RAD18-dependent Activation of FAN1 by Ub-PCNA to Rescue DNA Interstrand crosslinks along with TLS Pol η

Ph.D. Dissertation

Qiuzhen Li

Supervisor: Lajos Haracska, Ph.D, D.Sc.

Doctoral School of Biology

University of Szeged

Institute of Genetics, Biological Research Centre, Szeged

Szeged, 2021

Table of Contents

List of Abbreviations

ATM: Ataxia-telangiectasia-mutated

ATP: Adenosine triphosphate

ATR: Ataxia telangiectasia and Rad3-related

BER: Base excision repair

BiFC: Bimolecular fluorescence complementation

CDK1: Cyclin-dependent kinase 1

CHK1: Checkpoint kinase 1

DDR: DNA damage response

DDT: DNA damage tolerance

DMSO: Dimethyl sulfoxide

DNA: Deoxyribonucleic acid

DSB: Double-strand break

DTT: Dithiothreitol

EDTA: Ethylenediaminetetraacetic acid

ERCC1-XPF: Xeroderma pigmentosum complementation group F- Excision repair cross complementation group 1

FA: Fanconi anemia

FAM: 6-carboxyfluorescein

FAN1: FANCD2/FANCI-associated nuclease 1

FBS: Fetal bovine serum

FITC: Fluorescein

G0/G1 phase: Gap 0 (quiescent)/gap 1 phase

G4s: G-quadruplexes

GST: Glutathione S-transferase

HECT: Homologous to E6-associated Protein C-terminus

HLTF: Helicase-like transcription factor

HR: Homologous recombination

HRP: Horseradish peroxidase

ICL: Interstrand crosslink

ID2: FANCD2/FANCI

K164: Lysine 164 Residue of PCNA

MBL: Metallo β-lactamase

MSH2: DNA mismatch repair protein Msh2 (MutS homolog 2)

MMR: Mismatch repair

MRN: MRE11-RAD50-NBS1

NER: Nucleotide excision repair

NHEJ: Non-homologous end joining

OE: Overexpression

PAGE: Polyacrylamide gel

PCNA: Proliferative cell nuclear antigen

PMSF: Phenylmethylsulfonyl fluoride

PPi: Pyrophosphate

PTM: Post-translational modification

RFC: Replication factor C

RING: Really Interesting New Gene

RNA: Ribonucleic acid

ROS: Reactive oxygen species

RPA: Human replication protein A

RT: Room temperature

SDS: Sodium dodecyl sulfate

S phase: Synthesis phase

S/B: Signal-to-background ratio

SE: Standard error

SHPRH: Smf2 histone linker PHD ring helicase

SNM1A: Human Sensitive to Nitrogen Mustard 1A

STD: Saturation transfer difference

SUMO: Small ubiquitin-like modifier

TLS: TranLesion synthesis

TE: Tris and EDTA

TMP: 4,5′,8-Trimethylpsoralen

TPR: Tetratricopeptide repeat

UBA1: Ubiquitin-activating enzyme 1

UBM: Ubiquitin-binding motif

UBZ: Ubiquitin-binding zinc finger

UHRF1: Ubiquitin Like with PHD and Ring Finger Domains 1

UHRF2: Ubiquitin Like with PHD and Ring Finger Domains 2

UV: Ultraviolet

YFP: Yellow fluorescent protein

VRR_NUC: virus-type replication-repair nuclease

1. Introduction

1.1 DNA damage and repair

Deoxyribonucleic acid (DNA), our genetic material, is essential to insure the inheritance of genetic information from one generation to the next. It is crucial to protect DNA from any potential damage and maintain genome stability. DNA is wrapped around histone octamer (H2A, H2B, H3, and H4) forming the unit as the nucleosome. This structure does not only protect naked DNA from damage, but also regulates gene expression epigenetically. DNA is threatened by both endogenous and exogenous factors in our daily life. Endogenous mutagens include reactive oxygen species (ROS), aldehydes as metabolic products from β-oxidation of lipids, or mismatched incorporated deoxynucleotides during replication. Exogenous mutagens can be ultraviolet (UV) irradiation, chemotherapeutic drugs, alcohol consumption, tobacco smoking, etc. (Chatterjee and Walker, 2017). To protect fragile DNA, complex protection mechanisms, DNA damage responses (DDRs) have evolved in higher vertebrates. The first step of DDR is recognizing cellular DNA damage. When the damage is detected, a series of signal transduction is initiated. This cellular response can affect cell cycle progression, DNA replication, transcription, and activation of the DNA repair machinery. There are two cellular strategies to combat possible DNA damage: the DNA damage repair and the DNA damage tolerance (DDT) pathways. The former one operates in an error-free mode that can directly remove damaged DNA and does not introduce mutations after repair. DNA damage repair (such as base excision repair, nucleotide excision repair, mismatch repair, etc.) ensures the precise removal of damage in any stage of the cell cycle. DDT pathways, however, are responsible for tolerating DNA lesions in the S phase of the cell cycle and may be sources of mutagenesis. During DNA replication, DNA damage can hinder the ongoing process of replication. There are multiple steps in the DDT pathways to remove the obstacle from the stalled replication fork. Both DNA repair and DDT mechanisms are required to ensure faithful DNA replication.

1.2 Post-translational modification of proteins; phosphorylation and ubiquitination as DNA repair signals

Proteins are produced via the translation of RNA to protein in the ribosomes followed by folding and post-translational modification (PTM), which are important in determining the functions of proteins. PTMs are by covalent enzymatic modifications of certain domains in the

protein. Phosphorylation and ubiquitination play essential signal transduction roles in DNA repair pathways. For example, serine/tyrosine residues are often phosphorylated by protein kinases, which transfer a phosphoryl group onto the protein. In contrast, a phosphorylated protein can be dephosphorylated by phosphatases, which remove a phosphate group from the serine/tyrosine residue by hydrolysis. Both phosphorylation and dephosphorylation can alter the function of the protein, and the modified protein can act as a signal transducer to amplify cellular signals. The cell cycle is tightly governed by phosphorylation in DNA damage response so that the damaged DNA can be repaired or the affected cell can go into senescence or apoptosis. For instance, upon DNA damage, the MRE11-RAD50-NBS1 (MRN) complex serves as a platform and recruits the Ataxia-telangiectasia-mutated (ATM) protein, which undergoes autophosphorylation. Activated ATM can further phosphorylate the tumor suppressor checkpoint kinase 1 (CHK1) (Falck et al., 2005). The CHK1 protein can regulate the G2/M cell cycle checkpoint. The phosphorylated CHK1 can phosphorylate WEE1 kinase, and the activated WEE1 kinase can further phosphorylate tyrosine-15 residue on the Cyclin-Dependent Kinase 1 (CDK1) for preventing the cell from entering the S-phase (Zhang and Hunter, 2014). The phosphoryl group is added from one protein to another as a message, and by passing on this message, cell cycle regulation is well controlled.

Ubiquitin is a stable small regulatory protein, which can covalently bind to another protein to alter its function or to initiate its degradation. This enzymatic ATP-dependent PTM, the so-called ubiquitination, consists of a series of chain reactions, which are catalyzed by the E1 (ubiquitin-activating enzyme), the E2 (ubiquitin-conjugating enzyme), and the E3 (ubiquitin ligase enzyme) proteins (Pickart, 2001). Ubiquitin is activated and transferred to catalytic cysteine residue of E1 via ATP hydrolysis, then it is passed on to catalytic cysteine residue of E2. Eventually, E3 transfers the C-terminus of the ubiquitin protein to the NH_2 side chain of a lysine residue of the target protein (Di Fiore et al., 2003). There are over 500 E3 ligases, and these E3 ligases can be further categorized into RING (Really Interesting New Gene-finger), HECT (Homologous to the E6-AP Carboxyl Terminus), and U-box (a modified RING motif without the full complement of Zn^{2+}-binding residues) contains their domain structure. The most common E3 ligase contains RING-finger, which brings substrate and E2 together as a scaffold protein (U-box containing E3 has the same activity). HECT-E3 ligases can form a thioester bond between their cysteine residue and ubiquitin, and then the ubiquitin is passed to the target lysine residue of a protein through this additional thioester bond (Passmore and Barford, 2004) (Figure 1).

Monoubiquitination can be catalyzed on multiple lysine residues within one protein. The monoubiquitinated protein can be further ubiquitinated by the addition of more ubiquitin proteins on lysine residues of the ubiquitin. The polyubiquitin chain formation usually can happen on lysine 48 (K48) or lysine 63 (K63) of ubiquitin. This reaction is referred to as polyubiquitination. Polyubiquitinated proteins are usually degraded by the 26S proteasome such as in the case of p53 signal responses, thus, this polyubiquitination reaction is mainly characterized by a proteolytic reaction. However, monoubiquitination, and many times K63-polyubiquitination are not followed by a proteolytic reaction, and they serve as a signal for cellular DNA damage repair. Ubiquitination is a reversible enzymatic reaction; the ubiquitinated protein can be deubiquitinated by substrate-specific deubiquitinating enzymes (Nijman et al., 2005). The proteins that are involved in DNA repair often possess ubiquitin-binding domains (UBDs), which can recognize and non-covalently interact with other ubiquitinated proteins of cellular signal transduction cascades (Komander and Rape, 2012). For instance, upon DNA double-strand breaks (DBS), the histone variant H2AX is phosphorylated by ATM, and γH2AX can recruit the RING-finger protein RNF8 (E3-ligase). RNF8 catalyzes the monoubiquitination of histones H2A and H2AX, and these can recruit RNF168 to the site of damage (Doil et al., 2009). RNF168 can further promote polyubiquitination of H2AX as a ubiquitin ligase, which becomes a hub for the accumulation of BRCA1 or 53BP1 proteins (Al-Hakim et al., 2010)

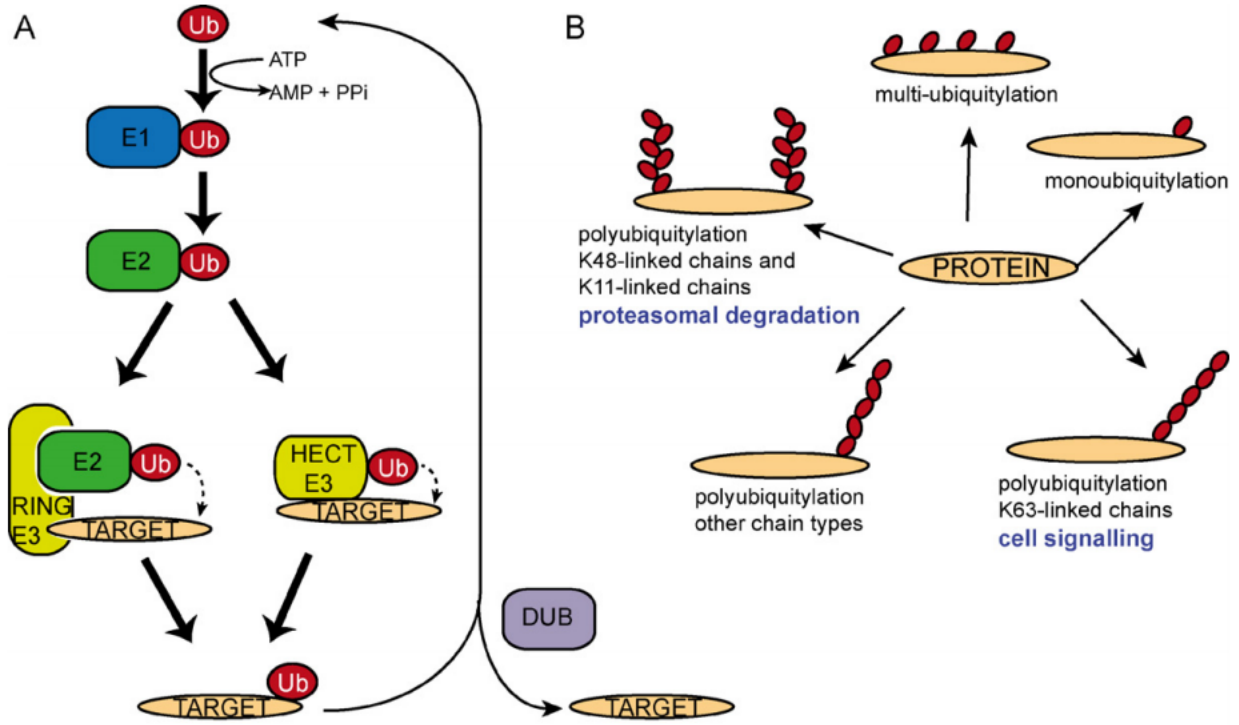


Figure 1. (A) Schematic review of the ubiquitination reaction. Ub: ubiquitin, HECT: homologous to E6-associated protein C-terminus, RING: really interesting new gene. DUB: deubiquitinating enzymes. (B) Schematic review of the functions of ubiquitylated proteins. (Al-Hakim et al., 2010)

1.3 The Fanconi anemia pathway

Fanconi anemia (FA) is a rare genetic disorder characterized by abnormalities of the bones and skin, bone marrow failure, endocrine disorders and hypersensitivity to crosslinking agents (Alter BP, Young NS., 1993). During DNA synthesis, the replication machinery may encounter various DNA lesions due to different damaging agents. Unrepaired lesions may lead to the stalling of the replication fork. Finally, it may result in genome instability, apoptosis or carcinogenesis. The Fanconi Anemia (FA) pathway is one of the main processes responsible for the repair of DNA interstrand crosslinks (ICLs) at the S/G2 cell cycle checkpoint. ICLs are covalent links between the two strands of the DNA helix. When the replication fork stalls due to an ICL, ssDNA is coated by the replication protein A (RPA), which is a signal for the autophosphorylation of the ataxia telangiectasia and Rad3-related (ATR) protein at Thr-1989. The activated phosphorylated ATR,

as a protein kinase, can further activate the serine 1045 residue of FANCM (FAAP250). Phosphorylated FANCM acts as an anchor protein for the FANC core complex assembly and further induces the monoubiquitination of FANCD2/FANCI (ID2) in the Fanconi anemia pathway (Singh et al., 2013). The monoubiquitinated ID2 complex is able to interact with ID2-associated nuclease 1 (FAN1) via its UBZ domain (Kim and D'Andrea, 2012), FAN1 can unhook the ICL, and this unhooked substrate can be further processed by TLS polymerases or nucleotide excision repair (Figure 2).

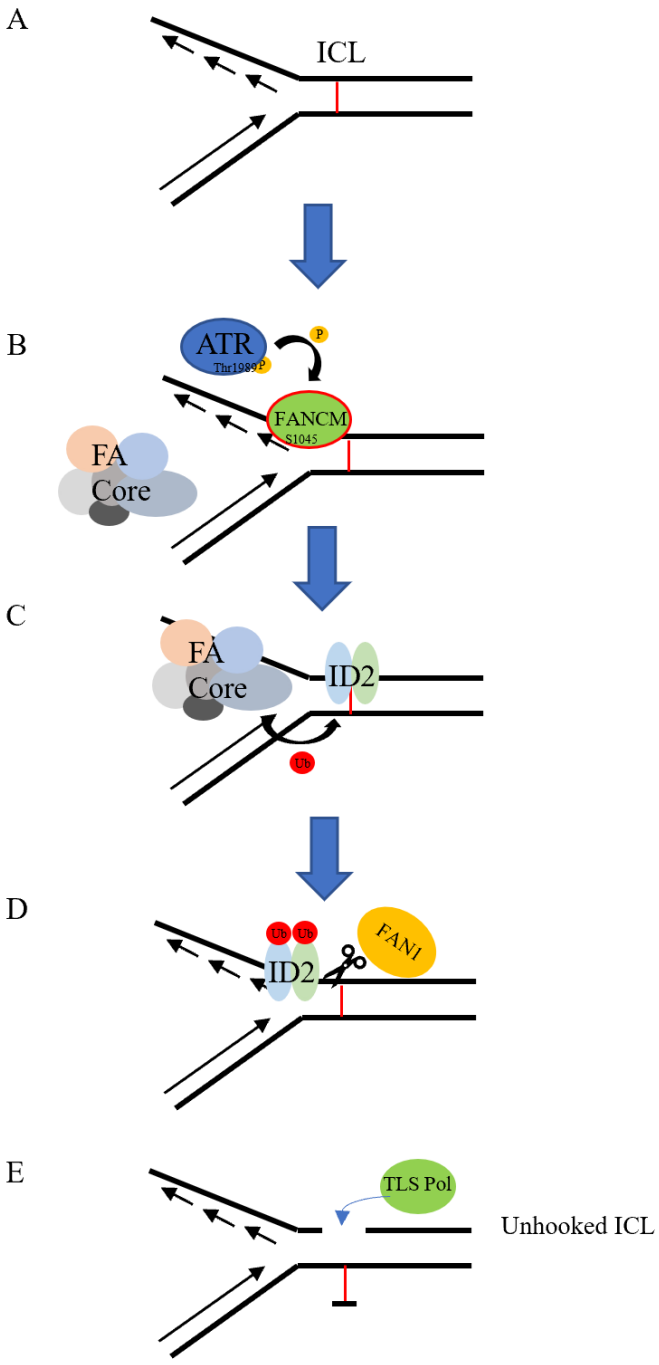


Figure 2. The ongoing replication fork is stalled by ICL (A). The activated ATR can phosphorylate FANCM, which can act as an ICL sensor and platform to recruit the Fanconi Anemia Core Complex to the ICL site (B). FANCL (E3) and UBE2T (E2) of the FA core complex can monoubiquitinate the FANCD2/FANCI heterodimer at the ICL site (C). The monoubiquitinated ID2 complex can recruit FAN1 to unhook ICL (D). The resolved ICL can be further processed by TLS polymerases.

1.4 Interstrand crosslink repair

ICLs hinder ongoing DNA replication and DNA transcription. There are two types of ICL repair mechanisms, the replication-dependent and independent ICL repair. When the ICL lesion is located on the DNA in quiescent cells, namely, in the G0/G1 phase, nucleotide excision repair (NER) is the main pathway employed to remove the lesion from the DNA (Wood, 2010). The XPC-HHR23B complex of the NER system can recruit the XPF-ERCC1 endonuclease to generate an ICL-containing adduct, which can be further processed by TLS polymerases (Sarkar et al., 2006) In non-dividing cells, ICLs are also recognized and repaired by the mismatch repair (MMR) machinery. MutSα (MSH2–MSH6) is critical for ICL recognition, while MutLα and EXO1 contribute to key downstream nucleolytic steps during ICL repair (Kato et al., 2017). ICL lesions can block both leading and lagging strand DNA synthesis, and the FA pathway is the main mechanism activated to remove ICL lesions from the stalled replication fork. This repair machinery is more complicated than ICL repair in the G0/G1 phase, because it involves ICL lesion recognition, recruitment of structure-specific endonucleases/exonucleases, polymerase switching, and, finally, homologous recombination (HR). Replication-associated ICL repair is tightly controlled by the phosphorylation and ubiquitination signal transduction mechanisms. UHRF1 was identified as an ICL sensor, and it can recruit nucleases to process ICL incisions (Liang et al., 2015; Tian et al., 2015). In a recent report, it was shown that UHRF2, as a paralogue of UHRF1, can recruit FANCD2 to the site of ICL and stimulate monoubiquitination of the ID2 complex with UHRF1 (Motnenko et al., 2018). Monoubiquitination of the ID2 complex can act as a platform to further recruit down-stream DNA repair players such as structure-specific nucleases. Structure-specific nucleases act as “scissors” in resolving DNA distortion caused by ICLs; some of them cut on the leading strand, while others make an incision on the lagging strand (Figure 3). This step is referred to as unhooking. The unhooked substrates can be further processed by TLS polymerases, a special class of polymerases that facilitate the bypass of the resolved ICL and the rescue of the stalled replication fork.

1.5 Structure-specific nucleases

1.5.1 ERCC1-XPF

The role of the Excision repair cross complementation group 1-Xeroderma pigmentosum complementation group F (ERCC1-XPF) is well characterized in global genome NER (GG-NER) and transcription-coupled NER (TC-NER) (Faridounnia et al., 2018). Both ERCC1 and XPF belong to the XPF nuclease family, also known as the XPF/MUS81 family. ERCC1-XPF forms a heterodimer complex, which functions as a structure-specific 3' flap endonuclease and is recruited to perform an incision at a ds/ss DNA junction on the 5' side of the damage, while XPG incises on the 3' side (Staresincic et al., 2009) (Figure. 3A). ERCC1-XPF also plays a role in the unhooking step of replication-coupled ICL repair, and ubiquitinated FANCD2 is able to recruit ERCC1-XPF along with SLX4 to the ICL site (Klein Douwel et al., 2014). This heterodimer can accommodate the psoralen-induced interstrand crosslink and process the ICL lesion by making an incision on both sides of the ICL site (Kuraoka et al., 2000).

1.5.2 SLX1-SLX4

SLX1-SLX4 forms a structure-specific endonuclease heterodimer, and it makes an incision when removing ICL lesions blocking the replication fork. SLX1 possesses a UvrC-intron-endonuclease domain (URI) and a PHD-type zinc finger domain, and it displays a weak endonuclease activity (Fricke and Brill, 2003). However, the nuclease activity of SLX1 can be extremely enhanced by its interaction with SLX4 via a helix-turn-helix motif. The heterodimeric SLX1-SLX4 prefers to cleave the 5'flap substrate, but it incises at the ssDNA-dsDNA junction (Fricke and Brill, 2003; Klein Douwel et al., 2014) (Figure. 3B). In the past decade, more and more studies have shown that SLX4 acts as a scaffold protein for recruiting not only SLX1 but other nucleases as well. For example, SLX4 is able to interact with XPF via an MLR motif (Fekairi et al., 2009; Klein Douwel et al., 2014), and this recruitment to the ICL site can promote ICL incision at various angles. Besides its role in ICL repair, SLX4 can also recruit MUS81-EME1 in homologous recombination repair, and the interaction between MUS81 and SLX4 is independent of ICL repair (Castor et al., 2013).

1.5.3 SNM1A

The Human Sensitive to Nitrogen Mustard 1A (SNM1A) protein is a member of the β-CASP family due to its metallo β-lactamase (MBL) and CASP homology domains. It is the orthologue of *Saccharomyces cerevisiae* Pso2 and, as such, it also possesses exonuclease activity

(Callebaut et al. 2002; Cattell et al., 2010). The SNM1A protein contains a UBZ domain at the N-terminal and a PIP-box in the middle region of the protein and is involved in the RAD18 DDT pathway (Yang et al., 2010). Recent research revealed that SNM1A has single-strand-specific endonuclease activity (5'- and 3'-overhangs, hairpins, flaps, and gapped substrates) as well (Figure. 3C). SNM1A might be able to process ICL-containing substrates but not bypass them because SNM1A stops at a short distance from the incision 3' to the ICL, possibly due to steric inhibition by the remnants of the ICL generated by further exonucleolytic degradation (Buzon et al., 2018). SNM1A may be a downstream processor of other structure-specific nucleases trimming the unhooked substrate in the ICL bypass repair.

1.5.4 FAN1

FANCD2/FANCI-associated nuclease 1 (FAN1) is a structure-specific nuclease that plays an important role in the unhooking step of the FA pathway when resolving ICL lesions (Pizzolato et al., 2015). FAN1 is composed of 1017 amino acids and its size is 114,225 daltons. FAN1 possesses a UBZ domain, an uncharacterized PIP-box at the N-terminal end and a TPR domain, a SAP domain (DNA-binding domain), and a VRR-NUC domain (catalytic domain) at the C-terminal end (Pennell et al., 2014; Porro et al., 2017; Smogorzewska et al., 2010). FAN1 is activated by monoubiquitinated FANCD2 via its UBZ domain for unhooking ICL lesions. The previous *in vivo* study also indicated that FAN1-knockdown cells are hypersensitive to interstrand crosslinking agents (Kratz et al., 2010). FAN1 possesses 5'-flap endonuclease activity and is able to incise at the 4th nucleotide after the replication junction on the 5' flap substrate (Figure. 3D). FAN1 has 5'→ 3' exonuclease activity as well, initiating cleavage 4 nt from the 5'-end on single- and double-stranded DNA (MacKay et al., 2010). FAN1 is able to unhook nitrogen mustard-induced interstrand crosslinks *in vitro* owing to its nuclease activity (Pizzolato et al., 2015).

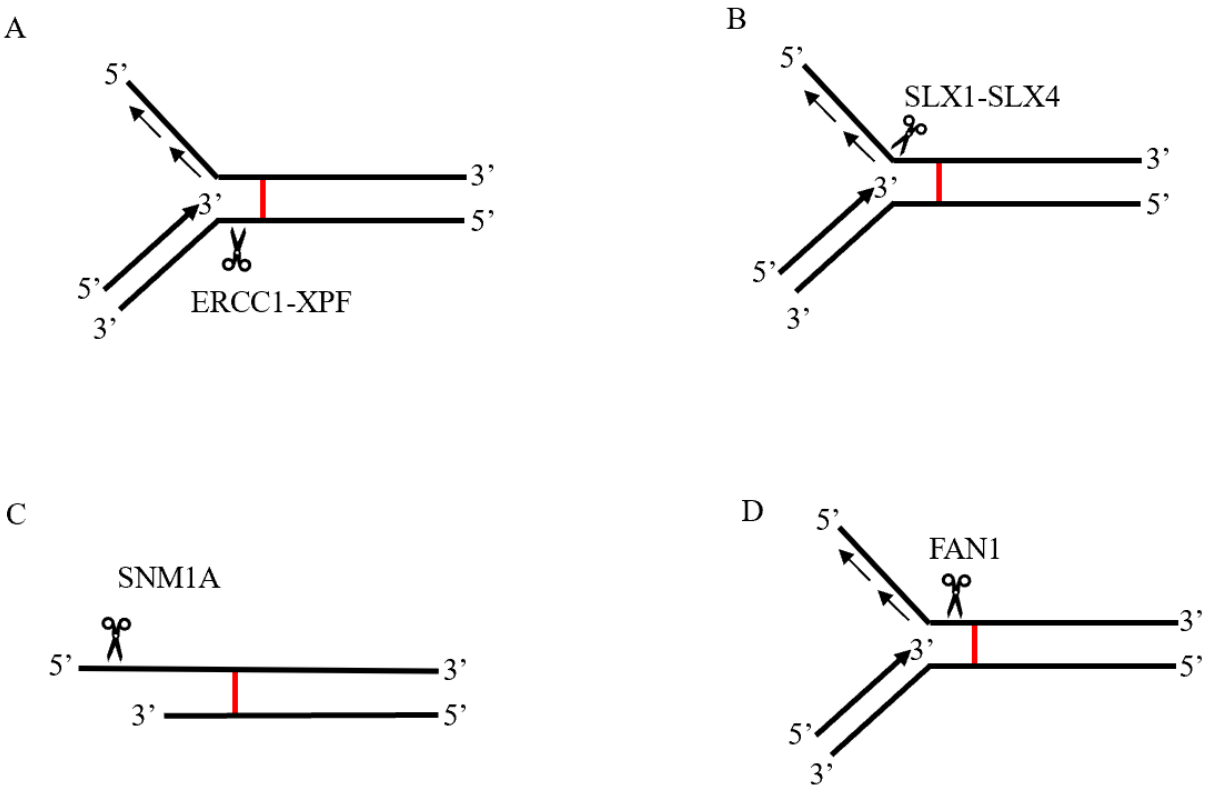


Figure 3. The incision site of structure-specific nucleases on the ICL-containing substrate. The schematic diagram indicates the location of the ICL (red line). (A). ERCC1-XPF nuclease can incise in double-stranded DNA on the 5′ side of such a junction (approximately two nucleotides away). (B). SLX1-SLX4 dimer is able to introduce a cut between ssDNA and dsDNA at replication fork junction. (C). SMN1A can trim on ssDNA with its 5‘-3’ endonuclease and exonuclease activity. (D). FAN1 is able to digest after replication fork junction 4nt.

1.6 Y-family Translesion DNA Polymerases

Replicative polymerases such as Pol δ are responsible for DNA replication in both the leading and the lagging strand (Johnson et al., 2015), and they possess 3' → 5' exonuclease activity as a proofreading function to ensure the fidelity of DNA replication. The replisome complex performs DNA replication in the 5' → 3' direction. If this process encounters any DNA lesions (such as base adducts, photoproducts, intrastrand or interstrand crosslinks), the DNA replication

fork may stall because the active site of the replicative polymerase δ cannot accommodate the ICL lesion (Bezalel-Buch et al., 2020). When the replication fork machinery is blocked, the helicases and replicative polymerases – the members of the replisome complex – can disassociate. The DDT pathway can be activated by replication protein A (RPA) to prevent further fork collapse (Davies et al., 2008). Y-family Translesion (TLS) DNA polymerases are a group of specialized polymerases that can be activated during DNA lesion bypass. The mammalian Y-family TLS polymerases includes REV1, Pol η, Pol ι, and Pol κ, and their protein structure is conserved. All of them contain the DNA-binding domain PAD and a catalytic site at the N-terminal domain and other protein-protein interaction domains at the C-terminus (Ling et al., 2001; Yang and Gao, 2018). TLS polymerases bypass DNA lesions in a "two-step fashion": a TLS polymerase adds a nucleotide opposite the DNA lesion and then another TLS polymerase extends the DNA from this point. Moreover, Pol η, Pol ι, and Pol κ possess a REV1-interacting-region (RIR), which enables them to interact with Rev1 in the bypass of DNA lesion (Yamanaka et al., 2017). This process is known as the Rev1-dependent pathway. REV1 is also able to serve as a platform to recruit TLS B-family Pol ζ for the extension part of lesion bypass. DNA damage bypass can also be performed via a REV1-independent pathway. Pol η, Pol ι, and Pol κ contain the proliferating cell nuclear antigen (PCNA)-interacting protein (PIP) and ubiquitin-binding motif (UBM)/ubiquitin-binding zinc finger (UBZ) domains, which allow them to interact with Ub-PCNA in the bypass of DNA damage (Farh et al., 2005; Haracska et al., 2001; Plosky et al., 2006).

1.7 Ubiquitinated PCNA constitutes a platform for replication fork rescue

Proliferating cell nuclear antigen (PCNA) is a homotrimer DNA clamp protein, which plays an essential role during DNA replication. PCNA is loaded onto the leading strand of the DNA replication machinery with the help of replication factor C (RFC). The RAD18 DNA damage tolerance (DDT) pathway can resolve stalled replication forks during DNA replication. Upon DNA lesion, the lysine 164 residue (K164) of PCNA can be monoubiquitinated by UBA1 (E1), RAD6 (E2), and RAD18 (E3). Monoubiquitinated PCNA acts as central hub to recruit UBZ domain-containing TransLesion synthesis (TLS) polymerases. Monoubiquitinated PCNA can also undergo polyubiquitination by MMS2–UBC13 through a lysine 63-linked ubiquitin chain, and HTLF can enhance the polyubiquitination reaction *in vitro* (Unk et al., 2008). Polyubiquitylated PCNA can

initiate the error-free DNA damage tolerance pathway template switching. The FA pathway is one of the main processes responsible for the repair of DNA ICL lesions at the S/G2 cell cycle checkpoint. In the FA pathway, FAN1 can be activated by the monoubiquitinated FANCI/FANCD2 heterodimer via its UBZ domain. FAN1 can incise the ICL-neighbouring region due to its structure-specific endonuclease activity and thus facilitate the bypass of the lesions. Surprisingly, depletion of FAN1 does not affect ICL-induced double-strand DNA break formation and does not lead to the development of FA either. Rather, germline FAN1 mutations cause caryomegalic interstitial nephritis (Lachaud et al., 2016). FAN1 also contains an uncharacterized PIP domain, which allows it to interact with PCNA to prevent replication fork collapse at G-quadruplexes (G4s) (Porro et al., 2017). DNA repair proteins are often recruited to specific sites of DNA damage through protein-protein interactions.

2. Goals and Objectives

The main goal of this study is to investigate whether FAN1 operates in the RAD18 DDT pathway and contributes to ICL bypass in cooperation with Ub-PCNA and TLS polymerases. The specific research questions are the followings:

-Does FAN1 interact with Ub-PCNA via its PIP box or UBZ domain?

-Can Ub-PCNA interaction affect the nuclease activity of FAN1?

- Are TLS polymerases able to cooperate with FAN1 for coordinated ICL lesion bypass and rescue of the stalled replication fork?

-Finally, to understand detailed molecular role of FAN1. I plan to reconstitute ICL bypass *in vitro* using substrates mimicking ICL and highly purified repair proteins.

In order to achieve the above goals, the following specific experiments are proposed:

- Purification of FAN1, FAN1 mutants' proteins and other DNA repair proteins (PCNA, TLS polymerases (η, κ, ι) by affinity chromatography
- Determining the domain that is responsible for the interaction between FAN1 and PCNA/Ub-PCNA by comparing FAN1 mutant derivatives:
    I. Examining physical protein-protein interactions with pull-down assays
    II. Investigating the effect of FAN1 nuclease activity in the presence of PCNA or Ub-PCNA and various DNA substrates
- Examining which Y-family TLS polymerase can cooperate with FAN1 in the rescue of the stalled replication *in vitro*
- Characterizing the function of TLS polymerase η with FAN1 in bypassing TMP-induced ICL substrates

3. Materials and Methods

3.1 Plasmids for protein expression

The pDONR201 plasmid carrying the FAN1-coding sequence was a gift from Dr. Jun Huang (Liu, T., et al. 2010). The point mutant derivatives of FAN1, with inactivated UBZ and/or PIP domains were obtained by the mutagenic PCR approach using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies) and appropriate pairs of mutagenic oligonucleotides carrying the desired mutations (Table 3 II).

For expressing FAN1 and its derivatives in yeast cells, they were cloned in fusion with glutathione S-transferase (GST) and FLAG tags under the control of the *Saccharomyces cerevisiae* galactose-inducible phosphoglycerate promoter using the Gateway cloning system (Invitrogen) with the LRII clonase enzyme and the pIL1844 (pBJ842 GST-FLAG-Destination) vector resulting in plasmids pIL2527, pIL3024, pIL2550, pIL3113 expressing the GST-FLAG-FAN1, and its derivatives carrying point mutations/deletions in their UBZ and/or PIP domains, respectively (Table 1). All plasmid constructs were verified by sequencing.

For localization studies and Western blot analyses, FAN1 and RAD18 cDNAs were cloned in fusion with N-terminal tags resulting in GFP-FAN1 (pIL2698), HA-FAN1 (pIL2528), FLAG-FAN1 (pIL2529), and FLAG-RAD18 (pIL2615).

Number	Name	Vector	Insert	Cloning method
PIL2519	FAN1	pDONR201	FAN1 WT	(Liu et al., 2010)
PIL3022	FAN1 PIP	pDONR201	FAN1 PIP (I30A/F34A, Y128A/F129A)	mutagenic PCR with O5635+O5636 on pIL2977
PIL2543	FAN1 UBZ	pDONR201	FAN1 ΔUBZ	(Liu et al., 2010)
PIL3110	FAN1 PIP/UBZ	pDONR201	FAN1 ΔUBZ- PIP (I30A/F34A, Y128A/F129A)	by mutagenic PCR with O5635+O5636
PIL2527	FAN1	pIL1844	FAN1 WT	LR recombination from pIL2519 to pIL1844

PIL3024	FAN1 PIP	pIL1844	FAN1 PIP (I30A/F34A, Y128A/F129A)	LR recombination from pIL3022 to pIL1844
PIL2550	FAN1 UBZ	pIL1844	FAN1 ΔUBZ	LR recombination from pIL2550 to pIL1844
PIL3113	FAN1 PIP/UBZ	pIL1844	GST-FLAG-FAN1-ΔUBZ- PIP (I30A/F34A, Y128A/F129A)	LR recombination from pIL3110 to pIL1844

Table 1 The list of plasmids used in this study.

3.2 Plasmid transformation

The pBJ842 yeast expression vector of GST- FLAG -FAN1 with – LEU selection marker was mixed with 300 µl 40 % Polyethylene glycol-lithium acetate and transformed into BJ5464 yeast competent cells. The mixture was incubated at 30 ºC for 30 minutes on the shaker. After incubation, 38.5 µl DMSO was added into the mixture for heat shock at 30 ºC for 15 minutes. After pelleting down the cells, the supernatant was removed and +12 media was added. The cells were incubated at 30 ºC overnight on the shaker. The cells were plated on the desired selection medium (such as -LEU).

3.3 Protein expression and purification

The GST-FLAG tagged FAN1 WT and mutant proteins were expressed in parallel in protease-deficient BJ5464 yeast competent cells. The cells containing the expression plasmid were collected after 7 hours of induction with 7 grams of D-galactose at 30 ºC. The cell pellets were resuspended in buffer A (150 mM Tris–HCl (pH 7.5), 500 mM NaCl, 150 mM KCl 8.5 mM 2-mercaptoethanol, 30 % sucrose and 1.5 mM EDTA) and were disrupted by a 6775 Freezer/Mill® Cryogenic Grinder. The supernatant was collected by centrifugation for 30 minutes at 50,377 g at 4 ºC. The supernatant was loaded into the filter containing 50 µl Glutathione Sepharose 4B resin and then incubated at 4 ºC for 90 minutes. The Glutathione Sepharose 4B resin was pre-

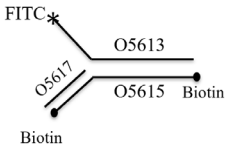
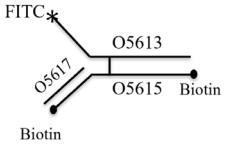
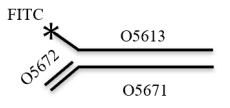
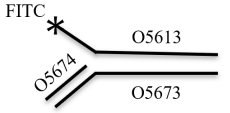
equilibriated with buffer A. Following the immobilization of the GST- FLAG tagged protein on the beads, the beads were washed with an 80 column-volume portion wash buffer containing decreasing salt concentration (20 mM Tris–HCl (pH 7.5), 500 mM -150 mM NaCl, 0.01 % NP 40, 0.1 mM and 10 % glycerol). Proteins were eluted with PreScission Protease from the beads in wash buffer containing 150 mM NaCl.

GST-fussed human PCNA protein and GST-tagged human DNA polymerases η, κ, ι, were expressed in protease-deficient BJ5464 yeast competent cells as GST-fusion proteins, from which GST was removed by Prescission protease during the elution step of purification as described (Haracska et al., 2001, 2001, 2002). The expression and purification of RFC complex was following the protocol from Gomes et al., 2000.

3.4 DNA substrates and DNA sequences

3.4.1 Preparation of structure-specific oligonucleotides

Non-single-stranded DNA was annealed in TE buffer. In order to get rid of the non-annealed oligonucleotides, the samples were run on a 4 % non-denaturing polyacrylamide gel and imaged using the Typhoon Trio Imager. The desired substrates were isolated by cutting them out from the gel. The isolated gel parts were soaked in water overnight at 4 ºC. The sequence of the oligonucleotides are listed below:

73 nt 5'-flap DNA	
73 nt 5'-flap interstrand-crosslinked DNA	
73 nt 5'-flap DNA with 20-nt flap length	
73 nt 5'-flap DNA with 25-nt flap length	

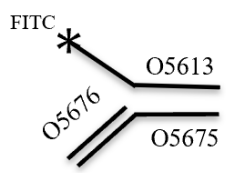
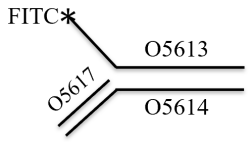
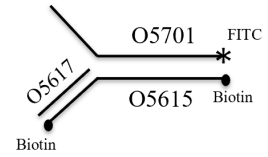
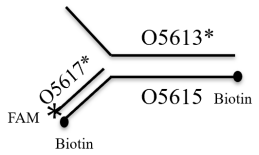
73 nt 5'-flap DNA with 40-nt flap length	
73 nt 5'-flap DNA with 30-nt flap length	
73 nt 5'-flap DNA with 30-nt flap length and 3' labeled with FITC	
73 nt 5'-flap DNA with 30-nt flap length, FAM on the leading strand	

Table 2 DNA substrates used in this study

3.4.2 Preparation of site-specific interstrand-crosslinked oligonucleotides

The site-specific interstrand-crosslinked substrate was prepared according to the following protocol: the 87.6 mM 4,5',8-trimethylpsoralen (TMP, Sigma, T6137) crosslinking agent was applied alone with 2.6 mM annealed DNA in crosslinking buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, and 50 mM NaCl) under a UVA source (4.2 mWm^{-2}) wavelength of 365 nm. The reaction was repeated for 5 cycles at 15 minutes intervals. Our protocol is a modification of a previously published method (Liang *et al*., 2016). ICL was confirmed by 8 M urea 8% denaturing polyacrylamide gel electrophoresis.

3.4.3 DNA sequences

I.

Name	Sequence	Length	5' labeling	3' labeling
O5613	5'GTTTTCCGAGTCACGACGCCGCTCCG GAACTCGCGTAGGCTTCTCACGACTTC TCGAGGAAGGCTCGGCGGCT 3'	73 nt	FITC	None
O5614	5'AGCCGCCGAGCCTTCCTCGAGAAGT CGTGAGAAGCCTACGCGACCGTTCTTC GCCTGGCGGACTGCCTTCCCG 3'	73 nt	None	None

O5615	5’AGCCGCCGAGCCTTCCTCGAGAAGT CGTGAGAAGCCTACGCGACCGTTCTTC GCCTGGCGGACTGCCTTCCCG 3'	73 nt	Biotin	Biotin
O5617	5’CGGGAAGGCAGTCCGCCAGGCGAAG AACGG 3’	30 nt	None	None
O5635	5’AAAAGCATCTAATTCTGCTATTTCGT GTGCTAACAATGCACCACCTGCTA ‘3	50 nt	None	None
O5636	5’TAGCAGGTGGTGCATTGTTAGCACAC GAAATAGCAGAATTAGATGCTTTT ‘3	50 nt	None	None
O5671	5’AGCCGCCGAGCCTTCCTCGAGAAGT CGTGAGAAGCCTACGCGAGTTCCGGA GCCCTGGCGGACTGCCTTCCCG 3’	73 nt	None	None
O5672	5’ CGGGAAGGCAGTCCGCCAGG 3’	20 nt	None	None
O5673	5’AGCCGCCGAGCCTTCCTCGAGAAGT CGTGAGAAGCCTACGCGAGTTCCCTTC GCCTGGCGGACTGCCTTCCCG 3’	73 nt	None	None
O5674	5’CGGGAAGGCAGTCCGCCAGGCGAAG 3’	25 nt	None	None
O5675	5’AGCCGCCGAGCCTTCCTCGAGAAGT CGTGAGAACTAAGGTAACCCGTTCTTC GCCTGGCGGACTGCCTTCCCG 3’	73 nt	None	None
O5676	5’CGGGAAGGCAGTCCGCCAGGCGAAG AACGGGTTACCTTAG 3’	40 nt	None	None
O5701	5’AGCCGCCGAGCCTTCCTCGAGAAGT CGTGAGAAGCCTACGCGAGTTCCGGA GCCCTGGCGGACTGCCTTCCCG 3’	73 nt	None	FITC
O5613*	5’GTTTTCCGAGTCACGACGCCGCTCCG GAACTCGCGTAGGCTTCTCACGACTTC TCGAGGAAGGCTCGGCGGCT 3’	73 nt	None	None
O5617*	5’CGGGAAGGCAGTCCGCCAGGCGAAG AACGG 3’	30 nt	FAM	None

II.

Name	Sequence	Purpose
O5635	5'AAAAGCATCTAATTCTGCTATTTCGT GTGCTAACAATGCACCACCTGCTA 3'	Forward mutagenic primer for the PIP I30A/F34A of FAN1
O5636	5'TAGCAGGTGGTGCATTGTTAGCACAC GAAATAGCAGAATTAGATGCTTTT3'	Reverse mutagenic primer for the PIP I30A/F34A of FAN1
O3378	5'GTAAAGCAGAAGATCAGTCCCGCCG CTAAAAGTAATGATGTGGTGTG3'	Forward mutagenic primer for PIP Y128A/F129A of FAN1
O3379	5'CACACCACATCATTACTTTTAGCGGC GGGACTGATCTTCTGCTTTAC 3'	Reverse mutagenic primer for the PIP Y128A/F129A of FAN1

Table 3 (I). The sequences used to prepare structure-specific oligonucleotides, and (II) oligos for mutagenic PCR.

3.4 Gels

In my studies, I used various polyacrylamide gels for different experimental purposes. SDS-PAGE: Sodium dodecyl sulfate (SDS) is able to denature the quaternary structure of protein, and the percentage of SDS-PAGE was set according to the size of the protein. 4 % native PAGE was used to separate and isolate DNA substrates of different size and structure. 10 % (8 M) denaturing urea PAGE was employed to separate the products of the nuclease and TLS assays.

3.5 Nuclease and TLS assays

3.5.1 *In vitro* nuclease assay

FAN1 nuclease assays were performed with incubation at 37 °C for 15 minutes in a reaction buffer containing 25 mM Tris pH 7.5, 150 mM sodium chloride, 15 mM potassium chloride, 7.5 mM magnesium chloride, 1 mM dithiothreitol. 30 nM of fluorescein-labeled substrate, a constant 740 nM of the FAN1 and mutant proteins, a constant 43 nM RFC, and increasing amounts of PCNA or mono-ub-PCNA (60 nM-600 nM) were applied in the reaction and then stopped with formamide loading buffer (95% formamide, 10 mM EDTA). The arm length of 73-nt-long 5'-fluorescein labeled flap substrate applied in the nuclease assay was 30 nt, or indicated otherwise. Products of the FAN1 WT and mutant proteins were separated using 16 % denaturing PAGE and detected at 526 nm using the Typhoon Trio Imager.

3.5.2 *In vitro* TLS polymerase assay

The TLS polymerase reactions were incubated at 37 ºC in the reaction buffer (25 mM Tris pH 7.5, 50 mM sodium chloride, 15 mM potassium chloride, 7.5 mM magnesium chloride, 1 mM dithiothreitol). The incubation time is indicated in the figure legend. The concentrations used for this assay were as follows: 740 nM FAN1, 43 nM RFC, 600 nM monoubiquitinated PCNA, 24 nM Pol η/Pol ι/Pol κ, 100 μM dNTPs. Aliquots were withdrawn at the indicated time points, stopped with formamide loading buffer (95 % formamide, 10 mM EDTA). The reactions were run in 10 % denaturing PAGE and detected at 526 nm using the Typhoon Trio Imager.

3.6 Immunoprecipitation and western blot analysis

The GST-FLAG-tagged FAN1 and GST-FLAG-tagged FAN1 UBZ mutants were expressed under the same conditions as described in 3.3. The cell lysate of the two proteins were loaded on the columns separately, and GST-FLAG-tagged FAN1 and GST-FLAG-tagged FAN1 UBZ proteins were immobilized on glutathione-Sepharose beads. The beads were washed with an 80x column-volume of gradually decreasing wash buffer B (20 mM HEPES–HCl (pH 7.4), 500 mM~150 mM NaCl, 0.01 % NP 40, 1 mM dithiothreitol, and 10 % glycerol). The GST-FLAG-tagged proteins bound to glutathione-Sepharose beads (15 μl) along with highly purified PCNA and monoubiquitinated PCNA (100 nM each) were incubated for 90 minutes at 4 ºC in buffer B containing 150 mM NaCl. Beads were washed three times with buffer B, and bound proteins were eluted in buffer B containing 150 mM NaCl and 20 mM reduced glutathione. Eluted fractions were analyzed by immunoblotting as indicated.

The GST-tagged PCNA and GST-Flag-tagged fusion-Ub-PCNA proteins were expressed under the same conditions as the FAN1 proteins. The cell lysates of the two proteins were loaded on the columns separately, and GST- tagged PCNA and GST-FLAG tagged fusion-Ub-PCNA proteins were immobilized on glutathione-Sepharose beads. The beads were washed with an 80x column-volume portion of gradually decreasing wash buffer B (20 mM HEPES–HCl (pH 7.4), 500 mM~150 mM NaCl, 0.01 % NP 40, 1 mM dithiothreitol, and 10% glycerol). The GST-tagged proteins bound to glutathione-Sepharose beads (15 μl) along with FAN1 and mutant proteins (100 nM each) were incubated for 90 minutes at 4°C in buffer B containing 150 mM NaCl. Beads were washed three times with buffer B, and bound proteins were eluted in buffer B containing 150 mM

NaCl and 20 mM reduced glutathione. Eluted fractions were analyzed by immunoblotting as indicated.

After running the SDS-PAGE gel, proteins were transferred onto blotting membrane. The blots were placed into 5 % BSA blocking solution for 1 hour at RT. Anti-FLAG M2 monoclonal antibody conjugated to HRP (Sigma-Aldrich) targeted the FAN1 proteins, and Anti-GST HRP conjugated antibody (GE healthcare) was able to recognize GST-tagged proteins.

3.7 Cell culture and cell transfection

HCT116 and HEK293 cells were grown in Dulbecco's modified Eagle's medium (Sigma) supplemented with 10% FCS (Sigma) at 37 ºC. Transfections were carried out using Lipofectamine 2000 transfection reagent (Invitrogen) according to the instruction of the manufacturer. After treatment with UV (20 J/m^2) or Cisplatin (50 μM), cells were incubated for 6 hours and immunostaining where indicated. For the cellular localization of FAN1 and PCNA, cells were treated with a detergent solution (10 mM TRIS-HCl, 2.5 mM $MgCl_2$, 1 % NP-40, 1 mM PMSF) and immunostaining by using anti-HA antibody (Cell signaling, C29F4) diluted 1:300 and Alexa Fluor 488 (Invitrogen, A11011) diluted 1:500, anti-PCNA antibody (Santa Cruz, sc-56) diluted 1:200, and Cy3-conjugated anti-mouse antibody (Sigma, catalogue number: C2181) diluted 1:1000.

RNAi and stable cell lines: HCT116 RAD18 -/- cell line was described previously (Shiomi et al., 2007). To complement the HCT116 RAD18 -/- cells, we used a DsRed-RAD18 plasmid construct.

HEK293-derived stable-silenced cell lines (shFAN1) were cultured in Dulbecco Modified Eagle Medium (DMEM) supplemented with 10 % fetal bovine serum (FCS), 100 U/ml penicillin, and 100 μg/ml streptomycin. Performing lipofectamine (Invitrogen)-mediated transfection according to the manufacturer's protocol and using 300 μg/ml G418 for selection, we established a cell line (shFAN1) stably expressing shRNA targeting FAN1. Silencing was detected by Western blot using anti-FLAG antibody (SIGMA M2A8592,1:3000) for detecting transiently expressed FLAG-FAN1.

3.8 Immunostaining and microscopy

Cells were cultured on coverslips, the medium was removed, and cells were washed two times with PBS. For detection of PCNA or UB-PCNA together with fluorescently labelled proteins, a pre-extraction step was performed (J. K. Hicks et al., 2010). Nuclei were pre-extracted by incubating the coverslips with pre-extraction buffer (20 mM HEPES pH 7.4, 50 mM NaCl, 1 mM EDTA, 3 mM $MgCl_2$, 300 mM sucrose, 0.5% Triton-X-100) for 15 minutes at 4 ºC. Nuclei were then fixed with 4 % formaldehyde/PBS for 10 minutes at room temperature, washed once with PBS, and then fixed again in ice-cold 100 % methanol to expose the PCNA (PC-10) epitope. Coverslips were incubated in 5 % BSA in PBS for minimum 2 hours. In other cases, after the paraformaldehyde fixation, cells were permeabilized with 0.2 % Triton X-100 in PBS for 5 minutes at 4 ºC. Samples were mounted with Fluoromount (PBS and analyzed with an Olympus FV1000 or Leica LS< microscope). Antibodies: anti-PCNA: Santa Cruz (sc56), anti-Ub-PCNA (lys164): Cell signaling D5C7P 13439S

3.9 Co-immunoprecipitation

Cells (2.5×10^6) were lysed in a buffer containing 50 mM Tris–HCl (pH 7.5), 5 mM EDTA, 150 mM NaCl, 0.1 % NP-40, 1 mM PMSF, and protease inhibitor cocktail (Sigma P8340). To reduce viscosity, the cell lysates were sonicated on ice. FLAG-tagged proteins were immunoprecipitated with immobilized M2 FLAG antibodies (Sigma). The precipitated proteins and the input lysates were analyzed by Western blot using mouse anti-PCNA HRP (Santa Cruz catalogue number: sc-56), mouse anti-FLAG HRP (Sigma M2 A8592), rat anti-HA HRP (Roche, catalogue number: 12013819001 clone 3F10), rabbit anti-Tubulin (Santa Cruz catalogue number: sc-9104) and goat anti-rabbit HRP (Millipore AP132P) antibodies.

3.10 Bimolecular fluorescence complementation assay

In order to examine the physical interactions between FAN1 and Pol η *in vivo*, the yellow fluorescent protein (YFP)-based bimolecular fluorescence complementation (BiFC) analysis was used (Kerppola, 2008). cDNA sequences coding for FAN1 derivatives and Pol η were fused to

Venus variant YFP fragments V1 and V2, respectively. Human cell lines (U2OS, HeLa) co-transfected with these constructs were analyzed with an ImageXpress high-throughput microscope using the MetaXpress software package (Molecular Device). First, 100 photos were taken using a 40× Plan Fluor ELWD objective. Total cell number was determined by counting DAPI-stained nuclei. In each experiment, at least 1000 cells were examined. For detection of interaction, YFP spots of 1-2 μm size with at least 2000× intensity above local background were counted. Spots located outside nuclei were omitted from further analysis. Then, the ratio (fold change) of YFP positive cells possessing at least 5 foci compared to cells carrying V1-FAN1 and V2-GUS constructs was calculated.

3.11 Resazurin-based cell viability assay

Cell viability was detected by a resazurin-based fluorometric assay. Briefly, HEK293 cells were seeded at 7×10^5 cells/well in 6-well plates. After 24 hours, cells were transfected with the appropriate shRNA-expressing construct, described above. One day after transfection, cells were harvested and seeded in a 96-well plate at a density of 4×10^3 cells/well in a volume of 100 μl of complete DMEM. Next day, cells were treated with different concentrations of cisplatin ranging from 0 to 40 μM (added in a volume of 100 μl to the 100 μl culturing media (2× dilution of cisplatin)) and incubated for an additional 48 hours. Experiments were carried out in hexaplicates for each drug concentration. After 2 days of treatment, 40 μl of resazurin solution in DMEM (0.15 mg/ml, Sigma, R7017-1G) was added to each well. After 5 hours of incubation at 37 ºC, cell viability was monitored by measuring fluorescence with excitation wavelength at 542 nm and emission wavelength at 590 nm in a Fluoroskan Ascent FL (Thermo Scientific) fluorimeter. The fluorescent signal generated from the assay is directly proportional to the number of living cells in the sample. Percentages of viability of cisplatin-treated cells were calculated relative to untreated control cells after calibration to background.

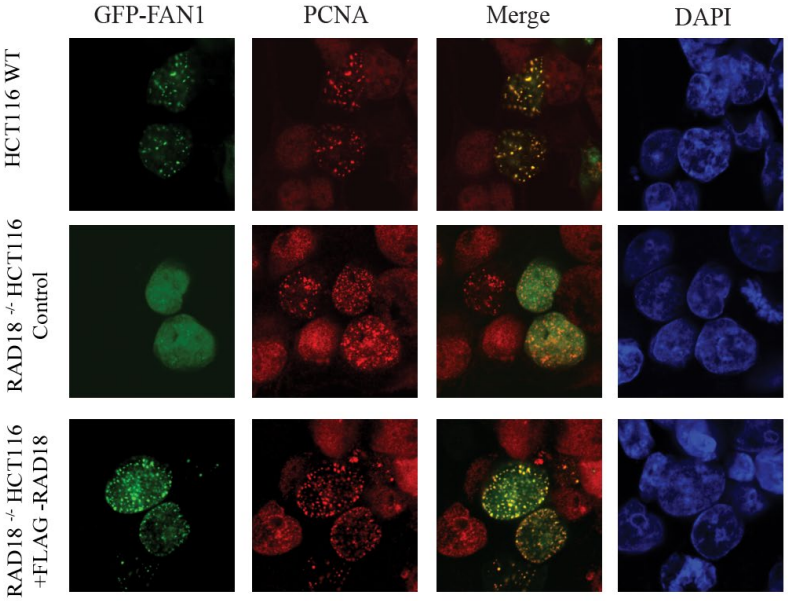
4. Results

4.1 FAN1 is involved in the RAD6/RAD18 DDT pathway

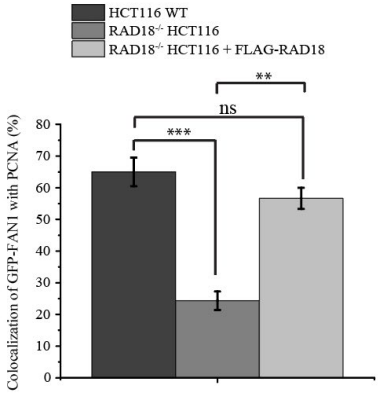
To examine whether FAN1 is involved in the RAD18 DDT pathway, we used the HCT116 cell line (colon cancer cell line) and UV irradiation as DNA damaging factor. After UV treatment, we could observe GFP-tagged FAN1 foci formation as FAN1 was recruited to the site of damage where it co-colocalized with PCNA (Figure 4A first line). In Rad18-/- knockout cells, however, GFP-FAN1 lost its ability to co-localize with PCNA. When a FLAG-RAD18-expressing construct was transfected into RAD18-/- cells, FAN1 foci formation could be detected again at the site of damage. Following image quantification, these results lead to the conclusion that FAN1 is regulated by the RAD18 DDT pathway and is able to co-localize with PCNA upon UV treatment (Figure 4B). Furthermore, we also performed co-immunoprecipitation of RAD18 by FLAG-FAN1 *in vivo*. HA-RAD18 and FLAG-FAN1 were co-expressed in HEK 293 cells. After cell lysis, FLAG-FAN1 immobilized on the anti-FLAG-beads along with RAD18 (Figure 4C). We conclude that FAN1 and RAD18 can act together in DNA repair.

In the RAD18 DDT pathway, the homotrimer PCNA can be monoubiquitinated by RAD6/RAD18 on K164. The ubiquitinated PCNA can further recruit UBZ-domain-containing DNA repair proteins. A previous report indicated that the ubiquitin-PCNA-dependent FAN1 stabilizes G4 structures to prevent replication fork collapse via its PIP-box and UBZ domains (Porro et al., 2017). FAN1 contains PIP-box and ubiquitin-interacting domains at the N-terminal, an SAP-type DNA-binding domain in the middle, and highly conserved VRR-nuclease domain at the C-terminus (Figure 4D). In this study, we mapped possible functional domains of FAN1 by using the VectorNTI program. In detail, we constructed a FAN1 PIP point mutant by combining a PIP box (I30A/F34A) from Porro et al and a mapped PIP box (Y128A/F129A). We also made FAN1 UBZ deletion mutant and FAN1 double mutant (PIP/UBZ). To investigate whether FAN1 is involved in the ubiquitin-PCNA-mediated RAD18 DDT pathway, we examined FAN1, FAN1 PIP, FAN1 UBZ, and FAN1 PIP/UBZ double mutants in the colocalization study. The results indicated that FAN1 does not form nuclear foci in undamaged cells, while it forms clear nuclear foci and colocalizes with ubiquitinated PCNA upon cisplatin treatment. However, the various mutants completely lost DNA-damage induced FAN1 foci formation and did not colocalize with ubiquitinated PCNA (Figure 4E).

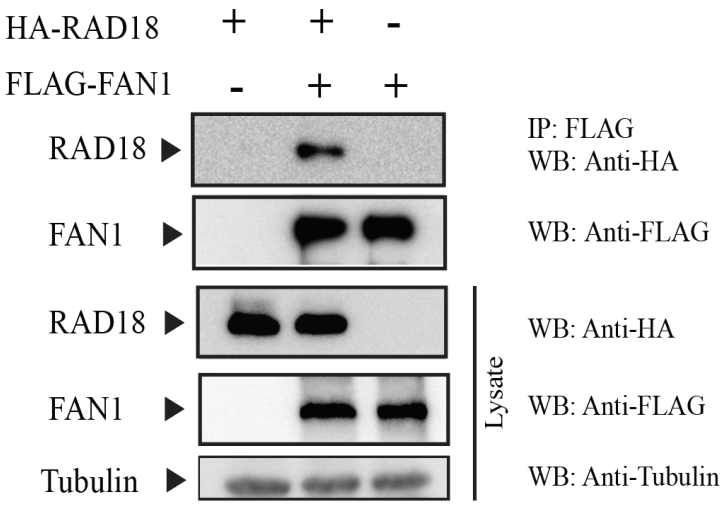
A


GFP-FAN1
PCNA
Merge
DAPI
HCT116 WT
RAD18 -/- HCT116 Control
RAD18 -/- HCT116 +FLAG -RAD18

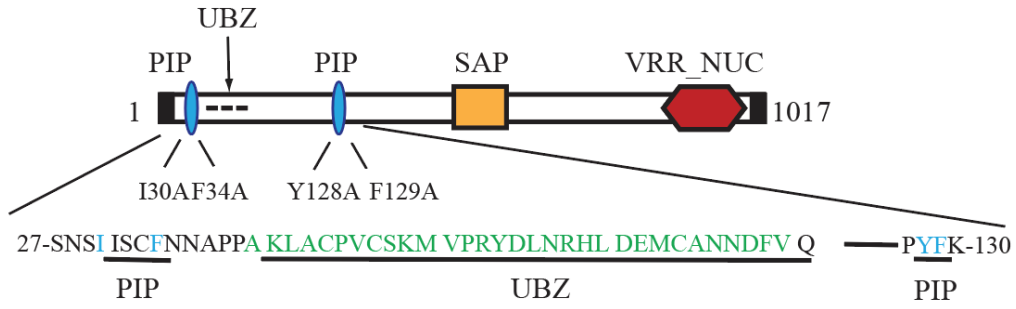
B


HCT116 WT
RAD18-/- HCT116
RAD18-/- HCT116 + FLAG-RAD18
**
ns
***
Colocalization of GFP-FAN1 with PCNA (%)
80
70
60
50
40
30
20
10
0

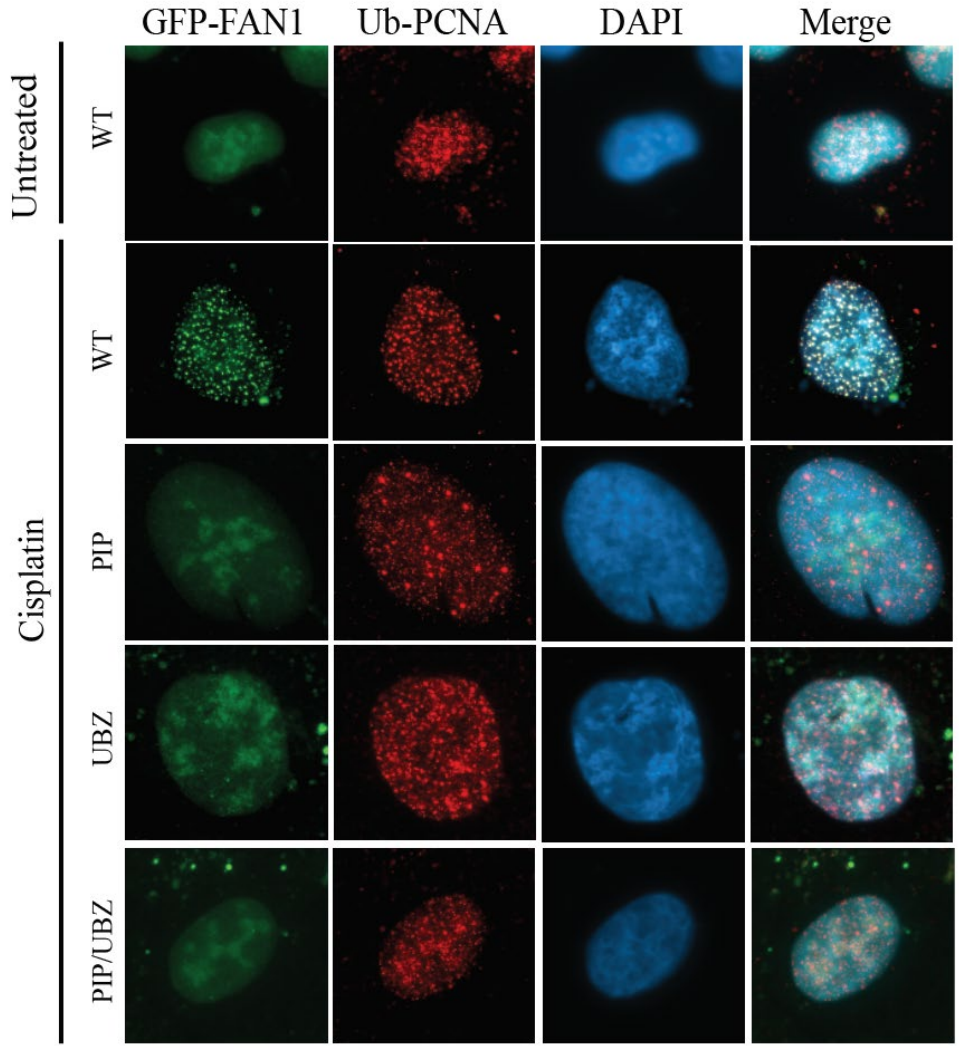
C.


HA-RAD18 + + -
FLAG-FAN1 - + +
RAD18
FAN1
RAD18
FAN1
Tubulin
IP: FLAG
WB: Anti-HA
WB: Anti-FLAG
Lysate
WB: Anti-HA
WB: Anti-FLAG
WB: Anti-Tubulin

D.


UBZ
PIP
PIP
SAP
VRR_NUC
1
1017
I30AF34A
Y128A F129A
27-SNSI ISCFNNAPPA KLACPVCSKM VPRYDLNRHL DEMCANNDFV Q
PYFK-130
PIP
UBZ
PIP

E.


GFP-FAN1
Ub-PCNA
DAPI
Merge
Untreated
WT
Cisplatin
WT
PIP
UBZ
PIP/UBZ

F.

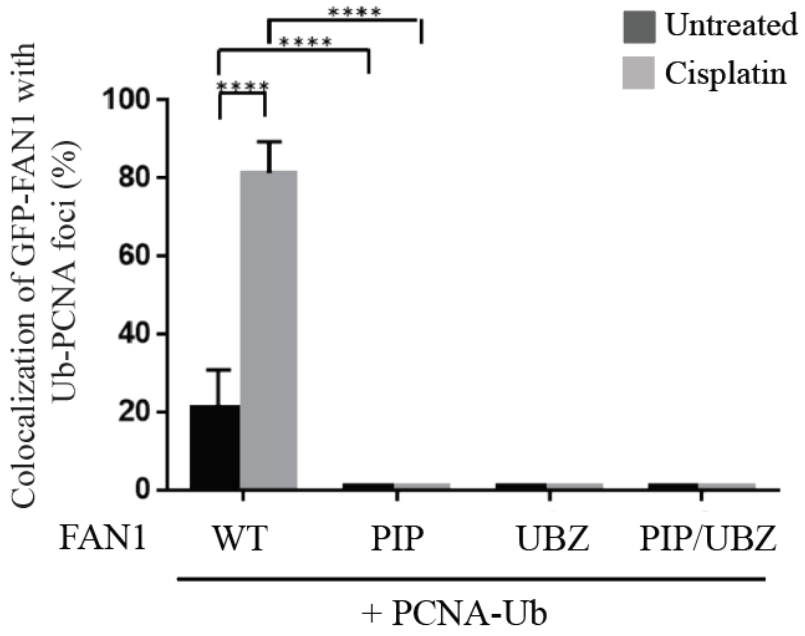


Figure 4. (A) The localization of FAN1 depends on RAD18. Knockout HCT116 RAD18-/- and HCT116 WT cells were transfected with plasmids expressing GFP-FAN1. FAN1 localization was compared in these two cell lines and in a cell line in which Rad18 was complemented with a FLAG-RAD18 plasmid construct. The cells were immunostained with anti-PCNA antibody and the localization of FAN1 was compared. (B) Quantification of FAN1 localization in RAD18 knockout cells. The percentage of PCNA – FAN1 colocalization was determined from three independent experiments, and standard deviation was calculated. (C) FAN1 associates with RAD18 *in vivo*. HEK293 cells were transfected with HA-RAD18 and FLAG-FAN1. Cell extract was subjected to immunoprecipitation with anti-FLAG antibody, and the co-immunoprecipitated RAD18 was detected by western blotting using anti-HA antibody. Tubulin was used as loading control. (D) The schematic representation of important residues of human FAN1 for interacting with mono-ubiquitinated PCNA in blue (PIP-box) and dash line (UBZ-domain). (E). Colocalization of GFP-FAN1 WT and FAN1 mutants with Ub-PCNA. Representative fluorescent microscopical images of colocalization of GFP-FAN1 or different GFP-tagged FAN1 mutants (untreated and cisplatin-treated) with Ubiquitin-PCNA. GFP-FAN1 and FAN1 mutants were transiently expressed in U2OS cells. Cells were treated with 30 μM cisplatin, washed and left to

recover for 1 h, fixed and immunostained with anti-Ub-PCNA antibody. (F). Quantitative measurement of experiment. Scanning and measurements were taken with an ImageXpress Micro Confocal microscope using a 40x ELWD objective and analyzed with the MetaXpress software package (Molecular Devices). Error bars represent standard deviations. Statistical analysis was done by 2way ANOVA with Turkey's multiple comparisons test (N=3). (****) $p<0.0001$; (**) $p<0.01$.

4.2 Pull-down assay and purification of the FAN1 protein and FAN1 mutants by affinity chromatography

The affinity chromatography technology is able to separate the biomolecule of interest from a mixture based on the specific type of binding interaction. In my experimental setup, the Glutathione Sepharose 4B resin was applied for binding the GST-FLAG-tagged FAN1 protein from the cell lysate. Affinity chromatography provides a highly specific, strong interaction between the resin and the target protein. After removing non-specific bound proteins with wash buffer, the GST-Flag-tagged FAN1 was eluted with PreScission Protease or 20 nM glutathione. Taking advantage of this affinity chromatography technique, I was able to investigate the direct interaction between FAN1 and PCNA/Ub-PCNA. I immobilized GST-Flag-tagged wild-type and UBZ-mutant FAN1 proteins as the "bait" protein on the glutathione agarose beads and then added highly purified PCNA or Ub-PCNA as the "prey" protein in the GST-pull down assay. This affinity system revealed that from a one-to-one mixture of PCNA and Ub-PCNA, FAN1 bound the ubiquitinated form of PCNA preferentially and that the mutation in the FAN1 UBZ domain completely disrupted this binding (Figure 5). Moreover, when we compared the binding affinity of FAN1 to GST-PCNA and GST-Ub-PCNA, we detected a higher affinity to Ub-PCNA than to PCNA. In addition, while the PIP mutations impaired FAN1 binding to PCNA, its ability to bind to Ub-PCNA remained partially, indicating that FAN1 can bind to PCNA via its two PIP domains and to ubiquitin via its UBZ domain (Figure 6). Taken together, we provide evidence that the PIP and UBZ binding sites of FAN1 provide direct and preferential Ub-PCNA binding, which is consistent with a model suggesting that after ICL-induced replication fork stalling, activated RAD18 can ubiquitinate PCNA, which can then recruit FAN1 to fork rescue.

Next, I purified FLAG-tagged wild-type FAN1 and its various mutant proteins (Figure 7). The cell lysate containing the protein of interest was loaded onto the Glutathione Sepharose 4B resin. After intensive washing of the beads, the GST-FLAG-tagged FAN1 protein was eluted with PreScission Protease. This specific protease is able to cleave between the Gln and Gly residues of the core amino acid sequence Leu-Phe-Gln/Gly-Pro. In my experimental setup, this core amino acid sequence is located between the GST tag and the FLAG tag, therefore, the GST tag will be removed in the last step of protein purification in the presence of PreScission Protease. The purified FAN1 proteins were subjected to further investigations.

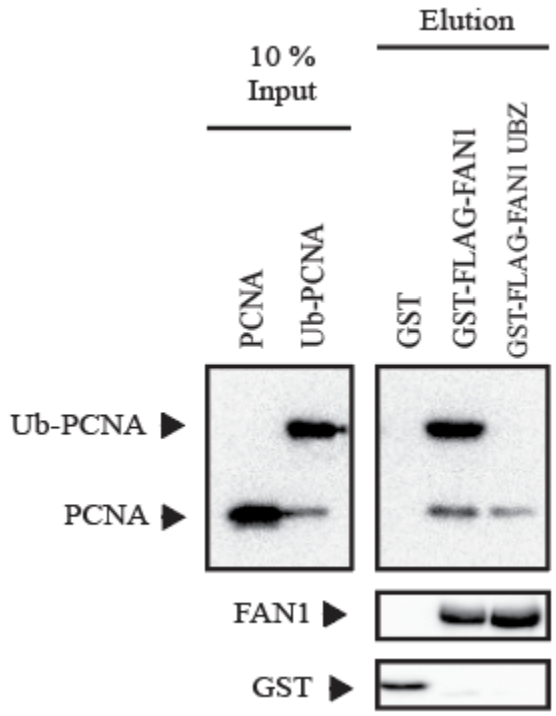


Figure 5. Pull-down assay. GST-FLAG-FAN1, and GST-FLAG-tagged FAN1 UBZ mutant bound to Glutathione Sepharose 4B resin were incubated with purified PCNA or monoubiquitinated PCNA in a 150 mM NaCl-containing buffer. After elution by 20 mM glutathione, samples were analyzed for direct physical interaction between FAN1 and PCNA/Ub-PCNA with anti-PCNA and anti-GST antibodies after western blotting.

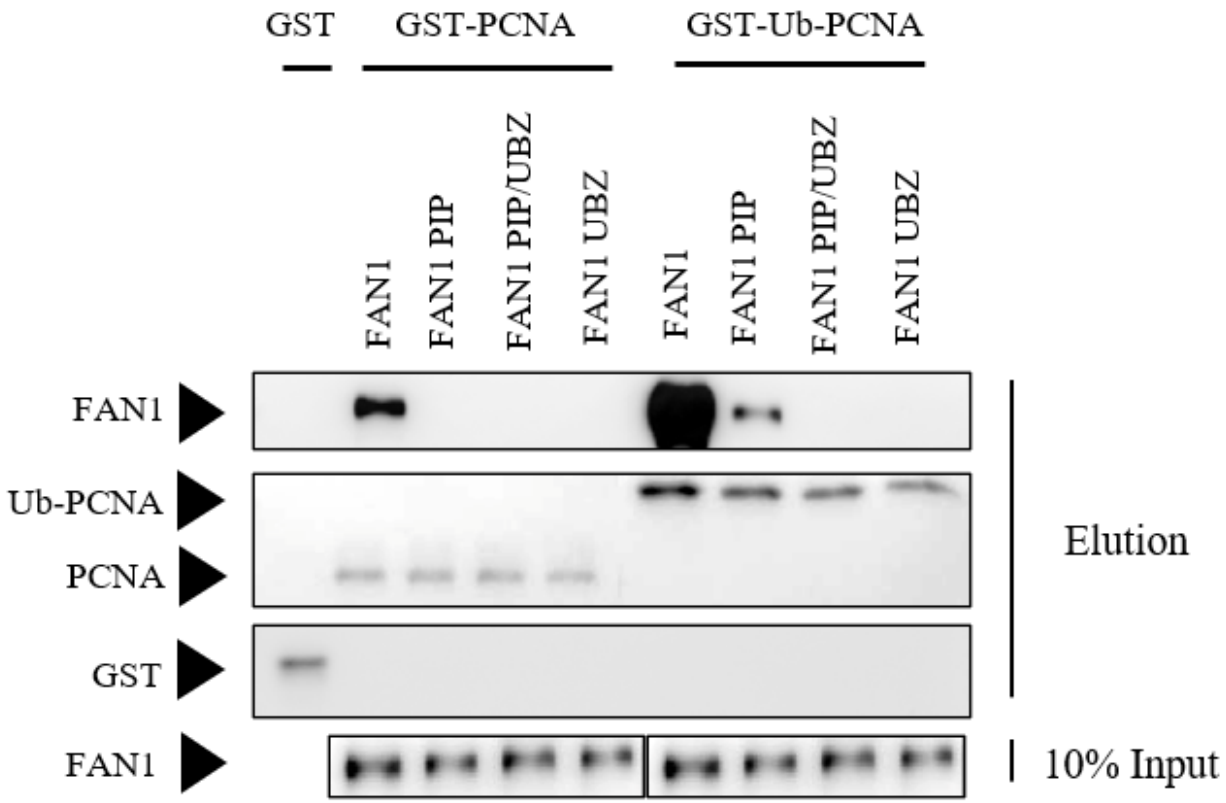


Figure 6. Pull-down assay: purified FLAG-FAN1 preferentially binds to GST-Ub-PCNA. GST-PCNA, and GST-Ub-PCNA bound to Glutathione Sepharose 4B resin were incubated with purified FLAG-FAN1 and its mutants in a 150 mM NaCl-containing buffer. After elution by 20 mM glutathione, samples were analyzed for direct physical interaction between FAN1 and PCNA/Ub-PCNA with anti-FLAG and anti-GST antibodies after western blotting.

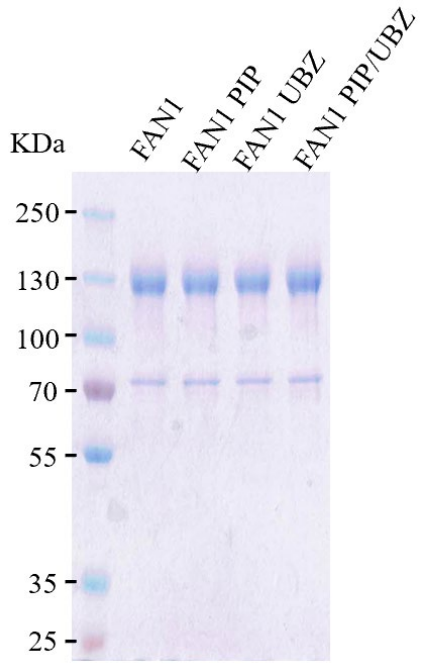


Figure 7. The purified human FAN1 WT and FAN1 mutants (800 ng) were subjected to 8% SDS–PAGE, and the gel was stained with Coomassie Brilliant Blue

4.3 The nuclease activity of FAN1 can be stimulated by PCNA via PIP-box

The function of FAN1 is considered as key step of ICL bypass. Because FAN1 possesses 5'flap endonuclease activity and 5' exonuclease activity by its VRR_nuclease domain (MacKay et al., 2010) , FAN1 is able to incise the ICL DNA to generate unhooked substrate for TLS polymerase. To investigate further these properties of FAN1, I designed a 73nt-long 5'-Fluorescein labeled flap substrate for nuclease assay (Figure 8). Firstly, I have compared the basic nuclease activity of FAN1 WT and its various mutants (Figure 9). All of the FAN1 preps exhibit the same level of nuclease activity and process the substrate by two cuts (indicated by "a" and "b"). Secondly, I introduced RFC and biotin-streptavidin on both sides of leading template strand to load increased concentration of purified PCNA and prevent its sliding off from the same DNA substrate (Figure 8). We could observe the enhanced nuclease activity of FAN1 WT with increased PCNA concentration (Figure 10, lane 4-6). We presumed that the stimulation of nuclease activity is due to the interaction of FAN 1 and PCNA via PIP-box. We have changed Y128 and F129 to alanine and mutated I30 and F34 to alanine as previously described by Antonio Porro *et al*. (2017).

These mutants allowed us to prove our hypothesis *in vitro*. The nuclease activity of PIP mutant FAN1 failed to be enhanced by PCNA (Figure 10, lane 8-10). So far, we obtain the insight that the nuclease activity of FAN1 can be stimulated by PCNA via PIP-box.

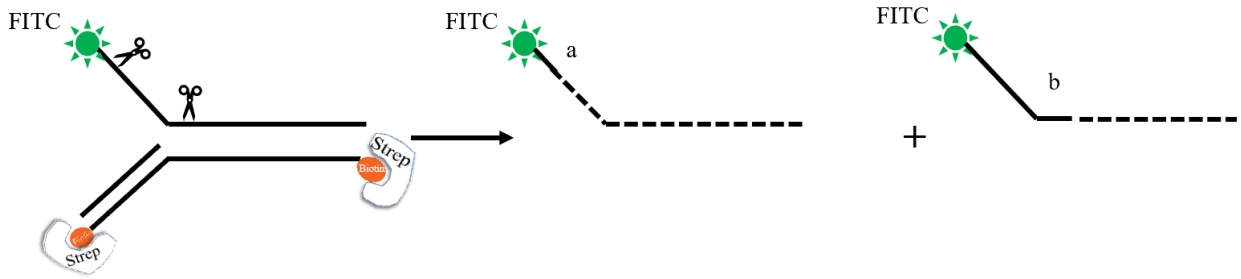


Figure 8. Schematic DNA substrate, red arrow indicates the incision site of FAN1. Strep: streptavidin.

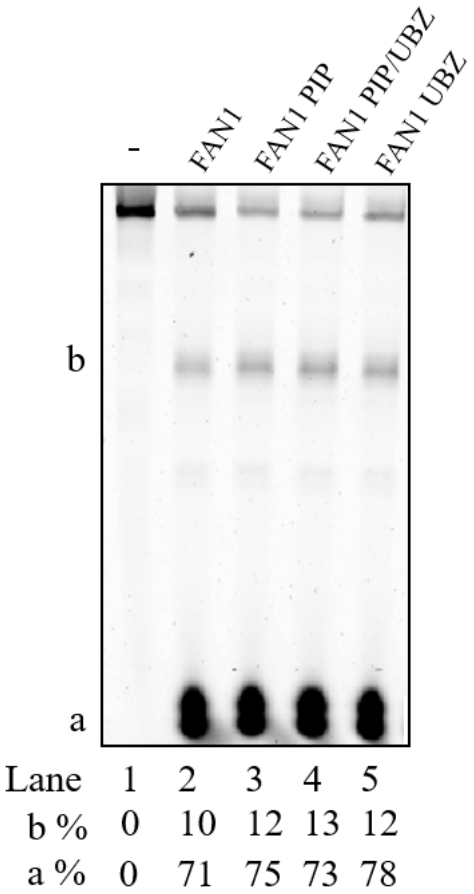
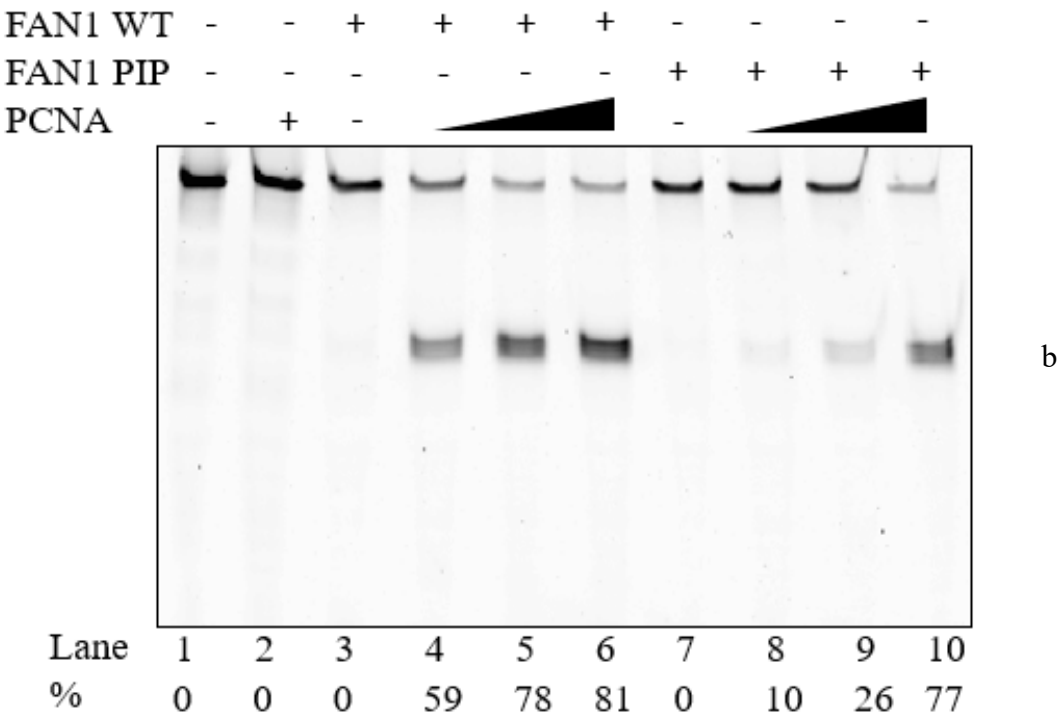


Figure 9. Comparison of the 5' flap endonuclease activities of FAN1 WT and various FAN1 mutants. 28 nM substrate and equal amounts of proteins (1 μM) were subjected to the reaction.

The nuclease reaction was run at 37 °C for 15 minutes. Reactions were resolved on 16 % urea polyacrylamide gels, and fluorescently labelled DNA was detected by Typhoon Trio Imager.



b

Figure 10. The nuclease activities of FAN1 and its PIP mutant were tested with increasing concentrations of PCNA. The nuclease buffer used in the experiments contained NaCl under 150 mM, and the reaction was incubated at 37 ºC for 15 min. Reactions were resolved on 16 % urea polyacrylamide gels and fluorescently labelled DNA was detected by Typhoon Trio Imager.

4.4 Both PIP-box and UBZ-domain of FAN1 is regulated by Ub-PCNA

To explore the UBZ domain function of FAN1, I have applied the same method for comparing the nuclease activity of FAN1 in the presence of PCNA and Ub-PCNA. When we increased the concentration of the purified monoubiquitinated PCNA, FAN1 nuclease activity was significantly more than in the case of unmodified PCNA (Figure 11, compare lanes 4-7 to 8-11). To confirm my finding, I have applied FAN1 PIP/UBZ in the nuclease assay. Mutations in the

UBZ domain of FAN1 impaired the robust Ub-PCNA-dependent stimulation of the FAN1 nuclease (Figure 12, lanes 8-10).

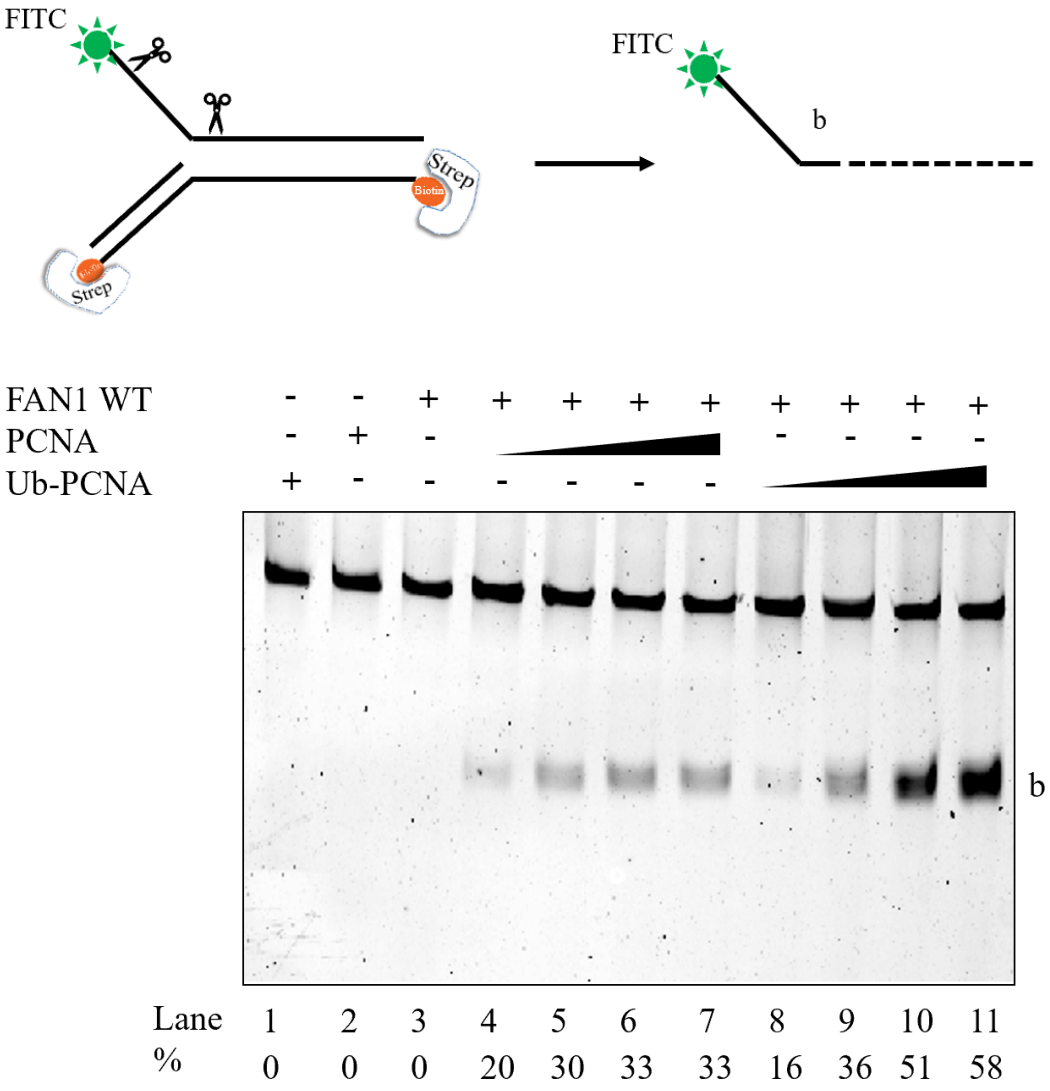


Figure 11. The enhanced nuclease activity of FAN1 was tested with increasing concentrations of PCAN and monoubiquitinated PCNA.

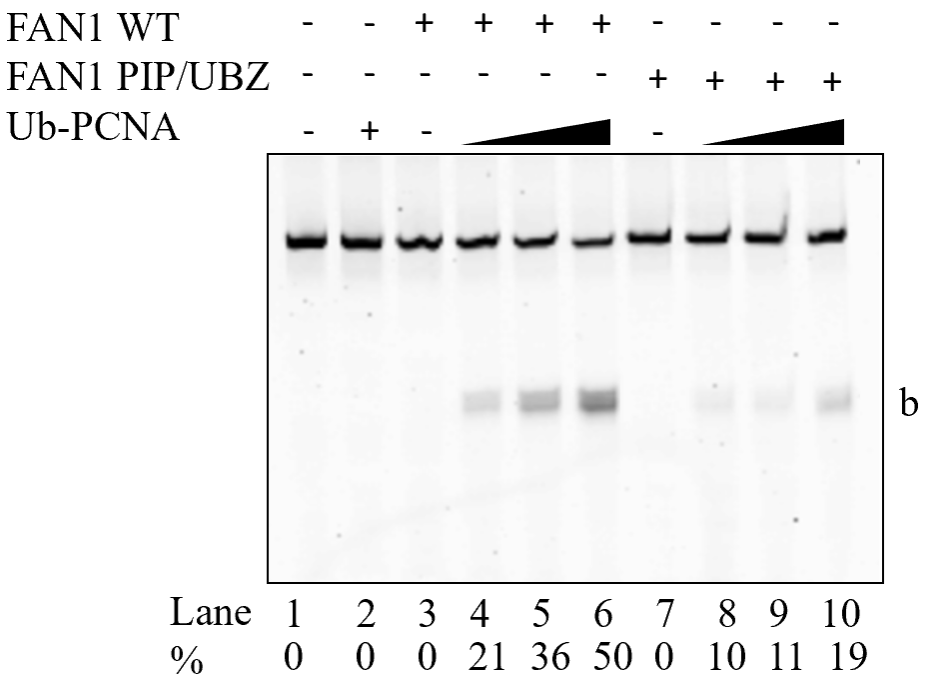


Figure 12. The enhanced nuclease activity of FAN1 and FAN1 PIP/UBZ mutants was tested with increasing concentrations of monoubiquitinated PCNA.

4.5 PCNA can spatially regulate FAN1 function

During the examination of the enhanced nuclease activity of FAN1 in the presence of PCNA, I have observed that the incision site of FAN1 on the substrate was dependent on the arm length of 5' flap substrate. When I applied FAN1 alone in the nuclease assay, I could observe two different sizes of products generated by FAN1 (Figure 9). FAN1 was able to digest 4 nt 3' to the branch point as described by Craig MacKay *et al.,* (2010) with 20 nt and 25 nt arm length. We hypothesized that the incision site of FAN1 might dependent on the length of 5' flap substrate. To confirm this theory, I have applied the same DNA sequence with various length of substrate arm from 20 nt to 40 nt. I have used an equal amount of FAN1 alone in the nuclease assay. I concluded that if the arm length was longer than 25 nt, FAN1 produced a short product close to the 5' end of the flap (Figure 13). Surprisingly, when I examined the nuclease activity of FAN1 in the presence of PCNA, we could observe that the PCNA was able to alter the specificity of FAN1 incision site near to the 5' end of the flap to 4 nt 3' to the branch point (Figure 14). It seems that PCNA is able to manipulate the specificity of FAN1 via its PIP-box interaction.

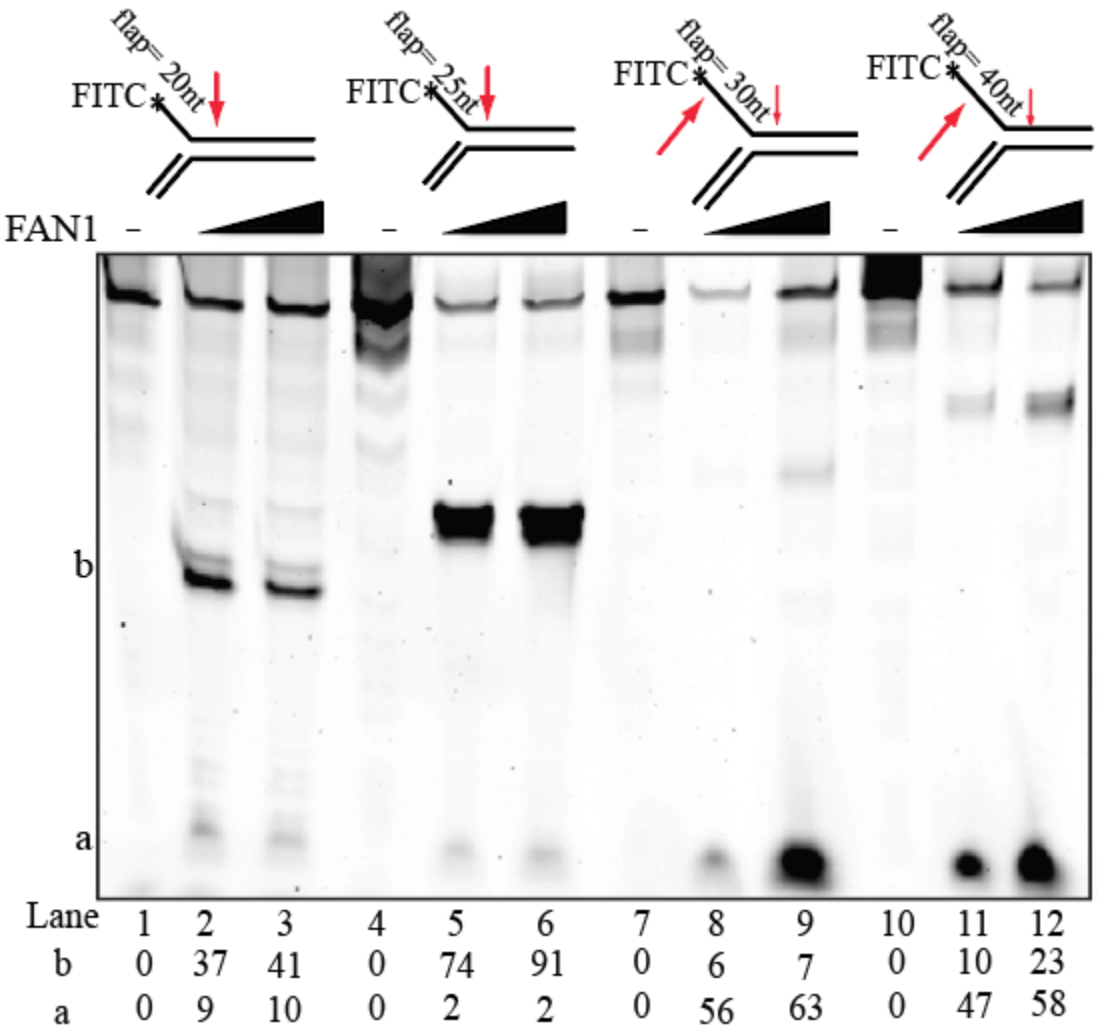


Figure 13. The comparison of FAN1 cutting site on substrates with different length of arm and same sequence. The red arrow indicates the incision site on the substrate. The reaction was carried in for 37 ºC in 15 minutes. Reactions were resolved on 16 % urea polyacrylamide gels, and fluorescently labelled DNA was detected by Typhoon Trio Imager.

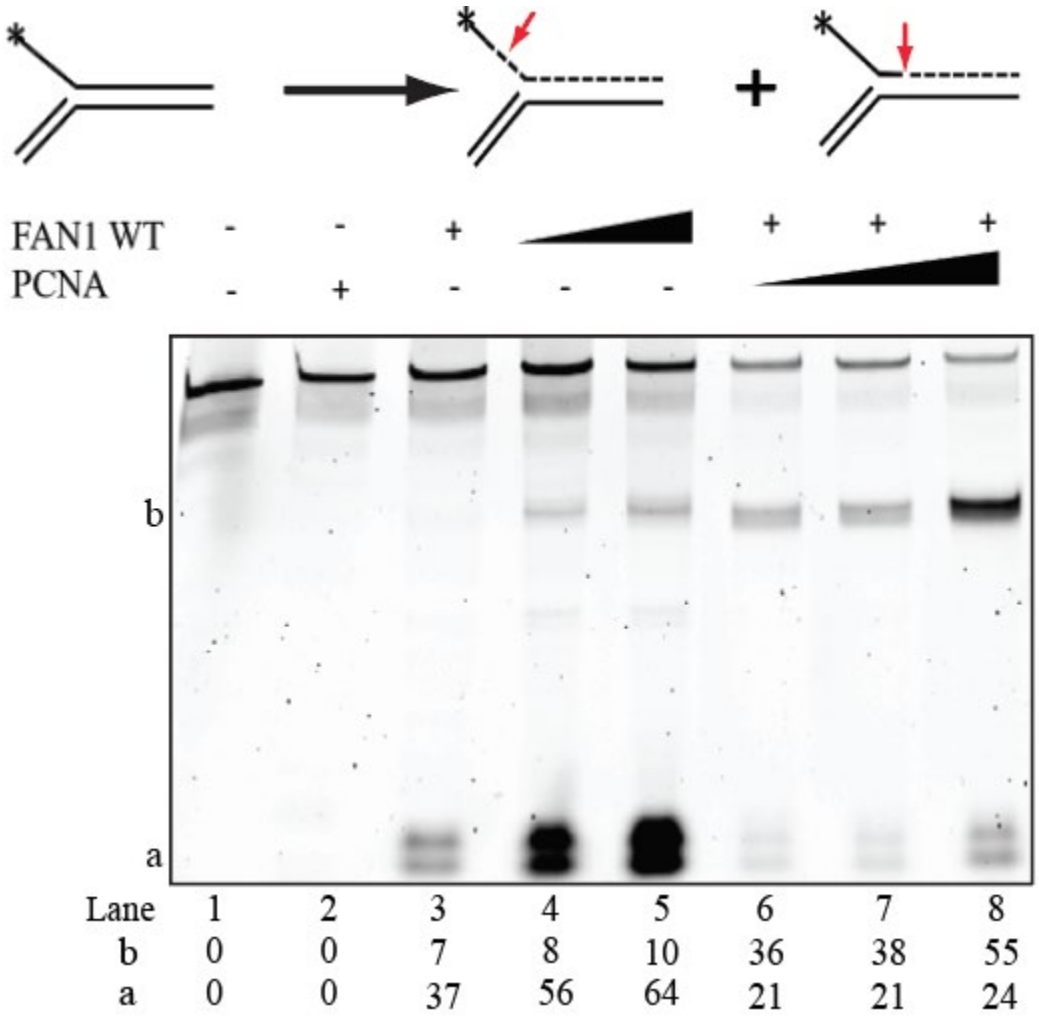


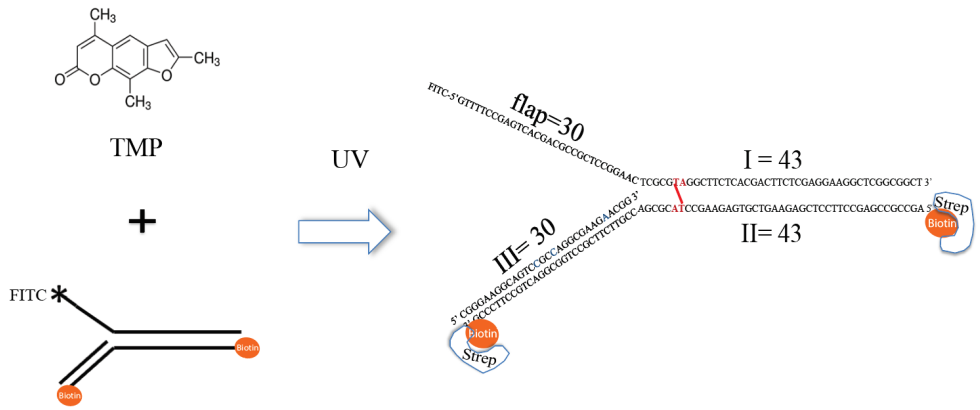
Figure 14. Schematics of substrate and products generated by FAN1 are shown next to the corresponding substrates, and the red arrow indicates the incision site of FAN1. The reaction was carried out at 37 ºC in 15 minutes. Reactions were resolved on 16 % urea polyacrylamide gels, and fluorescently labelled DNA was detected by Typhoon Trio Imager.

4.6 TMP-induced interstrand crosslink substrate

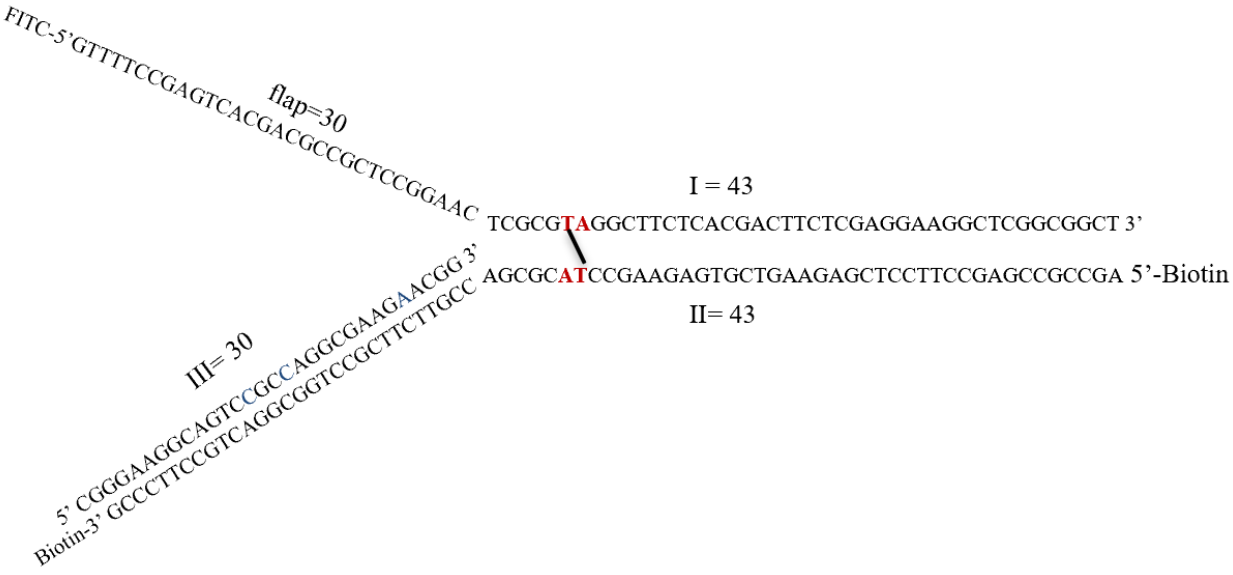
In order to mimic the process of FAN1 bypassing ICL, I have designed the DNA sequence of a substrate, which is suitable for making 4,5′,8-Trimethylpsoralen (TMP) induced ICL substrate. TMP is UV-induced photochemical DNA crosslinker, and it is able to promote 5'-AT dinucleotides as cross-link site (Espositos et al., 1988). Therefore, I have taken advantage of this compound to create of ICL-containing substrates for *in vitro* reconstitution assay, having cross-

link site 5 nt away from replication junction (Figure 15A). There are two kinds of ICL-containing substrates with the exactly same DNA sequence, namely, 73 nt 5'-FITC labeled flap substrate (Figure 15B) and 73 nt 5' (leading strand of replication fork) -FAM labeled flap (Figure 15C). The former one is to investigate the gap filling activity of TLS polymerases after FAN1 incision, and the latter substrate is to examine how TLS polymerases does the primer extension on FAN1 unhooked ICL substrates.

A



B



C

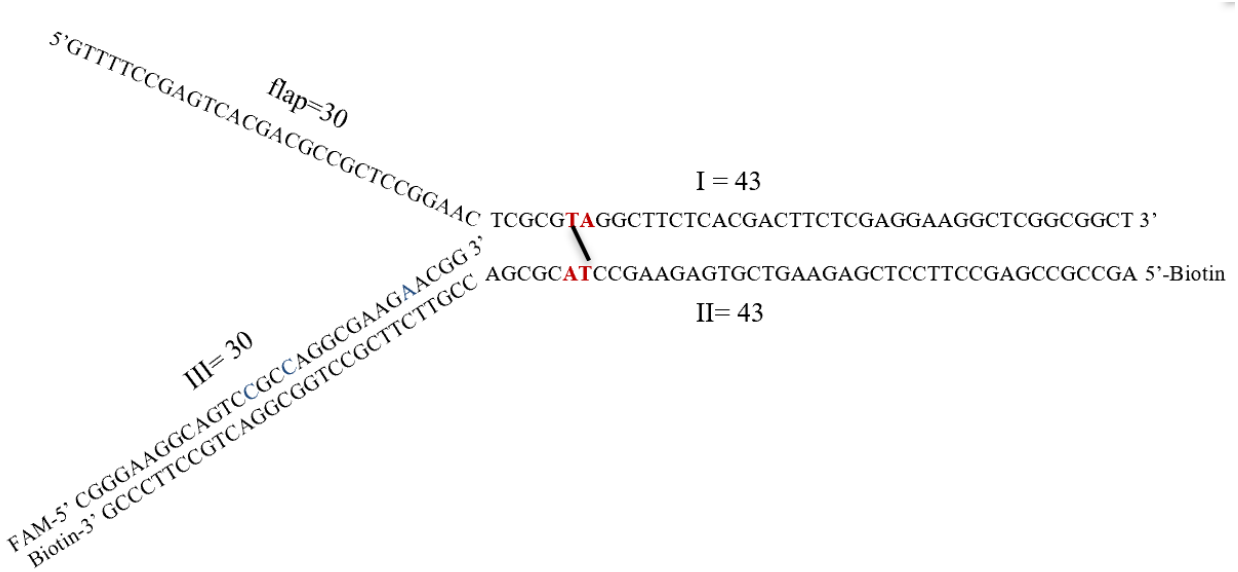


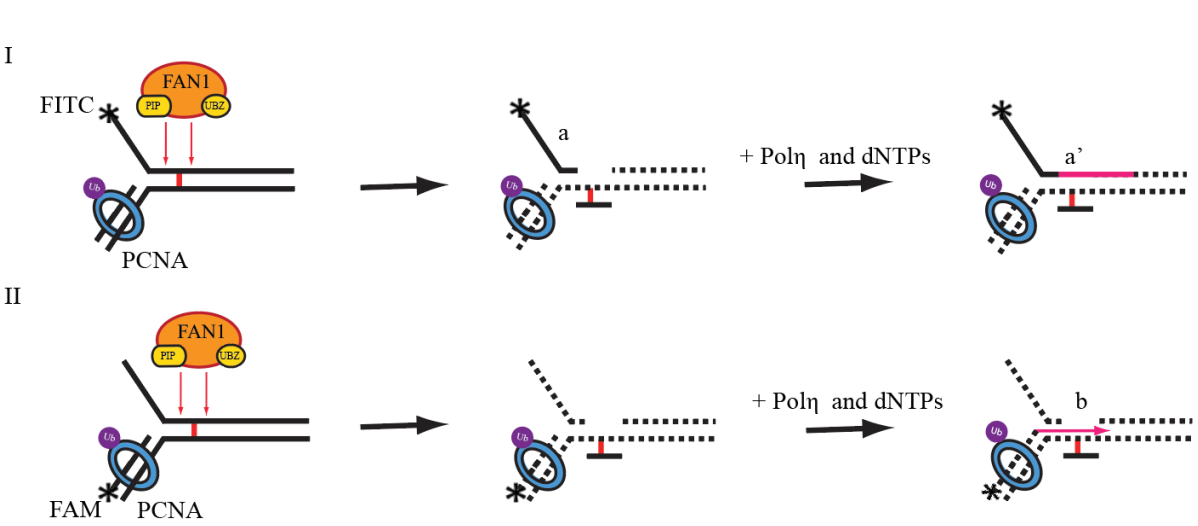
Figure 15. (A) Synthetic representative for 73 nt-long 5'-Fluorescein labeled ICL-containing flap. TMP: 4,5′,8-Trimethylpsoralen, UV: ultraviolet, Strep: streptavidin. See Materials and Methods for detail of making ICL-substrate. (B) 73nt 5'-FITC labeled flap substrate (C) 73 nt 5' (leading strand of replication fork) -FAM labeled flap substrate.

4.7 FAN1 and Pol η can bypass ICL in the presence of Ub-PCNA *in vitro*

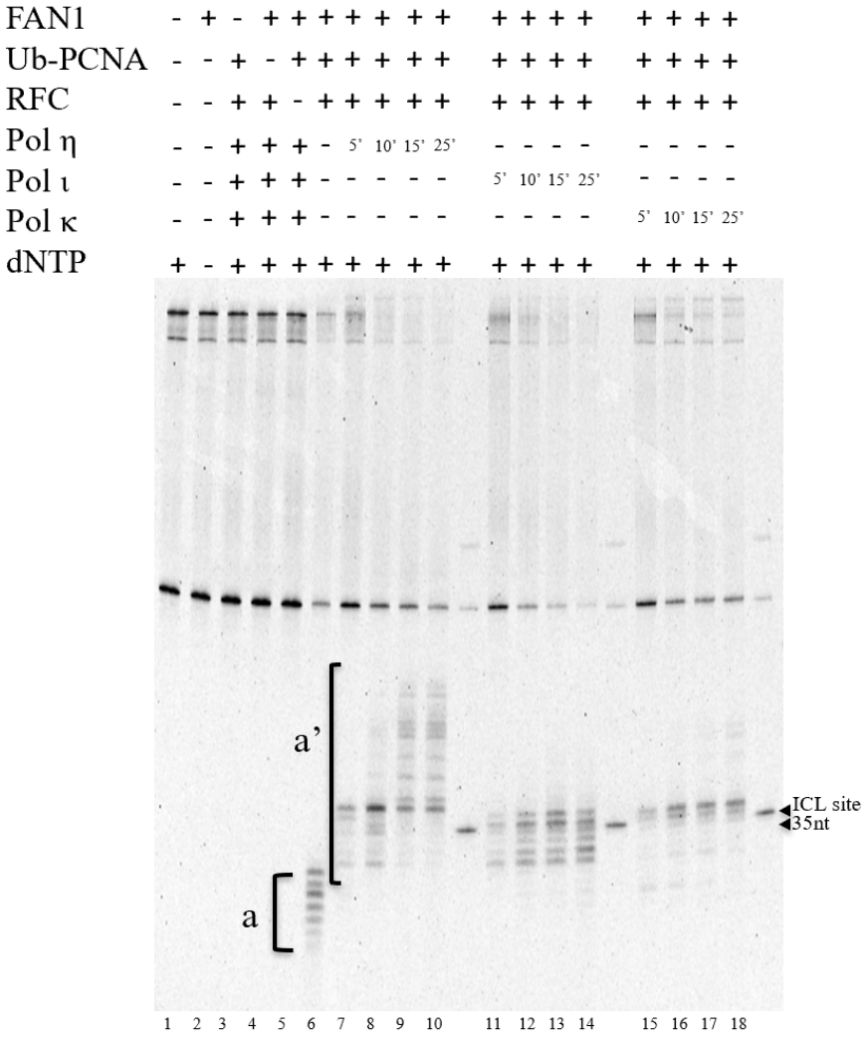
We have demonstrated the Ub-PCNA dependent function of FAN1 in DNA damage repair. In RAD18 DDT pathway, Ub-PCNA is able to recruit Y-family TLS polymerases via its UBZ domain on the DNA damage site, so that DNA lesion can be repaired. For example, Pol η is able to insert nucleotides opposite the UV-damaged cyclobutene pyrimidine dimers (Masutani et al., 1999). In this study, we are going to address how FAN1 facilitates to resolve the stalled replication fork with Y-family TLS polymerases. Since the TMP-induced ICL lesion affects both strands of the DNA, we applied two kinds of 5'flap substrates for monitoring the ICL lesion repair *in vitro* (Figure 16A). We are going to address which Y-family TLS polymerases is able to process FAN1-unhooked substrate. To do that, firstly we have applied 73 nt 5'-FITC labeled flap substrate in this reconstitution assay for examine the "gap filling" activity of TLS polymerases. We have reconstituted the replication fork rescue in the presence of FAN1, Ub-PCNA, and various TLS

polymerases *in vitro*. FAN1 can incise this ICL-containing substrate precisely after junction 4 nt site (which is exactly before the ICL) in the presence of Ub-PCNA (Figure 16B, lane 6). The unhooked substrate containing ICL attached covalently has become a substate for TLS polymerases. We have applied three different Y-family TLS polymerases to see which one can fulfill the gap that was created by FAN1. All of TLS polymerases accommodated FAN1 processed 3' hydroxyl group end substrate. Pol η extended substrate efficiently passing though ICL site (Figure 16B, lane 7-10). Other TLS polymerases (Pol ι and Pol κ), however, were able to insert dNTPs opposite of ICL site (Figure 16B, lane 12-15 and lane 17-20) only. Next, I also have applied 73 nt 5' (leading strand of replication fork) -FAM labeled flap DNA substrate for investigating extension of leading strand by Pol η in the presence of FAN1. We have subjected the reaction mixture with the same experimental condition as "gap filling". However, this experimental setup does not allow us to visualize FAN1 cleavage, its activity was controlled in a parallel reaction with the previous lagging-strand-labelled substrate (Figure 16D). We could observe that Pol η was able to bypass the ICL lesion and extend the labeled primer. In detail, the nuclease activity of FAN1 was stimulated by Ub-PCNA via its UBZ domain and PIP-box resulting in incision precisely 4nt after replication junction (in front of ICL site). The active site of Pol η was then able to accommodate the unhooked ICL adduct and inserted two nucleotides opposite of ICL lesion (Figure 16C, compare lane 9 and lane 10). In this study, we were able to reconstitute the whole process of ICL lesion bypass *in vitro* and found that bypass can occur by the coordinated action of FAN1, Ub-PCNA and Pol η. We conclude that Ub-PCNA dependent FAN1 nuclease can unhook the ICL and generate a bypassable substrate for Pol η indicating that FAN1 and Pol η might work together.

A


I
FAN1
PIP
UBZ
FITC
Ub
PCNA
a
+ Polη and dNTPs
a'
II
FAN1
PIP
UBZ
Ub
FAM
PCNA
+ Polη and dNTPs
b

B


FAN1 - + - + + + + + + + + + + + + + + + +
Ub-PCNA - - + - + + + + + + + + + + + + + +
RFC - - + + - + + + + + + + + + + + + + +
Pol η - - + + + - 5' 10' 15' 25' - - - - - - - -
Pol ι - - + + + - - - - - 5' 10' 15' 25' - - - -
Pol κ - - + + + - - - - - - - - - 5' 10' 15' 25'
dNTP + - + + + + + + + + + + + + + + + + +
a'
a
ICL site
35nt
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

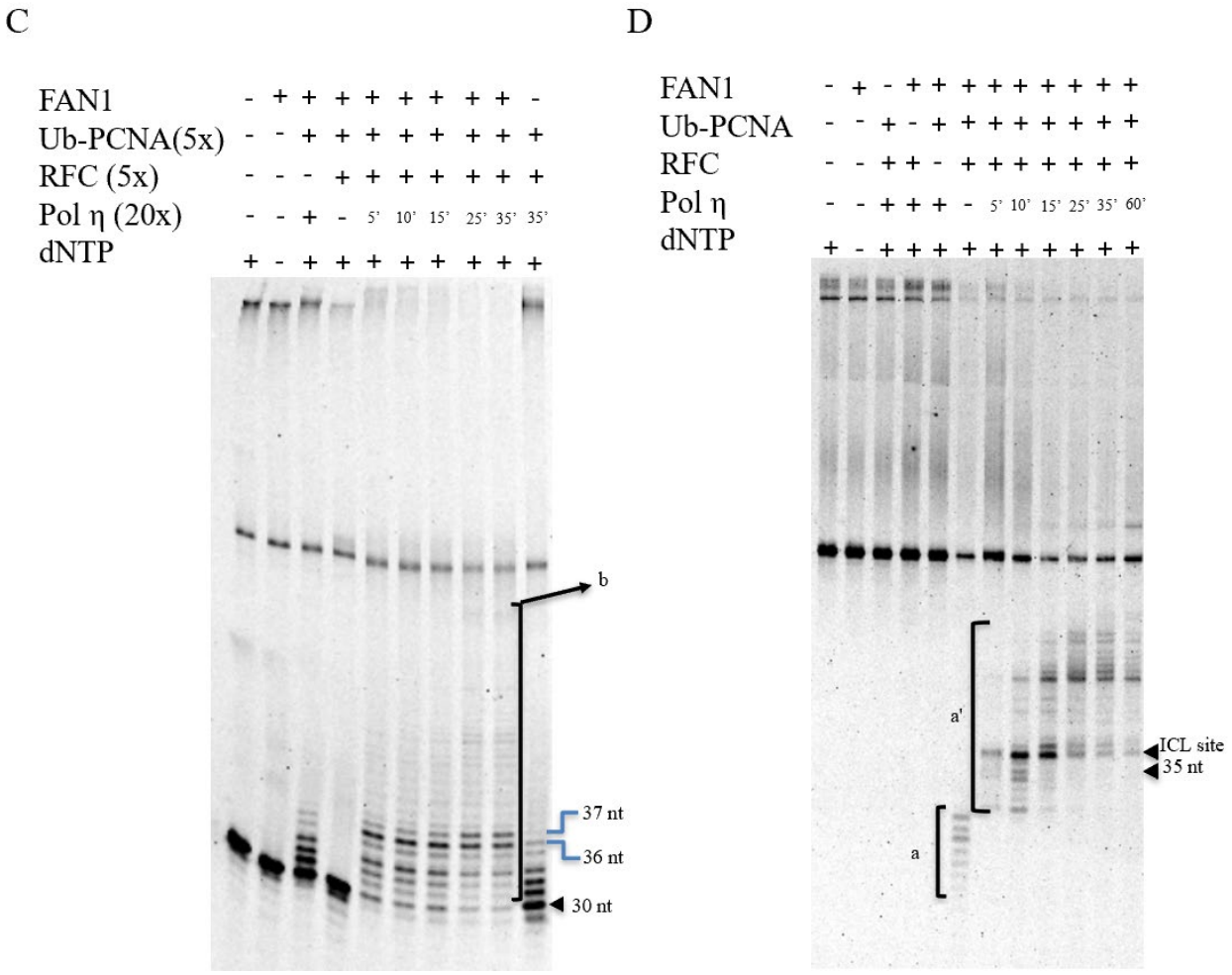


Figure 16. (A) Schematic representation of FAN1-Pol η bypass ICL reconstitution *in vitro*. I. The template of lagging strand was labeled at 5' prime, and FAN1 is able to generate the gap to unhook ICL; a: the products from FAN1 that can be visualize on the gel; a': the extension products from TLS Pol η that can be visualized on the gel. II. The 5' prime of leading strand was labeled, we could observe the primer extension after FAN1 unhooking; b: the extension products from TLS Pol η that can be visualized on the gel. (B) Comparison of the 5' flap ICL substrate bypass of different TLS polymerases *in vitro*. 28 nM of the ICL-containing substrate I was applied. The reaction was carried until various time points, which is indicated in the figure. Reactions were resolved on 10% urea polyacrylamide gels, and fluorescently labelled DNA was detected by Typhoon Trio Imager. (C) 28 nM of the ICL-containing substrate II was applied. The reaction was carried until various time points, which is indicated in the figure. Reactions were resolved on 10 % urea polyacrylamide gels, and fluorescently labelled DNA was detected by Typhoon Trio Imager. (D) 28 nM of the ICL-containing substrate was applied as parallel experiment as indicated in (C).

4.8 FAN1 processes ICL lesions and recruits TLS Pol η on the lesion site

My previous *in vitro* results indicated that FAN1 and Pol η may work together in the rescue of the stalled replication fork. In this experiment, we took advantage of bimolecular fluorescence complementation system (BiFC) to investigate the functional interaction of FAN1 and Pol η. BiFC assay is based on the ability of certain fluorophores to re-fold and emit fluorescent light when two truncated versions of the protein come to close proximity to each other. cDNA sequences of FAN1, its mutant derivatives, as well as Pol η were fused to fragments encoding the N- and C-terminal part (termed V1 and V2) of the YFP variant Venus, respectively. From the beginning of this experiment, we have carried out a systematic survey to detect the best combination of fusional constructs giving the highest fluorescent signal. Although all the eight possible combinations formed detectable fluorescent complexes, the V1-FAN1 and Pol η -V2 pair was selected for the further experiments. Consequently, the mutant variants of FAN1 were fused with the N-terminal V1 fragment only. As a positive control the Mek1-Trib1 complex was used (Guan et al., 2016). For setting up background fluorescence, V1-FAN1 was pared with V2-GUS construct carrying a beta-glucuronidase fusion. Consistent with the previous observations Mek1-V1/Trib1-V2 complexes showed strong, quite homogenous nuclear localization (Guan et al., 2016). Co-transfection efficiency calculated from the ratio of YFP positive/total cell number was between 10-15 % in each set of experiments. In the V1-FAN1/V2-GUS control samples rare cytoplasmic as well as nuclear bulbs were detected, considered as background fluorescence. The overlap of FAN1 and Pol η foci suggested that they might be localized to the same nuclear structures (Figure. 17). To evaluate that the two proteins came to close proximity to each other in response to cisplatin treatment, YFP positive cells possessing at least 5 nuclear foci were counted and compared to cells carrying V1-Fan1/V2-GUS constructs. Under normal conditions, the number of foci-forming cells was 4 times more in the V1-FAN1/V2-Pol η sample as compared to the negative control. However, FAN1 derivatives carrying mutations in the UBZ and/or PIP domains formed just as many foci as the negative control. Upon cisplatin treatment, the number of the positive cells increased to 11x in the V1-FAN1/V2-Pol η sample, while in the others carrying the mutants, the number of foci was not significantly different from the background (Figure 18). These results suggest that FAN1 and Pol η are closely associates upon cisplatin treatment.

The functional interaction between FAN1 and Pol η is further supported by ICL-induced sensitivity assay in HEK293 cell line with stable expressing a silencing construct against FAN1.

To test the sensitivity of shFAN1 cells to cisplatin induced DNA crosslinks, we transiently transfected shFAN1 cell line with an indifferent shDNA (shControl vector) or a vector expressing shPol η shDNA, creating double silenced shFAN1/shPol η cells. We also transiently transfected HEK293 cells with shPol η alone and after 48 h we treated all the silenced cell lines with cisplatin for 48 h and calculated the percentage of surviving cells by a resazurin based cell viability assay (Figure 19). In the absence of either Pol η or FAN1 alone the HEK293 cells showed increased sensitivity to cisplatin compared to the shControl cells, however silencing both Pol η and FAN1 did not sensitize the cells further, reflecting epistatic relationship between FAN1 and Pol η.

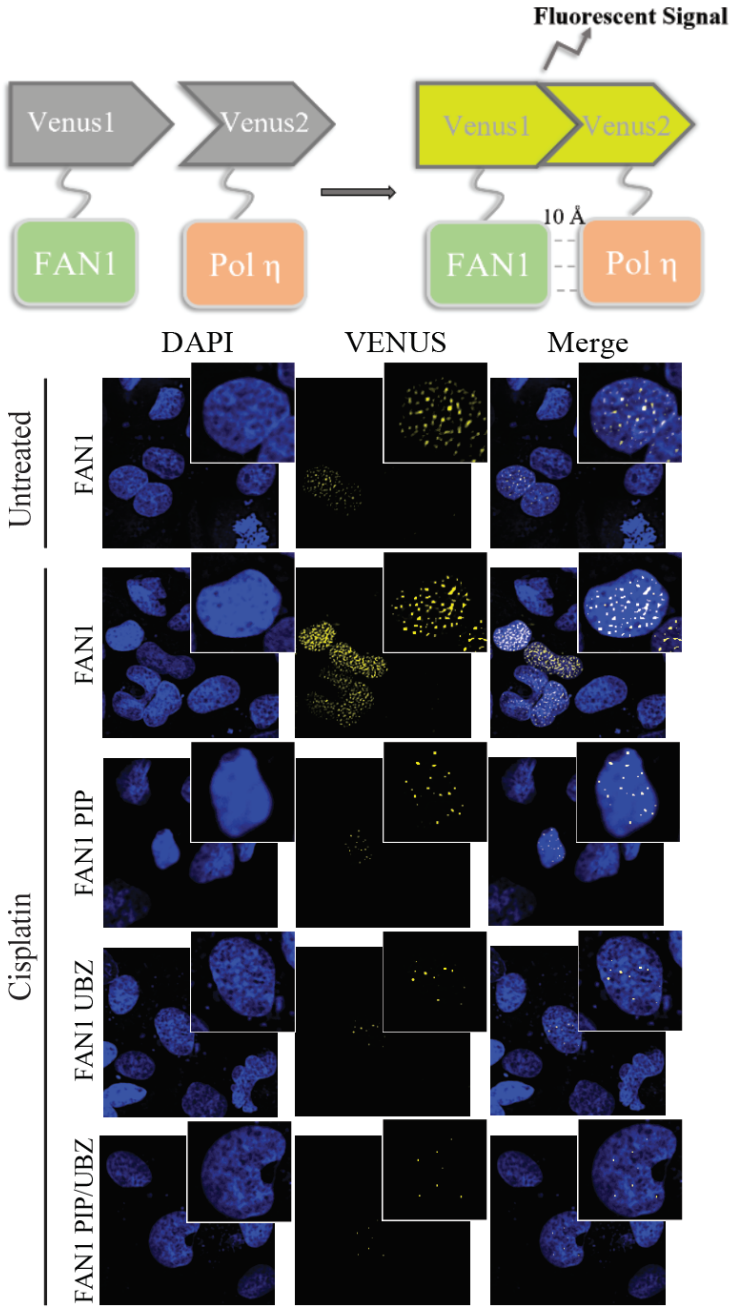


Figure 17. Schematic representation of the Split-Venus method to detect the interaction between FAN1 and Pol η. A fluorescent Venus protein is reassembled from its two complementary non-fluorescent fragments (Venus1 and Venus2) that are fused to the FAN1 and Pol η in appropriate combination (V1-FAN1 and Pol η-V2). The association of the Venus fragments depends on the FAN1-Pol η interaction. Fluorescence images of U2OS cells co-expressing FAN1 WT and Pol η without/with the treatment of a DNA crosslinking agent cisplatin. Followed by treated U2OS

cells co-expressing FAN1 PIP, FAN1 UBZ, FAN1 PIP, UBZ, together with Pol η. For all the experiments 30 µM cisplatin was used with 1 hour of treatment and 6 hours of recovery. Cells were fixed (4 % PFA) and DAPI stained. Fluorescent images were taken with an ImageXpress Micro Confocal microscope using a 40x ELWD objective and analyzed with the MetaXpress software package (Molecular Devices).

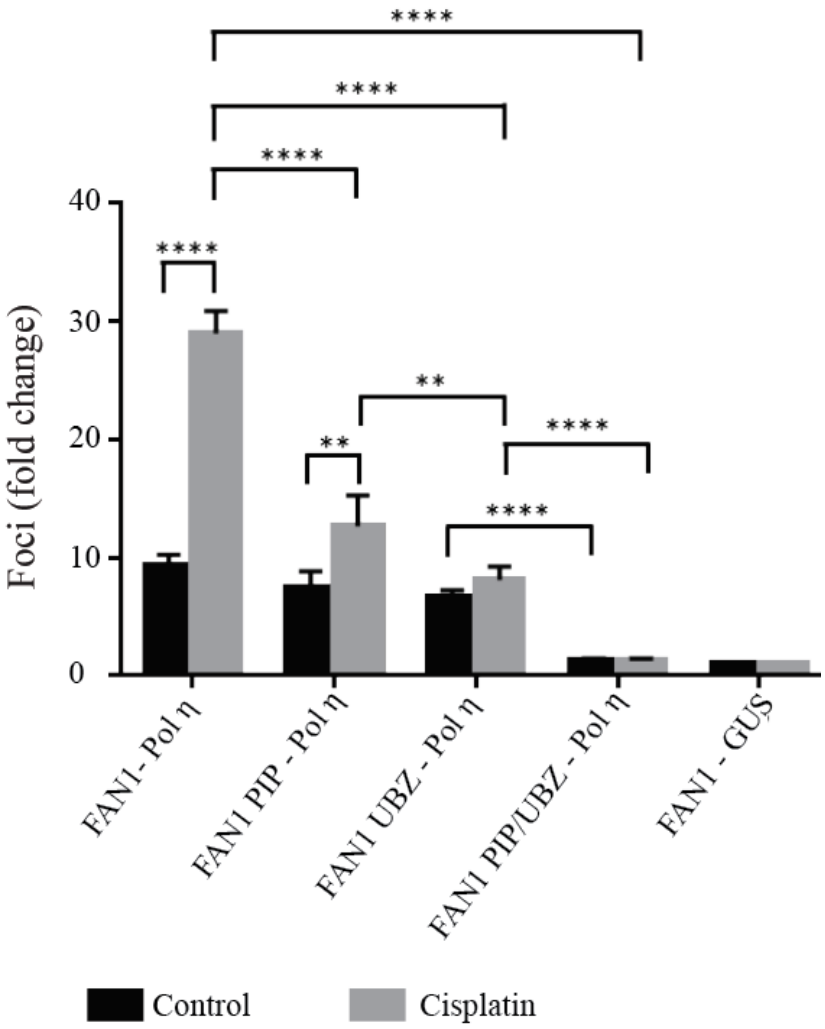


Figure 18. Quantitative analysis of cell numbers possessing more than 5 fluorescent nuclear foci. Ratios (fold changes) of mean cell numbers of samples relative to the controls. Error bars represent standard deviations. Statistical analysis was done by 2way ANOVA with Tukey's multiple comparisons test (N=3). (****) $p<0.0001$; (**) $p<0.01$.

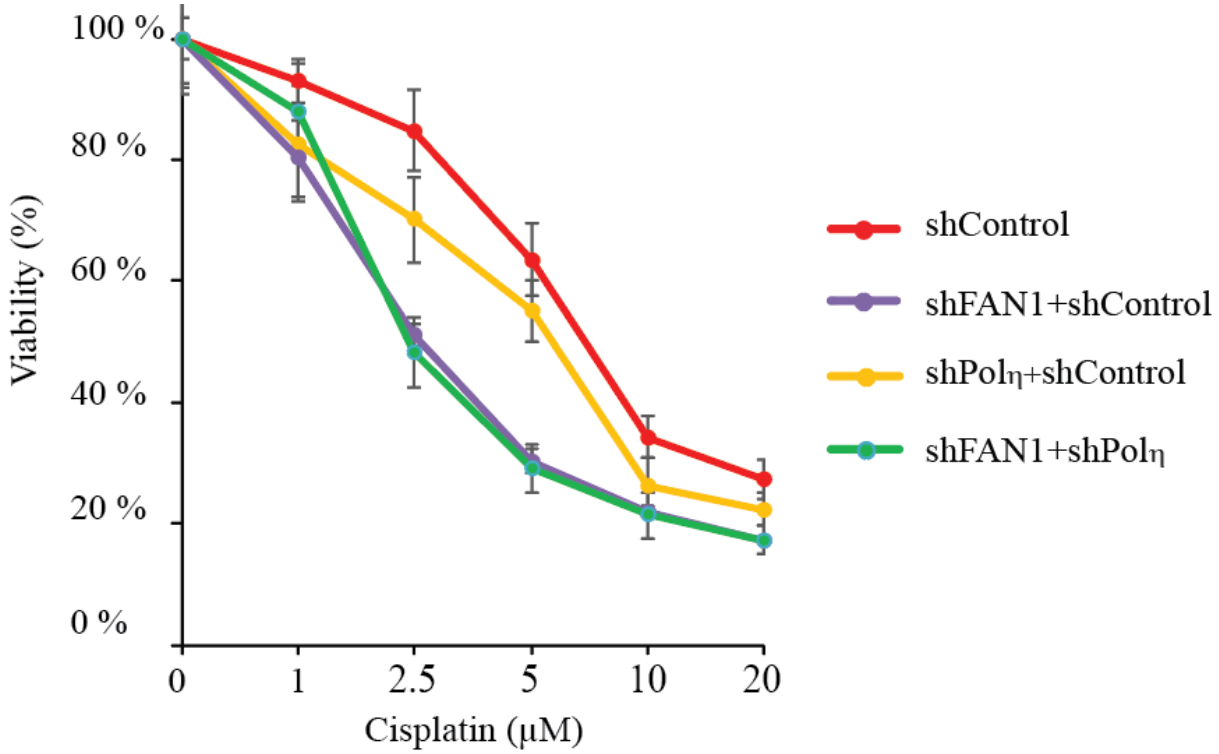


Figure 19. Simultaneous depletion of FAN1 and Pol η do not enhance the sensitivity of HEK293 cells against cisplatin. HEK293 cells were transiently transfected with plasmids encoding indifferent shRNA (shControl) or shRNA against Pol η. HEK293 stably expressing shRNA against FAN1 (shFAN1) was transiently transfected with a plasmid encoding shRNA against Pol η or simultaneously transfected with silencing constructs against FAN1 and Pol η. Two days after transfection cells were treated with different doses of cisplatin, incubated for 48 h and analyzed in resazurin viability assay.

5. Discussion

5.1 FAN1 resolves interstrand crosslink in cooperation with Ub-PCNA

The long history of studying the repair of replication-dependent ICL bypass has revealed that the unhooking step is crucial in resolving obstacles from the stalled replication fork and generating a suitable substrate for TLS polymerases. There are many structure-specific nucleases that play a role in unhooking of ICL. FAN1 was previously recognized as a nuclease in the FA pathway; it is activated and recruited to sites of DNA damage by monoubiquitinated FANCD2 (MacKay et al., 2010). FAN1, as DNA "scissors", can unhook ICL-induced DNA distortion *in vitro* and generate unhooked substrates for TLS polymerases (Pizzolato et al., 2015). Previous reports indicate that FAN1 also possesses a UBZ-domain and PIP-boxes at the N-terminus (Porro et al., 2017). FAN1 mutations cause much less DNA damage-induced genomic instability than other FANC protein mutations; and genetic findings indicated that instead of Fanconi Anemia FAN1 deficiency leads to karyomegalic interstitial nephritis (Yoshikiyo et al., 2010; Zhou et al., 2012). These features of FAN1 lay the groundwork for its role in replication fork rescue and allow us to explore the possible novel roles of FAN1 in the RAD18 DDT pathway.

The replication clamp PCNA homotrimer is a key platform of DNA replication processes. Upon stalling of the replication fork, monoubiquitination of PCNA on K164 by RAD6/18 is the key connecting link between upstream DNA damage sensor proteins and downstream DNA repair players in the RAD18 DDT pathway. The downstream recruitments are based on protein-protein interactions via UBZ-domains or/and PIP-boxes of PCNA-interacting proteins. In this study, I explore some new functions of FAN1 in the rescue of the stalled replication fork. First, I show the possible interaction between RAD18 and FAN1, and that FAN1 and Ub-PCNA can form DNA damage-induced nuclear foci *in vivo.* Second, based on cellular results, I investigate the manner of protein interaction between FAN1 and PCNA/Ub-PCNA. Protein-protein interactions (PPIs) are achieved via stable complex or transient interactions. By pull-down assays, we can observe the physical interaction between FAN1 and PCNA/Ub-PCNA via the PIP-boxes and the UBZ-domain, which is abolished between various FAN1 mutants and PCNA/Ub-PCNA. Our results from pull-down experiments are in agreement with the previous observations of Porro et al., (2017). I also show that the endonuclease activity of FAN1 is enhanced by Ub-PCNA via its PIP-box and UBZ-domain. Various FAN1 mutants (PIP-box and/or UBZ-domain) enabled the identification of the

functional domain responsible for the enhancement of nuclease activity. During my nuclease assay, I noticed that FAN1 is able to cleave at two sites on the branched DNA substrate (one near the 5'-FITC label and the other close to the junction point of the replication fork), and that FAN1's preferable incision site differs according to the length of the replication arm (Figure 13). Previous reports indicate the induced dimerization of FAN1 upon binding branched DNA. The mutation of the core α_9 helix region of the SAP domain abolishes dimerization and impacts the endonuclease activity of FAN1, but this mutant FAN1 still possesses exonuclease activity like the wild-type FAN1 (Zhao et al., 2014). These results illustrate that the structure-specific endonuclease activity of FAN1 is dependent on dimerization upon interaction with DNA. My experimental results reveal that when the 30-nt replication arm 5'flap substrate is applied in the presence of PCNA, FAN1 is able to incise after the junction preferably, exhibiting the endonuclease activity of FAN1. Therefore, we can conclude that PCNA is able to alter FAN1 cleavage specificity, and one of the possible explanations could be that PCNA may stimulate FAN1 dimerization via its PIP-boxes. Upon DNA damage, monoubiquitinated PCNA co-localizes with FAN1 at the site of damage, which further indicates that FAN1 might be a downstream player in resolving DNA lesions. To prove FAN1's unhooking ability *in vitro*, I employed a 73 nt 5'flap ICL-containing replication fork in the same nuclease assay experimental setup. In these experiments, FAN1 was able to bypass the ICL-containing substrate with Ub-PCNA, which indicates that the catalytic domain of FAN1 can accommodate TMP-induced ICLs. It has been described that FAN1 is able to bypass nitrogen mustard-induced ICLs (Pizzolato et al., 2015). Different experimental ICL antitumor drugs can interact with DNA to form various structures of ICL lesions. For example, ICL formation by nitrogen mustard was located exclusively at guanine residues in 5′-GNC-3′ sequences and caused distortion between the helix (Millard et al., 1990). Psoralen forms ICLs preferably with thymines in 5′-TA-3' sequences in the DNA, and these covalent adducts do not generate DNA distortion (Noll et al., 2006). Cisplatin causes mainly DNA-protein and DNA intrastrand crosslinks, but it is also able to induce DNA interstrand crosslinks at 5'-GC sites. As a result of CG interaction, a cytosine can be displaced from the DNA helix to reduce the distance from opposite the guanine (Coste et al., 1999). Therefore, cisplatin can also induce DNA distortion. FAN1 is able to process TMP and nitrogen mustard-induced ICLs *in vitro*, and there is a possibility for FAN1 to bypass ICLs induced by other anticancer drugs. FAN1 seems to be a powerful tool in resolving DNA lesions, and Ub-PCNA might be the guide for its function. This physical protein-protein interaction

can regulate the nuclease activity of FAN1 in unhooking the ICL substrate *in vitro*. Our data illustrate that FAN1 first incises exactly in front of the ICL site and secondly cuts after the ICL to generate a gap on the 5' to 3' lagging strand of the replication fork in the presence of Ub-PCNA (Figure 16A). The FAN1-unhooked substrate can be further processed by TLS polymerases. Unhooking is an essential step in ICL repair, and there is a cellular functional redundancy of endonucleases; other structure-specific nucleases may replace/incorporate FAN1's activity. Such a suitable candidate could be SNM1A, which also acts as "molecular scissors" and possesses a UBZ-domain and a PIP-box. Some reports indicate that SNMA1 is able to interact with PCNA, and it is involved in the RAD18 DDT pathway (Yang et al., 2010). Like FAN1, SNM1A also prefers 5' flap substrates, but it targets the single-strand region of this type of substrates. The exonuclease activity of SNM1A is significantly more robust than its endonuclease activity, and it stops before the ICL site (Buzon et al., 2018). During replication fork rescue, SNM1A mainly trims the substrate to create a nicked substrate which can be further processed by TLS polymerases.

In sum, the cellular redundancy of FAN1 and SNM1A can guarantee the removal of ICLs from the 5' terminus. The unhooking step also depends on the location of the ICL, whether it is at the ss/ds DNA replication junction or after the junction and also on the type of ICL, because different structure-specific nucleases can accommodate different anticancer drug-induced monoadducts.

5.2 TLS polymerase η can process the unhooked FAN1 products

TLS polymerases are a group of specialized DNA polymerases that can bypass DNA lesions without 3' to 5' exonuclease proofreading activity. This low-fidelity DNA replication may give rise to mutagenesis, but its function is very important in ensuring progressive DNA replication when it faces obstacles. The monoubiquitination of PCNA is a key step in regulating the action of TLS polymerases, and the ubiquitylation reaction is recognized as a molecular switch for DNA repair. It is known that Y-family TLS polymerases can cooperate with monoubiquitinated PCNA via its UBM/UBZ domain and PIP-box to bypass DNA lesions (Bienko M et al., 2005). More and more reports have emerged that these Y-family TLS polymerases are not only able to bypass DNA intrastand crosslinks (such as cyclobutene pyrimidine dimers), but can also process ICL-containing monoadducts (Prakash et al., 2005; Roy and Schärer, 2016)(Yang and Gao, 2018).

TLS polymerases often need to process unhooked substrates from structure-specific nucleases in interstrand crosslink DNA repair. In this study, I examine different Y-family TLS polymerases to determine which are able to process FAN1 unhooked substrates from two aspects, namely, gap filling and TMP-induced monoadduct bypass. Following Ub-PCNA-dependent incision, FAN1 can generate a gap on the template of the lagging strand and a monoadduct on the template of the leading strand (Figure 16A). In order to monitor the activity of TLS polymerases, we used a 73-nt-long 5'-Fluorescein-labeled flap substrate to examine gap filling and a 73-nt-long 5'-FAM leading strand 5' flap to observe TMP-induced monoadduct bypass. I applied Pol η, Pol ι and Pol κ separately to reconstitute FAN1-TLS polymerase bypass of ICLs *in vitro*. The novelty of this experimental setup is that I not only applied FAN1 as "scissors" to digest the ICL-containing substrate, but I also introduced highly purified Pol η, Pol ι, and Pol κ, separately, to examine how they process the products unhooked by FAN1. My results indicate that Pol η is the only TLS polymerase which is able to fill the gap generated by FAN1. In fact, Pol η can take the position at the 3' end of the FAN1 digestion site and extend it over 50 nt, but not fully, because the 5'-3'exonuclease activity of FAN1 digests the template strand of Pol η. In addition, by applying a 73-nt-long 5'-FAM leading strand 5' flap, we observed that Pol η is also able to process monoadducts generated by FAN1. It can insert nucleotides opposite the ICL lesion site. Because of the unique structure of its active site, it can accommodate two covalently linked and closely spaced DNA bases (Nair et al., 2004). Another possible reason could be that Pol η can process 5,6-position-modified thymines between the DNA duplex, so that ICLs induced by TMP (an analog of psoralen, which is able to crosslink 5'-TA-3') can also be processed by Pol η. TLS Pol ι, however, possesses a narrow active pocket (Smith et al., 2012). This may be the reason why it can only insert one nucleotide opposite the ICL site. TLS Pol κ displays the same repair activity as Pol ι under our experimental conditions. DNA polymerase κ was shown to act as an extender in the bypass of minor adducts (Smith et al., 2012). This could explain why Pol κ alone was not able to process substrates unhooked by FAN1, and it may require another TLS polymerase as an inserter for its extender function. Our results are consistent with the previous reports of Leigh A. Smith et al., 2012. They also stated that both Pol ι and Pol κ were able to insert one nucleotide opposite the ICL lesion.

The recruitment of TLS polymerases to the site of damage is coordinated by ubiquitinated PCNA in the RAD18 DDT pathway. From biochemical assay, we show that FAN1 and Pol η are

able to bypass ICLs in the ubiquitinated PCNA-dependent manner. We presume the possible explanation is that FAN1 and Pol η can form a protein complex along with ubiquitinated PCNA at the site of lesion at cellular context. Therefore, we employ bimolecular fluorescence complementation (BiFC) testing the functional interaction between FAN1 and Pol η upon cisplatin treatment. From BiFC assay, our data illustrate that FAN1 is able to recruit TLS Pol η to the cisplatin-induced DNA damage site as the signal of nuclear foci increases by 11-fold (by comparing to without cisplatin treatment). However, the signal of nuclear foci from UBZ and/or PIP domain mutants of FAN1 remain almost the same level as negative control. We can conclude that ubiquitinated PCNA can coordinate the recruitment of FAN1 and Pol η at ICL-induced nuclear foci. We also learned that silencing FAN1 and Pol η together did not increase cisplatin sensitivity compared to the FAN1-silenced sample. These results show an epistatic relationship between FAN1 and Pol η, suggesting that they act in the same pathway. It is not surprising that a structure-specific nuclease can cooperate with a TLS polymerase to bypass DNA lesions. The question remaining here is whether the recruitment of Pol η depends on a physical interaction with FAN1 or may be due to the catalytic processing of the gapped intermediate which generates a Pol η substrate. From my *in vitro* nuclease assay data, Pol η is indeed able to process the products from FAN1 digestion (gap filling activity and insertion of dNTP opposite the ICL lesion). Therefore, the possible answer here is that the recruitment of Pol η via FAN1 to the site of damage is due to the catalytic processing of the gapped intermediate which generates a Pol η substrate. The direct physical interaction between FAN1 and Pol η should be further investigated. In our model, we demonstrate that FAN1 is able to unhook the ICL-containing replication fork with the help of Ub-PCNA, and TLS Pol η is able to process the unhooked monoadduct by filling the gap on the lagging strand and by extending the primer on the leading strand (Figure 20).

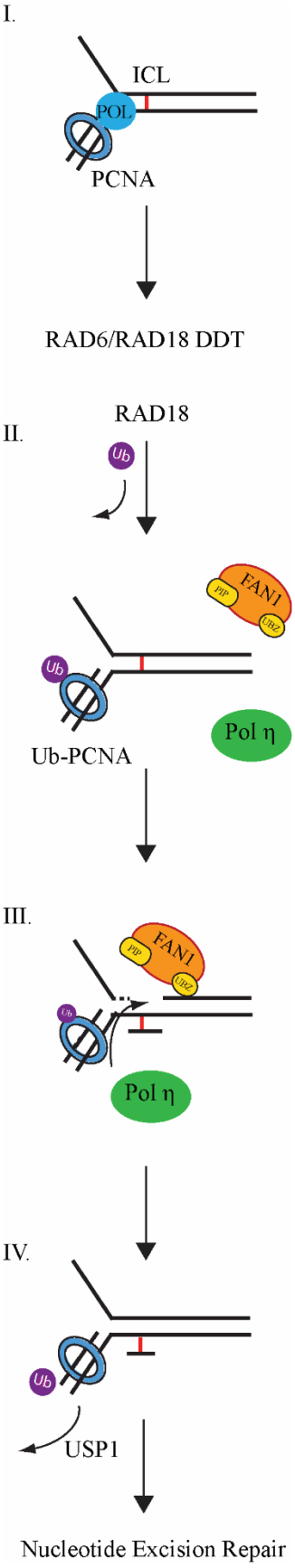

I.
ICL
POL
PCNA
RAD6/RAD18 DDT
II.
RAD18
Ub
FAN1
PIP
UBZ
Ub
Pol η
Ub-PCNA
III.
FAN1
PIP
UBZ
Ub
Pol η
IV.
Ub
USP1
Nucleotide Excision Repair

Figure 20. Model of RAD18-dependent FAN1- and Pol η-mediated bypass of DNA interstrand crosslinks. When DNA replication by canonical replicative DNA polymerases is blocked by an interstrand crosslink (I), the K164 residue of PCNA can be monoubiquitated in the RAD6/RAD18 DNA damage tolerance pathway (II). Monoubiquitated PCNA can further recruit FAN1 to unhook the ICL and Pol η to process the unhooked substrate (III), so that DNA replication can continue unhindered. (IV) Finally, Ubiquitin Specific Peptidase 1 (USP1) removes ubiquitin from the K164 residue of PCNA, and the ICL adduct can be removed through another DNA repair pathway such as nucleotide excision repair (NER).

6. Acknowledgement

I am extremely grateful to my supervisor, Prof. Lajos Haracska for his invaluable advice, continuous support, and patience during my PhD study. His immense knowledge and plentiful experience have guided me in all the time of my academic research. I would also like to thank Dr. Ernő Kiss, Lili Hegedűs, Kata Dudás, Dr. Mónika Krisztina Mórocz, Katalin Illésné Kovács, Katalin Vincze-Kontár and Dr. Ágnes Tóth, Lajos Pintér for their technical support on my study. I appreciate the help from Gabriella Tick for proofreading. I would like to thank all the members in the HCEMM-BRC Mutagenesis and Carcinogenesis Research Group, Delta Bio 2000 company and Dr. Ildikó Unk and Dr. Éva Bálint from DNA Repair Group for their assistance at every stage of the research project. I am grateful that I could spend five years with these smart scientists, and I learned lots of different aspects of knowledge from them.

I would like to offer my special thanks to Dr. Viktor Honti and Dr. László Bodai for reviewing my PhD dissertation. They provide lots of insightful comments and suggestions for me.

Finally, I would like to express my gratitude to my parents (Qing Li and Beiqing Wu), my wife (Orsolya Li). Without their tremendous understanding and encouragement in the past few years, it would be impossible for me to complete my study.

This work was supported by the National Research, Development and Innovation Office (GINOP-2.3.2–15-2016–00024 and GINOP-2.3.2–15-2016–00026) and also received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 739593.

7. References

Al-Hakim, A., Escribano-Diaz, C., Landry, M.C., Donnell, L., Panier, S., Szilard, R.K., and Durocher, D. (2010). The ubiquitous role of ubiquitin in the DNA damage response. DNA Repair 9, 1229–1240.

Alter BP, Young NS (1993): The bone marrow failure syndromes. In: Nathan DG, Oski FA, eds. Hematology of Infancy and Childhood, 4th ed. Philadelphia, PA: Saunders WB, Inc., pp. 216-316.

Aravind L, Koonin EV. SAP - a putative DNA-binding motif involved in chromosomal organization. Trends Biochem Sci. 2000; 25:112–114. [PubMed: 10694879]

Bezalel-Buch, R., Cheun, Y.K., Roy, U., Schärer, O.D., and Burgers, P.M. (2020). Bypass of DNA interstrand crosslinks by a Rev1–DNA polymerase ζ complex. Nucleic Acids Research.

Bienko M, Green CM, Crosetto N, Rudolf F, Zapart G, Coull B, Kannouche P, Wider G, Peter M, Lehmann AR, Hofmann K, Dikic I. Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. Science. 2005 Dec 16;310(5755):1821-4. doi: 10.1126/science.1120615. PMID: 16357261.

Buzon, B., Grainger, R., Huang, S., Rzadki, C., and Junop, M.S. (2018). Structure-specific endonuclease activity of SNM1A enables processing of a DNA interstrand crosslink. Nucleic Acids Research 46, 9057–9066.

Buzon, B., Grainger, R., Huang, S., Rzadki, C., and Junop, M.S. (2018). Structure-specific endonuclease activity of SNM1A enables processing of a DNA interstrand crosslink. Nucleic Acids Research 46, 9057–9066.

Callebaut, I., Moshous, D., Mornon, J.-P., and de Villartay, J.-P. (2002). Metallo-b-lactamase fold within nucleic acids processing enzymes: the b-CASP family.

Castor, D., Nair, N., Déclais, A.C., Lachaud, C., Toth, R., Macartney, T.J., Lilley, D.M.J., Arthur, J.S.C., and Rouse, J. (2013). Cooperative control of holliday junction resolution and DNA Repair by the SLX1 and MUS81-EME1 nucleases. Molecular Cell 52, 221–233.

Cattell, E., Sengerová, B., and McHugh, P.J. (2010). The SNM1/Pso2 family of ICL repair nucleases: From yeast to man. Environmental and Molecular Mutagenesis 51, 635–645.

Chatterjee, N., and Walker, G.C. (2017). Mechanisms of DNA damage, repair, and mutagenesis. Environmental and Molecular Mutagenesis 58, 235–263.

Coste, F., Malinge, J.-M., Serre, L., Shepard, W., Roth, M., Leng, M., and Zelwer, C. (1999). Crystal structure of a double-stranded DNA containing a cisplatin interstrand cross-link at 1.63 Å resolution: hydration at the platinated site. Nucleic Acids Research 27.

Davies, A.A., Huttner, D., Daigaku, Y., Chen, S., and Ulrich, H.D. (2008). Activation of Ubiquitin-Dependent DNA Damage Bypass Is Mediated by Replication Protein A. Molecular Cell 29, 625–636.

Di Fiore PP, Polo S, Hofmann K. When ubiquitin meets ubiquitin receptors: a signalling connection. Nat Rev Mol Cell Biol. 2003 Jun;4(6):491-7. doi: 10.1038/nrm1124. PMID: 12778128.

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.

Espositos, F., Brankamp, R.G., and Sinden, R.R. (1988). THE JOURNAL OF BIOLOGICAL CHEMISTRY DNA Sequence Specificity of 4,5',8-Trimethylpsoralen Cross-linking EFFECT OF NEIGHBORING BASES ON CROSS-LINKING THE 5"TA DINUCLEOTIDE*.

Falck, J., Coates, J., and Jackson, S.P. (2005). Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage.

Farh, K.K.H., Grimson, A., Jan, C., Lewis, B.P., Johnston, W.K., Lim, L.P., Burge, C.B., and Bartel, D.P. (2005). Biochemistry: The widespread impact of mammalian microRNAs on mRNA repression and evolution. Science 310, 1817–1821.

Faridounnia, M., Folkers, G.E., and Boelens, R. (2018). Function and interactions of ERCC1-XPF in DNA damage response. Molecules 23.

Fekairi, S., Scaglione, S., Chahwan, C., Taylor, E.R., Tissier, A., Coulon, S., Dong, M.Q., Ruse, C., Yates, J.R., Russell, P., et al. (2009). Human SLX4 Is a Holliday Junction Resolvase Subunit that Binds Multiple DNA Repair/Recombination Endonucleases. Cell 138, 78–89.

Fricke, W.M., and Brill, S.J. (2003). Slx1 - Slx4 is a second structure-specific endonuclease functionally redundant with Sgs1 - Top3. Genes and Development 17, 1768–1778.

Guan, H., Shuaib, A., Leon, D.D. de, Angyal, A., Salazar, M., Velasco, G., Holcombe, M., Dower, S.K., and Kiss-Toth, E. (2016). Competition between members of the tribbles pseudokinase protein family shapes their interactions with mitogen activated protein kinase pathways. Scientific Reports 6.

Gomes, X. v, Gary, S. L., and Burgers, P. M. J. (2000). Overproduction in Escherichia coli and Characterization of Yeast Replication Factor C Lacking the Ligase Homology Domain

Haracska, L., Johnson, R.E., Unk, I., Phillips, B., Hurwitz, J., Prakash, L., and Prakash, S. (2001). Physical and Functional Interactions of Human DNA Polymerase η with PCNA. Molecular and Cellular Biology 21, 7199–7206.

Haracska, L., Unk, I., Johnson, R.E., Phillips, B.B., Hurwitz, J., Prakash, L., and Prakash, S. (2002). Stimulation of DNA Synthesis Activity of Human DNA Polymerase κ by PCNA. Molecular and Cellular Biology 22, 784–791.

Haracska, L., Johnson, R.E., Unk, I., Phillips, B.B., Hurwitz, J., Prakash, L., and Prakash, S. (2001) Targeting of human DNA polymerase iota to the replication machinery via interaction with PCNA.

Hicks, J.K., Chute, C.L., Paulsen, M.T., Ragland, R.L., Howlett, N.G., Guéranger, Q., Glover, T.W., and Canman, C.E. (2010). Differential Roles for DNA Polymerases Eta, Zeta, and REV1 in Lesion Bypass of Intrastrand versus Interstrand DNA Cross-Links. Molecular and Cellular Biology 30, 1217–1230.

Johnson, R.E., Klassen, R., Prakash, L., and Prakash, S. (2015). A Major Role of DNA Polymerase δ in Replication of Both the Leading and Lagging DNA Strands. Molecular Cell 59, 163–175.

Julie T. Millard, Stanley Raucher, and Paul B. Hopkins Journal of the American Chemical Society 1990 112 (6), 2459-2460 DOI: 10.1021/ja00162a079

Kato, N., Kawasoe, Y., Williams, H., Coates, E., Roy, U., Shi, Y., Beese, L.S., Schärer, O.D., Yan, H., Gottesman, M.E., et al. (2017). Sensing and Processing of DNA Interstrand Crosslinks by the Mismatch Repair Pathway. Cell Reports 21, 1375–1385.

Kerppola, T. K. (2008). "Bimolecular fluorescence complementation: visualization of molecular interactions in living cells." Methods Cell Biol 85: 431-470.

Kim, H., and D'Andrea, A.D. (2012). Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. Genes and Development 26, 1393–1408.

Kim, H., and D'Andrea, A.D. (2012). Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. Genes and Development 26, 1393–1408.

Klein Douwel, D., Boonen, R.A.C.M., Long, D.T., Szypowska, A.A., Räschle, 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.

Komander, D., and Rape, M. (2012). The ubiquitin code. Annual Review of Biochemistry 81, 203–229.

Kratz, K., Schöpf, B., Kaden, S., Sendoel, A., Eberhard, R., Lademann, C., Cannavó, E., Sartori, A.A., Hengartner, M.O., and Jiricny, J. (2010). Deficiency of FANCD2-Associated Nuclease KIAA1018/FAN1 Sensitizes Cells to Interstrand Crosslinking Agents. Cell 142, 77–88.

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. Journal of Biological Chemistry 275, 26632–26636.

Lachaud, C., Slean, M., Marchesi, F., Lock, C., Odell, E., Castor, D., Toth, R., and Rouse, J. (2016). Karyomegalic interstitial nephritis and DNA damage-induced polyploidy in Fan1 nuclease-defective knock-in mice.

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

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 Reports 10, 1947–1956.

Ling, H., Ois Boudsocq, F., Woodgate, R., and Yang, W. (2001). Crystal Structure of a Y-Family DNA Polymerase in Action: A Mechanism for Error-Prone and Lesion-Bypass Replication fidelity polymerase (Woodgate (Goodman and Tippin).

MacKay, C., Déclais, A.C., Lundin, C., Agostinho, A., Deans, A.J., MacArtney, T.J., Hofmann, K., Gartner, A., West, S.C., Helleday, T., et al. (2010). Identification of KIAA1018/FAN1, a DNA Repair Nuclease Recruited to DNA Damage by Monoubiquitinated FANCD2. Cell 142, 65–76.

Masutani, C., Araki, M., Yamada, A., Kusumoto, R., Nogimori, T., Maekawa, T., Iwai, S., and Hanaoka, F. (1999). Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity.

Millard Julie T, Raucher Stanley, and Hopkins Paul B (1990). Millard 1990. Joural of the American Chemical Society.

Motnenko, A., Liang, C.C., Yang, D., Lopez-Martinez, D., Yoshikawa, Y., Zhan, B., Ward, K.E., Tian, J., Haas, W., Spingardi, P., et al. (2018). Identification of UHRF2 as a novel DNA interstrand crosslink sensor protein. PLoS Genetics 14.

Nair, D.T., Johnson, R.E., Prakash, S., Prakash, L., and Aggarwal, A.K. (2004). Replication by human DNA polymerase-i occurs by Hoogsteen base-pairing.

Nijman, S.M.B., Luna-Vargas, M.P.A., Velds, A., Brummelkamp, T.R., Dirac, A.M.G., Sixma, T.K., and Bernards, R. (2005). A genomic and functional inventory of deubiquitinating enzymes. Cell 123, 773–786.

Noll, D.M., Mason, T.M., and Miller, P.S. Formation and Repair of Interstrand Cross-Links in DNA.

Passmore, L.A., and Barford, D. (2004). Getting into position: the catalytic mechanisms of protein ubiquitylation.

Pennell, S., Déclais, A.C., Li, J., Haire, L.F., Berg, W., Saldanha, J., Taylor, I.A., Rouse, J., Lilley, D.M.J., and Smerdon, S.J. (2014). FAN1 activity on asymmetric repair intermediates is mediated by an atypical monomeric virus-type replication-repair nuclease domain. Cell Reports 8, 84–93.

Pickart, C.M. (2001). Mechanisms underlying ubiquitination

Pizzolato, J., Mukherjee, S., Schärer, O.D., and Jiricny, J. (2015). FANCD2-associated nuclease 1, but not exonuclease 1 or flap endonuclease 1, is able to unhook DNA interstrand cross-links in vitro. Journal of Biological Chemistry 290, 22602–22611.

Pizzolato, J., Mukherjee, S., Schärer, O.D., and Jiricny, J. (2015). FANCD2-associated nuclease 1, but not exonuclease 1 or flap endonuclease 1, is able to unhook DNA interstrand cross-links in vitro. Journal of Biological Chemistry 290, 22602–22611.

Plosky, B.S., Vidal, A.E., de Henestrosa, A.R.F., McLenigan, M.P., McDonald, J.P., Mead, S., and Woodgate, R. (2006). Controlling the subcellular localization of DNA polymerases ι and η via interactions with ubiquitin. EMBO Journal 25, 2847–2855.

Porro, A., Berti, M., Pizzolato, J., Bologna, S., Kaden, S., Saxer, A., Ma, Y., Nagasawa, K., Sartori, A.A., and Jiricny, J. (2017). FAN1 interaction with ubiquitylated PCNA alleviates replication stress and preserves genomic integrity independently of BRCA2. Nature Communications 8.

Prakash, S., Johnson, R.E., and Prakash, L. (2005). Eukaryotic translesion synthesis DNA polymerases: Specificity of structure and function. Annual Review of Biochemistry 74, 317–353.

Roy, U., and Schärer, O.D. (2016). Involvement of translesion synthesis DNA polymerases in DNA interstrand crosslink repair. DNA Repair 44, 33–41.

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 ζ. EMBO Journal 25, 1285–1294.

Singh, T.R., Ali, A.M., Paramasivam, M., Pradhan, A., Wahengbam, K., Seidman, M.M., and Meetei, A.R. (2013). ATR-dependent phosphorylation of FANCM at serine 1045 is essential for FANCM functions. Cancer Research 73, 4300–4310.

Smith, L.A., Makarova, A. v., Samson, L., Thiesen, K.E., Dhar, A., and Bessho, T. (2012). Bypass of a psoralen DNA interstrand cross-link by DNA polymerases β, l, and κ in Vitro. Biochemistry 51, 8931–8938.

Smogorzewska, A., Desetty, R., Saito, T.T., Schlabach, M., Lach, F.P., Sowa, M.E., Clark, A.B., Kunkel, T.A., Harper, J.W., Colaiácovo, M.P., et al. (2010). A Genetic Screen Identifies FAN1, a

Fanconi Anemia-Associated Nuclease Necessary for DNA Interstrand Crosslink Repair. Molecular Cell 39, 36–47.

Staresincic, L., Fagbemi, A.F., Enzlin, J.H., Gourdin, A.M., Wijgers, N., Dunand-Sauthier, I., Giglia-Mari, G., Clarkson, S.G., Vermeulen, W., and Schärer, O.D. (2009). Coordination of dual incision and repair synthesis in human nucleotide excision repair. EMBO Journal 28, 1111–1120.

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 Reports 10, 1957–1966.

Unk, I., Hajdú, I., Roly Fá, K., Hurwitz, J., Yoon, J.-H., Prakash, L., Prakash, S., and Haracska, L. (2008). Human HLTF functions as a ubiquitin ligase for proliferating cell nuclear antigen polyubiquitination (PNAS).

Wood, R.D. (2010). Mammalian nucleotide excision repair proteins and interstrand crosslink repair. Environmental and Molecular Mutagenesis 51, 520–526.

Yamanaka, K., Chatterjee, N., Hemann, M.T., and Walker, G.C. (2017). Inhibition of mutagenic translesion synthesis: A possible strategy for improving chemotherapy? PLoS Genetics 13.

Yang, K., Moldovan, G.L., and D'Andrea, A.D. (2010). RAD18-dependent recruitment of SNM1A to DNA repair complexes by a ubiquitin-binding zinc finger. Journal of Biological Chemistry 285, 19085–19091.

Yang, W., and Gao, Y. (2018). Translesion and Repair DNA Polymerases: Diverse Structure and Mechanism. Annual Review of Biochemistry 87, 239–261.

Yoshikiyo, K., Kratz, K., Hirota, K., Nishihara, K., Takata, M., Kurumizaka, H., Horimoto, S., Takeda, S., and Jiricny, J. (2010). KIAA1018/FAN1 nuclease protects cells against genomic instability induced by interstrand cross-linking agents. Proceedings of the National Academy of Sciences of the United States of America 107, 21553–21557.

Zhang, Y., and Hunter, T. (2014). Roles of Chk1 in cell biology and cancer therapy. International Journal of Cancer 134, 1013–1023.

Zhao, Q., Xue, X., Longerich, S., Sung, P., and Xiong, Y. (2014). Structural insights into 5 €2 flap DNA unwinding and incision by the human FAN1 dimer. Nature Communications 5.

Zhou, W., Otto, E.A., Cluckey, A., Airik, R., Hurd, T.W., Chaki, M., Diaz, K., Lach, F.P., Bennett, G.R., Gee, H.Y., et al. (2012). FAN1 mutations cause karyomegalic interstitial nephritis, linking chronic kidney failure to defective DNA damage repair. Nature Genetics 44, 910–915.

8. Summary

DNA damage can be generated by both external agents and internal metabolic processes. Once cells lose their ability to efficiently repair damaged DNA, there are three possible cellular responses. The cell may become apoptotic, malignant, or senescent. For every type of DNA damage, cells have evolved a specific method of repairing the damage or eliminating the damaging compound. For instance, UV-induced pyrimidine dimers can be repaired by nucleotide excision repair. RAD18 DNA damage tolerance (DDT) can resolve stalled replication forks during DNA replication. Upon DNA lesion, PCNA can be monoubiquitinated by RAD18/RAD6, and TransLesion Synthesis (TLS) polymerases can be recruited by the monoubiquitinated PCNA. Chemotherapeutic drugs such as cisplatin or mitomycin C and metabolites from lipid peroxidation can cause interstrand crosslinks (ICLs), covalent links between the opposite strands of the DNA (Stone et al., 2008). ICLs prevent strand separation, physically blocking replication and transcription. Stalled replication forks may collapse, causing DNA double-strand breaks, which can lead to chromosomal rearrangements, carcinogenesis, or cell death. Because ICLs pose such a risk to the human genome, cells have developed multiple pathways to repair this type of DNA lesion. The Fanconi anemia (FA) pathway is one of the main processes responsible for the repair of DNA interstrand crosslinks at the S/G2 cell cycle checkpoint. In the FA pathway, FANCD2/FANCI-associated nuclease 1 (FAN1) can be activated by the monoubiquitylated FANCI/FANCD2 heterodimer via its UBZ domain. FAN1 can digest the ICL-neighbouring region due to its structure-specific endonuclease activity and thus facilitate the bypass of the lesion. A recent report indicates that FAN1 can prevent fork collapse at G-quadruplexes with ubiquitylated PCNA via an uncharacterized PIP domain. Surprisingly, depletion of FAN1 neither affects ICL-induced double-strand DNA break formation nor leads to the development of FA. Rather, germline FAN1 mutations cause karyomegalic interstitial nephritis. Therefore, we asked whether FAN1 is regulated by the RAD18 DDT pathway and investigated the possible interaction between FAN1 and TLS polymerases.

In our current study, we first provide cellular results indicating that the recruitment of FAN1 to the stalled replication fork is regulated by RAD18. Our co-immunoprecipitation experiments also demonstrate that FAN1 and RAD18 can indeed be present in the same cellular multiprotein complex. Second, to further confirm the interplay between FAN1 and Ub-PCNA, we

mapped the protein domains of FAN1 required for the interaction with Ub-PCNA and then mutated these functional domains to use them as negative controls. After expressing FAN1 and its derivatives in yeast cells, we purified the protein of interest by affinity chromatography for *in vitro* functional studies. We subjected both FAN1 and its mutants to a nuclease assay and compared their nuclease activity. All the protein preparations displayed equal enzymatic activity. Subsequently, we examined the nuclease activity of FAN1 and its mutant derivatives in the presence of PCNA/Ub-PCNA. We loaded PCNA/Ub-PCNA by RFC and, to prevent its sliding off from the DNA, we introduced biotin-streptavidin on both sides of the leading template strand. After data analysis by ImageJ software, we concluded that the endonuclease activity of FAN1 can be enhanced by PCNA/Ub-PCNA via the PIP-box and the UBZ domain. To validate the physical interaction between FAN1 and PCNA or Ub-PCNA and to investigate the preferential binding ability of FAN1 to PCNA or Ub-PCNA, we also conducted GST pull down-assays. FAN1 is able to bind PCNA via its PIP-box, and it can interact with Ub-PCNA via its PIP-box and UBZ domain. During our FAN1 nuclease assays, we observed that the sites of the two incisions – one at the end of the 5' flap and the other after a junction 4 nt from the fork – depend on the length of the 5' arm of the substrate. This was revealed when we applied the same DNA sequences with various lengths of the 5' arm ranging from 20 nt to 40 nt. We found that FAN1 can only cleave after the junction on the substrate that has a short arm (< 30 nt) while it can also cut at the end of the 5' flap on the substrate that has a longer arm (≥ 30 nt). However, if we add PCNA to the substrate containing 30-nt arm in the nuclease assay, the cleavage specificity of FAN1 can be modified by PCNA via its PIP-box; PCNA is able to position FAN1 to cut at a specific site (4 nt after the junction) on the 5' flap substrate.

To investigate the role of FAN1 in the rescue of ICL-induced stalled replication forks, we *in vitro* synthesized a DNA molecule mimicking an ICL replication fork substrate with 4,5',8-Trimethylpsoralen (TMP) employing ultraviolet treatment (above 300 nm wavelength). TMP introduced a covalent bond between the 5'-TA-3' and 3'-AT'-5 sequences of the DNA helix at the 5th nucleotide after the junction. We reconstituted the ICL bypass for gap filling on the lagging strand. We added FAN1, Ub-PCNA, and different Y-family TLS polymerases employing the same experimental setup as for the nuclease assay. We discovered that Pol η is the only TLS polymerase that is able to process the substrate containing the ICL unhooked by FAN1. Pol η can fill the gap on the template of the lagging strand, and it is also able to insert dNTPs opposite the ICL lesion.

We also investigated the primer extension performed by Pol η in the presence of ICL blocks and revealed that following Ub-PCNA-dependent unhooking of the ICL by FAN1, Pol η was able to extend the primer through the ICL site. Furthermore, we provide *in vivo* evidence for the epistatic relationship between FAN1 and Pol η upon treatment with ICL-inducing DNA-damaging reagents. In addition, we took advantage of the bimolecular fluorescence complementation system (BiFC) to investigate the functional interaction between FAN1 and Pol η and showed direct binding between the two proteins following cisplatin treatment.

In conclusion, our current results indicate that FAN1 does not only play a role in the FA pathway, but it also participates in the RAD18 DDT pathway, acting in the rescue of stalled replication. Therefore, here we propose a new model for the function of FAN1: upon encountering ICL lesions, the Ub-PCNA-dependent FAN1 nuclease can unhook the covalent bond between the strands of the DNA helix, the unhooked substrate can be further processed by TLS Pol η, and thus stalled replication can be rescued. FAN1 may be a cross-talk protein which regulates two pathways in different cell cycle stages. Our model provides important scientific insight into the repair mechanism of DNA interstrand crosslinks and reveals fine details of the process.

9. Összefoglaló

Mind külső tényezők, mind belső metabolikus folyamatok okozhatnak DNS-károsodást. Ha a sejtek elveszítik azon képességüket, hogy a károsodott DNS-t hatékonyan ki tudják javítani, három lehetséges válaszút áll előttük: apoptotikussá, malignussá vagy szeneszcenssé válhatnak. A DNS-károsodások minden típusához specifikus mechanizmus alakult ki a hiba kijavítására vagy a hibás komponens eltávolítására. Az UV-indukálta pirimidin dimereket például a nukleotidkivágó hibajavítás folyamata javíthatja ki, míg a RAD18 DNS-hibatolerancia (DDT) útvonal a replikáció során elakadt replikációs villát menekíti. DNS-lézió kialakulásakor a RAD6/RAD18 fehérjekomplex monoubikvitinálja a PCNA-t, amely transzléziós polimerázokat (TLS) toboroz a károsodás helyére. Kemoterápiás szerek, mint például a cisplatin vagy a mitomycin C, illetve lipidperoxidációból származó metabolitok keresztkötéseket (ICL), kovalens kötéseket alakíthatnak ki a DNS egymással szemben lévő két szála között (Stone et al., 2008). Az ICL-ek megakadályozzák a szálak szétválását, fizikailag gátolva a replikációt és a transzkripciót. Az elakadt replikációs villa összeomolhat, kettősszálú DNS-töréseket okozva, melyek kromoszómaátrendeződésekhez, karcinogenezishez, illetve sejthalálhoz vezethetnek. Mivel az ICL-ek komoly veszélyt jelentenek a humán genomra, a sejtek többféle útvonalat alakítottak ki az ilyen típusú DNS-hibák kijavítására. A Fanconi Anemia (FA) útvonal képviseli az egyik legjelentősebbet a DNS-keresztkötések javításában, az S/G2 sejtciklus-ellenőrzési pontnál. Az FA útvonalban a FANCD2/FANCI-asszociált nukleáz 1 (FAN1) fehérjét a monoubikvitinált FANCI/FANCD2 heterodimer aktiválja, UBZ doménjén keresztül. A FAN1 szerkezetspecifikus endonukleáz aktivitása révén az ICL melletti régió hasítására képes, elősegítve ezzel a lézión történő áthaladást. Egy friss tanulmány szerint a FAN1 az ubikvitinált PCNA-vel együtt meg tudja akadályozni a replikációs villa G-quadruplexeknél történő összeomlását, egy még nem jellemzett PIP doménen keresztül. Meglepő módon, a FAN1 hiánya nincs hatással sem az ICL-ek által okozott kettősszálú DNS-törések, sem az FA kialakulására, ugyanakkor a csíravonalbeli FAN1 mutációk kariomegáliás intersticiális nefritiszt okoznak. Feltettük tehát a kérdést, hogy vajon a FAN1-et a RAD18 DDT útvonal szabályozza-e, és megvizsgáltuk a FAN1 fehérje és a TLS polimerázok közötti lehetséges interakciókat.

Jelen tanulmányunkban először a sejtbiológiai kísérletek eredményeit mutatjuk be, melyek azt bizonyítják, hogy a FAN1 elakadt replikációs villához történő toborzását a RAD18 szabályozza.

Koimmunoprecipitációs kísérleteink szintén azt igazolják, hogy a FAN1 és a RAD18 fehérje valóban ugyanabban a fehérjekomplexben van jelen a sejtben. Ezt követően, a FAN1 és az Ub-PCNA közötti interakció megerősítéséhez, feltérképeztük a FAN1 azon doménjeit, melyek az Ub-PCNA-vel kialakított interakcióhoz szükségesek, és mutációkat hoztunk létre ezekben a funkcionális doménekben, hogy negatív kontrollként szolgáljanak. A FAN1 és származékainak élesztőben történő expresszáltatását követően, a vizsgált fehérjét affinitás kromatográfiával tisztítottuk *in vitro* funkcióvizsgálatokhoz. Mind a FAN1-et, mind mutánsait nukleáz esszével vizsgáltuk, és összehasonlítottuk nukleáz aktivitásukat. Az összes fehérjepreparátum azonos enzimaktivitást mutatott. Ezt követően megvizsgáltuk a FAN1 és mutáns származékainak nukleáz aktivitását PCNA/Ub-PCNA jelenlétében. A PCNA/Ub-PCNA-t RFC segítségével a DNS-re juttattuk, és hogy megakadályozzuk lecsúszását, biotin-streptavidint vittünk fel a vezető templát szál mindkét oldalára. Az ImageJ szoftverrel végzett képelemzést követően arra a következtetésre jutottunk, hogy a FAN1 endonukleáz-aktivitása PCNA/Ub-PCNA-vel növelhető, a PIP-box és az UBZ domén révén. Ahhoz, hogy a FAN1 fehérje és a PCNA vagy az Ub-PCNA közötti fizikai interakciót igazoljuk és megvizsgáljuk a FAN1 PCNA-hez vagy Ub-PCNA-hez történő preferenciális kötődését, GST pull down kísérleteket is végeztünk. A FAN1 fehérje PIP-boxa révén tud kötődni a PCNA-hez; PIP-boxa és UBZ doménje révén pedig interakcióba tud lépni az Ub-PCNA-vel. A FAN1 nukleáz kísérletek során azt figyeltük meg, hogy a két hasítás helye – az egyik az 5’ túlnyúló végen, a másik pedig a villa elágazását követő 4. nukleotidnál – a szubsztrát 5’ karjának hosszától függ. Amint ez kiderült, amikor olyan DNS szubsztrátokat alkalmaztunk, melyek eltérő, 20-40 nt hosszúságú 5’ karral rendelkeztek. Azt tapasztaltuk, hogy a FAN1 fehérje a rövid karral (<30 nt) rendelkező szubsztráton csak az elágazás után tud hasítani, míg a hosszabb karral (≥30 nt) rendelkező szubsztráton az 5’ túlnyúló végen is. Ha azonban PCNA-t adunk a 30 nukleotid hosszúságú kart tartalmazó szubsztráthoz a nukleáz esszében, a PCNA a PIP-boxon keresztül módosítani képes a FAN1 hasítási specificitását; a FAN1 fehérjét úgy pozícionálja, hogy annak hasítása egy adott helyre (4 nukleotiddal az elágazás után) korlátozódjon.

Annak érdekében, hogy megvizsgáljuk a FAN1 fehérje szerepét az ICL által kiváltott replikációs villa elakadásának menekítésében, *in vitro* szintetizáltunk egy ICL-t tartalmazó replikációs villa szubsztrát DNS molekulát, 4,5’,8-Trimethylpsoralent (TMP) és ultraibolya sugárzást alkalmazva (300 nm hullámhossz felett). A TMP kovalens kötést hozott létre a DNS hélix 5’-TA-3’ és 3’-AT’-5 szekvenciái között az elágazást követő 5. nukleotidnál. Rekonstruáltuk

az ICL-hibajavítás során a lemaradó szálon történő résfeltöltést. A nukleáz esszével megegyező kísérleti elrendezésben FAN1-et, Ub-PCNA-t és különböző, az Y családhoz tartozó TLS polimerázokat alkalmaztunk. Kimutattuk, hogy a Pol η az egyetlen olyan TLS polimeráz, amely képes a FAN1 által kihurkolt ICL-t tartalmazó szubsztrát processzálására. A Pol η képes a lemaradó szál templátján lévő rés feltöltésére és az ICL-lézióval szemben is képes dNTP-ket beépíteni. Megvizsgáltuk a Pol η által, az ICL-blokk jelenlétében végzett primer extenziót is, és kimutattuk, hogy miután a FAN1 Ub-PCNA-függő módon kihurkolta az ICL-t, a Pol η képes volt továbbszintetizálni a primert, az ICL helyén keresztül. Ezen túlmenően, *in vivo* bizonyítékot is adunk a FAN1 és a Pol η közötti episztatikus kapcsolatára, ICL-t indukáló DNS-károsító reagensekkel történő kezelés során. A FAN1 és a Pol η közötti funkcionális interakciót is megvizsgáltuk Bimolekuláris Fluoreszcens Komplementáció (BiFC) segítségével, és közvetlen kötődést mutattunk ki a két fehérje között cisplatin kezelést követően.

Összegzésként, eredményeink azt mutatják, hogy a FAN1 nemcsak az FA, hanem a RAD18 DDT útvonalban is részt vesz, szerepet játszva az elakadt replikációs villa menekítésében. Következésképpen, egy a FAN1 működését leíró, új modellt mutatunk be: ICL-lel történő találkozás esetén, az Ub-PCNA-függő FAN1 nukleáz kihurkolja a DNS-hélix két szála között lévő kovalens kötést, az így létrejött szubsztrátot a TLS Pol η tovább feldolgozza, ezáltal menekítve az elakadt replikációs villát. A FAN1 fehérje lehet az az összekötő fehérje, amely a két különböző útvonalat különböző sejtciklus-stádiumokban szabályozza. Modellünk értékes betekintést nyújt a DNS-keresztkötések javításának mechanizmusába, és rávilágít a folyamat pontos részleteire.

10. List of Publications

MTMT number: 10061209

1. Mandatory peer-reviewed international publications for the fulfillment of the doctoral process

• Biophysical characterization of histone H3.3 K27M point mutation. Szabolcs Hetey, Beata Boros-Olah, Tímea Kuik-Rozsa, **Qiuzhen Li**, Zsolt Karanyi, Zoltan Szabo, Jason Roszik, Nikoletta Szaloki, Gyorgy Vamosi, Katalin Toth, Lorant Szekvolgyi. **Biochem Biophys Res Commun**. 2017 Aug 26;490(3):868-875. doi: 10.1016/j.bbrc.2017.06.133 (IF: 3.575)

• Coordinated Cut and Bypass: Replication of Interstrand Crosslink-Containing DNA. **Qiuzhen Li**, Kata Dudás, Gabriella Tick and Lajos Haracska. (2021) **Front. Cell Dev. Biol**. 9:699966. doi: 10.3389/fcell.2021.699966 (IF: 6.684)

Summarized IF: 10.259

2. Publications is submitted and in peer-reviewing process

- FAN1 and Pol η constitute a pathway for replicational bypass of DNA interstrand crosslinks under the control of RAD18-dependent PCNA-ubiquitylation. **Qiuzhen Li**, Mónika Mórocz, Szilvia Juhász, Lili Hegedűs, Alexandra Gráf, Ádám Sánta, Gaurav Sharma, Péter Burkovics, Ernő Kiss and Lajos Haracska
- Ubiquitylate for tolerance: regulatory mechanisms at the stalled replication fork. Lili Hegedűs, Kata Dudás, **Qiuzhen Li**, Indra Balogh, Gabriella Tick, Lajos Haracska. Peer reviewing in **Cancers** Journal (IF:6.639).

Declaration

I declare that the contribution of Qiuzhen Li was significant in the listed publications and the doctoral process is based on the publications listed. The results reported in the Ph.D. dissertation and the publications have not been used to acquire any PhD degree previously and will not be used in the future either.

Nyilatkozat

Kijelentem, hogy Qiuzhen Li hozzájárulása jelentős volt a felsorolt kiadványokban, és a doktori folyamat a felsorolt kiadványokon alapul. A Ph.D. disszertációban közölt eredményeket és a publikációkat korábban nem használták PhD fokozat megszerzésére, és a jövőben sem fogják.

Szeged, 2021.08.24

Lajos Haracska Ph.D., D.Sc

Kata Dudás