ubiquitin-proteasome system. *Journal of Biological Chemistry*, 284(15), pp.9796–9803.

Jasin, M. & Rothstein, R., 2013. Repair of strand breaks by homologous recombination. *Cold Spring Harbor perspectives in biology*, 5(11).

Jenuwein, T. et al., 1998. SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cellular and Molecular Life Sciences*, 54(1), pp.80–93.

Jia, Y. et al., 2016. Negative regulation of DNMT3A de novo DNA methylation by frequently overexpressed UHRF family proteins as a mechanism for widespread DNA hypomethylation in cancer. *Cell Discovery*, 2, p.16007.

Jones, P.A., 2001. The Role of DNA Methylation in Mammalian Epigenetics. *Science*, 293(5532), pp.1068–1070.

Kim, J.M. et al., 2008. Cell cycle dependent chromatin loading of the Fanconi anemia core complex by FANCM / FAAP24. *Blood*, 111(10), pp.5215–5222.

Kim, Y. et al., 2011. Mutations of the SLX4 gene in Fanconi anemia. *Nature Genetics*, 43(2), pp.142–146.

Klein Douwel, D. et al., 2014. XPF-ERCC1 Acts in Unhooking DNA Interstrand Crosslinks in Cooperation with FANCD2 and FANCP/SLX4. *Molecular Cell*, 54(3), pp.460–471.

Knies, K. et al., 2017. Biallelic mutations in the ubiquitin ligase RFWD3 cause Fanconi anemia. *Journal of Clinical Investigation*, 2, pp.1–15.

Kriaucionis, S. & Heintz, N., 2009. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science (New York, N.Y.)*, 324(5929), pp.929–30.

Krokan, H.E. et al., 2000. Base excision repair of DNA in mammalian cells. *FEBS Lett.*, 476(1–2), pp.73–77.

Langevin, F. et al., 2011. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature*, 475(7354), pp.53–58.

Lee, Y.C. et al., 2014. The relationships between XPC binding to conformationally diverse DNA adducts and their excision by the human NER system: Is there a correlation? *DNA Repair*, 19, pp.55–63.

Liang, C.-C. et al., 2016. The FANCD2-FANCI complex is recruited to DNA interstrand crosslinks before monoubiquitination of FANCD2. *Nature communications*, 7, p.12124.

Liang, C.C. et al., 2015. UHRF1 Is a sensor for DNA interstrand crosslinks and recruits FANCD2 to initiate the Fanconi Anemia pathway. *Cell Reports*, 10(12), pp.1947–1957.

Lieber, M.R., 2008. The mechanism of human nonhomologous DNA End joining. *Journal of Biological Chemistry*, 283(1), pp.1–5.

Lindahl, T., 1993. Instability and decay of the primary structure of DNA. *Nature*, 362.

Litman, R. et al., 2005. BACH1 is critical for homologous recombination and appears to be the Fanconi anemia gene product FANCJ. *Cancer Cell*, 8(3), pp.255–265.

Liu, X. et al., 2013. UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. *Nature communications*, 4, p.1563.

Liu, Y. et al., 2017. UHRF2 regulates local 5-methylcytosine and suppresses spontaneous seizures. *Epigenetics*, 12(7), pp.551–560.

Lu, H. et al., 2016. Loss of UHRF2 expression is associated with human neoplasia, promoter hypermethylation, decreased 5-hydroxymethylcytosine, and high proliferative activity. *Oncotarget*, 7(46), pp.76047–76061.

Luger, K. et al., 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, 389(6648), pp.251–260.

Luo, T. et al., 2013. Uhrf2 is important for DNA damage response in vascular smooth muscle cells. *Biochemical and Biophysical Research Communications*, 441(1), pp.65–70.

Malinge, J.M., Giraud-Panis, M.J. & Leng, M., 1999. Interstrand cross-links of cisplatin induce striking distortions in DNA. *Journal of Inorganic Biochemistry*, 77(1–2), pp.23–29.

Mari, P.-O. et al., 2006. Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4. *Proceedings of the National Academy of Sciences*, 103(49), pp.18597–18602.

Mathew, C.G., 2006. Fanconi anaemia genes and susceptibility to cancer. *Oncogene*, 25(43), pp.5875–5884.

Meetei, A.R. et al., 2005. A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nature Genetics*, 37(9), pp.958–963.

Meetei, A.R. et al., 2003. A novel ubiquitin ligase is deficient in Fanconi anemia. *Nature Genetics*, 35(2), pp.165–170.

Meetei, A.R. et al., 2004. X-linked inheritance of Fanconi anemia complementation group B. *Nature genetics*, 36(11), pp.1219–1224.

Melcher, M. et al., 2000. Structure-Function Analysis of SUV39H1 Reveals a Dominant Role in Heterochromatin Organization, Chromosome Segregation, and Mitotic Progression. *Molecular and Cellular Biology*, 20(10), pp.3728–3741.

Mimitou, E.P. & Symington, L.S., 2009. DNA end resection: Many nucleases make light work. *DNA Repair*, 8(9), pp.983–995.

Miura, M. et al., 2001. Dynamic Changes in Subnuclear NP95 Location during the Cell Cycle and Its Spatial Relationship with DNA Replication Foci. *Experimental Cell Research*, 263(2), pp.202–208.

Mori, T. et al., 2002. NIRF, a novel RING finger protein, is involved in cell-cycle regulation. *Biochemical and Biophysical Research Communications*, 296(3), pp.530–536.

Mori, T. et al., 2012. NIRF/UHRF2 occupies a central position in the cell cycle network and allows coupling with the epigenetic landscape. *FEBS Letters*, 586(11), pp.1570–1583.

Mori, T. et al., 2011. NIRF constitutes a nodal point in the cell cycle network and is a candidate tumor suppressor. *Cell Cycle*, 10(19), pp.3284–3299.

Mori, T. et al., 2004. NIRF is a ubiquitin ligase that is capable of ubiquitinating PCNP, a PEST-containing nuclear protein. *FEBS Letters*, 557(1–3), pp.209–214.

Mousli, M. et al., 2003. ICBP90 belongs to a new family of proteins with an expression that is deregulated in cancer cells. *British journal of cancer*, 89(1), pp.120–127.

Nakatani, Y. & Ogryzko, V., 2003. Immunoaffinity Purification of Mammalian Protein Complexes. *Methods in Enzymology*, 370, pp.430–444.

Nick McElhinny, S.A. et al., 2000. Ku recruits the XRCC4-ligase IV complex to DNA ends. *Molecular and Cellular Biology*, 20(9), pp.2996–3003.

Nowak, S.J. & Corces, V.G., 2004. Phosphorylation of histone H3: A balancing act between chromosome condensation and transcriptional activation. *Trends in Genetics*, 20(4), pp.214–220.

Okano, M. et al., 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, 99(3), pp.247–257.

Pang, Q. & Andreassen, P.R., 2009. Fanconi Anemia Proteins and Endogenous Stresses. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, 668(1–2), pp.42–53.

Park, J. et al., 2016. Complementation of hypersensitivity to DNA interstrand crosslinking agents

demonstrates that *XRCC2* is a Fanconi anaemia gene. *Journal of medical genetics*, 53(10), pp.1–9.

Parsons, J.L. & Dianov, G.L., 2013. Co-ordination of base excision repair and genome stability. *DNA Repair*, 12(5), pp.326–333.

Paull, T.T. & Gellert, M., 1998. The 3′ to 5′ Exonuclease Activity of Mre11 Facilitates Repair of DNA Double-Strand Breaks. *Molecular Cell*, 1(7), pp.969–979.

Pichler, G. et al., 2011. Cooperative DNA and histone binding by Uhrf2 links the two major repressive epigenetic pathways. *Journal of Cellular Biochemistry*, 112(9), pp.2585–2593.

Powell, S.N. & Kachnic, L. a, 2003. Roles of BRCA1 and BRCA2 in homologous recombination, DNA replication fidelity and the cellular response to ionizing radiation. *Oncogene*, 22(37), pp.5784–91.

Rajakumara, E. et al., 2016. Mechanistic insights into the recognition of 5-methylcytosine oxidation derivatives by the SUVH5 SRA domain. *Scientific Reports*, 6(September 2015), p.20161.

Ran, F.A. et al., 2013. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11), pp.2281–2308.

Razin, A. & Riggs, A.D., 1980. DNA methylation and gene function. *Science (New York, N.Y.)*, 210(4470), pp.604–10.

Rea, S. et al., 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, 406(6796), pp.593–599.

Reid, S. et al., 2007. Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nature Genetics*, 39(2), pp.162–164.

Rickman, K.A. et al., 2015. Deficiency of UBE2T, the E2Ubiquitin Ligase Necessary for

FANCD2 and FANCI Ubiquitination, Causes FA-T Subtype of Fanconi Anemia. *Cell Reports*, 12(1), pp.35–41.

Riggs, A.., 1975. X inactivation, differentiation, and DNA methylation. *Cytogenetic and Genome Research*, 25, pp.9–25.

Rothbart, S.B. et al., 2012. Association of UHRF1 with H3K9 methylation directs the maintenance of DNA methylation. *Nature Structural & Molecular Biology*, 19(11), pp.1155–1160.

Sartori, A.A. et al., 2007. Human CtIP promotes DNA end resection. *Nature*, 450(7169), pp.509–514.

Sawyer, S.L. et al., 2015. Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype. *Cancer Discovery*, 5(2), pp.135–142.

Sharif, J. et al., 2007. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature*, 450(7171), pp.908–912.

Sims, A.E. et al., 2007. FANCI is a second monoubiquitinated member of the Fanconi anemia pathway. *Nature structural & molecular biology*, 14(6), pp.564–567.

Singh, T.R. et al., 2009. Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M. *Blood*, 114(1), pp.174–180.

Song, J. et al., 2012. Structure-Based Mechanistic Insights into DNMT1-Mediated Maintenance DNA Methylation. *Science*, 335(6069), pp.709–712.

Spivak, G., 2015. Nucleotide excision repair in humans. *DNA Repair*, 36(1–2), pp.13–24.

Spruijt, C.G. et al., 2013. Dynamic readers for 5-(Hydroxy)methylcytosine and its oxidized derivatives. *Cell*, 152(5), pp.1146–1159.

Stoepker, C. et al., 2011. SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. *Nature Genetics*, 43(2), pp.138–141.

Strahl, B.D. et al., 1999. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. *Proc.Natl.Acad.Sci.USA*, 96(26), p.14967.

Strahl, B.D. & Allis, C.D., 2000. The language of covalent histone modifications. *Nature*, 403(6765), pp.41–45.

Strathee, C.A., Duncan, A.M.V. & Buchwald, M., 1992. Evidence for at least four Fanconi anaemia genes including FACC on chromosome 9. *Nature Genetics*, 1, pp.246–250.

Tachibana, M. et al., 2001. SET Domain-containing Protein, G9a, is a Novel Lysine-preferring Mammalian Histone Methyltransferase with Hyperactivity and Specific Selectivity to Lysines 9 and 27 of Histone H3. *Journal of Biological Chemistry*, 276(27), pp.25309–25317.

Tahiliani, M. et al., 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science (New York, N.Y.)*, 324(5929), pp.930–5.

Takai, D. & Jones, P.A., 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proceedings of the National Academy of Sciences of the United States of America*, 99(6), pp.3740–5.

Tartof, K.D. et al., 1989. Towards an understanding of position effect variegation. *Developmental Genetics*, 10(3), pp.162–176.

Lo Ten Foe, J.R. et al., 1996. Expression cloning of a cDNA for the major Fanconi anaemia gene, FAA. *Nat Genet*, 14(3), pp.320–323.

Thomas, C.J. et al., 2014. Kinase-Mediated Changes in Nucleosome Conformation Trigger

Chromatin Decondensation via Poly(ADP-Ribosyl)ation. *Molecular Cell*, 53(5), pp.831–842.

Tian, Y. et al., 2015. UHRF1 contributes to DNA damage repair as a lesion recognition factor and nuclease scaffold. *Cell reports*, 10(12), pp.1957–1966.

Timmers, C. et al., 2001. Positional cloning of a novel Fanconi anemia gene, FANCD2. *Molecular Cell*, 7(2), pp.241–248.

Tomasz, M., 1995. Mitomycin C: small, fast and deadly (but very selective). *Chemistry and Biology*, 2(9), pp.575–579.

Tubbs, A. & Nussenzweig, A., 2017. Endogenous DNA Damage as a Source of Genomic Instability in Cancer. *Cell*, 168, pp.644–656.

Uckelmann, M. & Sixma, T.K., 2017. Histone ubiquitination in the DNA damage response. *DNA Repair*, 56(June), pp.92–101.

Uematsu, N. et al., 2007. Autophosphorylation of DNA-PKCS regulates its dynamics at DNA double-strand breaks. *Journal of Cell Biology*, 177(2), pp.219–229.

Vaz, F. et al., 2010. Mutation of the RAD51C gene in a Fanconi anemia–like disorder. *Nature Genetics*, 42(5), pp.406–409.

Venter, J. et al., 2001. The Sequence of the Human Genome. *Science*, 291(5507), p.1304.

Voulgaridou, G.-P. et al., 2011. DNA damage induced by endogenous aldehydes: Current state of knowledge. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 711(1–2), pp.13–27.

Wang, A.T. et al., 2015. A Dominant Mutation in Human RAD51 Reveals Its Function in DNA Interstrand Crosslink Repair Independent of Homologous Recombination. *Molecular Cell*, 59(3), pp.478–490.

Wang, F. et al., 2012. NIRF is frequently upregulated in colorectal cancer and its oncogenicity can be suppressed by let-7a microRNA. *Cancer Letters*, 314(2), pp.223–231.

Wang, W., 2007. Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. *Nature Reviews Genetics*, 8(10), pp.735–748.

de Winter, J.P. et al., 1998. The Fanconi anaemia group G gene FANCG is identical with XRCC9. *Nature genetics*, 20(3), pp.281–3.

de Winter, J.P. et al., 2000. The Fanconi anemia protein FANCF forms a nuclear complex with FANCA, FANCC and FANCG. *Human molecular genetics*, 9(18), pp.2665–2674.

De Winter, J.P. et al., 2000. Isolation of a cDNA representing the fanconi anemia complementation group E gene. *American Journal of Human Genetics*, 67(5), pp.1306–1308.

Xia, B. et al., 2007. Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. *Nature Genetics*, 39(2), pp.159–161.

Yamamoto, K.N. et al., 2011. Involvement of SLX4 in interstrand cross-link repair is regulated by the Fanconi anemia pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 108(16), pp.6492–6496.

Yan, Z. et al., 2010. A Histone-Fold Complex and FANCM Form a Conserved DNA-Remodeling Complex to Maintain Genome Stability. *Molecular Cell*, 37(6), pp.865–878.

Yokoyama, A. et al., 2004. Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Molecular and cellular biology*, 24(13), pp.5639–5649.

Zeng, S. et al., 2016. E3 ligase UHRF2 stabilizes the acetyltransferase TIP60 and regulates H3K9ac and H3K14ac via RING finger domain. *Protein and Cell*, pp.1–17.

Zhang, J. et al., 2015. DNA interstrand cross-link repair requires replication-fork convergence. *Nature Structural and Molecular Biology*, 22(3).

Zhang, J. et al., 2011. S phase-dependent interaction with DNMT1 dictates the role of UHRF1 but not UHRF2 in DNA methylation maintenance. *Cell research*, 21(12), pp.1723–39.

Zhang, T. et al., 2016. UHRF2 decreases H3K9ac expression by interacting with it through the PHD and SRA/YDG domain in HepG2 hepatocellular carcinoma cells. *International Journal of Molecular Medicine*, pp.126–134.

Zhou, T. et al., 2014. Structural Basis for Hydroxymethylcytosine Recognition by the SRA Domain of UHRF2. *Molecular Cell*, 54(5), pp.879–886.

Zhou, Y. et al., 2016. Regulation of the DNA Damage Response by DNA-PKcs Inhibitory Phosphorylation of ATM. *Molecular Cell*, 65(1), pp.91–104.

Zou, L. & Elledge, S.J., 2003. Sensing DNA Damage Through ATRIP Recognition of RPA-ssDNA Complexes. *Science*, 300(5625), pp.1542–1548.