Studies of the Fanconi Anemia Pathway



Chih-Chao Liang

Department of Biochemistry

and

Somerville College

Trinity 2016

Acknowledgement

Most important of all, I would like to thank my mom and sister. Without their unconditional love and support, I would not have become the person I am today, and I would not have been able to complete my degree. I hope I have made you and dad proud.

Secondly, I would like to thank my friends/brothers, Gianluca Veggiani, Can Buldun and Christopher Schoene. I was lucky enough to meet you in Oxford. My research projects became accomplishable because of the idea and reagents we exchanged. The trips and parties were fun because I was with you. My Oxford life is colourful because of the time we shared.

Thirdly, I would like to thank my supervisor and mentor, Dr Martin Cohn, for the opportunity to work on different projects. I would not be able to become the scientist I am today without the discussions and arguments we had. I am grateful for your immense support and guidance during my DPhil study. The ups and downs we went through have made our relationship unique and unforgettable.

I was fortunately enough to receive enormous amount of additional guidance and assistance in my DPhil study. Without their help, I would not have been able to achieve as much. The FANCD2/FANCI complex structural project was fruitful because of the close collaboration with Dr Catherine Venien-Bryan. My thesis committee, Prof David Sherratt and Prof Lynne Cox, have provided me with valuable scientific and career advice. Dr Eva Wegel, former Micron assistant manager, provided training and helpful discussion for operating the microscopes in the Micron facility. David Lopez Marinez from our group has worked closely with me on many experiments and publications.

Abstract

The human genome is constantly challenged by various DNA damages, where several DNA damage repair pathways are required to guard the integrity of our genome. One type of lesion is DNA interstrand crosslink (ICL), which prevents any cellular process requiring the separation of the two DNA strands, and could lead to missegregation of chromosomes and mitotic catastrophe. The Fanconi anemia DNA damage repair pathway (FA pathway) has been shown to be responsible for the repair of ICLs. To date, there are 21 FA proteins identified that participate in the repair process. Among them, the FANCD2/FANCI heterodimer plays an important role in orchestrating the recruitment of the downstream repair proteins. Upon the appearance of ICLs, the FANCD2/FANCI complex is monoubiquitinated, and localised to the sites of DNA lesion. While the molecular mechanism of how the FA pathway operates is emerging, there are still questions remaining, such as if there is a sensor for ICLs being able to activate the FA pathway, or the timing of the recruitment of the FANCD2/FANCI heterodimer to damaged chromatin. In this thesis, first, I will describe the identification and characterisation of a novel ICL sensor protein, UHRF1, and its role in the activation of the FA pathway. Second, I will describe the first cryo-EM structure of the human FANCD2/FANCI complex, which led to the discovery of the functionally important Tower domain in the C-terminal of FANCD2. In addition, our work reveals that the FANCD2/FANCI complex is recruited to damaged chromatin prior to its monoubiquitination.

Table of Contents

Studies of the Fanconi Anemia Pathway ..... **i**

Acknowledgement ..... **ii**

Abstract ..... **iii**

Abbreviations **1**

Chapter 1 Introduction **4**

1.1 Post-translational modifications in the DNA damage response network **4**

1.2 Fanconi anemia and the Fanconi anemia DNA damage repair pathway **6**

1.3 DNA Interstrand crosslink repair **10**

Origins of ICLs 10

Replication-dependent ICL repair 11

Initiation of the FA pathway 11

Monoubiquitination and recruitment of the FANCD2/FANCI complex 13

Recruitment of the downstream repair proteins 13

1.4 Cryo-electron microscopy **14**

Chapter 2 UHRF1 Is a Sensor for DNA Interstrand Crosslinks and Recruits FANCD2 to Initiate the Fanconi anemia Pathway **17**

2.1 Introduction **17**

2.2 Specific aims **18**

2.3 Results **19**

2.3.1 UHRF1 directly interacts with ICL through its SRA domain 19

2.3.2 UHRF1 is required for cellular survival towards Mitomycin C and Trioxsalen/UVA treatments..... 23

2.3.3 UHRF1 recruitment to sites of damage is rapid and required for proper FANCD2 localisation..... 26

2.3.4 UHRF1 is required for the monoubiquitination of FANCD2 in response to Mitomycin C and Trioxsalen/UVA treatments 27

2.3.5 UHRF1 interacts directly with FANCD2 through its SRA domain 30

2.3.6 TTD and RING domain are dispensable for the role of UHRF1 in ICL repair 32

2.4 Discussion 35

UHRF1 Is a Sensor for ICLs 35

UHRF1 is required for the recruitment and monoubiquitination of FANCD2 36

Chapter 3 Structural Study of the full-length human FANCD2/FANCI complex utilizing cryo-electron microscopy 38

3.1 Introduction..... 38

3.2 Specific aims 39

3.3 Results..... 40

3.3.1 Cryo-EM structure of the full-length and C-terminal deletion of FANCD2/FANCI complex..... 40

3.3.2 The Tower domain of FANCD2 is functionally indispensable *in vivo* 42

3.3.3 The Tower domain of FANCD2 is required for its DNA binding activity..... 47

3.3.4 FANCD2/FANCI is first recruited to DNA, then ubiquitinated by FANCL.... 49

3.3.5 Electrostatic interactions underlie the binding of the Tower domain of FANCD2 with the ICL-containing replication fork..... 53

3.4 Discussion 58

The discrepancy between cryo-EM and crystal structure of FANCD2/FANCI heterodimer 58

The DNA binding activity of FANCD2 and FANCI 59

Revised model of the FA pathway activation 60

Chapter 4 Discussion 62

Different DNA damages 64

Replication-independent ICL repair 64

Emerging role of the FA pathway in stalled replication fork protection 65

Chapter 5 Material and methods 67

Cell lines and antibodies 67

Plasmids and transfection 67

CRISPR-Cas9 gene editing 67

Preparation of interstrand crosslinked DNA substrates 68

ICL-binding protein purification and UHRF1/DNA *in vitro* binding 69

Mass spectrometric analysis 69

Protein purification 70

Electrophoretic mobility shift assay (EMSA) 71

Fluorescence anisotropy 71

Cell fractionation and whole-cell lysate 72

Clonogenic survival assay 72

Co-immunoprecipitation 72

In vitro protein binding assay 73

Live-cell imaging 74

EM and 3D image processing 74

Flexible fitting of crystal structure into electron microscopy map using molecular dynamics 75

In vitro ubiquitination assay 76

Chapter 6 References 77

Appendix 98

Statement of previous acceptance or concurrent submission of thesis for degree 98

Statement of authorship 98

Published data and figures 98

Abbreviations

A-T	Ataxia telangiectasia
ATM	ataxia telangiectasia mutated
ATR	ataxia telangiectasia and Rad3-related
AML	acute myeloid leukaemia
BER	base excision repair
BRCA1	breast cancer type 1 susceptibility protein
BRCA2	breast cancer type 2 susceptibility protein
CCD	charge-coupled device
Cdc45	Cell division control protein 45 homolog
CHK1	checkpoint kinase 1
CMG	Cdc45, Mcm2–7 and GINS
CPD	cyclobutane pyrimidine dimer
Cryo-EM	cryo-electron microscopy
DDR	DNA damage response
DSB	DNA double-strand break
DNA-PKcs	DNA-dependent protein kinase
DNMT1	DNA (cytosine-5)-methyltransferase 1
FA	Fanconi anemia
FA pathway	Fanconi anemia DNA damage repair pathway
FANC	Fanconi anemia complementation group
FAAP	Fanconi anemia-associated protein
FAN1	Fanconi-associated nuclease 1
FHA	forkhead-associated

HA	Human influenza hemagglutinin
HR	Homologous recombination
HU	hydroxyurea
ICL	interstrand crosslink
IR	ionizing irradiation
IP	immunoprecipitation
Kd	dissociation constant
MCM	Minichromosome maintenance protein
MDC1	mediator of DNA damage checkpoint protein 1
MMC	Mitomycin C
MMS	mismatch repair
MRN	MRE11/Rad50/NBS1
NER	nucleotide excision repair
NHEJ	non-homologous end joining
NLS	nuclear localization signal
NMR	nuclear magnetic resonance
PCNA	proliferating cell nuclear antigen
PTM	post-translational modification
PHD	plant homeodomain
Pol ζ	DNA polymerase ζ
RAP80	receptor-associated protein 80
RING	Really Interesting New Gene
RNF8	ring finger protein 8
RNF168	ring finger protein 168
shRNA	short hairpin RNA

SLX4	Structure-specific endonuclease subunit SLX4
SRA	Set and Ring Finger Associated
Tip60	Histone acetyltransferase KAT5
TLS	translesion synthesis
TMP	4,5',8-trimethylpsoralen
TTD	tandem tudor
UBL	ubiquitin-like
UHRF1	Ubiquitin-like, containing PHD and RING finger domains, 1
UVA	ultraviolet A
UVC	ultraviolet C
XP	Xeroderma Pigmentosum

Chapter 1 Introduction

The human genome is constantly challenged by endogenous and exogenous source of DNA damaging agents. These damaging agents could lead to base damages, single- and double-strand breaks or intrastrand and interstrand crosslinks (Roos et al., 2016). If the damages are left unrepaired, they could lead to lose of genetic information and dysregulation of cellular signalling. There are several DNA damage repair and tolerance pathways in place to guard the genome integrity, such as (i) base and nucleotide excision repair and mismatch repair (BER, NER and MMR respectively) for base damages and single-strand breaks, (ii) translesion synthesis (TLS) for damaged base by-passing, (iii) homologous recombination and non-homologous end joining (HR and NHEJ, respectively) for double-strand breaks (DSB) and (iv) the Fanconi anemia DNA damage repair pathway (FA pathway) for the DNA interstrand crosslinks (ICL). These DNA damage repair pathways collectively form a highly complex DNA damage response (DDR) network with cross talks between different repair pathways to tackle every damage occurring in the cells. Mutations in any of the DNA damage repair proteins could result in accumulation of DNA lesions, which often associates with diseases and cancers (Helleday et al., 2008). Breast cancer type 1 susceptibility protein (BRCA1) and breast cancer type 2 susceptibility protein (BRCA2) are best known for their roles in HR, and their association with breast and ovarian cancers (Nielsen et al., 2016), mutation in ataxia telangiectasia mutated (ATM), a protein kinase involved in various DDR events, is known to cause Ataxia telangiectasia (A-T) (Lavin and Shiloh, 1996; Li and Swift, 2000), and loss of the NER factors gives rise to Xeroderma pigmentosum (XP) (Feltes and Bonatto, 2015; Karass et al., 2015).

1.1 Post-translational modifications in the DNA damage response network

The DDR activity is tightly regulated by post-translational modifications (PTMs) to ensure timely activation and deactivation. In contrast to transcriptional regulation, which takes time to adjust the expression level of proteins, PTMs could impose immediate effects on protein activities by altering the charging of the protein or causing conformational changes, which could facilitate additional protein/protein interaction (Prabakaran et al., 2012).

Among all types of PTMs, the roles of phosphorylation and ubiquitination in the DDR have been studied extensively. Often, phosphorylation and ubiquitination collaborate during signal transduction. One of the best-studied events is the double-strand break response. When double-strand breaks occur in cells, ATM, ataxia telangiectasia and Rad3-related (ATR) and DNA-dependent protein kinase (DNA-PK) are activated and phosphorylate histone H2A variant X (H2AX) on serine 139 (Rogakou et al., 1999). Phosphorylated H2AX (γH2AX) recruits the mediator of DNA damage checkpoint protein 1 (MDC1) to the DNA double-strand break (Stucki et al., 2005), and ring finger protein 8 (RNF8), an E3 ubiquitin ligase, is subsequently recruited to the break sites through the interaction between its forkhead-associated domain (FHA) and phosphorylated MDC1 (Huen et al., 2007; Mailand et al., 2007). RNF8 polyubiquitinates histone H1 creating a platform for the recruitment of ring finger protein 168 (RNF168) (Doil et al., 2009; Panier et al., 2012; Thorslund et al., 2015), and RNF168 further polyubiquitinates H2A and H2AX on lysines 13-15 at the break site (Mattiroli et al., 2012). Polyubiquitinated H2A and H2AX mediate the recruitments of downstream repair proteins to the breaks, e.g. 53BP1 (Fradet-Turcotte et al., 2013; Wilson et al., 2016) and receptor-associated protein 80 (RAP80)/BRCA1 (Kim et al., 2007; Wang et al., 2007; Wu et al., 2009). Furthermore, 53BP1, RAP80 and BRCA1 are also regulated by phosphorylation in response to DNA damages (Matsuoka et al., 2007), which further demonstrates the complexity of the regulation in the DDR.

In addition to phosphorylation and ubiquitination, the role of acetylation in the DDR is also emerging. It has been shown that acetylation of H2AX by histone acetyltransferase KAT5 (KAT5/TIP60) is required for appropriate H2A polyubiquitination (Ikura et al., 2007). A proteomic study provides some evidence that there are additional acetylation events in the DDR (Elia et al., 2015).

1.2 Fanconi anemia and the Fanconi anemia DNA damage repair pathway

Fanconi anemia (FA) is a rare genetic disorder, which is predominantly autosomal recessive (Dong et al., 2015). FA was first described by the Swiss paediatrician Guido Fanconi in 1927. The worldwide prevalence is 1 to 5 per million (Rosenberg et al., 2011). The current median lifespan for a FA patient is 33 years, while some FA individuals live into their 30s, 40s and 50s. The FA patients display broad and heterogeneous clinical manifestations. At least 60% of the FA patients are born with at least one physical anomaly, including short stature, thumb and arm anomalies, kidney problems, skin discoloration, small head or eyes and intellectual developmental delay. The FA patients show high level of genome instability and spontaneous chromosome breakage. The majority of FA patients develop bone marrow failure (aplastic anemia) and are susceptible to develop cancer, especially acute myeloid leukaemia (AML), squamous cell carcinomas and head and neck cancer (Auerbach, 2009). The treatment with androgens and hematopoietic growth factors could only help alleviate the symptoms of bone marrow failure temporarily. Bone marrow transplantation can cure the haematopoietic aspects of FA, but the patients may still develop solid tumour later (Joenje and Patel, 2001).

One hallmark of FA patients is their hypersensitivities towards DNA crosslinking agents, such as mitomycin C (MMC) and cisplatin. To date, 21 FA genes have been identified, where mutations in any of these genes can lead to Fanconi anemia (Bluteau et al., 2016; Lopez-

Martinez et al., 2016; Park et al., 2016; Shamseldin et al., 2012). These 21 FA genes encode proteins cooperating in a complex DNA damage repair pathway, namely the FA DNA damage repair pathway, that coordinate the repair of ICLs in the cells. These proteins can be grouped into three categories, (i) the FANCD2/FANCI complex, (ii) the FA core complex and (iii) downstream repair proteins (Table1).

FANCD2 and FANCI form a heterodimer (Dorsman et al., 2007; Sims et al., 2007), which plays a key role in orchestrating the repair events in the ICL repair. Abrogating the recruitment

	Name	Synonym	Main functions	Identification references
FANCD2/FANCI complex	FANCD2		Monoubiquitinated by the FA core complex; recruits nucleases and downstream repair proteins	(Timmers et al., 2001)
	FANCI		Phosphorylation by ATR activates the FA pathway; Monoubiquitinated by the FA core complex	(Dorsman et al., 2007; Sims et al., 2007; Smogorzewska et al., 2007)
FA core complex	FANCA		Component of the FA core complex	(Lo Ten Foe et al., 1996)
	FANCB		Component of the FA core complex	(Meetei et al., 2004)
	FANCC		Component of the FA core complex	(Strathdee et al., 1992)
	FANCE		Component of the FA core complex	(de Winter et al., 2000a)
	FANCF		Component of the FA core complex	(de Winter et al., 2000b)
	FANCG		Component of the FA core complex	(de Winter et al., 1998)
	FANCL		Component of the FA core complex; E3 ubiquitin ligase	(Meetei et al., 2003)
	FANCM		Component of the FA core complex; DNA translocase	(Meetei et al., 2005)
	FANCT	UBE2T	E2 ubiquitin-conjugating enzyme	(Hira et al., 2015; Machida et al., 2006; Rickman et al., 2015; Virts et al., 2015)
Downstream repair proteins	FANCD1	BRCA2	HR repair; facilitates RAD51 loading onto DNA	(Howlett et al., 2002)
	FANCJ	BRIP1	HR repair; 3' to 5' helicase	(Bridge et al., 2005; Levitus et al., 2005; Levran et al., 2005; Litman et al., 2005)
	FANCN	PALB2	HR repair; promotes BRCA2 functions	(Xia et al., 2007)
	FANCO	RAD51C	HR repair; promotes RAD51 filament stability	(Vaz et al., 2010)
	FANCP	SLX4	Scaffold protein for nucleases; interacts with SLX1, MUS81/EME1 and XPF/ERCC1	(Kim et al., 2011; Stoepker et al., 2011)
	FANCQ	XPF	Endonuclease in complex with ERCC1; unhooks ICLs	(Bogliolo et al., 2013; Kashiyama et al., 2013)
	FANCR	RAD51	HR repair; facilitates strand invasion	(Ameziane et al., 2015; Wang et al., 2015)
	FANCS	BRCA1	HR repair; promotes MCMs unloading at ICLs	(Sawyer et al., 2015)
	FANCU	XRCC2	HR repair	(Park et al., 2016; Shamseldin et al., 2012)
	FANCV	REV7	Subunit of translesion synthesis polymerase ζ; interacts with the catalytic subunit REV3 and inserter polymerase REV1	(Bluteau et al., 2016)

Table 1. Fanconi anemia proteins.
21 identified FA proteins are grouped into three categories, the FANCD2/FANCI complex, the FA core complex and the downstream repair proteins. Synonym, main known functions and the reference for the first identification of each protein are shown.

and retention of the FANCD2/FANCI heterodimer at the damaged chromatin leads to inefficient repair of ICLs, hence, hypersensitivity towards crosslinking agents. While some studies suggested that FANCD2 and FANCI function independently in response to DNA damage or replication stress in the *Xenopus* cell-free system (Chaudhury et al., 2013; Sareen et al., 2012), it is unclear if the observations would translate in other organisms. The FANCD2/FANCI heterodimer is monoubiquitinated by the FA core complex on lysine 561 (K561) and lysine 523 (K523), respectively, in response to various DNA damages (Garcia-Higuera et al., 2001; Sims et al., 2007; Smogorzewska et al., 2007). Monoubiquitination of FANCD2 is required for the recruitment and retention of FANCD2 at sites of damage (Garcia-Higuera et al., 2001; Smogorzewska et al., 2007) and is essential for the recruitment of nucleases to the sites of damage (Klein Douwel et al., 2014; Murina et al., 2014; Unno et al., 2014; Yamamoto et al., 2011). In addition to ubiquitination, phosphorylation also mediates the activities of the FANCD2/FANCI complex. Phosphorylations of serine 222 (S222), serine 331(S331), threonine 691 (T691) and serine 717 (S717) of FANCD2 by ATM, ATR and checkpoint kinase 1 (CHK1) have been shown to regulate monoubiquitination of the FANCD2 and the intra-S phase check point activity (Ho et al., 2006; Taniguchi et al., 2002b; Zhi et al., 2009). Similarly, ATR-mediated phosphorylations of serine 558 (S558), serine 561 (S561), threonine 567 (T567), serine 598 (S598), serine 619 (S619) and serine 631 (S631) of FANCI facilitate the monoubiquitination of FANCD2 (Ishiai et al., 2008).

The FA core complex is composed of FANCA, FANCB, FANCC, FANCE, FANCF, FANCG and FANCL (Ceccaldi et al., 2016). FANCL, an E3 ubiquitin ligase, is the catalytic subunit of the FA core complex (Meetei et al., 2003) that cooperates with FANCT, an E2 ubiquitin-conjugating enzyme best known as UBE2T, to monoubiquitinate the FANCD2/FANCI complex (Machida et al., 2006). It has been reported that some eukaryotic organisms express only FANCL instead of the full FA core complex, while the FA pathway is

still functional (Sugahara et al., 2012; Zhang et al., 2009). However, in vertebrates, loss of some FA core complex subunits leads to destabilised FA core complex and diminished monoubiquitination of FANCD2 (Garcia-Higuera et al., 1999; Pace et al., 2002), suggesting that an intact FA core complex is required in vertebrates. Furthermore, FANCM and Fanconi anemia-associated proteins (FAAP), FAAP10, FAAP16, FAAP20, FAAP24 and FAAP100, interact with the FA core complex to promote the chromatin loading of the FA core complex (Ceccaldi et al., 2016). It has been shown that there are several subcomplexes in the FA core complex and associating proteins. The A-G-20 subcomplex containing FANCA, FANCG and FAAP20 has been implied to contribute to the chromatin loading of the FA core complex through the interaction between FAAP20 and RNF168-mediated polyubiquitinated histone H2A (Garcia-Higuera et al., 1999; Kruyt et al., 1999; Leung et al., 2012; Reuter et al., 2000; Waisfisz et al., 1999; Yan et al., 2012). The C-E-F subcomplex containing FANCC, FANCE and FANCF is thought to facilitate the chromatin recruitment and substrate binding by the interactions between FANCE and FANCD2, and FANCF and FANCM (Deans and West, 2009; Polito et al., 2014). The FANCM-24-10-16 module containing FANCM, FAAP24, FAAP10 and FAAP16 has been shown to be important for the recruitment of the FA core complex to sites of damage (Ciccia et al., 2007; Kim et al., 2008). Last but not least, the B-L-100 subcomplex containing FANCB, FANCL and FAAP100 is the catalytic unit, in which FANCB and FAAP100 stimulate the enzymatic activity of FANCL (Huang et al., 2014; Rajendra et al., 2014).

Among the downstream repair proteins group, BRCA1, BRCA2, PALB2, RAD51, RAD51C, BRIP1 and XRCC2 are best known for their roles in the HR repair. SLX4 has been described as a Swiss army knife for its role as a nuclease scaffold protein (Cybulski and Howlett, 2011). SLX4 interacts with SLX1, MUS81/EME1 and XPF/ERCC1 endonucleases, where XPF has been shown to be the nuclease that unhooks the ICL (Andersen et al., 2009; Hodskinson et al., 2014; Klein Douwel et al., 2014; Munoz et al., 2009; Svendsen et al., 2009). REV7 is a

subunit of DNA polymerase ζ (pol ζ), which interacts with the catalytic subunit REV3 of pol ζ and inserter polymerase REV1 to promote DNA damage tolerance. (Acharya et al., 2005; Hara et al., 2010; Tomida et al., 2015).

1.3 DNA Interstrand crosslink repair

Origins of ICLs

An ICL covalently connects the DNA double helix, which can potentially block any cellular process that requires separation of the two DNA strands, such as replication and transcription. If an ICL is left unrepaired, it could lead to mutations, chromosome breakage, chromosome missegregation and mitotic catastrophe. ICLs could arise from endogenous and exogenous sources (Lopez-Martinez et al., 2016). There is evidence that reactive aldehydes could be the source of endogenous crosslinking agents (Langevin et al., 2011), which could be metabolic products from alcohol and lipid peroxidation (Langevin et al., 2011; Pang and Andreassen, 2009). Exogenous sources of crosslinking agents, such as MMC, cisplatin and nitrogen mustard derivatives, have been widely used as anti-cancer chemotherapeutic drugs and in the ICL repair studies. In addition to the common feature of blocking the DNA strand-separation, different crosslinking agents also pose distinctive distortions of the DNA double helix (Deans and West, 2011). For example, psoralen/UVA-induced ICL introduces minor distortion (Figure 1-1B; PDB: 1A2E), while cisplatin imposes major distortion on the DNA double helix (Figure 1-1C; PDB: 204D). The differently distorted DNA double helix could potentially result in the ICLs being recognised by different ICL sensing proteins in addition to stalled replication fork to initiate the ICL repair.

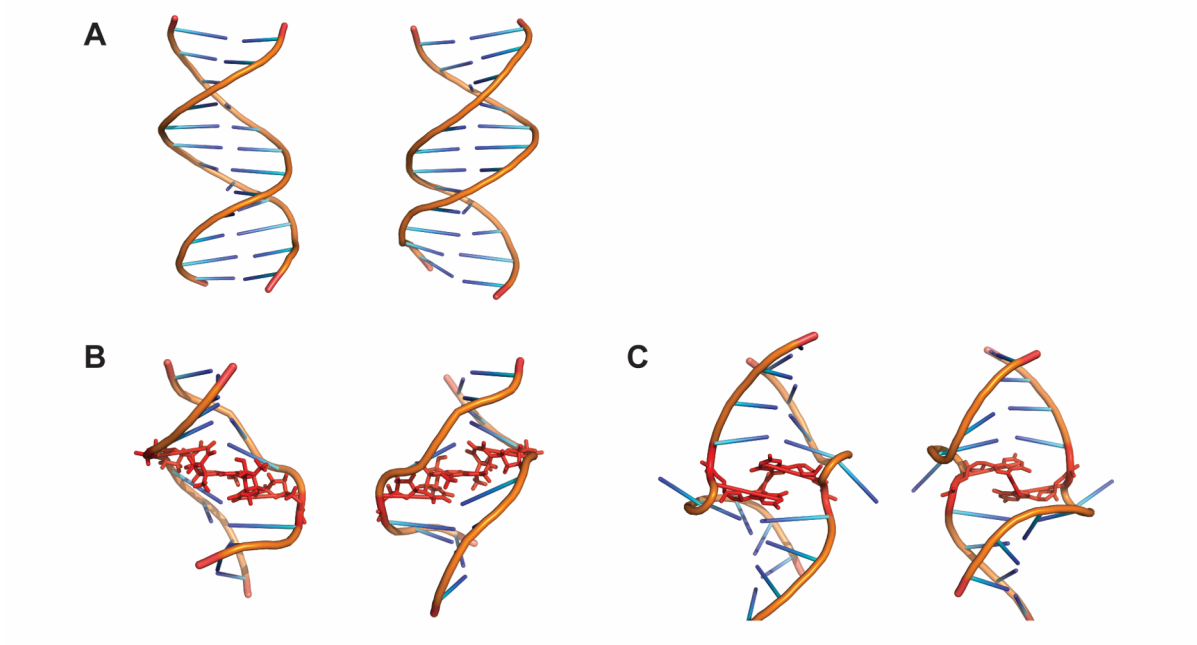


Figure 1-1. Structures of various ICLs.
A) B-DNA (PDB: 1-BNA). B) cisplatin-induced ICL (PDB: 1A2E). C) psoralen/UVA-induced ICL (PDB: 204D).

Replication-dependent ICL repair

The canonical FA-mediated ICL repair is believed to occur solely in S phase because of the observation of S phase specific FANCD2 foci co-localize with the downstream repair proteins, BRCA1, BRCA2 and RAD51 (Garcia-Higuera et al., 2001; Taniguchi et al., 2002a). Moreover, the evidence that replication is required for the ICL repair in the *Xenopus* cell free system further strengthen this mechanism (Raschle et al., 2008).

Initiation of the FA pathway

When a replication fork arrives at an ICL (Figure 1-2, step1), it can undergo FANCM mediated travers through the ICL, leaving the ICL to be repaired later (Figure 1-2, step 2) (Huang et al., 2013). The interaction between FANCM and proliferating cell nuclear antigen (PCNA) has been shown to be important for this process (Rohleder et al., 2016), although how the replication machinery migrates through the ICL is still elusive. Alternatively, in the *Xenopus* cell-free system, it has been shown that activation of the FA pathway requires the convergence of two replication forks at an ICL (Figure 1-2, step 3) (Zhang et al., 2015), and BRCA1

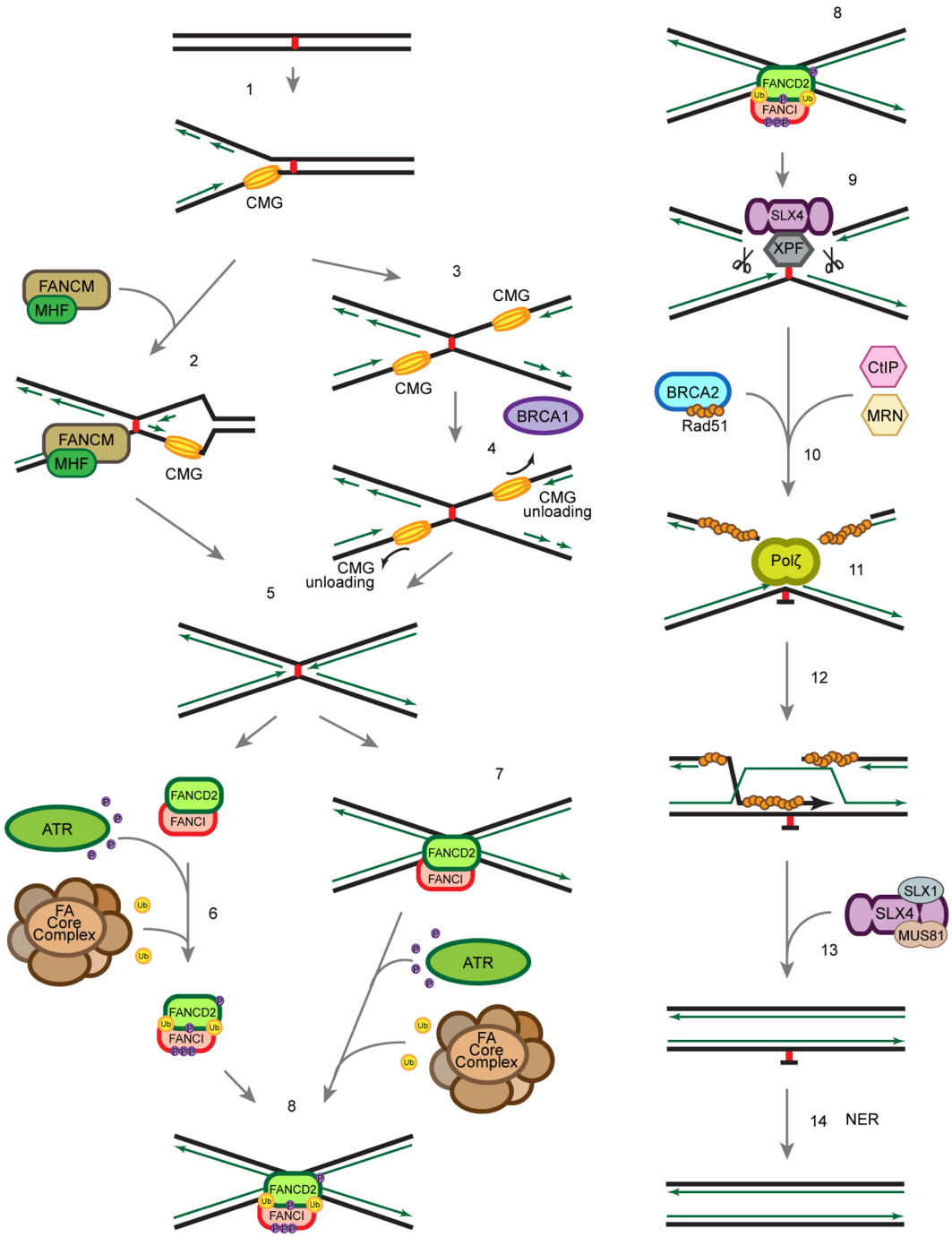


Figure 1-2. Schematic of ICL repair.
(1) Single replication fork arrives at the ICL. (2) FANCM/MHF complex mediates the traverse of the replication machinery through the ICL, which allows the replication fork to proceed, leaving the ICL for later repair. (3 and 4) Alternatively, BRCA1 (FANCS) facilitates the unloading of the CMG helicase complex when the second replication fork arrives at the ICL. (5) The replicative polymerase proceeds to the -1 position of the ICL, which leaves an X-shape structure similar to the traverse mechanism. (6) ATR phosphorylates the FANCD2/FANCI complex at multiple sites, whereas the FA core complex monoubiquitinates the FANCD2/FANCI complex at K561 and K523 respectively. (7) FANCD2/FANCI is recruited to the ICL at the replication fork. (8 and 9) Ubiquitinated FANCD2/FANCI complex recruits SLX4/XPF to ICL to unhook the ICL. (10) CtIP and the MRN complex resect the double-strand break ends generated by the incision in the previous step (step 5), and BRCA2 facilitates Rad51 filament formation on the ssDNA generated by the resection. (11) Polζ polymerizes new strand of DNA through the unhooked ICL. (12) Rad51 facilitates the strand invasion, which allows extension of the other strand. (13) SLX4 and nucleases resolve the double Holliday junction. (14) NER repair proteins remove the damaged nucleotide.

facilitates the unloading of the replicative helicases, Cdc45, Mcm2–7 and GINS, (CMG) (Figure 1-2, step 4) (Long et al., 2014). These two pathways leave a similar X-shaped DNA structure for repair (Figure 1-2, step 5).

Monoubiquitination and recruitment of the FANCD2/FANCI complex

The FANCM/FAAP20 complex has been shown to bind to a panel of DNA substrates, and together with the observation that FANCM and FAAP20 are required for overcoming the MMC hypersensitivity, a model was proposed, where the FANCM/FAAP20 complex recognises ICLs, and promotes recruitment of the FA core complex to sites of damage to monoubiquitinate the FANCD2/FANCI complex (Ciccia et al., 2007). Because loss of monoubiquitination of FANCD2 leads to impaired damaged FANCD2 foci (Garcia-Higuera et al., 2001), it is generally believed that the FANCD2/FANCI complex is monoubiquitinated before its recruitment to sites of damage (Figure 1-2, step 6). However, emerging evidence showing that DNA robustly stimulates the monoubiquitination of the FANCD2 *in vitro* (Longerich et al., 2014; Rajendra et al., 2014; Sato et al., 2012), hints that recruitment of the FANCD2/FANCI complex to the chromatin may precede monoubiquitination (Figure 1-2, step 7).

Recruitment of the downstream repair proteins

The monoubiquitinated FANCD2/FANCI complex recruits the XPF/ERCC1/SLX4 nuclease complex (Figure 1-2, step 8) and the nucleases carry out incisions at the ICL to unhook the crosslinked bases (Figure 1-2, step 9) (Klein Douwel et al., 2014; Kuraoka et al., 2000; Yamamoto et al., 2011). The TLS polymerase Rev1–polζ complex, which can accommodate bulky DNA substrates, is subsequently recruited to the lesion to carry out insertion through the unhooked ICL base pair (Figure 1-2, step 11) (Budzowska et al., 2015). Interestingly, the FA core complex, but not the FANCD2/FANCI complex, regulates recruitment of the Rev1–polζ complex (Budzowska et al., 2015). The double-strand break generated by the incision is resected

by CtIP, which is recruited through monoubiquitinated FANCD2 (Figure 1-2, step 10) (Murina et al., 2014; Unno et al., 2014), potentially together with the MRN complex (MRE11-RAD50-NBS1) and the nucleases EXO1 and DNA2. MRE11 has been shown to interact with FANCJ, which potentially regulates its nuclease activity ensuring a correct end resection (Suhasini et al., 2013). The DSB is, then, repaired by RAD51 mediated homologous recombination (Figure 1-2, step 12) (Long et al., 2011). To complete the repair of the ICL, the unhooked lesion needs to be removed (Figure 1-2, step 13). It is generally thought that the NER pathway is involved in the process (Figure 1-2, step 14) (Mouw and D'Andrea, 2014; Wood, 2010), though the precise molecular mechanism remains elusive.

1.4 Cryo-electron microscopy

Studying the structure of protein complexes often sheds light on the molecular mechanisms of their complex functions. X-ray crystallography is one of the popular methodologies for determining protein structure, which generally yields protein structures with near atomic resolution. However, truncation of the protein is often introduced to improve the crystallisation of target protein. While the resolution of protein structure determined by cryo-electron microscopy (cryo-EM) is not always comparable to X-ray crystallography, cryo-EM has the advantage of studying the structure of protein complex at its native state with the possibility of revealing different conformations within the same protein sample. Combining with atomic structures of subcomplexes or isolated components determined by X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy, we could generate crucial mechanistic models even with low resolution cryo-EM protein structure (Saibil, 2000). Moreover, the required protein amount is also relatively lower comparing to X-ray crystallography and NMR spectroscopy. In this thesis, the EM experiments and analysis were performed by our

collaborators, Dr. Catherine Vénien-Bryan and colleagues, Zhuolun Li and Dr. William Nicholson. Here, I described a simplified overview of the cryo-EM.

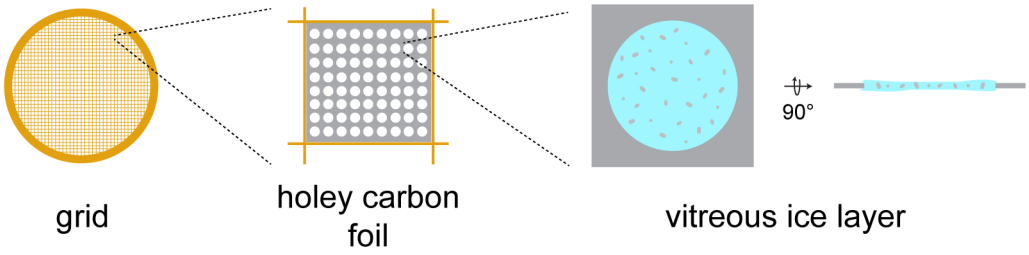


Figure 1-3. Schematic of cryo-EM grid with protein sample.
Protein sample is applied onto the metal (copper or gold) supported grid with a thin layer of holey carbon foil. The grid bearing a thin aqueous layer of protein sample is then plunged into liquid-nitrogen-cooled ethane to form a vitreous ice layer.

The purified protein sample is applied onto an EM grid covered with a thin holey carbon film as a thin liquid layer, and snap-frozen by plunging the grid into liquid-nitrogen-cooled ethane (Figure 1-3). The EM grid with a thin vitreous ice layer preserving the protein native structure is then inspected with a cryo-electron microscope, where the EM grid is irradiated with an electron beam, and the images are captured with a detector (Figure 1-4, step1) (Cheng, 2015). Traditionally, photographic film or charge-coupled device (CCD) camera are used as detectors. The CCD camera uses a thin layer of phosphor to convert electron signals to photons, which can then be detected by the camera sensor. However, this extra step of conversion adds extra noise to the cryo-EM images, lowering the resolution that can be reached after analysis. (Bai et al., 2015; Cheng, 2015).

The 2D micrographs of the protein sample are then analyzed to reconstitute the 3D model (Figure 1-4, step2). Each protein particle is selected, and their proximities to other particles are calculated to generate a 2D classification, where each class represents an orientation or a conformation of the protein. The data is then combined to generate a 3D model, and further refined with 3D classification (Figure 1-4, step3) (Nogales and Scheres, 2015).

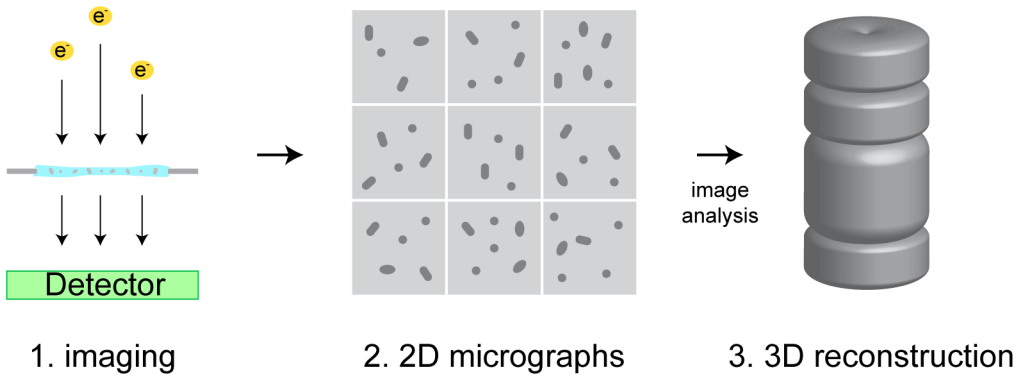


Figure 1-4. Schematic of cryo-EM imaging and analysis.
1) The cryo-EM grid bearing a thin layer of vitreous ice containing the protein sample is imaged on a cryo-electron microscope. 2) Particles are selected from 2D micrographs, and classified into different classes. If the sample is homogenous, each class should represent a projection of one orientation. The information of each orientation is then calculated to reconstructe the 3D model (3).

Chapter 2 UHRF1 Is a Sensor for DNA Interstrand Crosslinks and Recruits FANCD2 to Initiate the Fanconi anemia Pathway

2.1 Introduction

When DNA double-strand break occurs in the cells, MRE11/Rad50/NBS1 (MRN) complex and Ku70/Ku80 complex recognize and compete for the free DNA double-strand end (Khanna and Jackson, 2001; Uziel et al., 2003; Walker et al., 2001). This recognition event is accompanied by post-translational modifications, especially histone H2AX phosphorylation by ATM, which allows subsequent recruitment of MDC1, RNF8 and RNF168. RNF8 mediates ubiquitination of histone H1 and RNF168 mediates ubiquitination of histone H2A, creating a landing platform to recruit downstream repair factors (Doil et al., 2009; Huen et al., 2007; Mailand et al., 2007; Mattiroli et al., 2012; Thorslund et al., 2015).

However, it is unclear if there is an ICL sensor protein, which can recognise an ICL, and initiate the repair process in the cells. We hypothesize that there is a group of proteins that are capable of recognizing the distortion on the DNA double helix resulting from an ICL, and that such proteins may participate in the initiation process or involve directly in the repair process. In this chapter, we will show that we identified UHRF1 as a novel sensor for ICL, who also plays a role in the FA pathway activation.

UHRF1 is a multi-domain protein, containing an Ubiquitin-Like domain (UBL), a Tandem Tudor domain (TTD), a Plant Homeodomain (PHD), a Set and Ring Finger Associated domain (SRA) and a Really Interesting New Gene E3 ligase domain (RING). It is best known for its role in maintaining the DNA methylation in mammalian cells (Bostick et al., 2007; Sharif et al., 2007). It has been shown that UHRF1 recognizes newly replicated hemi-methylated DNA

through its SRA domain, and recruits DNMT1 to methylate the unmethylated DNA strand (Arita et al., 2008; Avvakumov et al., 2008; Hashimoto et al., 2008; Qian et al., 2008). In addition to the hemi-methylated DNA recognition, the SRA domain interacts directly with DNMT1 (Achour et al., 2008; Berkyurek et al., 2014; Bostick et al., 2007; Felle et al., 2011). The TTD and PHD domains of UHRF1 bind to tri-methylated lysine 9 containing histone H3 tail (H3K9me3) (Arita et al., 2012; Cheng et al., 2013; Xie et al., 2012), and its RING domain ubiquitinates histone H3 lysine 23 to facilitate the recruitment of DNMT1 to the DNA replication sites (Nishiyama et al., 2013).

There was no implication that UHRF1 is involved in the ICL repair in the literature. Here, we provide evidence that UHRF1 recognizes an ICL *in vitro* and *in vivo*, and it takes part in the FA pathway activation.

2.2 Specific aims

1. Identify novel ICL sensors.
2. Characterize the interaction between UHRF1 and ICL.
3. Understand the interplay between UHRF1 and the FA pathway.

2.3 Results

2.3.1 UHRF1 directly interacts with ICL through its SRA domain

In order to test our ICL sensor hypothesis, we established a biochemical purification scheme (Figure 2-1A). We designed a 5'-biotinylated double-stranded DNA molecule containing a unique central TA sequence (Figure 2-1B). A well-defined ICL can be generated by incubating the DNA molecule with the psoralen derivative 4,5',8-trimethylpsoralen (TMP), which intercalates specifically at the TA sequence, and can be photoactivated by irradiating with ultraviolet A light (UVA), generating an ICL. The ICL formation was then verified by running the double-stranded DNA onto a denaturing polyacrylamide gel (ICL-XL, Figure 2-1C). We coupled streptavidin agarose resin with biotinylated double-stranded DNA and biotinylated double-stranded DNA containing an ICL, and purified proteins specifically interacting with the two DNA molecules from Mitomycin C (MMC) treated HeLa nuclear extract (Figure 2-1D). Many of the polypeptides appeared to be present equally in both samples. However, one polypeptide was more abundant in the ICL containing DNA molecule sample. We excised the corresponding band, and identified 76 peptides representing UHRF1 (Figure 2-1E).

UHRF1, also known as NP95 and RNF106, contains 5 known domains (Figure 2-2A). UHRF1 is best known for its role in recognising hemi-methylated cytosine at newly replicated DNA, and recruiting DNMT1 to the sites to maintain DNA methylation marks in mammalian cells (Bostick et al., 2007; Sharif et al., 2007). It was shown that UHRF1 may be involved in DNA damage response as it was observed to be recruited to laser-induced DNA double-strand break (Tien et al., 2011). However, the mechanism was unclear, and little is known about the involvement of UHRF1 in other types of DNA damage response.

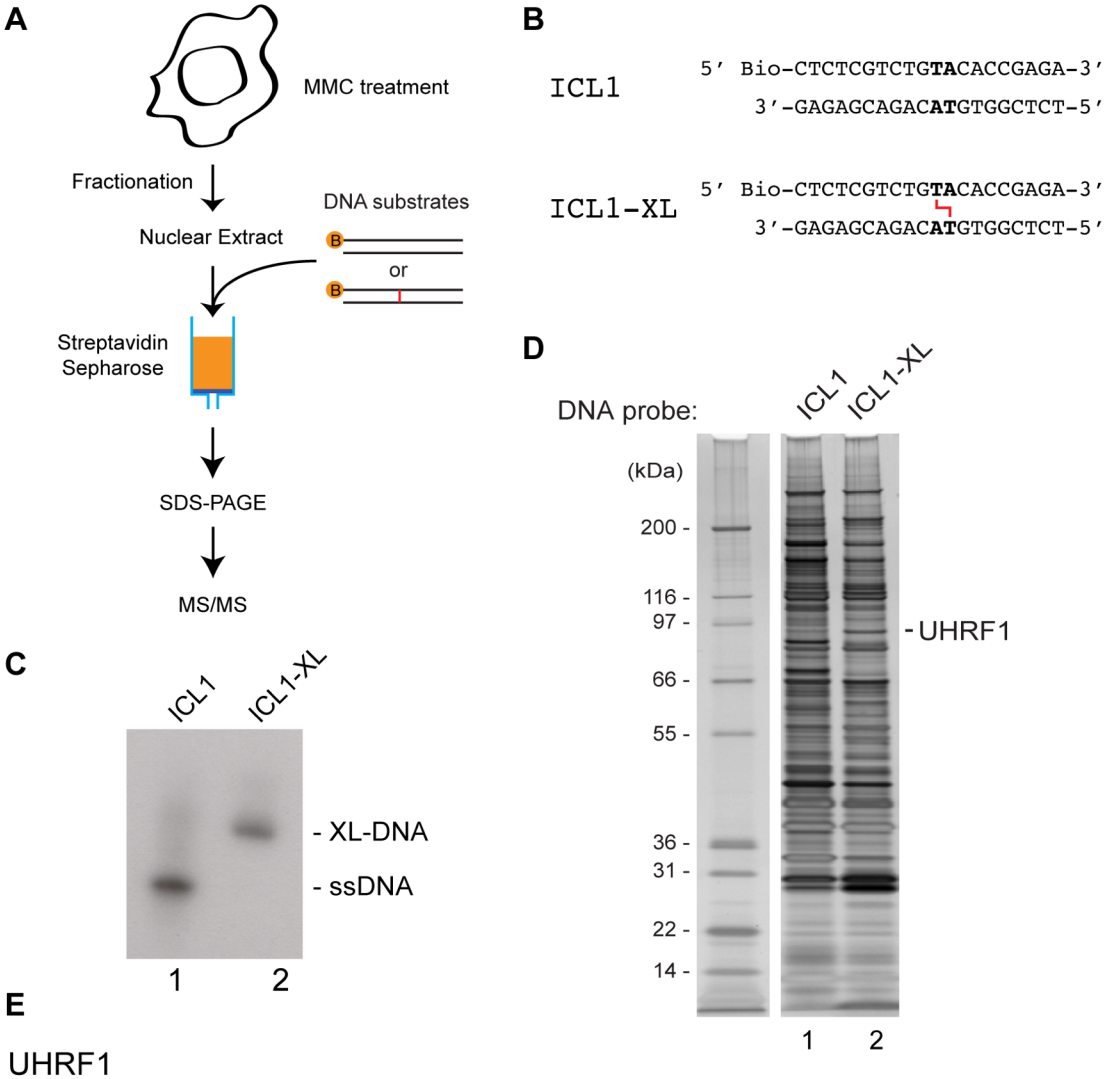


Figure 2-1. Purification of UHRF1.
A) ICL-interacting proteins purification scheme. Biotinylated DNA substrate is incubated with HeLa nuclear extract, and captured by Streptavidin-sepharose beads. The bound proteins are eluted, separated by SDS-PAGE and analyzed by silver stain. B) Schematic of the biotinylated ICLs used in the purification. C) Analysis of the cross-linked ICL1 on an 8M urea 20% polyacrylamide gel. D) UHRF1 was purified from HeLa nuclear extract and stained by silver stain. The polypeptide identified by mass spectrometry is indicated. E) List of identified peptides matching UHRF1. There were 76 peptides (in red) derived from UHRF1. Bao Zhan performed experiments shown in Figure 2-1 A-D, and Wilhelm Haas performed the mass spectrometry in Figure 2-1 E.

In order to test if UHRF1 interacts with ICL directly, we purified recombinant UHRF1 from *Sf9* insect cells (Figure 2-2B), and tested its DNA binding activities with several *in vitro* assays. We designed another set of DNA molecules, ICL8, which contains a unique TA site for introducing site specific ICL with TMP/UVA, and CPG3, which contains identical DNA sequence except for the TA/CG substitution in order to introduce a methylated cytosine (Figure 2-2C). With electrophoretic mobility shift assay (EMSA), we showed that our recombinant UHRF1 preferably binds to hemi-methylated CPG3 (Figure 2-2D, comparing lanes 1 and 2) as

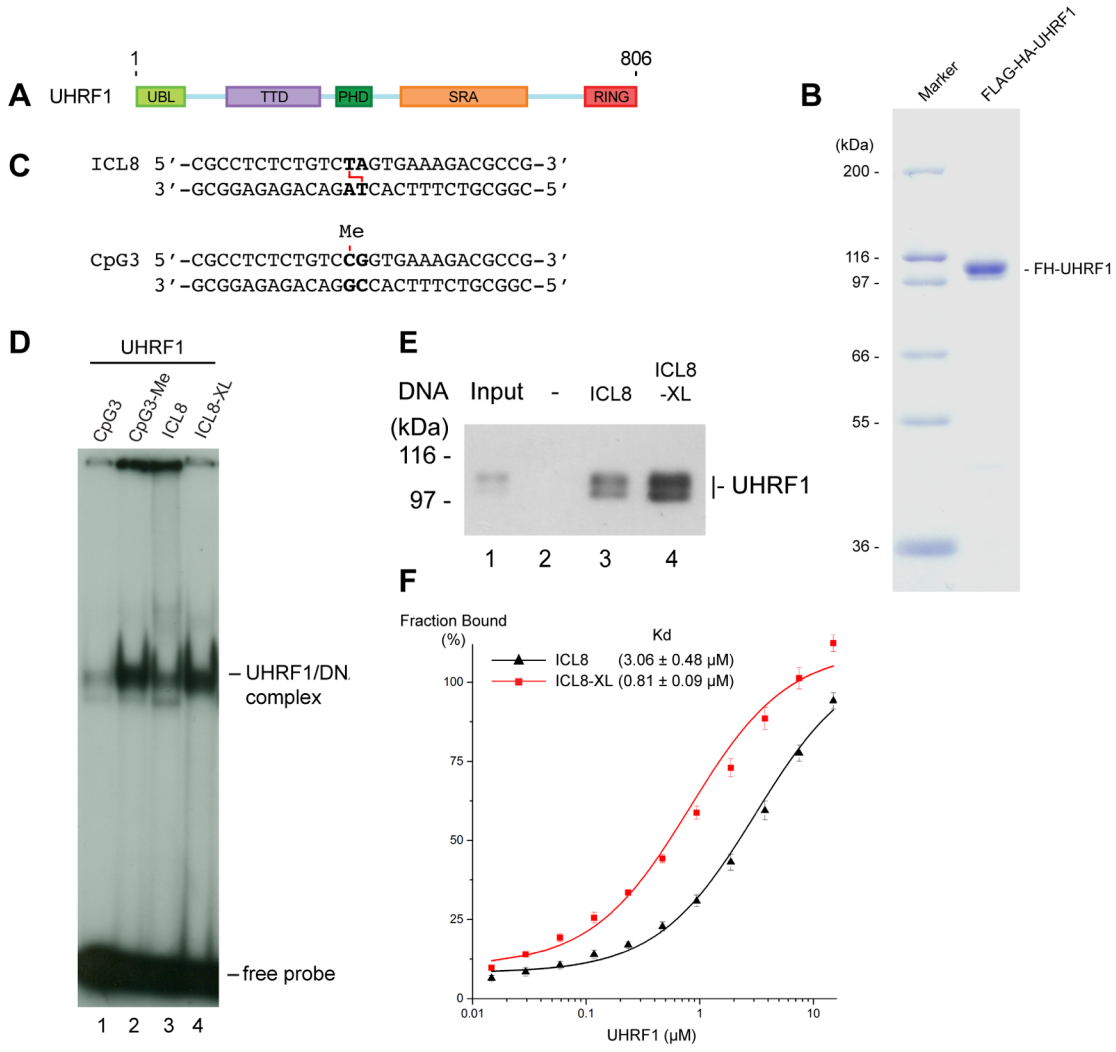


Figure 2-2. UHRF1 interacts directly with DNA interstrand cross-links.
A) Schematic of the UBL, TTD, PHD, SRA and RING domain in UHRF1 protein. B) Coomassie blue stain of recombinant FLAG-HA-tagged UHRF1 purified from Sf9 cells. C) Schematic of the ICL8 and CPG3 used in the *in vitro* DNA binding assay in (D, E and F). D) EMSA showing that UHRF1 binds better to cross-linked ICL8-XL and hemimethylated CPG3-Me DNA substrates than towards unmodified DNA. E) *In vitro* DNA binding assay showing that recombinant UHRF1 binds better to cross-linked ICL8-XL DNA than to normal ICL8 DNA. F) Fluorescence anisotropy assay determining the binding characteristics of UHRF1 to either ICL8 or ICL8-XL. Normalized and averaged anisotropy ± SEM.

shown in the literature (Arita et al., 2008; Avvakumov et al., 2008; Hashimoto et al., 2008; Qian et al., 2008) confirming that our recombinant UHRF1 is functionally active. Furthermore, UHRF1 showed higher affinity to crosslinked ICL8 (ICL8-XL) than its un-crosslinked counterpart (ICL8) (Figure 2-2D, comparing lanes 3 and 4), to a similar extent to the binding of hemi-methylated CPG3. To exclude the possibility that we are observing an artefact from EMSA, we tested the DNA binding activity of UHRF1 with a different approach. We incubated recombinant UHRF1 with 5'-biotinylated ICL8(-XL), and purified the protein/DNA complex with streptavidin agarose resin (Figure 2-2E). The pull-down type of assay showed again that UHRF1 directly interacts with ICL8, and it has higher affinity towards ICL8-XL. We determined the dissociation constant (Kd) of UHRF1 to ICL8(-XL) with fluorescence anisotropy using 5'-Alexa 488 labelled ICL8(-XL) (Figure 2-2F), which shows that the affinity of UHRF1 to ICL8-XL (0.81±0.09μM) is 3.7 fold higher than ICL8 (3.06±0.48μM).

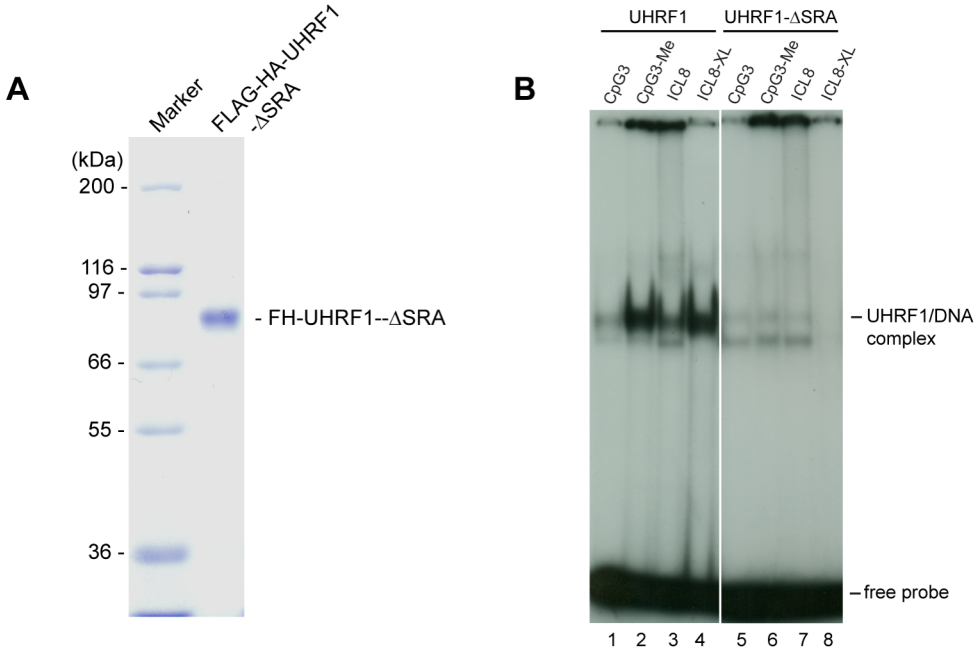


Figure 2-3. UHRF1 interacts directly with DNA interstrand cross-links through its SRA domain.
A) Coomassie blue stain of recombinant FLAG-HA-tagged UHRF1-ΔSRA purified from Sf9 cells. B) EMSA using recombinant UHRF1 lacking the SRA domain shows that the SRA domain of UHRF1 is responsible for the interaction with the ICL.

Previously, it was shown that the SRA domain is required for the interaction between UHRF1 and hemi-methylated DNA (Arita et al., 2008; Avvakumov et al., 2008; Hashimoto et al., 2008; Qian et al., 2008). Therefore, we hypothesised that the interaction between UHRF1 and crosslinked DNA may have similar mechanism. To test this hypothesis, we purified recombinant UHRF1 lacking the entire SRA domain (UHRF1-ΔSRA) (Figure 2-3A), and tested its DNA binding with EMSA (Figure 2-3B). As expected, deletion of the SRA domain prevents the interaction between UHRF1 and DNA. Taken together, these data showed that UHRF1 directly interacts with an ICL through its SRA domain.

2.3.2 UHRF1 is required for cellular survival towards Mitomycin C and Trioxsalen/UVA treatments

Since we previously demonstrated that UHRF1 directly interacts with crosslinked DNA *in vitro*, we set out to test if such interaction between UHRF1 and ICL we discovered has any biological significance in the cells. We down-regulated the expression level of endogenous UHRF1 to less than 5% by introducing a short hairpin RNA (shRNA) (Figure 2-4A). We tested if depletion of UHRF1 from the cells led to hypersensitivity towards different crosslinking agents, such as TMP/UVA, MMC and Cisplatin, with a clonogenic survival assay. Depletion of UHRF1 in the cells had significant impact on the cellular survival towards TMP/UVA and MMC treatment (Figure 2-4 B and C). In contrast, UHRF1 depleted cells showed little sensitivity towards Cisplatin (Figure 2-4D). These DNA crosslinking agents are generally considered posing similar genotoxic stress. However, the distortion on the DNA double helix created by these crosslinking agents are different. TMP/UVA and MMC-induced ICLs contain minor distortion, while Cisplatin-induced ICLs contain major distortion on the DNA double helix (Figure 1) (Deans and West, 2011; Lopez-Martinez et al., 2016).

In addition, we tested if UHRF1 is important for the repair for other types of DNA damages, such as double-strand breaks caused by ionized irradiation (IR) and cyclobutane pyrimidine dimer (CPD) caused by ultraviolet C (UVC). UHRF1 depleted cells showed little or no sensitivity toward IR (Figure 2-4E) and UVC (Figure 2-4F). To conclude, our survival assay results showed that depletion of UHRF1 sensitised the cells to TMP/UVA and MMC, but not to

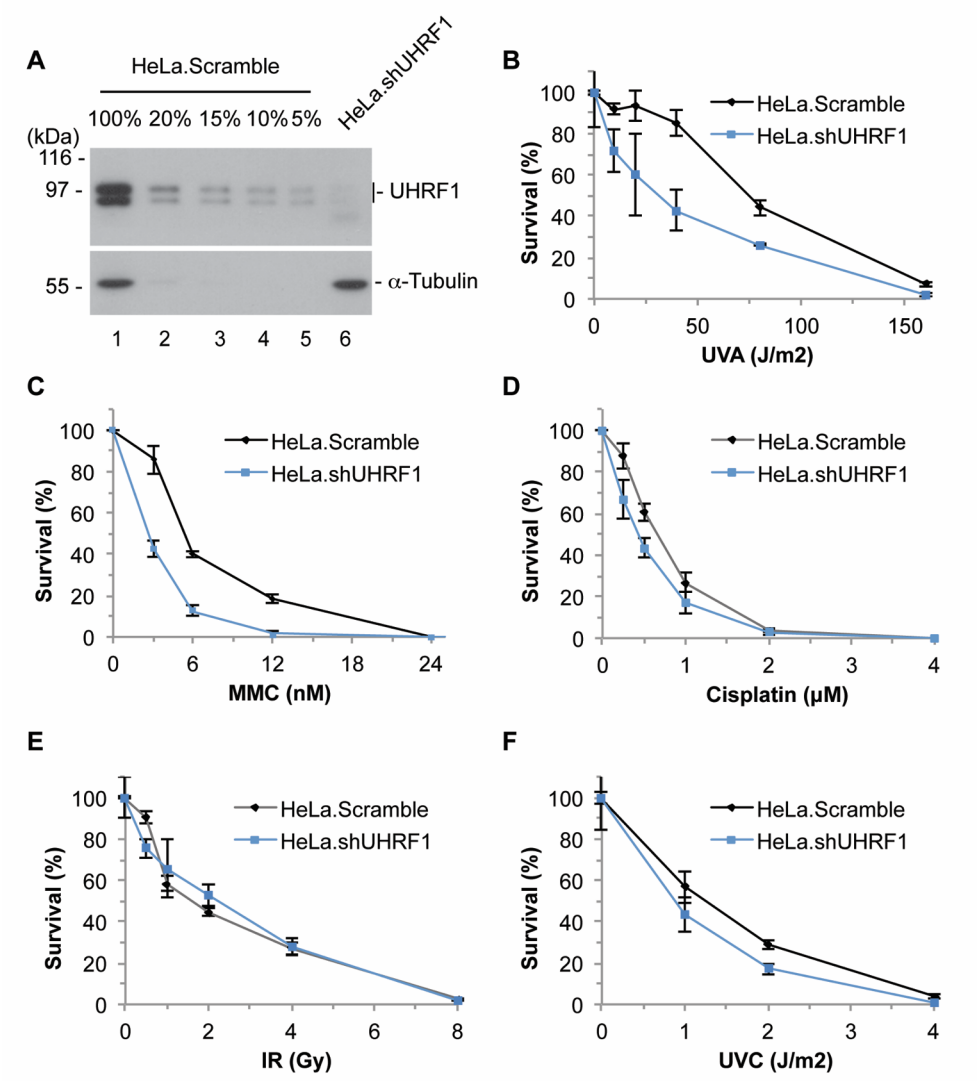


Figure 2-4. Knockdown of UHRF1 sensitizes cells to TMP/UVA and MMC.
A) Quantitative Western blot analysis comparing the serial dilution of the lysate of HeLa cells expressing non-targeting shRNA (HeLa.Scramble) and HeLa cells expressing shRNA targeting UHRF1 (HeLa.shUHRF1) confirms the efficiency of UHRF1 knockdown is around 95%. B) Clonogenic assay of HeLa.Scramble and HeLa.shUHRF1 cells shows that UHRF1 is required for cellular survival to TMP/UVA treatment. C) Clonogenic assay of HeLa.Scramble and HeLa.shUHRF1 cells shows that UHRF1 is required for cellular survival to MMC treatment. D) Clonogenic assay of HeLa.Scramble and HeLa.shUHRF1 cells shows UHRF1 is partly required for cellular survival to cisplatin treatment. E) Clonogenic assay of HeLa.Scramble and HeLa.shUHRF1 cells shows that UHRF1 is not required for cellular survival to IR treatment. F) Clonogenic assay of HeLa.Scramble and HeLa.shUHRF1 cells shows that UHRF1 is partly required for cellular survival to UVC treatment. For each survival assay, experiments were carried out at least two times, and each repeat was done in triplicate. Representative triplicate data is shown. Error bars in (B) - (F) based on individual triplicate show standard deviation.

other types of DNA damaging agents, which suggest that UHRF1 is functionally important for repairing TMP/UVA and MMC-induced ICLs in the cells, but not for cisplatin induced-ICL or other types of DNA lesion.

Intrigued by the results of the survival assays, we investigated further if UHRF1 collaborates with the FA pathway to accomplish ICL repair. We depleted UHRF1 and FANCD2 individually and in combination from HeLa cells using shRNA (Figure 2-5A), and tested their sensitivity towards MMC with clonogenic assay (Figure 2-5B). While it is consistent with previous results that UHRF1 depletion led to a decreased survival upon MMC treatment, FANCD2 knocked down cells were more sensitive towards MMC. Knocking down UHRF1 together with FANCD2 did not further sensitize the cells to MMC treatment, indicating that UHRF1 and FANCD2 are epistatic in the FA pathway. Interestingly, the double knock down led to a slight increase in the

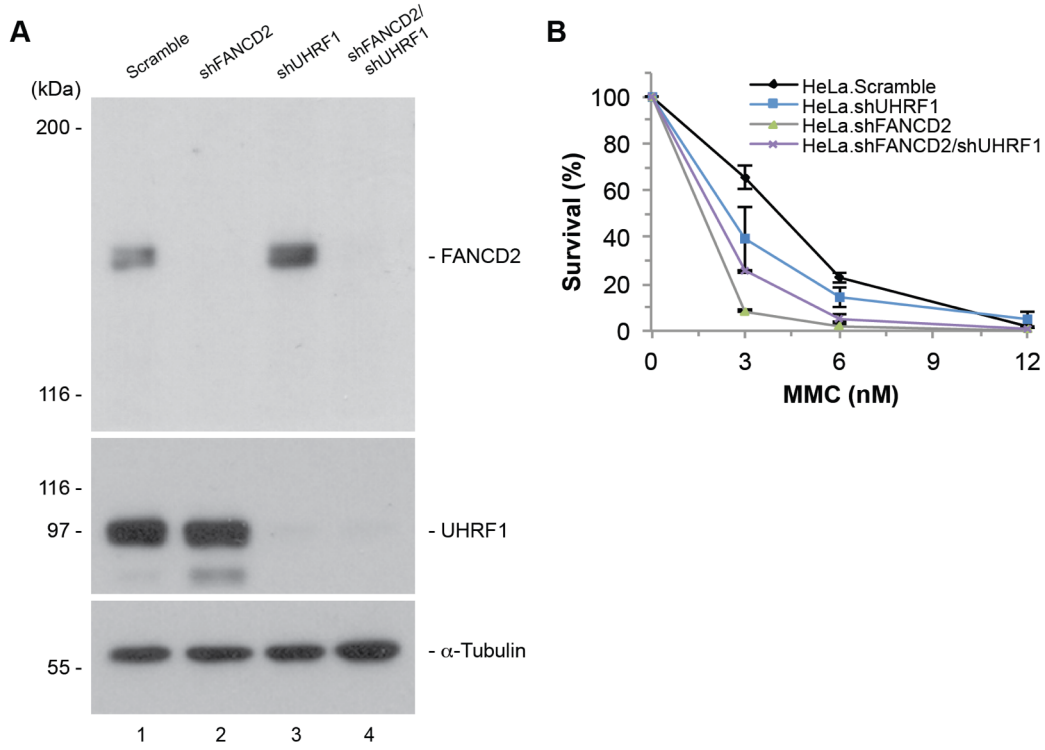


Figure 2-5. UHRF1 and FANCD2 are epistatic in the FA pathway.
A) Clonogenic assay of HeLa.Scramble, HeLa.shFANCD2, HeLa.shUHRF1 and HeLa.shFANCD2/shUHRF1 shows that UHRF1 and FANCD2 are epistatic. B) Western blot analysis confirming the efficiency of UHRF1 and FANCD2 knocked down in HeLa cells. For the survival assay, experiment was carried out two times, and each repeat was done in triplicate. Representative triplicate data is shown. Error bars in (B) based on individual triplicate show standard deviation.

survival assay. We could not rule out if it was due to some marginal difference in the UHRF1 or FANCD2 knock down efficiency, or if it was because abnormal repair initiated by UHRF1 in the absence of FANCD2.

2.3.3 UHRF1 recruitment to sites of damage is rapid and required for proper FANCD2 localisation

We have shown that UHRF1 interacts with ICLs *in vitro*, and it is functionally important for cellular survival toward crosslinking agents, such as TMP/UVA and MMC. To examine if UHRF1 is indeed recruited to the ICLs, we established a live-cell imaging protocol. By using a spinning disc confocal microscope, where we were able to introduce ICLs at desired area with TMP pre-treatment and laser activation, and observe the distribution of fluorophore-tagged UHRF1 and FANCD2 in HeLa cells over time. UHRF1 was recruited to ICLs within a minute (Figure 2-6A, top panel), suggesting UHRF1 could be one of the initial proteins recruited to the ICLs. The recruitment of FANCD2 to ICLs was detectable only after 10 - 15 minutes (Figure 2-6A, bottom panel). Little or no UHRF1 and FANCD2 recruitment was observed in the cells if they were not sensitized with TMP prior to laser irradiation (Figure 2-6B), which confirms the recruitment we observed was due to the presence of ICLs.

Given that the recruitment of UHRF1 to ICLs precedes FANCD2, we hypothesized that UHRF1 may be one of the players involved in the recruitment of FANCD2. We used the same shRNA described previously to reduce endogenous UHRF1 expression level, and observed the recruitment of EGFP-FANCD2 to ICLs. The rapid recruitment of FANCD2 to ICLs was no longer observed, if UHRF1 was not present (Figure 2-6C), which suggests the recruitment of UHRF1 to the ICLs is required for the subsequent rapid recruitment of FANCD2.

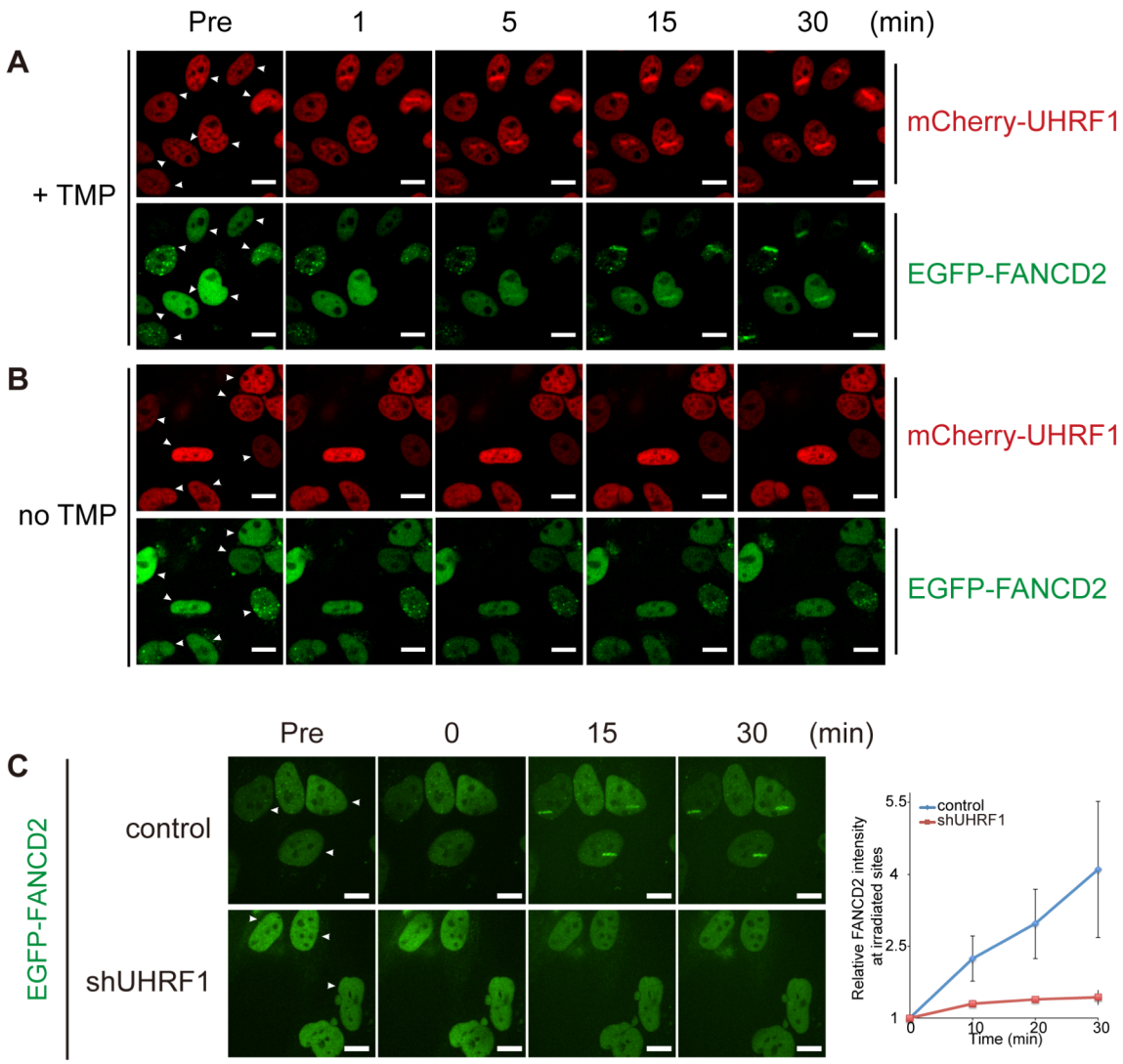


Figure 2-6. UHRF1 is rapidly recruited to ICLs *in vivo* and precedes the recruitment of FANCD2.
A-B) HeLa cells expressing mCherry-tagged UHRF1 and EGFP-tagged FANCD2 were (A) pre-treated with TMP, or (B) untreated, and microirradiated at the indicated area (white arrows). UHRF1 and FANCD2 were recruited to TMP induced ICLs sites (A), but not in the untreated cells (B). Scale bar: 10μm. C) HeLa cells expressing EGFP-tagged FANCD2 with or without UHRF1 knockdown were microirradiated at the indicated areas (white arrows). Depletion of UHRF1 abrogates the rapid accumulation of FANCD2 at the ICLs. Scale bar: 10μm. Error bars show standard deviation.

2.3.4 UHRF1 is required for the monoubiquitination of FANCD2 in response to Mitomycin C and Trioxsalen/UVA treatments

So far, we have shown that UHRF1 interacts with ICL *in vitro* through its SRA domain, and UHRF1 is functionally important for ICL repair in the cells, as its recruitment to ICLs precedes and is required for the recruitment of FANCD2. As mentioned previously in the introduction, monoubiquitination of FANCD2 is a key event determining the activation of the FA pathway.

To understand if UHRF1 is also involved in this step, we monitored the monoubiquitination of FANCD2 over time in control HeLa cells and UHRF1 depleted HeLa cells treated with various DNA damaging agents. In good agreement with the results from the survival assays, we observed a significant reduction of the amount of monoubiquitinated FANCD2 upon TMP/UVA and MMC treatment in the UHRF1 knocked down cell (Figure 2-7A and B), while cisplatin treated set only showed minimum difference in the FANCD2 monoubiquitination (Figure 2-7C). Similarly, UHRF1 depletion had little impact on the FANCD2 monoubiquitination upon UVC

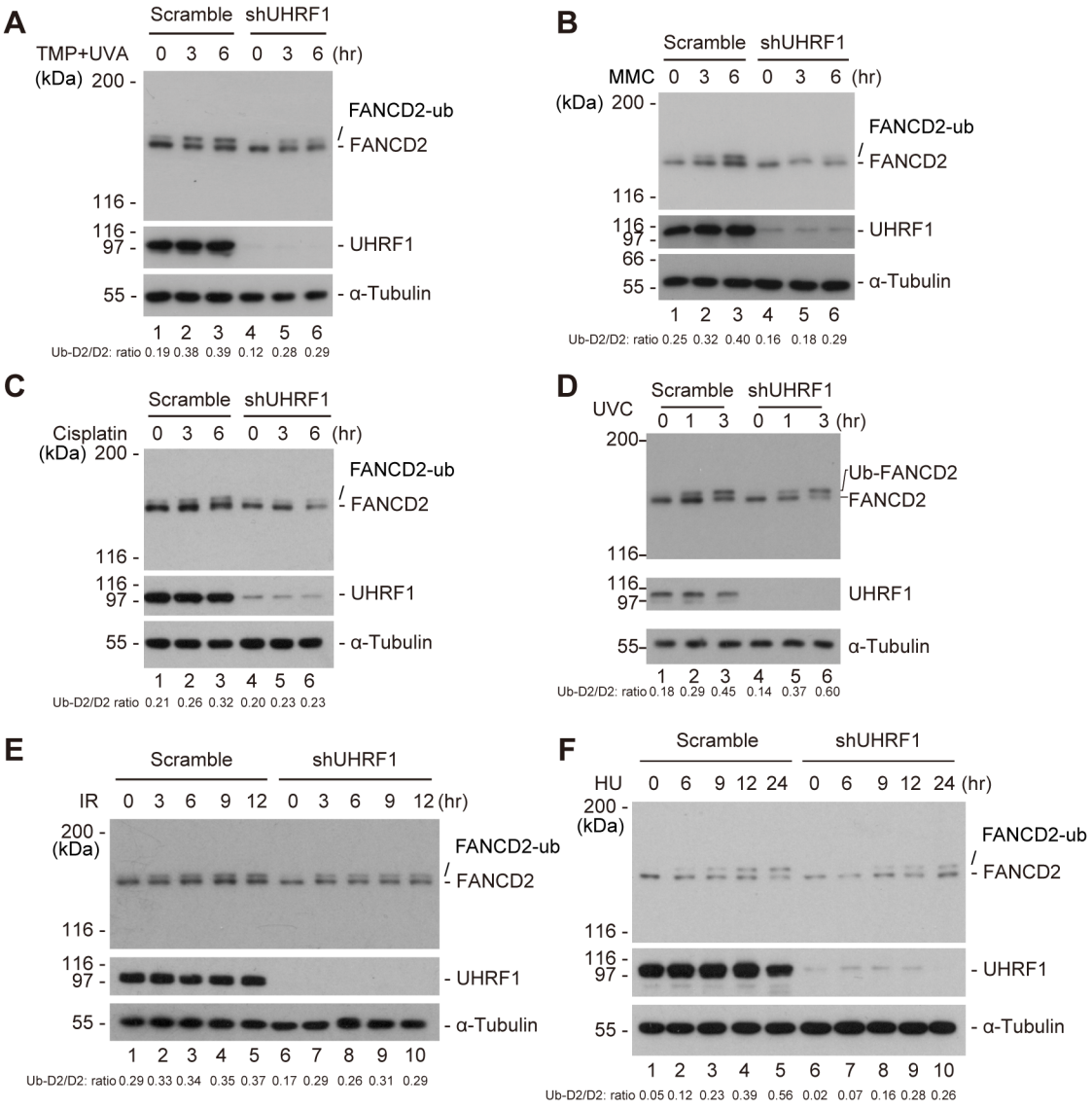


Figure 2-7. UHRF1 is required for normal FANCD2 ubiquitination *in vivo*.
A-F) FANCD2 monoubiquitination response in HeLa cells expressing scramble shRNA or shRNA targeting UHRF1 following the indicated time points after (A) 50 ng/ml TMP and 250 J/m^2 UVA, (B) 160 ng/ml MMC, (C) 10 μg/ml cisplatin, (D) 10 J/m^2 UVC, (E) 10Gy IR or (F) 4mM HU treatment. The ratios between ubiquitinated FANCD2 and FANCD2 are indicated below the lanes.

(Figure 2-7D) and IR (Figure 2-7E) treatment. Interestingly, knocking down UHRF1 affected FANCD2 monoubiquitination when the cells were treated with hydroxyurea (HU), which stalls replication fork (Koc et al., 2004). The difference is greater at later time points, when stalled replication forks start collapsing (Figure 2-7F, comparing lanes 5 to 10).

The FA pathway is known to be active only in S phase. One possible explanation for the lack of FANCD2 monoubiquitination when treated with TMP/UVA and MMC is that the depletion of UHRF1 led to cell cycle dysregulation. To test this hypothesis, we examined the cell cycle profile of the control HeLa and UHRF1 knocked down HeLa cells before and after TMP/UVA treatment. Knocking down UHRF1 in the HeLa cells did not have any significant impact on the cell cycle (Figure 2-8, left column), and there was no dramatic cell cycle arrest even after 3 hours from the TMP/UVA treatment, either (Figure 2-8, right column). To conclude, we have shown that UHRF1 is important for the FANCD2 monoubiquitination upon TMP/UVA and MMC treatment. Since this effect is not dependent on dysregulation of cell cycle, it is more likely that the presence of UHRF1 at the ICLs facilitate the monoubiquitination of FANCD2.

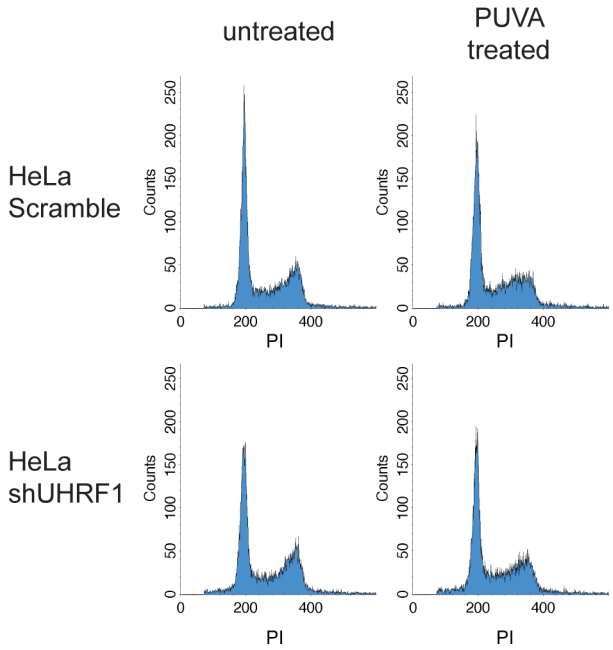


Figure 2-8. Knocking down UHRF1 in HeLa cells does not affect the cell cycle.
HeLa.Scramble and HeLa.shUHRF1 were untreated or treated with TMP/UVA. Cells were harvested after 3 hours, and stained with propidium iodide. Knock down of UHRF1 does not affect the cell cycle distribution in the absence of DNA damage, and TMP/UVA treatment does not cause dramatic cell cycle arrest 3 hours after treatment.

2.3.5 UHRF1 interacts directly with FANCD2 through its SRA domain

We have shown that UHRF1 plays an important role for the activation of the FA pathway, and that both UHRF1 and FANCD2 are recruited to ICLs *in vivo* with some time lag. We hypothesize that UHRF1 and FANCD2 are interacting with each other at the ICLs. To test if we could observe UHRF1 and FANCD2 in the same complex, we established a HeLa cell line stably expressing N-terminus FLAG-HA-tagged UHRF1, and immunoprecipitated (IP) against FLAG tag on UHRF1 after TMP/UVA treatment. UHRF1 was nicely immunoprecipitated (Figure 2-9A, lane 4), and we observed some enrichment of FANCD2 in the same sample (Figure 2-9A, comparing lanes 3 to 4). To validate this interaction, we established another HeLa cell line stably expressing N-terminus FLAG-HA-tagged FANCD2, and reciprocally immunoprecipitated against the FLAG tag on FANCD2 after TMP/UVA treatment. Similar to previous result, we successfully immunoprecipitated FANCD2, and observed some enrichment of UHRF1 (Figure 2-9B, comparing lanes 3 to 4). Our IP results show that UHRF1 interact with

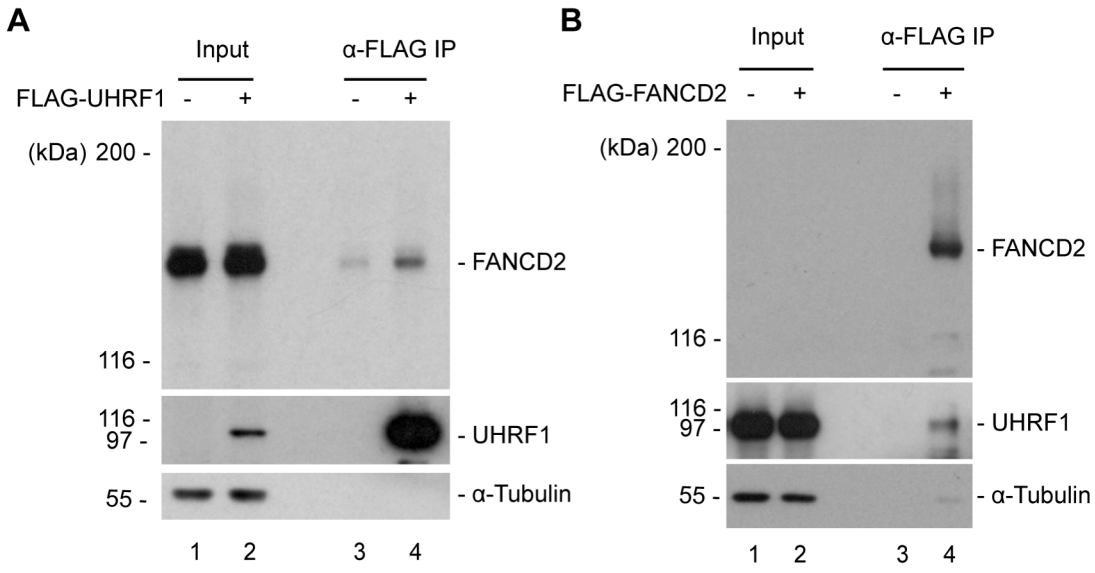


Figure 2-9. UHRF1 and FANCD2 are in the same protein complex *in vivo*
A) Immunoprecipitation of FLAG-UHRF1 from HeLa.shUHRF1 cells expressing exogenous FLAG-UHRF1. HeLa.Scramble was used as a negative control. Cells were treated with TMP/UVA and allowed to recover for 1 hour before lysis, immunoprecipitation and immunoblotting. B) Immunoprecipitation of FLAG-FANCD2 from HeLaS3 cells. HeLaS3 was used as a negative control. Cells were treated with TMP/UVA and allowed to recover for 1 hour before lysis, immunoprecipitation and immunoblotting.

FANCD2 *in vivo*, which suggests that they might interact with each other directly, or they might be in close proximity *in vivo* allowing them to be immunoprecipitated with each other.

To further investigate if UHRF1 and FANCD2 directly interact with each other, we purified recombinant StrepII-6xHis-tagged UHRF1 and FLAG-HA-tagged FANCD2, and evaluated their interaction by immunoprecipitating against the HA tag on FANCD2. 6xHis-MBP was used as a negative control in the same experiment. Consistent with the in vivo co-IP experiments, we observed a clear interaction between FANCD2 and UHRF1, but no interaction between

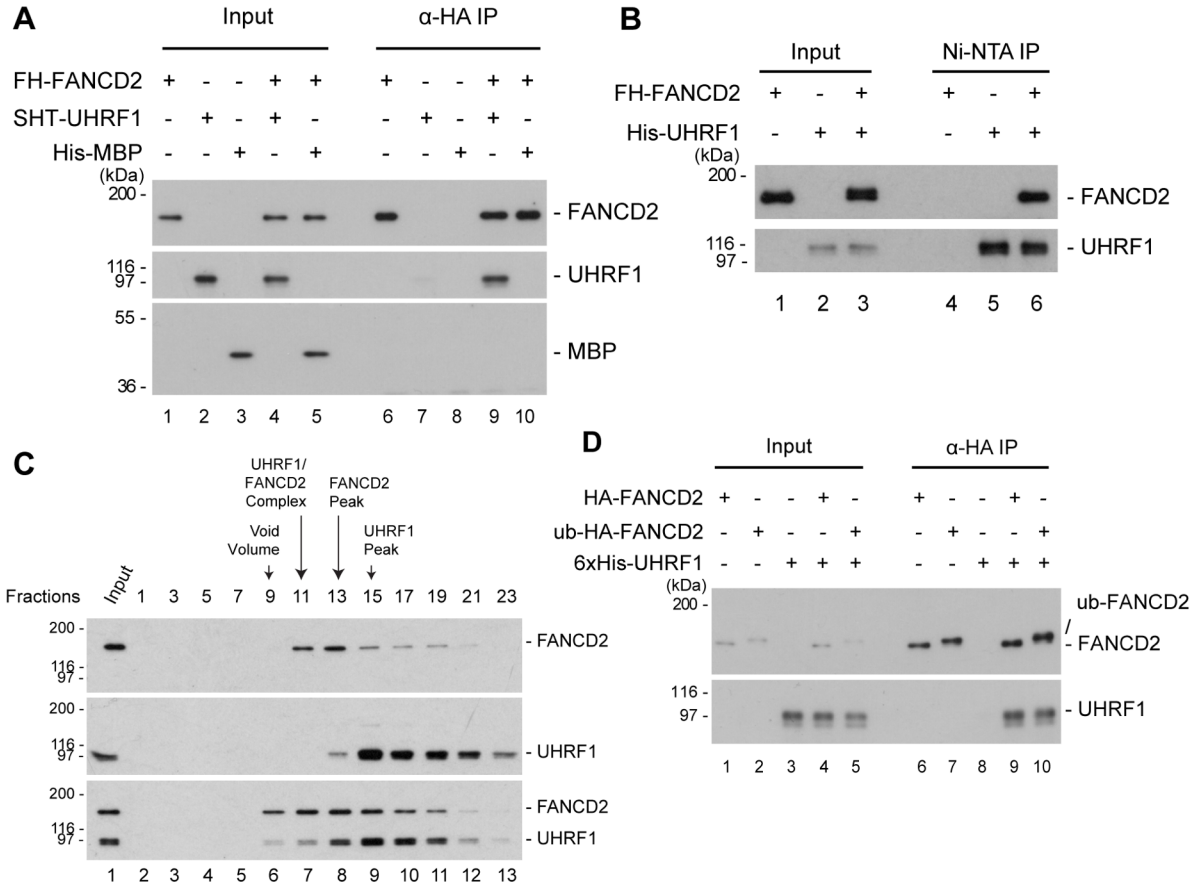


Figure 2-10. UHRF1 interacts directly with FANCD2.
A) *In vitro* binding assay using HA-FANCD2, 6xHis-UHRF1 (both purified from Sf9 cells) and 6xHis-MBP (purified from *E. coli*). HA-FANCD2 was immunoprecipitated and immunoblotting demonstrated co-immunoprecipitation of UHRF1 but not of MBP, which was used as a negative control. B) *In vitro* binding assay using FLAG-HA-FANCD2 and 6xHis-UHRF1 (both purified from Sf9 cells). 6xHis-UHRF1 was purified using NTA-agarose, and immunoblotting demonstrated co-purification of FANCD2. C) *In vitro* reconstituted FANCD2 and UHRF1 complex using recombinant proteins purified from Sf9 cells. Proteins were analyzed by gel filtration on a Superdex 200 5/150 GL chromatography column. The main FANCD2 peak is at fraction 13, and the main UHRF1 peak is at fraction 15. After complex formation, the peak shifted to fraction 11. D) *In vitro* binding assay of HA-FANCD2, monoubiquitinated HA-FANCD2 (ub-HA-FANCD2), both purified from HeLa cells, and 6xHis-UHRF1 purified from Sf9 cells. UHRF1 co-immunoprecipitates equally well with HA-FANCD2 and ub-HA-FANCD2, showing that FANCD2 interacts with UHRF1 independently of monoubiquitination.

FANCD2 and MBP (Figure 2-10A, lanes 9 and 10). Similarly, we could detect interaction between UHRF1 and FANCD2, when we pulled down UHRF1 against its 6xHis tag (Figure 2-10B). When we loaded the pre-incubated UHRF1 and FANCD2 protein mixture onto an analytical Superdex 200 size exclusion column, we observed a shift in the migration of the UHRF1/FANCD2 complex peak compared to the UHRF1 and FANCD2 monomer peaks (Figure 2-10C), which confirms the direct interaction between UHRF1 and FANCD2.

As shown previously, FANCD2 is monoubiquitinated when ICLs occur in the cells. Therefore, we purified unmodified and monoubiquitinated FANCD2 from HeLa cells to test if the monoubiquitination affects the interaction between UHRF1 and FANCD2. Both unmodified and monoubiquitinated FANCD2 interacted with UHRF1 (Figure 2-10D), which suggests that monoubiquitination of FANCD2 is not directly involved in the interaction with UHRF1.

In order to map the region in UHRF1 directly interacting with FANCD2, we designed a panel of UHRF1 deletion mutants removing individual domains (Figure 2-11A). We purified the recombinant FLAG-tagged UHRF1 domain deletion mutants (Figure 2-11B), and tested their interactions with FANCD2 using the same assay described previously. All the UHRF1 domain deletion mutants interacted with FANCD2 similarly to the full-length UHRF1 with the only exception of the SRA deletion mutant (Figure 2-11C), indicating that the SRA domain of UHRF1 is required for UHRF1 interaction with FANCD2 as well as binding to ICLs.

2.3.6 TTD and RING domain are dispensable for the role of UHRF1 in ICL repair

Some of the domains in UHRF1 are well-studied for their functions in the DNA methylation maintenance, e.g. the roles of SRA domain in hemi-methylated DNA recognition and recruiting DNMT1 to methylate the newly replicated DNA strand are well established (Arita et al., 2008; Avvakumov et al., 2008; Hashimoto et al., 2008; Qian et al., 2008). We stably expressed N-

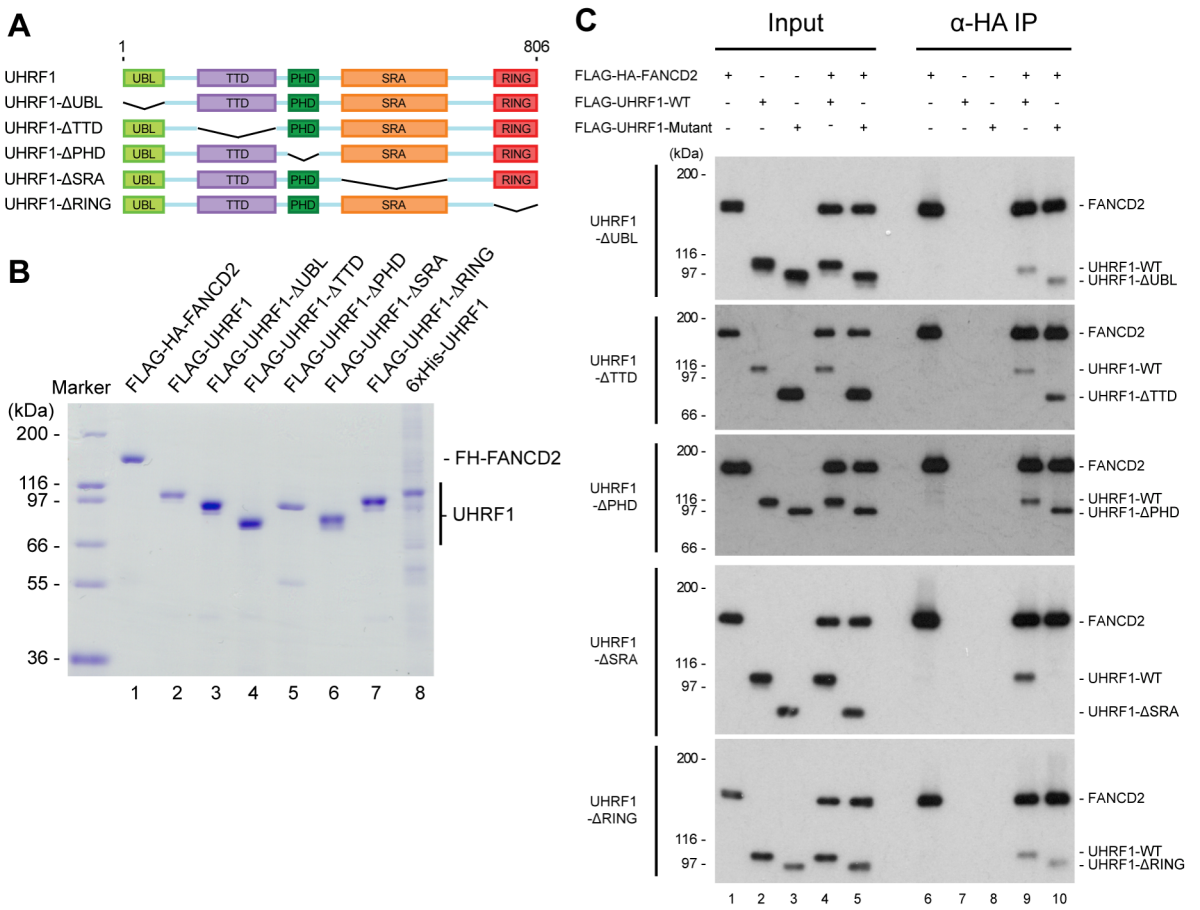


Figure 2-11. UHRF1 interacts with FANCD2 through its SRA domain.
A) Schematic of UHRF1 UBL, TTD, PHD, SRA and RING deletion mutants. B) Coomassie blue stain of recombinant FLAG-HA-tagged FANCD2, FLAG-tagged wild type and deletion mutants UHRF1 and 6xHis-tagged UHRF1 purified from Sf9 cells. C) *In vitro* binding assays of FLAG-HA-FANCD2 and wild type or deletion mutants of FLAG-UHRF1 (purified from Sf9 cells). HA-FANCD2 was immunoprecipitated and immunoblotting demonstrated that the SRA domain of UHRF1 is responsible for the protein-protein interaction between UHRF1 and FANCD2.

terminal mCherry-tagged UHRF1 bearing different domain deletion in HeLa UHRF1 knock down cells (Figure 2-12A and B), and tested with clonogenic assay if they can rescue the sensitivity towards MMC caused by UHRF1 depletion with clonogenic assay (Figure 2-12 C). Interestingly, only the TTD and RING deletion mutants together with full-length UHRF1 were able to rescue the sensitivity towards MMC. This result indicates that UBL, PHD and SRA domain are important for UHRF1's function in the ICL repair.

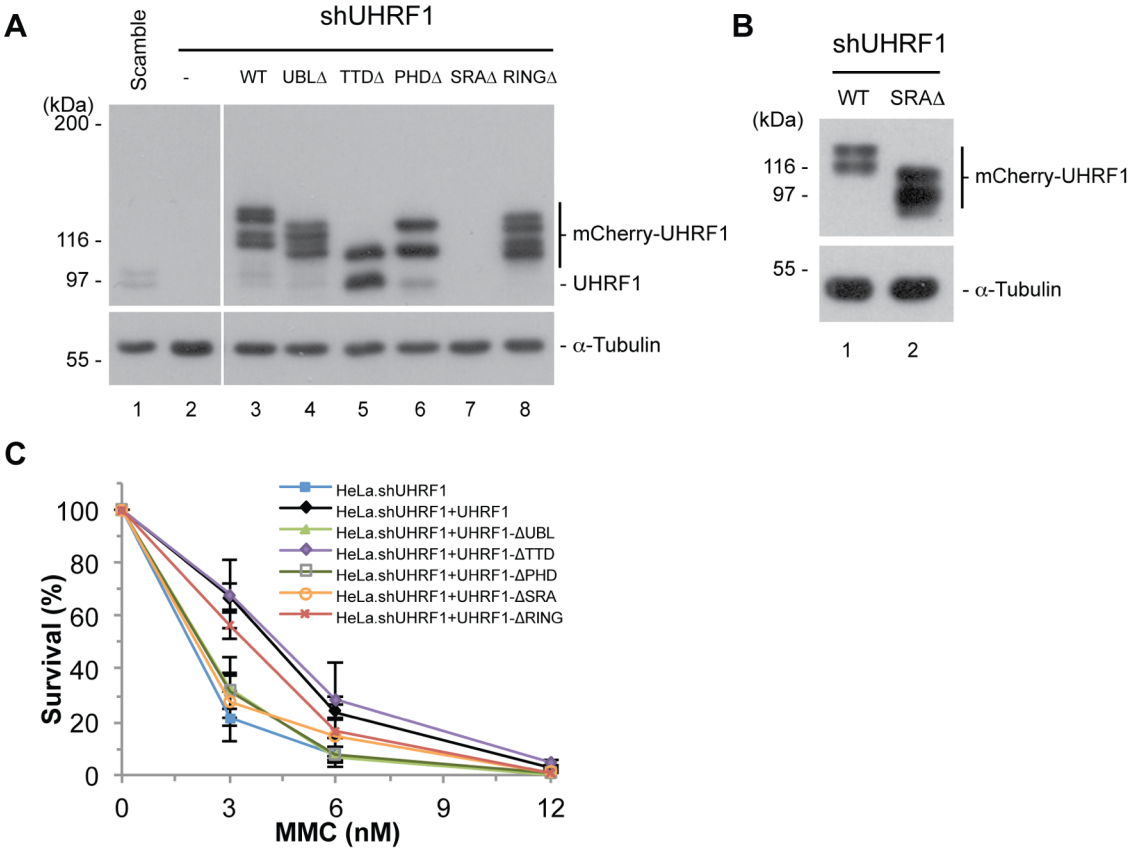


Figure 2-12. TTD and RING domains are dispensable for the roles of UHRF1 in the ICL repair
A) Western blot analysis confirms the knockdown level of endogenous UHRF1 and the expression levels of exogenous mCherry-tagged UHRF1 wild type and deletion mutants, cell lines which were used for the experiments described in C. The anti-UHRF1 antibody (sc-373; Santa Cruz Biotechnology) used recognizes an epitope in the SRA domain of UHRF1, so there was no detectable exogenous mCherry-UHRF1-ΔSRA in lane 7. The multiple species of mCherry-tagged UHRF1 was a result of insufficient denaturing. B) The same mCherry-UHRF1-WT and mCherry-UHRF1-ΔSRA as used in (A) were blotted using a different anti-UHRF1 antibody (sc-100606; Santa Cruz Biotechnology) confirming the expression level of mCherry-UHRF1-ΔSRA. The multiple species of mCherry-tagged UHRF1 was a result of insufficient denaturing. C) Clonogenic assay of HeLa.shUHRF1 cells expressing either mCherry-tagged UHRF1 wild type or deletion mutants as indicated. Only wild type, TTD and RING domain deletion mutants were able to rescue the MMC sensitivity.

2.4 Discussion

The repair of ICLs has been shown to be initiated through various mechanisms. It was shown that the FA pathway promotes the ICL repair in a replication dependent manner in the *Xenopus* cell-free system (Knipscheer et al., 2009; Raschle et al., 2008), which requires two replication forks converged at the ICL to initiate the repair (Zhang et al., 2015). However, it is also shown that when a replication fork encounters an ICL in mammalian cells, the majority undergoes FANCM/MHF dependent ICL traverse instead of being arrested at the ICL (Huang et al., 2013). Furthermore, there is evidence that ICLs are repaired outside of S phase of cell cycles, where no replication is present (Enoiu et al., 2012; Muniandy et al., 2009; Sarkar et al., 2006; Wang et al., 2001). Here, we have shown that UHRF1 is recruited to ICL very quickly *in vivo*, and plays an important role for the FA pathway activation for ICL repair.

UHRF1 Is a Sensor for ICLs

UHRF1 interacts with ICLs through its SRA domain. We have tested the interaction between UHRF1 and ICL with various deletion mutants removing individual domains, and we can only observe a defect in DNA binding when SRA domain is deleted or mutated. The crystal structure of the SRA domain interacting with hemi-methylated DNA shows rearrangement of the methylated cytosine to fit into the binding pocket of the SRA domain (Figure 2-13) (Arita et al., 2008). It is possible that the minor distortion on the DNA double helix created by TMP/UVA and MMC-induced ICL has a related structure to hemi-methylated DNA, which could fit into the binding pocket of SRA domain. Obtaining the structure of the UHRF1/ICL complex will help us understand how UHRF1 recognizes ICLs.

One fact that has been largely neglected in the literature so far is that different crosslinking agents lead to different degree of distortion on the DNA double helix (Deans and West, 2011; Lopez-Martinez et al., 2016). We have shown that UHRF1 serves as a sensor for TMP/UVA

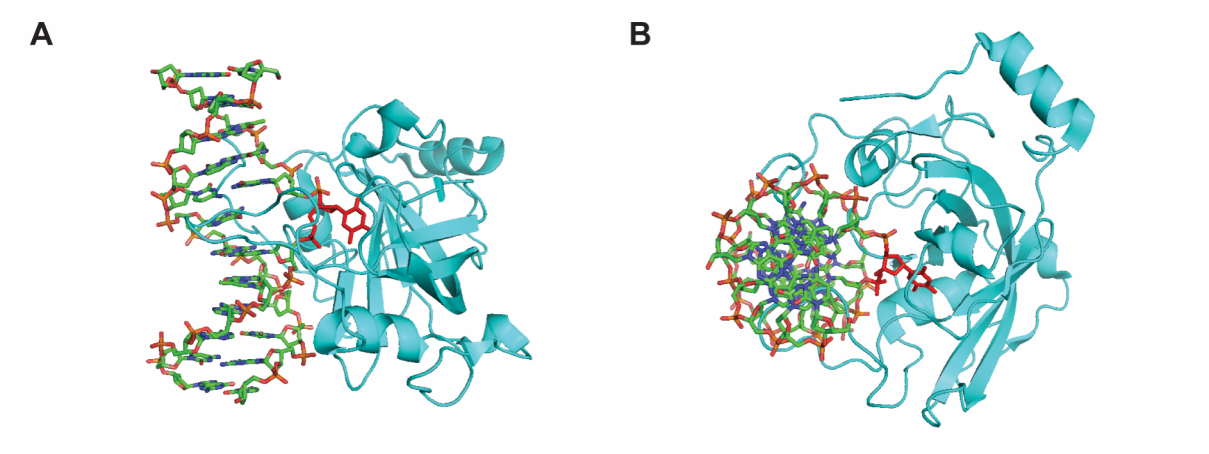


Figure 2-13. Crystal structure of SRA domain of UHRF1 in complex with hemi-methylated CpG DNA
The methylcytosine base at the hemi-methylated site is flipped out of the DNA helix in the SRA–DNA complex and fits tightly into a protein pocket on the concave surface. SRA domain Methylated cytosine is in red, and the SRA domain is in cyan. A) Side view. B) Top view. PDB: 2ZKD

and MMC-induced ICLs. It is likely that other proteins can take parts in other types of ICL recognition. Collision of the replication fork at the ICL has been shown to be a major event initiating the ICL repair (Zhang et al., 2015). Together, all these players could form a complex network guarding the genome from any ICL formed *in vivo*.

UHRF1 is required for the recruitment and monoubiquitination of FANCD2

We found that UHRF1 is recruited to ICLs within seconds of their appearance in the chromatin, which precedes and is required for the recruitment of FANCD2. Considering that UHRF1 directly interacts with ICLs via its SRA domain *in vitro*, we hypothesis that *in vivo* UHRF1 is recruited to ICLs directly through its SRA domain. However, FANCD2 is recruited to ICLs about 15 minutes later. While we have demonstrated that UHRF1 is important for this process, we cannot rule out that other proteins are also involved. The PHD domain deletion mutant failed to rescue the cells from the MMC sensitivity in our survival assay. As mentioned earlier, both TTD and PHD domains are involved in the recognition of histone H3K9me3. We speculate that some other proteins are recruited to the ICLs, which allow some modifications on the chromatin around the ICLs, and together with UHRF1, they form a landing platform for FANCD2 to be recruited to ICLs to initiate the repair (Figure 2-14).

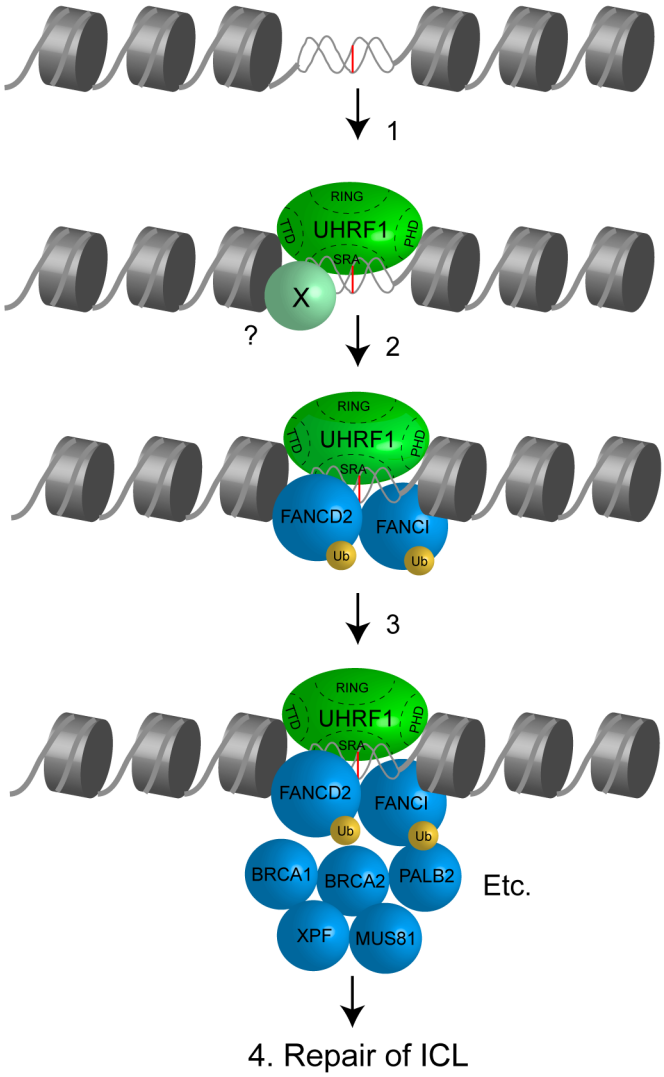


Figure 2-14. Proposed mechanism of UHRF1 in response to ICLs.
1) UHRF1 recognises ICLs in the chromatin. 2) UHRF1 facilitates the recruitment and monoubiquitination of FANCD2/FANCI heterodimer. It is unclear if there are other proteins (protein X) recruited to ICLs by UHRF1 to facilitate the recruitment and monoubiquitination of FANCD2/FANCI heterodimer. 3) Monoubiqutinated FANCD2/FANCI heterodimer recruits downstream repair proteins to damaged chromatin to repair the ICLs (step4).

Chapter 3 Structural Study of the full-length human FANCD2/FANCI complex utilizing cryo-electron microscopy

3.1 Introduction

The FANCD2/FANCI complex situates at the heart of the FA pathway. Monoubiquitination of the FANCD2 on lysine 561 (K561) and FANCI on lysine 523 (K523) determines the activation of the FA pathway. Because of the prominent role of FANCD2/FANCI complex in the FA pathway, FANCD2 and FANCI have been studied extensively since their first identification in 2001 and 2007 respectively (Dorsman et al., 2007; Sims et al., 2007; Smogorzewska et al., 2007; Timmers et al., 2001). It has been shown that the monoubiquitination of FANCD2/FANCI complex is required for its recruitment and retention at the ICLs (Garcia-Higuera et al., 2001; Smogorzewska et al., 2007), as well as for the recruitment of downstream repair proteins, such as SLX4/XPF nucleases complex and CtIP (Klein Douwel et al., 2014; Murina et al., 2014; Unno et al., 2014; Yamamoto et al., 2011).

The crystal structure of the mouse FANCD2/FANCI complex was determined (Joo et al., 2011), which the mouse FANCD2/FANCI complex forms a pseudo-two-fold symmetric structure (Figure 3-1; PDB: 3S4W). The inner groove of the FANCD2/FANCI complex is highly positively electrostatic, and it is proposed that the FANCD2/FANCI complex can bind to two DNA molecules through its pseudo-two-fold symmetric arrangement. Interestingly, the monoubiquitination sites are embedded in the heterodimer interface, which raised the question if FANCD2 and FANCI are monoubiquitinated prior to the heterodimer formation. In this chapter, we describe the structural insight of the human full-length FANCD2/FANCI complex using cryo-electron microscopy (cryo-EM). The pseudo-two-fold symmetry is lost in our cryo-EM structure, which the C-terminal of FANCD2 is positioned differently, and we named it as

Tower domain. We demonstrated that this Tower domain is indispensable for the function of FANCD2/FANCI complex through its DNA binding activity.

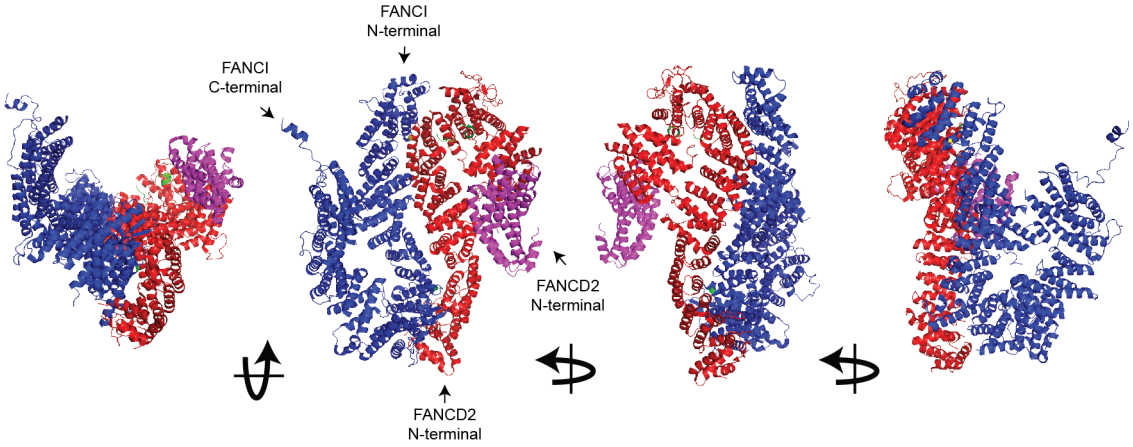


Figure 3-1. Crystal structure of mouse FANCD2/FANCI complex.
The pseudo-two-fold symmetric mouse FANCD2/FANCI structure in different orientations. FANCI is in blue, FANCD2 is in red, and the C-terminal of FANCD2 is in magenta. PDB: 3S4W

3.2 Specific aims

1. Study the structure of human full-length FANCD2/FANCI complex
2. Characterize the role of the Tower domain in the ICL repair

3.3 Results

3.3.1 Cryo-EM structure of the full-length and C-terminal deletion of FANCD2/FANCI complex

We co-expressed FLAG-HA-tagged FANCD2 and 6xHis-tagged FANCI in the Sf9 insect cells, and performed tandem affinity purification to purify the FANCD2/FANCI complex exploiting both the FLAG and His tags. Following the affinity purifications, the FANCD2/FANCI complex was further purified by size exclusion chromatography. The FANCD2/FANCI complex showed high purity level and was characterized by a near one to one ratio between FANCD2 and FANCI (Figure 3-2A). In collaboration with Dr. Catherine Vénien-Bryan and colleagues, Zhuolun Li and Dr. William Nicholson, we subjected the recombinant FANCD2/FANCI complex to negative stain electron microscopy using random conical tilt method (Figure 3-2B and C), and generated an initial 3D model. We then subjected the recombinant FANCD2/FANCI complex to cryo-EM, and further refined the initial model to a resolution of 22 Å. The FANCD2/FANCI complex is about 90 Å wide and 160 Å long, and it has a hollow central part. On the side, we also observed a characteristic 60 Å protruding domain, which we termed the Tower domain (Figure 3-2D). In order to compare our structure of the human FANCD2/FANCI complex (hFANCD2/FANCI) with the murine complex (mFANCD2/FANCI), we lowered the resolution of the published mouse crystal structure of FANCD2/FANCI complex to 20 Å (PDB: 3S4W) (Joo et al., 2011). Intriguingly, the pseudo-two-fold symmetry observed in the mouse crystal structure is not present in our cryo-EM structure, where the Tower domain is positioned differently (Figure 3-2E). We then docked our cryo-EM structure with the mFANCD2/FANCI crystal structure, allowing flexibility of each of the 8 solenoids in the crystal structure relative to each other, and observed good agreement between the two structures (Figure 3-2F). Considering that the human and mouse

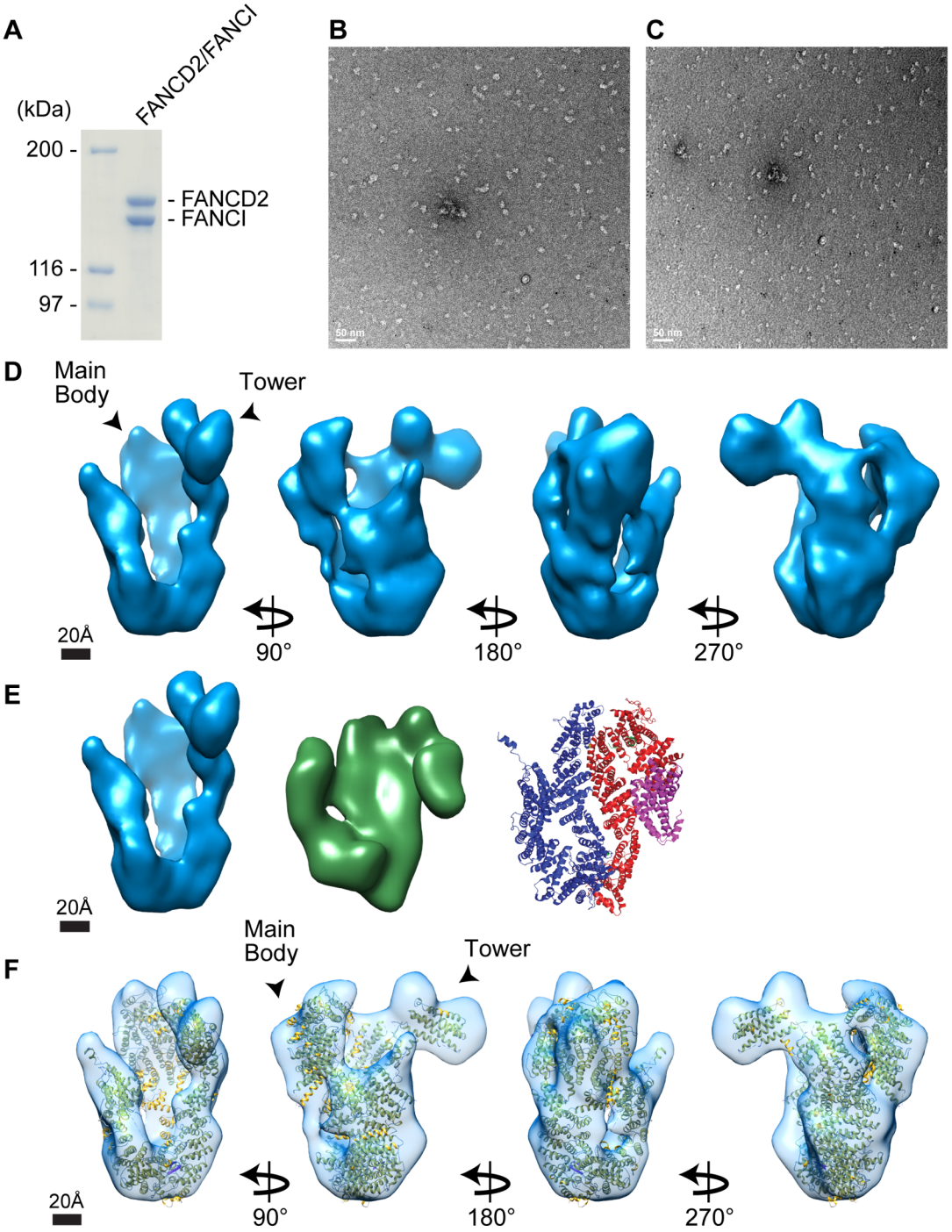


Figure 3-2. Cryo-EM structure of the FANCD2/FANCI complex.
A) Coomassie blue staining of the recombinant human full-length FANCD2/FANCI heterodimer purified from Sf9 cells. B) Micrograph of negative staining EM full-length FANCD2/FANCI complex 0° tilt. C) Tilted micrograph (50° tilt) of the same area as in (B). D) Cryo-EM density map of full-length hFANCD2/FANCI complex in different orientations. (Scale bar: 20 Å) E) Comparison of full-length hFANCD2/FANCI cryo-EM structure (left, blue) with the mFANCD2/FANCI crystal structure (PDB: 3S4W) (middle and right). The mFANCD2/FANCI crystal structure was filtered to the same resolution as the 3D cryo-EM model (middle, green). (Scale bar: 20 Å) F) Different orientations of cryo-EM density map of full-length hFANCD2/FANCI complex docked with the mFANCD2/FANCI crystal structure (PDB: 3S4W). (Scale bar: 20 Å) Dr. Catherine Vénien-Bryan and Dr. William Nicholson performed the experiments and image analysis for the negative staining and cryo-EM data shown in Figure 3-2 B-F.

FANCD2/FANCI complex are highly conserved, and they share 75% amino acid sequence identity, it is less likely that the difference in amino acid sequence gives rise to the loss of the pseudo-two-fold symmetry. While the resolution of our cryo-EM structure is limited, we speculate that the disagreement between the two structures may be a result of the deletions introduced in the mouse crystal structure, which will be discussed in detail in the discussion section of this chapter.

According to the docking of the two structures, the Tower domain is comprised of the C-terminal of FANCD2. To validate this, we generated a FANCD2 C-terminal deletion mutant (FANCD2-ΔTower), removing the last 305 amino acids of FANCD2 (Figure 3-3A). We co-expressed FLAG-HA-tagged FANCD2-ΔTower and 6xHis tagged-FANCI, and purified the heterodimer via tandem affinity purification as previously described (Figure 3-3B). The recombinant FANCD2-ΔTower/FANCI comple x was subjected to cryo-EM, resulting in a structure at 20 Å resolution (Figure 3-3C). The overall structure of the FANCD2-ΔTower/FANCI complex is comparable to that of the full-length complex, with the exception of the predicted disappearance of the Tower domain (Figure 3-3D). Docking of the mFANCD2/FANCI crystal structure lacking the last 305 amino acids to our cryo-EM FANCD2-ΔTower/FANCI complex structure further confirmed that the observed Tower domain is formed by the C-terminal of FANCD2 (Figure 3-3E).

3.3.2 The Tower domain of FANCD2 is functionally indispensable *in vivo*

Intrigued by the striking position of the Tower domain, we decided to investigate if the Tower domain has any functional implication in ICL repair. We introduced the full-length FANCD2 or FANCD2-ΔTower into PD20 cells, a patient-derived FANCD2-deficient cell line (Timmers et al., 2001) (Figure 3-4A), and tested if the FANCD2-ΔTower can rescue the MMC sensitivity phenotype of PD20 cells. In contrast to the full-length FANCD2, the PD20 cells

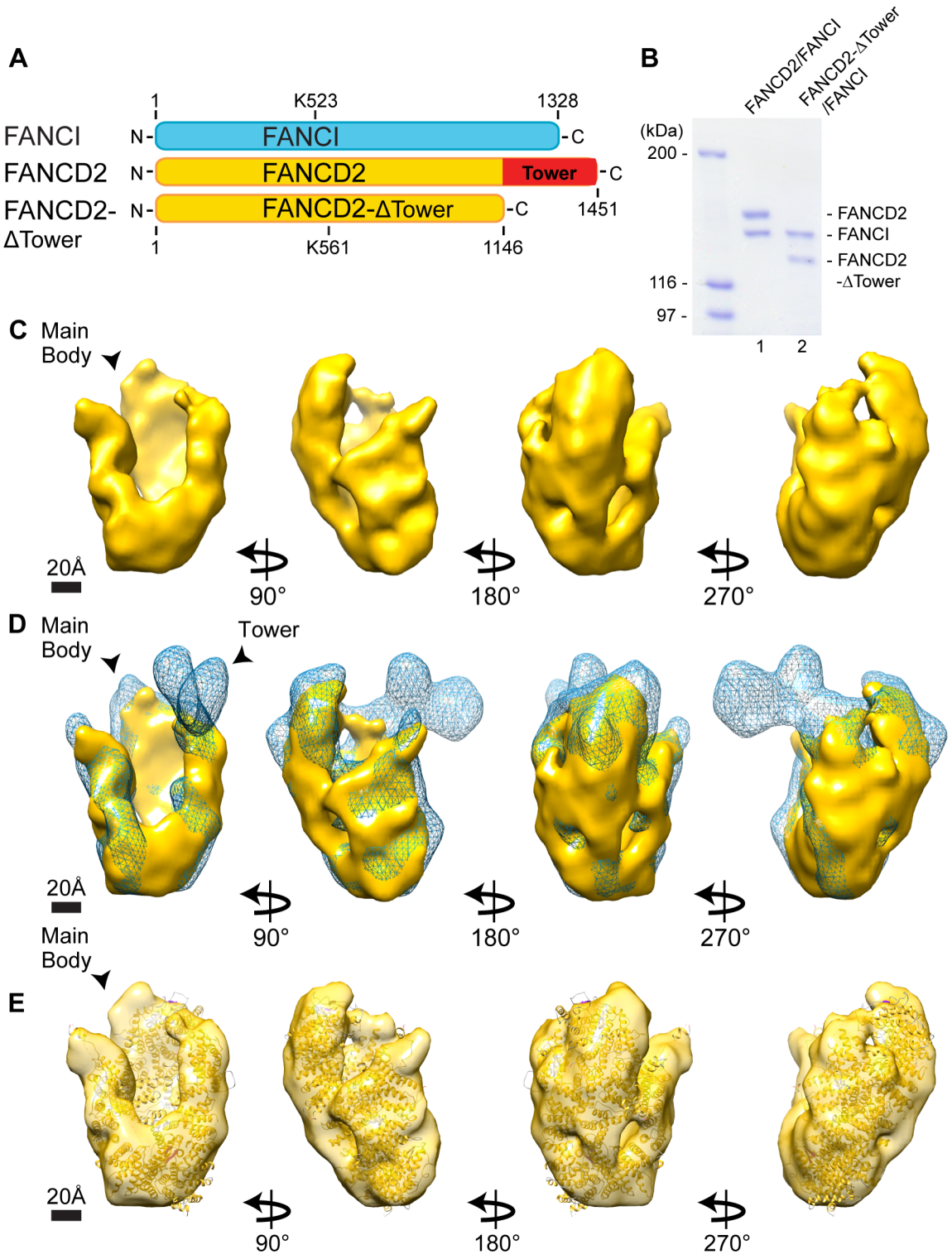


Figure 3-3. Analysis of the Tower domain of FANCD2.
A) Schematic of FANCI and FANCD2 indicating the thumb-like domain and Tower domain. B) Coomassie blue stain of recombinant full-length FANCD2/FANCI and FANCD2-ΔTower/FANCI heterodimer purified from Sf9 cells. C) Different orientations of cryo-EM density map of the FANCD2-ΔTower/FANCI complex. (Scale bar: 20 Å) D) Different orientations of cryo-EM density map of full-length FANCD2/FANCI complex (in blue mesh) superimposed with the FANCD2-ΔTower/FANCI complex (in gold). (Scale bar: 20 Å) E) Different orientations of cryo-EM density map of FANCD2-ΔTower/FANCI complex docked with the mouse FANCD2/FANCI crystal structure deleting amino acids 1144 – 1450 (PDB: 3S4W). (Scale bar: 20 Å) Dr.

expressing FANCD2-ΔTower are more sensitive to MMC treatment (Figure 3-4B). As demonstrated in Chapter 2, FANCD2 is recruited to the TMP/UVA induced ICLs. To test if the Tower domain is required for the recruitment of FANCD2 to ICLs, we established a FANCD2 knocked down cell line using shRNA targeting FANCD2, and complemented it with RNAi resistant (i) EGFP-tagged FANCD2 wildtype (EGFP-FANCD2), (ii) EGFP-FANCD2 containing the lysine 561 to arginine mutation (EGFP-FANCD2-K561R) and (iii) EGFP-tagged FANCD2-ΔTower (EGFP-FANCD2-ΔTower) (Figure 3-4C). An 8 amino acid nuclear localization signal (NLS) was added to the C-terminal of FANCD2-ΔTower to ensure proper nuclear localization. Recruitment of full-length FANCD2 to the damage site was detectable after 10 minutes, whereas, in the absence of the Tower domain, FANCD2 was no longer recruited to the ICLs (Figure 3-4D), suggesting that the defect of FANCD2-ΔTower in the clonogenic survival assay was due to inability of FANCD2 to relocate to the ICLs.

It has been shown that monoubiquitination of FANCD2 on K561 is important for its recruitment to the damage induced foci on chromatin (Taniguchi et al., 2002b). We hypothesized that maybe the lack of monoubiquitination of the FANCD2/FANCI complex was the reason for the defects we observed in the absence of the Tower domain. We treated the complemented PD20 cells shown in Figure 3-4A with MMC, and fractionated the cells into cytoplasmic and nuclear/chromatin fractions to enrich the monoubiquitinated FANCD2 in the nuclear/chromatin fraction. The full-length FANCD2 was able to support the monoubiquitination of both FANCD2 and FANCI upon MMC treatment (Figure 3-5A, lanes 9 and 10). As we expected, removal of the Tower domain prevents the monoubiquitination of both FANCD2 and FANCI (Figure 3-5A, lanes 11-12) underscoring the importance of the Tower domain.

One of the possibilities for the lack of monoubiquitination of both FANCD2 and FANCI when removing the C-terminal of FANCD2 is that the Tower domain is responsible for the

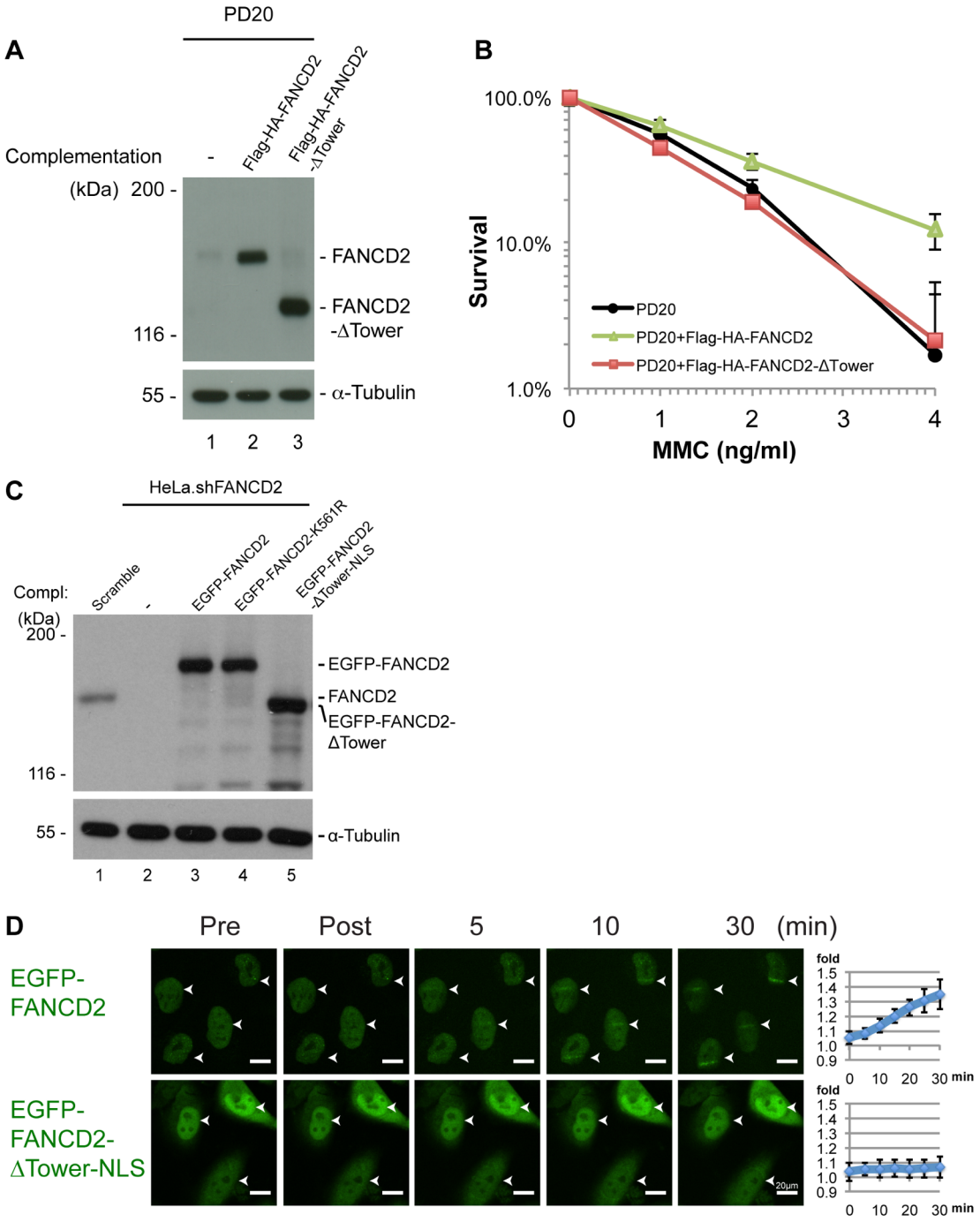


Figure 3-4. The Tower domain is critical for FANCD2 function in the FA pathway.
A) Western blot analysis of the expression levels of PD20 cells complemented with Flag-HA-FANCD2 and Flag-HA-FANCD2-ΔTower used in B. B) Clonogenic survival assay of PD20 cells complemented with Flag-HA-FANCD2 or Flag-HA-FANCD2-ΔTower. Experiment was carried out two times, and each repeat was done in triplicate. Representative triplicate data is shown. Error bars based on individual triplicate show standard deviation. C) Western blotting of the HeLa.Scramble, HeLa.shFANCD2 and HeLa.shFANCD2 complemented with EGFP-tagged wild type, K561R and ΔTower-NLS FANCD2 used in Figure 3-3D showing the expression level of FANCD2. D) HeLa.shFANCD2 cells expressing EGFP-FANCD2 or EGFP-FANCD2-ΔTower-NLS were pre-treated with TMP, and microirradiated at the indicated area (white arrows). Wild type FANCD2 was recruited to TMP-induced ICLs sites, while FANCD2-ΔTower-NLS was not. Representative fields are shown. Quantifications of EGFP-FANCD2 (6 cells quantified) or EGFP-FANCD2-ΔTower-NLS (6 cells quantified)

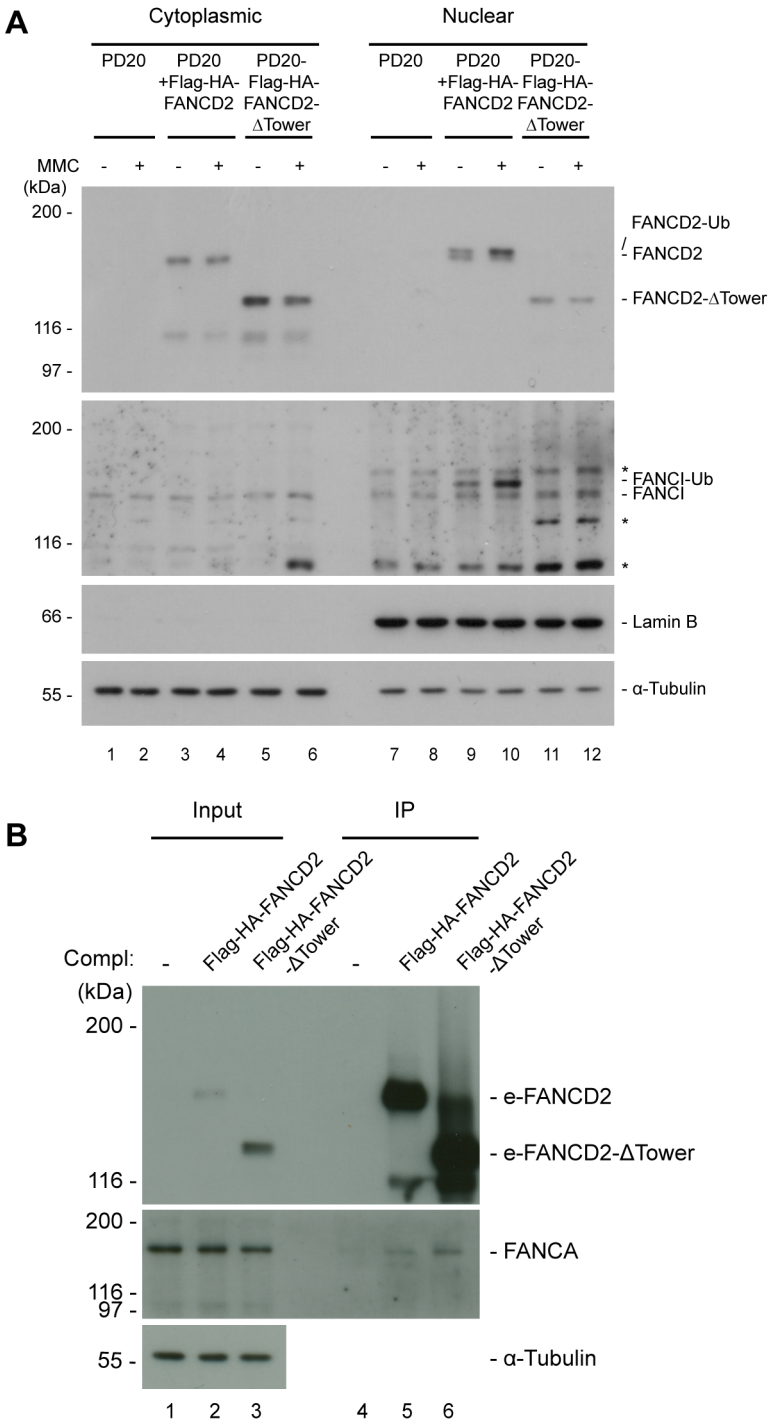


Figure 3-5. The Tower domain is required for the monoubiquitination of FANCD2 but dispensable for the interaction with the FA core complex.
A) FANCD2 monoubiquitination response in PD20 cells complemented with wild type Flag-HA-FANCD2 or Flag-HA-FANCD2-ΔTower after overnight treatment with 160ng/ml MMC. B) Immunoprecipitation of Flag-HA-FANCD2 and Flag-HA-FANCD2-ΔTower from PD20 cells. Immunoblotting was performed against FANCD2 and FANCA (component of the FA core complex). PD20 without complementation was used as a negative control. Cells were treated overnight with 160ng/ml MMC.

interaction between the FANCD2/FANCI complex and the FA core complex. To test this hypothesis, we performed a simple anti-FLAG immunoprecipitation from the FLAG-HA-tagged full-length FANCD2 and FANCD2-ΔTower complemented PD20 cells. In fact, should the FA core complex and the Tower domain interact, a decrease in the amount of co-immunoprecipitated subunits of the FA core complex, e.g. FANCA, should be observed. However, contradicting to this hypothesis, a comparable amount of co-immunoprecipitated FANCA was detected in both the full-length FANCD2 and FANCD2-ΔTower samples (Figure 3-5B, comparing lanes 5 and 6), indicating that the Tower domain is dispensable for the interaction between FANCD2 and the FA core complex.

3.3.3 The Tower domain of FANCD2 is required for its DNA binding activity

FANCD2 and FANCI have been shown to bind to various naked DNA substrates *in vitro* either as a monomer or heterodimer (Joo et al., 2011; Longerich et al., 2009; Park et al., 2005). However, it was not clear if the DNA binding activities have any functional implication in the ICL repair, except that addition of DNA robustly stimulates the monoubiquitination of FANCD2 *in vitro* (Sato et al., 2012). Since we previously observed that the Tower domain is required for proper functions of the FANCD2/FANCI complex *in vivo* (Figure 3-4B and Figure 3-5A), we hypothesized that FANCD2-ΔTower may be defective in binding DNA. To test this hypothesis, we tested the DNA binding activities of recombinant FANCD2/FANCI and FANCD2-ΔTower/FANCI heterodimer towards a substrate resembling a replication fork (ICL14FF). In good agreement with the literature, we observed weak interaction between FANCD2/FANCI and ICL14FF (Figure 3-6A, lane 2). Such interaction was confirmed by performing a super-shift using anti-HA antibody targeting FANCD2, and as expected, we observed a shift of the FANCD2/FANCI/DNA band (Figure 3-6A, lane 3). When the Tower domain is absent, the FANCD2/FANCI/DNA complex was not detectable (Figure 3-6A, lanes 4 and 5), confirming that the Tower domain is required for the DNA binding activity of FANCD2. To gain further

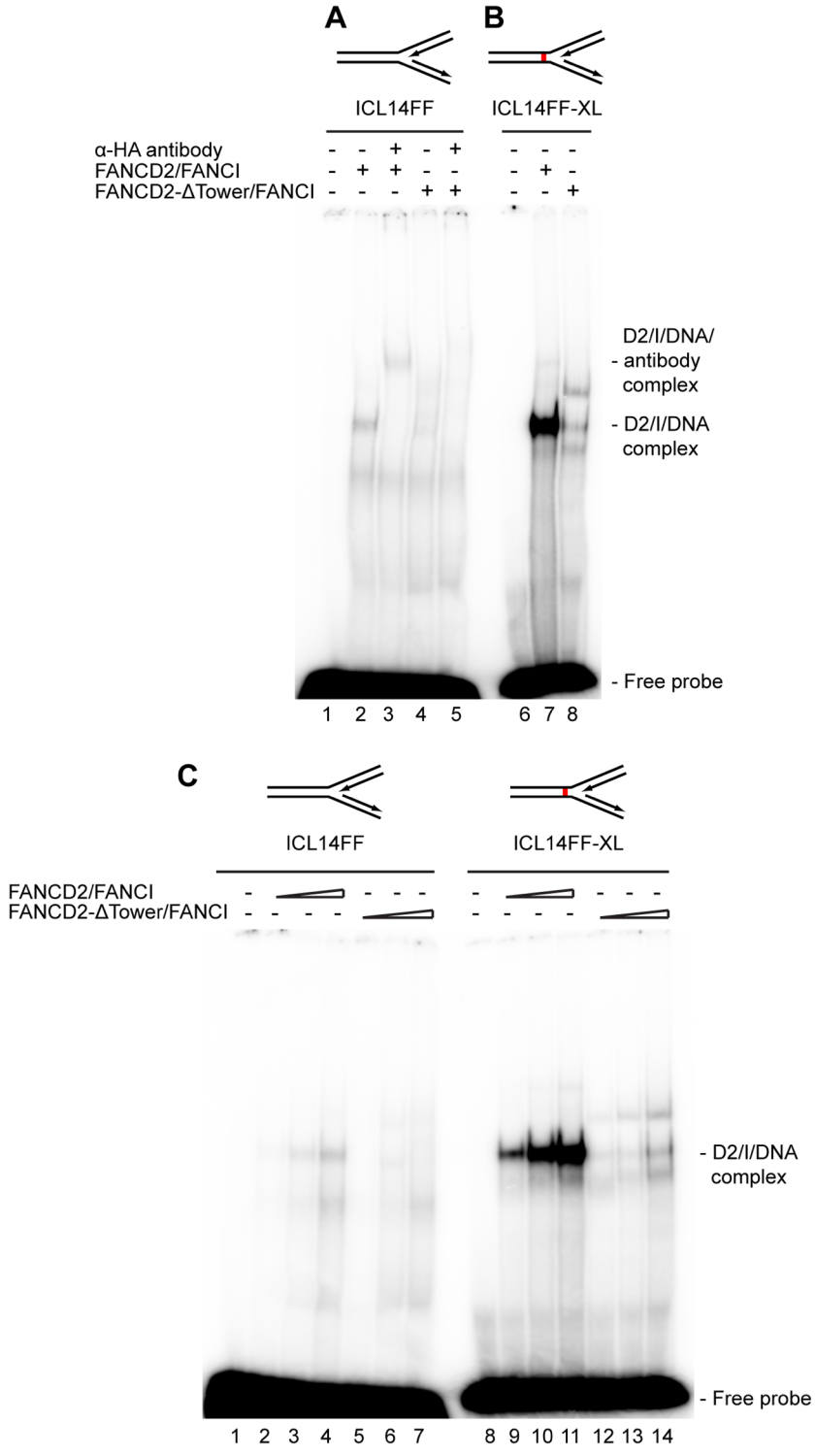


Figure 3-6. The Tower domain of FANCD2 is important for DNA binding.
A) EMSA showing full-length FANCD2/FANCI complex, but not the FANCD2-ΔTower/FANCI complex, forms specific protein-DNA complexes with replication fork substrate ICL14FF. B) EMSA showing full-length FANCD2/FANCI complex forms strong specific protein-DNA complexes with the crosslinked replication fork substrate, ICL14FF-XL. On the contrary, the FANCD2-ΔTower/FANCI complex bound weakly to ICL14FF-XL. C) Titration of protein concentration of full-length FANCD2/FANCI complex and FANCD2-ΔTower/FANCI complex in EMSA with replication fork substrate ICL14FF (left) and crosslinked replication fork substrate ICL14FF-XL.

insight into the mechanism of the interaction, we introduced an ICL into ICL14FF, which now resembled a replication fork arrested at an ICL (ICL14FF-XL). Consistent with previous result, removal of the Tower domain greatly attenuated the DNA binding activity of the FANCD2/FANCI heterodimer (Figure 3-6B, lanes 7 and 8). Intriguingly, the full-length FANCD2/FANCI heterodimer showed stronger affinity towards ICL14FF-XL than ICL14FF (Figure 3-6B, lane 7), implying that a stalled replication fork at an ICL could be the preferential substrate of the FANCD2/FANCI heterodimer *in vivo*. Titrating the amount of protein in the binding reaction confirmed the observed preference (Figure 3-6C).

3.3.4 FANCD2/FANCI is first recruited to DNA, then ubiquitinated by FANCL

We have demonstrated that FANCD2/FANCI directly interacts with a DNA molecule mimicking a replication fork arrested at an ICL, and that this protein/DNA interaction is dependent on the Tower domain of FANCD2. Our data also showed that the monoubiquitination of FANCD2/FANCI is abrogated when the Tower domain is missing. Therefore, we

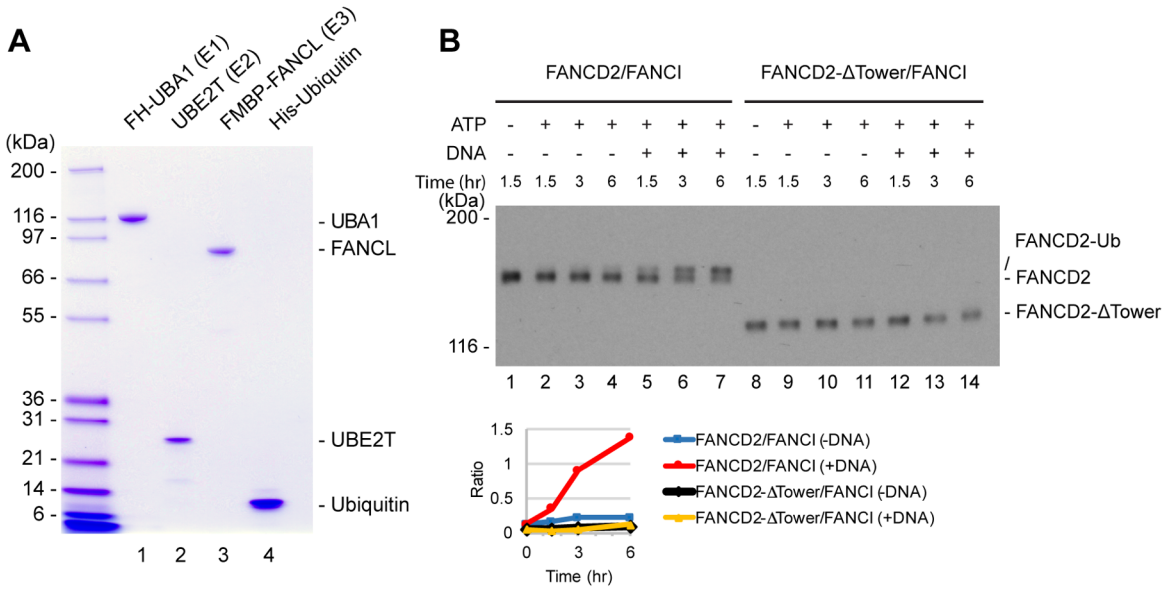


Figure 3-7. The Tower domain is critical for the monoubiquitination of FANCD2 after binding to DNA.
A) Coomassie blue staining of purified recombinant FLAG-HA-UBA1 (E1), UBE2T (E2), FLAG-MBP-FANCL (E3) and 6xHis-ubiquitin. B) *In vitro* ubiquitination assay of the full-length FANCD2/FANCI complex and FANCD2-ΔTower/FANCI complex. Immunoblot was performed using antibodies against FANCD2. Quantification was done by using ImageJ/Fiji software.

hypothesized that FANCD2 ubiquitination occurs upon the recruitment of FANCD2 to the DNA. To address this hypothesis an *in vitro* ubiquitination assay was performed. We reconstituted the *in vitro* ubiquitination assay by expressing and purifying recombinant UBA1 (E1), UBE2T (E2), FANCL (E3) and ubiquitin (Figure 3-7A). If we set up the reaction with only the enzymes and the FANCD2/FANCI heterodimer, we could only observe minimum amount of monoubiquitinated FANCD2 (Figure 3-7B, lanes 2 to 4). Conversely, in good agreement with the literature (Sato et al., 2012), addition of DNA into the *in vitro* ubiquitination reaction strongly stimulated the monoubiquitination of FANCD2 (Figure 3-7B, lanes 5 to 7). On the other hand, we could not detect any increased amount of monoubiquitination when we used the FANCD2-ΔTower/FANCI heterodimer as a substrate (Figure 3-7B, lanes 8 to 14). Moreover, the amount of monoubiquitination of FANCI in both FANCD2/FANCI and FANCD2-ΔTower/FANCI heterodimers was not comparable to that of FANCD2 and it was independent of the addition of DNA, in agreement with previously reported studies (Longerich et al., 2014;

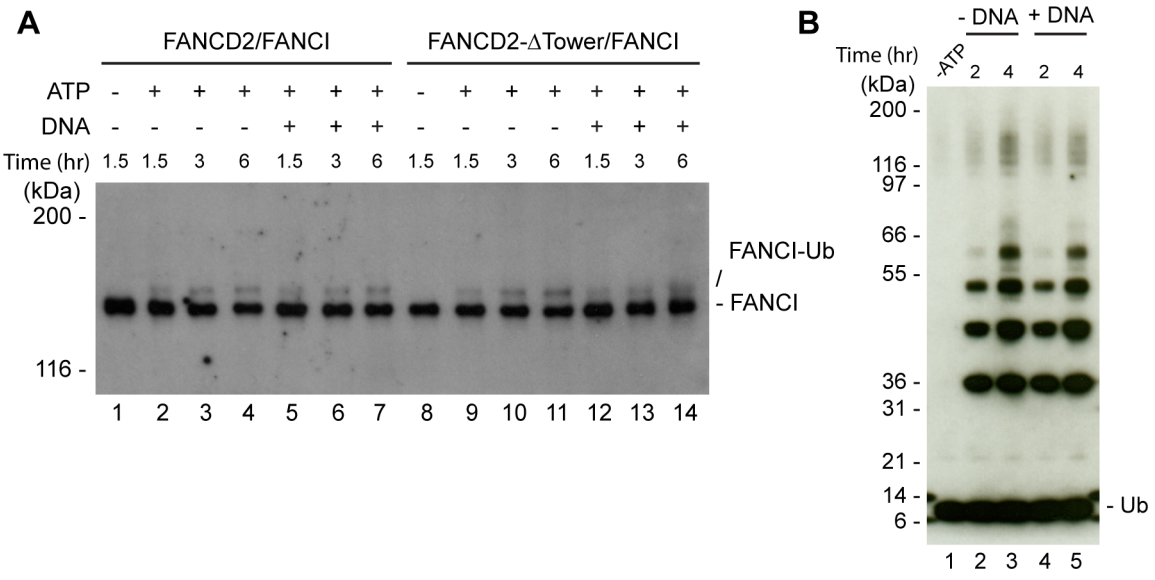


Figure 3-8. Addition of DNA in the *in vitro* ubiquitination assay has no impact on the monoubiquitination of FANCI and the activity of FANCL.
A) *In vitro* ubiquitination assay of the full-length FANCD2/FANCI and FANCD2-ΔTower/FANCI complexes. Immunoblot was performed using antibodies against FANCI. B) *In vitro* ubiquitination assay of full-length FANCD2/FANCI complex. Immunoblot was probed using antibodies against ubiquitin.

Sato et al., 2012) (Figure 3-8A). The increase in specific monoubiquitination was not due to increased general E3 ligase activity in the reaction (Figure 3-8B).

Taken together, our data suggests that FANCD2/FANCI complex needs to be bound to DNA to allow monoubiquitination to take place, and that the Tower domain is required for the FANCD2/FANCI/DNA interaction. If FANCD2/FANCI complex is recruited to the ICLs in its unmodified form, followed by the monoubiquitination to secure its retention on the chromatin, we should be able to detect some weak recruitment of FANCD2 in the absence of the ubiquitination. To test this theory, we substituted the lysine 561 of FANCD2 to arginine (FANCD2-K561R) preventing monoubiquitination, and introduced EGFP-tagged FANCD2, EGFP-tagged FANCD2-K561R or EGFP-tagged FANCD2-ΔTower together with mCherry-UHRF1 into CRISPR/Cas9-mediated FANCD2 knockout HeLa cells (HeLa-FANCD2 -/-) (Figure 3-9A). Indeed, we could detect some rapid and weak recruitment of the FANCD2-K561R mutant to ICLs, while there was little trace of recruitment of the FANCD2-ΔTower mutant in the same experiment (Figure 3-9, B and C). Collectively, we provided evidence not only that FANCD2/FANCI heterodimer is recruited to ICLs in its native form, and monoubiquitinated on the chromatin, but also that the Tower domain is crucial for the interaction between the FANCD2/FANCI complex and DNA. Intriguingly, we found several disease-causing mutations reside in the Tower domain, which result in short deletion or truncation of the Tower domain (Figure 3-10) (Kalb et al., 2007), underlining the importance of the Tower domain.

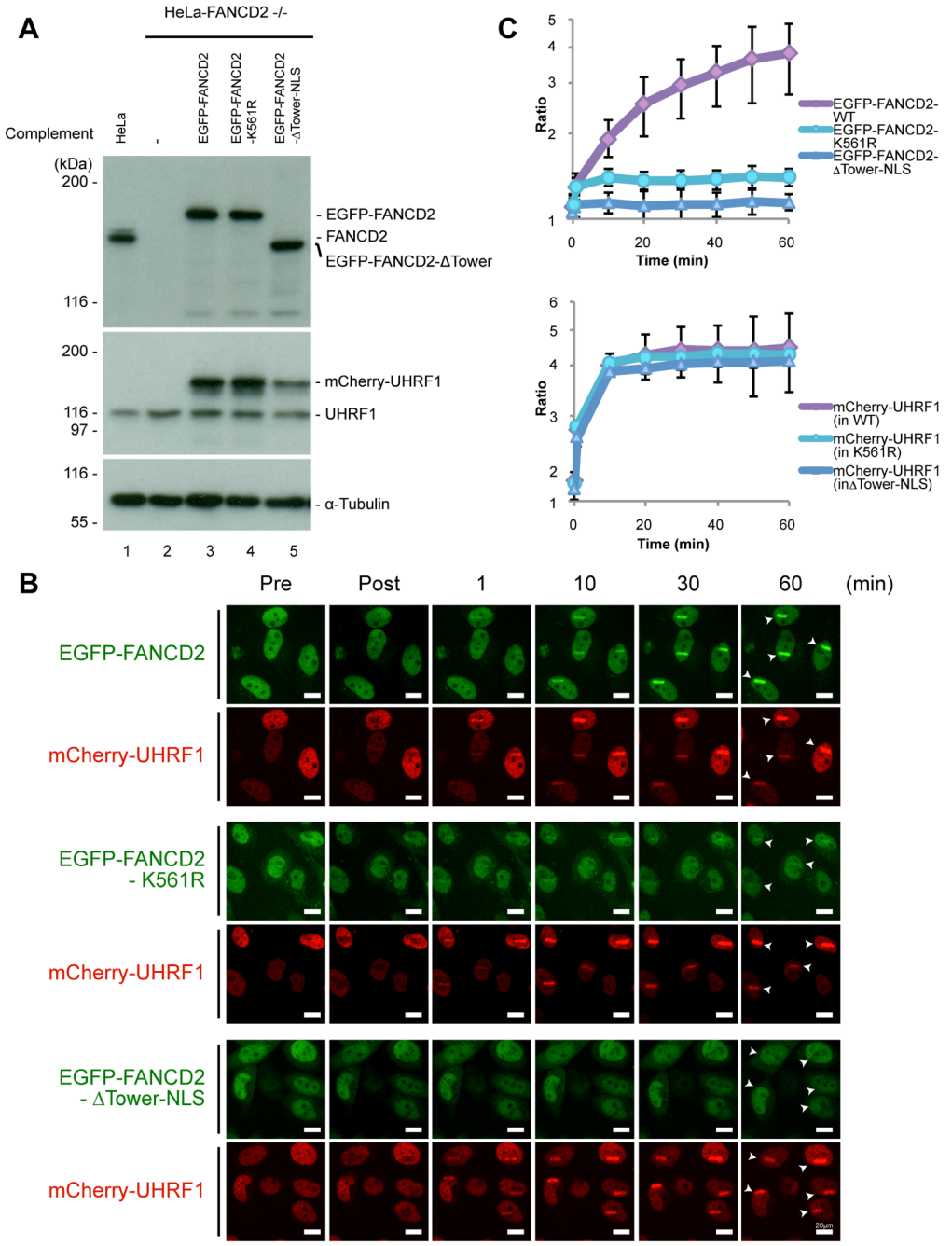


Figure 3-9. FANCD2 recruitment to DNA through the Tower domain precedes its monoubiquitination on lysine 561.

A) Western blot analysis of HeLa-FANCD2 -/- complemented with mCherry-UHRF1 and EGFP-tagged FANCD2 wild type, K561R or ΔTower-NLS. Blots showed comparable expression levels of UHRF1 and FANCD2 variants. B) HeLa-FANCD2 -/- cells expressing EGFP-tagged wild type, K561R and ΔTower-NLS FANCD2 were pre-treated with TMP (pre), and micro-irradiated at the indicated areas (white arrows). FANCD2 recruitment was monitored over time after damage induction. mCherry-UHRF1 was co-expressed in all three cell lines and used as positive controls for the introduction of ICLs. Wild type EGFP-FANCD2 was recruited to ICLs rapidly and strongly. EGFP-FANCD2-K561R was recruited rapidly but weakly, while no recruitment of EGFP-FANCD2-ΔTower-NLS was observed. Representative fields of view are shown. C) Quantification of the fluorophore-tagged proteins recruited to ICLs at the irradiated sites (8, 11 and 10 cells were quantified for EGFP-FANCD2, EGFP-FANCD2-K561R and EGFP-FANCD2-ΔTower-NLS, respectively). The ratio shows the mean intensity of the fluorophore at the damaged area divided by the intensity of fluorophore at rest of the nucleus. Scale bar: 20μm. Error bars show standard deviation. David Lopez-Martinez generated the FANCD2 CRISPR/Cas9-mediated knockout in HeLa cells.

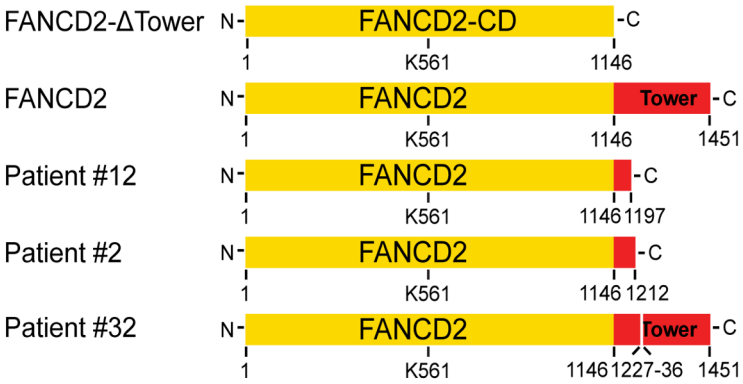


Figure 3-10. Patient-derived mutations in the Tower domain.
Schematic of patient derived mutations in the Tower domain that led to deletion of the Tower domain in the literature.

3.3.5 Electrostatic interactions underlie the binding of the Tower domain of FANCD2 with the ICL-containing replication fork

To strengthen our findings, we further investigated the role of the Tower domain in the DNA binding. Since ICLs could occur at any place on the chromatin, it is less likely that the Tower domain recognize any particular DNA sequence. We carefully examined the amino acid sequence of the Tower domain from 19 species, and found two highly conserved patches of positively-charged residues, one formed by arginine 1236 (R1236) and lysine 1247 (K1247), and the other composed of lysine 1296 (K1296), arginine 1299 (R1299) and lysine 1307 (K1307) (Figure 3-11A to C). Both patches of positively-charged residues are on the inner surface of the Tower domain in the crystal structure of mFANCD2/FANCI, potentially making contacts with the DNA (Figure 3-12A). When we had a closer look at the docking of our cryo-EM structure to the crystal structure of FANCD2/FANCI, we noted that K1296, R1299 and K1307 are still facing the inner groove of the FANCD2/FANCI complex, while K1247 is rotated (Figure 3-12B). To test whether these positively charged residues play a crucial role in the DNA binding of the Tower domain, we substituted them with the neutral amino acid alanine, purified the recombinant FANCD2-R1236A/K1247A/FANCI and FANCD2-K1296A/R1299A/K1307A/FANCI heterodimers (Figure 3-13A), and examined their DNA

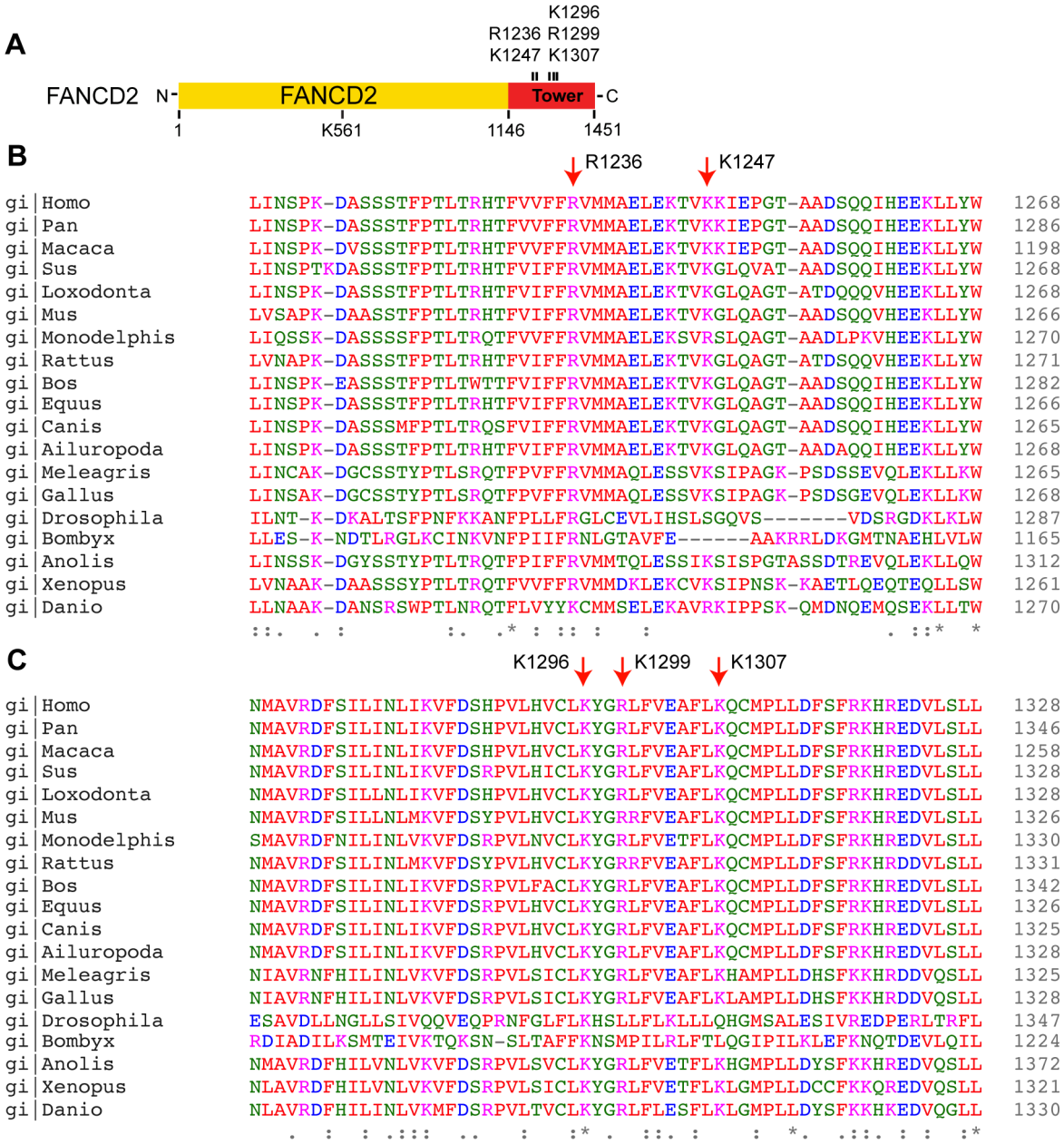


Figure 3-11. Positively charged residues in the Tower domain.
A) Schematic of FANCD2 indicating the positive residues, R1236, K1247, K1296, R1299 and K1307, in the Tower domain. B) Sequence alignment of the C-terminal region of FANCD2 showing the conservation of the positive residues, R1236 and K1247 in the Tower domain. C) Sequence alignment of the C-terminal region of FANCD2 showing the conservation of the positive residues, K1296, R1299 and K1307, in the Tower domain. Sequence alignments were performed using Clustal Omega.

binding activity. Mutation of the first positively-charged residue patch (R1236A/K1247A) led to an appreciable decrease in the DNA binding activity (Figure 3-13B, lanes 5 and 6). More strikingly, mutations in the second positively-charged residue patch (K1296A/R1299A/K1307A) were sufficient to drastically reduce the DNA binding activity of

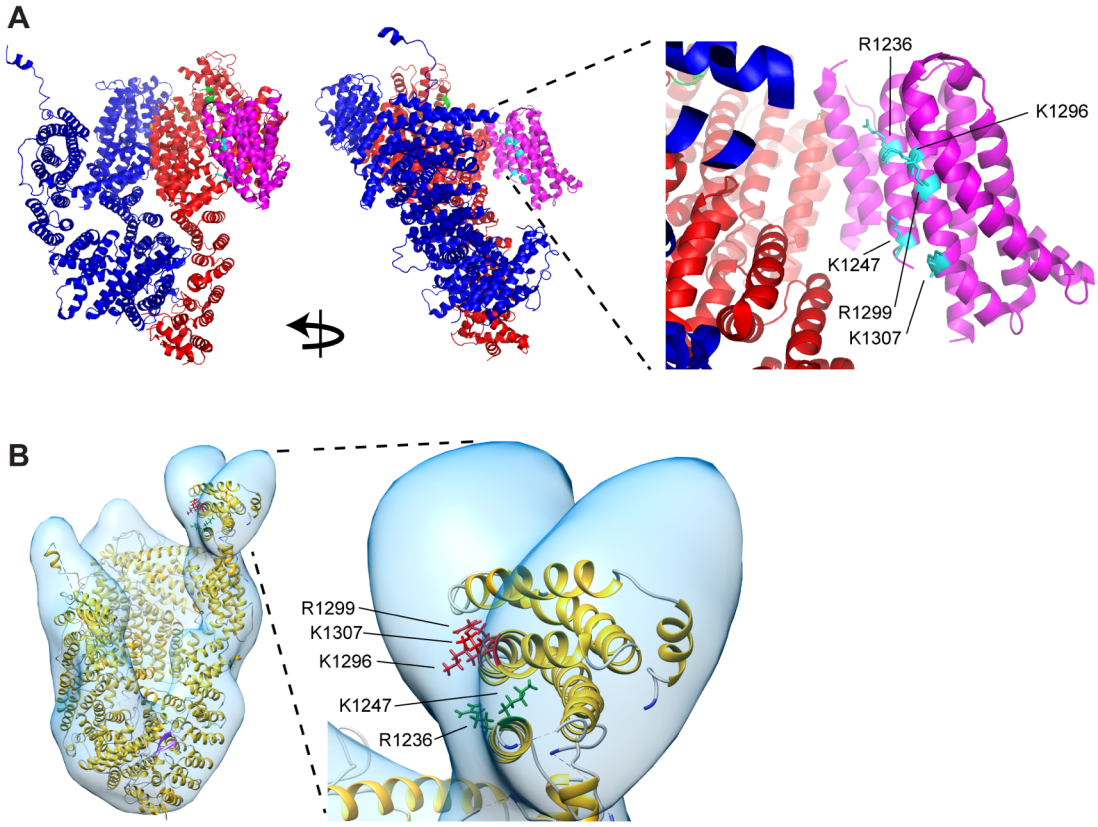


Figure 3-12. The position of the positively charged residues in the Tower domain.
A) Mapping of the position of the positively charged residues in the mouse FANCD2/FANCI crystal structure. Blue: FANCI; Red: FANCD2; Magenta: Tower domain; Cyan: positively charged residues. B) Zoom-in view of the Tower domain structure obtained by docking the cryo-EM structure of human FANCD2 into the crystal structure of the murine FANCD2/FANCI complex from Joo *et al.* Positively charged residues forming the two patches are shown as sticks indicating the position of the positive residues.

the FANCD2/FANCI complex (Figure 3-13B, lanes 7 and 8) to level similar to the Tower deletion mutant (Figure 3-6A and B), suggesting that these three positively-charged residues, K1296, R1299 and K1307, are required for DNA binding. When we tested the FANCD2-K1296A/R1299A/K1307A/FANCI complex in the *in vitro* ubiquitination assay, we observed dramatic reduction in the ubiquitination of the complex upon addition of DNA into the reaction (Figure 3-13C), confirming the crucial DNA binding role of the Tower domain in the monoubiquitination of FANCD2.

Encouraged by our *in vitro* biochemical data, we set out to test if these three positively-charged residues, K1296, R1299 and K1307, are functionally important for the activities of the

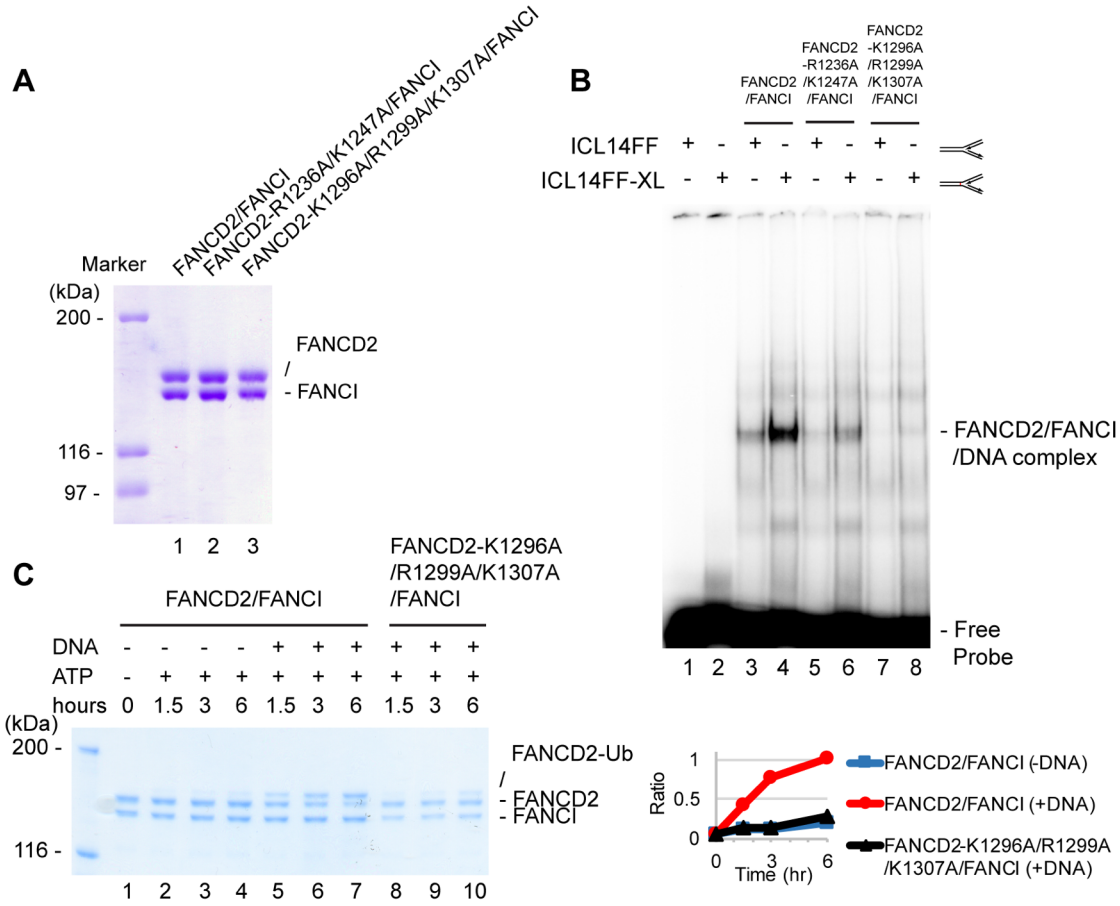


Figure 3-13. Electrostatic interactions underlie the binding of the FANCD2/FANCI complex to an arrested replication fork at an ICL.
A) Coomassie blue staining of purified recombinant wild type, R1236A/K1247A and K1296A/R1299A/K1307A FANCD2/FANCI complex. B) EMSA showing the R1236A and K1247A mutant has weaker DNA binding compared to the wild type complex, and K1296A, R1299A and K1307A mutant has even stronger defect in DNA binding. C) *In vitro* ubiquitination assay of full-length FANCD2/FANCI complex and FANCD2-K1296A/R1299A/K1307A/FANCI complex.

FANCD2/FANCI complex *in vivo*. We introduced EGFP-tagged FANCD2 and FANCD2-K1296A/R1299A/K1307A together with mCherry-tagged UHRF1 into CRISPR/Cas9-mediated FANCD2 knockout HeLa cells (HeLa-FANCD2 -/-). In the clonogenic survival assay, FANCD2-K1296A/R1299A/K1307A failed to rescue the MMC hypersensitivity caused by loss of FANCD2 (Figure 3-14A). In good agreement with the lack of recruitment of FANCD2-ΔTower, we could not detect any major fraction of FANCD2-K1296A/R1299A/K1307A recruited to ICLs (Figure 3-14B). Last but not least, we failed to detect any monoubiquitinated form of FANCD2-K1296A/R1299A/K1307A upon TMP/UVA treatment in whole cell lysate (Figure 3-14C).

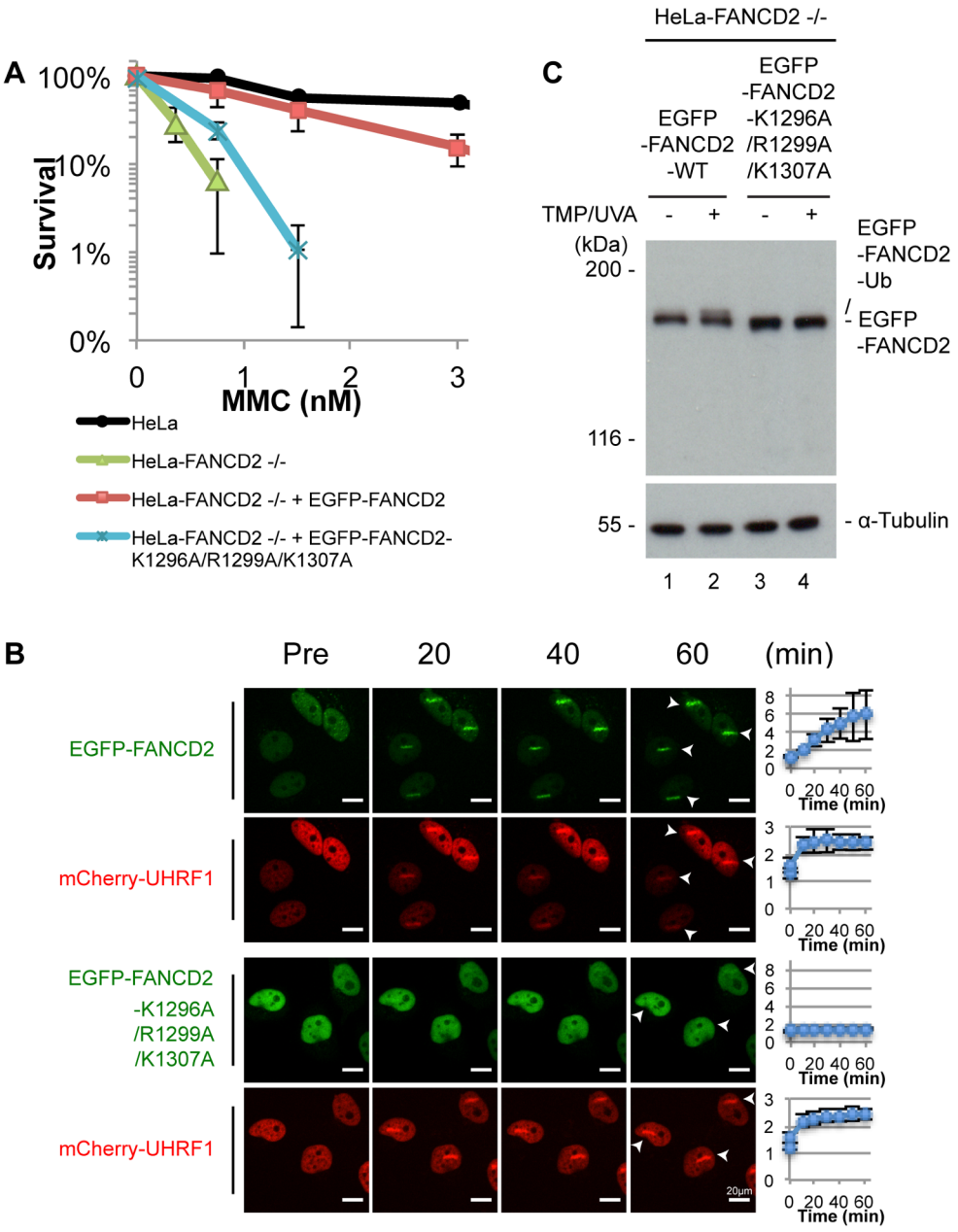


Figure 3-14. Residues K1296, R1299 and K1307 of the Tower domain of FANCD2 are highly critical for the function of FANCD2 *in vivo*.

A) Clonogenic survival assay of HeLa-FANCD2 -/- cells complemented with EGFP-FANCD2 or EGFP-FANCD2-K1296A/R1299A/K1307A. Experiment was done in triplicate. Error bars show standard deviation. B) HeLa-FANCD2 -/- cells expressing EGFP-FANCD2 or EGFP-FANCD2-K1296A/R1299A/K1307A were pre-treated with TMP (pre), and micro-irradiated at the indicated areas (white arrows). mCherry-UHRF1 was co-expressed in the same cells. EGFP-FANCD2 was recruited to ICLs, whereas the recruitment of EGFP-FANCD2-K1296A/R1299A/K1307A was strongly diminished. Representative fields of view are shown. Quantifications of the fluorophore-tagged proteins at the irradiated sites are shown on the right (7 cells for each cell line were quantified). Scale bar: 20μm. Error bars show standard deviation. C) HeLa-FANCD2 -/- cells complemented with EGFP-FANCD2 and EGFP-FANCD2 K1296A/R1299A/K1307A were treated with 500ng TMP, irradiated with 250J/m^2, 365nm UVA. 6 hours after irradiation, cells were harvested, lysed and analysed through immunoblotting using the indicated antibodies.

3.4 Discussion

In this chapter, we showed the cryo-EM structure of full-length human FANCD2/FANCI heterodimer at a modest resolution, which led to the identification of a functional Tower domain at the C-terminal of FANCD2. The Tower domain contains two conserved patches of positively-charged residues that can form a set of electrostatic interactions with negative charges of DNA. *In vitro* and *in vivo* analysis of the interaction between the Tower domain and DNA highlighted that the Tower domain plays an important role for the recruitment to ICLs and monoubiquitination of the FANCD2/FANCI complex. In addition, we demonstrated that the FANCD2/FANCI heterodimer preferably bound to a DNA substrate mimicking a replication fork arrested at an ICL. Collectively, our findings provide evidence that the FANCD2/FANCI complex is recruited to damaged chromatin prior to its monoubiquitination, in contrast to the current model, which foreshadowed the recruitment of FANCD2/FANCI to damaged chromatin only after monoubiquitination.

The discrepancy between cryo-EM and crystal structure of FANCD2/FANCI heterodimer

The main discrepancy between the crystal structure and cryo-EM structure is that the pseudo-two-fold symmetry observed in the mFANCD2/FANCI complex is lost in our cryo-EM structure of hFANCD2/FANCI complex, which leads to the re-position of the Tower domain. Interestingly, there were several deletions within FANCD2 and FANCI introduced to facilitate the crystallization, one of which includes an internal FANCD2 truncation (residues 851 to 910) (Joo et al., 2011). When we examine the docking of the crystal and cryo-EM structures, we found that this stretch of deleted amino acids situates at the base of the Tower domain (Figure 3-15). Therefore, it is possible that this truncation may affect the structure of FANCD2, thus justifying the discrepancy between the mouse crystal structure and the human full-length cryo-

EM structure of the FANCD2/FANCI complex. Another plausible reason for such disagreement between the two structures could be due to the intrinsic nature of the two techniques used in the studies, which may lead to bias against one conformation of the protein structure to another.

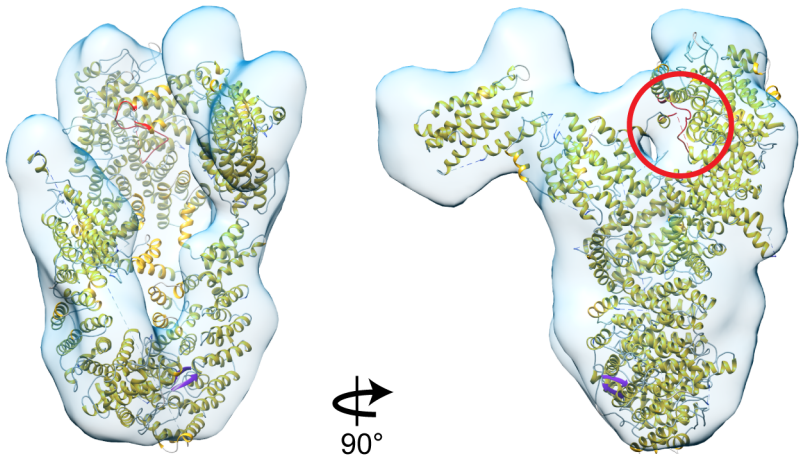


Figure 3-15. The position of the FANCD2 internal deletion in the mouse crystal structure of the FANCD2/FANCI complex.
Different orientations of cryo-EM density map of full-length hFANCD2/FANCI complex docked with the mFANCD2/FANCI crystal structure (PDB: 3S4W). The amino acid residues before and after the internal FANCD2 truncation (residues 851 to 910) are marked in red, and highlighted with red circle.

The DNA binding activity of FANCD2 and FANCI

It is known that both FANCD2 and FANCI possess some DNA binding activity as a monomer (Longerich et al., 2009; Park et al., 2005), and that the FANCD2/FANCI complex interacts with various of DNA substrates as well (Joo et al., 2011). However, it is commonly accepted that the FANCD2 and FANCI functions as a heterodimer in the ICL repair *in vivo*, disregarding the fact that the monomers and heterodimer have different DNA substrate preferences *in vitro*. In this chapter, we demonstrated that FANCD2/FANCI preferably bound to a DNA molecule resembling a replication fork arrested at an ICL. Such preference is supported by our observation that FANCD2 is recruited to the ICLs *in vivo*, and by the requirement of replication to activate the FA pathway to repair ICLs in the *Xenopus* cell free system (Raschle et al., 2008).

Revised model of the FA pathway activation

We have demonstrated the importance of the electrostatic interaction between the Tower domain and DNA for the activation of the FA pathway. Previous conclusions implicated that the monoubiquitination of FANCD2 precedes its recruitment to damaged chromatin. Such a mechanism was proposed following the observation that the FANCD2-K561R mutant, which cannot be monoubiquitinated, was unable to form damage-induced FANCD2 foci (Garcia-Higuera et al., 2001; Smogorzewska et al., 2007). However, we have shown that the FANCD2-K561R can be recruited to the laser directed ICLs. Therefore, by integrating our data and the literature together, we propose that the FANCD2/FANCI heterodimer is initially recruited to damaged chromatin. The interaction between FANCD2/FANCI complex and DNA could lead to a conformational change in the structure of the FANCD2/FANCI complex, allowing the heterodimer to be monoubiquitinated on DNA. Monoubiquitination of the FANCD2/FANCI complex may facilitate a consecutive structure conformational change of the heterodimer, ensuring its retention on chromatin (Figure 3-16). Other ubiquitin binding proteins may also be recruited to the damaged chromatin after the monoubiquitination of the FANCD2/FANCI complex to facilitate the retention of the FANCD2/FANCI heterodimer.

In our *in vitro* ubiquitination assay, we observed that, independent of the addition of DNA in the reaction, only a minor fraction of FANCI was ubiquitinated (Figure 3-7A), which is consistent with published data (Longerich et al., 2014; Rajendra et al., 2014; Sato et al., 2012), while FANCI is monoubiquitinated upon MMC treatment *in vivo* (Figure 3-4A). Intriguingly, it is also reported that the monoubiquitination of FANCI does not play a dominant role in the ICL repair *in vivo* (Smogorzewska et al., 2007). Therefore, we speculate that the monoubiquitination of FANCD2 is sufficient to lead to a conformational change to enhance the retention of the FANCD2/FANCI heterodimer on the chromatin. Further cryo-EM studies using state of the art

microscope could be informative, if we could capture different FANCD2/FANCI heterodimer structures depicted in our revised model.

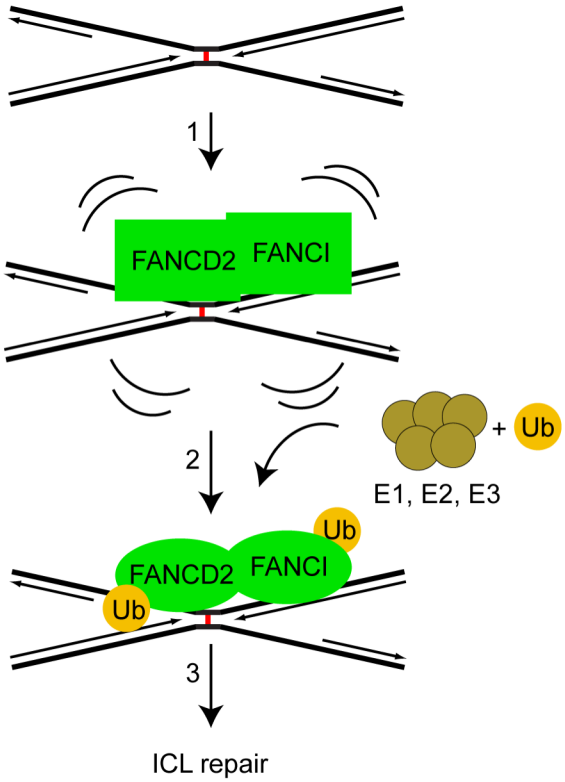


Figure 3-16. Proposed mechanism of FANCD2/FANCI *in vivo* mode of action
Schematic representation of the proposed mechanism of action of the FANCD2/FANCI heterodimer. The FANCD2/FANCI complex is (1) recruited to a crosslinked replication fork, and upon interaction with DNA, (2) it triggers the subsequent monoubiquitination of the complex. Monoubiquitination of FANCD2/FANCI (3) induces an additional conformational change that enhances the retention of complex on chromatin, favouring the ICL repair.

Chapter 4 Discussion

Since Fanconi anemia was first described by Swiss paediatrician Guido Fanconi in 1927, 21 FA genes have been identified. While the occurrence of FA is less than one in million worldwide, studies of FA have led to the discovery of an important DNA damage repair pathway, the FA pathway, which orchestrates homologous recombination and translesion synthesis to facility the repair of ICLs in the cells. Notably, DNA interstrand crosslinking agents, such as MMC and cisplatin, are widely used as anti-cancer drugs. In some cases, the initial treatment using DNA interstrand crosslinking agents successfully reduces the tumour size, but cancer relapses later due to expansion of the DNA interstrand crosslinking agents-resistant cancer cells. Our knowledge of the molecular basis of the FA pathway could help identify novel therapeutic targets, which by combination of inhibition of the FA pathway to re-sensitize the cancer cells to DNA interstrand crosslinking agents.

In this thesis, firstly, I described the identification and characterisation of a novel ICL sensor, UHRF1, which recognises ICLs *in vitro* and *in vivo*, and facilitates the recruitment of FANCD2 to ICLs (Liang et al., 2015). A study published around the same period suggests that UHRF1 plays a role in recruiting nucleases to ICLs independent of the FA pathway and SLX4 (Tian et al., 2015), while we showed that the activation of the FA pathway is defected in the absence of UHRF1. The discrepancy of the interplay between UHRF1 and the FA pathway in these two studies could result from different level of UHRF1 depletion achieved. Despite the difference in the conclusion, the two studies collectively demonstrate that UHRF1 is an ICL sensor, and plays an important role in the ICL repair. Secondly, we described the cryo-EM structure of the human FANCD2/FANCI complex, which led to the discovery of the functionally important Tower domain at the C-terminal of FANCD2 required for the DNA binding activities of the FANCD2/FANCI heterodimer. This study also reveals that the recruitment of the

FANCD2/FANCI complex to the ICLs precedes the monoubiquitination of the FANCD2/FANCI complex (Liang et al., 2016). I summarised our findings described in this thesis in an updated model focusing on the recognition of ICLs and the recruitment of the FANCD2/FANCI complex (Figure 4-1). There are still some missing pieces to complete the puzzle of the FA pathway. In this chapter, I will discuss some of the unaddressed questions in the FA pathway.

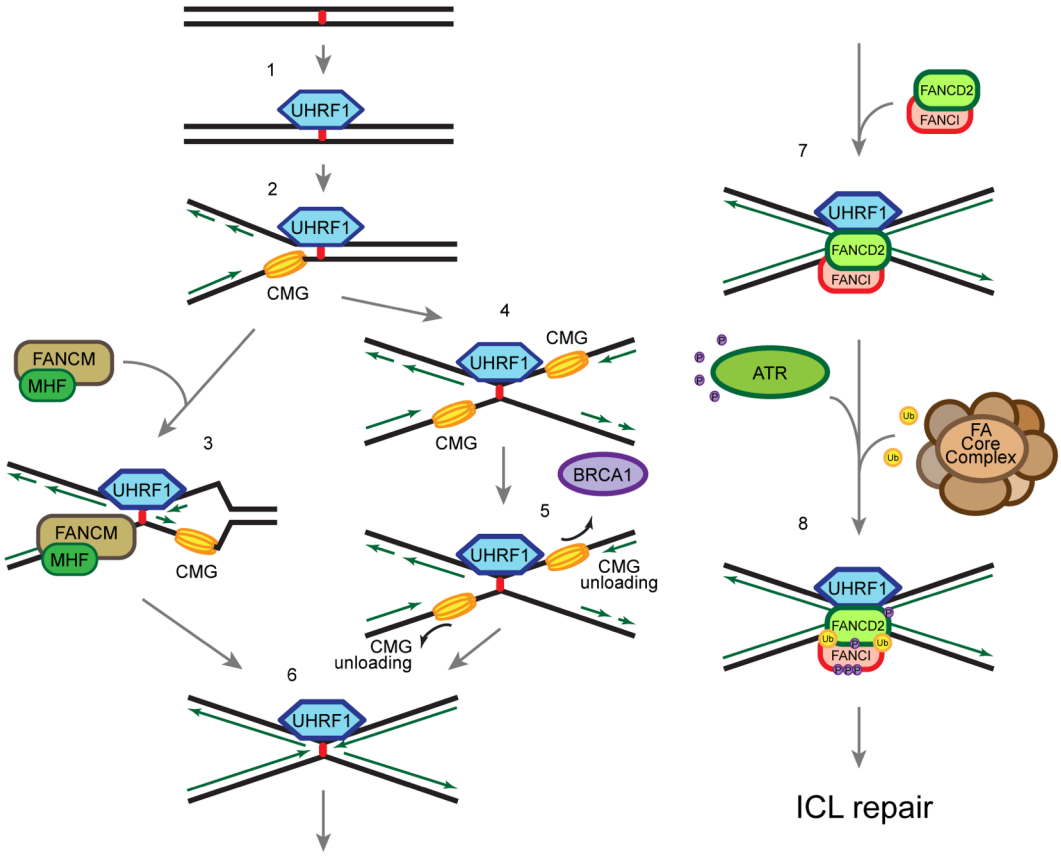


Figure 4-1. Revised model of ICL repair.
(1) UHRF1 recognises ICLs on chromatin. (2) Single replication fork arrives at the ICL. (3) FANCM/MHF complex mediates the traverse of the replication machinery through the ICL, which allows the replication fork to proceed, leaving the ICL for later repair. (4) Alternatively, BRCA1 (FANCS) facilitates the unloading of the CMG helicase complex when the second replication fork arrives at the ICL. (5) The replicative polymerase proceeds to the -1 position of the ICL, which leaves an X-shape structure similar to the traverse mechanism. (6) FANCD2/FANCI is recruited to the ICL at the replication fork. (7) ATR phosphorylates the FANCD2/FANCI complex at multiple sites, whereas the FA core complex monoubiquitinates the FANCD2/FANCI complex at K561 and K523 respectively. Ubiquitinated FANCD2/FANCI complex recruits downstream repair proteins to the damaged chromatin to remove the ICL.

Different DNA damages

The FA pathway is activated upon various DNA damaging agents, such as MMC, Cisplatin, HU, IR and UVC (Garcia-Higuera et al., 2001), and hence they have all been widely used to induce the monoubiquitination of FANCD2 in the literature. However, the damages induced by these agents are fundamentally different from each other. While FANCD2 is monoubiquitinated upon treatment, other PTMs could occur specifically in response to certain types of damage to facilitate different repair processes. One example is the distinct role of the FA pathway in the stalled replication fork protection upon HU treatment, which we will discuss in details later. Intriguingly, DNA interstrand crosslinking agents, TMP/UVA, MMC and Cisplatin, have been considered posing similar cellular toxicities. Yet, the distortion on the DNA double helix created by these chemicals are different, which potentially could be recognised by different ICL sensor(s). It would be interesting to explore if there are proteins recognising Cisplatin or other DNA crosslinking agents-induced ICLs, and facilitate their repair.

Replication-independent ICL repair

In addition to the canonical replication-coupled ICL repair, there is increasing evidence for the existence of a replication-independent ICL repair mechanism. It has been shown that ICLs can be removed via the NER pathway independently of replication (Enoiu et al., 2012; Muniandy et al., 2009; Sarkar et al., 2006; Wang et al., 2001). NER proteins XPA, XPB, XPC and XPF are recruited to psoralen/UVA induced ICLs and monoadducts within minutes in the G1-phase of the cell cycle and this ICL repair process is dependent on Polζ and XPC (Muniandy et al., 2009; Sarkar et al., 2006). It has also been shown that transcription-coupled NER (TC-NER) is crucial for cisplatin induced ICL repair (Enoiu et al., 2012; Wang et al., 2001).

The detailed molecular mechanism of the replication independent ICL repair remains elusive. It is likely that in the absence of arrested replication forks, the activation of ICL repair relies on

the proteins recognizing the distorted DNA helix and/or the collision of the RNA transcription machinery with an ICL. However, there are some contradictions in the literature whether XPC is involved in the replication independent repair process. The DNA crosslinking agents used in the studies mentioned above differ from study to study. As discussed previously, these crosslinking agents create structurally distinct ICLs, which could give rise to the discrepancies observed. It would be interesting to understand if the replication independent ICL repair communicates with the replication-coupled ICL repair mechanism, and if some of the FA proteins are also involved in the replication-independent ICL repair. Last but not least, replication independent ICL repair may play a significant role in maintaining the genome integrity in non-cycling cells, such as neurons.

Emerging role of the FA pathway in stalled replication fork protection

In addition to ICL repair, more evidence shows that the FA pathway cooperates with BRCA1/BRCA2, and plays an important role in the protection of stalled replication forks (Howlett et al., 2005; Kais et al., 2016; Michl et al., 2016; Schlacher et al., 2012; Wang et al., 2008). It has been shown that FANCD2 and FANCI interact with the MCM helicase at nascent DNA in response to replication arrest independent of the monoubiquitination on FANCD2 (Lossaint et al., 2013), and that ATR mediated phosphorylation of FANCI regulates dormant origin firing (Chen et al., 2015). Furthermore, Fanconi-associated nuclease 1 (FAN1) has also been shown to participate in the stalled replication fork protection, where the recruitment of FAN1 is dependent on monoubiquitinated FANCD2 (Lachaud et al., 2016). It is suggested that FANCD2 and BRCA1/2 protect the replication fork from the exonuclease MRE11 through facilitating the RAD51 filament formation (Schlacher et al., 2011; Schlacher et al., 2012). It is still unclear how FANCD2 orchestrates the order of events of the fork protection, and if FANCI plays a similar role in this process. Furthermore, it would also be interesting to know if other

PTMs on FANCD2 are required for this process, and if other repair proteins participate in the fork protection in addition to their already known roles.

Chapter 5 Material and methods

Cell lines and antibodies

HeLa (originally from ATCC) and PD20 (kindly provided by the FARF repository) cells were grown in DMEM (D5796, Sigma) supplemented with 2.5–10% fetal bovine serum. Antibodies used were as follows: anti-FANCD2, 1:100 dilution (sc-20022, Santa Cruz Biotechnology); anti-FANCI, 1:1,000 dilution (FARF); anti-FANCA, 1:1,000 dilution (FARF); anti-Lamin B, 1:1,000 dilution (sc-6216, Santa Cruz Biotechnology); anti-Flag, 1:1,000 dilution (M5, F4042, Sigma-Aldrich); anti-HA (mouse monoclonal antibody clone 12CA5); anti-Ubiquitin, 1:400 dilution (FK2, BML-PW8810, Enzo Life Sciences); anti-UHRF1, 1:1,000 dilution (sc-373750, Santa Cruz Biotechnology); and anti-a-Tubulin, 1:2,000 dilution (5829, Millipore).

Plasmids and transfection

UHRF1 and FANCD2 were expressed using the pOZ-N (FLAG-HA tag), pOZ-EGFP and pOZ-mCherry plasmid (Nakatani and Ogryzko, 2003). shRNA-mediated knockdown of the UHRF1 and FANCD2 gene was achieved by expressing the target sequence 5'-AGATATAACGTTAGGGTTT-3' and 5'-GAGCAAAGCCACTGAGGTA-3', respectively, in the pSuper.retro vector (Clontech). The deleted amino acids in UHRF1-ΔUBL, ΔTTD, ΔPHD, ΔSRA and ΔRING were 14-91, 146-296, 323-379, 427-630 and 724-806, respectively, and amino acids 1,147-1,451 was deleted in the FANCD2-ΔTower. The domain deletion plasmids were generated by a reverse PCR removing desired sequence followed by a blunt-end ligation. Transfections of plasmid DNA were carried out using FuGENE6 (Promega) according to the manufacturer's instructions.

CRISPR-Cas9 gene editing

The targeting sequence used in the sgRNA was: 5'-GTTTGTCTTGTGAGCGTCTGC-3' targeting the fourth exon of FANCD2. HeLa-FANCD2 -/- cells were generated using plasmid pX459 (Addgene #48139, expressing Cas9 from S. pyogenes with 2A-Puro) as follows 20: primers, 5'-CACCGTTTGTCTTGTGAGCGTCTGC-3' and 5'-AAACGCAGACGCTCAC AAGACAAAC-3', were annealed and introduced into the pX459 plasmid through its BbsI restriction site. HeLa cells were transfected with 2 mg of the resulting pX459 plasmid and selected with 4 mg/ml puromycin after 24 h. After another 24 h, cells were plated at low density and clones were picked after 2 weeks. Individual clones were tested, and the knocked out of FANCD2 was confirmed by western blotting.

Preparation of interstrand crosslinked DNA substrates

The DNA oligos were annealed in a buffer containing 10mM Tris-HCl pH 7.5, 100mM NaCl and 1mM EDTA. 4,50,8-trimethylpsoralen (TMP, Sigma, T6137) crosslinking was carried out using 2.6 mM annealed DNA in crosslinking buffer (10mM Tris-HCl pH 7.5, 1mM EDTA, and 50mM NaCl) with 87.6 mM TMP. The reaction was performed using a UVA source ($4.2mW/m^2$) wavelength of 365nm for 6 cycles at 15-min intervals. Our protocol is a modification of a previously published method (Esposito et al., 1988). ICL was confirmed by 8M urea 12% denaturing polyacrylmide gel electrophoresis. ICL14FF was created by annealing the following DNA oligonucleotides (the bases forming of the ICL are indicated in bold):

ICL14 (+):

5'-CATTGTGAATTCGCCTCTCTGTC**TA**GCCGAAGCTCGAAACGATCTTGTGC-3'

ICL14 (-):

5'-GTCCATCAAAGTTCGACTGTGCGGC**TA**GACAGAGAGGCGAATTCACAATG-3'

ICL14 leading: 5'-GCACAAGATCGTTTCGAGCTT-3'

ICL14 lagging: 5'-GTCGAACTTTGATGGAC-3'

ICL-binding protein purification and UHRF1/DNA *in vitro* binding

Nuclear proteins were extracted as described (Dignam et al., 1983). 1mg nuclear extract from HeLa cells or 3 µg of recombinant UHRF1/FANCL from Sf9 cells were incubated with 25pmol biotin-labeled DNA substrates (ICL1 and ICL1-XL were used for purification from nuclear extract; ICL8 and ICL8-XL were used for UHRF1 *in vitro* binding experiments). Binding buffer contained 5mM Tris, pH 7.9, 30mM KCl, 1mM dithiothreitol (DTT), 10mM Hepes-KOH pH 7.9, 1mM EDTA, 5% glycerol and 0.3mg/ml BSA (New England Biolabs). The protein and DNA probe were mixed together and the reaction mixture was incubated at 30°C for 15min, and mixed with 10µl Streptavadin sepharose 50% slurry (GE Healthcare). The matrix with streptavadin beads were washed with several column volumes before elution. Eluted proteins were analyzed by electrophoresis on a 4-12% NuPage Bis-Tris gradient gel (Life Technologies), and visualized by silver stain (Silver Quest, Life Technologies). For UHRF1 and FANCL *in vitro* DNA binding, the proteins were analyzed by electrophoresis on 10% SDS-PAGE gel followed by Western blot analysis.

Mass spectrometric analysis

Proteins were reduced with DTT, cysteine residues were acetylated with iodoacetamide, and the proteins were separated by SDS-PAGE followed by silver staining. Proteins from silver stained gel bands were in-gel digested with trypsin (Shevchenko et al., 1996). The generated peptide mixtures were subjected to LC-MS/MS using a hybrid linear ion trap/ FT-ICR mass spectrometer (LTQ FT, Thermo Electron) essentially as described previously (Haas et al., 2006). MS/MS spectra were assigned by using the SEQUEST algorithm (Eng et al., 1994) against the human International Protein Index sequence database.

Protein purification

UHRF1, FANCD2, FANCI and UBA1 were expressed using the pFastBac1 vector (Life Technologies) with an engineered N-terminal Flag-HA tag or pFastBac-HT vector (6xHis tag). FANCL was expressed using the pFastBac1 vector (Life Technologies) with an engineered N-terminal Flag-MBP tag or a N-terminal Flag-HA tag. For FANCD2/FANCI complex, Sf9 cell pellets were re-suspended in lysis buffer (20mM Tris-HCl pH 8.0, 0.1M KCl, 10% glycerol and 0.2mM PMSF), and sonicated. Lysates were clarified by centrifugation (17,000g), and the supernatants were incubated with the M2 anti-FLAG agarose resin (A2220, Sigma) for 2 h. The resin was washed carefully, and the protein was eluted in the lysis buffer containing 0.5 mg/ml FLAG peptide. The Flag eluate was supplemented with 20mM Imidazole and incubated with Ni2.-NTA (30310, QIAGEN) at 4°C for 2 h with end-over-end rotation. The resin was washed carefully, and retained protein was eluted with buffer containing 20mM Tris (pH 8.0), 0.1M KCl, 250mM Imidazole, 0.2mM PMSF and 10% glycerol. The Ni^{2+}-NTA eluate was injected into pre-equilibrated size-exclusion chromatography column, Superdex 200, and proteins were eluted with the base buffer containing 20mM Tris (pH 8.0), 0.1M KCl and 5% glycerol. The FANCD2-ΔTower/FANCI complex wass purified as the full-length FANCD2/FANCI except for the size-exclusion chromatography step. UHRF1, FANCD2, UBA1 and FANCL were purified against FLAG tag as described above. UBE2T expressing plasmid was a kind gift from Dr. Helen Walden, where in brief, UBE2T was cloned into a pET based expression vector containing a N-terminal 6xHis-Smt3 tag. UBE2T was expressed in E. coli BL21(DE3) (Novagen). Cells were grown in Lysogeny Broth supplemented with antibiotics at 37°C. Expression of UBE2T was induced at OD600 0.6 with 0.5mM IPTG, and cells were cultured overnight at 16°C. Harvested cells were lysed by using a french press in the buffer containing 0.5M NaCl, 0.1M Tris (pH 8), 0.02M Imidazole, and 0.25mM TCEP. The lysate was clarified by centrifugation (48,000g), and the supernatant was incubated with Ni^{2+}-NTA agarose

(QIAGEN) for 2 hours at 4°C. 6xHis-Smt3 tag was removed overnight at 4°C by Ulp1 protease at a Ulp1:UBE2T w/w ratio of 1:15, Ulp1:UBE2T. UBE2T was concentrated, and loaded onto a Superdex 200 column with 200mM NaCl, 0.1M Tris (pH 8) and 10% glycerol. The described UBE2T purification is a derivative of a previously published method (Hodson et al., 2014).

Electrophoretic mobility shift assay (EMSA)

The UHRF1/DNA binding reaction that contained 1 μg of UHRF1 and 1 nM radiolabeled DNA, was performed in 10μl containing 14mM Tris-HCl pH 8.0, 100mM NaCl, 3.4% glycerol, 1mM dithiothreitol (DTT), 20ng poly(dI·dC)-poly(dI·dC) and 1μg bovine serum albumin (BSA, New England Biolabs), and 2 μg anti-HA antibody was added for super-shift.

The FANCD2/FANCI/DNA-binding reaction that contained 0.5 mg of FANCD2/FANCI complex, or derivative thereof, and 1 nM radiolabeled DNA, was performed in a final volume of 10 μl containing 25mM Tris-HCl pH 7.5, 100mM NaCl, 6% glycerol and 1mM DTT, and 0.5 μg anti-HA antibody was added for super-shift. For the experiment in Fig. 2-5C the following amounts of protein were used: 0, 0.125, 0.25 and 0.5 mg.

For both binding experiments, the reaction mixture was incubated at room temperature for 1 h, and incubated for another 1 h for super-shift. After the incubation, the reaction mixtures were loaded onto a 0.4xTBE, 4% polyacrylamide gel, and run with 0.4xTBE buffer at 4°C. The gel was then fixed with 10% methanol 10% acetic acid for 5min, dried and exposed to film or photostimulable phosphor imaging plate.

Fluorescence anisotropy

Recombinant Flag-tagged UHRF1 was incubated with 10nM ICL8 or ICL-XL (Eurofins Genomics) labeled with Alexa Flour 488 at the 5’ terminus in a buffer containing 20mM Tris-HCl pH8.0, 100mM KCl, 1mM DTT and 0.2 mg/ml BSA in 40μl on ice for 1 hour. A

PHERASTAR FS fluorimeter (BMG Labtech, λex= 490 nm, λem = 520 nm) was used to record the fluorescence polarization. Data were fitted using the Origin software (Origin Lab) with a one-site specific binding equation.

Cell fractionation and whole-cell lysate

Collected cell pellets were permeabilized with CSK buffer containing 200mM NaCl, 10 mM PIPES, 300 mM Sucrose, 1 mM $MgCl_2$, 1 mM EDTA and 0.5% Triton X-100 on ice for 10 min. CSK fraction (supernatant) and nuclear pellet were separated by centrifugation at 900g at 4°C for 10 min. Nuclear pellet was processed the same way as whole-cell lysate described below. Preparation of whole-cell lysate was performed by scraping off cells from dishes followed by centrifugation at 100g for 5 min. Cell pellets were re-suspended and incubated in Benzonase buffer (2 mM MgCl2, 20 mM Tris-HCl (pH 8.0), 10% glycerol, 1% Triton X-100 and 12.5U/ml Benzonase (E1014, Sigma-Aldrich) on ice for 10 min. Cells were then lysed by addition of an equal volume of 2% SDS to the final concentration of 1%. Samples were heated at 70°C for 2min. Protein concentration was determined by the Bradford assay (Bio-Rad Life Science).

Clonogenic survival assay

Cells (250–4,000) were plated in six-well plates and treated with different dosages of damaging agents indicated on the next day. Colony formation was scored after 10–14 days using 1% (w/v) crystal violet in methanol.

Co-immunoprecipitation

HeLa.Scramble, HeLa.shUHRF1 cells expressing FLAG-tagged UHRF1 or HeLa.shFANCD2 cells expressing FLAG-tagged FANCD2 cells were treated with TMP/UVA as described. Cell pellets were incubated with Buffer A (0.1% Triton X-100, 20mM Hepes pH7.9, 5mM $MgCl_2$, 10% Glycerol, 1unit/µl Benzonase and 2.5mg/ml DSP (D3669, Sigma)) for 30 minutes on ice. After 30 min of incubation Tris pH 8.0 was added to the mixture at the

final concentration of 0.2M to quench DSP. 10 times pellet volume of Buffer B (0.15% Triton X-100, 20mM Tris-HCl pH 8.0, 5mM $MgCl_2$, 10% Glycerol, 300mM KCl and 0.2mM PMSF) was added to the mixture and incubated for 10 minutes for extraction. Lysates were clarified by centrifugation, and supernatant was used for immunoprecipitation. M2 agarose beads were added to the lysates, and incubated for 2 hours. The resin was washed extensively, and eluted with 0.5mg/ml FLAG peptide.

PD20 and PD20-expressing FLAG-HA-tagged FANCD2 and FANCD2DTower cells were treated with 160 ng/ml MMC overnight. Cell pellets were incubated with Buffer A (0.5% Triton X-100, 20mM Tris pH 8.0, 2mM $MgCl_2$, 5mM $MgCl_2$, 10% Glycerol, 50U/ml Benzonase for 10 min on ice. Ten times the pellet volume of Buffer B (0.5% Triton X-100, 20mM Tris-HCl pH 8.0, 2mM $MgCl_2$, 10% Glycerol, 150mM KCl and 0.2mM PMSF) was added to the mixture and incubated for 10 min for extraction. Lysates were clarified by centrifugation (17,000g), and supernatant was used for immunoprecipitation. M2 agarose beads were added to the lysates, and incubated for 2 h. The resin was washed extensively, and eluted with 0.5 mg/ml FLAG peptide.

In vitro protein binding assay

The recombinant proteins were expressed and purified from Sf9 insect cells as indicated above. 1μg of each protein was mixed in the reaction buffer containing 10μg BSA (NEB), 20mM Tris-HCl pH 8.0, 150mM KCl, 10% Glycerol, 2mM β-ME and 0.2mM PMSF in 10μl. The mixture was first incubated at 37°C for 1h for protein complex formation. Subsequently, anti-HA sepharose beads 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. Proteins were eluted in a buffer contained 100mM Tris-HCl pH 6.8, 100mM KCl, 0.1% Tween 20, 0.2mM EDTA, 10% glycerol and 0.5 mg/ml HA peptide (Sigma).

Live-cell imaging

EGFP-fused wild-type and mutant FANCD2 or mCherry-fused UHRF1 cDNA were inserted into the pOZ vector as described above. Live-cell imaging was 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 was excited with 488-nm laser lines. Throughout the experiment, cells were maintained at 5% CO2, and 37°C using a live-cell environmental chamber (Tokai hit). Confocal image series were typically recorded with a frame size of 512 x 512 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 10% fetal bovine serum and 20 mg/ml 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) pre-selected 3–5 stripes (50 x 3 pixels) within the nucleus 10–60 times with a 405-nm laser set to 100% laser power. The EGFP and mCherry intensities at micro-irradiated sites were quantified using ImageJ with Fiji, and normalized by their intensities before microirradiation.

EM and 3D image processing

Negatively stained freshly prepared full length FANCD2/FANCI complex (200 ng/µl) was applied to glow-discharged, carbon-coated grids and allowed to adsorb for 15–60 s. Specimen was then stained with 2% uranyl acetate. Vitrified full length FANCD2/FANCI (200 ng/µl) and C-terminus deleted FANCD2/FANCI complex (150 ng/µl) was prepared on glow discharged carbon-coated Quantifoil grids. Specimen were imaged at a nominal magnification of 30,000× with a 2k×2k Gatan CCD camera (corresponding to a pixel size of 3.5 Å at the specimen level) in a JEOL 2100, LaB_6 operating at 200kV. A preliminary full length FANCD2/FANCI 3D model was calculated using the random conical tilt method (RCT) (Radermacher et al., 1987) and the WEB and SPIDER software package (Frank et al., 1996). For this first 3D model, 4082

pairs of particles of the full length FANCD2/FANCI complex were picked from 80 tilt pair images recorded at 50° and 0°. Refinement of the 3D volume obtained from the RCT method was performed using images from frozen hydrated full length FANCD2/FANCI complex. 5058 particles were selected after using Roseman's algorithm (Roseman, 2003) in SPIDER (Frank et al., 1996) procedure and subsequently manual selection in WEB (Frank et al., 1996). Defocus was determined using CTFTILT (Mindell and Grigorieff, 2003). 3D classification and refinement were carried out in RELION (Scheres, 2012), first with 1 class (25 iterations) and then with 4 classes (25 iteratiions) 3D classification. For the C-terminus deleted FANCD2/FANCI complex, frozen particles were selected using the Roseman's algorithm (Roseman, 2003) in a SPIDER (Frank et al., 1996) procedure using re-projection of the full-length FANCD2/FANCI structure as template, and then manually screened using the EMAN (Ludtke et al., 1999) program Boxer. 8547 particles were selected from 71 micrographs with a defocus range of 3-5.2 μm. Defocus was determined using CTFFIND3 (Mindell and Grigorieff, 2003). Refinement was carried out using the EMAN software package (Ludtke et al., 1999). The full-length FANCD2/FANCI complex structure was used as the starting model and refined against the C-terminus deleted FANCD2/FANCI particles. At each iteration, particles with bad cross-correlation values were temporally removed, and intermediate volume was band-pass filtered between 10 and 150 Å. For the resolution limit estimation of 3D-reconstruction volumes, two independent reconstructions were carried out and compared in reciprocal space using increasing shells with the FSC (Fourier shell correlation) technique. Visualization of 3D density map was done in UCSF Chimera (Pettersen et al., 2004).

Flexible fitting of crystal structure into electron microscopy map using molecular dynamics

The crystal structure of mFANCD2-CD/FANCI (3S4W) was fitted flexibly into the cryo-EM structure of full-length FANCD2/FANCI using a flexible fitting software MDFF (Molecular

Dynamics Flexible Fitting) (Trabuco et al., 2008) where solenoid (S1 to S4), HD1, and HD2 domains are set as individual rigid domains.

In vitro ubiquitination assay

Reaction volumes of 25 µl contained 17 nM UBA1, 0.64 mM UBE2T, 0.372 mM FANCL or FLAG-HA-FANCL, 4.2 mM 6xHis-Ub, 0.25 mM FANCD2/FANCI complex or derivatives thereof, 20mM pBlueScript SKII(+) when indicated, in the following reaction buffer: 50mM Tris (pH 7.5), 100mM KCl2, 2mM MgCl2, 0.5mM dithiothreitol and 2mM ATP. Reactions were incubated at room temperature for the indicated time. In all samples, 5 µl of 6x SDS loading buffer containing BME was used to terminate reactions. Samples were loaded onto an SDS-PAGE gel and subjected to Coomassie blue staining or immunoblotting. The described *in vitro* ubiquitination assay is a derivative of a previously published method (Hodson et al., 2014).

Chapter 6 References

Acharya, N., Haracska, L., Johnson, R.E., Unk, I., Prakash, S., and Prakash, L. (2005). Complex formation of yeast Rev1 and Rev7 proteins: a novel role for the polymerase-associated domain. Molecular and cellular biology *25*, 9734-9740.

Achour, M., Jacq, X., Ronde, P., Alhosin, M., Charlot, C., Chataigneau, T., Jeanblanc, M., Macaluso, M., Giordano, A., Hughes, A.D.*, 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*, 2187-2197.

Ameziane, N., May, P., Haitjema, A., van de Vrugt, H.J., van Rossum-Fikkert, S.E., Ristic, D., Williams, G.J., Balk, J., Rockx, D., Li, H.*, et al.* (2015). A novel Fanconi anaemia subtype associated with a dominant-negative mutation in RAD51. Nature communications *6*, 8829.

Andersen, S.L., Bergstralh, D.T., Kohl, K.P., LaRocque, J.R., Moore, C.B., and Sekelsky, J. (2009). Drosophila MUS312 and the vertebrate ortholog BTBD12 interact with DNA structure-specific endonucleases in DNA repair and recombination. Molecular cell *35*, 128-135.

Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y., and Shirakawa, M. (2008). Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. Nature *455*, 818-821.

Arita, K., Isogai, S., Oda, T., Unoki, M., Sugita, K., Sekiyama, N., Kuwata, K., Hamamoto, R., Tochio, H., Sato, M.*, et al.* (2012). Recognition of modification status on a histone H3 tail by linked histone reader modules of the epigenetic regulator UHRF1. Proc Natl Acad Sci U S A *109*, 12950-12955.

Auerbach, A.D. (2009). Fanconi anemia and its diagnosis. Mutat Res *668*, 4-10.

Avvakumov, G.V., Walker, J.R., Xue, S., Li, Y., Duan, S., Bronner, C., Arrowsmith, C.H., and Dhe-Paganon, S. (2008). Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. Nature *455*, 822-825.

Bai, X.C., McMullan, G., and Scheres, S.H. (2015). How cryo-EM is revolutionizing structural biology. Trends Biochem Sci *40*, 49-57.

Berkyurek, A.C., Suetake, I., Arita, K., Takeshita, K., Nakagawa, A., Shirakawa, M., and Tajima, S. (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. The Journal of biological chemistry *289*, 379-386.

Bluteau, D., Masliah-Planchon, J., Clairmont, C., Rousseau, A., Ceccaldi, R., Dubois d'Enghien, C., Bluteau, O., Cuccuini, W., Gachet, S., Peffault de Latour, R.*, et al.* (2016). Biallelic inactivation of REV7 is associated with Fanconi anemia. J Clin Invest *126*, 3580-3584.

Bogliolo, M., Schuster, B., Stoepker, C., Derkunt, B., Su, Y., Raams, A., Trujillo, J.P., Minguillon, J., Ramirez, M.J., Pujol, R.*, et al.* (2013). Mutations in ERCC4, Encoding the DNA-Repair Endonuclease XPF, Cause Fanconi Anemia. American journal of human genetics *92*, 800-806.

Bostick, M., Kim, J.K., Esteve, P.O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science *317*, 1760-1764.

Bridge, W.L., Vandenberg, C.J., Franklin, R.J., and Hiom, K. (2005). The BRIP1 helicase functions independently of BRCA1 in the Fanconi anemia pathway for DNA crosslink repair. Nature genetics *37*, 953-957.

Budzowska, M., Graham, T.G., Sobeck, A., Waga, S., and Walter, J.C. (2015). Regulation of the Rev1-pol zeta complex during bypass of a DNA interstrand cross-link. The EMBO journal *34*, 1971-1985.

Ceccaldi, R., Sarangi, P., and D'Andrea, A.D. (2016). The Fanconi anaemia pathway: new players and new functions. Nat Rev Mol Cell Biol *17*, 337-349.

Chaudhury, I., Sareen, A., Raghunandan, M., and Sobeck, A. (2013). FANCD2 regulates BLM complex functions independently of FANCI to promote replication fork recovery. Nucleic acids research *41*, 6444-6459.

Chen, Y.H., Jones, M.J., Yin, Y., Crist, S.B., Colnaghi, L., Sims, R.J., 3rd, Rothenberg, E., Jallepalli, P.V., and Huang, T.T. (2015). ATR-mediated phosphorylation of FANCI regulates dormant origin firing in response to replication stress. Molecular cell *58*, 323-338.

Cheng, J., Yang, Y., Fang, J., Xiao, J., Zhu, T., Chen, F., Wang, P., Li, Z., Yang, H., and Xu, Y. (2013). Structural insight into coordinated recognition of trimethylated histone H3 lysine 9 (H3K9me3) by the plant homeodomain (PHD) and tandem tudor domain (TTD) of UHRF1 (ubiquitin-like, containing PHD and RING finger domains, 1) protein. The Journal of biological chemistry *288*, 1329-1339.

Cheng, Y. (2015). Single-Particle Cryo-EM at Crystallographic Resolution. Cell *161*, 450-457.

Ciccia, A., Ling, C., Coulthard, R., Yan, Z., Xue, Y., Meetei, A.R., Laghmani el, H., Joenje, H., McDonald, N., de Winter, J.P.*, et al.* (2007). Identification of FAAP24, a Fanconi anemia core complex protein that interacts with FANCM. Molecular cell *25*, 331-343.

Cybulski, K.E., and Howlett, N.G. (2011). FANCP/SLX4: a Swiss army knife of DNA interstrand crosslink repair. Cell Cycle *10*, 1757-1763.

de Winter, J.P., Leveille, F., van Berkel, C.G., Rooimans, M.A., van Der Weel, L., Steltenpool, J., Demuth, I., Morgan, N.V., Alon, N., Bosnoyan-Collins, L.*, et al.* (2000a). Isolation of a cDNA representing the Fanconi anemia complementation group E gene. American journal of human genetics *67*, 1306-1308.

de Winter, J.P., Rooimans, M.A., van Der Weel, L., van Berkel, C.G., Alon, N., Bosnoyan-Collins, L., de Groot, J., Zhi, Y., Waisfisz, Q., Pronk, J.C.*, et al.* (2000b). The Fanconi anaemia gene FANCF encodes a novel protein with homology to ROM. Nature genetics *24*, 15-16.

de Winter, J.P., Waisfisz, Q., Rooimans, M.A., van Berkel, C.G.M., Bosnoyan-Collins, L., Alon, N., Carreau, M., Bender, O., Demuth, I., Schindler, D.*, et al.* (1998). The Fanconi anaemia group G gene FANCG is identical with XRCC9. Nature genetics *20*, 281-283.

Deans, A.J., and West, S.C. (2009). FANCM connects the genome instability disorders Bloom's Syndrome and Fanconi Anemia. Molecular cell *36*, 943-953.

Deans, A.J., and West, S.C. (2011). DNA interstrand crosslink repair and cancer. Nat Rev Cancer *11*, 467-480.

Dignam, J.D., Martin, P.L., Shastry, B.S., and Roeder, R.G. (1983). Eukaryotic gene transcription with purified components. Methods Enzymol *101*, 582-598.

Doil, C., Mailand, N., Bekker-Jensen, S., Menard, P., Larsen, D.H., Pepperkok, R., Ellenberg, J., Panier, S., Durocher, D., Bartek, J.*, et al.* (2009). RNF168 Binds and Amplifies Ubiquitin Conjugates on Damaged Chromosomes to Allow Accumulation of Repair Proteins. Cell *136*, 435-446.

Dong, H., Nebert, D.W., Bruford, E.A., Thompson, D.C., Joenje, H., and Vasiliou, V. (2015). Update of the human and mouse Fanconi anemia genes. Hum Genomics *9*, 32.

Dorsman, J.C., Levitus, M., Rockx, D., Rooimans, M.A., Oostra, A.B., Haitjema, A., Bakker, S.T., Steltenpool, J., Schuler, D., Mohan, S.*, et al.* (2007). Identification of the Fanconi anemia complementation group I gene, FANCI. Cellular Oncology *29*, 211-218.

Elia, A.E., Boardman, A.P., Wang, D.C., Huttlin, E.L., Everley, R.A., Dephoure, N., Zhou, C., Koren, I., Gygi, S.P., and Elledge, S.J. (2015). Quantitative Proteomic Atlas of Ubiquitination and Acetylation in the DNA Damage Response. Molecular cell *59*, 867-881.

Eng, J.K., McCormack, A.L., and Yates, J.R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom *5*, 976-989.

Enoiu, M., Jiricny, J., and Scharer, O.D. (2012). Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. Nucleic Acids Res *40*, 8953-8964.

Esposito, F., Brankamp, R.G., and Sinden, R.R. (1988). DNA sequence specificity of 4,5',8-trimethylpsoralen cross-linking. Effect of neighboring bases on cross-linking the 5'-TA dinucleotide. The Journal of biological chemistry *263*, 11466-11472.

Felle, M., Joppien, S., Nemeth, A., Diermeier, S., Thalhammer, V., Dobner, T., Kremmer, E., Kappler, R., and Langst, G. (2011). The USP7/Dnmt1 complex stimulates the DNA methylation activity of Dnmt1 and regulates the stability of UHRF1. Nucleic Acids Res *39*, 8355-8365.

Feltes, B.C., and Bonatto, D. (2015). Overview of xeroderma pigmentosum proteins architecture, mutations and post-translational modifications. Mutat Res Rev Mutat Res *763*, 306-320.

Fradet-Turcotte, A., Canny, M.D., Escribano-Diaz, C., Orthwein, A., Leung, C.C., Huang, H., Landry, M.C., Kitevski-LeBlanc, J., Noordermeer, S.M., Sicheri, F.*, et al.* (2013). 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark. Nature *499*, 50-54.

Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. Journal of structural biology *116*, 190-199.

Garcia-Higuera, I., Kuang, Y., Naf, D., Wasik, J., and D'Andrea, A.D. (1999). Fanconi anemia proteins FANCA, FANCC, and FANCG/XRCC9 interact in a functional nuclear complex. Molecular and cellular biology *19*, 4866-4873.

Garcia-Higuera, I., Taniguchi, T., Ganesan, S., Meyn, M.S., Timmers, C., Hejna, J., Grompe, M., and D'Andrea, A.D. (2001). Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. Molecular cell *7*, 249-262.

Haas, W., Faherty, B.K., Gerber, S.A., Elias, J.E., Beausoleil, S.A., Bakalarski, C.E., Li, X., Villen, J., and Gygi, S.P. (2006). Optimization and use of peptide mass measurement accuracy in shotgun proteomics. Mol Cell Proteomics *5*, 1326-1337.

Hara, K., Hashimoto, H., Murakumo, Y., Kobayashi, S., Kogame, T., Unzai, S., Akashi, S., Takeda, S., Shimizu, T., and Sato, M. (2010). Crystal structure of human REV7 in complex with a human REV3 fragment and structural implication of the interaction between DNA polymerase zeta and REV1. The Journal of biological chemistry *285*, 12299-12307.

Hashimoto, H., Horton, J.R., Zhang, X., Bostick, M., Jacobsen, S.E., and Cheng, X. (2008). The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. Nature *455*, 826-829.

Helleday, T., Petermann, E., Lundin, C., Hodgson, B., and Sharma, R.A. (2008). DNA repair pathways as targets for cancer therapy. Nat Rev Cancer *8*, 193-204.

Hira, A., Yoshida, K., Sato, K., Okuno, Y., Shiraishi, Y., Chiba, K., Tanaka, H., Miyano, S., Shimamoto, A., Tahara, H.*, et al.* (2015). Mutations in the Gene Encoding the E2 Conjugating Enzyme UBE2T Cause Fanconi Anemia. American journal of human genetics *96*, 1001-1007.

Ho, G.P., Margossian, S., Taniguchi, T., and D'Andrea, A.D. (2006). Phosphorylation of FANCD2 on two novel sites is required for mitomycin C resistance. Molecular and cellular biology *26*, 7005-7015.

Hodskinson, M.R.G., Silhan, J., Crossan, G.P., Garaycoechea, J.I., Mukherjee, S., Johnson, C.M., Scharer, O.D., and Patel, K.J. (2014). Mouse SLX4 Is a Tumor Suppressor that Stimulates the Activity of the Nuclease XPF-ERCC1 in DNA Crosslink Repair. Molecular cell *54*, 472-484.

Hodson, C., Purkiss, A., Miles, J.A., and Walden, H. (2014). Structure of the human FANCL RING-Ube2T complex reveals determinants of cognate E3-E2 selection. Structure *22*, 337-344.

Howlett, N.G., Taniguchi, T., Durkin, S.G., D'Andrea, A.D., and Glover, T.W. (2005). The Fanconi anemia pathway is required for the DNA replication stress response and for the regulation of common fragile site stability. Human molecular genetics *14*, 693-701.

Howlett, N.G., Taniguchi, T., Olson, S., Cox, B., Waisfisz, Q., de Die-Smulders, C., Persky, N., Grompe, M., Joenje, H., Pals, G.*, et al.* (2002). Biallelic inactivation of BRCA2 in Fanconi anemia. Science *297*, 606-609.

Huang, J., Liu, S., Bellani, M.A., Thazhathveetil, A.K., Ling, C., de Winter, J.P., Wang, Y., Wang, W., and Seidman, M.M. (2013). The DNA translocase FANCM/MHF promotes replication traverse of DNA interstrand crosslinks. Molecular cell *52*, 434-446.

Huang, Y., Leung, J.W., Lowery, M., Matsushita, N., Wang, Y., Shen, X., Huong, D., Takata, M., Chen, J., and Li, L. (2014). Modularized functions of the Fanconi anemia core complex. Cell Rep *7*, 1849-1857.

Huen, M.S.Y., Grant, R., Manke, I., Minn, K., Yu, X.C., Yaffe, M.B., and Chen, J.J. (2007). RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly. Cell *131*, 901-914.

Ikura, T., Tashiro, S., Kakino, A., Shima, H., Jacob, N., Amunugama, R., Yoder, K., Izumi, S., Kuraoka, I., Tanaka, K.*, et al.* (2007). DNA damage-dependent acetylation and ubiquitination of H2AX enhances chromatin dynamics. Molecular and cellular biology *27*, 7028-7040.

Ishiai, M., Kitao, H., Smogorzewska, A., Tomida, J., Kinomura, A., Uchida, E., Saberi, A., Kinoshita, E., Kinoshita-Kikuta, E., Koike, T.*, et al.* (2008). FANCI phosphorylation functions as a molecular switch to turn on the Fanconi anemia pathway. Nature structural & molecular biology *15*, 1138-1146.

Joenje, H., and Patel, K.J. (2001). The emerging genetic and molecular basis of Fanconi anaemia. Nat Rev Genet *2*, 446-457.

Joo, W., Xu, G., Persky, N.S., Smogorzewska, A., Rudge, D.G., Buzovetsky, O., Elledge, S.J., and Pavletich, N.P. (2011). Structure of the FANCI-FANCD2 complex: insights into the Fanconi anemia DNA repair pathway. Science *333*, 312-316.

Kais, Z., Rondinelli, B., Holmes, A., O'Leary, C., Kozono, D., D'Andrea, A.D., and Ceccaldi, R. (2016). FANCD2 Maintains Fork Stability in BRCA1/2-Deficient Tumors and Promotes Alternative End-Joining DNA Repair. Cell Rep *15*, 2488-2499.

Kalb, R., Neveling, K., Hoehn, H., Schneider, H., Linka, Y., Batish, S.D., Hunt, C., Berwick, M., Callen, E., Surralles, J.*, et al.* (2007). Hypomorphic mutations in the gene encoding a key Fanconi anemia protein, FANCD2, sustain a significant group of FA-D2 patients with severe phenotype. American journal of human genetics *80*, 895-910.

Karass, M., Naguib, M.M., Elawabdeh, N., Cundiff, C.A., Thomason, J., Steelman, C.K., Cone, R., Schwenkter, A., Jordan, C., and Shehata, B.M. (2015). Xeroderma pigmentosa: three new cases with an in depth review of the genetic and clinical characteristics of the disease. Fetal Pediatr Pathol *34*, 120-127.

Kashiyama, K., Nakazawa, Y., Pilz, D.T., Guo, C.W., Shimada, M., Sasaki, K., Fawcett, H., Wing, J.F., Lewin, S.O., Carr, L.*, et al.* (2013). Malfunction of Nuclease ERCC1-XPF Results in Diverse Clinical Manifestations and Causes Cockayne Syndrome, Xeroderma Pigmentosum, and Fanconi Anemia. American journal of human genetics *92*, 807-819.

Khanna, K.K., and Jackson, S.P. (2001). DNA double-strand breaks: signaling, repair and the cancer connection. Nature genetics *27*, 247-254.

Kim, H., Chen, J., and Yu, X. (2007). Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response. Science *316*, 1202-1205.

Kim, J.M., Kee, Y., Gurtan, A., and D'Andrea, A.D. (2008). Cell cycle-dependent chromatin loading of the Fanconi anemia core complex by FANCM/FAAP24. Blood *111*, 5215-5222.

Kim, Y., Lach, F.P., Desetty, R., Hanenberg, H., Auerbach, A.D., and Smogorzewska, A. (2011). Mutations of the SLX4 gene in Fanconi anemia. Nature genetics *43*, 142-U191.

Klein Douwel, D., Boonen, R.A., Long, D.T., Szypowska, A.A., Raschle, M., Walter, J.C., and Knipscheer, P. (2014). XPF-ERCC1 acts in Unhooking DNA interstrand crosslinks in cooperation with FANCD2 and FANCP/SLX4. Molecular cell *54*, 460-471.

Knipscheer, P., Raschle, M., Smogorzewska, A., Enoiu, M., Ho, T.V., Scharer, O.D., Elledge, S.J., and Walter, J.C. (2009). The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. Science *326*, 1698-1701.

Koc, A., Wheeler, L.J., Mathews, C.K., and Merrill, G.F. (2004). Hydroxyurea arrests DNA replication by a mechanism that preserves basal dNTP pools. The Journal of biological chemistry *279*, 223-230.

Kruyt, F.A., Abou-Zahr, F., Mok, H., and Youssoufian, H. (1999). Resistance to mitomycin C requires direct interaction between the Fanconi anemia proteins FANCA and FANCG in the nucleus through an arginine-rich domain. The Journal of biological chemistry *274*, 34212-34218.

Kuraoka, I., Kobertz, W.R., Ariza, R.R., Biggerstaff, M., Essigmann, J.M., and Wood, R.D. (2000). Repair of an interstrand DNA cross-link initiated by ERCC1-XPF repair/recombination nuclease. The Journal of biological chemistry *275*, 26632-26636.

Lachaud, C., Moreno, A., Marchesi, F., Toth, R., Blow, J.J., and Rouse, J. (2016). Ubiquitinated Fancd2 recruits Fan1 to stalled replication forks to prevent genome instability. Science *351*, 846-849.

Langevin, F., Crossan, G.P., Rosado, I.V., Arends, M.J., and Patel, K.J. (2011). Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. Nature *475*, 53-58.

Lavin, M.F., and Shiloh, Y. (1996). Ataxia-telangiectasia: a multifaceted genetic disorder associated with defective signal transduction. Curr Opin Immunol *8*, 459-464.

Leung, J.W., Wang, Y., Fong, K.W., Huen, M.S., Li, L., and Chen, J. (2012). Fanconi anemia (FA) binding protein FAAP20 stabilizes FA complementation group A (FANCA) and participates in interstrand cross-link repair. Proc Natl Acad Sci U S A *109*, 4491-4496.

Levitus, M., Waisfisz, Q., Godthelp, B.C., de Vries, Y., Hussain, S., Wiegant, W.W., Elghalbzouri-Maghrani, E., Steltenpool, J., Rooimans, M.A., Pals, G.*, et al.* (2005). The DNA helicase BRIP1 is defective in Fanconi anemia complementation group J. Nature genetics *37*, 934-935.

Levran, O., Attwooll, C., Henry, R.T., Milton, K.L., Neveling, K., Rio, P., Batish, S.D., Kalb, R., Velleuer, E., Barral, S.*, et al.* (2005). The BRCA1-interacting helicase BRIP1 is deficient in Fanconi anemia (vol 37, pg 931, 2005). Nature genetics *37*, 1296-1296.

Li, A., and Swift, M. (2000). Mutations at the ataxia-telangiectasia locus and clinical phenotypes of A-T patients. Am J Med Genet *92*, 170-177.

Liang, C.C., Li, Z., Lopez-Martinez, D., Nicholson, W.V., Venien-Bryan, C., and Cohn, M.A. (2016). The FANCD2-FANCI complex is recruited to DNA interstrand crosslinks before monoubiquitination of FANCD2. Nature communications *7*, 12124.

Liang, C.C., Zhan, B., Yoshikawa, Y., Haas, W., Gygi, S.P., and Cohn, M.A. (2015). UHRF1 is a sensor for DNA interstrand crosslinks and recruits FANCD2 to initiate the Fanconi anemia pathway. Cell Rep *10*, 1947-1956.

Litman, R., Peng, M., Jin, Z., Zhang, F., Zhang, J.R., Powell, S., Andreassen, P.R., and Cantor, S.B. (2005). BACH1 is critical for homologous recombination and appears to be the Fanconi anemia gene product FANCJ. Cancer cell *8*, 255-265.

Lo Ten Foe, J.R., Rooimans, M.A., Bosnoyan-Collins, L., Alon, N., Wijker, M., Parker, L., Lightfoot, J., Carreau, M., Callen, D.F., Savoia, A.*, et al.* (1996). Expression cloning of a cDNA for the major Fanconi anaemia gene, FAA. Nature genetics *14*, 320-323.

Long, D.T., Joukov, V., Budzowska, M., and Walter, J.C. (2014). BRCA1 promotes unloading of the CMG helicase from a stalled DNA replication fork. Molecular cell *56*, 174-185.

Long, D.T., Raschle, M., Joukov, V., and Walter, J.C. (2011). Mechanism of RAD51-dependent DNA interstrand cross-link repair. Science *333*, 84-87.

Longerich, S., Kwon, Y., Tsai, M.S., Hlaing, A.S., Kupfer, G.M., and Sung, P. (2014). Regulation of FANCD2 and FANCI monoubiquitination by their interaction and by DNA. Nucleic Acids Res *42*, 5657-5670.

Longerich, S., San Filippo, J., Liu, D., and Sung, P. (2009). FANCI binds branched DNA and is monoubiquitinated by UBE2T-FANCL. The Journal of biological chemistry *284*, 23182-23186.

Lopez-Martinez, D., Liang, C.C., and Cohn, M.A. (2016). Cellular response to DNA interstrand crosslinks: the Fanconi anemia pathway. Cellular and molecular life sciences : CMLS *73*, 3097-3114.

Lossaint, G., Larroque, M., Ribeyre, C., Bec, N., Larroque, C., Decaillet, C., Gari, K., and Constantinou, A. (2013). FANCD2 Binds MCM Proteins and Controls Replisome Function upon Activation of S Phase Checkpoint Signaling. Molecular cell *51*, 678-690.

Ludtke, S.J., Baldwin, P.R., and Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. Journal of structural biology *128*, 82-97.

Machida, Y.J., Machida, Y., Chen, Y., Gurtan, A.M., Kupfer, G.M., D'Andrea, A.D., and Dutta, A. (2006). UBE2T is the E2 in the Fanconi anemia pathway and undergoes negative autoregulation. Molecular cell *23*, 589-596.

Mailand, N., Bekker-Jensen, S., Faustrup, H., Melander, F., Bartek, J., Lukas, C., and Lukas, J. (2007). RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. Cell *131*, 887-900.

Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y.*, et al.* (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science *316*, 1160-1166.

Mattiroli, F., Vissers, J.H.A., van Dijk, W.J., Ikpa, P., Citterio, E., Vermeulen, W., Marteijn, J.A., and Sixma, T.K. (2012). RNF168 Ubiquitinates K13-15 on H2A/H2AX to Drive DNA Damage Signaling. Cell *150*, 1182-1195.

Meetei, A.R., de Winter, J.P., Medhurst, A.L., Wallisch, M., Waisfisz, Q., van de Vrugt, H.J., Oostra, A.B., Yan, Z.J., Ling, C., Bishop, C.E.*, et al.* (2003). A novel ubiquitin ligase is deficient in Fanconi anemia. Nature genetics *35*, 165-170.

Meetei, A.R., Levitus, M., Xue, Y., Medhurst, A.L., Zwaan, M., Ling, C., Rooimans, M.A., Bier, P., Hoatlin, M., Pals, G.*, et al.* (2004). X-linked inheritance of Fanconi anemia complementation group B. Nature genetics *36*, 1219-1224.

Meetei, A.R., Medhurst, A.L., Ling, C., Xue, Y.T., Singh, T.R., Bier, P., Steltenpool, J., Stone, S., Dokal, I., Mathew, C.G.*, et al.* (2005). A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. Nature genetics *37*, 958-963.

Michl, J., Zimmer, J., Buffa, F.M., McDermott, U., and Tarsounas, M. (2016). FANCD2 limits replication stress and genome instability in cells lacking BRCA2. Nature structural & molecular biology *23*, 755-757.

Mindell, J.A., and Grigorieff, N. (2003). Accurate determination of local defocus and specimen tilt in electron microscopy. Journal of structural biology *142*, 334-347.

Mouw, K.W., and D'Andrea, A.D. (2014). Crosstalk between the nucleotide excision repair and Fanconi anemia/BRCA pathways. DNA repair *19*, 130-134.

Muniandy, P.A., Thapa, D., Thazhathveetil, A.K., Liu, S.T., and Seidman, M.M. (2009). Repair of laser-localized DNA interstrand cross-links in G1 phase mammalian cells. The Journal of biological chemistry *284*, 27908-27917.

Munoz, I.M., Hain, K., Declais, A.C., Gardiner, M., Toh, G.W., Sanchez-Pulido, L., Heuckmann, J.M., Toth, R., Macartney, T., Eppink, B.*, et al.* (2009). Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair. Molecular cell *35*, 116-127.

Murina, O., von Aesch, C., Karakus, U., Ferretti, L.P., Bolck, H.A., Hanggi, K., and Sartori, A.A. (2014). FANCD2 and CtIP cooperate to repair DNA interstrand crosslinks. Cell Rep *7*, 1030-1038.

Nakatani, Y., and Ogryzko, V. (2003). Immunoaffinity purification of mammalian protein complexes. Methods Enzymol *370*, 430-444.

Nielsen, F.C., van Overeem Hansen, T., and Sorensen, C.S. (2016). Hereditary breast and ovarian cancer: new genes in confined pathways. Nat Rev Cancer *16*, 599-612.

Nishiyama, A., Yamaguchi, L., Sharif, J., Johmura, Y., Kawamura, T., Nakanishi, K., Shimamura, S., Arita, K., Kodama, T., Ishikawa, F.*, et al.* (2013). Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. Nature *502*, 249-253.

Nogales, E., and Scheres, S.H. (2015). Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. Molecular cell *58*, 677-689.

Pace, P., Johnson, M., Tan, W.M., Mosedale, G., Sng, C., Hoatlin, M., de Winter, J., Joenje, H., Gergely, F., and Patel, K.J. (2002). FANCE: the link between Fanconi anaemia complex assembly and activity. Embo Journal *21*, 3414-3423.

Pang, Q., and Andreassen, P.R. (2009). Fanconi anemia proteins and endogenous stresses. Mutat Res *668*, 42-53.

Panier, S., Ichijima, Y., Fradet-Turcotte, A., Leung, C.C.Y., Kaustov, L., Arrowsmith, C.H., and Durocher, D. (2012). Tandem Protein Interaction Modules Organize the Ubiquitin-Dependent Response to DNA Double-Strand Breaks. Molecular cell *47*, 383-395.

Park, J.Y., Virts, E.L., Jankowska, A., Wiek, C., Othman, M., Chakraborty, S.C., Vance, G.H., Alkuraya, F.S., Hanenberg, H., and Andreassen, P.R. (2016). Complementation of hypersensitivity to DNA interstrand crosslinking agents demonstrates that XRCC2 is a Fanconi anaemia gene. J Med Genet.

Park, W.H., Margossian, S., Horwitz, A.A., Simons, A.M., D'Andrea, A.D., and Parvin, J.D. (2005). Direct DNA binding activity of the Fanconi anemia D2 protein. The Journal of biological chemistry *280*, 23593-23598.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. Journal of computational chemistry *25*, 1605-1612.

Polito, D., Cukras, S., Wang, X., Spence, P., Moreau, L., D'Andrea, A.D., and Kee, Y. (2014). The carboxyl terminus of FANCE recruits FANCD2 to the Fanconi Anemia (FA) E3 ligase complex to promote the FA DNA repair pathway. The Journal of biological chemistry *289*, 7003-7010.

Prabakaran, S., Lippens, G., Steen, H., and Gunawardena, J. (2012). Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding. Wiley Interdiscip Rev Syst Biol Med *4*, 565-583.

Qian, C., Li, S., Jakoncic, J., Zeng, L., Walsh, M.J., and Zhou, M.M. (2008). Structure and hemimethylated CpG binding of the SRA domain from human UHRF1. The Journal of biological chemistry *283*, 34490-34494.

Radermacher, M., Wagenknecht, T., Verschoor, A., and Frank, J. (1987). Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of Escherichia coli. Journal of microscopy *146*, 113-136.

Rajendra, E., Oestergaard, V.H., Langevin, F., Wang, M., Dornan, G.L., Patel, K.J., and Passmore, L.A. (2014). The genetic and biochemical basis of FANCD2 monoubiquitination. Molecular cell *54*, 858-869.

Raschle, M., Knipscheer, P., Enoiu, M., Angelov, T., Sun, J., Griffith, J.D., Ellenberger, T.E., Scharer, O.D., and Walter, J.C. (2008). Mechanism of replication-coupled DNA interstrand crosslink repair. Cell *134*, 969-980.

Reuter, T., Herterich, S., Bernhard, O., Hoehn, H., and Gross, H.J. (2000). Strong FANCA/FANCG but weak FANCA/FANCC interaction in the yeast 2-hybrid system. Blood *95*, 719-720.

Rickman, K.A., Lach, F.P., Abhyankar, A., Donovan, F.X., Sanborn, E.M., Kennedy, J.A., Sougnez, C., Gabriel, S.B., Elemento, O., Chandrasekharappa, S.C*., et al.* (2015). Deficiency of UBE2T, the E2 Ubiquitin Ligase Necessary for FANCD2 and FANCI Ubiquitination, Causes FA-T Subtype of Fanconi Anemia. Cell Reports *12*, 35-41.

Rogakou, E.P., Boon, C., Redon, C., and Bonner, W.M. (1999). Megabase chromatin domains involved in DNA double-strand breaks in vivo. J Cell Biol *146*, 905-916.

Rohleder, F., Huang, J., Xue, Y., Kuper, J., Round, A., Seidman, M., Wang, W., and Kisker, C. (2016). FANCM interacts with PCNA to promote replication traverse of DNA interstrand crosslinks. Nucleic Acids Res *44*, 3219-3232.

Roos, W.P., Thomas, A.D., and Kaina, B. (2016). DNA damage and the balance between survival and death in cancer biology. Nat Rev Cancer *16*, 20-33.

Roseman, A.M. (2003). Particle finding in electron micrographs using a fast local correlation algorithm. Ultramicroscopy *94*, 225-236.

Rosenberg, P.S., Tamary, H., and Alter, B.P. (2011). How high are carrier frequencies of rare recessive syndromes? Contemporary estimates for Fanconi Anemia in the United States and Israel. Am J Med Genet A *155A*, 1877-1883.

Saibil, H.R. (2000). Conformational changes studied by cryo-electron microscopy. Nat Struct Biol *7*, 711-714.

Sareen, A., Chaudhury, I., Adams, N., and Sobeck, A. (2012). Fanconi anemia proteins FANCD2 and FANCI exhibit different DNA damage responses during S-phase. Nucleic acids research *40*, 8425-8439.

Sarkar, S., Davies, A.A., Ulrich, H.D., and McHugh, P.J. (2006). DNA interstrand crosslink repair during G1 involves nucleotide excision repair and DNA polymerase zeta. The EMBO journal *25*, 1285-1294.

Sato, K., Toda, K., Ishiai, M., Takata, M., and Kurumizaka, H. (2012). DNA robustly stimulates FANCD2 monoubiquitylation in the complex with FANCI. Nucleic Acids Res *40*, 4553-4561.

Sawyer, S.L., Tian, L., Kahkonen, M., Schwartzentruber, J., Kircher, M., Majewski, J., Dyment, D.A., Innes, A.M., Boycott, K.M., Moreau, L.A.*, et al.* (2015). Biallelic Mutations in BRCA1 Cause a New Fanconi Anemia Subtype. Cancer discovery *5*, 135-142.

Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. Journal of structural biology *180*, 519-530.

Schlacher, K., Christ, N., Siaud, N., Egashira, A., Wu, H., and Jasin, M. (2011). Double-Strand Break Repair-Independent Role for BRCA2 in Blocking Stalled Replication Fork Degradation by MRE11 (vol 145, pg 529, 2011). Cell *145*, 993-993.

Schlacher, K., Wu, H., and Jasin, M. (2012). A Distinct Replication Fork Protection Pathway Connects Fanconi Anemia Tumor Suppressors to RAD51-BRCA1/2. Cancer cell *22*, 106-116.

Shamseldin, H.E., Elfaki, M., and Alkuraya, F.S. (2012). Exome sequencing reveals a novel Fanconi group defined by XRCC2 mutation. Journal of Medical Genetics *49*, 184-186.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K.*, et al.* (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature *450*, 908-912.

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem *68*, 850-858.

Sims, A.E., Spiteri, E., Sims, R.J., Arita, A.G., Lach, F.P., Landers, T., Wurm, M., Freund, M., Neveling, K., Hanenberg, H.*, et al.* (2007). FANCI is a second monoubiquitinated member of the Fanconi anemia pathway. Nature structural & molecular biology *14*, 564-567.

Smogorzewska, A., Matsuoka, S., Vinciguerra, P., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Ballif, B.A., Gygi, S.P., Hofmann, K., D'Andrea, A.D.*, et al.* (2007). Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. Cell *129*, 289-301.

Stoepker, C., Hain, K., Schuster, B., Hilhorst-Hofstee, Y., Rooimans, M.A., Steltenpool, J., Oostra, A.B., Eirich, K., Korthof, E.T., Nieuwint, A.W.M.*, et al.* (2011). SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. Nature genetics *43*, 138-U185.

Strathdee, C.A., Gavish, H., Shannon, W.R., and Buchwald, M. (1992). Cloning of cDNAs for Fanconi's anaemia by functional complementation. Nature *356*, 763-767.

Stucki, M., Clapperton, J.A., Mohammad, D., Yaffe, M.B., Smerdon, S.J., and Jackson, S.P. (2005). MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. Cell *123*, 1213-1226.

Sugahara, R., Mon, H., Lee, J.M., and Kusakabe, T. (2012). Monoubiquitination-dependent chromatin loading of FancD2 in silkworms, a species lacking the FA core complex. Gene *501*, 180-187.

Suhasini, A.N., Sommers, J.A., Muniandy, P.A., Coulombe, Y., Cantor, S.B., Masson, J.Y., Seidman, M.M., and Brosh, R.M., Jr. (2013). Fanconi anemia group J helicase and MRE11 nuclease interact to facilitate the DNA damage response. Molecular and cellular biology *33*, 2212-2227.

Svendsen, J.M., Smogorzewska, A., Sowa, M.E., O'Connell, B.C., Gygi, S.P., Elledge, S.J., and Harper, J.W. (2009). Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. Cell *138*, 63-77.

Taniguchi, T., Garcia-Higuera, I., Andreassen, P.R., Gregory, R.C., Grompe, M., and D'Andrea, A.D. (2002a). S-phase-specific interaction of the Fanconi anemia protein, FANCD2, with BRCA1 and RAD51. Blood *100*, 2414-2420.

Taniguchi, T., Garcia-Higuera, I., Xu, B., Andreassen, P.R., Gregory, R.C., Kim, S.-T.T., Lane, W.S., Kastan, M.B., and D'Andrea, A.D. (2002b). Convergence of the fanconi anemia and ataxia telangiectasia signaling pathways. Cell *109*, 459-472.

Thorslund, T., Ripplinger, A., Hoffmann, S., Wild, T., Uckelmann, M., Villumsen, B., Narita, T., Sixma, T.K., Choudhary, C., Bekker-Jensen, S.*, et al.* (2015). Histone H1 couples initiation and amplification of ubiquitin signalling after DNA damage. Nature *527*, 389-393.

Tian, Y., Paramasivam, M., Ghosal, G., Chen, D., Shen, X., Huang, Y., Akhter, S., Legerski, R., Chen, J., Seidman, M.M.*, et al.* (2015). UHRF1 contributes to DNA damage repair as a lesion recognition factor and nuclease scaffold. Cell Rep *10*, 1957-1966.

Tien, A.L., Senbanerjee, S., Kulkarni, A., Mudbhary, R., Goudreau, B., Ganesan, S., Sadler, K.C., and Ukomadu, C. (2011). UHRF1 depletion causes a G2/M arrest, activation of DNA damage response and apoptosis. Biochem J *435*, 175-185.

Timmers, C., Taniguchi, T., Hejna, J., Reifsteck, C., Lucas, L., Bruun, D., Thayer, M., Cox, B., Olson, S., D'Andrea, A.D.*, et al.* (2001). Positional cloning of a novel fanconi anemia gene, FANCD2. Molecular cell *7*, 241-248.

Tomida, J., Takata, K., Lange, S.S., Schibler, A.C., Yousefzadeh, M.J., Bhetawal, S., Dent, S.Y.R., and Wood, R.D. (2015). REV7 is essential for DNA damage tolerance via two REV3L binding sites in mammalian DNA polymerase zeta. Nucleic acids research *43*, 1000-1011.

Trabuco, L.G., Villa, E., Mitra, K., Frank, J., and Schulten, K. (2008). Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. Structure *16*, 673-683.

Unno, J., Itaya, A., Taoka, M., Sato, K., Tomida, J., Sakai, W., Sugasawa, K., Ishiai, M., Ikura, T., Isobe, T.*, et al.* (2014). FANCD2 binds CtIP and regulates DNA-end resection during DNA interstrand crosslink repair. Cell Rep *7*, 1039-1047.

Uziel, T., Lerenthal, Y., Moyal, L., Andegeko, Y., Mittelman, L., and Shiloh, Y. (2003). Requirement of the MRN complex for ATM activation by DNA damage. The EMBO journal *22*, 5612-5621.

Vaz, F., Hanenberg, H., Schuster, B., Barker, K., Wiek, C., Erven, V., Neveling, K., Endt, D., Kesterton, I., Autore, F.*, et al.* (2010). Mutation of the RAD51C gene in a Fanconi anemia-like disorder. Nature genetics *42*, 406-409.

Virts, E.L., Jankowska, A., Mackay, C., Glaas, M.F., Wiek, C., Kelich, S.L., Lottmann, N., Kennedy, F.M., Marchal, C., Lehnert, E.*, et al.* (2015). AluY-mediated germline deletion, duplication and somatic stem cell reversion in UBE2T defines a new subtype of Fanconi anemia. Human molecular genetics *24*, 5093-5108.

Waisfisz, Q., de Winter, J.P., Kruyt, F.A., de Groot, J., van der Weel, L., Dijkmans, L.M., Zhi, Y., Arwert, F., Scheper, R.J., Youssoufian, H.*, et al.* (1999). A physical complex of the Fanconi anemia proteins FANCG/XRCC9 and FANCA. Proc Natl Acad Sci U S A *96*, 10320-10325.

Walker, J.R., Corpina, R.A., and Goldberg, J. (2001). Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature *412*, 607-614.

Wang, A.T., Kim, T., Wagner, J.E., Conti, B.A., Lach, F.P., Huang, A.L., Molina, H., Sanborn, E.M., Zierhut, H., Cornes, B.K.*, et al.* (2015). A Dominant Mutation in Human RAD51 Reveals

Its Function in DNA Interstrand Crosslink Repair Independent of Homologous Recombination. Molecular cell *59*, 478-490.

Wang, B., Matsuoka, S., Ballif, B.A., Zhang, D., Smogorzewska, A., Gygi, S.P., and Elledge, S.J. (2007). Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. Science *316*, 1194-1198.

Wang, L.C., Stone, S., Hoatlin, M.E., and Gautier, J. (2008). Fanconi anemia proteins stabilize replication forks. DNA repair *7*, 1973-1981.

Wang, X., Peterson, C.A., Zheng, H., Nairn, R.S., Legerski, R.J., and Li, L. (2001). Involvement of nucleotide excision repair in a recombination-independent and error-prone pathway of DNA interstrand cross-link repair. Molecular and cellular biology *21*, 713-720.

Wilson, M.D., Benlekbir, S., Fradet-Turcotte, A., Sherker, A., Julien, J.P., McEwan, A., Noordermeer, S.M., Sicheri, F., Rubinstein, J.L., and Durocher, D. (2016). The structural basis of modified nucleosome recognition by 53BP1. Nature *536*, 100-103.

Wood, R.D. (2010). Mammalian nucleotide excision repair proteins and interstrand crosslink repair. Environmental and molecular mutagenesis *51*, 520-526.

Wu, J., Huen, M.S., Lu, L.Y., Ye, L., Dou, Y., Ljungman, M., Chen, J., and Yu, X. (2009). Histone ubiquitination associates with BRCA1-dependent DNA damage response. Molecular and cellular biology *29*, 849-860.

Xia, B., Dorsman, J.C., Ameziane, N., de Vries, Y., Rooimans, M.A., Sheng, Q., Pals, G., Errami, A., Gluckman, E., Llera, J.*, et al.* (2007). Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. Nature genetics *39*, 159-161.

Xie, S., Jakoncic, J., and Qian, C. (2012). UHRF1 double tudor domain and the adjacent PHD finger act together to recognize K9me3-containing histone H3 tail. J Mol Biol *415*, 318-328.

Yamamoto, K.N., Kobayashi, S., Tsuda, M., Kurumizaka, H., Takata, M., Kono, K., Jiricny, J., Takeda, S., and Hirota, K. (2011). Involvement of SLX4 in interstrand cross-link repair is regulated by the Fanconi anemia pathway. Proc Natl Acad Sci U S A *108*, 6492-6496.

Yan, Z., Guo, R., Paramasivam, M., Shen, W., Ling, C., Fox, D., 3rd, Wang, Y., Oostra, A.B., Kuehl, J., Lee, D.Y.*, et al.* (2012). A ubiquitin-binding protein, FAAP20, links RNF8-mediated ubiquitination to the Fanconi anemia DNA repair network. Molecular cell *47*, 61-75.

Zhang, J., Dewar, J.M., Budzowska, M., Motnenko, A., Cohn, M.A., and Walter, J.C. (2015). DNA interstrand cross-link repair requires replication-fork convergence. Nature structural & molecular biology *22*, 242-247.

Zhang, X.Y., Langenick, J., Traynor, D., Babu, M.M., Kay, R.R., and Patel, K.J. (2009). Xpf and not the Fanconi anaemia proteins or Rev3 accounts for the extreme resistance to cisplatin in Dictyostelium discoideum. PLoS genetics *5*, e1000645.

Zhi, G., Wilson, J.B., Chen, X., Krause, D.S., Xiao, Y., Jones, N.J., and Kupfer, G.M. (2009). Fanconi anemia complementation group FANCD2 protein serine 331 phosphorylation is important for fanconi anemia pathway function and BRCA2 interaction. Cancer Res *69*, 8775-8783.

Appendix

Statement of previous acceptance or concurrent submission of thesis for degree

The purification and identification of UHRF1 as an interstrand crosslinked DNA sensor (Figure 2-1) was published in the DPhil thesis of Bao Zjan in 2013.

Statement of authorship

The purification and identification of UHRF1 as an interstrand crosslinked DNA sensor (Figure 2-1 A-D) was done by Bao Zjan, Wilhelm Haas and Steven Gygi (Figure 2-1E).

The Cryo-EM structure determination of full-length FANCD2/FANCI complex and Tower domain deletion FANCD2/FANCI complex was done by Zhuolun Li, William Nicholson and Catherine Vénien-Bryan (Figure 3-1 D-F and Figure 3-2 C-E)). The FANCD2 CRISPR/Cas9-mediated knockout in HeLa cells was generated by David Lopez-Martinez.

Published data and figures

Some works described in this thesis are published (Liang et al., 2016; Liang et al., 2015; Lopez-Martinez et al., 2016), and are re-used under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/ or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Liang, C.C., Li, Z., Lopez-Martinez, D., Nicholson, W.V., Venien-Bryan, C., and Cohn, M.A. (2016). The FANCD2-FANCI complex is recruited to DNA interstrand crosslinks before monoubiquitination of FANCD2. Nature communications *7*, 12124.

Liang, C.C., Zhan, B., Yoshikawa, Y., Haas, W., Gygi, S.P., and Cohn, M.A. (2015). UHRF1 is a sensor for DNA interstrand crosslinks and recruits FANCD2 to initiate the Fanconi anemia pathway. Cell Rep *10*, 1947-1956.

Lopez-Martinez, D., Liang, C.C., and Cohn, M.A. (2016). Cellular response to DNA interstrand crosslinks: the Fanconi anemia pathway. Cellular and molecular life sciences : CMLS *73*, 3097-3114.