A role of Werner Helicase Interacting Protein 1 in the Fanconi Anemia DNA repair pathway.

Anna Katarzyna Socha

St John's College
University of Oxford

A thesis submitted for the degree of Doctor of Philosophy Trinity Term 2017

Abstract

A role of Werner Helicase Interacting Protein 1 in the Fanconi Anemia DNA repair pathway.

Anna Katarzyna Socha
St John's College, submitted for the degree of Doctor of Philosophy, Trinity Term 2017

This doctoral thesis investigates whether Werner Helicase Interacting Protein 1 (WRNIP1) is involved in the Fanconi Anemia (FA) pathway.

The FA pathway is responsible for resolving DNA interstrand crosslinks (ICLs). While the exact mechanism is unknown, there are 21 confirmed proteins that belong to the pathway, and they can be all divided into three functional groups: the core complex, the FANCD2-FANCI complex and downstream effector proteins.

It was demonstrated here that WRNIP1 depletion results in higher sensitivity to cross-linking agents (Mitomycin C (MMC) and UVA light after trioxalen (TMP) sensitization). The complementation of WRNIP1 – null cell lines with WRNIP1-EGFP resulted in sensitivity comparable with the wild type cell line.

It was also shown *in vivo* that WRNIP1-EGFP is recruited to the chromatin, after TMP-UVA treatment, within first five minutes after irradiation. It was also demonstrated that WRNIP1 is recruited before FANCD2 and that while FANCD2 depletion does not affect WRNIP1 recruitment in any way, WRNIP1 depletion slightly reduces the levels of FANCD2 recruitment, which suggests that WRNIP1 acts upstream from FANCD2.

Since WRNIP1 migrates on the SDS-PAGE gel as two species, the upper form, was analysed and confirmed as ubiquitinated WRNIP1. It was then purified and analysed with mass spectrometry, which allowed the identification of several potential ubiquitination sites. None of the found residues was shown to be the necessary for WRNIP1 upper form to appear and none of them was crucial for WRNIP1 recruitment to the ICL damage site.

To further investigate the roles of each of the domains found within the structure of human WRNIP1, each of the domains was abrogated using one specific point mutation designed to disrupt the function of the whole domain or motif. The analysis of WRNIP1-EGFP point mutants have shown that UBZ domain is responsible for WRNIP1 ubiquitination and while its activity is abrogated, WRNIP1-EGFP is not recruited to the damage site.

Finally, it was confirmed that WRNIP1 interacts *in vitro* with FANCD2, and the analysis of the deletion mutants have shown that none of the deleted sequences was alone necessary for this interaction. The similar analysis performed with WRNIP1 divided into five short fragments, suggests that the sequence between 370 and 510 aa participates in WRNIP1-FANCD2 binding, though it is not conclusive if it is the only part of WRNIP1 that is participating in this interaction.

Acknowledgements

First and foremost, I would like to express my very great appreciation to my supervisor Dr Martin Cohn for his patient guidance and enthusiastic encouragement throughout the course of my doctoral studies. Without his help and constructive suggestions, the completion of this project would be simply impossible.

I would also like to express my deep gratitude to my second supervisor Professor David Sherratt as well as to Professor Nicholas Lakin, who were both the members of my thesis committee. Their advice and assistance have been very much appreciated. Besides my thesis committee members, I would like to thank Professor Petros Ligoxygakis who along with Professor Nicholas Lakin assessed my progress during both my transfer of status viva as well as my confirmation of status viva, providing me with his feedback and the constructive criticism of my data. I am very grateful as well to my college advisor Professor Jason Schnell who was always willing to offer me his time and help.

I would like to offer my thanks to the members of the Micron group as well, especially, Dr Eva Wegel, Dr Ian Dobbie, Dr Ilan Davis and Dr Andrew Jefferson for their assistance with obtaining all the microscopy data presented in this thesis.

I am also very grateful to Dr Fumiko Esashi and Dr Conrad Nieduszynski as well as to all the present and past members of their groups for providing me with advice on my research during our monthly joint meetings.

I would like to thank all the members, past and present, of Dr Cohn's group for their help, support and advice on my experimental work. In particular, I am grateful to Alicja Bulsiewicz, Eric Chih-Chao Liang, Anna Motnenko, David Lopez-Martinez, Dr Yasunaga Yoshikawa, Marian Kupculak and Ganesha Pitchai. Moreover, I would like to thank members of Pears, Lakin and Howarth laboratories for their assistance with my experimental work and necessary laboratory equipment.

Also, I would like to express my gratitude to my parents (and also my sister)- Elżbieta, Cezary and Magda Socha as well as to my grandparents Anna Kot, Marian Kot, Hanna Socha and Oskar Socha.

Finally, I would like to thank all my friends in Oxford and at home. This refers especially to Alex Thorne, Hanna Biernacka, Paweł Lewandowski, Robert Stuglik, Justyna Suwala, Jarek Łukow, Maciek Aleksander, Hossy Ramzy and Moe Eldawi but also to Ollie Vipond, Ambrose Yim, Todd Liebenschutz-Jones, Andreas Gobel, Tim Bourns, Lisa Choi and Rita Nissim.

And in the end special thanks to Omar Kidwai who didn't ditch me when I was unhappy and obnoxious.

Declaration:
I confirm that this DPhil thesis represents my own work. External contributions to the research are acknowledged.

Abbreviations

-/- - throughout this thesis this symbol is used to refer to cell lines deficient in a given protein, which does NOT mean that there are only two alleles of a given gene present in the cell line

53BP1 - p53-binding protein 1

aa – amino acid

ATCC- American Type Culture Collection

ATMIN - ATM Interactor

AP site - apurinic/apyrimidinic site.

BER – base excision repair

BRCA1 - breast cancer type 1 susceptibility protein

CETN2 - centrin 2

CRISPR - clustered regularly interspaced short palindromic repeats

DAPI - 4',6-diamidino-2-phenylindole

DDR- DNA damage response

DMEM – Dulbecco's Modified Eagle's Medium

DNA-PKcs - DNA-dependent protein kinase catalytic subunit

dNTP - deoxyribonucleotide triphosphates

dox - doxycyclin

DSB – double strand break

EDTA - Ethylenediaminetetraacetic acid

FA – Fanconi Anemia

FAN1 - Fanconi-associated nuclease 1

FBS – Foetal Bovine Serum

GG-NER - global genome NER

HeLa - Human cervical cancer cells

HR – homologous recombination

HU - hydroxyurea

ICL – interstrand crosslink

IPTG - Isopropyl β-D-1-thiogalactopyranoside

MDC1 - mediator of DNA damage checkpoint 1

Mgs1 – Maintenance of Genome Stability 1

Milli-Q – ultrapure water

MMC – mitomycin C

MMR – mismatch repair

MRE11- meiotic recombination protein-11

NEM- N-Ethylmaleimide

NBS1- Nijmegen breakage syndrome protein-1

NER – nucleotide excision repair

NHEJ – non-homologous end joining

NLS – nuclear localization signal

PBS – phosphate- buffered saline

PCNA - proliferating cell nuclear antigen

PMSF - phenylmethylsulfonyl fluoride

RAD23B - RAD23 homologue B

RNA Pol II - RNA polymerase II

ROS – reactive oxygen species

RPA - Replication Protein A

Sf9 - *Spodoptera frugiperda*

SDS – sodium dodecyl sulfate

ssDNA – single stranded DNA

TBS-T- Tris- buffered saline with 0.1% Tween20

TCA - trichloroacetic acid (CCl_3COOH)

TC-NER-transcription coupled NER

TLS - translesion synthesis

TMP – trioxalen (4,5′,8-Trimethylpsoralen)

UV – ultraviolet

WB – Western Blot

WRNIP1 – Werner helicase interacting protein 1

XLF - XRCC4-like factor

XRCC4 - X-ray cross-complementing protein 4

Table of Contents

List of figures

List of Tables

1 Introduction

The findings presented in this doctoral thesis provide the preliminary evidence that Werner helicase interacting protein 1 (WRNIP1) plays a role in Fanconi anemia (FA) DNA repair pathway. This work places WRNIP1 and the FA pathway in a broader context of DNA repair mechanisms present in mammalian cells.

1.1 DNA repair mechanisms

Since mammalian chromosomes are constantly exposed to all kinds of damaging agents (both endo- and exogenous) cells protect genomic stability using a number of highly specialized mechanisms that ensure that DNA lesions are efficiently repaired during all stages of cell cycle. The complicated enzymatic machinery responsible for both DNA damage signalling and repair is known as DNA damage response (DDR) (***Jackson and Bartek, 2009***).

DDR includes such important repair mechanisms as homologous recombination (HR), non-homologous end joining (NHEJ), base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), and the FA pathway.

1.1.1 Homologous recombination

HR pathway and NHEJ are the two DNA repair pathways that target double strand breaks (DSBs). The choice between these two mechanisms depends on a number of conditions, such as the stage of cell cycle (HR is active during S and G2 phases of the cell cycle, when the sister chromatids are present (***Branzei et al., 2008***)). On the molecular level 53BP1 (p53-binding protein 1) and BRCA1 (Breast Cancer 1) have been shown to be responsible for this decision. BRCA1 promotes HR while 53BP1 acts in the opposite direction, promoting NHEJ (***Daley et al., 2015***).

HR is a mostly error-free repair process. The nucleotide sequences are simply exchanged between two similar or identical molecules of DNA, so the damaged part of one chromatid can be

replaced with the sequence from the non-damaged one.

When the DSB is introduced, the MRN complex is formed, which in turn activates the ATM (ataxia-telangiectasia mutated) kinase and recruits it to the DSB (***Lee and Paull, 2005***). To form the MRN complex, first meiotic recombination protein-11 (MRE11) binds to DNA together with RAD50 (***Trujillo and Sung, 2001***) in the presence of the third protein - Nijmegen breakage syndrome protein-1 (NBS1) (***Paull and Gellert, 1999; Williams et al., 2009***). MRE11 has a nuclease activity while RAD50 is an ATPase (***Hopfner et al., 2001***).

The activated ATM kinase phosphorylates histone H2AX (at serine 139), which is a specific molecular marker of DNA damage (***Mah et al., 2010***). The phosphorylation of H2AX is spread along the DNA after the mediator of DNA damage checkpoint 1 (MDC1) protein gets recruited to the damage site in response to the H2AX phosphorylation (***Zeman and Cimprich, 2014***). At this point the MRN complex initiates the DNA resection (***Ciccia and Elledge, 2010***), which is the no-return point in the process of deciding between HR and NHEJ because it is impossible to join the resected DNA ends by NHEJ (***Shibata et al., 2014***).

DNA end resection happens in two stages (**Zhu *et al*., 2008**), first the slow initial step (catalysed by the MRN complex together with CtIP protein) and then the long-range resection, which digests DNA away from DSBs (in 5'-3' direction). Long range resection occurs via two routes: the ATP dependant BLM−DNA2−RPA−MRN complex and EXO1−BLM−RPA−MRN. In the first mechanism MRN recruits BLM helicase to the DNA to unwind the double helix and allows DNA2 exonuclease to resect it while Replication Protein A (RPA) binds to single-stranded DNA. In the EXO1−BLM−RPA−MRN case EXO1 is recruited by MRN to cut DNA in 5'-3' direction, while BLM enhances its binding to DNA (***Nimonkar et al., 2011***).

In both cases the result is resection of thousands of nucleotides with the remaining 3' ssDNA overhangs coated in RPA, which is subsequently replaced by RAD51 (a eukaryotic homologue of *E. coli* RecA protein, that binds to the ssDNA and in an ATP-dependent reaction, catalyses strand

exchange) in a process mediated by BRCA1 (***Sung and Klein, 2006***). Once the ssDNA is covered in RAD51 it invades the sister chromatid to use it as a template for error-free repair (***Farmer et al., 2005***).

1.1.2 Non homologous end joining

NHEJ is the alternative DSB repair pathway to HR. It does not require any homologous region to be used as a template and instead allows for the broken ends of DNA to be directly ligated. Because of this it can be functional in all phases of cell cycle but is also error prone (***Daley et al., 2015***).

The NHEJ repair happens in three stages: DSB recognition, processing and ligation (***Yang et al., 2016***). DSBs are recognized by Ku70/80 heterodimer (***Mimori and Hardin, 1986***), which forms a complex with DNA. This complex limits the mobility of DNA ends and keeps them close to each other (***Downs and Jackson, 2004***).

After binding to DNA, Ku70/80 changes its conformation (***Yang et al., 2016***) and then forms a complex with DNA-PKcs (**DNA**-dependent **p**rotein **k**inase, **c**atalytic **s**ubunit), which phosphorylates a number of NHEJ downstream proteins such as the endonuclease ARTEMIS (***Davis et al., 2014***). ARTEMIS, which forms a complex with DNA-PKcs, is a necessary factor for the processing of DSBs before the ligation (***De Ioannes et al., 2012***).

Prior to ligation DNA ends are stabilized by XRCC4 (X-ray cross-complementing protein 4), XLF (XRCC4-like factor) and PAXX (***Yang et al., 2016***) and then after processing and stabilization they are finally ligated by LIG4 (recruited and activated by XRCC4-XLF complex) (***Lu et al., 2007***).

1.1.3 Nucleotide excision repair

NER removes bulky DNA adducts induced by ultraviolet (UV) light, that may otherwise

block DNA polymerases. Defects in NER may cause genetic disorders such as for example Xeroderma pigmentosum and Cockayne syndrome.

There are two main NER lesion recognition mechanisms - global genome NER (GG-NER) and transcription-coupled NER (TC-NER). GG-NER prevents mutagenesis by probing the genome for helix-distorting lesions (the scanning is performed by the damage sensor XPC, in complex with UV excision repair protein RAD23 homologue B (RAD23B) and centrin 2 (CETN2)), while TC-NER is activated when RNA polymerase II (RNA Pol II) is stalled during transcript elongation by a lesion in the template strand (***Marteijn et al., 2014***).

After the lesion is detected XPG endonuclease is recruited and at the same time the TFIIH (Transcription factor II H) helicase unwinds double-helix in the vicinity of the lesion. After the double helix is separated into two strands RPA coats the undamaged one and recruits XPF-ERCC1 heterodimer which, makes a 5' incision to the lesion (***Marteijn et al*., 2014**). On the other side of the lesion the 3' incision is made by XPG. After the lesion is cut out from the sequence PCNA gets recruited onto DNA allowing the recruitment of one of the DNA polymerases, which fills in the gap. In the end the DNA nicks are ligated with either ligase 1 or 3 (***Moser et al., 2007***).

1.1.4 Base excision repair

BER pathway comes into play when it is necessary to remove damaged bases. The removal is conducted by specific glycosylases which leave the apurinic/apyrimidinic (AP) site.

The base removal process always starts with APE1 endonuclease recruitment but the mechanism that follows depends on the number of bases that need to be removed (***Petermann et al., 2006***). If it is only one base, APE1 cuts the phosphodiester bond on the 5' side of the AP site, while Pol β is recruited to fill in the lesion. Afterwards the ligation is performed by Lig3/XRCC1. If 2-10 bases have to be removed, after APE1 recruitment the RFC/PCNA-Polδ/ε complex conducts repair synthesis and nick translation while the FEN1 nuclease removes the flap structure and the

final ligation is carried out by Ligase 1 (***Parsons and Dianov, 2013***).

1.1.5 Mismatch Repair

The strand-directed mismatch repair system detects the DNA helix distortion resulting from the non-complementary base matching.

In *E. coli* the mismatch repair mechanism starts from MutS protein binding to a mismatched base pair (as a homodimer). In the next step the MutS-DNA complex binds the MutL protein and MutS-MutL proteins move along the DNA in two opposite directions while still in complex with each other. This translocation results in a DNA loop.

In bacteria the newly synthesised strand is recognized by its methylation. Upon recognizing a hemimetylated GATC sequence MutH endonuclease introduces a nick, and UvrD helicase separates the strands. The nicked strand is degraded and the resulting gap can be filled by a polymerase.

In eukaryotes the principle of the mismatch repair is similar, but the pathway is more complicated with six homologs of MutS and five homologs of MutL. The newly synthesised strand is recognised by nicks, that are not present in the template strand. In fact, a strand-specific nick or gap is sufficient to direct mismatch repair in extracts of mammalian and Drosophila cells, as well as Xenopus egg extracts (***Modrich 2006***).

1.1.6 Interstrand crosslink repair during and outside of S-phase

ICLs are one of the most cytotoxic DNA lesions and as such require repair mechanisms throughout the cell cycle. Therefore they can be repaired either in the presence of DNA replication or independently. Both pathways share some common important steps such as the recognition of ICL, signalling and recruitment of the proteins that ultimately resolve the lesion. One of the very characteristic steps during ICL repair is so-called unhooking step, performed by the nucleases that

cut the DNA on the same strand both 3' and 5' from the ICL ***(Williams et al., 2013)***.

In replication dependant repair pathways such unhooking step results in a formation of a double strand break intermediate, that will need to be repaired later by HR. In the next step the unhooked lesion is bypassed by translesion polymerases, and it is believed that the remaining covalently linked nucleotides are removed by NER.

One of the most important differences between the replication-dependent and independent repair is that while replication-dependent repair relies on a blocked DNA polymerase to signal ICL damage, replication-independent repair had to evolve some alternative mechanisms to recognize the lesions.

FA proteins, that are discussed in this thesis play an important role in replication-dependent ICL repair and are mostly dispensable for replication-independent ICL repair.

While the replication-dependent ICL repair mechanisms are very important in rapidly dividing cells, replication independent repair is more important in post-mitotic cells such as neurons, or cells that do not divide very often (so for example stem cells).

While the detection of ICLs in the replication-dependent pathways relies on stalled replication forks, it is obviously not possible to employ the same sensing mechanism in replication-independent pathway. One potential alternative mechanism could be simply a collision between a transcribing RNA polymerase and the ICL. It is suggested by the fact that ICLs generated by psoralen or cisplatin are repaired preferentially in actively transcribed genes ***(Larminat et al., 1993; Islas et al., 1991).***

Another possibility is that even in the absence of RNA polymerase, DNA distortion caused by ICL damage can be enough to initiate some ICL repair mechanisms.

1.2 Fanconi anemia

Fanconi anemia (FA) is a rare autosomal, recessive disorder, with incidence estimated to be

between 1/200,000–1/400,000 in the general population (***Dong et al., 2015***). While it is primarily presenting itself as a bone marrow failure disease (which leads to aplastic anemia), it has been also shown that it is associated with a number of other clinical problems such as immune deficiency (***Fagerlie and Bagby, 2006***), endocrine dysfunction (***Giri at al., 2007***), osteoporosis (***Giri et al., 2007***) and finally cancer.

It has been observed that FA cells present many aspects of accelerated ageing process and therefore FA was proposed to be one of the so called segmental progerias – genetic disorders which do not fully recapitulate the normal ageing phenotype but cause a subset of pathologies typical of ageing at much younger age than they occur in patients not affected by the mutation. Other examples of such genetic syndromes are Werner Syndrome, Bloom Syndrome, Cockayne Syndrome, Huntington Guilford Progeria and Xeroderma pigmentosum.

The molecular basis for the accelerated ageing phenotype in FA patients is a mutation in one of the 21 genes associated with DNA cross-link repair, although it remains a subject of discussion if all the problems FA patients experience are directly stemming from the impaired repair of DNA damage or if possibly there are other not fully investigated causes. For example, it has been observed that FA cells can be characterized by a high-oxidative state (***Pagano et al., 2004***; ***Pagano et al., 2012***), which can be observed by the fact that DNA obtained from FA patients cells shows higher levels of a marker of base oxidation (8-oxo-deoxyguanosine), than DNA obtained from control cells. Additionally, the production of hydroxyl radicals from FA leukocytes is increased, in part due to mitochondrial dysfunction which leads to elevated levels of reactive oxygen species (ROS) (***Kumari et. al., 2014***, ***Pagano et al., 2014***).

Increased production of ROS combined with impaired DNA damage repair leads to an increase in unrepaired DNA lesions that result in symptoms of progeria suffered by FA patients.

1.2.1 FA pathway

The pathway responsible for Fanconi anemia consists of 21 genes, discovered so far. These genes are responsible for 95% of all FA cases and new genes are still being discovered. FA proteins cooperate with non-FA proteins, including proteins that belong to other DNA repair pathways, such as homologous recombination (HR), nucleotide excision repair (NER) and translesion synthesis (TLS). Together, they first detect and then repair ICLs.

There are three functional groups of FA proteins: the core complex, the FANCD2/FANCI complex and the downstream effector proteins (***Figure 1***). The core complex is composed of eight FANC proteins (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL and FANCM) and three associated proteins (FAAP20, FAAP24 and FAAP100). FANCL has the ubiquitin ligase activity and performs the ubiquitination of the FANCD2/FANCI complex (***Longerich et al., 2009***).

While the exact early stage molecular events that trigger FA pathway, after ICL is formed, are not clear, there is evidence that the UHRF1 protein acts as one of the early sensors and plays a role in recruiting the FANCD2/FANCI complex to the damage site (***Liang et al., 2015***). It was shown to interact with ICLs both *in vivo* and *in vitro*, and also to be quickly recruited to the damage site within seconds to be followed by FANCD2 that appears with a delay of approximately 10 minutes (***Liang et al., 2015***; ***Tian et al., 2015***). UHRF1 recognizes ICLs through its SET and RING finger associated (SRA) domain.

The FANCD2/FANCI complex is the central element of the entire FA pathway. After its recruitment to the ICL damage site it is monoubiquitinated by the core complex (FANCD2 at lysine 561 and FANCI at lysine 523) **(*Liang et al., 2016*)**. Without this recruitment the ICL will not be repaired. The FANCD2/FANCI was also shown to be multiphosphorylated by the ATR and ATM kinases (***Matsuoka et al., 2007***). While the exact function of the FANCD2/FANCI complex is not fully known yet, the most likely possibility is that it is required to regulate the order and timing of recruitment of the downstream effector proteins to the ICL damage site.

Downstream effector proteins such as: FANCD1 (BRCA2), FANCJ (BRIP1 - (BRCA1-interacting protein 1)), FANCN (PALB2 - Partner and localizer of BRCA2), FANCO (RAD51C), FANCP (SLX4), FANCQ (XPF), FANCR (RAD51) and FANCS (BRCA1) are the proteins that contribute to the repair of the ICL damage at later stages of the process. Their main goal is to unhook the ICL during the S-phase of the cell cycle (***Knipsheer et al., 2009***). Some of them are associated with homologous recombination (for example BRCA1, BRCA2, BRIP1, PALB2, RAD51 and RAD51C). Mutations in many genes encoding downstream effector proteins have been shown to significantly increase the risk of cancers. The most well-known ones are BRCA1 and BRCA2 that increase the risk of both breast and ovarian cancer (***King et al., 2003***). Also, less known ones such as BRIP1, PALB2 and RAD51C have been shown to increase the risk of breast cancer (***Meindl at al., 2010***, ***Turnbull et al., 2008***)

So far, it has been shown that FANCJ interacts with BLM (the protein product of the gene whose mutation leads to the Bloom Syndrome), promoting its stability (***Suhasini et al., 2011***; ***Suhasini et al., 2012***). There are several nucleases among the downstream effector proteins, such as XPF/ERCC1, MUS81/EME1 and SLX1 (XPF is possibly especially important in the ICL unhooking step). The nucleases interact with SLX4, which acts as a scaffold for their respective activities. Also, FAN1 (Fanconi-associated nuclease 1) (***Kratz et al., 2010***; ***Smogorzewska et al., 2010***; ***MacKay et al., 2010***; ***Liu et al., 2010***; ***Yoshikiyo et al., 2010***) and SNM1A nucleases have been implicated in the ICL repair (***Ishiai et al., 2004***, ***Wang et al., 2011***). While it is not completely clear at this point which nuclease is responsible for which incision, the outcome of their recruitment to the ICL damage site are incisions at both – 5' and 3' sides of the lesion, a process that is known as the unhooking.

After the ICL is unhooked, matching nucleotide is inserted across the lesion (insertion by Polζ) and then TLS DNA polymerases REV1 and Polζ are recruited (in a core complex-dependent manner) to extend the leading strand beyond the damage site (extension) (***Budzowska et al., 2015***).

At this point monoubiquitinated FANCD2 recruits CtIP, which together with MRN starts DNA resection on the other strand (***Unno et al., 2014***). HR pathway follows, using the sister chromatid for the error-free repair of the incision generated DSB.

The final step of the described ICL repair process includes the removal of the now unhooked lesion by the NER pathway (***Mouw and D'Andrea, 2014***) (***Figure 2***).

When the ICL damage is repaired and ubiquitinated FANCD2/FANCI complex is no longer needed it is deubiquitinated by the USP1-UAF1 complex. It is crucial for genomic stability as unlimited FANCD2/FANCI ubiquitination results in ineffective ICL resolution (Cohn et al., 2009).

The ICL repair during G1 and G0 phases of the cell cycle (replication independent ICL repair) is critical to allow transcription in non-dividing or slowly dividing cells but the full mechanism for the ICL recognition of and repair in the absence of DNA replication remains to be elucidated. It is possible that the molecular machinery required for the repair is less complex than in case of replication dependent case (**Figure 3**).

Finally, it is important to mention that FA proteins seem to have some alternative roles apart from repairing ICLs. For example there is growing evidence to support the role of FA proteins in replication fork protection and recovery after stalling, even if caused by other genomic stresses and not necessarily ICLs (***Chaudhury et al., 2014***; ***Lachaud et al., 2016***).

FANCD2 and FANCI are also associated with the ultra-fine DNA bridges that interlink chromosomes during mitosis causing genome instability (***Chan at al., 2009***) and other FA proteins (mainly FANCM) have been associated with the processing of transcription associated DNA - RNA hybrids, that are also known as R-loops, and cause stalling of the replication forks (***Schwab et al., 2015***).

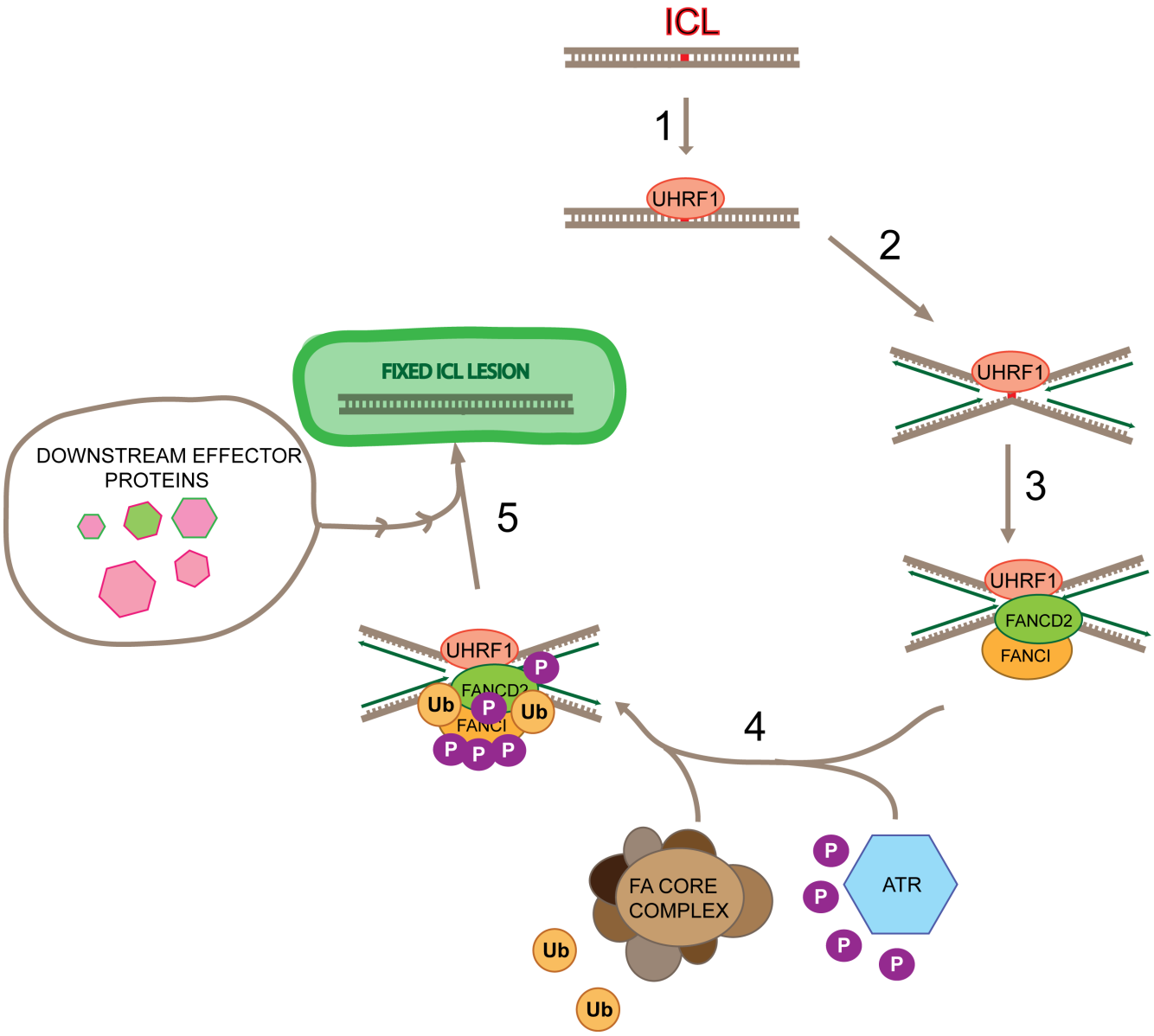


Figure 1: Schematic representation of the FA pathway. After ICL is formed UHRF1 is recruited to the damage site (1), the double helix is unwinded on both sides of the lesion (2) and then UHRF1 recruits FANCD2-FANCI complex to the chromatin (3). FANCD2-FANCI complex undergoes both phosphorylation by the ATR kinase as well as ubiquitination by the Core Complex (4). At this point the Downstream Effector Proteins are recruited to the lesion to repair the damage.

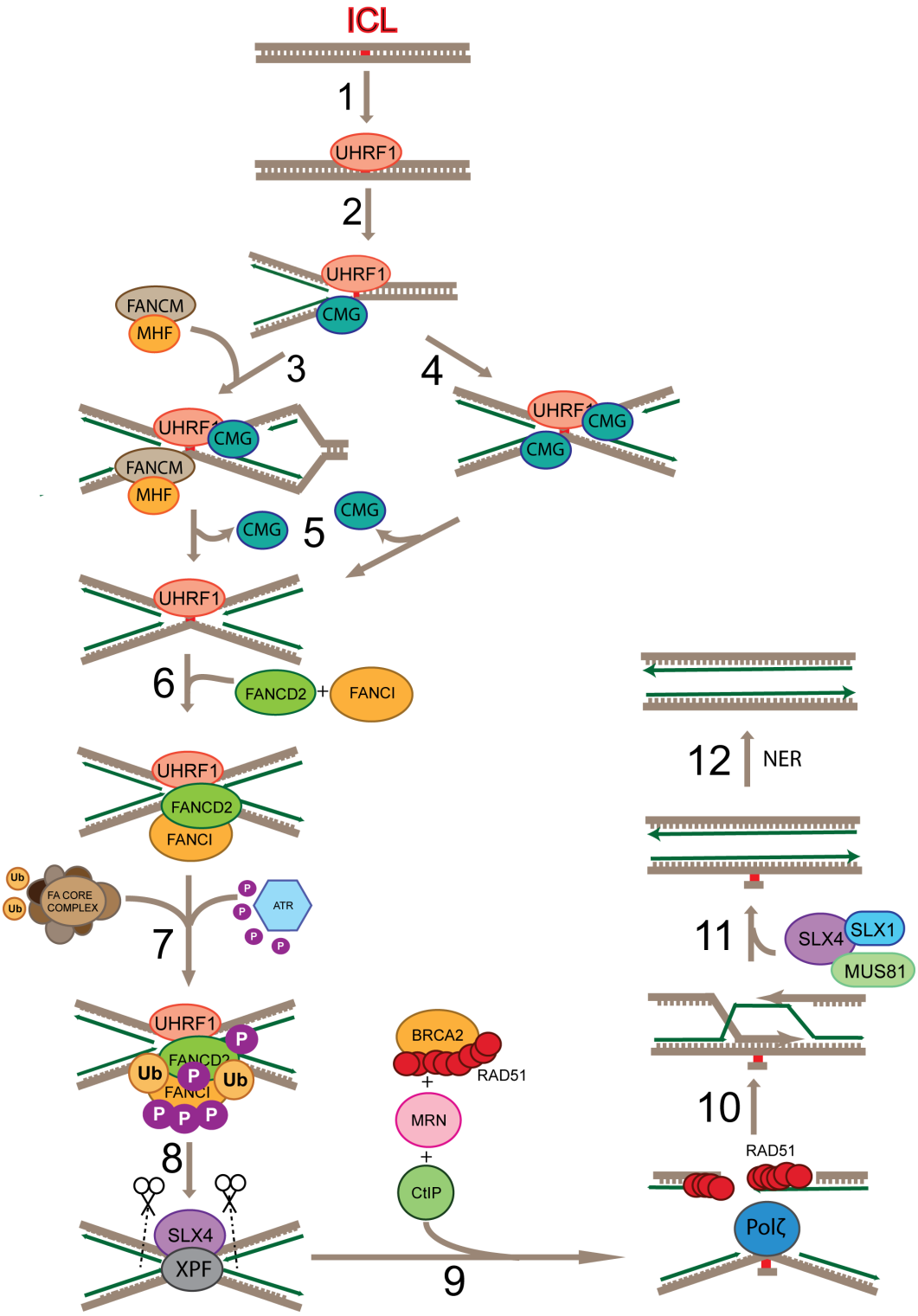


Figure 2: The molecular mechanism of the FA pathway. 1) UHRF1 is recruited to ICLs shortly after they are formed 2) Replication for arrives at the ICL 3) FANCM/MHF complex mediates the traverse of the replication complex through the ICL 4) Alternatively, second replication fork arrives at the ICL 5) The CMG complex is unloaded 6) FANCD2-FANCI complex is recruited 7) FANCD2 and FANCI undergo ubiquitination and phosphorylation 8) FANCD2/FANCI complex recruits SLX4/XPF to ICL to unhook the ICL. 9) CtIP and the MRN complex resect the DSB ends generated by the incision, BRCA2 facilitates Rad51 filament formation on the ssDNA 10) Polζ polymerizes new strand of DNA through the unhooked ICL, Rad51 initiates the strand invasion, 11) SLX4 and nucleases resolve the Holliday junctions. 12) NER pathway removes the damaged nucleotide

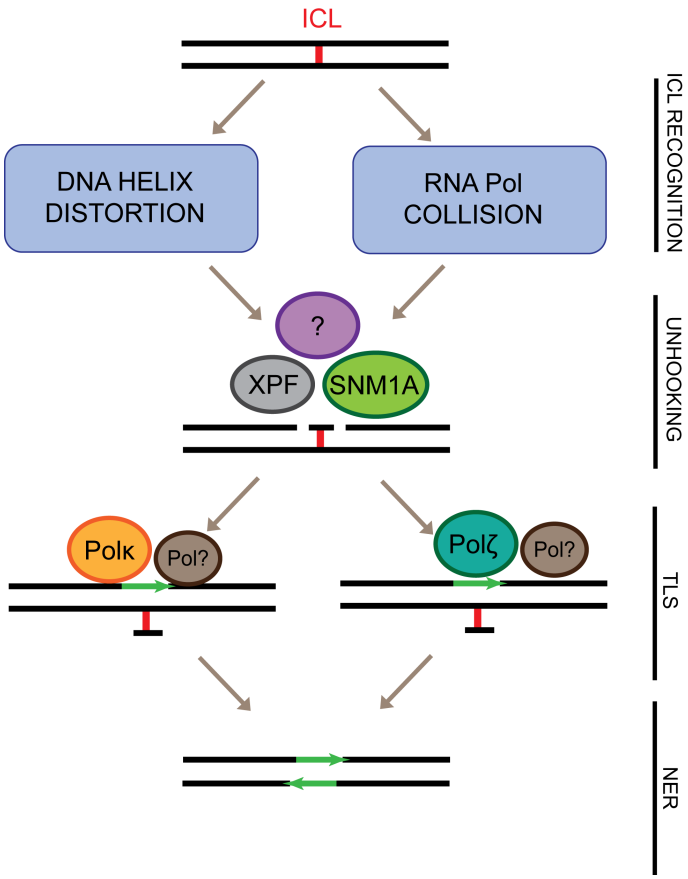


Figure 3: Schematic representation of the replication independent ICL repair. First ICL is unhooked, then TLS fills in the gap resulting in unhooking and NER repair pathway removes the crosslinked nucleotide.

1.2.2 Ubiquitination signalling in FA pathway

Ubiquitination is a post-translational modification, where 8.5 kDa protein ubiquitin is covalently linked to one of the lysine residues present within the sequence of a modified protein. There are three types of ubiquitination: mono- (when a single ubiquitin molecule is attached), multi- (when several monoubiquitinations happen within one protein at different lysine residues), and poly- (when ubiquitin chain is attached to the lysine residue instead of a single ubiquitin molecule).

There are many types of polyubiquitination. Each type differs by the lysine residue that links the ubiquitins. For example, in the K48 polyubiquitin chain, every next ubiquitin in the chain is linked to the K48 residue of the previous one. In case of K63 – to the K63 residue. K48 chains are marking the proteins for proteasome degradation, while K63 polyubiquitination are associated with

DNA repair pathways regulation and also inflammation (***Ulrich and Walden, 2010***). There are also other types of linkages, such as K11, which was implied in cell cycle control (***Komander and Rape, 2012***) or K27 which was reported to signal DNA damage (***Gatti et al., 2015***). While the exact function of the noncanonical ubiquitin linkages is not fully elucidated yet, it is becoming more and more evident that they are far more important and common among ubiquitinated proteins, than it was initially assumed.

It is important to note, that apart from K48 chains, that are consistently indicated in proteasome degradation, it would be very difficult to assign one single role to any of the other ubiquitin chains. Like many other posttranslational modifications, ubiquitination alters the structure of the modified proteins, therefore changing their molecular properties, allowing them to interact with different proteins and different pathways than their non-modified versions. Such modification can be used to regulate variety of pathways and functions, thus making the ubiquitination one of the most common mechanisms of the regulation of the cellular processes.

Ubiquitination is a complex process that requires three classes of enzymes: E1 (ubiquitin-activating enzymes), E2 (ubiquitin-conjugating enzymes) and E3 (ubiquitin ligases) (***Huang and D'Andrea, 2006***). In the FA pathway FANCT acts as E2, and FANCL as E3 enzyme.

When the ubiquitin or ubiquitin chain is no longer needed it can be removed by the deubiquitinating enzymes (DUBs), therefore making ubiquitination reversible. In the FA pathway this role is performed by the USP1-UAF1 complex (***Cohn et al., 2009***). Alternatively, in case of the polyubiquitin chains that target the protein to the proteasome, the chain may be digested along with the protein, which results in free ubiquitin being recycled.

Many of the proteins which participate in ubiquitin metabolism contain ubiquitin binding domains (UBD). There are more than twenty types of UBDs, which interact specifically with either free ubiquitin or ubiquitin chains. Especially common sub-set of UBDs among proteins that participate in FA pathway are UBZ domains (ubiquitin binding zinc finger domains) (***Kim and***

D'Andrea, 2012). They were originally discovered in the Y-family of translesion DNA polymerases and were shown to participate in postreplicative DNA repair in response to the UV light (***Bienko et al., 2005***). They coordinate a zinc ion with cysteine or histidine residues, depending on their amino acid sequence (***Hofmann, 2009***).

There are four types of UBZ domains: UBZ1, UBZ2, UBZ3 (majority of the UBZ domains belongs to this type) and UBZ4. The last two are especially important for DNA repair. UBZ3 contains a conserved C2H2 motif, while UBZ4 a conserved C2HC motif. UBZ3 can be found for example in Polymerase η, while type 4 UBZ is present in proteins such as FAN1, SLX4, FAAP20 and Rad18 and many other FA associated proteins.

WRNIP1, the main focus of this thesis, contains a UBZ domain of type 4 as well, which is one of the reasons why it seemed worthwhile to investigate it in context of ICL repair.

1.2.3 Interstrand cross links

Human genome is constantly exposed to a number of damaging factors (both endogenous and exogenous). Protection of its stability is therefore one of the most crucial biological processes. Cellular responses to genotoxic stress need to be effective, quick and highly specific to ensure that all the lesions are resolved in time for replication and transcription.

Among other types of DNA damage one of the most cytotoxic ones are interstrand crosslinks (ICLs). An interstrand crosslink is created when two complimentary DNA strands (Watson and Crick) are covalently linked together. Since this kind of obstacle is blocking the separation of the two strands, when left unresolved it can lead to very serious consequences such as chromosome breakage or missegragation.

An ICL lesion appears when the two strands get bound together via a linker molecule. The linker molecules are known as crosslinking agents, and not all of them introduce the same kind of ICLs. While the general nature of the lesion remains similar (two strands are covalently bound),

different crosslinking agents have different base specificity, and also cause different levels of DNA double helix distortion. The toxicity of the crosslinking agents depends also on factors such as cellular uptake and metabolic activation.

One of the crosslinking agents that are widely used in research is mitomycin C (MMC). It is a natural antitumour drug that forms adducts at N-7 and N-2 of guanine, and forms ICLs between the N-2 of guanines at d(CpG) minor groove sequences (***Tomasz, 1995***; ***Kumar at al., 1997***). The MMC-induced ICLs cause only a minor DNA distortion (***Tomasz, 1995***). As with all crosslinking agents ICLs are not the only adducts formed (they are usually between 5-13% of all adducts formed) (***Warren et al., 1998***).

Another group of important crosslinking agents consists of psoralens. They do not cause damage until they get activated by UVA light and after activation they cause thymine monoadducts and ICLs between thymines in d(TpA) fragments of DNA sequence. Since the damage depends on UVA, the amount of ICL lesions compared to other types of DNA damage can vary depending on exact wavelength and dose. In general, ICLs will be approximately 30-40% of all lesions introduced by the psoralen-UVA treatment (***Brendel and Ruhland, 1984***). The DNA distortion is minor in this case. It is also worth noting that this type of cross-link is asymmetric because one thymine is bound via furan ring and the other via pyrone (***Brendel and Ruhland, 1984***).

ICLs can be also formed by nitrosoureas (which cause cytosine and guanine monoadducts, guanine intrastrand crosslinks and guanine-cytosine ICLs), nitrogen mustards (these ICLs cause major DNA distortion, but they constitute only 5% of the damage while the rest is mostly monoadducts and intrastrand breaks (***Rink and Hopkins, 1995***), and finally platinum compounds. One of the most well studied platinum compounds is cisplatin, that mainly reacts with guanines, which causes mostly intrastrand crosslink damage and only around 5-8% of ICLs.

Endogenous ICLs are especially difficult to study. Most of them have shown mutagenicity in *in vitro* studies. The most common endogenous cross-linking agents belong to the family of reactive

aldehydes. One example of such aldehyde is acetaldehyde (which is formed during the metabolism of ethanol) and it was proven that FA- deficient cell lines are comparably sensitive to it as to other well studied crosslinking agents, which proves that the FA pathway is important in repairing acetaldehyde damage (***Langevin et al. 2011***). On the molecular level one molecule of acetaldehyde reacts with guanine, and forms N2-ethyl-2'-deoxyguanosine adducts, or if there are any basic molecules around (for example histones) two molecules may react with guanine and form 1-N2-propano-2'-deoxyguanosine. While the first adduct cannot form ICLs the one formed after reaction with two acetaldehyde molecules can exist in either cyclic or open chain configuration and in the open one free aldehyde group can induce ICLs or alternatively DNA-protein cross-links (***Brooks, Theruvathu 2005***). ICLs formed in that way are formed in the minor groove. Another type of aldehyde that can cause similar DNA damage is crotonaldehyde.

Another endogenous source of ICLs is lipid peroxidation in the presence of oxidative stress, which is one of the consequences of a bad diet rich in fat. One of the molecules that appears in cells as a consequence of ongoing lipid peroxidation is malondialdehyde, that can react not only with guanine but also adenine and cytosine. The main products of the reaction are ICLs formed between the guanines in the CpG sequence (***Niedernhofer at al., 2003***).

Finally, ICLs can be also induced by nitric oxide (NO), which generates them between two guanines in the CpG sequence. The residues are bound by N2 amine group. It has been shown that this reaction is more likely to occur in the presence of methylated cytosines (***Caulfield et al., 2003***).

Since different cross-linking agents introduce different types of damage on top of ICLs (which constitute only a fraction of the changes in DNA structure), it is important to remember that no treatment that can be used in potential experimental procedures is giving a pure condition with only one type of damage.

ICLs are not the most common DNA damage but nevertheless they have a lot of clinical relevance and it is very important to fully understand their formation and the repair mechanisms of

this type of damage. The main reason for this is that many of the cross-linking agents are used in the treatment of various cancers, so good understanding of the underlying repair processes is crucial for designing the most efficient therapies.

ICL-inducing agents are generally very good for the purpose of fighting cancer because they often allow the selective killing of cancer cells while healthy cells are mostly spared due to better DNA damage protection mechanisms. Because ICLs interfere with both replication and transcription (processes that are generally more active in rapidly dividing cells), and tumour cells tend to divide much more rapidly, they are likely to be affected much more than non-tumourigenic cells. Some kind of tumours, though are either inherently resistant to certain agents or become resistant after treatment.

1.3 Werner Helicase Interacting Protein 1 (WRNIP1)

WRNIP1 is a 665 aa long AAA+ ATPase that is highly conserved from *E. coli* to human. It consists of three separate domains (**Figure 4**). The ubiquitin-binding zinc-finger (UBZ) domain is responsible for binding both mono- **(*Crosetto et al., 2008*)** and poly-ubiquitin (***Bish and Myers, 2007*)**. The ATPase domain consists of Walker A and Walker B motifs that trap and hydrolyse ATP. This domain was proven to interact with WRN helicase **(*Kawabe et al., 2006*)**. Finally, at the C-terminus of the protein there is the leucine zipper domain (consisting of two predicted leucine zipper motifs, between 519 and 655 aa).

WRNIP1 is monoubiquitinated, polyubiquitinated, and also sumoylated (***Bish and Myers, 2007*)**, but the exact function of these post-translational modifications remains elusive. It was observed though that both ubiquitination of WRNIP1 as well as its binding to ubiquitin were completely abrogated after deletion of UBZ domain (or after introducing D37A point mutation in this domain) (***Bish and Myers, 2007; Bish et al., 2008***). The aspartic acid 37 of WRNIP1 interacts *via* noncovalent interactions with R42 residue of ubiquitin.

It was reported that WRNIP1 predominantly forms a homo-octameric complex (***Tsurimoto et al., 2005***). This data however was obtained using human WRNIP1, expressed and purified from insect cells, and so far it was not confirmed using WRNIP1 expressed in human cells.

WRNIP1 was first identified by a yeast two hybrid assay as a binding partner of the mouse Werner helicase **(*Kawabe et al., 2001*)**. It was confirmed in the same study that this interaction is direct, which raised a question if WRNIP1 and WRN act in the same pathway. This hypothesis was disproved by ***Kawabe et al. (2006).***

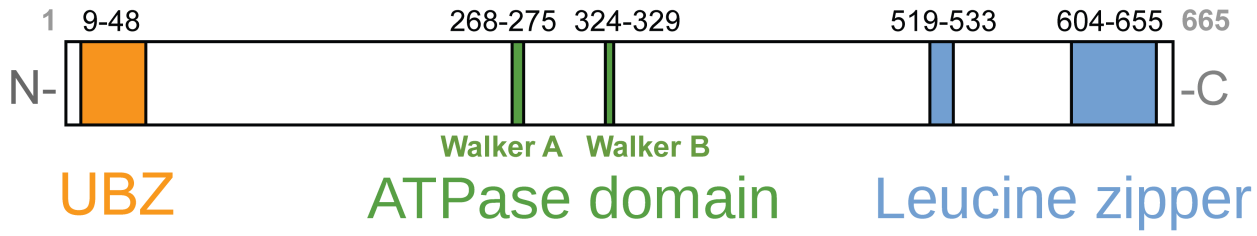


Figure 4: Domain architecture of WRNIP1. WRNIP1 is a 665 aa long protein, which contains a UBZ type 4 domain on its N-terminus (orange), ATPase domain, consisting of two Walker motifs: A and B (green), and leucine zipper domain at the C-terminus, containing two leucine zipper motifs (blue).

Apart from WRN helicase there are also other known interacting partners of WRNIP1. It was found to interact with three of the four subunits of DNA Polymerase δ (and to stimulate its activity) (***Tsurimoto et al., 2005***), to bind to the E3 ligase Rad18 (***Yoshimura et al., 2009***), and finally to interact with Polymerase η ***(Yoshimura et al., 2014***). Recently another interacting partner of WRNIP1 was discovered - ATM cofactor – ATMIN (ATM Interactor) (***Kanu et al., 2015***).

The best characterized homolog of WRNIP1 – yeast protein Mgs1 (Maintenance of genome stability) was identified by two independent studies (Kawabe *et al*., 2001 and Hishida *et al*., 2001). It interacts physically and functionally with Pol31, the second subunit of polymerase delta **(*Vijeh Motlagh et al. 2006*)** and was also reported to bind PCNA **(*Saugar et al., 2012*),** (ubiquitinated

PCNA targets Mgs1 to the DNA lesions (Saugar *et al.,* 2011)) and it has been suggested to play a role in bypass of stalled replication forks. Interestingly yeast lack the FA pathway, so if WRNIP1 indeed plays a role in the FA pathway, it must have evolved from some similar DNA- related pathway.

Mgs1 also has DNA-dependent ATPase and single-strand DNA annealing activities. Deletion of the gene *MGS1* results in elevated rates of mitotic recombination, and it was shown that this phenotype is synergistic with mutation in *SGS1 (*a gene encoding the yeast homolog of WRN): the *mgs1 sgs1* double mutant presents more severe genome instability than any of the single-gene mutants (***Yoshimura, 2017***).

While genetic analyses of Mgs1 have shown that the protein is required for preventing genome instability **(*Hishida et al., 2001*)**, the role of WRNIP1 remains elusive.

WRNIP1 also shares sequence similarity with *E. coli* RuvB protein, which also exhibits ATPase activity, was observed to form a hexameric ring complex, and is known to drive DNA strand migration at Holliday junction structures while forming a complex with RuvA protein (***Yamada et al. 2002***). So far, no functional involvement of WRNIP1 in Holliday junction migration in eukaryotes has been reported, though.

Experiments with laser irradiation have shown that WRNIP1 accumulates at the DNA damage sites such as double strand breaks **(*Nomura et al. 2012*)** and it was suggested that UVC treatment causes the increase in the number of its nuclear foci **(*Crossetto et al. 2008*)**.

In vitro analysis has shown that WRNIP1 binds in an ATP-dependent manner to forked DNA that mimics stalled replication forks (***Yoshimura et al., 2009***) as well as short and long DNA fragments bound together which resemble the template-primer pair (***Kanamori et al., 2011***).

Additionally, microscopy data proved that WRNIP1 foci tend to overlap with replication factories, which reinforces the hypothesis of its function at replication forks **(*Crossetto et al. 2008*)**.

The latest findings suggest that WRNIP1 may have an important role in protecting stalled

replication forks from degradation and promoting fork restart after the damage is repaired, which is a function that it performs in cooperation with RAD51. WRNIP1 seems to promote stabilization of RAD51 on ssDNA, which in turn prevents uncontrolled MRE11- mediated degradation of stalled replication forks. Interestingly the study suggests that while the protection does not require WRNIP1 to have ATPase activity, but this activity is needed for the recovery of the stalled fork (***Leuzzi et al., 2016***).

1.4 Aims of the thesis

Since WRNIP1 has an interesting domain architecture (previously described UBZ domain at N-terminus and ATPase domain in the middle) and it was also observed to be recruited to the DNA damage site (after laser irradiation, which results in both single and double strand breaks) it seemed likely that WRNIP1 is involved in DNA repair.

The general goal of this doctoral project was to either confirm or reject the hypothesis that WRNIP1 is involved in the FA pathway. The detailed aims were as follows:

a) to test if WRNIP1 depleted cell lines are more sensitive to ICL damage than wild type cell lines

b) to determine whether WRNIP1 function is associated with FANCD2 under ICL-inducing genotoxic stress

c) to investigate WRNIP1 recruitment to the chromatin, after introducing ICL damage, especially in relation to FANCD2 recruitment

d) to investigate the nature of WRNIP1 ubiquitinated form (in particular to find the exact lysine residue which is monoubiquitinated, and then to see if abrogation of this modification has any functional implications)

e) to determine the role each of the WRNIP1 domains plays in its molecular function

f) to see if WRNIP1 and FANCD2 interact directly with each other

2. Materials and Methods

2.1 Materials

All reagents, solvents and buffers are described in detail in methods section, when the procedure during which they were used is mentioned. They were all purchased from Sigma-Aldrich unless indicated otherwise.

Many cell lines were obtained during the course of this doctoral project, by modification of other basic cell lines, such as HeLa, Phoenix A or *Spodoptera frugiperda* Sf9 insect cells. Each of these basic non-modified cell lines is described in section 2.1.1 (Cell lines).

Similarly, the plasmids obtained during the course of this study are described in the Methods section, unless they were provided by external sources. The ones obtained from the external sources are described in section 2.1.2 (Plasmids).

All the primers used in this study were purchased from Eurofins Genomics.

2.1.1 Cell lines

The HeLa cells, HeLa S3 cells and Phoenix A cells were obtained from American Type Culture Collection (ATCC). Flp-in T-REx HeLa cells were purchased from Thermo Fisher Scientific.

HeLa (FANCD2 $^{-/-}$) cell line was obtained from David Lopez Martinez (Dr Cohn's lab), who obtained it using the CRISPR/Cas9 gene editing technology. The targeting sgRNA sequence used was: 5'-GTTTGTCTTGTGAGCGTCTGC-3' and it was targeting the forth exon of FANCD2. The following two primers were annealed and introduced into the pX459 plasmid through its BbsI site: 5'-CACCGTTTGTCTTGTGAGCGTCTGC-3' and 5'-AAACGCAGACGCTCACAAGACAAAC-3'. Then 2 μg of the modified pX459 plasmid was transfected into HeLa cells and selected with 4 μg/ml of puromycin for 24 hours. The surviving cells were plated at low density incubated for two weeks to allow the formation of single clone colonies. Later the clones were analysed by Western

blot analysis (***Liang et al. 2016***) to confirm the full depletion of FANCD2.

2.1.2 Plasmids

The plasmids used during the course of this doctoral project, are listed in ***Table 1***.

Name	Procedure
pcDNA5-FRT/TO-EGFP (kindly provided by prof Francis Barr)	obtaining doxycyclin (dox) promotor regulated, WRNIP1-EGFP cell line for both survival assays and live cell microscopy (as well as all WRNIP1-EGFP point mutants)
pOG44 (kindly provided by dr Fumiko Esashi)	Flp-recombinase expression vector (helper vector that needs to be transfected along with pcDNA5-FRT/TO-EGFP)
pOZ-N (FLAG-HA tag)	expressing, and purifying FLAG-HA-WRNIP1 to purify upper form for mass spectrometry analysis
pG+P	helper plasmid encoding viral polymerase (Phoenix A transfection system)
pEnv	helper plasmid encoding envelope protein, needed for assembly of viral particles (Phoenix A transfection protocol)
pX459 (Addgene #48139) kindly provided by dr Andrew Bassett)	plasmid used for introduction of the target sequence into cells, during CRISPR/Cas9 genome editing
pBlueScript	plasmid used for introduction of the homology cassette into cells, during CRISPR/Cas9 genome editing
pET28-c	plasmid used (after cloning the sequence of WRNIP1 fragments inside) to transform *E. coli* BL21 (λDE3) cells, for expression and purification of WRNIP1 fragments

Table 1: *Plasmids used during the course of this doctoral project*

2.1.3 Antibodies

Primary antibodies used in this doctoral project were as follows: α-WRNIP1, 1:3000 or 1:10 000 dilution depending on the experiment (WHIP A8, sc-376438, Santa Cruz Biotechnology); α-Tubulin, 1:2000 dilution (5829, Millipore); α-FANCD2, 1:100 dilution (sc-20022, Santa Cruz Biotechnology); α-HA, 1:1000 dilution (11583816001, Roche); FK2, 1:1000 dilution (04-263, Emdmillipore); and α-His, 1:1000 dilution (H1029, Sigma)**.** Secondary antibody used in this project was α-mouse (Horseradish Peroxidase conjugated), 1:2000 dilution (NA9310V, GE Healthcare).

2.2 Methods

2.2.1 Western Blot (WB) whole lysate sample preparation

For all WB whole cell lysate tests cells were grown in 9 cm plates in an incubator overnight, then scraped off with a plastic spatula. The medium with the scraped off cells was transferred to a 15 ml Falcon tube and spun down (1000 RPM; 5 min), then resuspended in 1 ml of Phosphate-buffered saline (PBS) buffer, transferred into an Eppendorf tube and spun down (4000 RPM; 2 min; 4°C). PBS was aspirated, spun down again and the remains of the supernatant were aspirated once again. The volume of the pellet was estimated and equal volume of benzonase buffer (20 mM Tris (pH 8.0), 10% Glycerol, 2 mM $MgCl_2$, 1% Triton X-100) with Benzonase (12.5 units/ml) was added. The sample was vortexed then put on ice for 10 min. After on ice incubation equal volume of 2 % sodium dodecyl sulfate (SDS) (twice the volume of the pellet) was added then the sample was vortexed and heated up at 70°C for 2 min. The amount of protein in each lysate sample was measured using the Bradford assay and ready to load samples were prepared with the concentration of 1 μg/μl.

2.2.2 Tissue culture

All cell lines were grown in standard Dulbecco's Modified Eagle's Medium - DMEM (Sigma, D5796) supplemented with 10% Foetal Bovine Serum - FBS. The temperature in the incubator was kept at 37°C at all times and the concentration of CO_2 was kept at 5%. All cells were passaged every two to three days in order to maintain their density between approximately 30% and 70% and keep them from overgrowing. During each passage the cells were washed with pre-warmed PBS buffer supplemented with 1 mM EDTA (Ethylenediaminetetraacetic acid), then incubated for 3 min in PBS-trypsin solution. After the incubation equal volume of DMEM (supplemented with 10% FBS) was added to the dish, the cells were detached, spun down (1000

RPM, 5 min), the medium was aspirated and the cell pellet was resuspended in fresh medium. The solution (diluted accordingly depending on the number of cells) was plated into a new dish.

2.2.3 Obtaining Flp-in T-REx HeLa WRNIP1-EGFP cell lines

Flp-in T-REx is a HeLa cell line modified to contain an FRT integration site conferring Zeocin resistance and a tetracycline repressor gene selected by Blasticidine resistance. A transgene of choice (WRNIP1-EGFP and its many point mutants in this case) cloned into a pcDNA5-FRT/TO-EGFP vector can be integrated specifically into the FRT site by co-transfecting the Flp-in recombinase (on the pOG44 plasmid).

WRNIP1 sequence was amplified by PCR reaction (***Table 2***) from the pFB-WRNIP1 plasmid provided by Alicja Bulsiewicz (Dr Cohn's group), using the following primers:

Forward: WRNIP1BamHIpcD

5' - AGTCG↓GATCCACCATGGAGGTGAGCGGGCCGGAAG – 3'

Reverse: WRNIP1SalIpcD

5' – AGTCG↓TCGACGCA CCT CCT CTG CTTGAA GAA ATC TAC C – 3'

In both cases the ↓ symbol indicates the restriction site via which the amplified fragment was later inserted (BamHI in the first case and SalI in the other). pcDNA5-FRT/TO-EGFP was digested at 37°C for 1 h using XhoI and BamHI enzymes, to ensure that the endings of both insert and vector are compatible and can be ligated later. After the digestion the enzymes were deactivated at 65°C for 20 min and the vector was dephosphorylated using the alkaline phosphatase CIP (37°C, 1 h incubation), followed by inactivation of the phosphatase (65°C, 20 min incubation) and gel purification.

The insert was amplified from the template using the following PCR solution:

SAMPLE	**VOLUME [µl]**
milliQ water	35.0
10x buffer Pfu (Ultra II Buffer)	5
dNTP (10 mM)	0.5
DNA template (pFB-WRNIP1) (5 ng/µl)	4
Forward Primer (10 µM)	1
Reverse Primer (10 µM)	1
Pfu polymerase (2.5 U/µl)	1
DMSO	2.5

T*able 2: PCR solution used to initially amplify WRNIP1 sequence from the pFB-WRNIP1 vector*

The following PCR conditions were used:

	2 min – 95°C	(additional denaturation)
x 20 cycles	30 s – 95°C	(denaturation)
	30 s – 72°C	(annealing)
	3 min – 72°C	(extension)
	10 min – 72°C	(additional extension)

After amplifying and digesting the insert and digesting and dephosphorylating the vector, both DNA particles were separated using electrophoresis on agarose gel, then carefully cut out from the gel using scalpel, and gel extracted using the *QIAquick* gel extraction Kit. Finally, they were ligated overnight at 16°C using the T4 ligase, and the ligated sample was used to transform *E. coli Top 10* cells. As cDNA5-FRT/TO-EGFP carries Ampillicin resistance, the cells were plated on Ampicillin supplemented LB-agar plates, and then colonies that appeared after 12 hours were used to inoculate overnight liquid culture (LB broth supplemented with Ampillicin). The liquid colonies were incubated at 37°C overnight and then plasmids were purified from them using Miniprep alkaline lysis method. All obtained plasmids were tested with sequencing to confirm that we obtained pcDNA5-FRT/TO-[WRNIP1]-EGFP construct and then one of the successfully obtained samples was used in future procedure.

After the cloning step, both pcDNA5-FRT/TO-[WRNIP1]-EGFP and pOG44 plasmids were transfected into the HeLa Flp-in T-Rex cells using the following transfection mix: 200 µl of DMEM

+ 6 µl Fugene6 (Promega) + 200 ng of the construct (pcDNA5-FRT/TO-EGFP with insert of choice (either WRNIP1 or WRNIP1 with a point mutation)) + 1800 ng pOG44. The recombination event destroys the Zeocin resistance and introduces Hygromycin resistance (the Hygromycin resistance gene encoded by pcDNA5-FRT/TO-EGFP only works once integrated into the Flp-in cell line since the gene on pcDNA5-FRT/TO-EGFP lacks the start ATG codon which is provided by the FRT integration site in the cellular genome. pcDNA5 plasmids can therefore not be used for stable expression in other cell lines. The tetracycline repression of the transgene operates unless switched off by adding tetracycline (or the more stable doxycycline (dox)) which then permits expression of the integrated epitope-tagged transgene.

After transfection cells were plated in low concentration, and selected with hygromycin to allow single clone selection after the formation of visible colonies. For both WRNIP1-EGFP and point mutation variants of WRNIP1-EGFP two clones were always selected, to exclude the possibility of a random clonal effect (where the obtained cell line presents changed phenotype not because of the transfected protein but because of some random mutation which occurred in the first cell which started the new cell line).

The Flp-in T-REx HeLa cells were grown in the standard DMEM medium supplemented with 10% FBS and passaged every two to three days according to the standard HeLa cells splitting protocol.

2.2.4 Phoenix A transfections

The Phoenix A cells are based on a human embryonic kidney line transformed with adenovirus E1a and carrying a temperature sensitive T antigen co-selected with neomycin (the 293T cell line). They are second-generation retrovirus producer line and they are able to generate helper free ecotropic and amphotropic retroviruses.

To use the system first Phoenix A cells need to be transfected with the construct, then

incubated for 2-3 days to generate retrovirus and the supernatant (DMEM +10% FBS medium) in which they were growing needs to be filtered with 0.45 μm syringe filter (to make sure that no Phoenix A cells are contaminating the cells that are being infected) onto regular HeLa cells.

HeLa transfected with the construct DNA have to be selected with magnetic beads. pOZ-N vector has the IL-2 receptor encoded within its sequence. By mixing the cells with IL-2 antibody connected to magnetic beads it is possible to select transfected cells from untransfected.

2.2.5 Western Blot assays

The samples were loaded onto previously prepared polyacrylamide-SDS gel along with the Mark12 Protein Standard (Thermo Fisher Scientific). The percentage of polyacrylamide was adjusted for each protein in such way that the marker will have at least one band of higher molecular weight than the analysed protein and at least one lower.

The voltage was usually kept around 100 V while the protein of interest was passing through the separating gel, and then the gel was removed from its covers and applied onto a nitrocellulose membrane. Both the membrane and the gel where then placed in the transfer buffer and the proteins were transported onto the membrane via electroblotting. Once the proteins were transferred on the membrane, the membrane was stained with the Ponceau S dye and the bands of Mark12 were marked with the regular pen before immersing the membrane in the 5% milk in Tris-buffered saline with 0.1% Tween 20 (TBS-T) solution for an hour of blocking to prevent nonspecific binding.

After the blocking stage the membranes were covered with the same 5% milk TBS-T buffer but supplemented with the indicated dilution of primary antibody. The membranes were then left overnight at 4°C.

The next day they were washed in the TBS-T solution three times for 6 minutes each wash and secondary antibody solution was applied on each of the membranes for an hour. After the secondary binding was finished the membranes were again washed three times for 6 minutes in

TBS-T buffer, and then ECL Western Blotting Substrate was applied on the washed membranes.

The now fluorescent protein bands were visualized using the chemiluminescence film (Amersham Hyperfilm ECL). The previously marked Mark12 bands were copied onto the acquired image.

2.2.6 Fixed Cells Microscopy

The cells prepared for fixed cells microscopy were passaged using the standard protocol and transferred into the new dish at desired confluency. Thin microscopy cover glass was placed in the dish beforehand, so some cells were forced to attach themselves to the glass instead of the bottom of the dish. After the cells were properly attached dox was added to the medium at the time and concentration indicated for each sample.

When the incubation with dox was over, the medium was aspirated and the cells were washed with PBS buffer (unsupplemented with EDTA). After the initial wash the cells were fixed in 4% paraformaldehyde/ PBS at room-temperature for 10 min and then washed three times for 5 min in 0.1% PBS-T.

After the cells were fixed they were stained with 0.05 µg/ml DAPI (4',6-diamidino-2-phenylindole) for exactly 5 min (covered from light during the procedure), then washed once with ultrapure water (Milli-Q) water. The microscopy slide was picked up with tweezers quickly submerged into the falcon tube with fresh Milli-Q water. The excessive liquid was removed with a tissue gently (without touching the surface covered with now fixed and stained cells) and the glass was placed on the drop of 7 µl Vectashield (unsupplemented with DAPI) placed on the microscopy sample glass stand in such way that the cells were facing downwards (submerged in Vectashield).

The cover glass was sealed with standard nail polish, and then the samples were stored at 4°C and protected from light.

For microscopy analysis images were acquired using 60x oil objective (Olympus 60X/1.42,

PlanApo, N) on the Delta Vision ELITE microscope installed with Resolve 3D SoftWoRx (Acquire Precision) Version 4.0.0 and the images were captured by the camera (CoolSNAP_HQ/ICX285) in both DAPI and FITC channels. The images were deconvolved using SoftWoRx (Applied Precision) software (ratio=conservative) and processed using FiJi.

2.2.7 Live-cell microscopy and chromatin recruitment assay

Cells used for live-cell microscopy were plated into microscopy dishes (*MatTek Glass Bottom Microwell Dishes P35G-1.5-20-C*) and pre-treated accordingly (in case of WRNIP1-EGFP cells and the variants of WRNIP1-EGFP with point mutations they were supplemented with 0.005 μg/ml of dox for 3 h). Then the medium was supplemented with 500 ng/ml of TMP (Sigma, T6137) and the dish was placed in the stand of the PE Ultraview spinning disk microscope for 15 min. The used objective properties were as follows: magnification: 60x, numerical aperture: 1.4, immersion medium: oil (N=1.514). 405 nm wave length laser irradiation was used to irradiate the samples along indicated lines and the before and after images were acquired for each sample. The images were acquired in the green (488 nm fluorescence channel and/or red (561 nm) channels. The used camera was Hamamatsu C9100-13. All images were acquired and processed using Volocity software.

After acquisition the images were opened with FiJi software and the following measurements were taken. First the places where the cells were irradiated were circled by hand using the free-select tool. It was necessary to do it by hand for each image because the irradiated areas tend to move along with the cell during the long acquisition and they are not always exactly on the same spot on the field of view as they were introduced. Then the measurement of the brightness of the selected field was performed and written down. After that the rest of the cell nucleus was selected avoiding the irradiated part, and the same measurement was performed. At the end part of the background was selected and its intensity was measured. Then for each cell the

following formula was used:

$$S_x=(A_x-B)/(C_x-B)$$

Where A_x represents the brightness of the stripe, C_x the brightness of the rest of the rest of the cell nucleus and B the brightness of the background in given field of view. The formula is returning S_x, which represents the relative brightness of the stripe compared to the rest of the nucleus in the given cell. After obtaining S_x values for each of the cells in the view field and for each time point of interest, the average S value for the experiment was calculated. Since for each presented quantification at least two independent experiments were combined, the standard error for each point equals standard deviation of the average values calculated for each experiment, divided by a square root of N- the number of all the independent experiments that were quantified (so for example N=2 or N=3). It is important to note that N is not equal to the total number of cells that are behind the calculated values, since for example if during the first independent experiment there were 3 cells quantified, and during the second there were 5 of them, the total number of cells would be 8, but N would still equal 2 since, there were only two independent experiments quantified.

To make the comparison between the values easier each of them was then normalized in such way that the first time point (before the irradiation) would always equal 1. To normalize, the difference between the first timepoint and 1 was added or subtracted in all later measurements (so for example if the average value for all cells in the first timepoint was giving the value of 1.12, 0.12 would be subtracted from this and all future timepoints). It was a necessary adjustment since, the distribution of the protein was not uniform in the nucleus, with dark spots and sometimes bright foci.

Since the measurements depend heavily on the precision of the hand-selection of the areas of interest it is of course to certain extent prone to unavoidable human error but given the length of acquisition time and the fact that we were observing live cells that were likely to change their position during the time of the experiment (sometimes even bending the stripe slightly) it was the

most accurate way of measuring the brightness of the irradiated chromatin.

2.2.8 Obtaining CRISPR/Cas9 gene disruption cell lines

One of the simplest ways to investigate the role of a given protein is to block its expression and observe the phenotype of a cell line in its absence. The RNA interference is relatively effective and gives good results but it does not ensure the complete lack of expression that is possible to achieve with the CRISPR/Cas9 gene disruption technique (***Wang et al., 2013; Cong et al., 2013; Mali et al., 2013***).

CRISPR - clustered regularly interspaced short palindromic repeats – (initially named SRSRs – short regularly spaced repeats (***Mojica et al., 2000***)) are bacterial DNA sequences containing short repetitions. Each repetition is followed by a short segment of spacer DNA. CRISPR loci are also associated with Cas genes, and CRISPR/Cas system provides a form of an acquired immunity to foreign genetic elements in prokaryotic organisms. The mechanism is as follows: spacers from invading viruses are incorporated at CRISPR loci to play the role of the memory of the previous infections. Re-infection triggers the complementary CRISPR RNA (crRNA) to find a matching sequence. This provides the CRISPR-associated (Cas) nuclease the specificity to form a double- strand break at specific "foreign" DNA sequences.

CRISPR/Cas system can be used as a tool for RNA-guided genome editing.

First the genomic sequence of the WRNIP1 gene was submitted into the bioinformatic tool available online (*crispr.mit.edu*) that specializes in identifying potential CRISPR target sequences.

The sequence identified by the software (5'- CGAGTTGATGTGCGCGGC – 3') belongs to the first exon in WRNIP1 sequence (there are seven exons in total). Two complementary oligos were designed, both flanked with short sequences matching the ends of pX459 vector after introducing the cut with BbsI enzyme.

(BbsI cutting site: `5'-GAAGACNN/ -3'`
`3'-CTTCTGNNNNNN/-3').`

The designed oligos had the final sequence of:

WRNIP1_CR1+ 5' CAC CGG CGA GTT GAT GTG CGC GGC 3'

WRNIP1_CR1- 5' AAA CGC CGC GCA CAT CAA CTC GCC 3'

10 µl of 2x annealing buffer (20 mM Tris, 2 mM EDTA, 100 mM NaCl, pH 8.0) was then mixed with 5 µl each of 100 µM forward and reverse oligos. The oligos were then annealed in the PCR machine using the following program:

98˚C	1 min
98-88˚C	5 s, decrease 0.1˚C/cycle x99 cycles
88-78˚C	10 s, decrease 0.1˚C/cycle x99 cycles
78-68˚C	10 s, decrease 0.1˚C/cycle x99 cycles
68-58˚C	10 s, decrease 0.1˚C/cycle x99 cycles
58-48˚C	10 s, decrease 0.1˚C/cycle x99 cycles
48-38˚C	10 s, decrease 0.1˚C/cycle x99 cycles
38-18˚C	1 s, decrease 0.2˚C/cycle x99 cycles
4˚C	Forever

Then the oligos were phosphorylated in the following mixture:

1 µl	Annealed oligo
1 µl	Roche T4 DNA ligase buffer (contains ATP)
1 µl	T4 PNK
7 µl	ddH_2O
10 µl	Total

Afterwards they were incubated for 30 min at 37˚C, and diluted 10 -fold in Milli-Q H_2O. At the same time 2 µg of the pX459 vector was digested in the following mixture:

2 µg	pX459 vector
5 µl	10x buffer
x µl	Milli-Q H_2O
2 µl	Bbs I
50 µl	Total

1 ul of CIP was added for last 10 min of digestion (and afterwards CIP was heat inactivated). The digested vector was loaded into agarose gel, separated from DNA fragments of different size using electrophoresis, cut out with a clean scalpel and extracted from the agarose gel using the QIAquick Gel Extraction Kit (QIAGEN). The purified vector was ligated with the oligos in the

following mixture overnight at 16°C:

1 µl	Vector
2 µl	10x diluted oligo
1 µl	Roche T4 DNA ligase buffer
4 µl	Milli-Q H_2O
1 µl	T4 DNA ligase
10 µl	Total

The plasmids obtained in such way were transformed into *E. coli Top 10* strain and the clones that were picked after incubation were tested by sequencing. One construct with successfully inserted target sequence was picked and transfected into HeLa cell line of choice.

To transfect, 2 µg of the purified plasmid was mixed with 200 µl of the unsupplemented DMEM medium, and 6 µl of the Fugene6 reagent (Promega) was added to the solution. The transfection mix was mixed afterwards and applied onto a small dish with a 5 ml of DMEM medium (supplemented with 10% of the FBS serum) and the cells of the cell line we intended to transfect. After 24 h the procedure was repeated and after another 24 h the medium was exchanged to the medium containing 1 µg/ml puromycin (since the cells that were successfully transfected should have temporary resistance to the drug). Then after selection the cells were split into bigger 15 cm dish in several dilutions in such way that allows picking single clones after several days. The clones were later picked grown to higher confluency in separate dishes and tested for presence of WRNIP1 using Western Blot analysis.

FANCD2 $^{-/-}$ CRISPR/Cas9 cell line was obtained (according to analogous protocol) by David Lopez Martinez (Dr Cohn's group), and tested using Western blot assay to confirm the gene disruption of the FANCD2.

2.2.9 CRISPR/Cas9 insertional complementation of WRNIP1-EGFP

The partial CRISPR/Cas9 insertional complementation of WRNIP1-EGFP was obtained using the analogous protocol to the one used for CRISPR/Cas9 gene disruption, with several modifications. Firstly, the insertional complementation cell line required the target site to be located

near the C-terminus of the WRNIP1 sequence instead of N-terminus (as this time the aim was to tag WRNIP1 with EGFP at its C-terminus).

The target sequence was identified and cloned into the pX459 vector using the same software and identical procedure as was described for the gene disruption protocol. The oligonucleotides used to introduce the target sequence were as follows:

WRNIP1_C_CR+: 5' – CACCGCCTAGTGGGCGCGTTCCTC – 3'

WRNIP1_C_CR-: 5' – AAACGAGGAACGCGCCCACTAGGC – 3'

But since this time the procedure required not only cutting the genome but also introducing EGFP sequence into it (via HR repair), it required introduction of homologous region as well. The homologous region had to contain both a region upstream of the target sequence (approximately 1 kb), downstream of the target sequence (also 1 kb) and EGFP sequence in between to allow it to be introduced by HR. It was assembled via Gibson assembly, using pBlueScript vector as a backbone (Figure 5).

Gibson assembly is a molecular cloning method which allows for several DNA fragments to join in a single reaction. Fragments need to have 20-40 bp overlaps and are mixed with three enzymes: exonuclease that digests DNA from the 5' end to form ssDNA fragments that can anneal, DNA polymerase to fill in gaps and ligase.

To obtain the fragments needed for the Gibson assembly first the three needed parts needed to be amplified from the genomic DNA (in case of the upstream, and downstream fragment), or from the pcDNA5-FRT/TO-EGFP vector in case of the EGFP fragment.

Figure 5: The region homologous to WRNIP1 sequence (with inserted EGFP tag sequence) cloned into pBlueScript vector. The homologous region, contains both the sequence upstream to the target site (represented in red) and downstream (represented in blue). Between them there is EGFP sequence (represented in yellow), which will be inserted into genome, via HR repair. The backbone (grey) is pBlueScript vector, and the whole construct is 5698 bp long. The black lines indicated on the schematic of the construct represent the primers needed to amplify each of the three fragments that the homologous region is assembled from.

The following primers were designed to allow the amplification of each of the three fragments, flanked with sequences needed for Gibson assembly (**Table 3**):

No.	Primer Name	Sequence (5'→3')
1	KI_WRNIP1up+	ATTGGGTACCGGGCCCCCCCTGTGTGTCATGACTGAAACTC
2	KI_WRNIP1up-	CCATCTCGACGCACCTCCTCTGCTTGAAG
3	KI_WRNIP1down+	GACTCAGTAGCTCCTCAGGGCACGACAG
4	KI_WRNIP1down-	GCTTATCGATACCGTCGACCCACTCAGAATAAAGGCAGAGGAC
5	KI_WRNIP1gfp+	GAGGAGGTGCGTCGAGATGGTGAGCAAGG
6	KI_WRNIP1gfp-	CCCTGAGGAGCTACTGAGTCCGGCCGGA

Table 3: Primers used to obtain three fragments needed for Gibson assembly

All the fragments were obtained using the PCR reaction, while the pBlueScript backbone vector was digested with XhoI enzyme (37°C, 1 h), dephosphorylated (CIP enzyme, 37°C, 20 min), separated on the agarose gel with electrophoresis, and extracted from the gel using the *QIAquick* gel extraction Kit.

Once all four fragments needed for the Gibson assembly were ready, they were assembled following the *HiFi DNA Assembly Protocol (NEB)* using the NEB Master Mix solution (2x).

The obtained construct was amplified using the *E. coli* transformation and Miniprep protocol, and then it was confirmed with sequencing that the sequence is correct. Before the construct could be used for the HeLa cells transfection, it had to be mutated in such way as to make it resistant to the CRISPR/Cas9 targeting, because otherwise it would be at risk of being targeted again after being already successfully modified via the CRISPR/Cas9 editing. To achieve it it was edited with a point mutation, using the standard PCR reaction and the following primers:

KI_WRNIP1r+: 5' – CCGTGCCCCTGCATCTGAGGAACGCGCCCACTAG – 3'

KI_WRNIP1r-: 5' – CTAGTGGGCGCGTTCCTCAGATGCAGGGGCACGG – 3'

After the mutation was successfully introduced, it was checked with the sequencing, and then the ready homologous construct, along with the targeting construct were used to transfect HeLa cells (protocol analogous to the one used to obtain the gene disruption CRISPR/Cas9 cell line – the only difference being, that instead of 2 μg of the targeting construct, the transfection mix was supplemented with 0.15 μg of the targeting construct and 1.85 μg of the homologous region construct).

2.2.10 Point mutations introduction into WRNIP1-EGFP construct

To introduce point mutations into WRNIP1-EGFP construct, previously obtained pcDNA5-FRT/TO-WRNIP1-EGFP construct (chapter 2.2.3) was amplified using the primers designed to introduce desired point mutations. The used primers were as follows:

For N33A mutation highlighted gcc sequence was inserted in place of acc:

N33A+: 5'- ccgccgcgcacatc**gcc**tcgcacctggaccg - 3'
N33A-: 5'- cggtccaggtgcga**ggc**gatgtgcgcggcgg - 3'

For D37A mutation highlighted gcc sequence was inserted in place of gac:

D37A+: 5'-gcacatcaactcgcacctg**gcc**cgctgtctgctgctccacc -3'
D37A-: 5'-ggtggagcagcagacagcg**ggc**caggtgcgagttgatgtgc -3'

For A45 V mutation highlighted gtg sequence was inserted in place of gcg:

A45V+: 5'-ctgctccacccg**gtg**gggcacgcggag -3'
A45V-: 5'-ctccgcgtgccc**cac**cgggtggagcag -3'

For K70R mutation highlighted agg sequence was inserted in place of aag:

K70R+: 5'- ccgcccggcgcc**agg**aggcggcggctg -3'
K70R-: 5'- cagccgccgcct**cct**ggcgccgggcgg -3'

For K274R mutation highlighted agg sequence was inserted in place of aag:

K274R+: 5' - gccgggctgcggc**agg**accactctggctc - 3'
K274R-: 5' - gagccagagtggt**cct**gccgcagcccggc -3'

For K301R mutation highlighted agg sequence was inserted in place of aag:

K301R+: 5'- ctgcaacaaatgcc**agg**acaaatgatgtgcgagatg -3'
K301R-: 5'- catctcgcacatcatttgt**cct**ggcatttgttgcag –3'

For K310R mutation highlighted aga sequence was inserted in place of aaa:

K310R+: 5' - gtgcgagatgtcata**aga**caagctcaaaatgaaaag - 3'
K310R-: 5' - cttttcattttgagcttg**tct**tatgacatctcgcac – 3'

For E329Q mutation highlighted cag sequence was inserted in place of gag:

E329Q+: 5'-gaaaaccatccttttattgat**cag**attcatcggttcaataaatctc - 3'
E329Q-: 5'-gagatttattgaaccgatgaat**ctg**atcaataaaaaggatggttttc - 3'

For K335R mutation highlighted aga sequence was inserted in place of aaa:

K335R+: 5'- gagattcatcggttcaat**aga**tctcagcaggacactttcc - 3'
K335R-: 5'- ggaaagtgtcctgctgaga**tct**attgaaccgatgaatctc - 3'

For L526P mutation highlighted ccc sequence was inserted in place of ctc:

L526P+: 5'- ctactggctggctcgcatg**ccc**gagggaggagaggac -3'
L526P-: 5'- gtcctctcctccctc**ggg**catgcgagccagccagtag -3'

For L625P mutation highlighted ccg sequence was inserted in place of ctg:

L625P+: 5'- cgcgcccactagg**ccg**atgaaggatttgg - 3'
L625P-: 5'- ccaaatccttcat**cgg**cctagtgggcgcg – 3

For K633R mutation highlighted aga sequence was inserted in place of aaa:

```
K633R+: 5'- ggatttgggctatggcaggctacaagtacaacc -3'
K633R-: 5'- ggttgtacttgtagcctctgccatagcccaaatcc -3'
```

For K636R mutation highlighted agg sequence was inserted in place of aag:

```
K636R+: 5'- ggctatggcaaaggctacaggtacaaccccatgtac -3'
K636R-: 5'- gtacatggggttgtacctgtagcctttgccatagcc -3'
```

The vector was amplified in the following PCR solution (***Table 4***), where Primer 1 and Primer 2 were the pair of primers designed for each mutation as indicated above:

sample	volume [µl]
Milli-Q water	15.3
10x Pfu buffer (Ultra II Buffer)	2
dNTP (10 mM)	0.5
DNA template (20 ng/µl)	1
Primer 1 (10 µM)	0.4
Primer 2 (10 µM)	0.4
Pfu pol (2.5 U/µl)	0.4

Table 4: PCR solution used for introducing WRNIP1 point mutations

The PCR conditions were as follows:

	2 min – 95°C	(additional denaturation)
	20 s – 95°C	(denaturation)
x 19 cycles	20 s – 55°C	(annealing)
	3 min – 72°C	(extension)
	5 min – 72°C	(additional extension)

After PCR, 1 µl of Dpn1 (enzyme digesting methylated DNA) was added to each reaction, and after an hour all volume (now 21 µl) of the reaction was used to transform competent *Top 10 E. coli* cells, that were selected by plating on LB-agar supplemented with Ampicillin.

The colonies obtained from this process were sent to sequencing to confirm that the desired mutations were introduced, and then the ones that screened positively were transfected into HeLa

Flp-in T-REx cell lines with previously introduced CRISPR gene disruption of the endogenous WRNIP1.

2.2.11 Plasmid amplification using *E. coli Top10* strain followed by Miniprep alkaline lysis

Plates containing the antibiotic of choice were warmed up to 37°C in the incubator (to get rid of the condensation). The competent *Top10 E. coli* cells were taken from -80°C freezer. The sample of DNA was spun down and put on ice. The DNA sample (between 20 and 100 ng) was transported to the tube with the competent cells and the tube was incubated on ice for 30 min.

The sample was put to the 42°C water bath for 45 s, then put back on ice for 2 min (heat shock). 1 ml of LB medium was added to the tube and the sample was incubated at 37°C for an hour. After incubation the sample was spun down on max speed for 1 min, 900 µl of the supernatant was removed, the cells were resuspended in the remaining medium, transported onto the LB-agar plate and left in the 37°C incubator to grow overnight.

The colonies formed after incubation were used to inoculate 3 ml of liquid LB broth medium (one bacterial colony at a time), and left overnight in a shaker (37°C). The LB medium with the bacterial colony was kept in a 14-ml loosely-capped tube (to ensure air circulation). The following day, the cultures were centrifuged at 13 000 RPM speed at 4°C for 1 minute. After centrifugation, the supernatant was aspirated and the pellet was resuspended in 100 µl of ice cold Solution I (50 mM glucose; 25 mM Tris-Cl pH 8.0; 10 mM EDTA pH 8.0, followed by adding, 200 µl of freshly prepared Solution II (200 mM NaOH; 1% SDS). At this stage the tube was closed and the solution was mixed by inverting the tube several times. Finally, 150 µl of ice-cold Solution III (3 M potassium acetate, 11.5 % glacial acetic acid) was added and the tube was closed and vortexed for 10 seconds. After vortexing the tube was stored on ice for 3-5 minutes and then spun down at 13 000 RPM for 5 min (at 4°C). The supernatant was transferred into a fresh 1.5 ml Eppendorf tube.

To achieve better purity of the samples they were subjected to phenol-chloroform extraction at

this stage. First equal volume of phenol: chloroform: isoamyl (1: 1: 1) solution was added to the supernatant and the sample was vortexed. Then, the sample was centrifuged at 13 000 RPM (4°C; 20 min). The aqueous phase was transferred into a new Eppendorf tube and mixed with equal volume of chloroform. The sample was vortexed and spun down (13 000 RPM, 4°C, 1 min) and then the aqueous phase was transferred into a fresh 1.5 ml Eppendorf tube, centrifuged again and the upper (aqueous) layer was transferred into a new tube.

At this stage, the double-stranded DNA was precipitated with two volumes of room temperature ethanol and then incubated in the freezer (-20°C) for 1 h. After the precipitation step was completed the sample was centrifuged (13 000 RPM, 4°C, 20 min). The supernatant was aspirated and the pellet was washed with 1 ml of 70% (volume percentage) of 4°C ethanol. The sample was centrifuged (13 000 RPM, 10 minutes, 4°C), then the supernatant was aspirated, the pellet was dried and dissolved in 50 µl of TE buffer at pH 8.0. Finally, the sample was vortexed, and stored at -20°C.

2.2.12 Mitomycin C clonogenic survival assays

For the Mitomycin C (MMC) survival assays cells of the selected cell lines were split into fresh 9 cm dishes with approximately 50% confluency and left overnight to recover. The next day (if the cells to be used in the experiment were expressing WRNIP1-EGFP under doxycyclin promotor) they were supplemented with dox (0.005 µg/ml for 3 h) and then after incubation passaged, counted and plated into 6-well plates. The number of cell plated depended on the cell line as some cell lines respond to treatment better than others, and was usually around 200 cells for HeLa wild type, 200 cells for HeLa-(WRNIP1 $^{-/-}$), 250 for any HeLa Flp-in T-REx cells, and 1000 for both Hela (FANCD2 $^{-/-}$) and HeLa (FANCD2 $^{-/-}$ and WRNIP1 $^{-/-}$), for the control untreated wells. Then as the concentration of the MMC was raising the number was adjusted so it was still possible to count the colonies even with higher concentrations of the drug.

For each independent experiment three technical repeats were performed, and then each experiment was repeated at least 2 times to confirm the obtained data.

The cells were then left overnight and treated with raising concentrations of MMC on the next day. The drug was left in the medium for the whole incubation period (10 -14 days to allow the formation of the colonies), and after that period of time the plates were recovered from the incubator and stained as follows: first, the medium was aspirated and all wells were washed with room-temperature PBS (no EDTA), then PBS was aspirated and the colonies were fixed for 10 min with the fixing solution (10% methanol + 10 % acetic acid) that was then aspirated to allow staining all wells with crystal violet solution for 15 min (10 g of crystal violet diluted in 1 l of methanol). After staining the plates were carefully washed with tap water and dried in the oven.

The stained colonies were then counted. For statistical analysis, each technical repeat was translated into percentage values of cells that survived the treatment in comparison to the untreated well, so we could have for example three repeats with following values:

I: 100%, 80%, 60%, 40%, 10%

II: 100%, 75%, 55%, 33%, 12%

III: 100%, 82 %, 66%, 45%, 11%.

Then we would count the average value for each data point so in this case for the second point it would be for example: (80%+75%+82%)/3=79%, which would be the data point for the second used concentration.

After counting the percentage of cells that survived at given MMC concentration, we combined the obtained data from two or more independent experiments, and the standard error in each case was calculated as standard deviation of the average values for each point in each independent experiment divided by a square root of the number of all the independent experiments combined.

2.2.13 TMP-UVA clonogenic survival assays

For TMP-UVA assay the cells were pre-treated and plated according to the same principles

as in case of the MMC assay. Each cell line was also plated in three technical repeats.

The next day TMP solution (Sigma, T6137) was added to all wells to the final concentration of 50 ng/ml (or 100 ng/ml in indicated cases) and the cells were incubated for 30 min in 37°C incubator. After this period of incubation, the cells were irradiated with indicated doses of UVA light (at the 365 nm wavelength) using the UVA XL-1500 UV CROSSLINKER, and then left for 10-14 days to form the colonies.

When the colonies were formed the cells were fixed, stained and counted using identical protocol as in case of the MMC clonogenic assay.

2.2.14 UVC clonogenic survival assays

For UVC clonogenic survival assay the cells were pre-treated and plated according to the same principles as in case of the MMC assay. Each cell line was plated in three repeats.

The next day after the cells attached to the bottom of each well, the plates were irradiated with indicated doses of UVC light, using the UVC XL-1500 UV CROSSLINKER.

After the irradiation the cells were the cells were left in the incubator (37°C, 5% CO_2) for 12 days to allow the formation of the colonies.

After the colonies formed the cells were fixed, stained and counted according to identical protocol as in case of the MMC survival assay.

2.2.15 Hydroxyurea clonogenic survival assays

For HU clonogenic survival assay the cells were pre-treated and plated according to the same principles as in case of the MMC assay.

The next day, after the cells attached themselves to the bottom of the plates, HU was added to each well at indicated concentration, and the cells were left in the incubator (37°C, 5% CO_2) for 12 days to allow the formation of the colonies. Similarly, as in case of MMC treatment the drug was

kept in the medium throughout the whole incubation period (the acute treatment protocol was not tested).

After the colonies formed the cells were fixed, stained and counted according to identical protocol as in case of the MMC survival assay.

2.2.16 Purification of FLAG-HA-WRNIP1 from HeLa S3 cells

S3 cells expressing Flag-HA-WRNIP1 were split into three 15 cm dishes with approximately 70% confluency. All three dishes were left in 37°C incubator overnight.

The next day the cells were harvested (the cells were detached with a spatula and the medium was spun on 1000 RPM for 5 min, the pellet was resuspended in 1 ml room-temperature PBS (no EDTA), transported to eppendorf tube, and spun on 4000 RPM for 2 min, supernatant was aspirated). At this stage the volume of the pellet was close to 250 µl**.** The pellet was resuspended in 250 µl of Buffer A (***Table 5***) and left on ice for 10 min:

INGREDIENT	CONCENTRATION	FUNCTION
Tris pH 8.0	20 mM	pH buffer
Triton X-100	0.1-0.5%	non-ionic detergent to permeabilize the cell membrane
Glycerol	10%	to prevent protein precipitation
Benzonase	50 Unit/ml	nuclease
β-mercaptoethanol	2 mM	reducing agent
PMSF (phenylmethylsulfonyl fluoride)	0.2 mM	protease inhibitor

Table 5: *Buffer A for protein purification*

2.5 ml (10x pellet volume) of Buffer B (***Table 6***) was added to the solution, and the falcon was incubated on ice for 10 min.

INGREDIENT	CONCENTRATION	FUNCTION
Tris pH 8.0	20 mM	pH buffer
Triton X-100	0.1-0.5%	non-ionic detergent to permeabilize the cell membrane
Glycerol	10%	to prevent protein precipitation
KCl	200 mM	to reduce background binding and stabilise the protein
β-Me	2 mM	reducing agent
PMSF	0.2 mM	protease inhibitor

Table 6: Buffer B for protein purification

The solution was spun down on 18 000 RPM for 10 min at 4°C. The supernatant was transferred to new tubes and spun again (18 000 RPM, 10 min, 4°C), then it was collected, 100 µl of M2 beads (α-FLAG) slurry (Sigma-Aldrich) was added, and the falcon was incubated for 2 h on the rotor (4°C). The lysate/beads slurry was transferred to column (10 ml volume, equilibrated with 1 ml of Buffer B beforehand) and washed with 5 ml of Buffer B.

The column was closed, and 50 µl of the FLAG peptide solution was added (0.5 µg/µl FLAG peptide in buffer B). The lysate/beads slurry was incubated with FLAG peptide in the cold room for 1.5 h (mixed every 15 min). The eluate was collected and the elution step was repeated (the same volume of the FLAG peptide solution was added, and incubated for 1.5 h), the eluate from the second elution was collected to the same tube as eluate from the first elution.

The samples taken during the procedure were then analysed with SDS-PAGE gel followed by Coomassie blue staining and Western Blot (α-WRNIP1, α-HA and α-ubiquitin (FK2)), assays.

After confirming that the purification was successful in order to further purify the protein, CCl_3COOH (**tri**chloroacetic **a**cid= TCA) precipitation was performed. First one volume (75 µl) of 20% TCA was added to 1 volume of eluate. The sample was incubated on ice for 4 h, and then spun down (max speed, 30 min, 4°C). Supernatant was removed, and the pellet was washed with 800 µl ice cold acetone. The sample was spun down again (max speed, 10 min, 4°C). The pellet was dried on air for 10 min and resuspended in 10 µl of the LDS loading buffer.

Then 9% polyacrylamide-SDS gel was used for electrophoresis, all 10 µl of the sample was loaded next to 10 µl of the Mark12 protein marker, the proteins were separated in the process of electrophoresis, after which the gel was removed from its covers and placed in the Coomassie Blue stain, on a shaker (room temperature, overnight). The next day the Coomassie Blue staining buffer was exchanged to 45% methanol/10% acetic acid destaining buffer, and after 3 h to 5% methanol / 7.5% acetic acid destaining buffer in which the gel was left overnight.

After destaining the upper form of WRNIP1 was carefully cut out of the gel using scalpel and sent to the mass spec facility (Dunn School of University of Oxford).

2.2.17 Obtaining the constructs for *E. coli* expression of the His-tagged WRNIP1 fragments

The five short His-tagged WRNIP1 fragments were designed by simply dividing WRNIP1 sequence into five parts, in such a way that each domain remains as a whole, and that all the pieces are more or less comparable in size so they can all be tested using one SDS-PAGE polyacrylamide gel. Additionally, there was always one amino acid overlapping to ensure that all the sequence is covered. The first fragment was equivalent to the first 120 aa of WRNIP1, the second was covering amino acids 120-240, the third one 240- 370 aa, the fourth one 370-510 aa and finally the fifth one 510 – 665. All the fragments had NdeI and EcoRI restriction sites introduced at the ends to allow ligation to the backbone plasmid later (NdeI and EcoRI were a very good choice, as they do not have restriction sites inside WRNIP1 sequence).

The fragments were amplified from the previously obtained from Alicja Bulsiewicz (Dr Cohn's group) pFB-WRNIP1 construct using the following PCR primers (***Table 7***):

No.	Primer Name	Sequence (5'→3')
1	WRNIP1_1_120+	ATCGCATATGATGGAGGTGAGCGGGCCG
2	WRNIP1_1_120-	ATCGGAATTCTCAGGCGCCGCTGGGTGTG
3	WRNIP1_120_240+	ATCGCATATGGCCCGCCTTATCCCCGAC
4	WRNIP1_120_240-	ATCGGAATTCTCAGAAGTAATCCTGCAGCGTGTC
5	WRNIP1_240_370+	ATCGCATATGTTCGGGCAGAGCAAGGCC

6	WRNIP1_240_370-	ATCGGAATTCTCAGCTCAGAAGAGCAGCGTTGAC
7	WRNIP1_370_510+	ATCGCATATGAGCCGCTGTCGAGTGATTG
8	WRNIP1_370_510-	ATCGGAATTCTCACATGGACTTGTGCAGGGCG
9	WRNIP1_510_665+	ATCGCATATGATGCGGGGCTCAGACCAG
10	WRNIP1_510_665-	ATCGGAATTCTCAGCACCTCCTCTGCTTGAAG

Table 7: Primers used for amplifying WRNIP1 fragments needed to obtain His-tagged constructs

The PCR solution was identical to the solution described in Table 2, and the PCR conditions used were almost the same with the only difference being the shortened extension time as the fragments expressed this time were not as long as full length WRNIP1.

After the PCR reaction, the fragments were purified using the *QIAquick PCR Purification Kit* (QIAGEN), and the sizes of the fragments were checked using agarose gel electrophoresis. After confirming that we obtained DNA particles of the right sizes, they were all cut with NdeI and EcoRI enzymes (37℃, 1 h).

At the same time 2 µg of the backbone plasmid pET28-c was also cut by NdeI and EcoRI by incubating on 37℃ for 1 h. After the incubation was over the restriction enzymes were deactivated (65℃, 20 min) and the backbone plasmid (but not the five fragments) was dephosphorylated (CIP enzyme, 37℃, 20 min) followed with deactivation of the CIP (37℃, 1 h).

Later all of the digested DNA was separated via electrophoresis on the agarose gel, cut out with the clean scalpel and purified from the gel using the *QIAquick* gel extraction Kit. The concentration of the purified DNA was measured by testing 10% of the obtained sample on the agarose gel again next to the DNA marker with known amount of the DNA, and then five ligation reactions (each of the fragments with the backbone plasmid) were prepared using the T4 ligase in the provided T4 buffer (16℃, overnight).

After the ligation reaction was over the obtained solution was transformed into *E. coli Top 10* cells according to the standard protocol, and plated on the Kanamycin (as this is the resistance encoded on the pET28-c plasmid) supplemented agar-LB plates.

Later the formed colonies were used to obtain the plasmid samples using the standard

Miniprep protocol, and then tested with sequencing to obtain that they all have desired sequence.

2.2.18 Expression and purification of the His-tagged WRNIP1 fragments

The five plasmids obtained as described in chapter 2.2.17, were used to transform *E. coli* BL21 (λDE3) cells (identical protocol as in case of *E. coli Top 10*, just with different strain of bacteria), and once the colonies were formed, they were used to inoculate 3 ml of the liquid LB (+0.1% glucose) bacterial culture, that was incubated with vigorous shaking overnight at 37°C. The following day the overnight culture was used to inoculate larger scale LB (+0.1% glucose) culture (0.5 l) that was grown to 0.9 OD (optical density measured at 600 nm) at 37°C with shaking, and then IPTG (Isopropyl β-D-1-thiogalactopyranoside) was added, to induce the expression of the WRNIP1 fragments. Each of the fragments was expressed using 0.1 mM of IPTG, but then some of them had to be expressed under different conditions than others, to ensure optimal expression. Fragment 1 (1-120 aa) was expressed at 25°C overnight, fragment 2 (120-240 aa) at 30°C for 4 h, fragment 3 (240-370 aa) at 16°C overnight, fragment 4 (370-510 aa) at 37°C for 2 hours, and finally fragment 5 (510-665 aa) at 16°C overnight. Each of the conditions had to be optimized separately as these protein fragments are not bacterial proteins and are undergoing degradation if expressed under wrong conditions.

After the proteins were expressed the cells were spun down (6000 RPM, 30 min, 4°C), resuspended in lysis buffer (150 mM KCl; 10% glycerol; 20 mM Tris-HCl pH 8.0; 0.1% Tween-20, 0.2 mM PMSF; 2 mM β-mercapthoethanol; 20 mM imidazole), incubated on ice for 30 min, and then sonicated (10 min, 1 s on/off, 10 W) to fully lyse the cells.

After the cells were lysed the solution was centrifuged (10 min, 4°C, 13 000 RPM), and the supernatant was transferred into the fresh falcon tube.

The Ni^{2+}-NTA beads (Qiagen) were washed on the 15 ml column first with water and then with the buffer used for cell lysis, then the column was closed, the beads were resuspended in equal

volume of the buffer and 150 µl of the slurry was added to each of the five lysates.

The samples were incubated with beads for 2 h on the rotor (at 4°C) and then placed in the column and washed with previously used lysis buffer. After the beads were washed, the column was closed and the proteins were eluted using 75 µl of the elution buffer (150 mM KCl; 10 % glycerol; 20 mM Tris-HCl pH 8.0; 0.1% Tween-20, 0.2 mM PMSF; 2 mM β-mercapthoethanol; 250 mM imidazole).

The purified proteins were later tested by electrophoresis on the 15% SDS-PAGE polyacrylamide gel, and then visualized both using Coomassie Blue stain and WB analysis.

2.2.19 WRNIP1-FANCD2 pulldown assays

Flag-HA-FANCD2 in Flag peptide eluate and either Strep(II)-WRNIP1 (wild type or deletion mutant in desthiobiotin eluate) or one of the five His-tagged WRNIP1 fragments were used in this experiment. Flag-HA-FANCD2 and Strep(II)-WRNIP1 variants were all purified, analysed and provided by Alicja Bulsiewicz (Dr Cohn's group). 0.5 µg of WRNIP1 (or WRNIP1 deletion mutants/fragments) and FANCD2 were combined in the Eppendorf tube. FANCD2, WRNIP1, 1 µl BSA (so the final concentration was 100 µg/ml) and binding buffer (150 mM KCl, 20 mM Tris-HCl pH 8.0, 0.2 mM PMSF, 2 mM β-Mercapthoethanol) were added to the final volume of 10 µl.

The sample was mixed by vortexing and spun quickly. 1.5 µl of the input sample was taken from from each Eppendorf for WB analysis. The sample was incubated at 30°C for 1 h. After the incubation the binding buffer was supplemented with 0.1% of Tween-20 (v/v). α-HA sepharose beads were washed two times in the Tween-20 supplemented binding buffer (resuspeded and spun down at 2000 rpm for 5 min). 10 µl of α-HA sepharose bead slurry was added to each Eppendorf tube. The samples were incubated for half an hour on the rotor in the cold room. Afterwards the samples were transported into the 3 ml purification columns and washed with 3 ml of the binding buffer in total. After the washing was finished, 30 µl of glycine buffer (0.1 M glycine; pH 2.5) was added to the column. The column was centrifuged at 1000 rpm for 2 min. Afterwards the pH was neutralized with Tris pH 8.0 buffer, LDS

loading buffer was added to the samples, they were vortexed, heated at 70°C for 10 min and analysed by WB with α-WRNIP1, α-FANCD2 or α-His antibodies accordingly.

2.2.20 Sequencing

All the sequencing described in this study was performed externally by Source BioScience.

2.2.21 Statistical analysis

For each statistical analysis the t-Student test was used. In each case the average of values obtained in each independent experiment was calculated and then the standard deviation of these values was calculated. To translate the standard deviation into standard error, it was divided by the square root of the number of all independent experiments (so usually the square root of 2 or 3). To use the t-Student test, the standard error was multiplied by the t value obtained from the Student two -tailed t-Table, for p=0.05.

3 Results

3.1 Depletion of WRNIP1 affects the cell survival after treatment with ICL inducing agents

3.1.1 Initial theoretical analysis

WRNIP1 seemed like a good candidate to be involved in an ICL repair. It was detected previously among proteins that interact with ubiquitinated FANCD2. The proteins that bind to monoubiquitinated FANCD2 were identified using mass spectrometry. In order to prepare samples for mass spectrometry assay, the FLAG-tagged FANCD2 was expressed in HeLa cells (Chih-Chao Liang, Dr Martin Cohn's laboratory) and the cells were treated with MMC. In the next step the FLAG-tagged FANCD2 was purified using α-FLAG M2 agarose chromatography. The immunoprecipitate was analysed by mass spectrometry as described by ***Liang et al. (2015)*** resulting in 21 peptides from WRNIP1 protein (preliminary data obtained previously by Chih-Chao Liang, dr Cohn's Laboratory).

Additionally, WRNIP1 has also an interesting domain architecture and sequence similarities to proteins associated with DNA repair such as Rad18 (E3 ubiquitin-protein ligase that is, among other functions, involved in postreplication repair of the UV-damaged DNA and is also known for guiding Polη to replication stalling sites through physical interaction and PCNA monoubiquitination (***Watanabe et al., 2004***). WRNIP1 also shares a functional domain with polymerase κ, and polymerase η, both polymerases specifically involved in DNA repair. The similarities are between the UBZ ubiquitin-binding-zinc finger) domains of WRNIP1 and the mentioned proteins. In case of WRNIP1 the UBZ domain is located on its N terminus.

There are four sub-types of the UBZ domains (***Table 8***). The majority of all the UBZ domains identified so far belong to either the C2H2 (type 3 Polη-like) or C2HC (type 4 Rad18-like) family. The UBZ domain present in WRNIP1 belongs to the UBZ4 type, meaning that it is most

similar to the one that can be found in Rad18.

PROTEIN	TYPE	SEQUENCE
TAX1BP1 (1)	UBZ1	722----SSFDVHKK**C**PL**C**EL---MFPPNYDQSKFEE**H**V-ES**H**W----------754
TAX1BP1 (2)	UBZ1	755----------KV**C**PM**C**SE---QFPPDYDQQVFER**H**V-QT**H**FDQNVLNFD--789
FAAP20	UBZ2	139----EGAAALRS**C**PM**C**QK---EFAPRLTQLDVDS**H**L-AQ**C**LAESTEDVTW-180
Pol η	UBZ3	628-----AAEDQVP**C**EK**C**GS---LVPV----WDMPE**H**M-DY**H**FALELQKS---662
Pol κ	UBZ4	620--------QILT**C**PV**C**FRAQGCISL----EALNK**H**V-DE**C**LDGPSIS----653
Rad18	UBZ4	200--------TKVD**C**PV**C**GV---NIPE----SHINK**H**L-DS**C**LSREEKK----230
WRNIP1	UBZ4	9-DPFLSQLHQVQ**C**PV**C**QQ---MMPA----AHINS**H**L-DR**C**LLLHPAGHA--48
FAN1	UBZ4	41---------KLA**C**PV**C**SK---MVPR----YDLNR**H**LDEM**C**ANN--------67
SLX4 (1)	UBZ4	286--DSLEEKGLFF**C**QI**C**QKNLSAMNV----TRREQ**H**V-NR**C**LDEAEKT----325
SLX4 (2)	UBZ4	326--LRPSVPQIPE**C**PI**C**GK---PFLTL---KSRTS**H**L-KQ**C**AVKMEVGPQL-366

Table 8: Examples of known UBZ domains. For proteins with two tandem UBZ domains (TAX1BP1 and SLX4) the one on the N-side is indicated with (1) and the one on the C-side with (2)

The fact that UBZ domain is present in WRNIP1 would make a lot of sense if the protein was discovered to be involved in the cellular DNA damage response since ubiquitin signalling is critical for it. Ubiquitin signalling is also important in the Fanconi anemia pathway so it seemed plausible that WRNIP1 is somehow associated in the ICL repair process as either crucial part of the pathway or one of the regulatory proteins. Using the publicly accessible protein model and Pymol software, we identified potential polar interaction sites between the WRNIP1 UBZ domain and the ubiquitin molecule.

Pymol is an open-source molecular visualization system based on the Python programming language. It is available for free for non-commercial scientific purposes. It can both render and animate three dimensional crystal protein structures which makes it fairly easy to find the residues of interest. In this particular case, the model of the UBZ domain of WRNIP1 in complex with ubiquitin was already accessible in the RCSB PDB database (listed as “Crystal structure of GFP-Wrnip1 UBZ domain fusion protein in complex with ubiquitin” - code name: 3VHT) (***Suzuki et al., 2016***). The model was downloaded from the database, and then the WRNIP1 UBZ fragment and ubiquitin particle were set to visible mode, while the GFP tag attached to the UBZ domain was

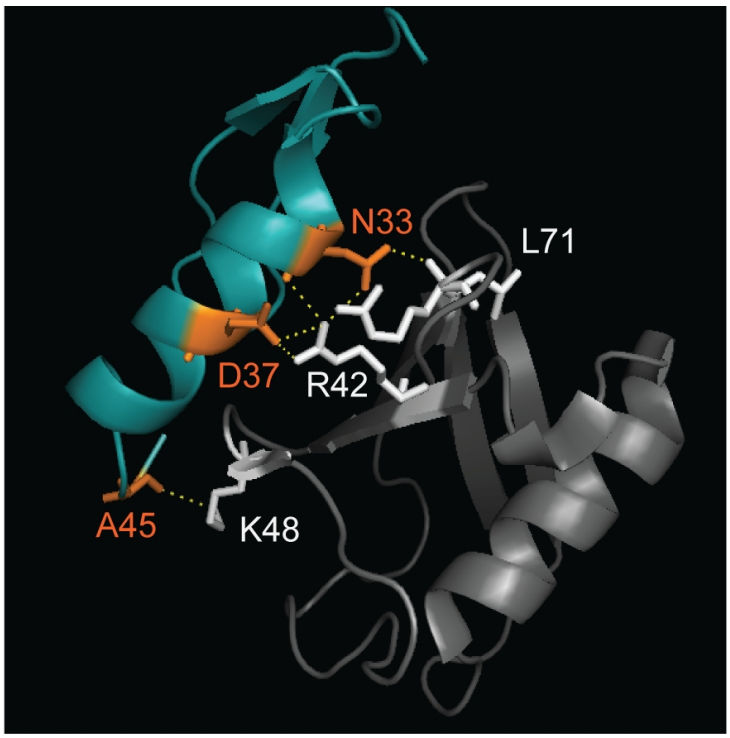
hidden to not distract from the important part of the model. Then the software was set to display all polar interactions between the residues. Of course, as a result of this display mode not only interactions between ubiquitin and UBZ were highlighted but also all interactions within the UBZ domain for example between its different residues. At this point all the interactions that were not connecting the UBZ with ubiquitin were hidden from the model (**Figure 6A**).

There were three identified inter-molecule polar interaction residues. One of them was D37 interacting with R42 residue in the ubiquitin sequence. D37 was previously mentioned in literature (***Crosetto et al., 2008***) as a residue necessary for WRNIP1 recruitment to the DNA damage site. Other two were N33 (interacting with L71 in the ubiquitin sequence) and A45 (interacting with K48 in the ubiquitin sequence). WRNIP1 UBZ domain was aligned with UBZ domains present in other proteins using the Clustal Omega tool (accessible online: http://www.ebi.ac.uk/Tools/msa/clustalo/) (***Table 8***).

It is a bioinformatic multiple protein alignment algorithm that uses seeded guide trees and HMM (Hidden Markov Model) profile-profile techniques to generate alignments. HMM algorithms allow the alignment of the sequences derived from each other with the assumption that between the visible stages there are also invisible temporary stages. The visibility of the final stage allows the prediction of stages that led to it. This tool allows accurate prediction of which sequences had the lowest number of temporary stages (where the Markov chain is the shortest) which corresponds to the sequences which are the most evolutionarily similar to each other. It calculates the best match for the selected sequences, and lines them up so that the similarities and differences can be seen.

According to the performed analysis both D37 and N33 seemed to be conserved among the UBZ domains that were the most similar to the one observed in WRNIP1, while A45 seemed to be specific only to WRNIP1 (**Figure 6B**).

A.



B.

```
                                                    33  37      45
WRNIP1          9 DPFLSQLHQVQCPVC---QQMMPAAHINSHLDRCLLLHPAGHA- 48
Rad18         200 -------TKVDCPVC---GVNIPESHINKHLDSCLSREEKK--- 230
Pol_Kappa     620 -------QILTCPVCFRAQGCISLEALNKHVDECLDGPSIS--- 653
Pol_Eta       628 ----AAEDQVPCEKC---GSLVPVWDMPEHMDYHFALELQKS-- 662
```

Figure 6: There are three potential ubiquitin interaction sites in WRNIP1 UBZ domain. **A)** Pymol molecular visualization software predicted that residues N33, D37 and A45 (orange) of WRNIP1 UBZ domain (blue) interact with L71, R42 and K48 (light grey) residues on ubiquitin particle (dark grey). The model that was used for this visualization is publicly accessible in RCSB PDB database as "Crystal structure of GFP-Wrnip1 UBZ domain fusion protein in complex with ubiquitin". **B)** WRNIP1 aligned with proteins that have closest homology to it such as RAD18, Pol κ, and Pol η show that while both N33 and D37 are well conserved, A45 is specific to WRNIP1 only.

3.1.2 WRNIP1 depletion results in increased sensitivity to MMC and TMP/UVA treatment but not UVC and HU

While the bioinformatic analysis seemed to provide quite promising reasons to believe that WRNIP1 may be involved in some form of DNA repair, to either reject or confirm the hypothesis that it may be an ICL damage repair protein, simple clonogenic survival assay experiments were performed. Initially WRNIP1 $^{-/-}$ (CRISPR gene disruption) cell line was obtained and then WRNIP1 expression was complemented with exogenous WRNIP1-EGFP protein expressed under doxycyclin promotor (**Figure 7**).

In the survival experiment, all three cell lines (wild type cells, cells with WRNIP1 gene disruption and cells where the gene disruption was complimented with exogenous WRNIP1-EGFP) were used. Having all three cell lines allows to compare the phenotype of the cells that are WRNIP1-deficient, with wild type cells, while at the same time comparison to the cell line with rescued WRNIP1-EGFP expression allows to exclude the possibility that any other cellular event that took place during obtaining the CRISPR/Cas9 cell line was responsible for the change.

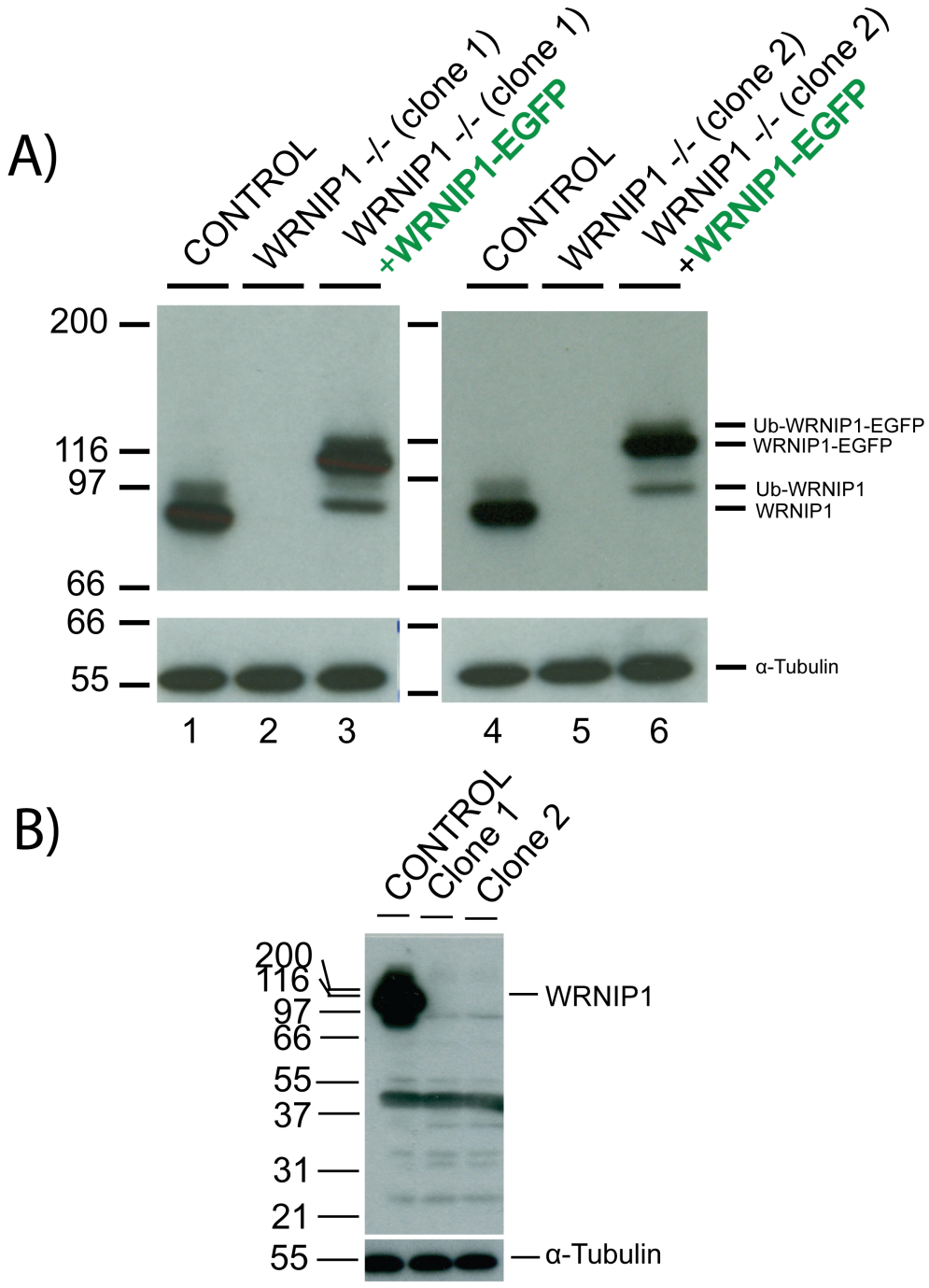


Figure 7: Western Blot analysis of the WRNIP1 deficient cell lines complemented with exogenous WRNIP1-EGFP. A) Standard HeLa Flp-in T-REx cells (CONTROL) were modified with CRISPR/Cas9 genome editing technology to abrogate WRNIP1 expression, and then two of the obtained clones were independently complemented with exogenous WRNIP1-EGFP. To rule out the clonal effect, both sets of cells (clone 1 and its complemented variant as well as clone 2 and its complemented variant) will be used to analyse the effects of WRNIP1 depletion. WRNIP1-EGFP expressing cells were pretreated with 50 ng/ml dox for 3 h before harvesting. 9% SDS-PAGE gel was used during WB analysis **B**) To make sure that there are no truncated WRNIP1 variants, and no larger fusion proteins the same samples were tested with higher percentage gel, and this time the full length of the gel was analysed (15% polyacrylamide SDS gel). Some minor truncated proteins are observed, but their amount is very low in comparison to wild type WRNIP1.

The number of the cells was known (the cells were counted for each cell line) and adjusted in such way as to allow forming colonies. Then the cells were allowed to attach themselves to the bottom of each well and recover after splitting for 24 hours and then they were treated with raising concentrations of MMC. After two weeks, when the colonies were formed it was possible to count which fraction of the initially plated cells survived the treatment and given this number it was possible to estimate the survival curves for each cell line (see materials and methods for statistical analysis). The plotted graphs are the result of a combination of at least two independent experiments (where n is the number of independent experiments).

Two independently obtained WRNIP1 -deficient CRISPR/Cas9 cell lines were tested to eliminate the possibility of the clonal effect, and also the complementation with WRNIP1-EGFP was done in separate procedures. In both cases the phenotype of WRNIP1 deficient cells was that they were slightly more sensitive to the MMC treatment, while complementation of the phenotype with exogenous WRNIP1-EGFP was bringing the phenotype the normal (wild type) level of sensitivity (**Figure 8**).

There is a minor difference between the two clones in the general survival of all the cell lines tested (for second clone all the curves are steeper), which was most likely caused by different potency of the MMC stock solution used (between testing the two cell lines new MMC stock solution was prepared, and most likely the second solution was containing slightly more active drug), since the HeLa control cell line used in both cases is the same, and is still presenting lower survival rate during the experiments corresponding to the second clone. The general tendency though was identical in both cases giving a clear drop in the WRNIP1- deficient cell line survival and equally statistically significant rescue of the phenotype in the complemented cell line.

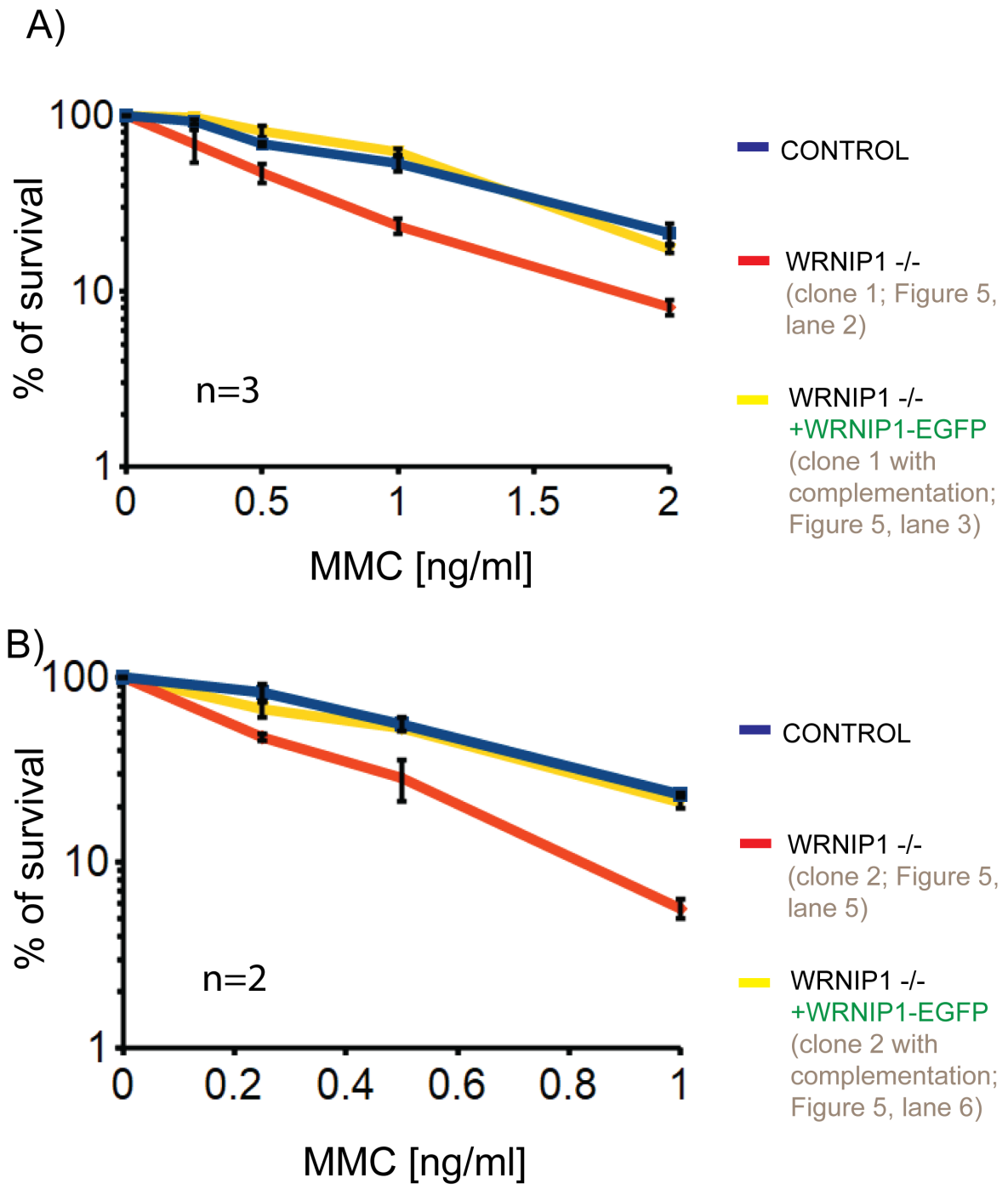


Figure 8: WRNIP1 deficient cell line (derived from HeLa Flp-in T-REx cells) is more sensitive to MMC treatment than wild type cell line. The cells were treated with indicated doses of MMC for 2 weeks and then the survival rates were calculated. Complementing the WRNIP1 deficient cell line with WRNIP1-EGFP brings the sensitivity back to the wild type levels. **A)** Clonogenic survival assay performed using the cell lines corresponding to Figure 7 lanes 1-3, n=3 **B)** Clonogenic survival assay performed using the cell lines corresponding to Figure 7 lanes 4-5, n=2. For both survival curves we calculated the t-Student confidence limits (p=0.05), corresponding to the last MMC dose (2 and 1 ng/ml respectively). In both cases, there was a significant difference between the survival of the WRNIP1 deficient cell line, and both wild type and complemented cell line, while there was no significant difference between the wild type and complemented cell line.

To see if WRNIP1 is involved in repair of other types of ICL damage, similar experiments were performed with two cell lines (wild type and WRNIP1 deficient one) and different drugs.

Since stable complementation with exogenous WRNIP1-EGFP required using HeLa Flp-in T-REx cell line, which is generally difficult to work with, as the cells are forming smaller colonies than regular HeLa cells, and are also more sensitive, the remaining clonogenic survival assays presented in this chapter were performed without the complemented cell line, using regular HeLa cells instead. First two HeLa (WRNIP1 $^{-/-}$) cell lines were obtained independently, using the CRISPR/Cas9 genome editing technique. Lack of WRNIP1 was confirmed using WB assays. Then HeLa (FANCD2 $^{-/-}$) cell line was obtained (kindly provided by David Lopez Martinez, Dr Cohn's group) and from this cell line, two more WRNIP1 $^{-/-}$ mutants were obtained, using the same CRISPR/Cas9 technique as before.

As a result, we had one regular HeLa cell line (from which all the other cell lines were derived), two independently obtained HeLa (WRNIP1 $^{-/-}$) cell lines, one HeLa (FANCD2 $^{-/-}$) cell line (obtained by David Lopez Martinez (Dr Cohn's group) from the same initial HeLa cell line, and two double depletion cell lines, where a WRNIP1 depletion was introduced on top of the one for FANCD2 (**Figure 9**).

Regular HeLa cells and HeLa (WRNIP1 $^{-/-}$) cells were used to analyse the effect of WRNIP1 depletion on cell sensitivity to TMP/UVA treatment, UVC treatment, and HU treatment (described below), while the HeLa FANCD2 $^{-/-}$ cell line and the HeLa (FANCD2 $^{-/-}$, WRNIP $^{-/-}$) cell lines were used for the epistasis experiments (both with MMC and TMP/UVA damage) as described in the next chapter.

To test the effects of other DNA damaging agents on WRNIP1 depleted cells first, tests were performed using the combination of the psoralen sensitization (in this particular case trioxsalen (TMP) was used) and UVA irradiation, and as predicted the WRNIP1 deficient cell line was more sensitive to the damage (**Figure 8**). Interestingly neither HU nor UVC treatment has resulted in any

statistically significant difference of sensitivity between the two tested cell lines (**Figure 9 A and B**).

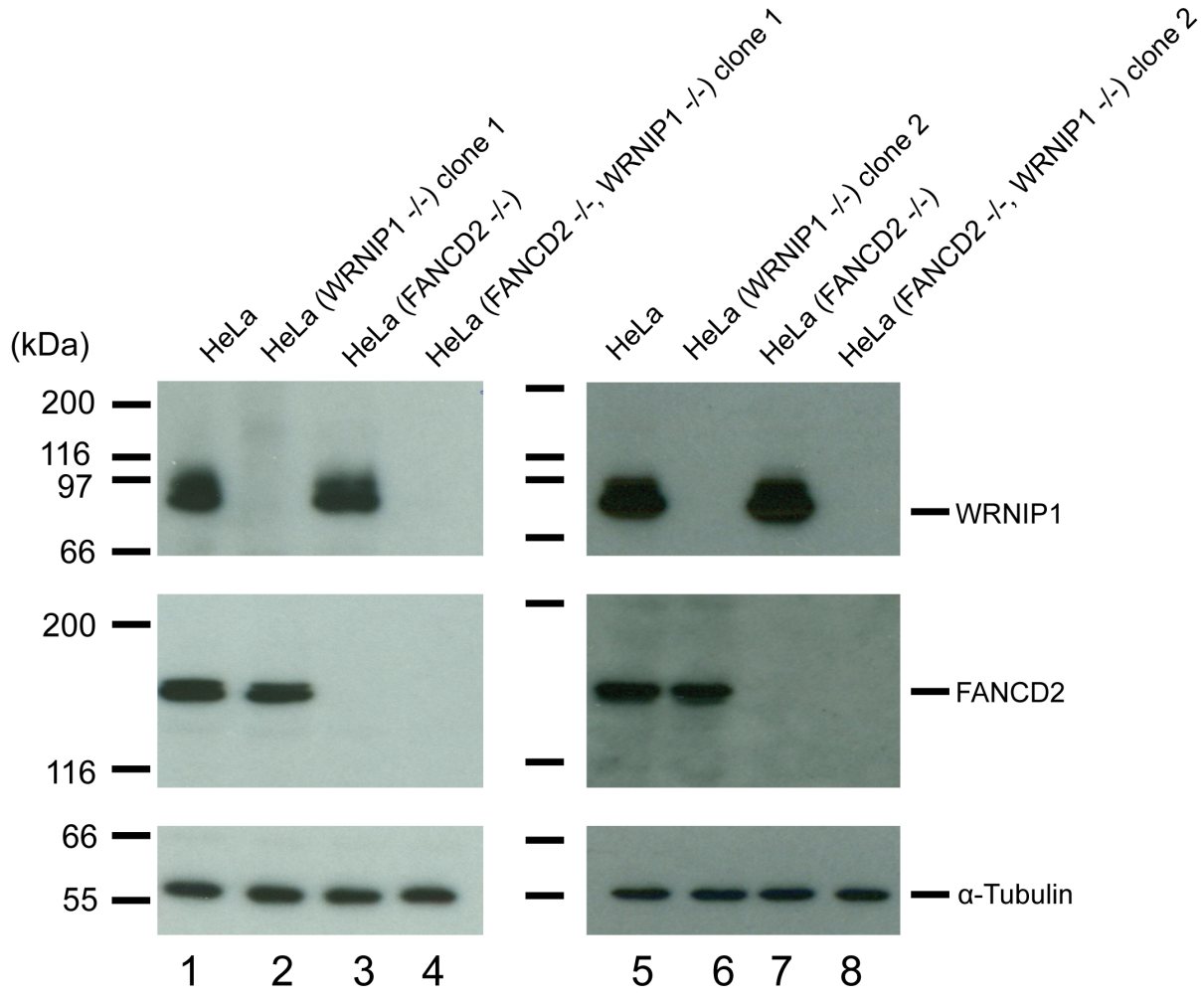


Figure 9: Western Blot analysis of the cells used for clonogenic survival assays. The presented HeLa derived cell lines were used for TMP-UVA, UVC and HU survival assays (cell lines corresponding to the lanes 1 and 2, and also each experiment was confirmed with the WRNIP1 gene disruption cell line, corresponding to the lane 6, to rule out the clonal effect). All of the presented cell lines were used for the epistasis experiments (with both MMC and TMP-UVA), as described in the next chapter. Lanes 1 and 5 are corresponding to the same cell line (the regular HeLa cells), as well as lanes 3 and 7 (FANCD2 gene disrupted cell line provided by David Lopez Martinez, Dr Cohn's group). The cells were lysed and the obtained lysate was tested in WB assay against anti-WRNIP1, anti-FANCD2 and anti-tubulin, antibodies. The concentration of the polyacrylamide in the SDS gel was 9% for the WRNIP1 and tubulin assay and 6% for the FANCD2 assay.

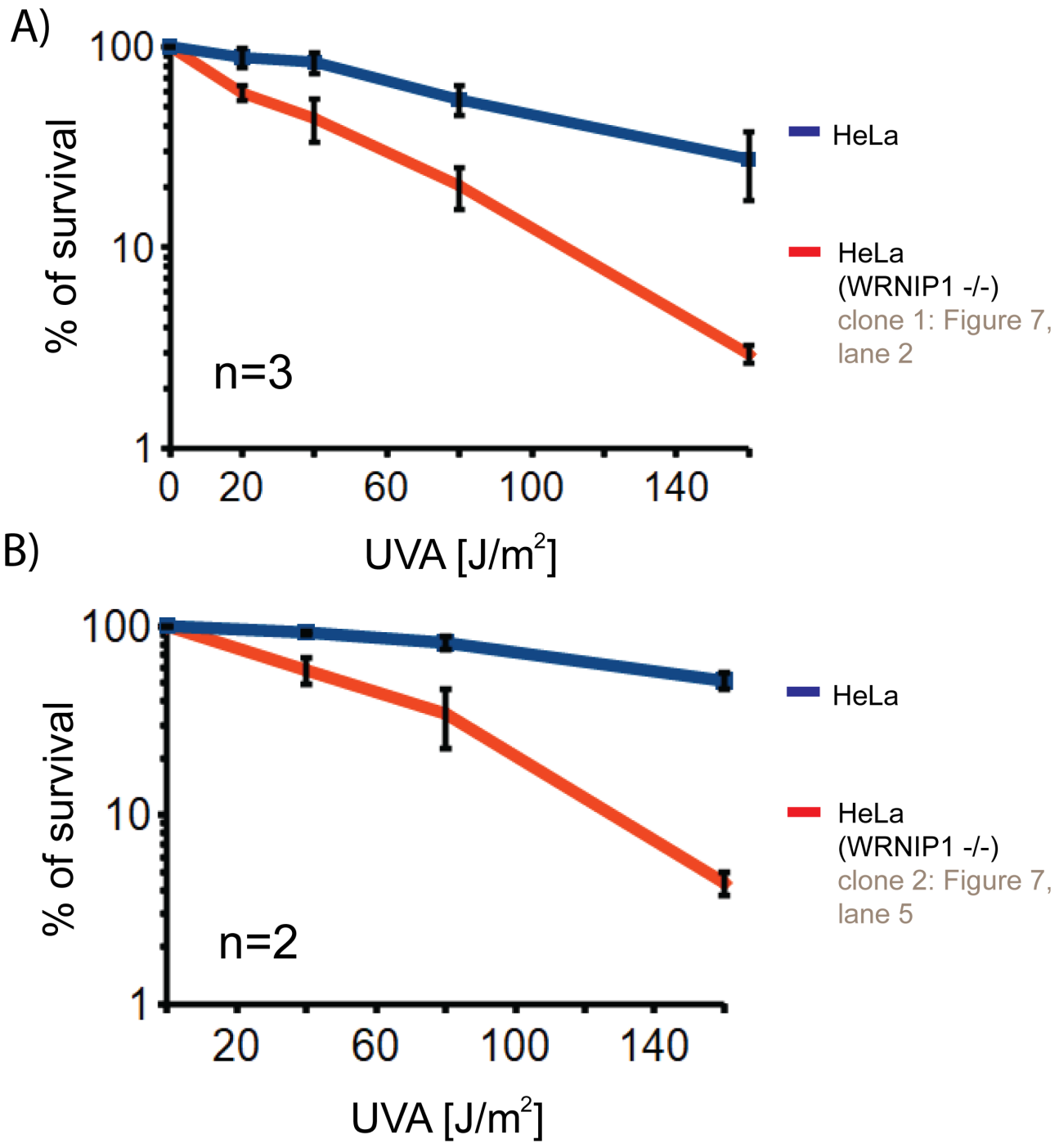


Figure 10: WRNIP1 deficient cell line is more sensitive to the TMP-UVA treatment than wild type HeLa cell line. Both cell lines were pre-treated with TMP and irradiated with indicated doses of UVA light as described in chapter "Materials and Methods". Then they were left to form the colonies for 10-14 days. **A)** Clone 1 was tested along with control HeLa cells (lane 2; Figure 9). Three independent experiments were combined together (n=3) **B)** Clone 2 was tested along with control HeLa cells (lane 6; Figure 9). Two independent experiments were combined together (n=2). For both survival curves we calculated the t-Student confidence limits (p=0.05), corresponding to the last UVA dose. In both cases, there was a significant difference between the survival of the WRNIP1 deficient cell line, and the wild type cell line.

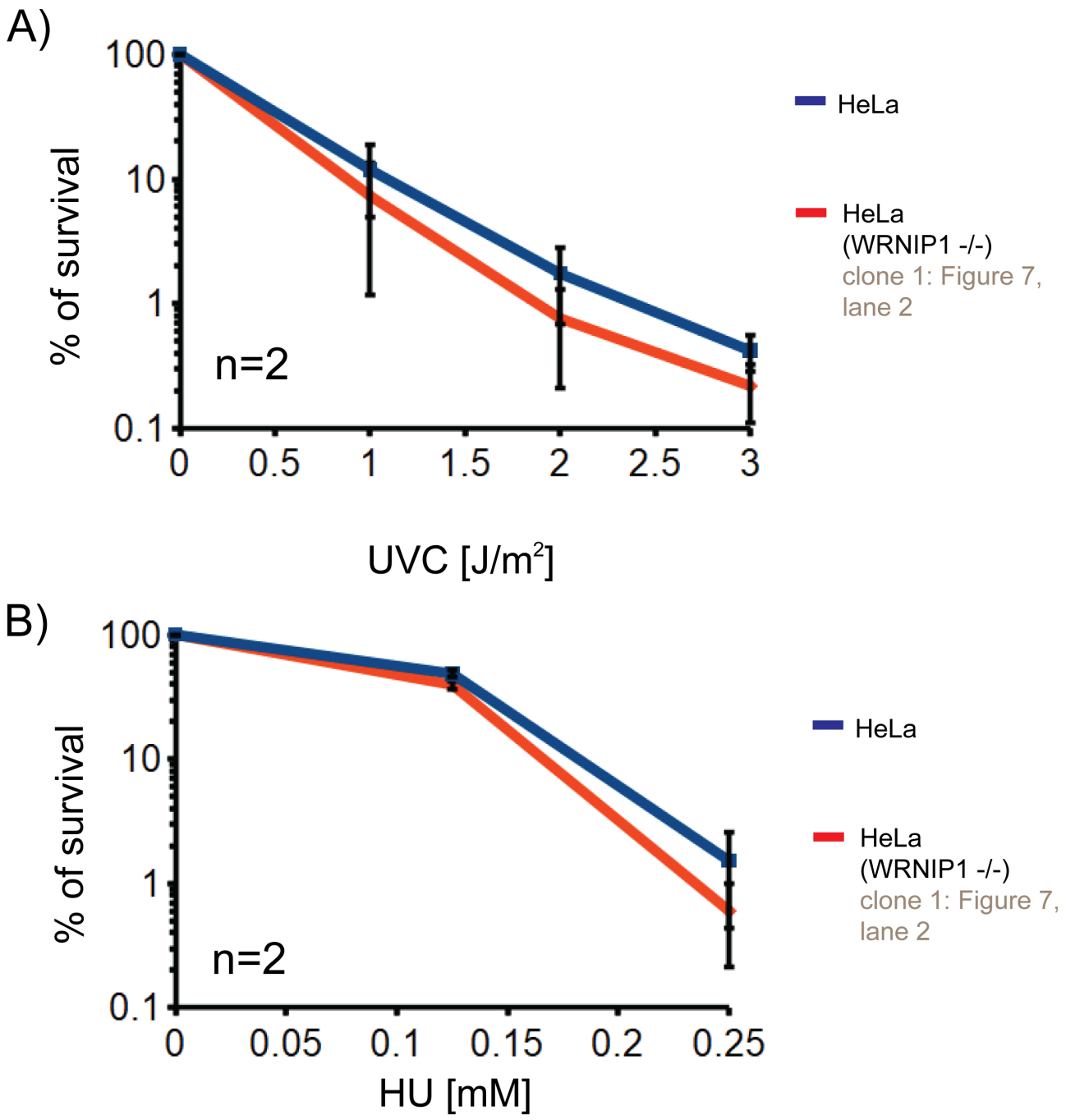


Figure 11: WRNIP1 depletion cells are not significantly more sensitive to A) UVC irradiation or B) HU treatment, than HeLa wild type cells. Cell lines were treated with indicated does of DNA damaging agent and left in the incubator to form colonies. In each case two independent experiments were combined to compare the differences, and the sensitivity of the WRNIP1 deficient cell line and the wild type cell line were no further than one standard error away from each other at each tested point.

3.1.3 WRNIP1 depletion does not increase the sensitivity of the FANCD2 $^{-/-}$ cell line to ICL inducing agents

After confirming that WRNIP1 deficient cells are sensitive to, both MMC and TMP-UVA treatment, both of these DNA damage inducing agents were used in the epistasis experiments. For these particular survival assays a combination of four different cell lines was used: HeLa wild type, with no modifications, HeLa WRNIP1 $^{-/-}$, HeLa FANCD2 $^{-/-}$ and the cell line where both WRNIP1 and FANCD2 were independently disrupted using the same CRISPR genome editing technique twice. To rule out the possibility of the clonal effect each cell line was prepared in two independently obtained variants. Since both, depleting FANCD2 as well as depleting WRNIP1 makes the cell lines more sensitive to ICL inducing agents we wanted to see what happens when the two gene disruptions are induced in the same cell line – whether the double depletion will cause cells to be more sensitive than any of the single ones (additive effect), or not.

Several outcomes were possible. One would be the case in which disrupting WRNIP1 gene rescues to certain extent the sensitivity of the FANCD2 deficient cell line. It would indicate that while FANCD2 is necessary during later stages of the DNA damage repair, lack of WRNIP1 during earlier stages allows for the majority of the damage signals to be redirected into different repair mechanisms that are not activated in the presence of WRNIP1. It was proven to not be the case here.

Two other options were either that WRNIP1 depletion will make cells even more sensitive to the ICL damage, which would suggest that there are at least some WRNIP1 functions that are independent of FANCD2, or that WRNIP1 depletion does not increase the sensitivity of FANCD2 deficient cell line which would suggest that whatever its function is its absence is not making the cells that are already lacking the FA pathway mechanism visibly more vulnerable to ICL damage. Both clonogenic survival assay (MMC treatment (**Figure 12**) as well as TMP/UVA (**Figure 13)**) have shown that adding WRNIP1 depletion on top of FANCD2 depletion is not statistically increasing the sensitivity of the cells to the induced damage.

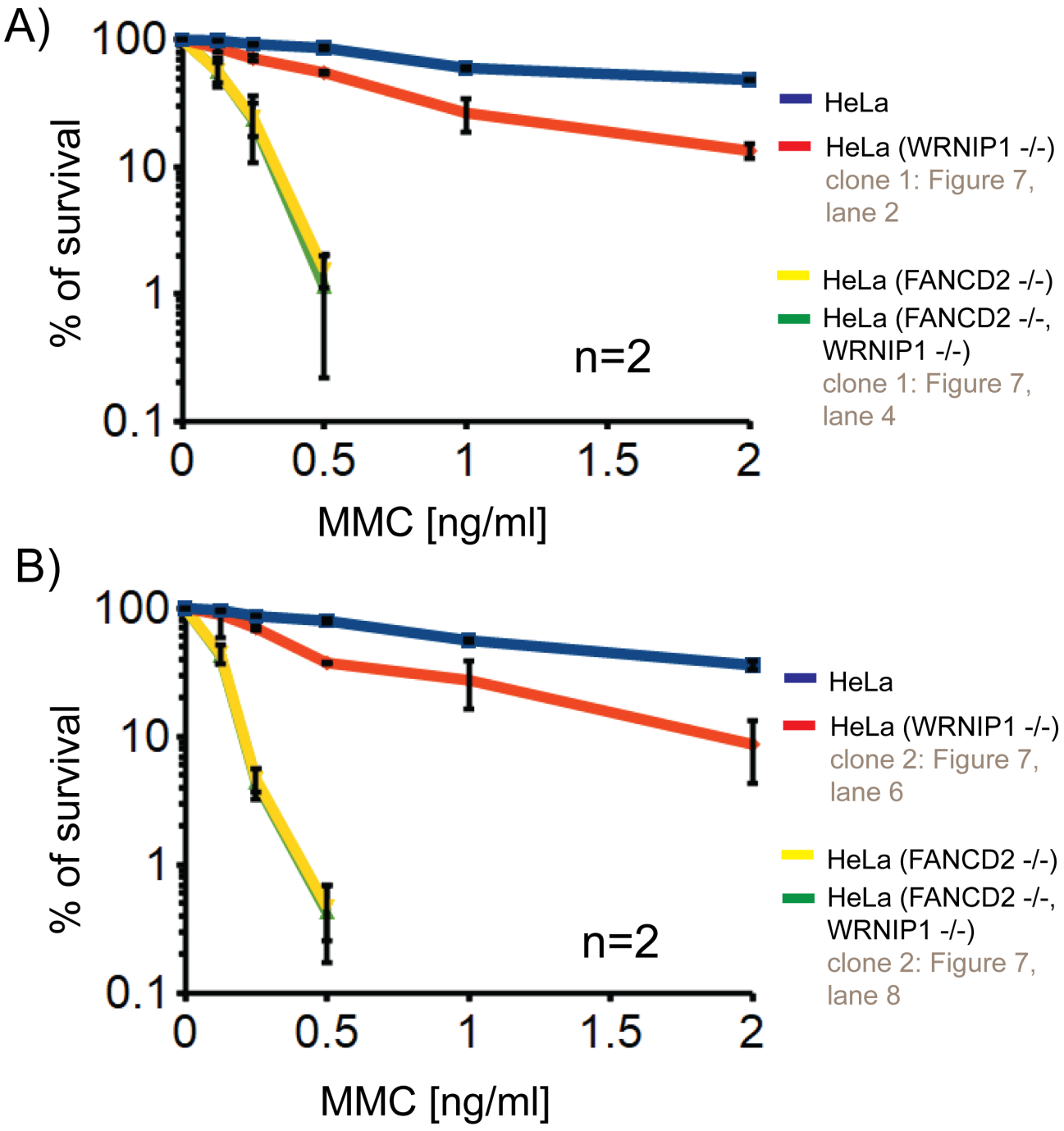


Figure 12: Depletion of WRNIP1 gene on top of FANCD2 gene does not further increase the MMC sensitivity of the cell line. The survival of two independently obtained sets of clones was tested after treatment with indicated doses of MMC. For each graph two independent experiments were combined (n=2) A) Assays performed with cell lines corresponding to the lanes 1-4 (Figure 9) B) Assays performed using cell lines corresponding to the lanes 5-8 (Figure 9). For both graphs, t-Student test was performed to confirm if the difference between the survival of WRNIP1 deficient cell line and wt HeLa cells is significantly different after treatment with 0.5 ng/ml of MMC (the highest concentration at which there are still some countable colonies of the FANCD2 deficient cells and WRNIP1 and FANCD2 deficient cells. The t-Student test has confirmed that the difference between the two cell lines is significant (p=0.05).

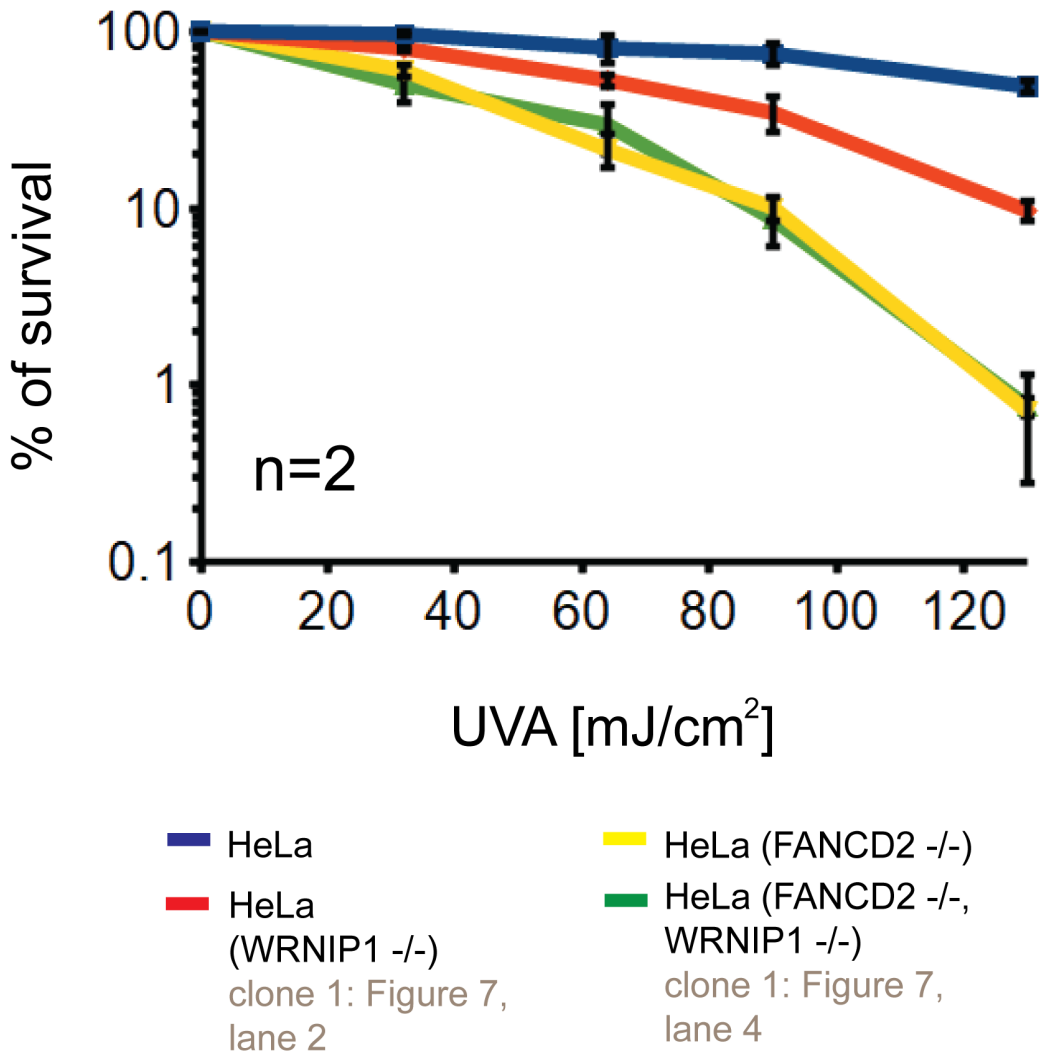


Figure 13: Disrupting WRNIP1 gene on top of FANCD2 gene does not further increase the TMP-UVA sensitivity of the cell line. The results obtained with MMC (**Figure 12**) were additionally confirmed using the TMP-UVA treatment. This time only the first set of clones was used (lanes 1-4; Figure 9). The cells were plated, left for one day to recover, then sensitized with TMP, and treated with increasing doses of UVA treatment. Two independent experiments were combined to plot the graph. To confirm if the difference in survival between WRNIP1 deficient cell line and control cell line is statistically significant at the highest dose of the treatment t-Student test was used. The difference is significant ($p=0.05$). HeLa (FANCD2 $^{-/-}$) and HeLa (FANCD2 $^{-/-}$, WRNIP1 $^{-/-}$) cell line survival rate at this dose are no further from each other than one standard error.

3.1.4 Discussion

The analysis of the literature suggests that it is likely that the protein is involved in some kind of DNA damage repair activity. Given the presence of the UBZ4 domain at its N-terminus as well as the fact that ubiquitin signalling is known to play an important role in Fanconi Anemia pathway the involvement of WRNIP1 in the FA pathway seemed like a hypothesis worth pursuing. The initial survival tests have shown that while lack of WRNIP1 makes cells more vulnerable to the DNA damage induced by the MMC treatment, adding the exogenous WRNIP1-EGFP brings the levels of survival to the point where its statistically comparable with the wild type. This data confirms the initial hypothesis of WRNIP1 being somewhat important in resolving the ICL damage. MMC induces other types of damage too, so to additionally test the hypothesis WRNIP1 deficient cells were also tested against the psoralen/UVA, UVC and hydroxyurea treatment. While both UVC and HU cell survival assays showed no significant difference in WRNIP1 $^{-/-}$ cells sensitivity, TMP-UVA treatment has shown that WRNIP1 is in some way important for the survival of the cells.

The results of the epistasis experiments suggest that whatever is the function of WRNIP1 in repairing the damage caused by MMC or TMP/UVA, it is redundant once the FA pathway is disrupted. This may be explained by either WRNIP1 having some kind of regulatory function during early stages of the pathway, in which case the lack of the protein does not make any difference in the absence of the proteins it is supposed to regulate, or alternatively at this stage it is equally probable that WRNIP1 is involved at later stages of the ICL repair and that lack of FANCD2 that would otherwise regulate its function in some way is enough to ensure WRNIP1 will not perform its biological tasks. That possibility would indicate that lack of WRNIP1 by itself is partially affecting the survival, but since FANCD2 is still present there are still other alternative ways of resolving the damage. However, when FANCD2 gene is disrupted WRNIP1 (though correctly expressed and present at physiological levels), remains inactive and in this case its depletion makes no difference for the cell survival. At this stage of the project it was impossible to

establish which of the two options is more likely and more experiments including observing the timing of the recruitment of both WRNIP1 and FANCD2 to the damage site were needed to suggest any reasonable molecular model.

3.2 WRNIP1 is recruited to the TMP-UVA induced ICL damage site

The next question that we wanted to answer is whether or not WRNIP1 is recruited to the ICL damage site and if it is then what is the relationship between the recruitment of WRNIP1 and FANCD2. In order to address this problem we used the HeLa Flp-in T-REx WRNIP1-EGFP expressing cell line, that was previously mentioned in the context of the clonogenic survival assays (Chapter 3.1 Figure 7, lane 3). In this cell line the tetracycline repression of the transgene operates unless switched off by adding tetracycline (or the more stable doxycycline) which then permits expression of the integrated epitope-tagged transgene. The HeLa Flp-in T-REx cell line is modified to have an FRT integration site in its genome, conferring Zeocin resistance and tetracycline repressor gene selected by Blasticidine resistance. A transgene of choice (WRNIP1-EGFP in this case) cloned into a pcDNA5-FRT/TO-WRNIP1-EGFP vector can be integrated specifically into the FRT site by co-transfecting Flp-in recombinase (on the pOG44 plasmid). The recombination event destroys the Zeocin resistance and introduces Hygromycin resistance (the Hygromycin resistance gene encoded on the pcDNA5-FRT/TO-WRNIP1-EGFP only works once integrated into the Flp-in cell line since the gene on the vector lacks the start ATG codon which is provided by the FRT integration site in the cellular genome. pcDNA5-FRT/TO-WRNIP1- EGFP plasmids can therefore not be used for stable expression in other cell lines. The tetracycline repression of the transgene operates unless switched off by adding tetracycline (or the more stable doxycycline) which then permits expression of the integrated epitope-tagged transgene.

3.2.1 WRNIP1-EGFP was expressed at levels close to the physiological levels of endogenous protein

We used doxycycline in order to control the expression of WRNIP1-EGFP but even as low concentration as 0.5 ng/ml turned out to cause overexpression. In order to find a concentration that gives physiological levels of WRNIP1-EGFP a titration was performed where the concentrations of

0.1, 0.2, 0.3, 0.4 and 0.5 ng/ml were tested. WB analysis of the titration provided us with the estimation which of the five tested concentrations is the best one. It seems that the best concentration which gives the ratio 1:1 is 0.2 ng/ml (**Figure 14**).

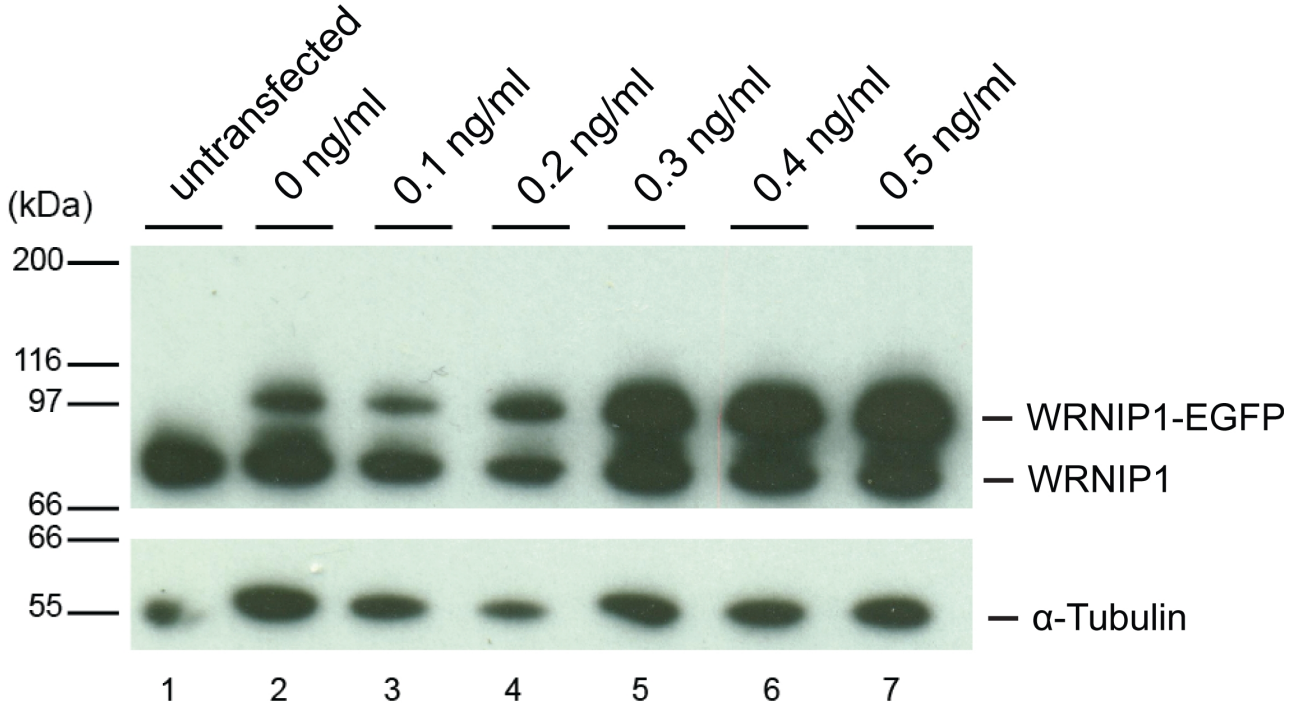


Figure 14: Induction with 0.2 ng/ml dox results in 1:1 ratio of exogenous to endogenous WRNIP1. Untransfected cells were tested against the WRNIP1 antibody next to the cells untreated with doxycyclin, and then cells treated with 0.1, 0.2, 0.3, 0.4 and 0.5 ng/ml respectively. The cells were lysed and whole cell lysates were tested against anti-WRNIP1 antibody. 9% gel was used for the WB analysis

While on the Western Blot the cell line treated with this amount of doxycyclin seemed to express physiological levels of WRNIP1, the microscopy analysis seemed to suggest something else. The same cells compared after 24 h treatment with 0.2 ng/ml, 1 μg/ml and no treatment at all seem to have very different fraction of cells expressing WRNIP1-EGFP (**Figure 15**). It seems that the expression of WRNIP1-EGFP is not uniform, and becomes even less uniform when the cells are

treated with low doses of dox. While we could not fully fix this issue, we decided to treat cells with higher level, for shorter time than 24 h, which resulted in more uniform (though still not perfect) expression of WRNIP1-EGFP. It is important to note, that while this technique allowed us to express WRNIP1-EGFP levels similar to physiological, it was not accurate enough to allow us the perfect physiological expression, and still the population of cells had some cells expressing almost no WRNIP1 and some cells overexpressing it.

The cells were grown overnight on the microscopy coverslips, and then stained with DAPI to see the nuclei. Multiple fields of view were photographed, for each slide. The number of cells with DAPI stained nuclei was compared with the number of cells visible in the green (FITC) channel giving the following numbers (**Table 9**).

SAMPLE [µg/ml of dox]	ALL CELLS	EGFP EXPRESSING CELLS	% OF EXPRESSION
0	185	0	0
0.0002	160	3	1.9
1	108	62	57.4

Table 9: Comparison of the percentage of the cells expressing EGFP, depending on the concentration of dox that they were treated with beforehand

This analysis has shown that such a drastic reduction in the amount of added doxycyclin reduced the fraction of the cells expressing WRNIP1-EGFP, rather than reducing the amount expressed by each cell.

To achieve desired physiological levels of WRNIP1-EGFP, the cells were treated with higher concentrations of doxycyclin (50 ng/ml) but for shorter time, and then the media was removed and the cells were harvested after accordingly 3, 6 or 24 h incubation period to see if the levels of WRNIP1 are stable or if they are rapidly dropping after dox removal (**Figure 16**).

Surprisingly the levels of protein were physiological after 3 hours of incubation but then even though the drug was already removed, the protein was still expressed, reaching the highest levels 6 hours after media exchange.

The data was confirmed with microscopy images to make sure that this time the majority of cells are expressing WRNIP1-EGFP (**Figure 17**).

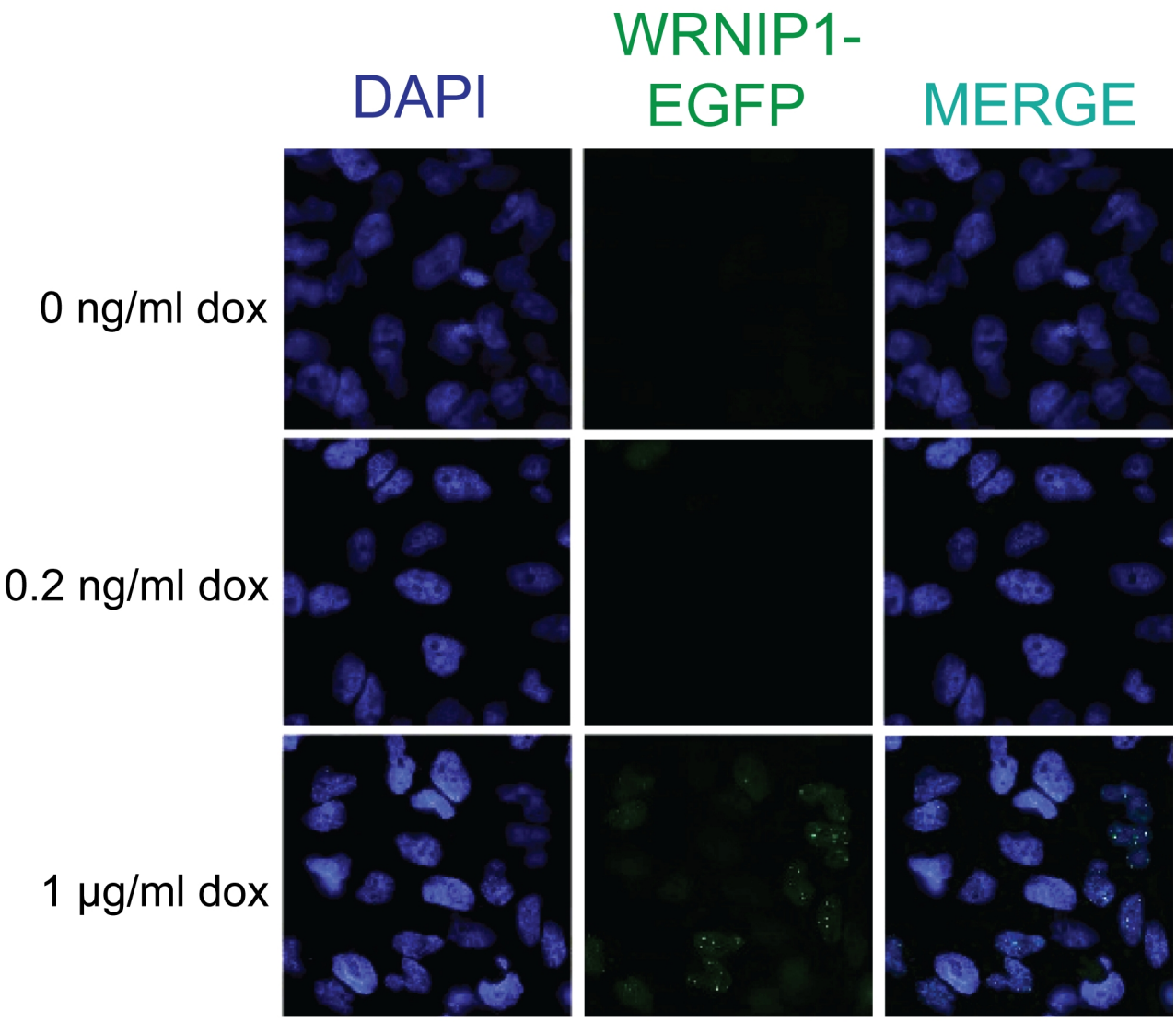


Figure 15: Very low concentration of dox is inducing WRNIP1-EGFP expression in a very small fraction of cells only. The same cell line was treated for 24 h with 0.2 ng/ml dox and 1 µg/ml dox. At the same time some cells were not treated with dox at all to be used as a reference. The cells were fixed on the coverslips, stained with DAPI and photographed using both DAPI and FITC channels.

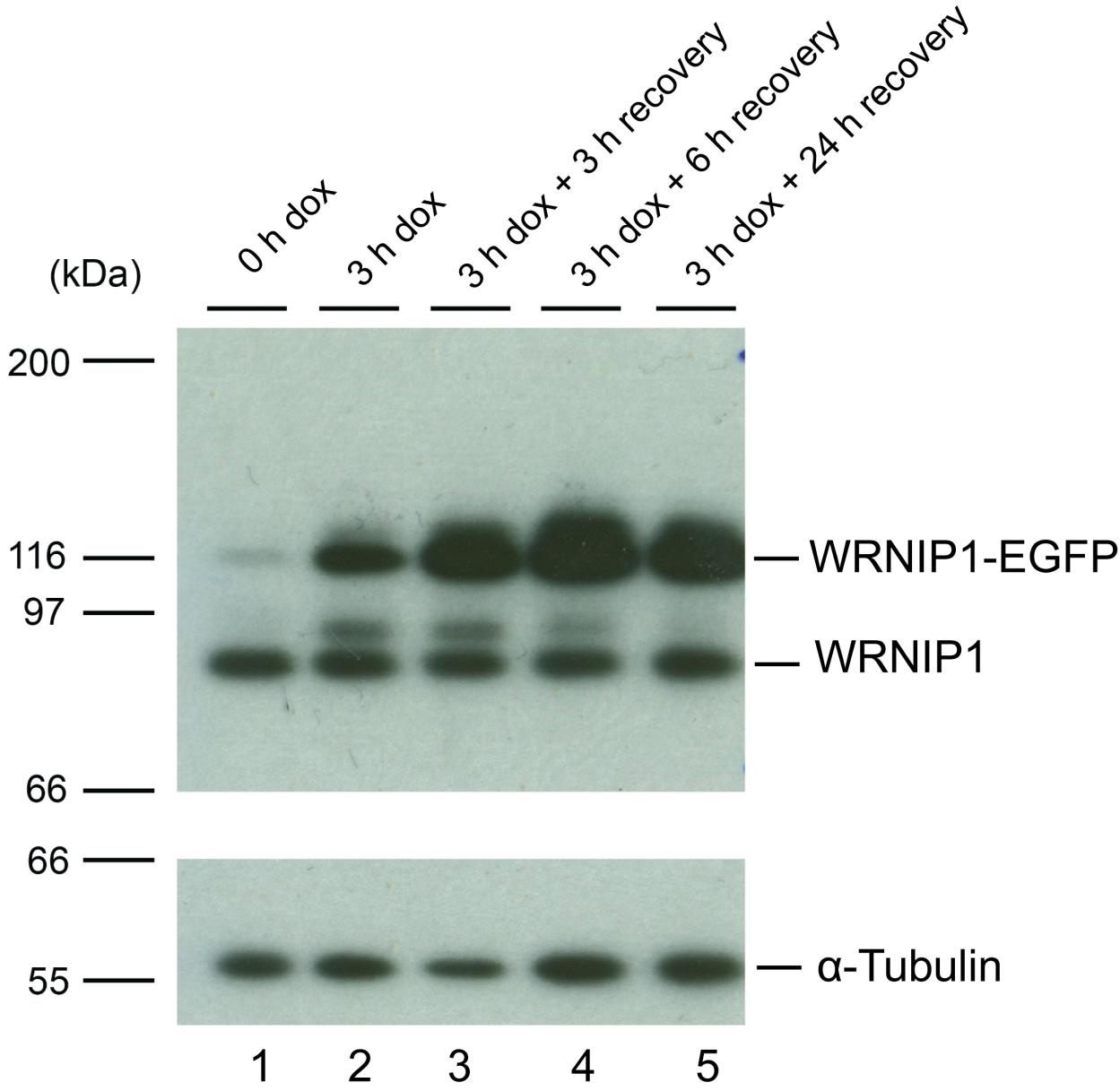


Figure 16: After 3 h of dox treatment at the concentration of 50 ng/ml the levels of WRNIP1-EGFP are comparable to the physiological levels of WRNIP1. Western Blot analysis of the cells that were treated with doxycyclin for 3 h and then harvested at different time points after removing it (0 h, 3 h, 6 h and 24 h after media exchange). The cells were harvested and lysed, and then tested against anti-WRNIP1 and anti-tubulin antibodies. The gel used for the WB analysis was 9% polyacrylamide.

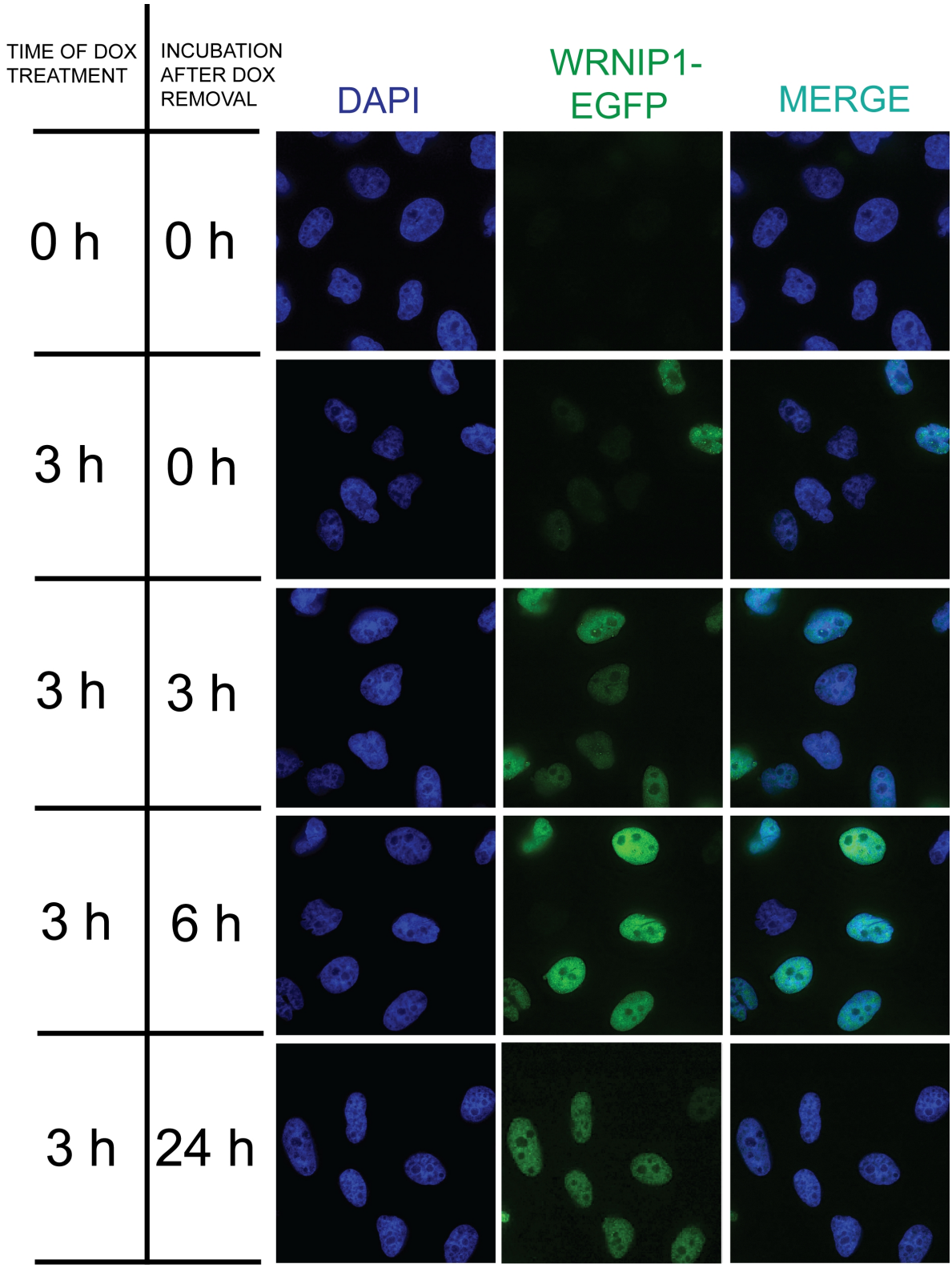


Figure 17: 50 ng/ml of dox is enough to ensure universal expression of WRNIP1-EGFP. The microscopy analysis of the time course titration analysis. The first row shows images acquired for the control cells that were transfected with the WRNIP1-EGFP plasmid but not induced with doxycyclin. The remaining panels were treated with dox for 3 h, and then incubated in fresh medium for as long as indicated.

3.2.2 WRNIP1 is recruited to ICLs after TMP-UVA treatment and not in absence of TMP

The obtained cell line and the tested conditions were used to perform live cell imaging microscopy. The technique we used consists of first sensitizing cells with trioxalen (TMP), and then drawing a small stripe through the selected nuclei with the 405 nm UVA laser. Trioxalen is a psoralen derivative, and in combination with UVA irradiation is known to introduce ICL damage in DNA. This technique allows to see a quick response of the proteins tagged with a number of fluorescent tags (in this case WRNIP1 tagged with EGFP), because if the given protein is being recruited to the damage site then after a while the part of the nucleus that was irradiated will show stronger fluorescence than the rest of the nucleus.

Additionally, to exclude the possibility that the laser power that we used was too high we also tested the same sample but untreated with TMP. Only the combination of TMP and UVA resulted in the recruitment of WRNIP1 proving that WRNIP1 responded specifically to ICL damage (**Figure 18**).

Both experiments were quantified using FiJi (ImageJ) software. For each cell the brightness of the stripe was measured (for details see materials and methods), and plotted on the graph. While the cells that were not sensitized with TMP showed no WRNIP1 recruitment, the ones treated with it have shown approximately 2 fold increase in the amount of WRNIP-EGFP on the irradiated site and the recruitment was visible as early as 3 min after the irradiation.

Interestingly many cells observed during the live microscopy experiments (and some fixed cells experiments, for example Figure 13, when the cells were treated with a very high dose of dox) contained foci. We cannot see them in the cells observed during the fixed cell imaging, that were treated with lower doses of dox, most likely due to different treatment of the samples than in case of the live cell imaging. While we do not have enough data to establish the function of these foci, we can speculate that they are either some sort of a DNA repair complex, or an artefact due to non-physiological expression levels of WRNIP1.

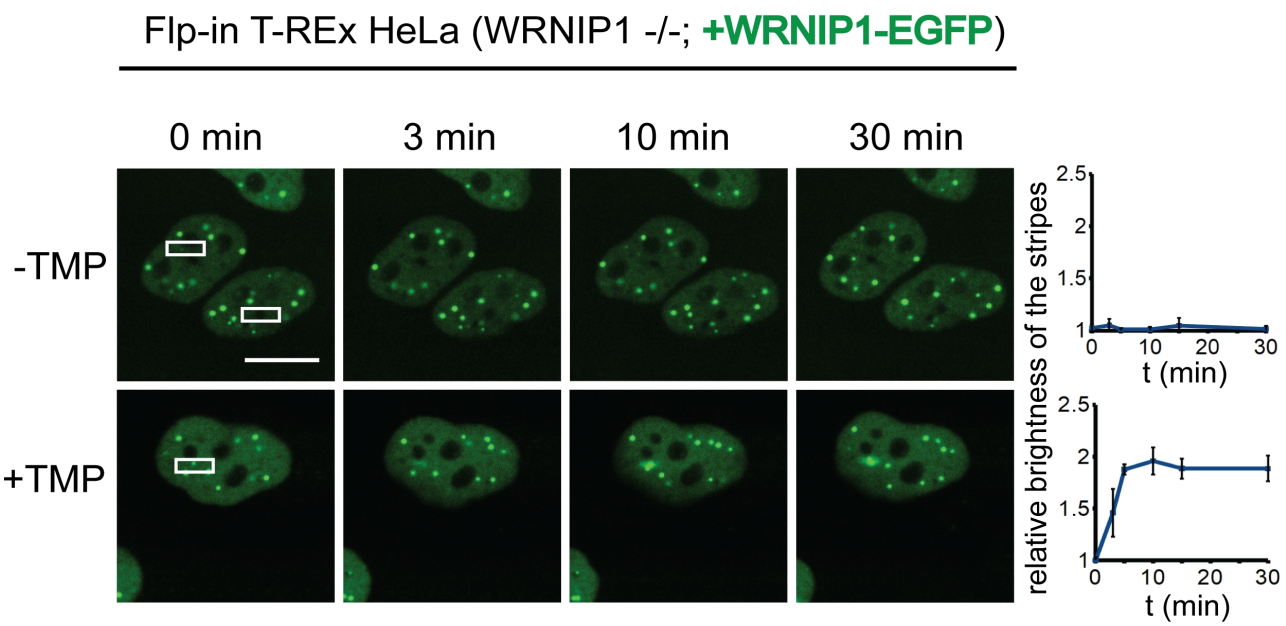


Figure 18: WRNIP1-EGFP is being recruited to the ICL damage site after UVA (405 nm) laser irradiation if the cells have been sensitized with TMP beforehand. There is no observable recruitment if the cells. The scale bar is equivalent to 10 μm. Note that the images of the cells treated with TMP, were acquired during the experiment during which some images presented in Figure 29 and Figure 38 were acquired, and therefore they will be presented again. 3 independent experiments were used for the quantification (N=3), 18 cells in total, 9 for cells untreated with TMP and 9 for the cells treated with TMP.

3.2.3 WRNIP1 is recruited before FANCD2 recruitment

Since WRNIP1 recruitment was observable as early as 3 min after the irradiation and FANCD2 was known to be recruited to ICLs as well, but usually its accumulation was starting to be visible around 15 min after the irradiation it was reasonable to assume that WRNIP1 is recruited before FANCD2. To confirm this hypothesis mCherry tagged FANCD2 was transfected into the HeLa Flp-in T-REx cells that were already transfected with WRNIP1-EGFP, using the Phoenix A system (see Materials and Methods) (**Figure 19 A**). The cell line expressing both proteins with different fluorescent tags was sensitized with TMP, and irradiated with 405 nm laser, and then the images were acquired for the next hour in both green and red fluorescent channel (**Figure 19 B**).

After the experiment stripes in both EGFP and mCherry tagged proteins were quantified using the standard quantification protocol and compared with each other giving a clear proof that WRNIP1 is indeed recruited before FANCD2 reaches observable levels at the ICL damage site (**Figure 19 C**).

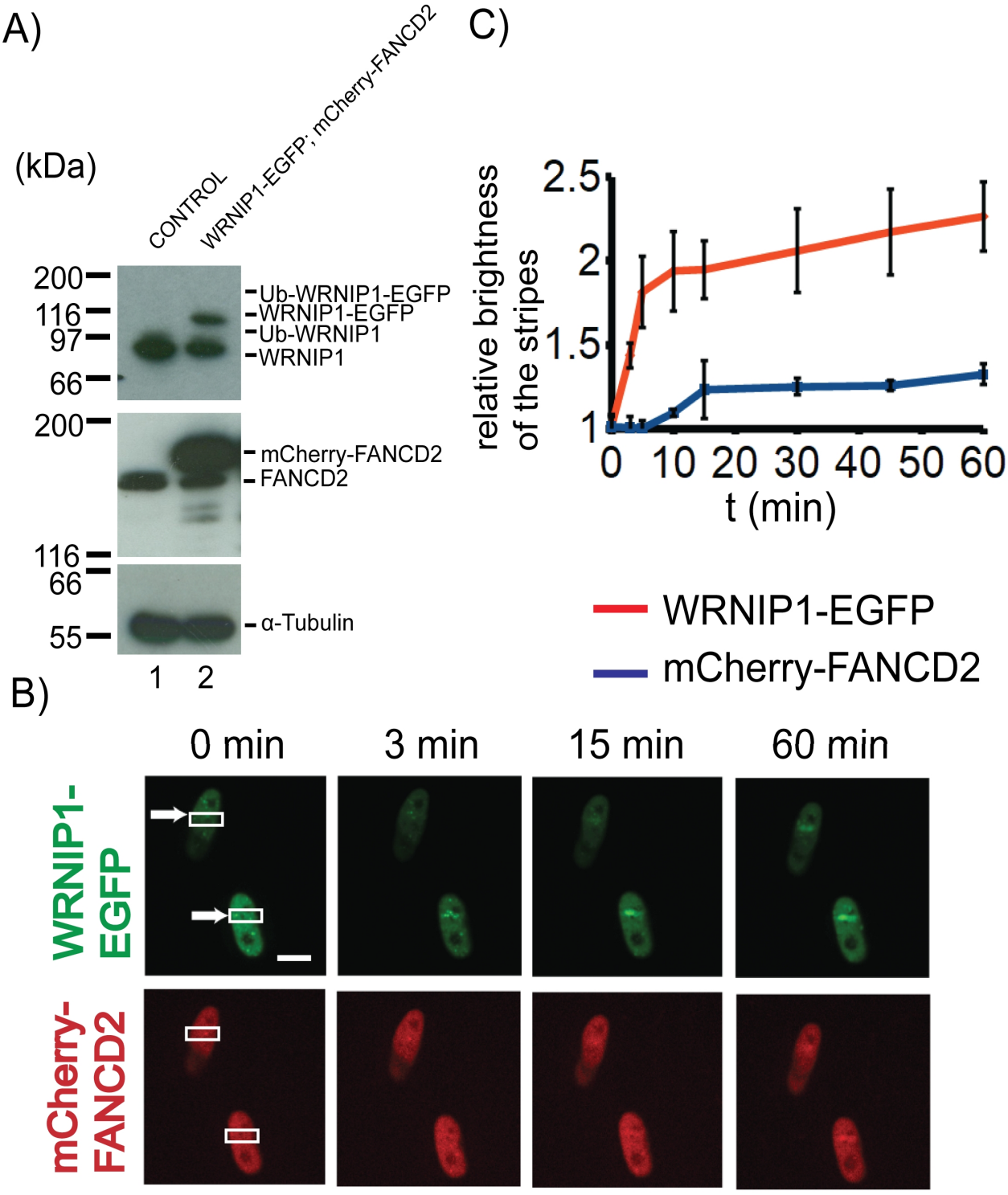


Figure 19: WRNIP1 recruitment at the ICL damage site can be observed 3 min and FANCD2 can be observed 15 min after the damage is introduced into the genome. WRNIP1-EGFP and mCherry-FANCD2 expressing cell line was irradiated with 405 nm laser after pretreatment with TMP, and the protein recruitment was observed for an hour after the irradiation. A) WB analysis of the used cell lines. The scale bar is equivalent to 10 μm. B) Microscopy images of the irradiated cells taken in both EGFP and mCherry channels, C) Quantification of the results, 3 independent experiements (N=3), 9 cells in total.

This observation would indicate that if WRNIP1 is indeed participating in ICL repair it is likely to play some role upstream to FANCD2 rather than downstream.

To further investigate the relationship between FANCD2 and WRNIP1 we designed four new HeLa Flp-in T-REx cell lines expressing fluorescent tagged proteins (**Table 10**).

No.	endogenous WRNIP1	WRNIP1-EGFP	endogenous FANCD2	mCherry-FANCD2
1	Yes	-	Yes	Yes
2	-	-	Yes	Yes
3	Yes	Yes	Yes	-
4	Yes	Yes	-	-

Table 10: Cell lines designed for further analysis of WRNIP1 recruitment in relation to FANCD2 recruitment.

As indicated in Table 10 the first two cell lines were both expressing EGFP-tagged FANCD2 on top of physiological levels of FANCD2, but while one of the cell lines was expressing physiological levels of endogenous WRNIP1, the other one had WRNIP1 gene disrupted using CRISPR/Cas9 technique.

The other two cell lines, were designed in analogous fashion, but this time the proteins were reversed, so both of them were expressing WRNIP1-EGFP on top of physiological levels of endogenous WRNIP1, and one of the cell lines was expressing wild type FANCD2, while the other one had FANCD2 gene disrupted, using the CRISPR/Cas9 technique.

By performing the chromatin recruitment analysis on all of these cell lines and comparing the results between them, we were hoping to establish whether there is any connection between WRNIP1 recruitment and FANCD2 recruitment.

3.2.4 FANCD2 depletion does not affect WRNIP1 recruitment to the ICL damage site

To further investigate the relationship between WRNIP1 and FANCD2, we observed and compared WRNIP1-EGFP recruitment to the ICL damage site in two cell lines: one expressing wild

type FANCD2 and WRNIP1-EGFP and the other one also expressing WRNIP1-EGFP but in the absence of FANCD2 (FANCD2 depletion) (**Figure 20 A**). Both cell lines were sensitized with TMP and then irradiated with 405 nm laser. The images were acquired for half an hour in each case to observe WRNIP1 being recruited to the damage site (Figure 20 B), and after the acquisition was over the results were quantified (see Materials and Methods) (Figure 20 C).

Since it was already proven at this point that WRNIP1 is being recruited before FANCD2 it was not surprising that lack of FANCD2 did not affect in any visible way the recruitment of WRNIP1-EGFP.

3.2.5 WRNIP1 depletion results in decreased FANCD2 recruitment

To test the hypothesis that WRNIP1 acts upstream to FANCD2 we observed and compared FANCD2 recruitment in two other cell lines: both expressing mCherry tagged FANCD2, but one with WRNIP1 gene disrupted and the other expressing wild type WRNIP1 **(Figure 21 A**). Both cell lines were tested using the standard live-cell chromatin recruitment assay (first cells were sensitized with TMP and then laser irradiated with UVA) (**Figure 21 B**), and then the protein recruitment was quantified as described previously (**Figure 21 C**).

While the difference between the two graphs was not dramatic, it was statistically significant as measured at 60 min after irradiation ($p=0.05$) which confirms that WRNIP1 acts upstream to FANCD2

3.2.6 WRNIP1 recruitment varies slightly depending on the stage of the cell cycle

The doxycyclin dependent HeLa Flp-in T-REx cell line, expressing WRNIP-EGFP was relatively easy to obtain and work with, but could not be used to estimate if there are any differences in WRNIP1 recruitment between cell cycle stages. The variation of WRNIP1-EGFP expression was regulated by dox levels and not by natural molecular mechanism of the cell.

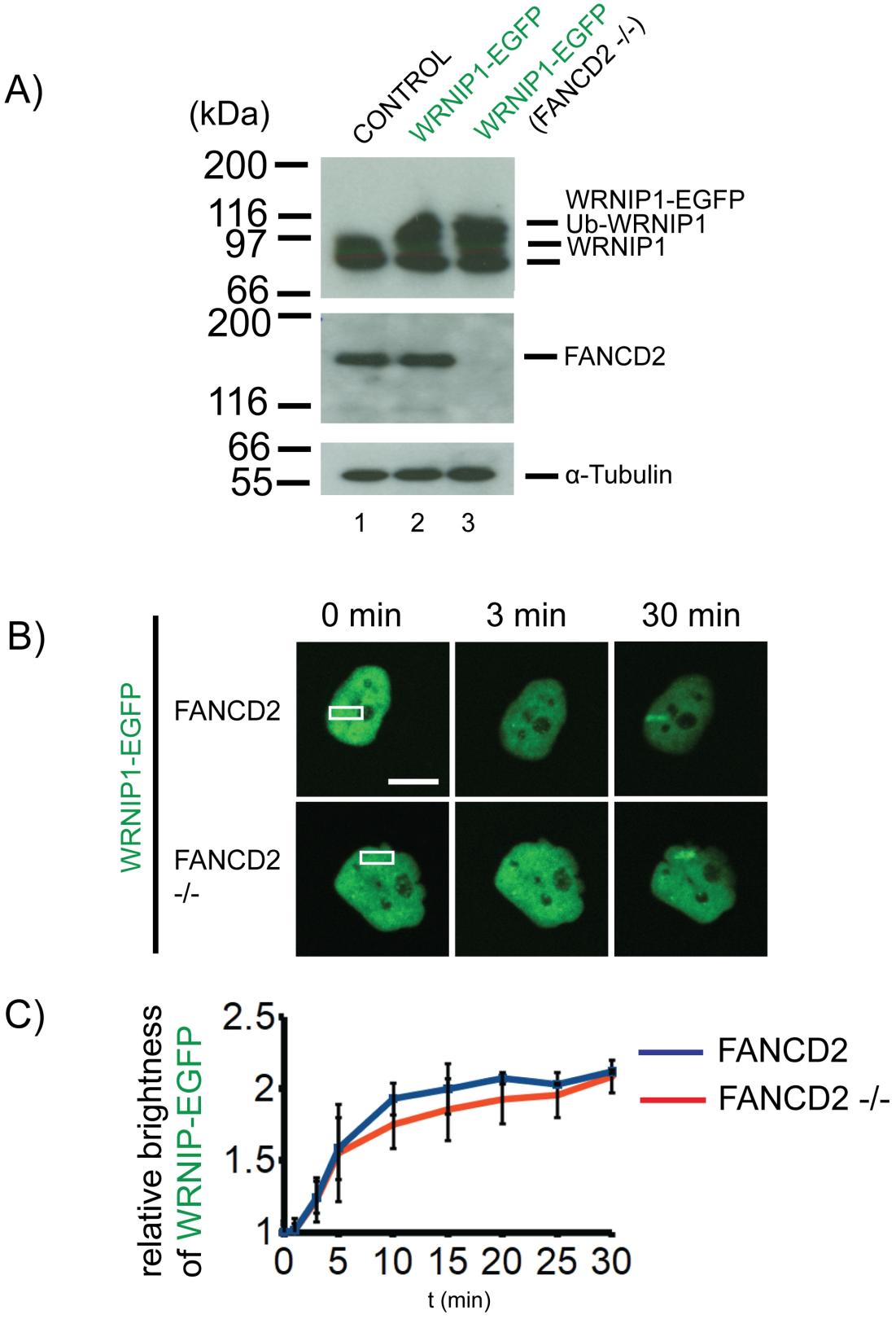


Figure 20: In the absence of FANCD2 WRNIP-EGFP recruitment is not affected. WRNIP1-EGFP recruitment after TMP-UVA treatment was compared in WT and FANCD2 deficient cell lines A) WB analysis of the cell lines B) Microscopy imaging after irradiating cells with 405 nm laser. The scale bar is equivalent to 10 µm. C) Quantification of the results: 2 independent experiments (N=2), 14 cells in total, 9 control cells, 5 FANCD2 deficient cells.

To avoid this obstacle, and to obtain the cell line, with the most physiological levels of WRNIP1, the partial insertional complementation cell lines were obtained, using the CRISPR/Cas9 technique (**Figure 22**). The obtained cell lines were there transfected with mCherry-tagged geminin fragment, that undergoes degradation in G1 stage of the cell cycle (the construct was provided by Chih-Chao Liang, dr Cohn's Laboratory), so it was possible to observe WRNIP1 recruitment and to determine if the cell is in G1 stage of the cell cycle (lack of the geminin fragment) or not (the geminin fragment is expressed).

Four separate experiments were performed, each time the cells were sensitized with TMP and then irradiated with 405 nm laser (as described in Materials and Methods) and the images were acquired in both mCherry and EGFP channels (**Figure 23 A**). Later the cells were segregated into two groups, depending on whether they express the m-Cherry tagged geminin fragment or not, WRNIP1-EGFP recruitment levels were quantified and compared between the G1 phase and the rest of the cell cycle (**Figure 23 B**). While the average recruitment was higher in the G1 phase, after calculating the confidence limits using the t-Student test (n=4, p=0.05), the difference turned out to be not statistically significant at any of the measured time points.

3.2.7 Discussion

The presented data confirms that WRNIP1 is recruited to the ICL damage site within first 3 minutes after ICL appears in the genome. It is a much quicker recruitment than in the case of FANCD2, which strongly suggested that WRNIP1 may be one of the proteins acting upstream to it.

To further confirm that WRNIP1 is recruited before FANCD2 both proteins were introduced into the same cell line, while being tagged with different fluorescent tags (EGFP and mCherry), which gave a very clear proof that WRNIP1 is indeed recruited first.

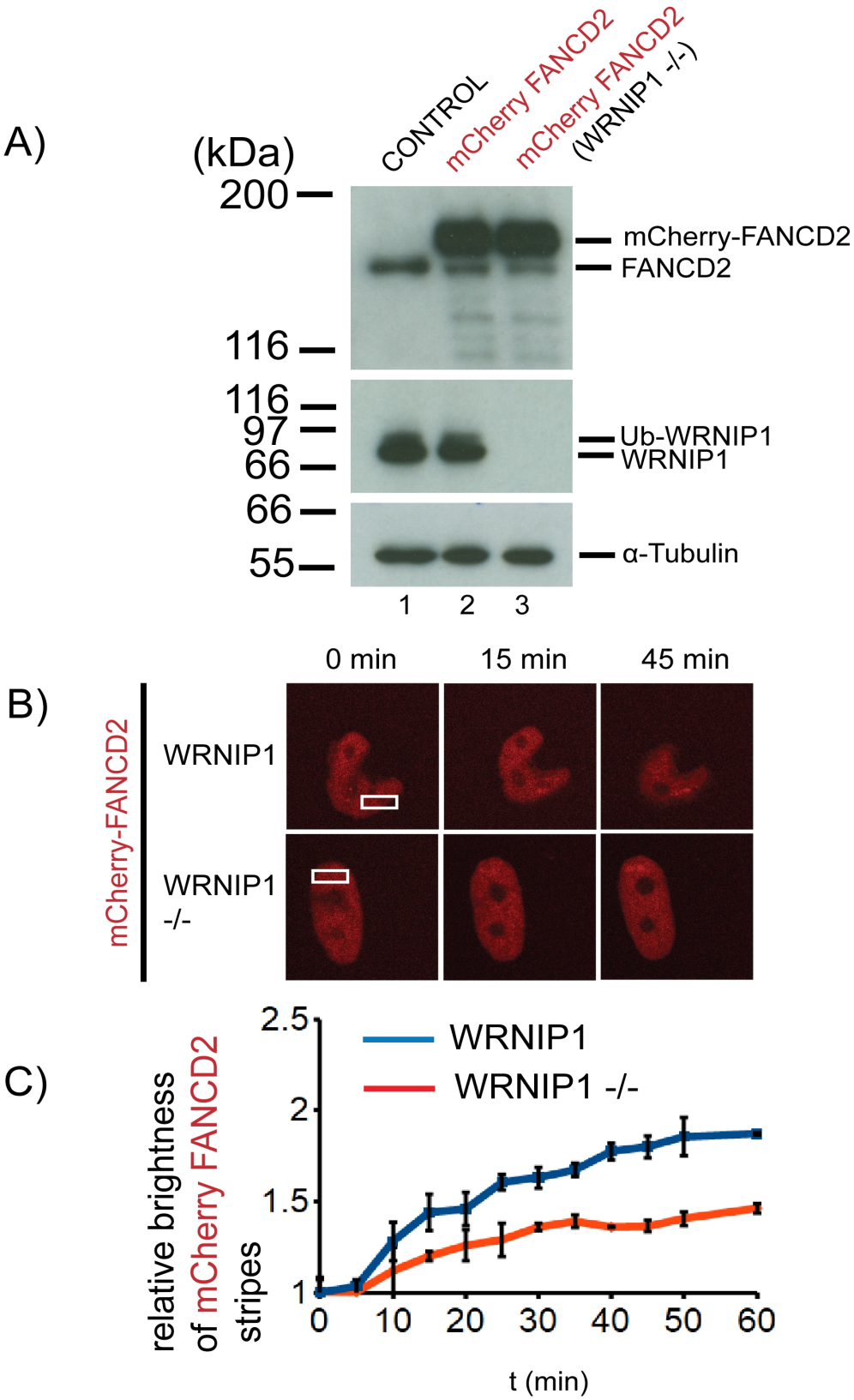


Figure 21: In the absence of WRNIP1 mCherry-FANCD2 recruitment is slightly less intense. mCherry-FANCD2 recruitment was compared in control and WRNIP1 deficient cell lines. A) WB analysis of the cell lines B) Microscopy imaging after irradiating cells with 405 nm laser. The scale bar is equivalent to 10 μm C) Quantification of the results: N=2, 9 cells in total, 4 control cells, and 5 WRNIP1 deficient cells, for p= 0.05 the differences are significant for 40 and 60 min, for p=0.1 all points above 25 min are significantly different.

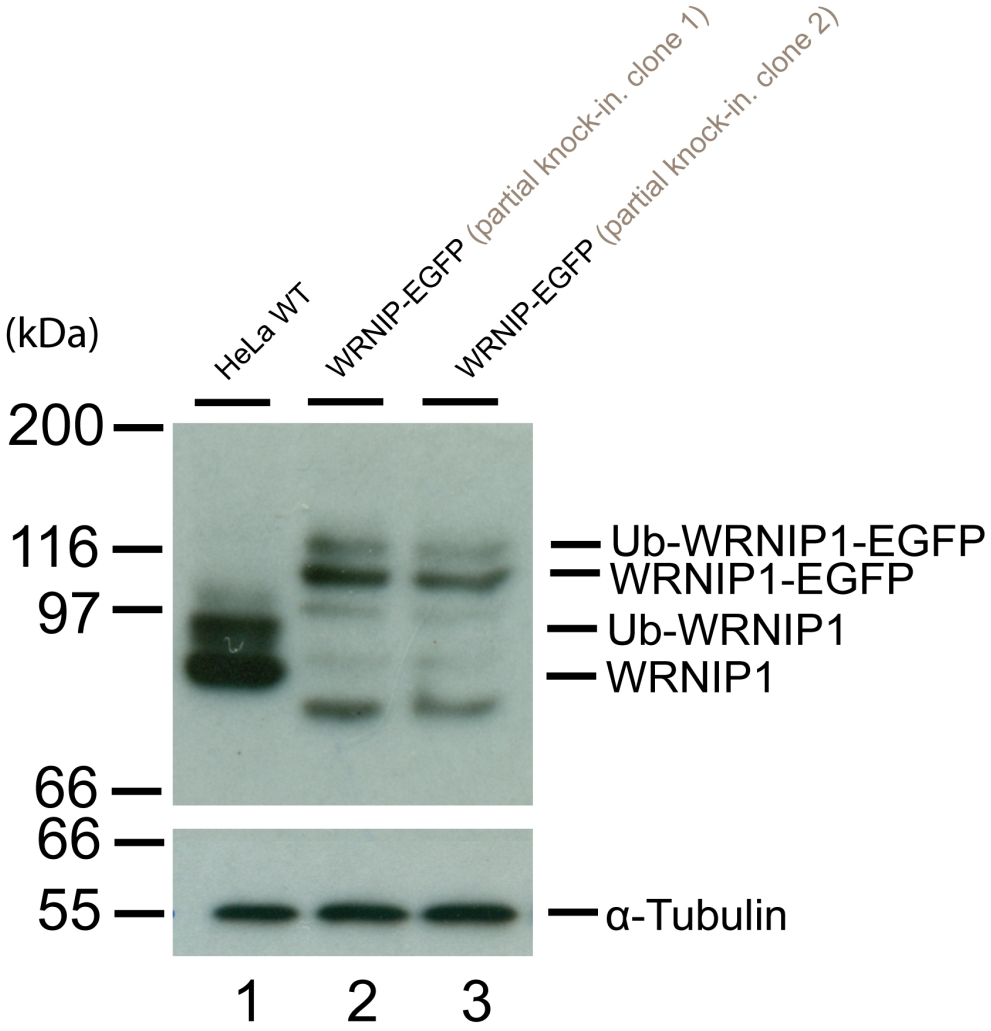


Figure 22: WB analysis of the two obtained WRNIP1-EGFP partial insertional complementation cell lines. Both cell lines were obtained using the CRISPR/Cas9 genome editing technique, and both show slightly lower than WT levels of WRNIP1-EGFP, suggesting that not all of the alleles were successfully targeted. In case of both partial insertional complementations there are several bands below the WRNIP1-EGFP band, that indicate high levels of protein degradation. The cells were harvested, lysed and tested with WB assay against anti-WRNIP1 and anti-tubulin antibodies, the concentration of acrylamide in the SDS-PAGE gel was 9%.

Moreover, while the cell line expressing WRNIP1-EGFP and no FANCD2 showed no difference in WRNIP1-EGFP recruitment, compared to the one expressing normal levels of wild type FANCD2, the cell line expressing mCherry-FANCD2 with WRNIP1 depletion had slightly (but statistically significant) decrease in FANCD2 recruitment compared to the cell line expressing normal levels of wild type WRNIP1.

The comparison of WRNIP1-EGFP recruitment to the damage site during G1 phase and the

rest of the cell cycle, suggested that WRNIP1 recruitment levels may be slightly higher during the G1 phase, but the difference was not statistically significant and relatively small though, so these results did not link WRNIP1 to either replication dependent or independent DNA repair. So far, it seems that WRNIP1 is recruited throughout the cell cycle, and independently of FANCD2 recruitment.

These results are suggesting that WRNIP1 acts upstream to FANCD2. This dependency does not have to necessarily be direct and at this stage it was unclear whether WRNIP1 is recruited to the damage site and then interacts with some proteins that are affecting the recruitment of FANCD2 or whether WRNIP1 interaction with FANCD2 is direct.

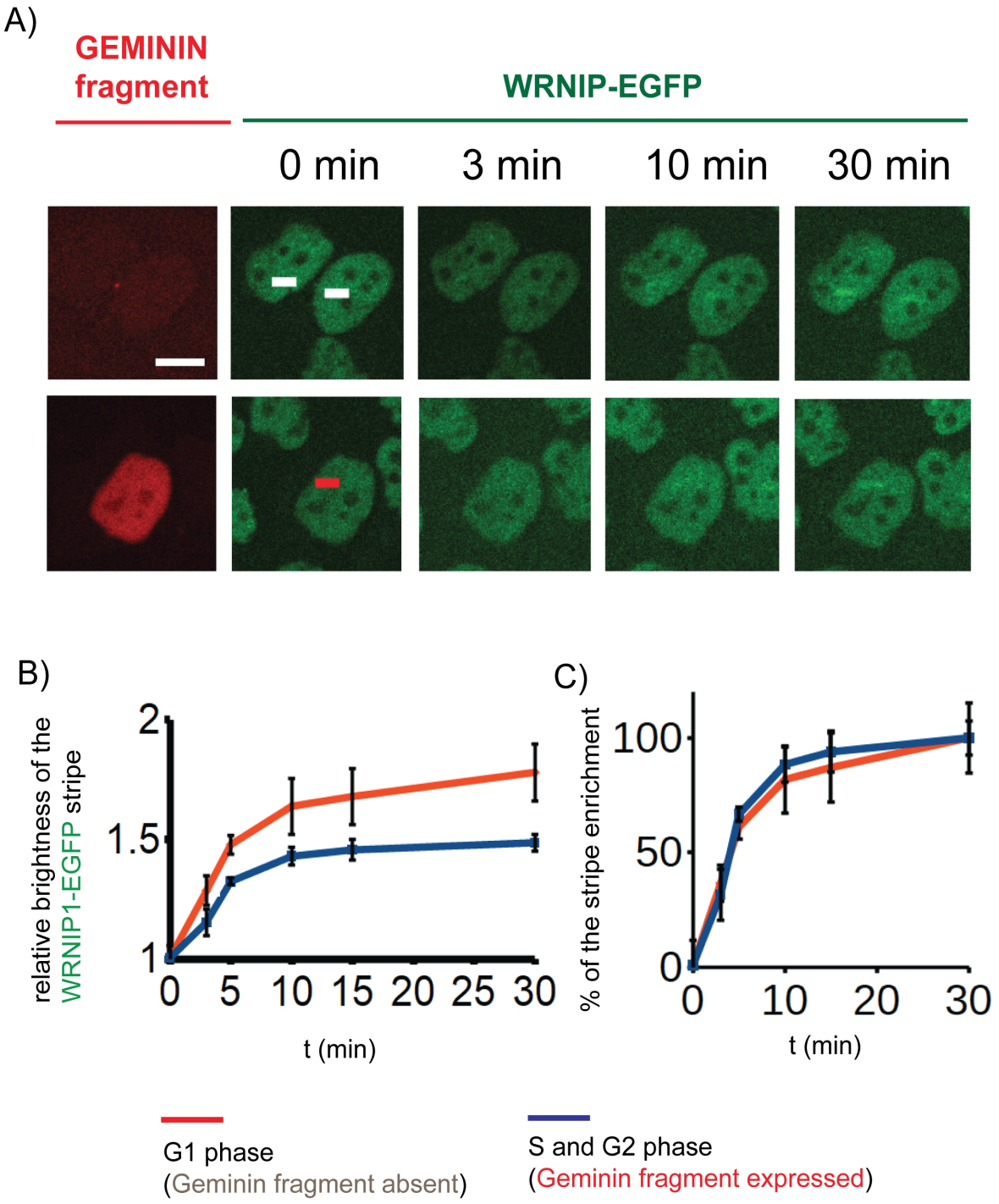


Figure 23: The levels of WRNIP-EGFP recruitment to the ICL damage site are slightly higher in G1 phase in comparison to S and G2 phase. A) Cells expressing both mCherry-Geminin fragment and WRNIP1-EGFP were sensitized with TMP, irradiated with 405 nm laser and the images were acquired for 30 min in both red and green channels. The cells were classified either as G1 phase cells (the ones where the irradiation site is marked in white), or as G2 or S phase (irradiation site marked in red); the scale bar is equivalent to 10 μm B) The results were quantified and plotted on the graph, N=4, 14 geminin fragment positive cells and 24 geminin fragment negative cells were quantified. The difference was not significant for p=0.05 (t-Student test). C) The same results was plotted on the graph but this time maximum recruitment was plotted as 100% for each set of cells. This allowed us to see the kinetics of the recruitment, which turned out to not be significantly different between the two sets.

3.3 Analysis of WRNIP1 upper form

3.3.1 WRNIP1 migrates as two species on SDS-PAGE gel

WRNIP1 antibody was usually picking up not one but two bands during the Western Blot analysis. The lower one was always migrating just under the 97 kDa marker, which corresponds to the 72 kDa molecular mass of WRNIP1. The upper form though migrates between 97 kDa and 116 kDa, and the distance suggests that it must be some kind of modification that shifts it more than monoubiquitination in one place would.

To further investigate this upper form, we prepared several dilutions of the whole cell lysate. The most concentrated sample contained 10 μg of protein in the whole cell lysate loading sample, and the other samples were accordingly diluted: 2, 4, 8, 16, 32 and 64 fold (**Figure 24**). By comparing the intensity of the bands, we could observe that the upper form in the sample that was diluted two fold (so there was a total of 5 μg of the proteins loaded on the SDS-PAGE gel) is slightly less visible as the lower form in the sample that was diluted 16-fold but more visible than lower form band of the sample diluted 32-fold, which indicates that the ratio of upper to lower form is somewhere between 1:8 and 1:16 (approximately 1:10 ratio).

3.3.2 Purification of the upper form

To further analyse the upper form of WRNIP1, we decided to express and purify FLAG-HA-WRNIP1 and purify specifically only the upper form in order to analyse it by the mass spectrometry analysis. In order to do it we obtained HeLa S3 cell line with stable expression of the FLAG-HA-WRNIP1 modified protein, using the Phoenix A transfection protocol.

The expressed protein was bound to the anti-FLAG beads, washed with the buffer and eluted and to test if the purification was successful initial lysate, insoluble faction (pellet), flow-through (FT) and eluate samples were analysed by Western Blot using anti-HA antibody, anti-WRNIP1 antibody and anti-Tubulin antibody (**Figure 25**).

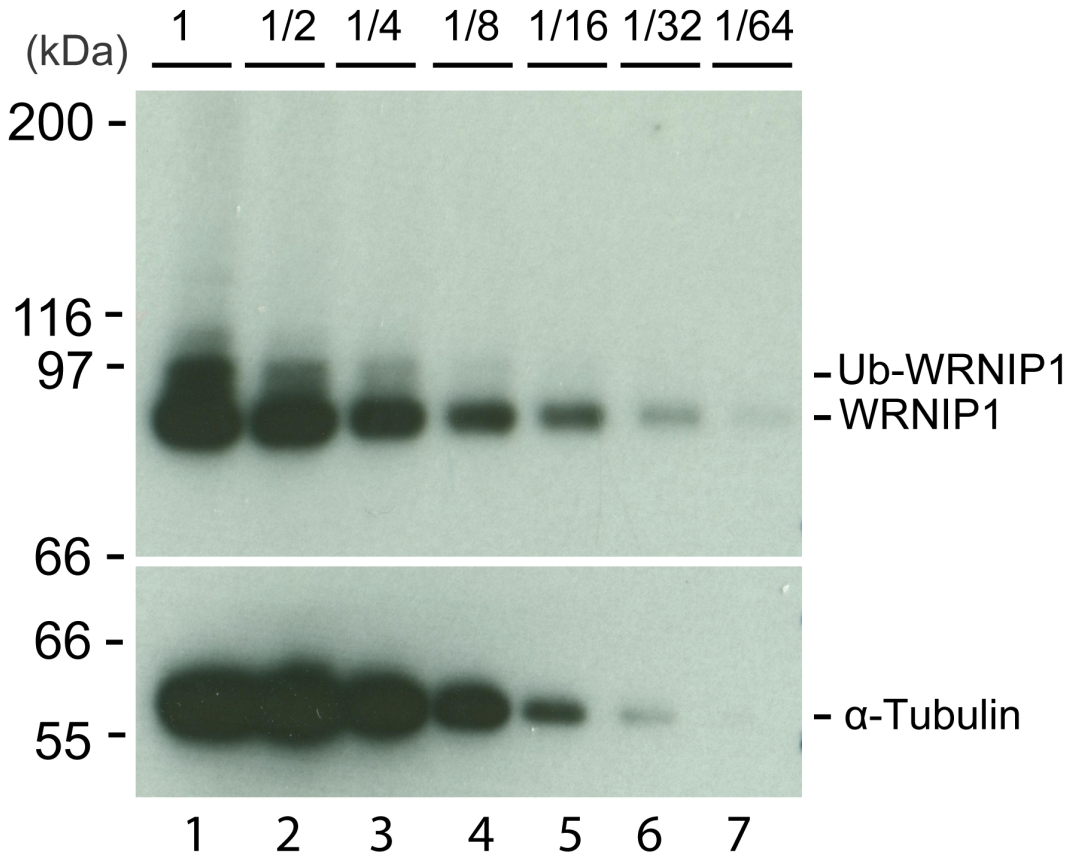


Figure 24: WRNIP migrates as two species during electrophoresis, and the ratio of upper to lower form is approximately 1:10. Titration of WRNIP1 from wild type HeLa cells. In the first lane there was 10 μg of the cell lysate loaded, and in each following lane the amount of lysate was diluted with equal volume of loading buffer. The FiJi software quantified the intensity of the upper band in the 3rd lane to be equal to 120% of intensity of the lower band in the 6th lane. Since the sample loaded into the 6th lane is 8 times more diluted than the sample in the 3rd lane, it means that, the approximate ratio of upper to lower form of WRNIP1 is close to 120%/ (8x100%)=120/800=0.15, so the amount of upper form is equal to approximately 15% of the lower form (roughly 1:10 ratio). The cells were harvested, lysed and tested in WB analysis against anti-WRNIP1 and anti-Tubulin antibody. The percentage of the SDS-PAGE gel used was 9%.

In the next step all of the obtained eluate was precipitated using **t**ri**c**hloroacetic **a**cid (CCl_3COOH), and then resuspended in loading buffer and loaded on a 9% polyacrylamide gel. After electrophoresis the gel was stained using the Coomassie dye, and then only the upper form of WRNIP1 was cut out from the gel and sent to the mass spectrometry analysis in order to find the residues that were ubiquitinated (**Figure 26**).

Later the whole purification was repeated according to the analogous protocol, but in smaller scale and one half of the purification was performed in the presence of N-Ethylmaleimide (NEM) added to all of the buffers at concentration of 10 mM. N-Ethylmaleimide is an organic compound that is a maleic acid derivative. One of its properties is inhibition of deubiquitinases, which would allow us to prevent potential deubiquitination of WRNIP1 during the purification process (which would occur *in vitro* and not during physiological processes regulated by the cell, therefore we would end up with less physiological pattern of WRNIP1 and its ubiquitinatated forms during the Western Blot analysis).

The obtained eluates were not precipitated this time but instead diluted with the loading buffer and analysed using three different antibodies: anti-WRNIP1, anti-HA tag and anti-mono and poly-ubiquitin (FK2).

It was previously reported by Bish et al. (2008), that WRNIP1 is polyubiquitinated with ubiquitin linkages K11, K48, and K63. Our analysis is consistent with these findings. While both anti-WRNIP1 and anti-HA antibodies, confirmed that the purified protein is indeed WRNIP1, FK2 antibody confirmed that the upper species is ubiquitinated. The FK2 antibody has also detected some longer ubiquitin chains attached to WRNIP1, visible as a smear above the upper form of WRNIP1, that would correspond to the polyubiquitinated WRNIP1 (***Bish et al., 2008)***. Since there is a barely visible shift of the bands obtained from the sample treated with NEM (visible especially in the FK2 blot), which is characteristic for this chemical, it is an additional proof that the samples were not accidentally swapped (**Figure 27**).

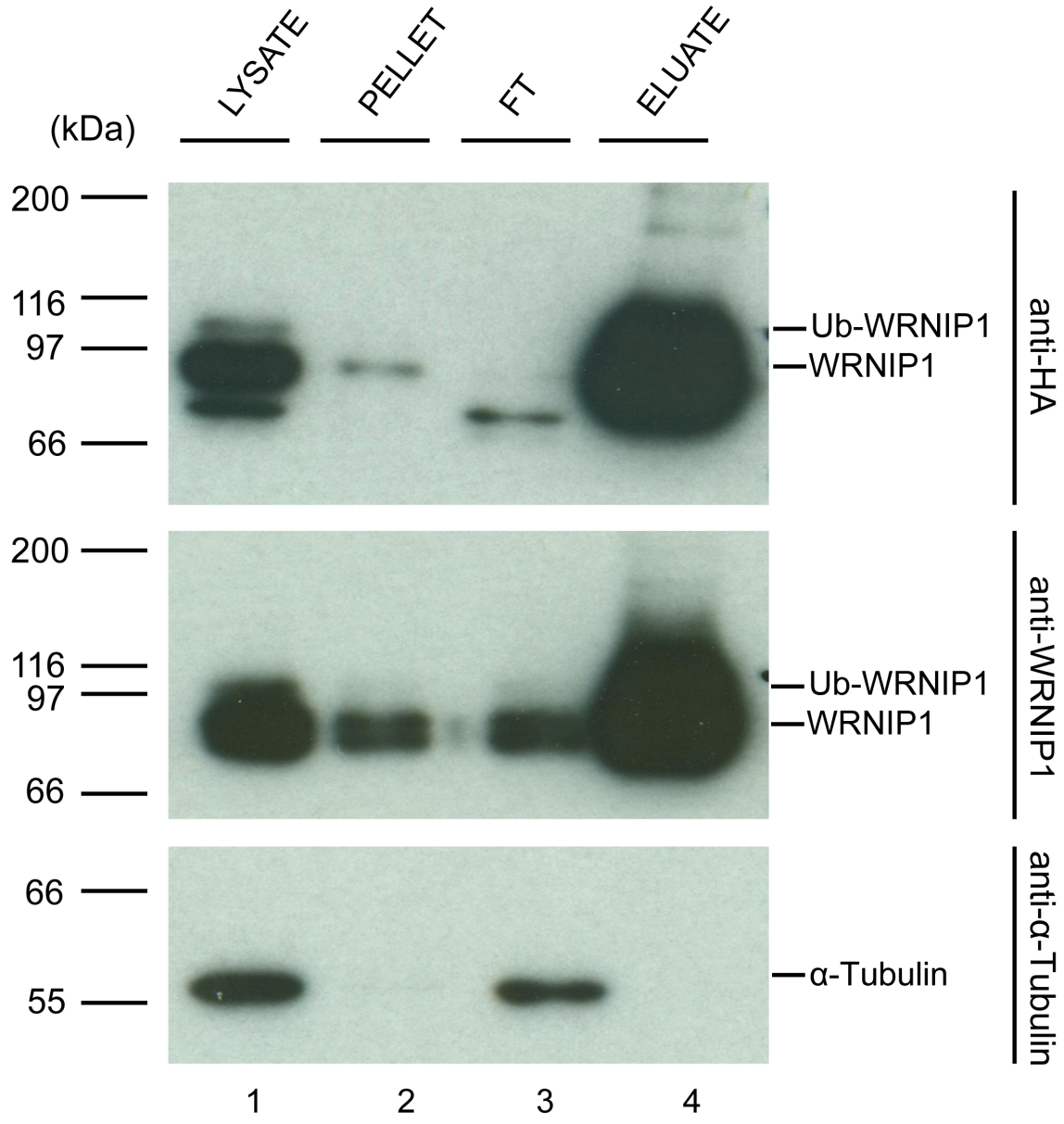


Figure 25: Western Blot analysis of the FLAG-HA-WRNIP1 purification. Samples taken during FLAG-HA-WRNIP1 purification from the HeLa S3 cells were analysed with WB assay. Three 15 cm dishes, with 80% confluency of Hela S3 cells were harvested for the experiment. The expressed protein was bound to the anti-FLAG beads, washed with the buffer and eluted and to test if the purification was successful initial lysate, insoluble faction (pellet), flow-through (FT) and eluate samples were analysed by Western Blot using anti-HA antibody, anti-WRNIP1 antibody and anti-Tubulin antibody 10 µg of lysate and flow through (FT) were loaded next to 10% of the pellet and 10 µl of the eluate. 9% SDS-PAGE gel was used for the WB analysis.

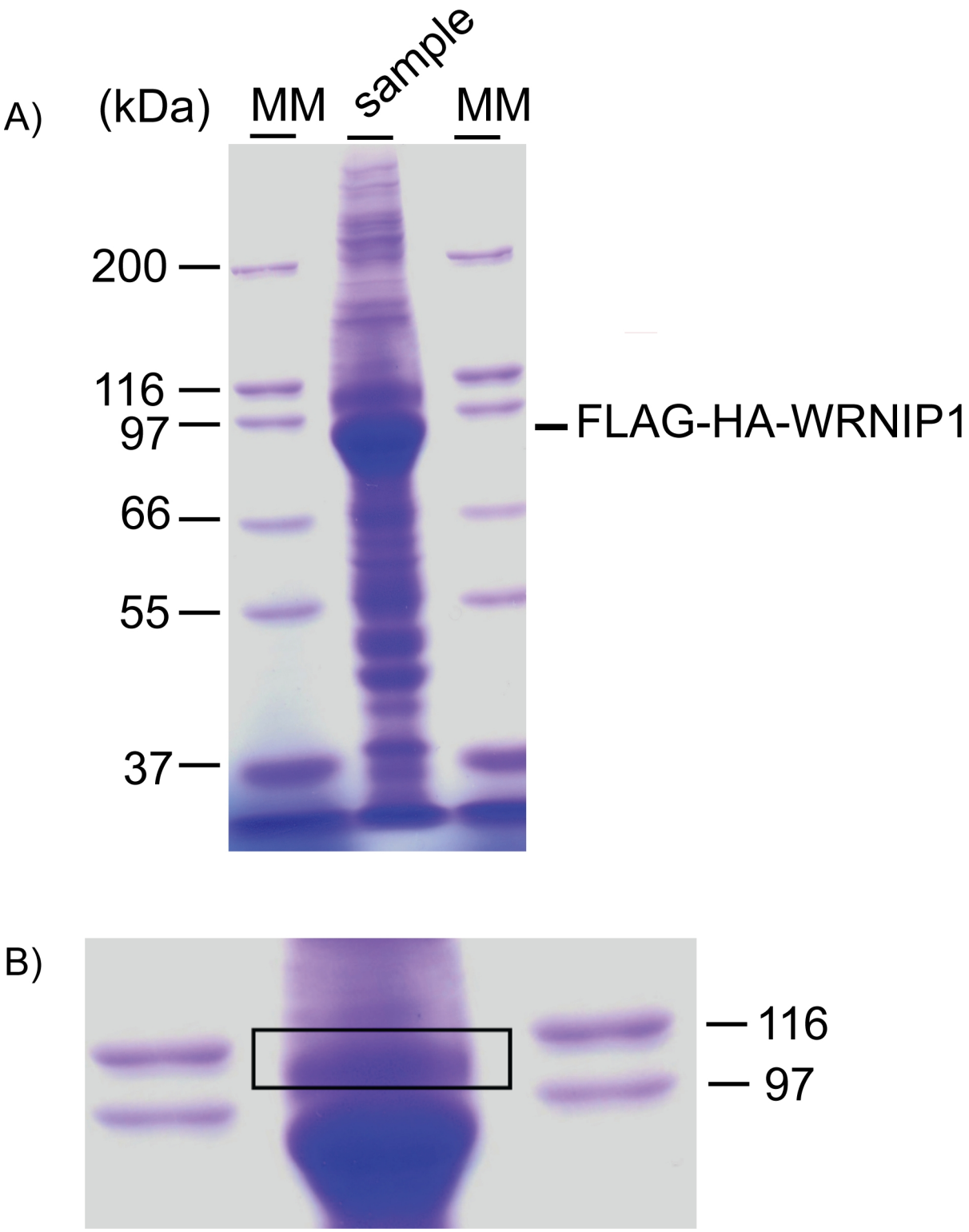


Figure 26: Purification of FLAG-HA-WRNIP1. A) All of the sample (purification of which was presented in Figure 25: the expressed from HeLa S3 cells protein was bound to the anti-FLAG beads, washed with the buffer and eluted) was loaded on the 9% SDS-PAGE gel and stained with Coomassie Blue dye. B) WRNIP1 upper form fragment was enlarged to show the exact fragment of the gel that was analysed by mass spectrometry. Black rectangle indicates the part that was cut out.

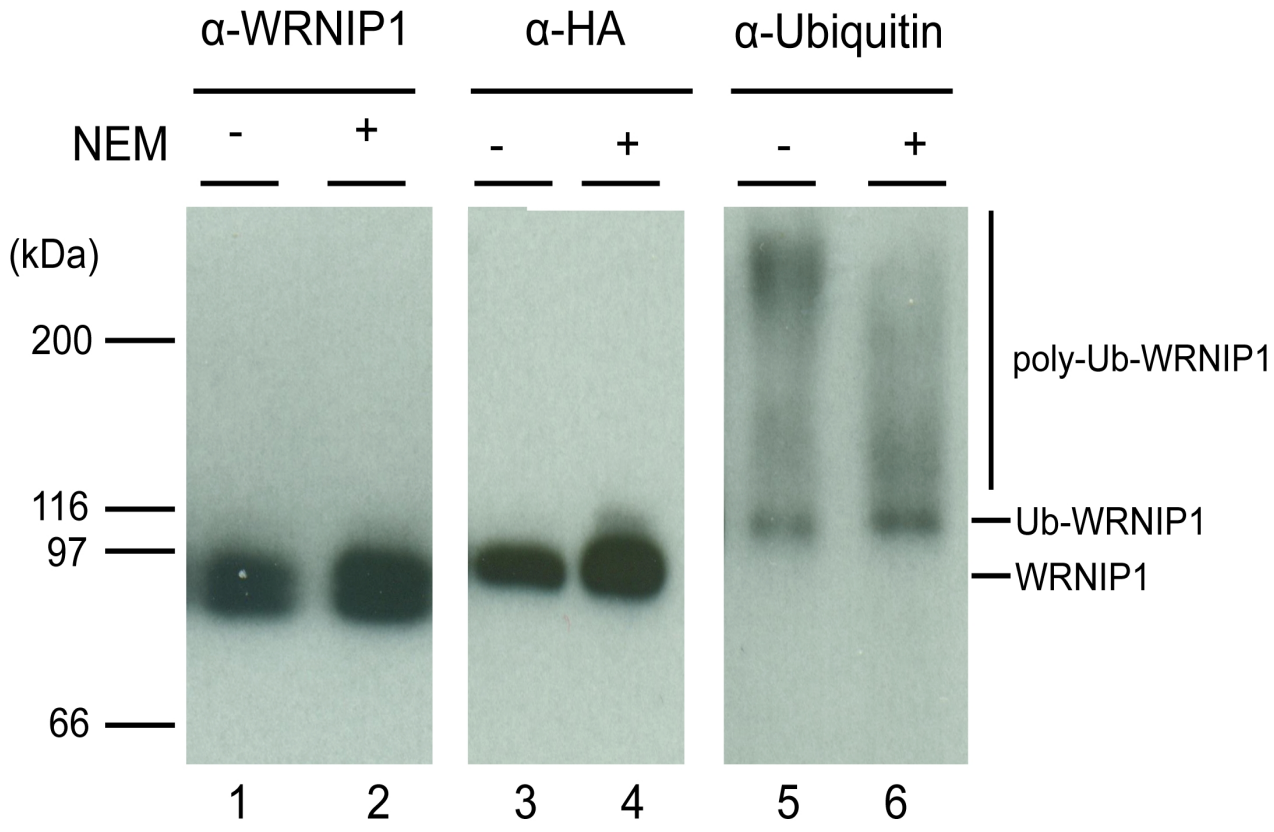


Figure 27: The upper form of WRNIP1 is its multi- or polyubiquitinated form. FLAG-HA-WRNIP1 was purified with and without NEM, and the results were analysed by WB analysis against anti-WRNIP1, anti-HA and anti-Ubiquitin antibodies. WRNIP1 was purified from HeLa S3 cells (one 80% confluent 15 cm dish for each purification), using the buffers supplemented with 10 mM NEM, at the same time the control cells were purified in the buffer unsupplemented with NEM. In both cases the expressed protein was bound to the anti-FLAG beads, washed with the buffer and eluted. Eluate samples were analysed by Western Blot using anti-HA antibody, anti-WRNIP1 antibody and anti-Ubiquitin antibody. 9% SDS-PAGE gel was used for the WB analysis.

3.3.3 Mass spectrometry results

Mass spectrometry was performed using the so called “bottom-up“ approach, where the proteins prior to the analysis are enzymatically digested into smaller peptides using trypsin, and then identified using tandem mass spectrometry (also known as MS/MS). During this process the peptides are ionized and then separated depending on the charge to mass ratio, which allows a very precise identification of different peptides and the modifications they underwent, while still being a

part of WRNIP1.

All the mass spectrometry data are analysed using the Mascot software available through the Oxford University Proteomics Facility.

Surprisingly the mass spectrometry analysis identified not one but eleven ubiquitinated lysines in the WRNIP1 sequence. Each of the hits had its own score, which was high or low depending on how certain the hit was, and also each of them had its own expect value (E-value), which represents the probability that given hit was obtained completely randomly, and is actually not significant. E-values could vary anywhere between 0 and 1, where 0 would mean that there is 100% chance the hit was real and 1 would mean 100% chance the hit is just an artifact. The threshold for E-value that is low enough, is of course arbitrary, but for the purpose of our analysis, we decided to reject all hits with E-value higher than 10^{-3}.

Additionally, we also took into consideration how many times given residue was detected (how many peptides containing particular residue were reported) and in how many of them the analysed residue was ubiquitinated.

In the end we decided to concentrate only on the ubiquitination sites that looked most promising given the combination of values that the Mascot server provided. The best hit was K301 with three ubiquitinated peptides, one of which gave a very high score (over one hundred) and E-value as low as $1.5 \cdot 10^{-10}$. The other ones with relatively high score and low E-value were residues K70 and K310 (**Table 11**).

RESIDUE	NUMBER OF IDENTIFIED UBIQUITINATED PEPTIDES/ ALL DETECTED PEPTIDES WITH GIVEN RESIDUE	SCORES	E-VALUES
K62	2/3	hit 1) 5 hit 2) 12	hit 1) 0.31 hit 2) 0.32
K70	1/2	51	$3.3 \cdot 10^{-5}$
K287	2/15	hit 1) 23 hit 2) 18	hit 1) 0.0064 hit 2) 0.019
K301	3/11	hit 1) 105 hit 2) 27 hit 3) 16	hit 1) $1.5 \cdot 10^{-10}$ hit 2) 0.0031 hit 3) 0.03
K310	1/6	33	$8.6 \cdot 10^{-4}$
K322	1/9	1	0.87
K464	2/5	hit 1) 26 hit 2) 28	hit 1) 0.018 hit 2) 0.011
K482	1/18	18	0.019
K601	1/4	25	0.0042
K627	1/4	33	$7.9 \cdot 10^{-3}$
K661	1/4	21	0.056

Table 11: All ubiquitinated residues identified by the Mascot software, with indicated ratio of ubiquitinated peptides to non-ubiquitinated ones, score and E-values for each hit. The residues highlighted with green, are the ones we decided to further investigate, based on the combination of relatively high score and low E-value. Some residues (K62, K287, K301 and K464) were detected in more than one peptide. These residues have E-value and scores indicated separately for each hit.

3.3.4 Point mutation analysis

After identifying three potential ubiquitination sites in the mass spectrometry analysis (K70, K310 and K310), we decided to introduce them into our WRNIP1-EGFP construct in order to test whether abrogating the ubiquitintion in the indicated residues affects the phenotype of WRNIP1.

In addition to the residues that we found in our mass spectrometry we decided to include several residues that were previously identified in the literature (*Bish et al. 2008*) as potential

ubiquitination sites. The experiment performed by the authors of the mentioned paper was quite different from the one we designed, because all ubiquitinated species (including the ones with very long ubiquitin chains) were tested, while we just analysed the upper form of WRNIP1 – so only WRNIP1 forms lighter than 116 kDa were purified and sent to mass spectrometry analysis. Still we decided that testing the lysines that were most often ubiquitinated according to their analysis maybe worthwhile. The lysines with highest number of identified ubiquitinations were K301 (20 identified peptides), K335 (29 identified peptides), K633 (14 identified peptides) and K636 (14 identified peptides). Since K301 was already identified in our own analysis, only three additional residues were introduced to the set of point mutations we decided to introduce to the WRNIP1-EGFP sequence.

The final set of mutants we designed consisted of: K70R, K301R, K310R, K335R, K633R and K636R (**Figure 28**).

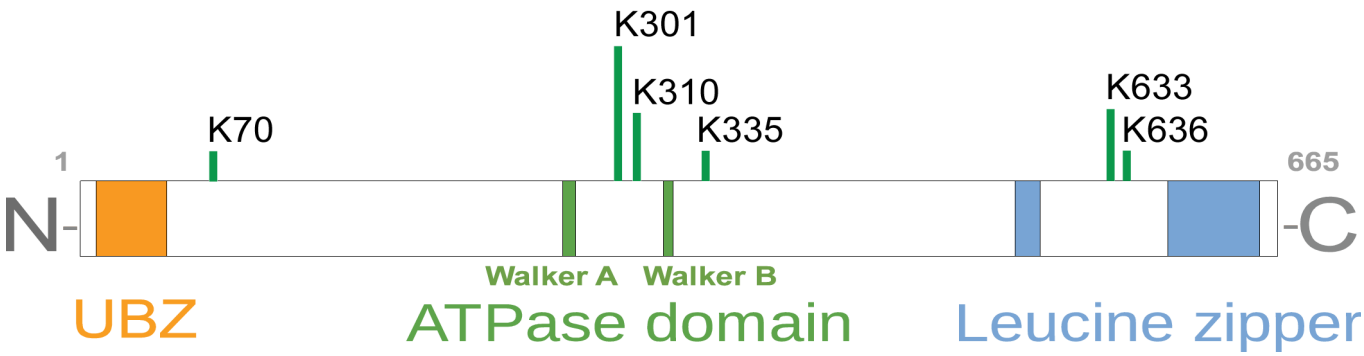


Figure 28: Schematic of all the lysine mutations introduced into the WRNIP1 sequence. The orange, green and blue parts represent accordingly: Ubiquitin binding zinc finger domain (UBZ), ATPase domain and Leucin Zipper at the C-terminus of WRNIP1.

We introduced these point mutations into the WRNIP1-EGFP construct, and then used it to transfect the HeLa Flp-in T-REx cell line, with previously disrupted WRNIP1 gene, using the CRISPR/Cas9 technique (see Materials and Methods).

The clones obtained in such a way were then expressing only exogenous WRNIP1 tagged with EGFP, and provided we treated the cells with doxycyclin under the conditions we previously established we were able to obtain it at levels comparable with physiological.

For each of the mutants two independent clones were selected and they were compared against the wild type both in Western Blot analysis and live-cell microscopy imaging with the previously described TMP/UVA treatment, where the cells are first sensitized with TMP, then irradiated with 405 nm laser, which results in ICL formation.

The first of the mutants: K70R has shown an interesting difference in phenotype compared to the wild type protein: it seemed that the protein was not properly localized to the nuclei (**Figure 29 A**). Interestingly, the WB analysis of the mutants (the cells were treated with dox for 3 h to achieve physiological levels of WRNIP1 and then harvested, lysed, the lysates were mixed with the LDS loading buffer and loaded on the 7% SDS-PAGE gel) has shown that the levels of WRNIP1 degradation were higher (**Figure 29 B**).

Since K70R was not even localized in the nuclei it was impossible to tell whether or not it is being recruited to the DNA damage site. The bioinformatic software: cNLS Mapper (*Kosugi et al., 2008; Kosugi et al., 2009*) predicted that K70 is part of the nuclear localization sequence (NLS), which starts at residue P66 and ends at residue E76 which suggests that rather than being a result of K70 not being ubiquitinated, lack of WRNIP1-EGFP in the nuclei may be caused by disrupting NLS. To further investigate the effects of the mutation at K70 residue it would be necessary to attach the NLS signal either at the N- or C-terminus of WRNIP1 and then introduce the mutation at the lysine 70, to ensure the correct localization of the protein despite the mutated NLS sequence.

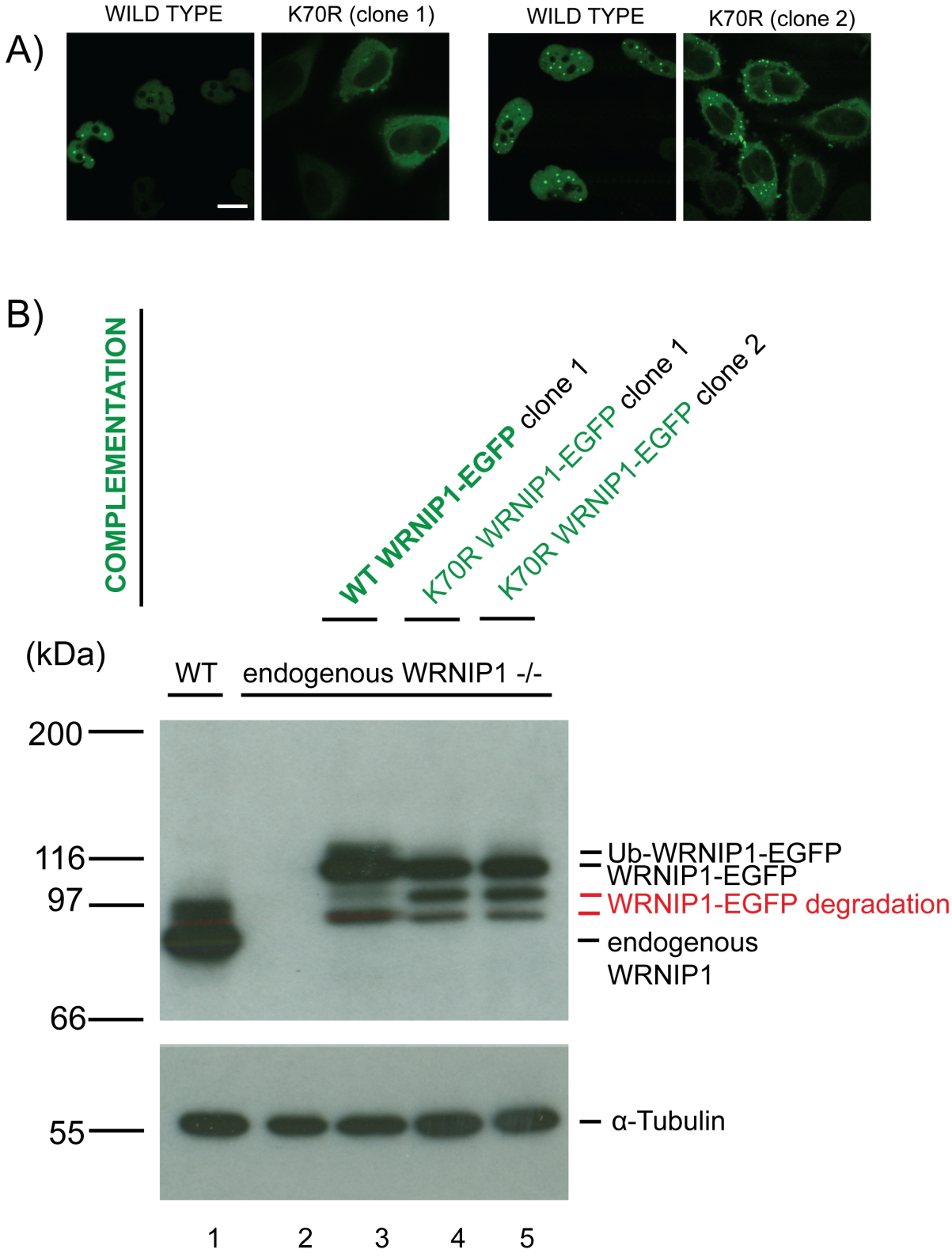


Figure 29: K70R mutant is not properly localized into the nucleus. Two different clones were tested to exclude the possibility that it is caused by some kind of a random clonal effect. A) microscopy images. Wild type WRNIP1-EGFP is shown for comparison (the protein is visibly contained in the nuclei), while in case of the K70R mutant, the EGFP-tagged proteins are abundant in cytoplasm and almost not present inside the nuclei, the scale bar is equivalent to 10 μm. Note that the images of the control cells used for the right panels, were acquired during the same experiment session as the images presented in Figure 18 and also later in Figure 38. B) WB analysis, using 9% SDS-PAGE gel.

K301R, K310R and K335R WRNIP1-EGFP mutants have shown a phenotype very similar to the wild type protein. Just as the wild type they had visible foci in some cells, and after ICLs were introduced, WRNIP1 was recruited to the damage site within the first 3-5 min (**Figure 30**). It confirms that none of these three ubiquitination sites alone, are necessary for WRNIP1 recruitment to ICLs.

K633R and K636R gave similar results (**Figure 31**). Both were recruited to the damage site after introducing ICL damage. In both cases the abundance of the recruited protein was comparable with the wild type, as well as timing. Both mutants were recruited between 3 and 5 min, which is a typical pattern for the wild type protein as well.

Additional WB analysis of all these lysine mutants (K301R, K310, K335, K633 and K636), has shown that all of them have upper ubiquitinated form (**Figure 32**). This suggests that none of these residues alone, are necessary for the existence of ubiquitinated WRNIP1.

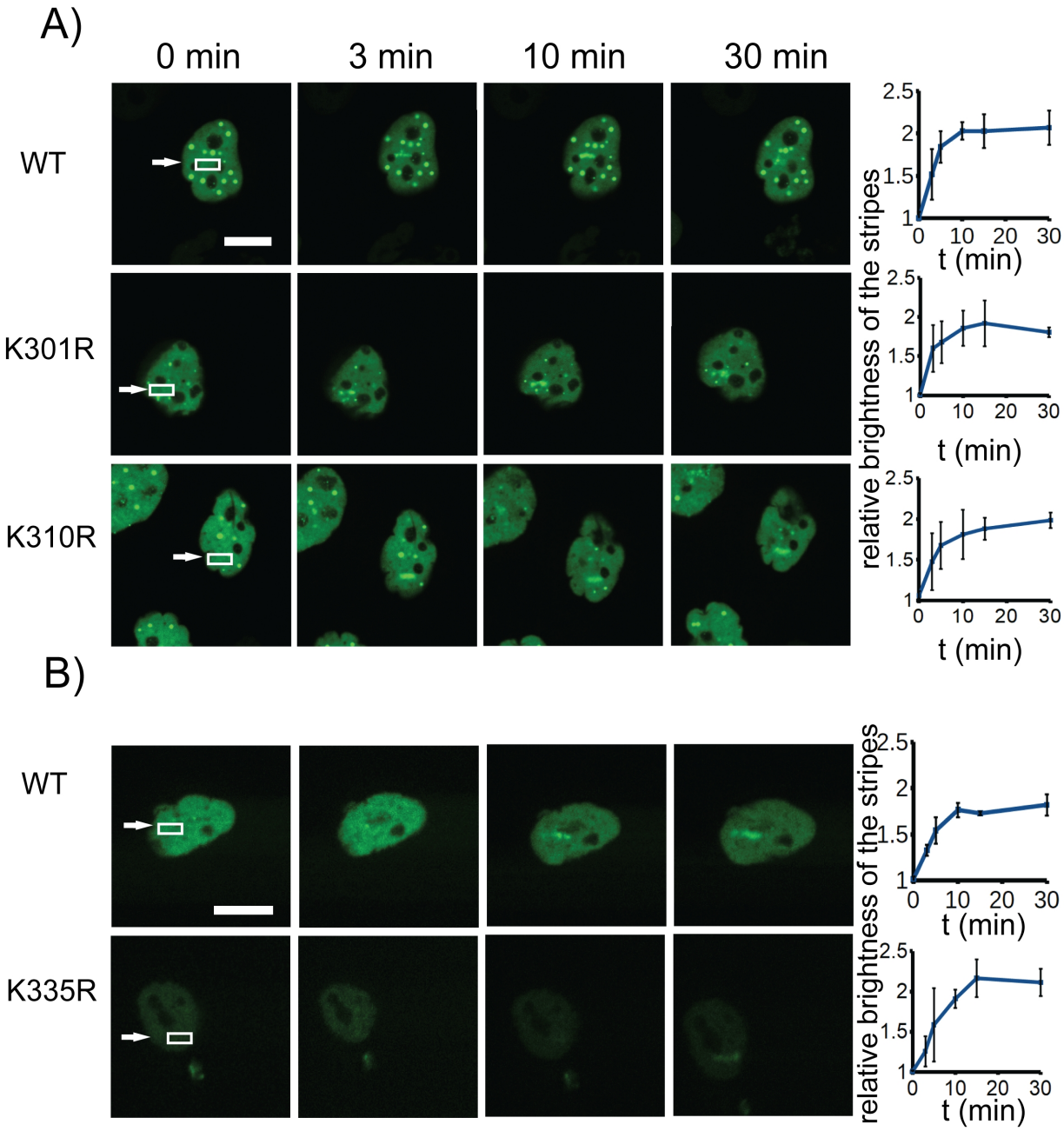


Figure 30: K301R, K310R and K335R WRNIP1 mutants are recruited to the damage site. The cells expressing mutated WRNIP1 were irradiated with 405 nm laser after 15 min incubation with TMP. The white arrows indicate where the cells were irradiated, while the graphs on the right side show the quantification of the recruitment, of each mutant. The scale bar is equivalent to 10 μm. A) K301R and K310R mutants tested with WT WRNIP1-EGFP as control, using 60x objective (the acquisition was performed during the same experiment as for mutant K636R (Figure 31), therefore the control is identical); B) K335R tested with WT WRNIP1-EGFP as control, using identical conditions but with 40x objective. Two independent experiments were quantified for each clone (N=2) and same two independent experiments were used to quantify the control. In total: 7 cells for K301R, 6 cells for K310R, 6 cells for K335R, 5 cells for the control in the upper panel and 6 cells for the control in the lower panel.

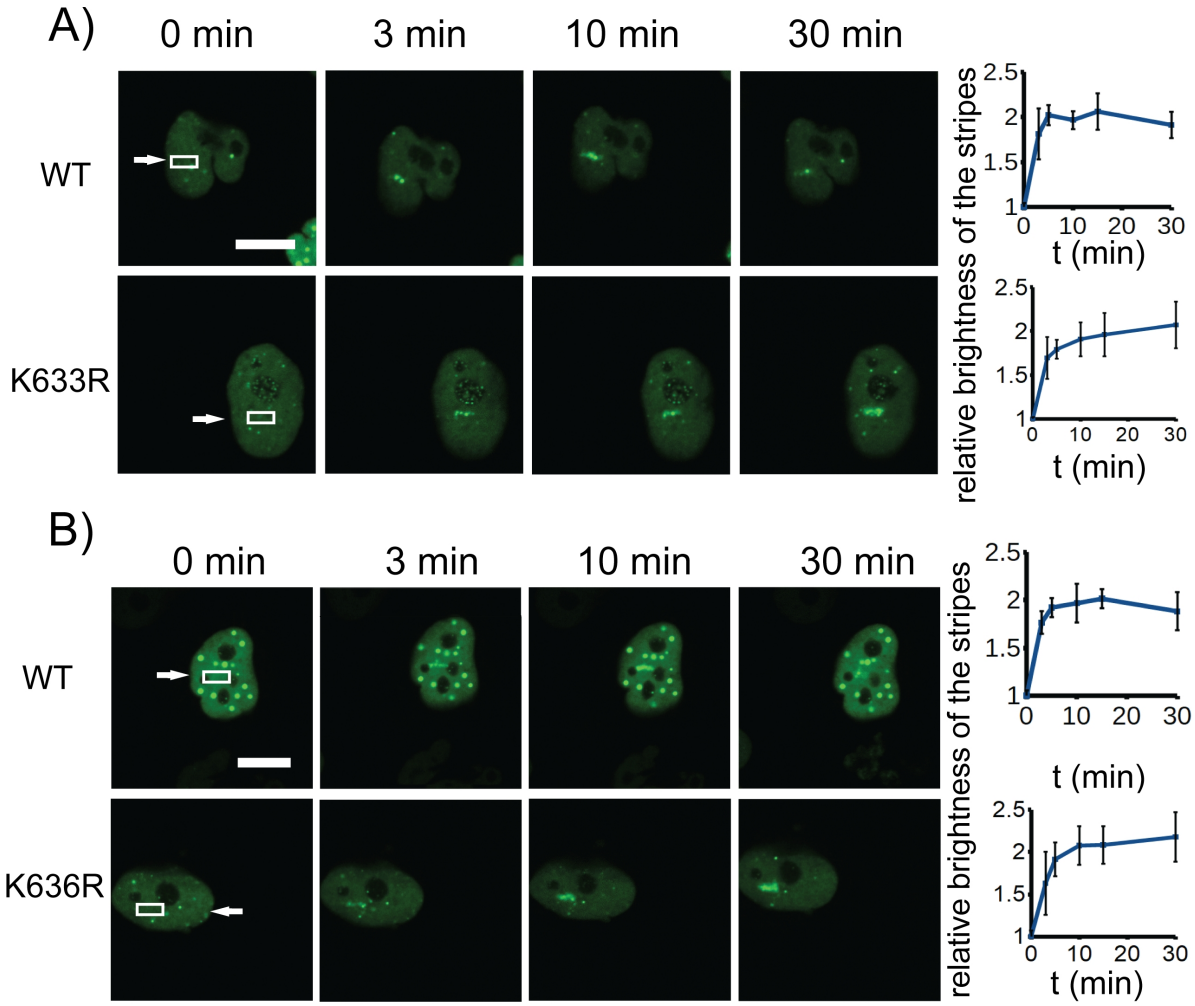


Figure 31: K633R and K636R WRNIP1-EGFP mutants are recruited to the damage site. The cells expressing mutated WRNIP1 were irradiated with 405 nm laser after 15 min incubation with TMP. The white arrows indicate where the cells were irradiated, while the graphs on the right side show the quantification of the recruitment, of each mutant. The scale bar is equivalent to 10 μm. For both experiments 60x objective was used A) K633R recruitment, B) K636R recruitment (the acquisition was performed during the same experiment as for mutants K301R and K310R (Figure 30 A), therefore the control is identical). Two independent experiments were quantified for each clone (N=2), 6 cells were quantified for K633R, 5 for K636R, and same two independent experiments were used for each control cells (6 cells for the upper panel and 8 cells for the lower panel).

3.3.5 Discussion

WRNIP1 migrates on SDS-PAGE gel in two species: upper and lower. Upper is approximately ten times less abundant than lower, and is some form of ubiquitinated WRNIP1, as it is recognized by anti-ubiquitin FK2 antibody. Above upper form there is also a smear of higher molecular weight ubiqutitinated WRNIP1, which raised the possibility that maybe *in vivo* WRNIP1 is modified with longer ubiquitin chains but during the cell lysis, when all the deubiquitinating enzymes are released into the solution, the long chains are cleaved off and the upper form is just a result of this process. The purification performed in the presence of NEM (the inhibitor of deubiquitinating enzymes) excluded such possibility, because the upper form was present in both samples, the one treated with NEM and the control purified without the inhibitor.

In order to further investigate the upper form it was purified and analysed by mass spectrometry, which suggested several potential ubiquitination sites. The most likely candidates seemed to be either K70, K301 or K310. We also decided to test three other residues known to be ubiquitinated from literature: K335, K633 and K636. The experiments with 405 nm irradiation after TMP sensitization has shown that the abrogation of ubiquitination at K301, K310, K335, K633 or K636 alone is not enough to abrogate WRNIP1-EGFP recruitment to the damage site, which means that none of the mentioned ubiquititnations are necessary for WRNIP1 ubiquitination.

K70 was not localized to the nuclei, as it is one of the residues that form the nuclear localization signal.

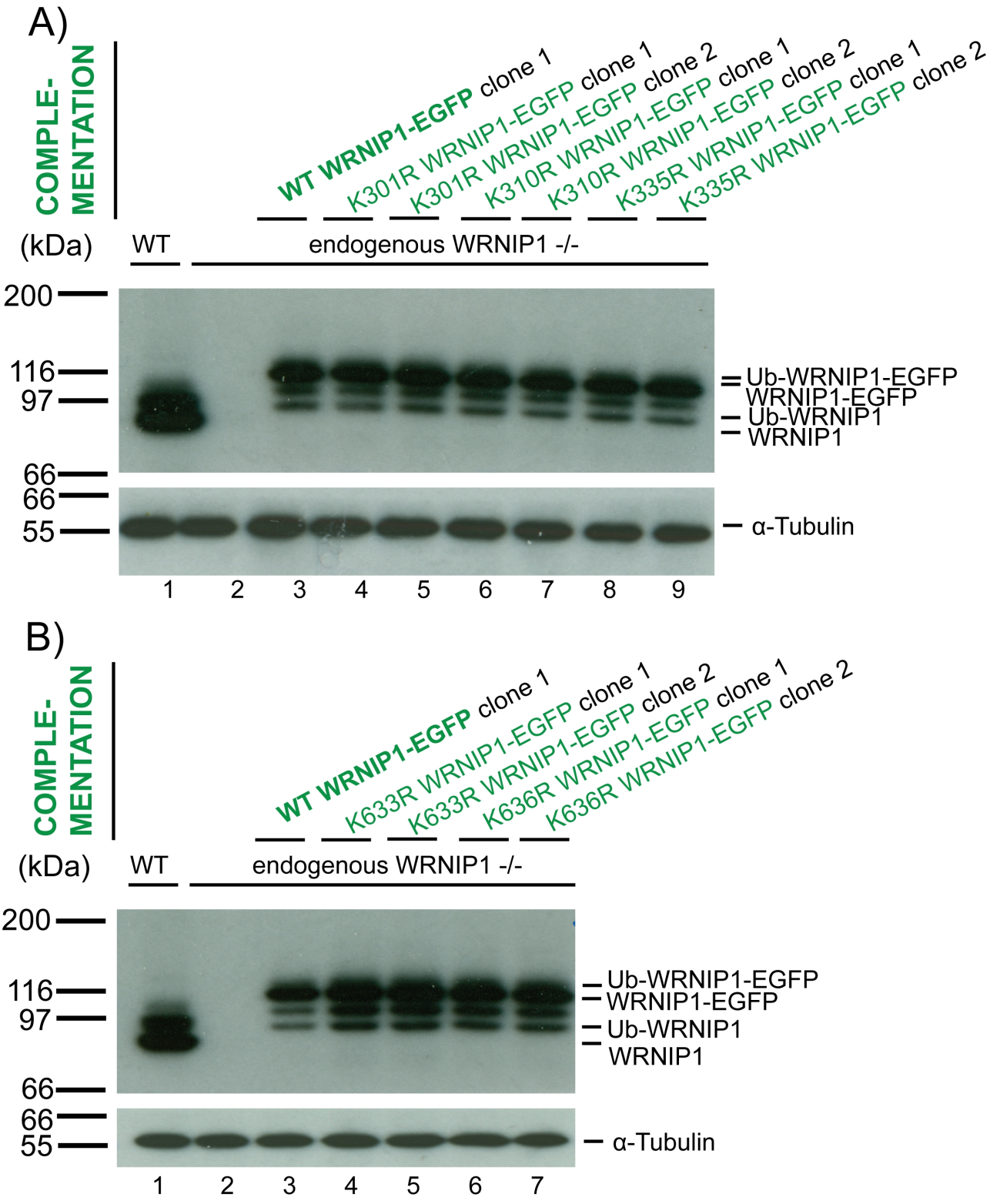


Figure 32: WB analysis of the lysine point mutants. Each cell line was treated with doxycyclin for 3 hours and then harvested A) K301, K310 and K335 clones were tested next to wild type cells, WRNIP1 $^{-/-}$ cell line and the cell line expressing WT WRNIP1-EGFP A) K633 and K636 clones were tested next to wild type cells, WRNIP1 $^{-/-}$ cell line and the cell line expressing WT WRNIP1-EGFP. 7% SDS- PAGE gel was used during the WB analysis for the anti-WRNIP1 analysis, and 9% gel for the anti-tubulin analysis.

3.4 Analysis of the domains of WRNIP1

3.4.1 Designing point mutations to disrupt the activity of three WRNIP1 domains

In order to further investigate the importance of each domain present in WRNIP1 we designed a set of point mutations (**Figure 33**). Every point mutation we designed was supposed to abrogate the activity of a chosen domain.

First we used three potential polar interaction with ubiquitin sites – N33, D37 and A45 (described previously in chapter 1) to mutate them into A33, A37 and V45. While D37 was previously known in literature to abrogate the activity of UBZ domain, N33 and A45 seemed like good candidates to potentially allow us to confirm this result.

The next domain we wanted to investigate was ATPase domain. Since it is divided into Walker A and Walker B motifs we decided to disrupt each of the motifs separately.

Walker A has a consensus of G-XXXX-GK-[TS]. In case of WRNIP1 it is located between 268 and 275 residues with the sequence: GPPGCGKTT (the underlined K274 was mutated):

CONSENSUS:	**5' – G X X X X G K [TS] – 3'**
WRNIP1 (268-275 aa):	**5' – G P P G C G <u>K</u> T – 3'.**

The lysine (K) residue that is always at the seventh position within the motif (in case of WRNIP1 it is K274) is crucial for nucleotide binding (***Hanson and Whiteheart, 2005***), therefore it seemed like a good candidate to abrogate the ATPase activity via the Walker A motif.

The Walker B motif consists of a stretch of bulky hydrophobic amino acids (at least four) followed by aspartic acid and then glutamic acid (hhhhDE). The aspartate residue co-ordinates magnesium ions, and the glutamate is essential for ATP hydrolysis (***Hanson and Whiteheart, 2005***). A E329Q mutation within the Walker B motif was supposed to abrogate the ATPase activity, without disrupting the structure of Walker A motif – so with the least amount of potential side effects for the protein structure (the underlined E329 residue was mutated):

CONSENSUS:	**5' – h h h h D E – 3'**
WRNIP1 (324-329 aa):	**5' – I L F I D <u>E</u> – 3'.**

To abrogate the activity of the two Leucine Zipper motifs at the C-terminus, we targeted two

leucines: one in the first short leucine zipper motif: L526, and the second one in the second leucine zipper motif: L625, mutating them into accordingly: L526P and L625P, as proline has the biggest potential to disrupt the structure of the leucine zipper due to its side chain.

The mutated constructs were then transfected into HeLa Flp-in T-REx cell line with the endogenous WRNIP1 depletion, obtained using the CRISPR/Cas9 technique, so using the right timing and doxycyclin concentration we were able to obtain levels of mutated proteins comparable with endogenous.

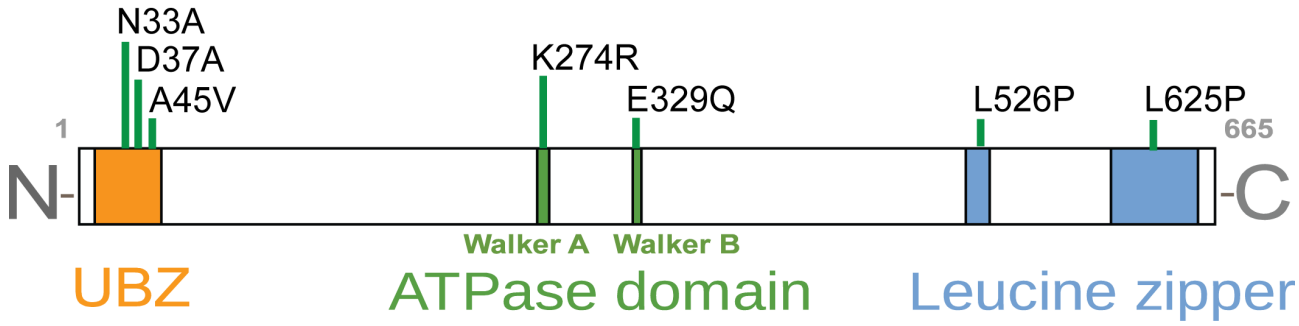


Figure 33: Point mutations designed to disrupt the activity of each domain: UBZ, ATPase and Leucine zipper. The green indicators above the schematic represent sites where each of the mutations was introduced.

3.4.2 Point mutations in the crucial residues of the UBZ domain, abrogate the ubiquitination of WRNIP1 as well as its recruitment to the ICL damage site

The N33A and D37A mutants had no upper form visible during Western Blot analysis (**Figure 34 lanes 3, 4, 5, 6 and 9**), confirming that these two mutations abrogate WRNIP1 ubiquitination.

To establish whether UBZ domain activity is needed for WRNIP1 recruitment to the damage site, we sensitized the HeLa Flp-in T-REx cells expressing WRNIP1-EGFP(N33A), WRNIP1-EGFP(D37A) and WRNIP1-EGFP(A45V) with TMP, and then we irradiated them with 405 nm laser.

In case of N33A mutant as well as D37A there was no observable recruitment (**Figure 34**). The mutated proteins seemed to not form any visible foci nor aggregates and were instead

distributed uniformly in the nucleus.

A45V did have normal levels of upper ubiquitinated form (compared to the wild type WRNIP1-EGFP) (**Figure 34 lanes 7, 8 and 9**). And its level did not seem to be lower nor higher than the level observed in the wild type WRNIP1.

While sensitized with TMP and irradiated with the 405 nm laser A45V mutant protein was recruited to the ICL damage site and again the kinetics of this recruitment were not significantly different from the kinetics of the wild type WRNIP1-EGFP (**Figure 35**).

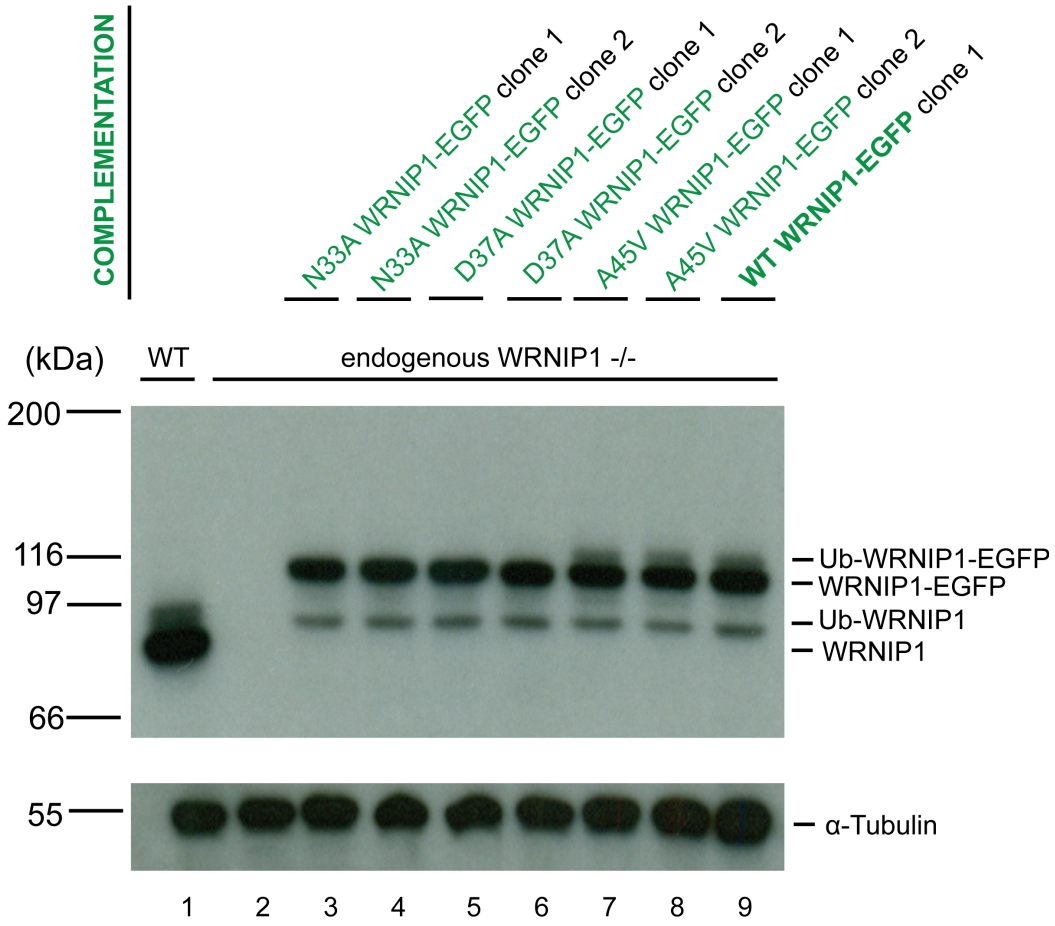


Figure 34: D37A and N33A WRNIP1-EGFP point mutants do not have upper ubiquitinated form. A45V has normal levels of upper form. In the first lane wild type (WT) HeLa Flp-in T-REx cells. In lane 2 the same cell line but with endogenous WRNIP1 gene disrupted using CRISPR/Cas9 technique. Lanes 3-8 are presenting point mutants of WRNIP1-EGFP expressed after 3 h of a standard dox treatment, and lane 9 wild type WRNIP1-EGFP. 7% SDS- PAGE gel was used during the WB analysis for the anti-WRNIP1 analysis, and 9% gel for the anti-tubulin analysis.

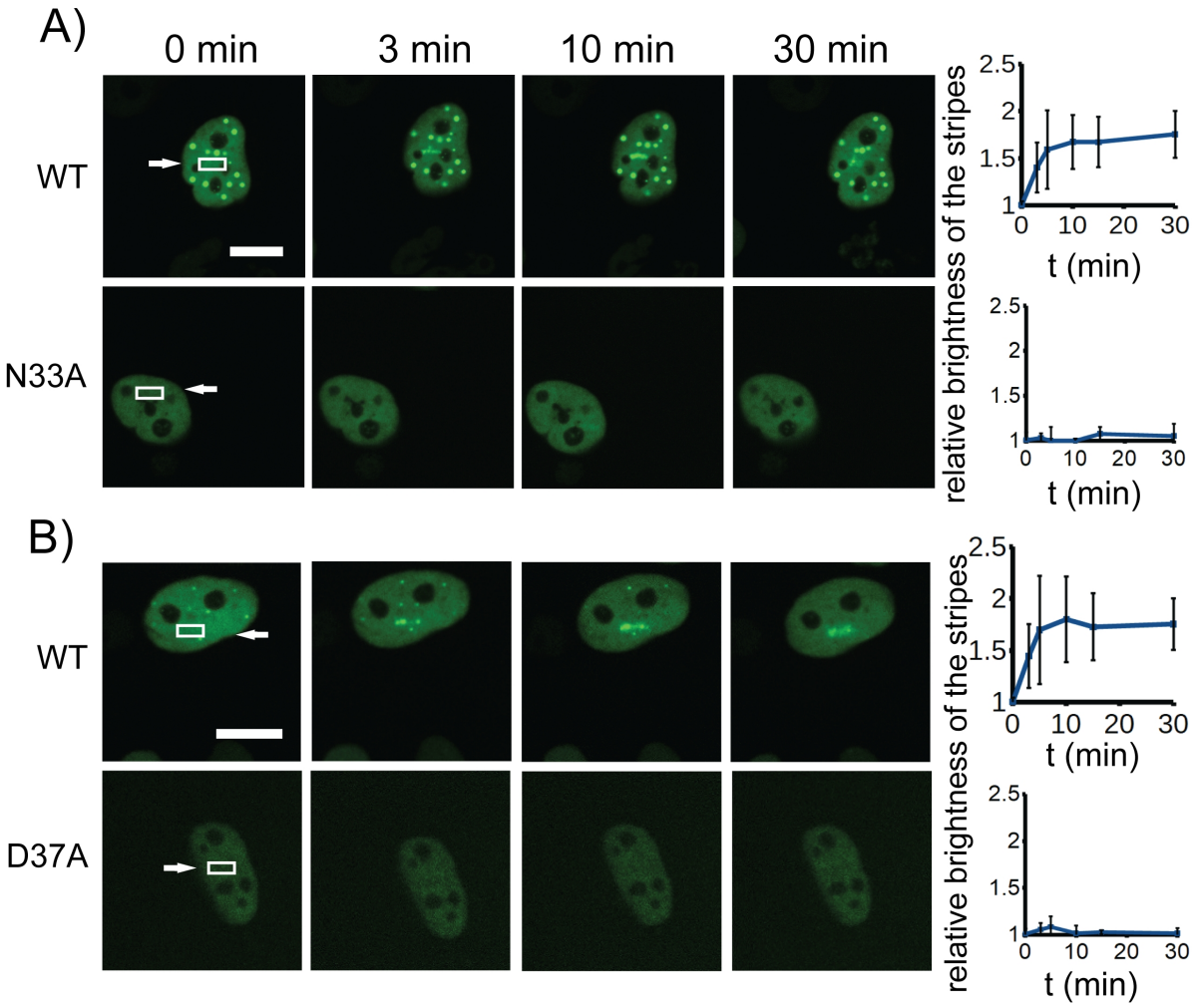


Figure 35: Both N33A and D37A mutations abrogate WRNIP1 recruitment to the ICL damage site. The cells expressing mutated WRNIP1 were irradiated with 405 nm laser after 15 min incubation with TMP. The upper panel presents the microscopy data of N33A, the lower of D37A. Both experiments were performed using the same microscope and the same laser intensity but N33A images were acquired using the 60x objective while D37A using 40x objective. The scale bar is equivalent to 10 µm. The acquisition of the N33A mutant was performed during the same experiment as for mutants K301R, K310R and K636R (Figures 30 and 31), therefore the control is identical. The same applies to the D37A mutant that shared control with K274R mutant (Figure 36). 2 independent experiments were quantified for each cell line (N=2), 9 cells for N33A and 5 cells for D37A, and for the control: 6 cells for the upper panel and 5 cells for the lower.

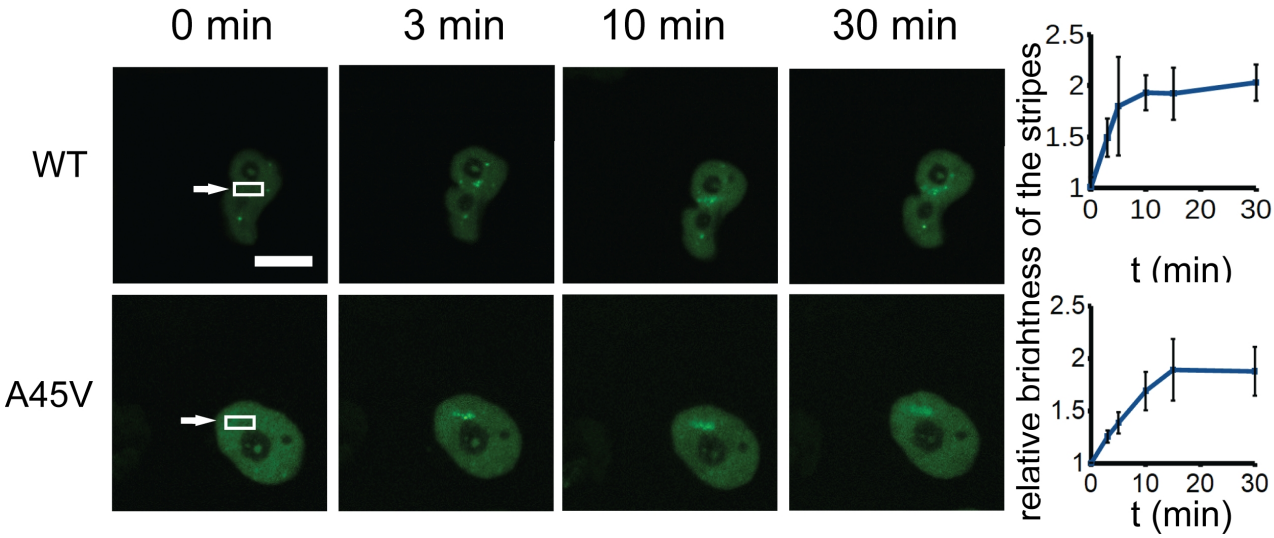


Figure 36: A45V point mutant was recruited to the ICL damage site. The cells expressing mutated WRNIP1 were irradiated with 405 nm laser after 15 min incubation with TMP. The kinetics of the recruitment are at the levels comparable with the wild type WRNIP1-EGFP. The pictures were acquired using the 40x objective. The scale bar is equivalent to 10 µm. Note that A45V mutant images were acquired during the same experiment as images corresponding to the L526P mutant, therefore both cell lines share control images (Figure 36). Two independent experiments were quantified (N=2) including 8 cells in total, and for the control: 6 cells in total were quantified.

3.4.3 Point mutations in Walker A and Walker B motifs do not abrogate the recruitment of WRNIP1 to the ICL damage site but do change the kinetics of the recruitment

Both K274R and E329Q WRNIP1-EGFP mutants did present the ubiquitinated (upper) form of WRNIP1 while tested using the Western Blot analysis **(Figure 37)**. Both of them also showed an altered band pattern on the WB that may be due to altered degradation.

While sensitized with TMP and irradiated with 405 nm laser, both K274R and E329Q were recruited to the ICL damage site but the kinetics of this recruitment was slightly different than in case of the wild type EGFP – the recruitment of both mutants seemed to be slower (Figure 36). The greatest difference was observable at early time points (up to 10 min) where the protein recruited to the ICL was significantly less abundant than in case of the wild type. In case of E329Q after 30 min the levels of its recruitment were comparable with the wild type, while K274R remained lower.

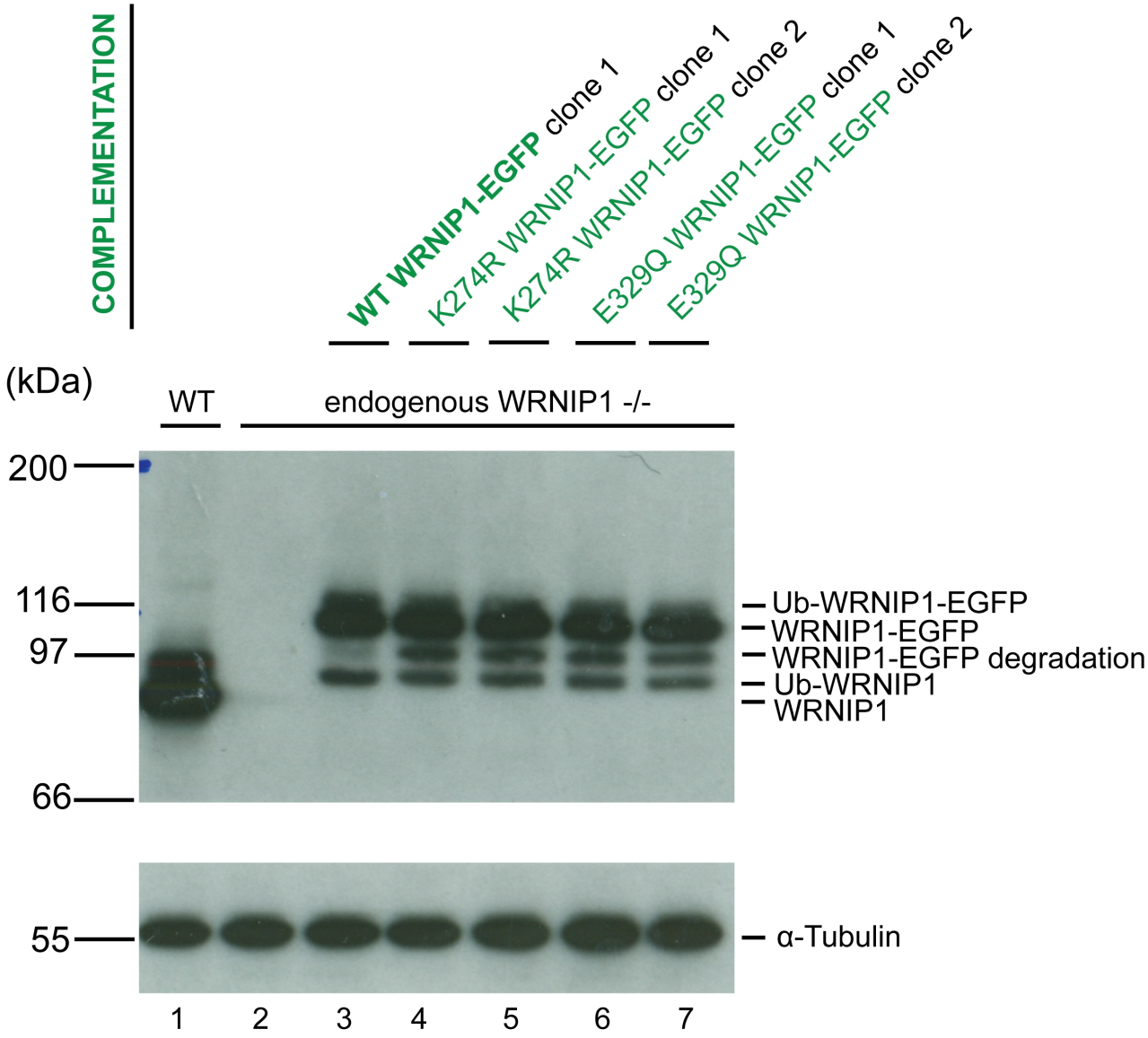


Figure 37: K274R and E329Q WRNIP1-EGFP point mutants both have upper ubiquitinated form, but also altered band pattern on the WB (possibly due to altered degradation). In lane 1 wild type (WT) HeLa Flp-in T-REx cells. In lane 2 the same cell line but with endogenous WRNIP1 gene disrupted using CRISPR/Cas9 technique. Lane 3 is presenting the WRNIP1 deficient cell line complemented with wild type WRNIP1-EGFP expressed under doxycyclin promotor and lanes 4-7 correspond to the same endogenous WRNIP1 deficient cell line complemented with point mutants of WRNIP1-EGFP that were then expressed under the same conditions as the wild type WRNIP1-EGFP shown in lane 3. 7% SDS- PAGE gel was used during the WB analysis for the anti-WRNIP1 analysis, and 9% gel for the anti-tubulin analysis.

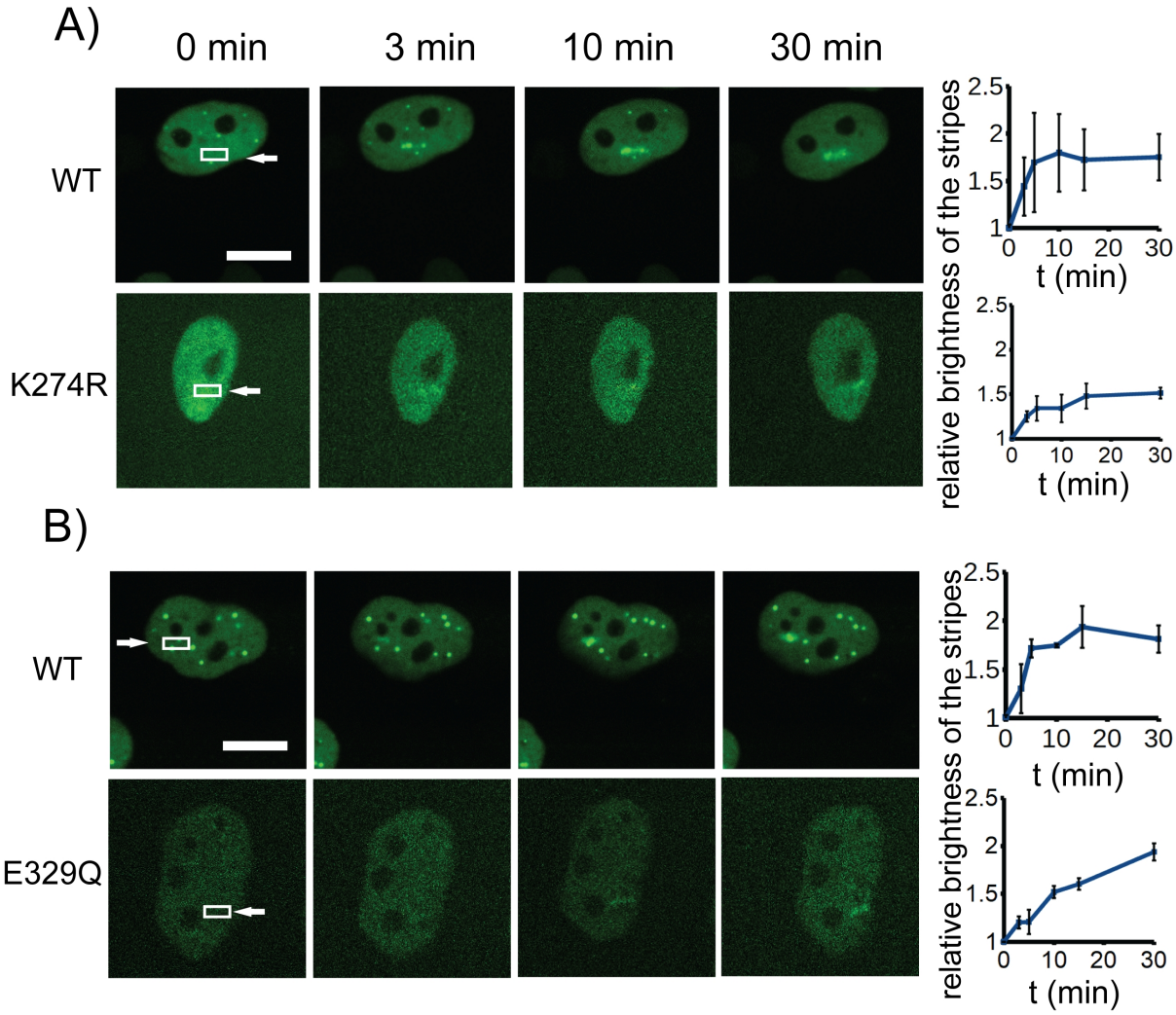


Figure 38: Both K274R and E329Q mutations still allow WRNIP1-EGFP recruitment to the ICL damage site but the kinetics are different to what is observed in the wild type WRNIP1-EGFP. The cells expressing mutated WRNIP1 were irradiated with 405 nm laser after 15 min incubation with TMP. The upper panel presents the microscopy data of K274R, the lower of E329Q. Both experiments were performed using the same microscope and the same laser intensity but K274R images were acquired using the 40x objective while E329Q using 60x objective. The scale bars are equivalent to 10 μm. Note that the images of the control cells corresponding to the E329Q mutant, were also used in Figure 16 and Figure 27 as they were all acquired during the same experiment and the same applies to the control cells corresponding to the K274R mutant, that were also used as control for D37A cells (Figure 35). 2 independent experiments were quantified for each cell line (N=2), 6 cells in total for each mutant. During the same two independent experiments control cells were analysed, 7 in the upper panel and 6 in the lower in total.

3.4.4 Point mutations in the predicted leucin zipper

At the C terminus WRNIP1 has two regions that may potentially form leucine zippers. Leucine zippers have a characteristic heptad leucine repeat: Leu-X6-Leu-X6-Leu-X6-Leu (where X may be any residue). The periodic repetition of leucine at every seventh position covers a distance of eight helical turns. However, many sequences contain the leucine repeat, but do not adopt the zipper structure. Since WRNIP1 has two regions that fit (or almost fit) the consensus we decided to mutate one of the leucines in the heptad in each of the two fragments to abrogate the formation of the structure in case the structure is actually formed. We mutated them into proline, as the structure of proline is the most disruptive to the secondary structure of the protein.

The first sequence that resembles leucine zipper consensus starts at leucine 519 and ends at 539. This fragment is not perfectly maintaining the consensus as there are only five amino acids between the last two leucines, while there should be six. But since the fragment resembles the heptad repeat and also all the remaining leucines have correct distance from each other we decided to not exclude the possibility that there is a structure that resembles leucine zipper and we introduced a mutation at the second leucine of the repeat (L526).

The second fragment maintains the ideal consensus with four leucines separated by six amino acids in between each of them. It starts at leucine 604 and ends at leucine 625. To disrupt the structure of this potential leucine zipper we introduced a mutation at the last position (L625P).

Both L526P and L625P (**Table 12**) mutants were transfected into HeLa Flp-in T-REx cells (with endogenous WRNIP1 depletion obtained using the CRISPR/Cas9 technique) as all the other cell lines described in this chapter and the same timing and doxycyclin concentration were used to obtain levels of proteins comparable with endogenous. While L625P has shown the phenotype very similar to the wild type protein (both on western blot (**Figure 39**) and when the cells were sensitized with TMP and irradiated with 405 nm laser (**Figure 40**), L526P has shown both very reduced levels of ubiquitination and very reduced recruitment to the ICL damage site.

FRAGMENT	SEQUENCE	MUTATION
CONSENSUS	**L**xxxxxx **L**xxxxxx **L**xxxxxx **L**	-
519-539	**L**YWLARM **L**EGGEDP **L**YVARR **L**	L526P
604-625	**L**RNHQGP **L**PPVPLH **L**RNAPTR **L**	L625P

Table 12: Comparison of the two potential leucine zipper structures to the leucine zipper consensus

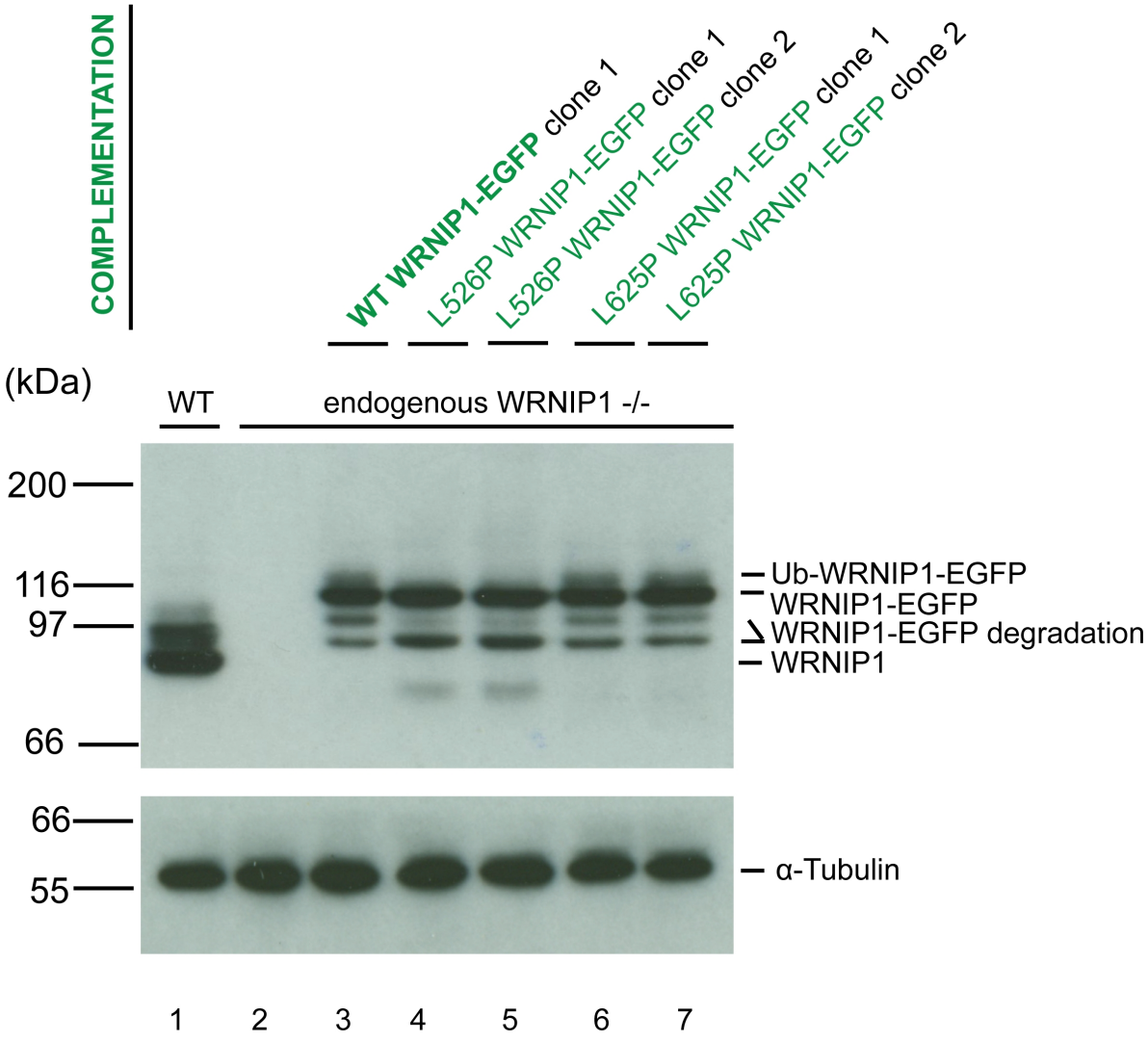


Figure 39: L526P and L625P WRNIP1-EGFP point mutants are expressed at the similar levels to wild type WRNIP1-EGFP and while L625 has normal levels of upper ubiquitinated form, it seems like L526P has very low levels of ubiquitinated upper species, and slightly different pattern of degradation. In lane 1 wild type (WT) HeLa Flp-in T-Rex cells. In lane 2 the same cell line but with endogenous WRNIP1 gene disrupted using CRISPR/Cas9 technique. Lane 3 is presenting the cell line with disrupted endogenous WRNIP1 gene complemented with wild type WRNIP1-EGFP expressed under doxycyclin promotor and lanes 4-7 correspond to the same gene disruption cell line complemented with point mutants of WRNIP1-EGFP that were then expressed under the same conditions as the wild type WRNIP1-EGFP from the lane 3. 7% SDS- PAGE gel was used during the WB analysis for the anti-WRNIP1 analysis, and 9% gel for the anti-tubulin analysis.

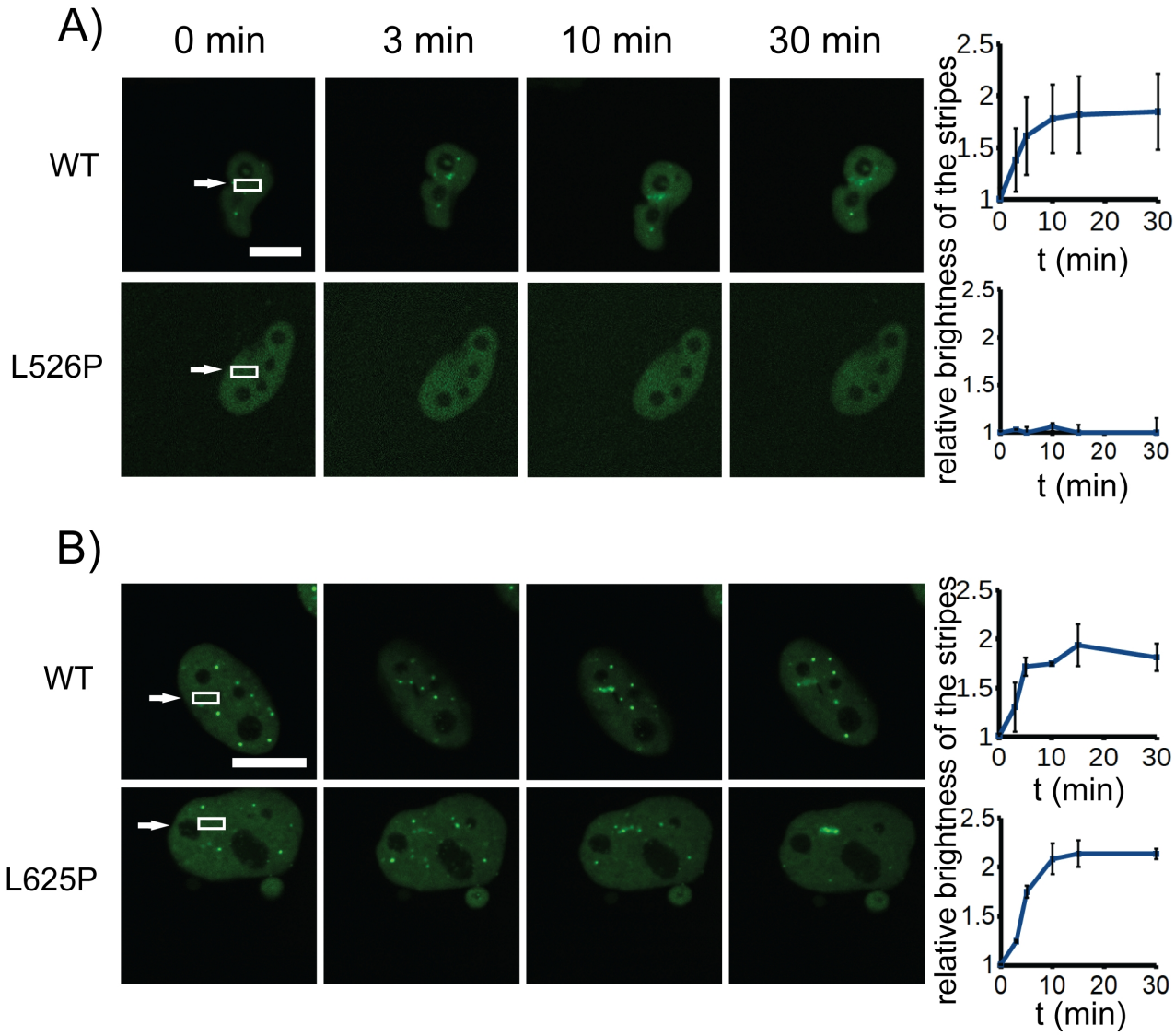


Figure 40: While L625P still allows WRNIP1-EGFP recruitment to the ICL damage site and the kinetics are comparable to what is observed in the wild type WRNIP1-EGFP, L526P has very low (if any) levels of recruitment to chromatin, which goes in line with our previous observation, that the levels of upper ubiquitinated form are very low. The cells expressing mutated WRNIP1 were irradiated with 405 nm laser after 15 min incubation with TMP. The upper panel presents the microscopy data of N33A, the lower of D37A. Both experiments were performed using the same microscope and the same laser intensity, L526P images were acquired using the 40x microscope and L625P were acquired using the 60x laser. The scale bars are equivalent to 10 µm. Note that L526P mutant images were acquired during the same experiment as images corresponding to the A45V mutant, therefore both cell lines share control images (Figure 36). Two independent experiments were quantified for each clone (N=2), 10 cells in total for L526P and 6 cells in total for L625P. For the control cells 6 cells in total were quantified during the same two independent experiments for each clone.

3.4.5 The abrogation of UBZ domain activity, but not ATPase or Leucine zipper domain, affects the survival of the cells after TMP-UVA treatment

To further investigate the functional role of each of the WRNIP1 domains, all the domain point mutants (apart from A45V, which was previously proven to not be crucial for WRNIP1recruitment to the ICL damage site), were tested (in all cases the first out of two clones obtained for each mutation was tested) in a clonogenic survival assay next to the control cell line (unmodified HeLa Flp-in T-Rex cells), previously obtained HeLa Flp-in T-Rex cells with the disruption of endogenous WRNIP1 gene and the same WRNIP1 deficient HeLa Flp-in T-Rex cell line but complemented with exogenous WRNIP1-EGFP (all the used lanes were previously demonstrated analysed with WB analysis: Figure 7, lanes 1-3).

All the cell lines were plated in three repeats, 500 cells per one well, and treated with 100 ng/ml TMP. After 30 min incubation with the drug all cell lines were irradiated with UVA light at indicated doses (between 0 and 130 J/m^2), and left in the incubator for several days to allow the formation of the colonies.

After this time, the colonies were stained and counted, and statistical analysis was conducted as described in Materials and methods.

As previously, the WT control cells had demonstrated higher survival rates than WRNIP1 deficient cells, and WRNIP1-EGFP complementation was bringing the survival back to the WT levels. As for the mutants: both of the UBZ mutations (N33A and D37A) resulted in the phenotype similar to the WRNIP1 deficient cell line, indicating that UBZ domain is necessary for WRNIP1 to perform its function. Both the ATPase mutants (K274R and E329Q) as well as Leucine zipper mutants (L526P and L625P), brought the survival back to the WT levels just as much as WT WRNIP1-EGFP complementation did, which shows that only mutations in UBZ domain were crucial for WRNIP1 function (**Figure 41**).

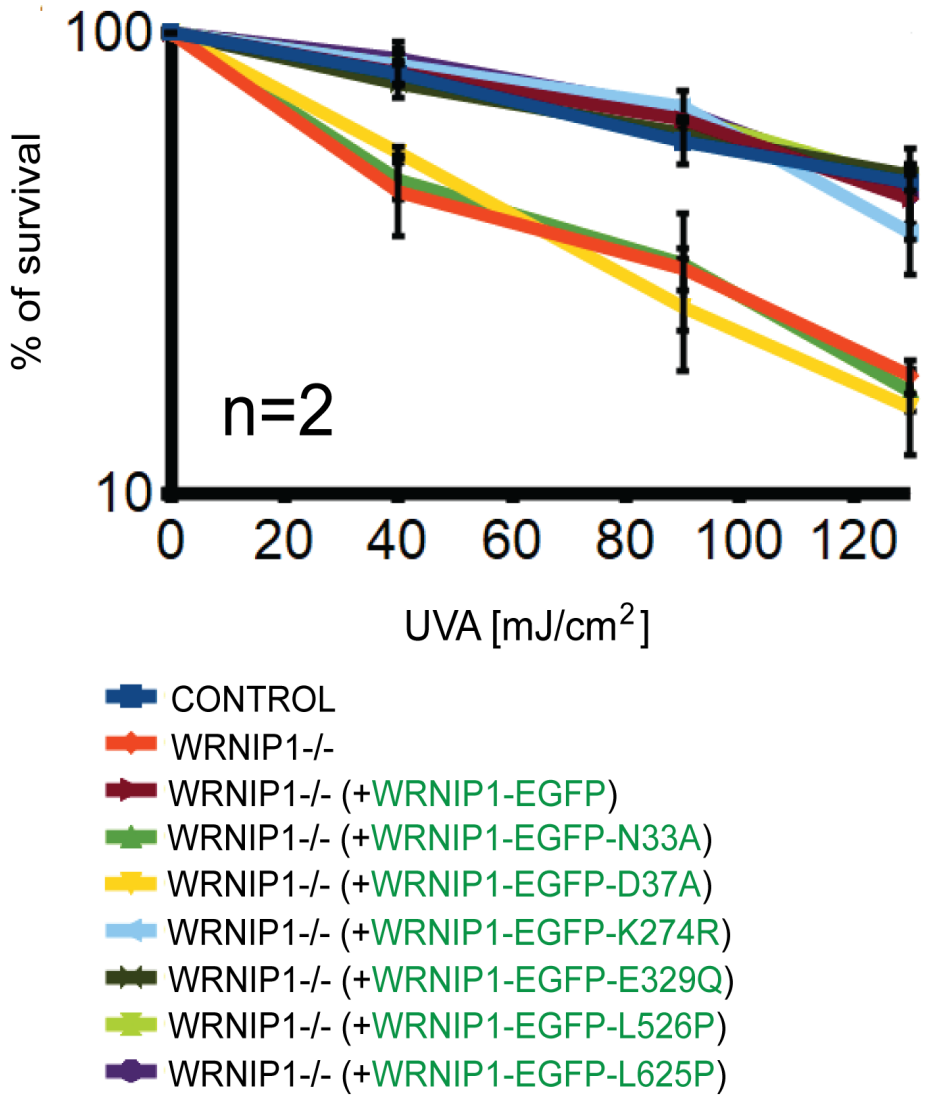


Figure 41: The abrogation of UBZ domain activity results in decreased survival of the cells in response to TMP-UVA treatment. The complementation of WRNIP1 deficient HeLa Flp-in T-Rex cells with WRNIP1-EGFP-N33A or WRNIP1-EGFP-D37A did not bring the cell survival back to the levels comparable with the cells complemented with WT WRNIP1-EGFP. The complementation with K274R, E329Q, L526P and L625P resulted in survival rates at the levels observed in cells complemented with WT WRNIP1-EGFP. The presented graph is a result of combining two independent experiments, with 3 technical repeats each.

3.4.6 Looking for the part of WRNIP1 sequence that binds FANCD2

At this stage of the project thanks to the previous experiments performed by Alicja Bulsiewicz (Dr Cohn's group), we already knew that *in vitro* WRNIP1 interacts directly with FANDC2. To further investigate the role of each domain within WRNIP1 we decided to express respectively; WRNIP1 domain deletion mutants.

We designed five Strep-tagged WRNIP1 mutants. They were missing accordingly: the UBZ domain (17-40 aa), the protein fragment between UBZ and ATPase domain (41-244 aa), ATPase domain – 246-276 aa, the protein fragment between the ATPase and predicted leucine zippers at the C terminus of WRNIP1 (378-511 aa), and finally predicted leucine zippers present at the C terminus of the protein (512-665 aa) (**Figure 42**).

The proteins were expressed in Sf9 cells, and purified by Alicja Bulsiewicz (Dr Cohn's group), who then provided me with purified samples of the mutant proteins (at known concentrations).

To confirm her previous results at the same time as the binding experiment for the mutants we also included wild type Strep-WRNIP1 to see if it will bind to FANCD2.

Surprisingly all the deletion mutants interacted with FANCD2 (**Figure 43**), so to approach the problem from the other angle we divided WRNIP1 into five smaller fragments (**Figure 42**), and we tagged each of them with His-tag. Each of the fragments was cloned into the pET28-c vector, which can be used for *E. coli* expression and purification.

The modified plasmids were checked with DNA Sanger sequencing and transformed into *E. coli* cells, which were grown up to 0.8 optical density (measured at 600 nm wavelength) at 37°C with shaking, and later induced with 0.1 mM IPTG. The expression step was conducted at different temperature and lasted different amount of time for each fragment (see Material and Methods) because as it turned out each of the fragment needed different expression conditions to yield the best results.

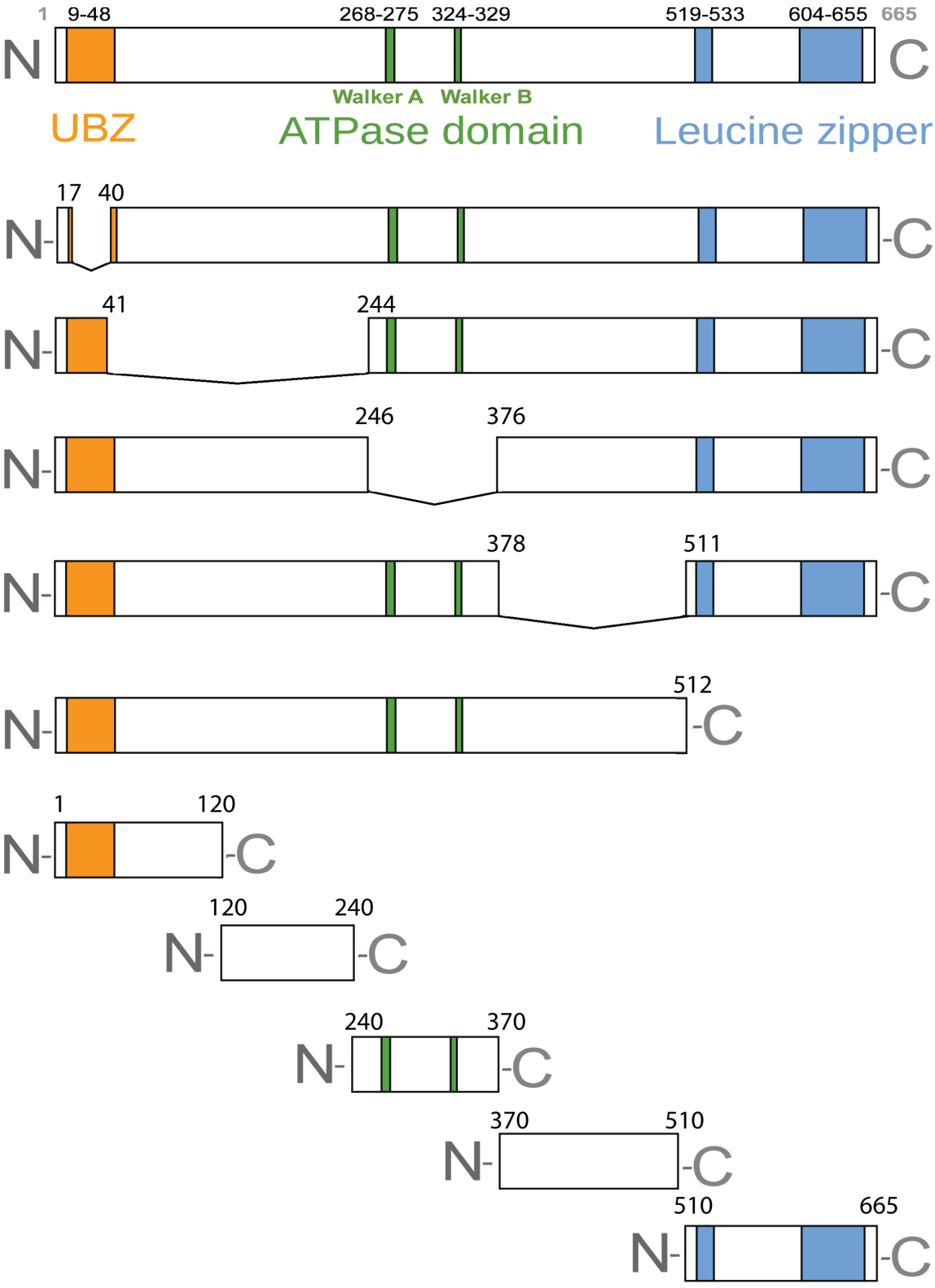


Figure 42: Schematic representation of all the WRNIP1 fragments and deletion mutants tested for FANCD2 binding. All the deletion mutants were Strep-tagged, and all the fragments were His-tagged (bottom five drawings). The schematic of unmodified WRNIP1 is shown at the very top of the figure for comparison.

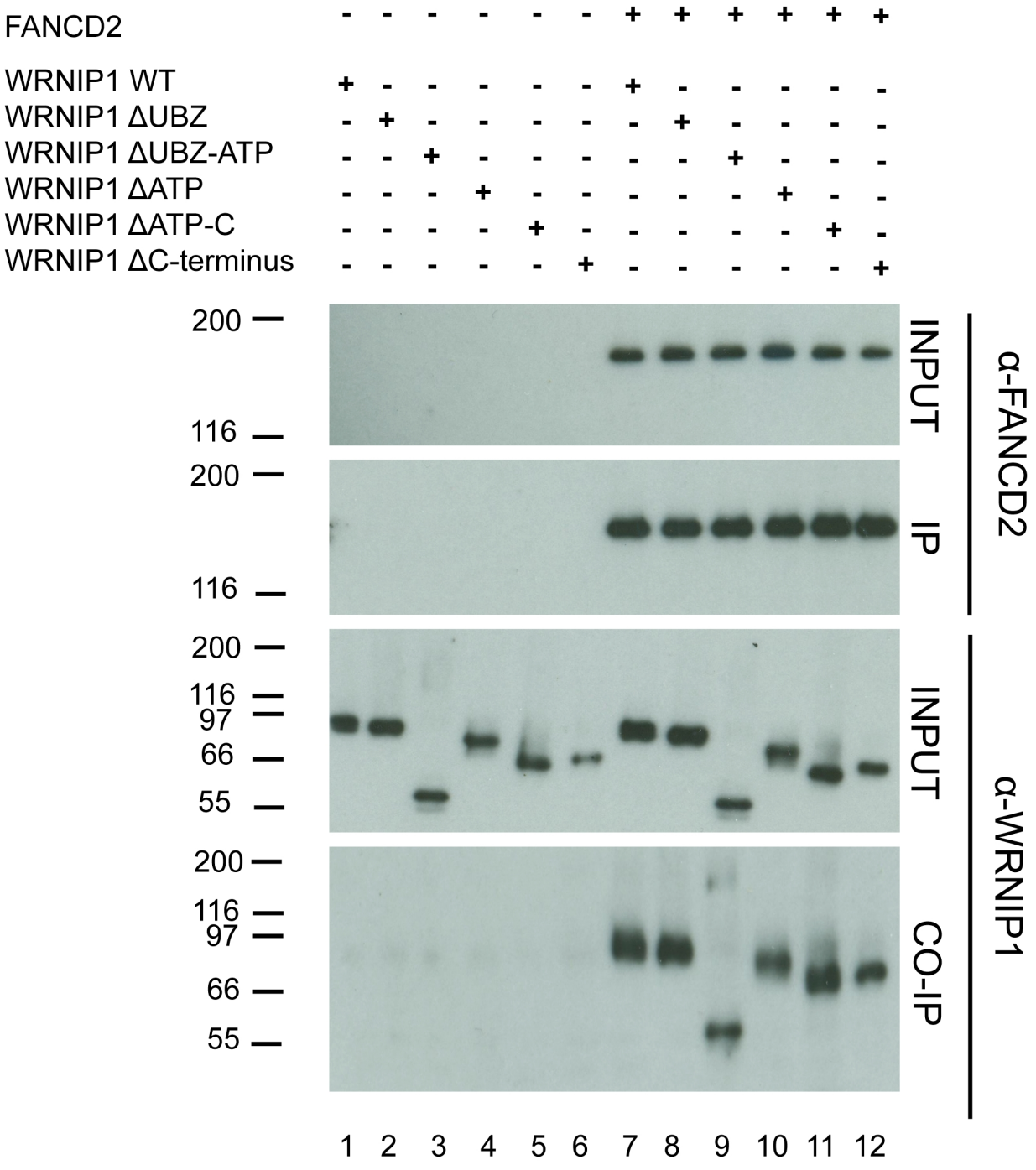


Figure 43: All WRNIP1 deletion mutants bind to FANCD2. Twelve 20 µl solutions were prepared and 2 µl of each solution was taken before the binding reaction to test the input samples, hence the upper α-FANCD2 panel and the upper α-WRNIP1 panel correspond to the input samples. The lower panels for each antibody show the samples after the binding reaction (first pre-binding to allow the proteins to bind and then beads binding so the FLAG-tagged FANCD2 can bind to the beads) and elution from the beads. Each lane corresponds to one sample. In lanes 1-6 the reaction was not supplemented with FANCD2 at all but the same amounts of WRNIP1 mutants were added in order to make sure no unspecific binding takes place. Samples 7-8 were supplemented with both FANCD2 and WRNIP1 deletion mutants or full length protein respectively. All the protein samples were provided by Alicja Bulsiewicz (Dr Cohn's group).

After expression the cells were harvested and the proteins were purified according to the His-purification protocol (see Materials and Methods). Each of the samples was obtained using slightly modified protocol. 0.1 μg of each of the purified protein fragments was tested using WB assay, to test the quality of the samples (Figure 44 A). If higher levels of proteins were possible to obtain, it would be ideal to test the samples using Coomassie Blue staining instead, as it would allow to visualize other proteins contaminating the sample as well, but unfortunately the fragments were possible to purify only at very low concentrations that were possible to detect only using WB analysis.

The co-immunoprecipitation of the purified His-tagged fragments of WRNIP1 was performed using the standard WRNIP1 – FANCD2 binding protocol (see Materials and Methods) and as an additional control the full length Strep-tagged WRNIP1 was also bound to FANCD2 under the same conditions. Because the control full-length WRNIP1 used in this experiment was not tagged with the same tag as the fragments and also was not purified from the same organism and under the same conditions (for practical reasons), it is important to note that all these data can only be regarded as preliminary and to fully confirm the observed binding, additional experiments with full length WRNIP1 tagged with identical tags and expressed using the same expression system should be performed.

Two other control samples included one sample that had only FANCD2 but no WRNIP1 and the other where only WRNIP1 was added with no FANCD2, to exclude the possibility that WRNIP1 binds to the α-HA sepharose beads on its own.

The result of this experiment confirmed once again that full length Strep-WRNIP1 binds to FANDC2, but it also showed that fragment number four (370-510) interacts with FANCD2 *in vitro* (**Figure 44**). None of the other fragments interacted with FLAG-FANCD2 and also the control sample where sample number four was included but none of the FANCD2 was added has shown no interaction between the fragment and the beads, which suggests that the interaction is specific.

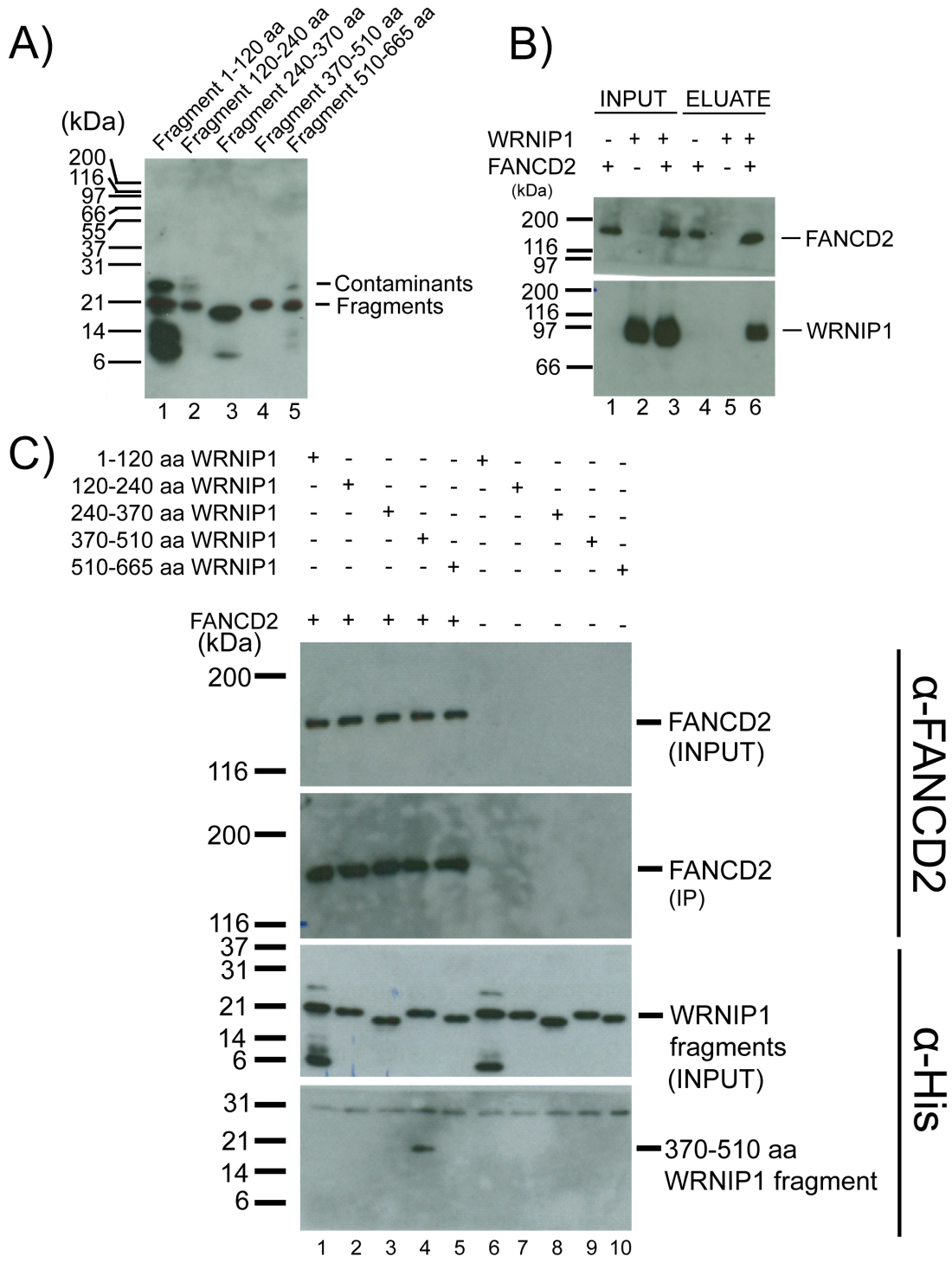


Figure 44: WRNIP1 fragment consisting of amino acids 370-510 binds to FANCD2. A) 0.1 µl of each purified sample was tested on WB against α-His antibody, B) Control reactions: full length Strep-WRNIP1 and FLAG-FANCD2 were proven to interact (lanes 3 – input and 6 – elution). The sample not supplemented with FANCD2 shows almost no unspecific binding of WRNIP1 to the beads (lanes 2 - input and 5 – elution). C) Under the same conditions 5 His tagged fragments were tested. Lanes 1-5 correspond to samples supplemented with FANCD2, lanes 6-10 correspond to the control samples that were not supplemented with FANCD2. None of the samples 6-10 have shown any unspecific binding. Among the samples supplemented with FANCD2 fragment 370-510 aa did bind to FANCD2 at the levels detectable by the Western Blot assay (lane 4).

3.4.7 Discussion

The analysis of WRNIP1 point mutants has shown that the UBZ domain present at the N-terminus of WRNIP1 is crucial for WRNIP1 ubiquitination as well as its recruitment to the ICL damage site. Two out of the three mutants introduced within this domain have shown no upper form in Western Blot analysis, and the same two were not recruited to the ICL damage site (after 405 nm laser irradiation of the TMP sensitized cells).

The mutations introduced in the ATPase domain both resulted in similar changes. Firstly, both mutants had altered band pattern on the WB, which may be the result of altered degradation (visible as additional bands of lower molecular weight than WRNIP1-EGFP) both ATP-ase mutants have also shown similar ICL recruitment pattern, where the EGFP-tagged WRNIP1 was recruited to the damage site, but with slightly different kinetics- slower recruitment to the chromatin.

One of the two potential leucine zipper mutations- L625P- did not differ in any significant way to the unmodified WRNIP1, which means that either there are no leucine zippers corresponding to this residue, or that this leucine zipper is formed, but its lack is not affecting the levels of expressed WRNIP1-EGFP, its upper ubiquitinated form nor its recruitment to the ICL damage site.

The second mutation associated with leucine zippers at the C-terminus did differ from unmodified WRNIP1-EGFP though. Firstly we observed significantly reduced levels of ubiquitinated WRNIP1-EGFP, and secondly the mutated protein was not recruited to the ICL damage site, which confirms that WRNIP1-recruitment is strongly linked to the presence of its upper form. Why the mutation at this residue results in such striking phenotype remains to be further investigated.

The survival clonogenic assay has shown that of all the three investigated domains only UBZ is crucial for WRNIP1 function (both N33A and D37A mutants have shown survival as low as WRNIP1 deficient cells, while all the other mutations presented survival identical to the cells

supplemented with unmodified WRNIP1-EGFP). These observed differences in recruitment and survival, would suggest that there is some alternative pathway that can be used to resolve ICL damage, and perhaps it is activated when WRNIP1 is mutated at L526 residue, but not activated when the mutations are introduced in the UBZ domain.

The *in vitro* co-immunoprecipitation experiments have suggested that none of the deleted fragments alone is crucial for WRNIP1-FANCD2 binding, which suggests that there is more than one part of the protein that interacting with FANCD2. It is important though, to consider the possibility of potential artefacts in this experiment- possibly we are observing some form of a non-specific interaction, and more research will be required to fully exclude such possibility.

In the second part of the binding experiments, we have presented preliminary data that one of the binding fragments is somewhere between 370-510 amino acids, as the fragment containing the sequence did specifically bind to FANCD2.

It is important to stress that *in vitro* data may sometimes give the wrong idea about what is happening *in vivo* under physiological conditions and should be treated only as useful preliminary data. To fully understand the mechanism of WRNIP1-FANCD2 interaction more work would be required, possibly starting with *in vivo* binding assays.

4 Discussion

4.1 Conclusions

WRNIP1 seemed to be likely associated with the Fanconi anemia pathway, due to its interesting domain architecture (UBZ4 domain next to the ATPase domain), the fact that its recruitment to the DNA damage site was previously observed by ***Nomura et al., 2012*** and our preliminary FANCD2 binding data.

At first the clonogenic survival assays have indeed shown, that cells that are not expressing WRNIP1 (after CRISPR/Cas9 gene disruption) are more sensitive to both MMC and TMP-UVA treatment than wild type cells. To confirm this result the CRISPR/Cas9 modified cell line was transfected with WRNIP1-EGFP (obtaining stable cell line), and it was shown that adding exogenous WRNIP1 rescues the sensitivity, to the levels observed in the wild type cell line.

In this thesis only MMC and TMP-UVA crosslinking was investigated and it is important to note that not all types of ICLs have to be necessarily resolved using identical molecular machinery. In fact, it seems unlikely given that each type involves different nucleotides and causes different DNA helix distortion, which means that after all they are quite varied chemical entities. It was demonstrated that the level of DNA distortion does indeed affect the amount of ICL unhooking in G1 phase of the cell cycle (***Smeaton et al., 2008***). It was also observed that lesions that result in little helix distortion are repaired only in replication-competent extracts, while the repair of helix-distorting ICLs is detectable also in non-replicating DNA (***Raschle et al., 2008; Ben-Yehoyada et al., 2009***), which further supports the hypothesis that there are differences in the way cells resolve different kinds of ICLs. Both MMC and TMP-UVA cause ICLs that result in similar minor helix distortion, while cisplatin introduces major helix distortion. HU on the other hand causes replication stall and fork collapse while not introducing any ICLs.

It was impossible to exclude at this stage though, that this observed sensitivity of the WRNIP1 deficient cell lines is caused not by ICLs but by some other type of DNA damage, as all

ICL inducing agents are also introducing many other types of lesions (***Deans and West., 2011***).

To further investigate links between WRNIP1 and FA pathway, WRNIP1 and FANCD2 double gene disruption cell line was obtained and tested in clonogenic survival assay next to WRNIP1 and FANCD2 single gene disruption cell lines. The tests were conducted both using MMC and TMP-UVA assays. In both cases WRNIP1 deficient cell line was more sensitive than wild type cell line (consistent with previous observations), but less sensitive than FANCD2 deficient cell line, which was not unexpected given the central role of FANCD2 in the ICL repair machinery. The double gene disruption cell line was not more sensitive than FANCD2 deficient one, showing that WRNIP1 and FANCD2 deficiency does not result in even more sensitivity to the ICL agents than FANCD2 deficiency alone. This observation can be interpreted as suggesting WRNIP1 and FANCD2 operate within the same pathway as if they were completely independent from each other a more likely outcome would be the additive effect of both gene disruptions.

In the next stage of the project it was shown that WRNIP1 is recruited to the damage site after TMP treatment followed by 405 nm laser irradiation through a microscope lens. This technique is a very effective tool for investigating protein recruitment to the DNA damage site in real time *in vivo (**Lan et al., 2004**)*. After introducing exogenous protein tagged with fluorescent tag one area of chromatin can be irradiated with the laser using the wavelength of choice, after some time, if the investigated protein is being recruited to the damage site, the irradiated area will have increased intensity of the fluorescence due to the accumulation of the protein.

Similar experiments involving WRNIP1-EGFP were performed previously by ***Nomura et al. (2012)***. In this previous study cells were treated only with UV laser (405 nm wavelength), with no previous psoralen sensitization. This procedure would result in all kinds of DNA damage, with large percentage of double strand breaks among the lesions. In this study, cells were incubated with TMP for 15 min before the irradiation, in order to ensure that the treatment will introduce ICL damage.

To find the conditions that result only in TMP related damage we gradually decreased the levels of laser irradiation until the chromatin recruitment was observed only in the sample pre-treated with the drug, while the control sample showed no visible fluorescence enrichment in the irradiated area. It is important to note however, that while optimizing the experimental conditions in such way can increase the ratio of ICL damage it cannot completely eliminate other types of lesions.

The chromatin recruitment assays have all confirmed that WRNIP1 is relocalizing to the DNA damage site in response to ICL, and they also suggested that in fact, it is recruited faster than FANCD2. Both FANCD2 and WRNIP1 recruitment assays were performed under identical conditions, yet WRNIP1 recruitment was observed within first 3 min, while the increase in FANCD2 levels at the damage site was usually not visible for the first 10-15 min. To confirm, WRNIP1-EGFP and mCherry-tagged FANCD2 were expressed in the same cell line, and the recruitment was observed in two channels. As expected WRNIP1 was recruited to the damage site before FANCD2 suggesting that if it plays some role in the FA pathway it would be involved before the recruitment of the FANCD2-FANCI complex, at the initial stages of the ICL recognition, upstream of FANCD2.

This was partially confirmed while FANCD2 recruitment in the WRNIP1 deficient cell line was compared with the recruitment in the wild type cell line, showing that FANCD2 is indeed slightly affected by the lack of WRNIP1, and that its recruitment is delayed and less abundant. The similar experiment but with WRNIP1-EGFP and one of the cell lines, having FANCD2 gene disrupted instead has shown no differences in WRNIP1 recruitment.

Another important question that required further investigation at this stage, was whether or not WRNIP1 recruitment is affected by the cell cycle in any way. It is well documented that the FA pathway activity depends on the cell cycle. Core complex localization is regulated in a cell cycle dependent way and it associates more strongly with chromatin during the S phase (***Kim et al., 2008***). Also FANCD2 interactions were shown to be dependent on the cell cycle and for example it

was demonstrated that it interacts with both RAD51 and BRCA1 in the S-phase specific manner (***Taniguchi et al., 2002***).

As majority of FA proteins are active primarily in the S-phase of the cell cycle **(*Jo and Kim, 2015***) we speculated that maybe the recruitment of WRNIP1 to the laser irradiated chromatin will be also higher during S and G2 phases of the cell cycle in comparison to the G1 phase. Surprisingly the chromatin recruitment assay performed with cells expressing WRNIP1-EGFP and mCherry tagged geminin fragment (a replication inhibitor, almost absent during G1 phase and present in S and G2 phases), has shown that, the levels of WRNIP1-EGFP recruitment are in fact higher during the G1 phase of the cell cycle, but the difference was not statistically significant. This finding seems counter-intuitive as covalent linkages between the two DNA strands need to be resolved most urgently during the S phase of the cell cycle, during DNA replication, but ICL repair outside of the S phase is also critical to maintain genomic integrity and to allow transcription in non-dividing or slowly dividing cells (***Williams et al., 2012***). It is important to note though, that transcription levels were not checked in WRNIP1 deficient cell lines, which is potentially one more future direction to explore in this project.

Moreover, this observation is in line with previous findings that WRNIP1 interacts and stimulates the activity of DNA Polymerase δ (***Tsurimoto et al., 2005***), and then that Polymerase δ is involved in G1 ICL repair **(*Sarkar et al., 2006*)**.

While the live-cell microscopy experiments allowed us to observe WRNIP1 response to the DNA damage, we still did not have much insight into the exact molecular mechanism of this response. In order to further investigate the nature of WRNIP1 we decided to focus some more on WRNIP1 post-translational modifications. It was known at this stage that while analysed using Western Blot assay, WRNIP1 always shows two protein bands. The lower one that migrates at 72 kDa and the upper one migrates slightly above the main band around 90-95 kDa (the molecular weight that would indicate ubiquitination with either two or three ubiquitin molecules).

Simple titration of the whole cell lysate, followed by WB analysis, allowed us to roughly estimate that the upper form constitutes roughly 10% of all WRNIP1 detected by the antibody. Then after testing the purified sample of WRNIP1 with both α-WRNIP1 and α-ubiquitin antibodies we were able to show that the upper form is in fact ubiquitinated WRNIP1.

This was quite encouraging as ubiquitination plays a large role in the FA pathway signalling. Both FANCD2 and FANCI undergo ubiquitination. FANCD2 monoubiquitination at lysine 561 is a crucial step in FA response to ICLs, as is FANCI monoubiquitination at lysine 523.

Extrapolating from these two examples we were expecting WRNIP1 to also have one lysine which undergoes monoubiquitination, possibly enabling WRNIP1 interaction with some other FA pathway players. To identify this lysine, we purified the upper form, and analysed it with mass spectrometry. The outcome of the analysis was very surprising though. Not one but many potential ubiquitination sites were found.

To further analyse the found ubiquitination sites, we introduced several point mutations, substituting lysines with arginines (which cannot undergo ubiquitination). The point mutations were successfully introduced into WRNIP1-EGFP sequence and transfected into HeLa cells, yet the subsequent WB analysis had provided the outcome as puzzling as the initial mass spectrometry: none of the tested point mutations was abrogating the formation of the ubiquitinated upper form.

The chromatin recruitment assay results were in line with our WB observations. None of the lysine point mutants (apart from K70R, which was not localized correctly because of mutation within the NLS sequence) have presented phenotype different to the wild type.

This result may have a trivial explanation. The real lysine responsible for the upper form may have been simply not detected by the mass spectrometry or if it was then it might have come with score low enough to be overlooked during the data analysis stage. This possibility is impossible to exclude, as mass spectrometry is not always a perfectly accurate technique and there is always a chance that some peptide was simply not detected. Another possibility would be that the

upper form is not the result of one ubiquitinated lysine but two or more (multiubiquitination) and that abrogating the ubiquitination at one of these residues simply increases the ratio of all the other ubiqutinated WRNIP1 forms. We can speculate for example that maybe WRNIP1 has three lysine residues where it can be monoubiquitinated, and maybe ubiquitination of the two of them is enough to make it perform its function. In such case even if we did abrogate one of the three it would still have the other two and the effect of the mutation would not be even detectable using the WB assay.

In the next stage of the project we wanted to gain more insight into the function that each of WRNIP1 domains performs. To achieve this goal, we designed several point mutations, each of them aimed at disrupting a specific domain (or subdomain) of the protein.

We found two mutations within the UBZ domain that completely abrogate WRNIP1 ubiquitination as well as its recruitment to the laser irradiation site in our chromatin recruitment assay. One of them was D37A mutation (previously described by ***Crosetto et al., 2008***), the other one was N33A which resulted in identical phenotype as D37A. Our observations were similar to the findings previously described by ***Crosetto et al. (2008)***. In that previous study wild type WRNIP1-EGFP was shown to accumulate in many small foci within the cell nucleus, following the UVC irradiation. D37A mutation (as well as UBZ domain deletion) was abrogating the formation of such foci. Of course, the nature of DNA damage introduced by UVC light is very different to the damage introduced by the TMP-UVA treatment, and also foci formation is not the same process as chromatin site-specific recruitment, but in both cases functional UBZ domain was necessary for the protein to respond to the damage.

The next examined domain was ATPase domain. We introduced one mutation to abrogate Walker A motif (K274R) and one to abrogate Walker B motif (E329Q). Both point mutations resulted in a very similar phenotype. Both mutated proteins seemed to have altered band pattern on WB (possibly due to altered degradation), and both of them seemed to be recruited to the ICL damage site, but the kinetics of this recruitment were different to the kinetics of the wild type

protein. Especially during the first 15 min after the irradiation, mutated WRNIP1-EGFP levels were raising more slowly than wild type levels. This may suggest that ATPase activity is necessary for quick efficient WRNIP1 recruitment (and retention) to the damage site. It was shown by ***Tsurimoto et al. (2005)*** that WRNIP1 has ATPase activity and that primer/template structures and double-stranded termini stimulated it. It seems tempting to speculate at this point that maybe ATPase domain acts as a sensor of DNA damage and, upon detecting ICL or other type of DNA damage that WRNIP1 is associated with, it starts ATP hydrolysis, which in turn stimulates the activity of some unknown yet WRNIP1 partners. As a result, WRNIP1 binding to this unknown yet protein (or complex) may be mitigated and more WRNIP1 molecules may be recruited to the damage site. In such case lack of ATPase activity would still allow WRNIP1 to be relocalized to the irradiated chromatin, either on its own or in a complex with unknown yet binding partners, but since the interaction will not be facilitated, the recruitment will be less efficient, delayed and possibly reaching lower maximum levels.

Also it is important to mention that since it was reported before that WRNIP1 binds DNA in an ATP-dependent manner (***Yoshimura et al., 2009***), and that lack of ATPase activity abolishes the ATP binding activity of WRNIP1, the fact that both WRNIP1(K274R) and WRNIP1(E329Q) still accumulate at irradiated sites means that WRNIP1 recruitment should not be mediated by binding of WRNIP1 to DNA.

Finally, we identified one of the leucine residues at the C-terminus of WRNIP1 (L526) to be almost as important for WRNIP1 ubiquitination and recruitment to the damage site as the functional UBZ domain. The L526P mutant cells have shown both none (or almost none) of the upper ubiquitinated form as well as almost no recruitment to the damage site.

The fact that L526 WRNIP1-EGFP mutant have responded with almost no recruitment to the ICL damage, was surprising given the results presented previously by ***Nomura et al. (2012)***. It was shown in this study that N-terminal region containing the UBZ domain and the

C-terminal region containing the two potential leucine zippers but not the middle region containing ATPase domain was required for the accumulation of WRNIP1 at the DNA damage sites (consisting mostly of double strand breaks). These results so far were consistent with our observations. But then to further investigate the mechanism of WRNIP1 recruitment the authors introduce the deletion of the first potential leucine zipper (Δ502-548) into the protein sequence, and the mutant is still recruited to the irradiated chromatin, even though the residue L526 is not present within this protein. There may be several explanations of why our data are not supporting these previous reports. One possibility is that WRNIP1 responds differently to ICL damage and to DSB damage (type of lesions introduced into the DNA by the laser irradiation with no previous psoralen treatment), and while ICL response requires L526 residue, DSB does not. Another possibility is that if both of the predicted leucine zippers are indeed at the C-terminus of the protein, possibly the first one is the one required for recruitment to the chromatin, while the second one serves some other purpose. We can speculate that the deletion mutant Δ502-548, folded in such way that the second leucine zipper, substituted the first one (since there were almost 50 aa missing, and there might have been enough space for such re-arrangement) and the resulting protein acted like WRNIP1 without the second leucine zipper, rather than the first, while our mutant, could not have folded in such way as we only introduced a point mutation and there was a long potentially incorrectly folded part of the protein in the way.

It is also possible that our L526P WRNIP1 mutant had a far more disrupted structure than the deletion mutant Δ502-548. Even though our mutant had only one point mutation, while the mutant tested by ***Nomura et al. (2012)*** was missing almost 50 amino acids, it is important to note that proline can be very disruptive for the protein structure as it is likely to induce a kink.

In the last part of the project we tried to confirm that there is a physical interaction between WRNIP1 and FANCD2. The co-immunoprecipitation of the tagged proteins expressed and purified from Sf9 cells by Alicja Bulsiewicz (Dr Cohn's laboratory) have confirmed that the interaction is

indeed direct. More puzzling were the results obtained with WRNIP1 deletion mutants, which have shown that no matter which part of WRNIP1 was deleted it was still interacting with FANCD2.

To address this issue from another angle, we designed five truncated His-tagged WRNIP1 mutants, each being a short part of WRNIP1 sequence, and we found that only one fragment binds to FANCD2 under identical conditions as wild type WRNIP1, namely the fragment consisting of residues 370-510. Interestingly the deletion mutant missing this part of the sequence have shown no decreased levels of binding in comparison to the full length WRNIP1. We speculated that maybe there are two parts of WRNIP1 that bind to FANCD2 and each of them alone is enough to achieve the same levels of interaction as both of them together, but in the second experiment where only short fragments were expressed only one of them (370-510 aa) was intact, while the other one was unintentionally divided into two fragments and could no longer bind FANCD2.

Given all the previously discussed data it seems that WRNIP1 is likely involved in ICL repair, either by FA pathway or some other DNA repair mechanism. It was proven that its recruitment to the ICL damage site precedes FANCD2 recruitment and based on the measured timing it also occurs after UHRF1 recruitment (**Figure 45**). It is not yet established whether it is WRNIP1 ubiquitination that is necessary for the recruitment or maybe just functional UBZ domain, but given our observation that D37A and N33A mutations abrogate all ubiquitination and recruitment to the ICL damage site, we can speculate that WRNIP1 ubiquitination precedes its recruitment, and that likely it is ubiquitinated form that ends up on the lesion site (further experiments needed).

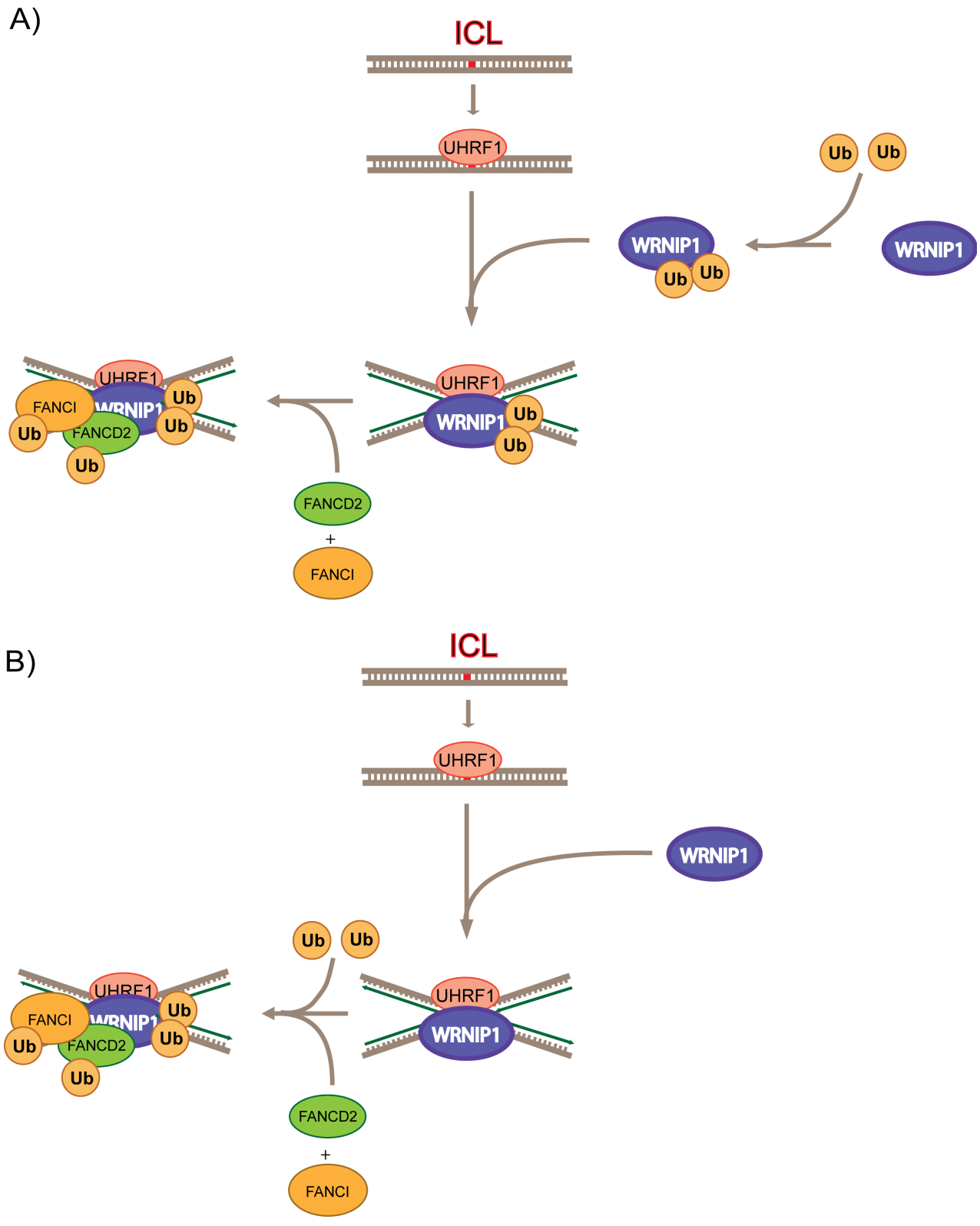


Figure 45: Proposed model of WRNIP1 recruitment to the ICL damage site. We can speculate that WRNIP1 is multiubiquitinated before being recruited to the damage site via its UBZ domain **(A)** or after **(B)**. Based on our chromatin recruitment live-cell microscopy observations, this event occurs later than UHRF1 recruitment but before FANCD2 accumulation can be observed at the lesion.

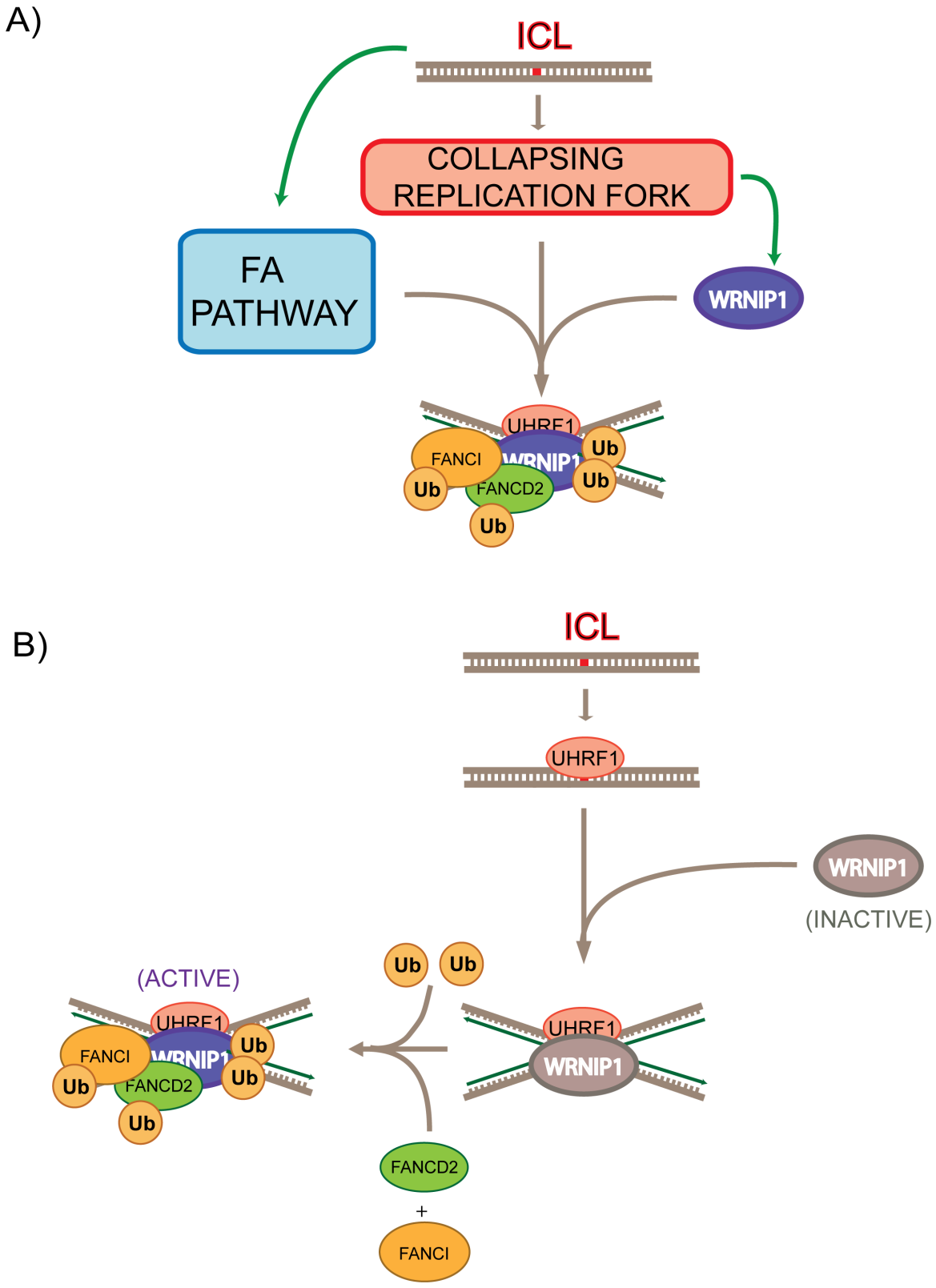


Figure 46: Alternative theories. A) WRNIP1 is independent of the FA pathway, and while the FA pathways proteins are recruited in response to the ICL, WRNIP1 is recruited in response to the stalled replication fork, which is a result of the ICL damage **B)** WRNIP1 acts downstream to FANCD2 but is recruited before it. It may be activated during later stages of ICL repair.

While the functional interaction between WRNIP1 and FANCD2 remains to be further investigated, there are several most likely possibilities. First, WRNIP1 may be acting upstream of FANCD2, and its ATPase domain may be serving as an early stage DNA damage marker. While not being crucial for WRNIP1 recruitment to the ICL lesion, it may allow WRNIP1 to boost the efficiency of the DNA repair machinery, which would explain why the levels of FANCD2 recruitment to the damage site are lower in the absence of WRNIP1.

Another possibility may be that WRNIP1 is not necessarily involved with the FA pathway (**Figure 46 A**), but has some important role in ensuring the progression of the replication fork. We can speculate that in such case it would be evolutionarily beneficial for such protein to interact physically with many proteins that are recruited to the DNA damage site, to ensure that it is always abundant at the lesion site, when the replication fork collapses. This hypothesis would be supported by the data presented by ***Leuzzi et al. (2016)***, which demonstrated that loss of WRNIP1 leads to nascent DNA strand degradation after HU- induced replication stress and also that one of WRNIP1 functions is to stabilize RAD51 on stalled replication forks. It is also not necessarily excluded by our own survival assay where the cells were treated with HU, and WRNIP1 depleted cell line has shown no more sensitivity than the wild type cell line. HU depletes the cells of dNTPs, which initially results in stalled replication forks that, after prolonged treatment, collapse into DSBs. It seems entirely possible that WRNIP1 is simply a first response protein, and that long time effects of its depletion are not adverse, because the function it performs can be redirected into some other pathway. The discrepancies we see between TMP-UVA, MMC, and HU treatment may simply mean that depending on the character of the DNA lesion, different pathways will come into play to compensate lack of WRNIP1 and potentially some of them are less effective substitutions than others.

The third option can be that despite being recruited to the ICL damage site early and independently to FANCD2-FANCI complex, WRNIP1 does not perform its role until later, being in

fact mostly a downstream protein, with very minor potential to slightly boost FANCD2 recruitment to the damage site (potentially even only as a side effect of the physical interaction between the two proteins) (**Figure 46 B**).

Regardless of the speculations, at this stage of the project more work is needed to establish the real function of WRNIP1 as well as the precise mechanism of its chromatin recruitment and interaction with the FANCD2-FANCI complex.

4.2 Future Directions

All the data presented in this thesis is mostly preliminary and there are several possible steps that would allow to corroborate our findings. Firstly, it would be useful to test WRNIP1-EGFP recruitment to the damage site in relation to UHRF1 recruitment in analogous experiments as the ones performed with FANCD2. At this stage it seems that WRNIP1 is recruited shortly after UHRF1 within first 5 min after the introduction of the DNA damage (UHRF1 can be observed around 30 s after irradiation). It would be interesting to see if UHRF1 depletion affects WRNIP1 recruitment in any way and vice versa.

It would also be interesting to see if the recruitment of FANCD2 is affected in anyway in the cell line expressing WRNIP1-EGFP with one of the two mutations that abrogate ubiquitination of the protein. While the recruitment of the FANCD2 was only slightly decreased in the cell line with WRNIP1 gene disrupted, it is impossible to exclude that under such conditions, lack of WRNIP1 is quickly compensated by some other protein and the overall phenotype is not strongly affected. Possibly under conditions where there is WRNIP1 with functional both ATPase and leucine zipper domains, lack of its ubiquitination will have more adverse effect than simple WRNIP1 depletion, as the cells may not be able to detect lack of functional WRNIP1 as efficiently to redirect the pathway early enough.

Another possible direction is investigating the relationship between WRNIP1-EGFP

recruitment and FANCI. It was shown in previous *in vitro* assays that there is an interaction between the two proteins (Alicja Bulsiewicz, Dr Cohn's laboratory: personal communication), and maybe it would be possible to show some more functional interaction if we observe the recruitment of the two proteins in the same cell line and also see if depletion of one has any effect on the recruitment of the other.

The in vitro binding assays themselves present several possibilities for future improvement. It would be helpful to corroborate WRNIP1 binding to both FANCD2 and FANCI by checking whether WRNIP1 is able to bind FANCD2 as well as FANCI *in vitro*.

As to mapping of the fragment of WRNIP1 that interacts with FANCD2, it seems that more work will be needed to find the exact amino acids responsible for the interaction. One of the possibilities would be to divide WRNIP1 into three parts, and then exclude one part at the time. As a result, we would have one truncated WRNIP1, with only UBZ and ATPase domain, one with UBZ and Leucine zipper domain, and one with ATPase and leucine zipper (double-fragments). At the same time each of the excluded fragments should be expressed with the same tag (single-fragments), and then all six versions of WRNIP1 should be co-immunoprecipitated with FANCD2. If one of the three single-fragment variants of WRNIP1 would not bind to the FANCD2 while two other ones would we would have a very strong evidence that indeed two fragments of WRNIP1 are responsible for the interaction with FANCD2. In such situation we would expect all three double-fragment mutants to interact with FANCD2. If we see that only one of the three single-fragment mutants interacts with FANCD2 and at the same time the double-fragment mutant where this piece is deleted does not interact then, we will know where in the WRNIP1 sequence is the fragment or several fragments responsible for FANCD2 interaction.

It is important to note that even if only one of the three fragments is found to interact with FANCD2 it would not exclude the possibility that in fact there are several interaction sites, as one third of the entire 665 aa sequence is still long enough to contain multiple binding motifs.

The analysis of the upper form of WRNIP1 may also require more work. We found many potential ubiquitinated lysines, and none of them seemed to be necessary for the upper form to exist. ***Nomura et al. (2012)*** investigated if multiple lysine mutants are still recruited to the DNA damage site, finding out that indeed they are. In this study not all potential ubiquitinated lysines were mutated (the highest number of lysines mutated in one variant of the protein was four) but various combinations were tested and all of them were still recruited to the damage site. It would be interesting to see if maybe by introducing all the lysine mutations into one construct (apart from K70R which would prevent WRNIP1 localization to the nucleus), we would be able to abrogate its ubiquitination. If then none of the upper form would be detectable in the WB assay, but the protein would be still recruited to the damage site, that would strongly implicate that it is not ubiquitination but rather interaction with ubiquitin via UBZ domain that is necessary for WRNIP1 recruitment to the ICL damage site. If what we see is that such WRNIP1-EGFP mutant is not recruited to the damage site, despite having a functional UBZ domain, then we can conclude that abrogating UBZ function did not have a direct impact on lack of WRNIP1 recruitment, and that it was lack of ubiquitinated form that resulted in such outcome.

Finally, it would be useful to further investigate the upper ubiquitinated form of WRNIP1. There are many commercially available antibodies that are able to detect different types of ubiquitination. It would be helpful to purify WRNIP1 from human cell line and establish the exact type of WRNIP1 polyubiquitination, or to confirm that the upper form is either mono- or multiubiquitinated.

Bibliography

Ben-Yehoyada, M., Wang, L. C., Kozekov, I. D., Rizzo, C. J., Gottesman, M. E., & Gautier, J (2009). Checkpoint Signaling from a Single DNA Interstrand Crosslink. *Molecular Cell*, *35*(5), 704–715. https://doi.org/10.1016/j.molcel.2009.08.014

Bienko, M.; Green, C.; Crosetto, N.; Rudolf, F.; Zapart, G.; Coull, B.; Kannouche, P.; Wider, G.; Peter, M., Lehmann, A.; Hofmann, K.; Dikic, I. (2005). Ubiquitin-Binding Domains in Y-Family Polymerases Regulate Translesion Synthesis. *Science*, *310*(5755), 1821–1824. https://doi.org/10.1126/science.1120615

Bish, R. A., Fregoso, O. I., Piccini, A., & Myers, M. P. (2008). Conjugation of complex polyubiquitin chains to WRNIP1. *Journal of Proteome Research*, *7*(8), 3481–3489. https://doi.org/10.1021/pr800217q

Bish, R. A., & Myers, M. P. (2007). Werner helicase-interacting protein 1 binds polyubiquitin via its zinc finger domain. *Journal of Biological Chemistry*, *282*(32), 23184–23193. https://doi.org/10.1074/jbc.M701042200

Bornberg-Bauer, E., Rivals, E., & Vingron, M. (1998). Computational approaches to identify leucine zippers. *Nucleic Acids Research*, *26*(11), 2740–2746. https://doi.org/10.1093/nar/26.11.2740

Branzei, D., & Foiani, M. (2008). Regulation of DNA repair throughout the cell cycle. *Nature Reviews. Molecular Cell Biology*, *9*(4), 297–308. https://doi.org/10.1038/nrm2351

Brendel, M., & Ruhland, A. (1984). Relationships between functionality and genetic toxicology of selected DNA-damaging agents. *Mutation Research/Reviews in Genetic Toxicology*, *133*(1), 51–85. https://doi.org/10.1016/0165-1110(84)90003-4

Brooks, P. J., & Theruvathu, J. A. (2005). DNA adducts from acetaldehyde: Implications for alcohol-related carcinogenesis. *Alcohol*, *35*(3), 187–193. https://doi.org/10.1016/j.alcohol.2005.03.009

Budzowska, M., Graham, T. G., Sobeck, A., Waga, S., & Walter, J. C. (2015). Regulation of the Rev1–polζ complex during bypass of a DNA interstrand cross-link. *The EMBO Journal*, *34*(14), 1971–1985. https://doi.org/10.15252/embj

Caulfield, J. L., Wishnok, J. S., & Tannenbaum, S. R. (2003). Nitric Oxide-Induced Interstrand Cross-Links in DNA. *Chemical Research in Toxicology*, *16*(5).

Chan, K. L., Palmai-Pallag, T., Ying, S., & Hickson, I. D. (2009). Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. *Nature Cell Biology*, *11*(6), 753–760. https://doi.org/10.1038/ncb1882

Chaudhury, I., Stroik, D. R., & Sobeck, A. (2014). FANCD2-controlled chromatin access of the Fanconi-associated nuclease FAN1 is crucial for the recovery of stalled replication forks. *Molecular and Cellular Biology*, *34*(21), 3939–54. https://doi.org/10.1128/MCB.00457-14

Ciccia, A., & Elledge, S. J. (2010). The DNA Damage Response: Making It Safe to Play with Knives. *Molecular Cell*, *40*(2), 179–204. https://doi.org/10.1016/j.molcel.2010.09.019

Cohn, M. A., Kee, Y., Haas, W., Gygi, S. P., & D'Andrea, A. D. (2009). UAF1 is a subunit of multiple deubiquitinating enzyme complexes. *Journal of Biological Chemistry*, *284*(8), 5343–5351. https://doi.org/10.1074/jbc.M808430200

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., Zhang F. (2013). Multiplex Genome Engineering Using CRISPR/VCas Systems. *Science*, *339*(6121), 819–823. https://doi.org/10.1126/science.1231143.Multiplex

Costanzo, V., Paull, T., Gottesman, M., & Gautier, J. (2004). Mre11 assembles linear DNA fragments into DNA damage signaling complexes. *PLoS Biology*, *2*(5). https://doi.org/10.1371/journal.pbio.0020110

Crosetto, N., Bienko, M., Hibbert, R. G., Perica, T., Ambrogio, C., Kensche, T., Hofmann, K., Sixma, T. K, Dikic, I. (2008). Human Wrnip1 is localized in replication factories in a ubiquitin-binding zinc finger-dependent manner. *Journal of Biological Chemistry*, *283*(50), 35173–35185. https://doi.org/10.1074/jbc.M803219200

Daley, J. M., Niu, H., Miller, A. S., & Sung, P. (2015). Biochemical mechanism of DSB end resection and its regulation. *DNA Repair, 32*, 66–74. https://doi.org/10.1016/j.dnarep.2015.04.015

Deans, A. J., & West, S. C. (2011). DNA interstrand crosslink repair and cancer. *Nature Reviews Cancer, 11*(7), 467–480. https://doi.org/10.1038/nrc3088

De Ioannes, P., Malu, S., Cortes, P., Aggarwal, A., (2012). Structural Basis of DNA Ligase IV-Artemis Interaction in Nonhomologous End-Joining. *Cell Reports*, Vol2, Issue 6, 1505–1512. http://dx.doi.org/10.1016/j.celrep.2012.11.004

Dong, H., Nebert, D. W., Bruford, E. A., Thompson, D. C., Joenje, H., & Vasiliou, V. (2015). Update of the human and mouse Fanconi anemia genes. *Human Genomics*, *9*, 32. https://doi.org/10.1186/s40246-015-0054-y

Fagerlie, S. R., & Bagby, G. C. (2006). Immune defects in Fanconi anemia. *Critical Reviews in Immunology*, *26*(1), 81–96. https://doi.org/10.1615/CritRevImmunol.v26.i1.40

Farmer, H., McCabe, N., Lord, C. J., Tutt, A. N. J., Johnson, D. a, Richardson, T. B., Santarosa, M., Dillon, K. J., Hickson, I., Knights, C., Martin, N. M., Jackson, S. P., Smith, G. C., Ashworth, A. (2005). Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature*, *434*(7035), 917–21. https://doi.org/10.1038/nature03445

Gatti M, Pinato S, Maiolica A, Rocchio F, Prato MG, Aebersold R, Penengo L. RNF168 promotes noncanonical K27 ubiquitination to signal DNA damage. Cell reports. 2015 Jan 13;10(2):226-38.

Giri, N., Batista, D. L., Alter, B. P., & Stratakis, C. A. (2007). Endocrine abnormalities in patients with fanconi anemia. *Journal of Clinical Endocrinology and Metabolism*, *92*(7), 2624–2631.

https://doi.org/10.1210/jc.2007-0135

Hanson, P. I., & Whiteheart, S. W. (2005). AAA+ proteins: have engine, will work. *Nature Reviews Molecular Cell Biology*, *6*(7), 519–529. https://doi.org/10.1038/nrm1684

Hishida, T., Iwasaki, H., Ohno, T., Morishita, T., & Shinagawa, H. (2001). A yeast gene, MGS1, encoding a DNA-dependent AAA(+) ATPase is required to maintain genome stability. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(15), 8283–9. https://doi.org/10.1073/pnas.121009098

Hofmann, K. (2009). Ubiquitin-binding domains and their role in the DNA damage response. *DNA Repair*, *8*(4), 544–556. https://doi.org/10.1016/j.dnarep.2009.01.003

Hopfner, K. P., Karcher, A., Craig, L., Woo, T. T., Carney, J. P., & Tainer, J. A. (2001). Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. *Cell*, *105*(4), 473–485. https://doi.org/10.1016/S0092-8674(01)00335-X

Huang, T. T., & D'Andrea, A. D. (2006). Regulation of DNA repair by ubiquitylation. *Nat Rev Mol Cell Biol*, *7*(5), 323–334. https://doi.org/nrm1908 [pii]\r10.1038/nrm1908

Ishiai, M., Kimura, M., Namikoshi, K., Yamazoe, M., Yamamoto, K., Arakawa, H., Agematsu, K., Matsushita, N., Takeda, S., Buerstedde, J. M., Takata, M. (2004). DNA Cross-Link Repair Protein SNM1A Interacts with PIAS1 in Nuclear Focus Formation. *Molecular and Cellular Biology*, *24*(24), 10733–10741. https://doi.org/10.1128/MCB.24.24.10733–10741.2004

Islas AL, Vos JM, Hanawalt PC. (1991) Differential introduction and repair of psoralen photoadducts to DNA in specific human genes. Cancer research; 51(11):2867–73. [PubMed: 2032227]

Jackson, S., & Bartek, J. (2009). The DNA-damage response in human biology and disease. *Nature*, *461*(7267), 1071–1078. https://doi.org/10.1038/nature08467

Jo, U., & Kim, H. (2015). Exploiting the Fanconi Anemia Pathway for Targeted Anti-Cancer Therapy. *Molecules and Cells*, *38*(8), 669–676. https://doi.org/10.14348/molcells.2015.0175

Kanamori, M., Seki, M., Yoshimura, A., Tsurimoto, T., Tada, S., & Enomoto, T. (2011). Werner interacting protein 1 promotes binding of Werner protein to template-primer DNA. *Biological & Pharmaceutical Bulletin*, *34*(8), 1314–1318. https://doi.org/10.1248/bpb.34.1314

Kanu, N., Zhang, T., Burrell, R. A., Chakraborty, A., Cronshaw, J., Costa, C. D., DaCosta, C., Grönroos, E., Pemberton, H. N., Anderton, E., Gonzalez, L., Sabbioneda, S., Ulrich, H. D., Swanton, C., Behrens, A. (2015). RAD18, WRNIP1 and ATMIN promote ATM signalling in response to replication stress. *Oncogene*, (October), 1–11. https://doi.org/10.1038/onc.2015.427

Kawabe, Y. I., Branzei, D., Hayashi, T., Suzuki, H., Masuko, T., Onoda, F., Heo, S. J., Ikeda, H., Shimamoto, A., Furuichi, Y., Seki, M., Enomoto, T. (2001). A Novel Protein Interacts with the Werner's Syndrome Gene Product Physically and Functionally. *Journal of Biological*

Chemistry, *276*(23), 20364–20369. https://doi.org/10.1074/jbc.C100035200

Kawabe, Y. I., Seki, M., Yoshimura, A., Nishino, K., Hayashi, T., Takeuchi, T., Iguchi, S., Kusa, Y., Ohtsuki, M., Tsuyama, T., Imamura, O., Matsumoto, T., Furuichi, Y., Tada, S., Enomoto, T. (2006). Analyses of the interaction of WRNIP1 with Werner syndrome protein (WRN) in vitro and in the cell. *DNA Repair*, *5*(7), 816–828. https://doi.org/10.1016/j.dnarep.2006.04.006

Kim, J. M., Kee, Y., Gurtan, A., & Andrea, A. D. D. (2008). Cell cycle – dependent chromatin loading of the Fanconi anemia core complex by FANCM / FAAP24. *Blood*, *111*(10), 5215–5222. https://doi.org/10.1182/blood-2007-09-113092.An

Kim, Y., Spitz, G. S., Veturi, U., Lach, F. P., & Auerbach, A. D. (2012). Regulation of multiple DNA repair pathways by the Fanconi anemia protein. *Genes and Development*, (Id), 1393–1408. https://doi.org/10.1182/blood-2012-07-441212

King, M. C., Marks, J. H., & Mandell, J. B. (2003). Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. *Science*, *302*(5645), 643–646. https://doi.org/10.1126/science.1088759

Knipscheer, P., Räschle, M., Smogorzewska, A., Enoiu, M., Ho, T. V., Schärer, O. D., Elledge, S. J., Walter, J. C. (2009). The Fanconi anemia pathway promotes replication-dependent DNA Interstrand Cross-Link Repair. *Science*, *326*(5960), 1698–1701.

Komander, D. and Rape, M., 2012. The ubiquitin code. *Annual review of biochemistry*, *81*, pp.203-229.

Kosugi, S., Hasebe, M., Entani, T., Takayama, S., Tomita, M., & Yanagawa, H. (2008). Design of Peptide Inhibitors for the Importin α/β Nuclear Import Pathway by Activity-Based Profiling. *Chemistry and Biology*, *15*(9), 940–949. https://doi.org/10.1016/j.chembiol.2008.07.019

Kosugi, S., Hasebe, M., Matsumura, N., Takashima, H., Miyamoto-Sato, E., Tomita, M., & Yanagawa, H. (2009). Six classes of nuclear localization signals specific to different binding grooves of importin alpha. *Journal of Biological Chemistry*, *284*(1), 478–485. https://doi.org/10.1074/jbc.M807017200

Kratz, K., Schöpf, B., Kaden, S., Sendoel, A., Eberhard, R., Lademann, C., Cannavó, E., Sartori, A. A., Hengartner, M. O., Jiricny, J. (2010). Deficiency of FANCD2-Associated Nuclease KIAA1018/FAN1 Sensitizes Cells to Interstrand Crosslinking Agents. *Cell*, *142*(1), 77–88. https://doi.org/10.1016/j.cell.2010.06.022

Kumar, G. S., Lipman, R., Cummings, J., & Tomasz, M. (1997). Mitomycin C-DNA adducts generated by DT-diaphorase. Revised mechanism of the enzymatic reductive activation of mitomycin C. *Biochemistry*, *36*(46), 14128–14136. https://doi.org/10.1021/bi971394i

Kumari, U., Ya Jun, W., Huat Bay, B., & Lyakhovich, a. (2014). Evidence of mitochondrial dysfunction and impaired ROS detoxifying machinery in Fanconi anemia cells. *Oncogene*, *33*(2), 165–72. https://doi.org/10.1038/onc.2012.583

Lachaud, C., Moreno, A., Marchesi, F., Toth, R., Blow, J. J., & Rouse, J. (2016). Ubiquitinated

Fancd2 recruits Fan1 to stalled replication forks to prevent genome instability. *Scientific Reports*, *351*(62), 846–849. https://doi.org/10.1126/science.aad5634

Lan, L., Nakajima, S., Oohata, Y., Takao, M., Okano, S., Masutani, M., Wilson, S. H, Yasui, A. (2004). In situ analysis of repair processes for oxidative DNA damage in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(38), 13738–43. https://doi.org/10.1073/pnas.0406048101

Langevin, F., Crossan, G. P., Rosado, I. V, Arends, M. J., & Patel, K. J. (2011). Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature*, *475*(7354), 53–58. https://doi.org/10.1038/nature10192

Larminat, F., Zhen, W., Bohr, V.A. (1993). Gene-specific DNA repair of interstrand cross-links induced by chemotherapeutic agents can be preferential. The Journal of biological chemistry. 268(4): 2649–54. [PubMed: 8428941]

Lee, J.-H., & Paull, T. T. (2005). ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. *Science (New York, NY)*, *308*(5721), 551–554. https://doi.org/10.1126/science.1108297

Leuzzi, G., Marabitti, V., Pichierri, P., & Franchitto, A. (2016). WRNIP 1 protects stalled forks from degradation and promotes fork restart after replication stress. *The EMBO Journal*, *35*(13), 1–15. https://doi.org/10.15252/embj.201593265

Liang, C. C., Zhan, B., Yoshikawa, Y., Haas, W., Gygi, S., & Cohn, M. (2015). UHRF1 Is a sensor for DNA interstrand crosslinks and recruits FANCD2 to initiate the Fanconi Anemia pathway. *Cell Reports*, *10*(12), 1947–1957. https://doi.org/10.1016/j.celrep.2015.02.053

Liang, C.-C., Li, Z., Lopez-Martinez, D., Nicholson, W. V, Vénien-Bryan, C., & Cohn, M. A. (2016). The FANCD2-FANCI complex is recruited to DNA interstrand crosslinks before monoubiquitination of FANCD2. *Nature Communications*, *7*, 12124. https://doi.org/10.1038/ncomms12124

Liu, T., Ghosal, G., Yuan, J., Chen, J., & Huang, J. (2010). FAN1 acts with FANCI-FANCD2 to promote DNA interstrand cross-link repair. *Science*, *329*(5992), 1192656. https://doi.org/10.1126/science.1192656.

Longerich, S., San Filippo, J., Liu, D., & Sung, P. (2009). FANCI binds branched DNA and is monoubiquitinated by UBE2T-FANCL. *Journal of Biological Chemistry*, *284*(35), 23182–23186. https://doi.org/10.1074/jbc.C109.038075

MacKay, C., Déclais, A. C., Lundin, C., Agostinho, A., Deans, A. J., MacArtney, T. J., Hofmann, K., Gartner, A., West, S. C., Helleday, T., Lilley, D. M., Rouse, J. (2010). Identification of KIAA1018/FAN1, a DNA Repair Nuclease Recruited to DNA Damage by Monoubiquitinated FANCD2. *Cell*, *142*(1), 65–76. https://doi.org/10.1016/j.cell.2010.06.021

Mah, L.-J., El-Osta, A., & Karagiannis, T. C. (2010). γH2AX: a sensitive molecular marker of DNA damage and repair. *Leukemia : Official Journal of the Leukemia Society of America, Leukemia*

Research Fund, U.K, *24*(4), 679–686. https://doi.org/10.1038/leu.2010.6

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., Church, G. M. (2013). RNA-Guided Human Genome Engineering via Cas9. *Science*, *339(6121)*(February), 823–827.

Marteijn, J. a, Lans, H., Vermeulen, W., & Hoeijmakers, J. H. J. (2014). Understanding nucleotide excision repair and its roles in cancer and ageing. *Nature Reviews. Molecular Cell Biology*, *15*(7), 465–81. https://doi.org/10.1038/nrm3822

Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., Hurov, K. E., Luo, J., Bakalarski, C. E., Zhao, Z., Solimini, N., Lerenthal, Y., Shiloh, Y., Gygi, S. P., Elledge, S. J. (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science*, *316*(5828), 1160–1166. https://doi.org/10.1126/science.1140321

Meindl, A., Hellebrand, H., Wiek, C., Erven, V., Wappenschmidt, B., Niederacher, D., Freund, M., Lichtner, P., Hartmann, L., Schaal, H., Ramser, J., Honisch, E., Kubisch, C., Wichmann, H. E., Kast, K., Deissler, H., Engel, C., Müller-Myhsok, B., Neveling, K., Kiechle, M., Mathew, C. G., Schindler, D., Schmutzler, R. K., Hanenberg, H. (2010). Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nature Genetics*, *42*(5), 410–4. https://doi.org/10.1038/ng.569

Modrich, P., 2006. Mechanisms in eukaryotic mismatch repair. *Journal of Biological Chemistry*, *281*(41), pp.30305-30309.

Mojica, F. J. M., Díez-Villaseñor, C., Soria, E., & Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Molecular Microbiology*, *36*(1), 244–246. https://doi.org/10.1046/j.1365-2958.2000.01838.x

Moser, J., Kool, H., Giakzidis, I., Caldecott, K., Mullenders, L. H. F., & Fousteri, M. I. (2007). Sealing of Chromosomal DNA Nicks during Nucleotide Excision Repair Requires XRCC1 and DNA Ligase III?? in a Cell-Cycle-Specific Manner. *Molecular Cell*, *27*(2), 311–323. https://doi.org/10.1016/j.molcel.2007.06.014

Mouw, K. W., & D'Andrea, A. D. (2014). Crosstalk between the nucleotide excision repair and Fanconi anemia/BRCA pathways. *DNA Repair*, *19*, 130–134. https://doi.org/10.1016/j.dnarep.2014.03.019

Niedernhofer, L. J., Daniels, J. S., Rouzer, C. A., Greene, R. E., & Marnett, L. J. (2003). Malondialdehyde, a product of lipid peroxidation, is mutagenic in human cells. *Journal of Biological Chemistry*, *278*(33), 31426–31433. https://doi.org/10.1074/jbc.M212549200

Nimonkar, A. V., Genschel, J., Kinoshita, E., Polaczek, P., Campbell, J. L., Wyman, C., Modrich, P., Kowalczykowski, S. C. (2011). BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes and Development*, *25*(4), 350–362. https://doi.org/10.1101/gad.2003811

Nomura, H., Yoshimura, A., Edo, T., Kanno, S. I., Tada, S., Seki, M., Yasui, A., Enomoto, T. (2012).

WRNIP1 accumulates at laser light irradiated sites rapidly via its ubiquitin-binding zinc finger domain and independently from its ATPase domain. *Biochemical and Biophysical Research Communications*, *417*(4), 1145–1150. https://doi.org/10.1016/j.bbrc.2011.12.080

Pagano, G., Degan, P., D'Ischia, M., Kelly, F. J., Pallardo, F. V., Zatterale, A., Anak, S. S., Akisik, E. E., Beneduce, G., Calzone, R., De Nicola, E., Dunster, C., Lloret, A., Manini, P., Nobili, B., Saviano, A., Vuttariello, E., Warnau, M. (2004). Gender- and age-related distinctions for the in vivo prooxidant state in Fanconi anaemia patients. *Carcinogenesis*, *25*(10), 1899–1909. https://doi.org/10.1093/carcin/bgh194

Pagano, G., Shyamsunder, P., Verma, R. S., & Lyakhovich, A. (2014). Damaged mitochondria in Fanconi anemia - an isolated event or a general phenomenon? *Oncoscience*, *1*(4), 287–95. https://doi.org/10.18632/oncoscience.33

Pagano, G., Talamanca, A. A., Castello, G., Pallardó, F. V., Zatterale, A., & Degan, P. (2012). Oxidative stress in Fanconi anaemia: From cells and molecules towards prospects in clinical management. *Biological Chemistry*, *393*(1–2), 11–21. https://doi.org/10.1515/BC-2011-227

Parsons, J. L., & Dianov, G. L. (2013). Co-ordination of base excision repair and genome stability. *DNA Repair*, *12*(5), 326–333. https://doi.org/10.1016/j.dnarep.2013.02.001

Paull, T. T., & Gellert, M. (1999). Nbs1 potentiates ATP-driven DNA unwinding and endonuclease cleavage by the Mre11 / Rad50 complex. *Genes and Development*, 1276–1288. https://doi.org/10.1101/gad.13.10.1276

Petermann, E., Keil, C., & Oei, S. L. (2006). Roles of DNA ligase III and XRCC1 in regulating the switch between short patch and long patch BER. *DNA Repair*, *5*(5), 544–555. https://doi.org/10.1016/j.dnarep.2005.12.008

Räschle, M., Knipsheer, P., Enoiu, M., Angelov, T., Sun, J., Griffith, J. D., Ellenberger, T. E., Schärer, O. D., Walter, J. C. (2008). Mechanism of Replication-Coupled DNA Interstrand Crosslink Repair. *Cell*, *134*(6), 969–980. https://doi.org/10.1016/j.cell.2008.08.030

Rink, S. M., & Hopkins, P. B. (1995). A mechlorethamine-induced DNA interstrand cross-link bends duplex DNA. *Biochemistry*, *34*, 1439–1445. https://doi.org/10.1021/bi00004a039

Sarkar, S., Davies, A. A., Ulrich, H. D., & McHugh, P. J. (2006). DNA interstrand crosslink repair during G1 involves nucleotide excision repair and DNA polymerase zeta. *The EMBO Journal*, *25*(6), 1285–94. https://doi.org/10.1038/sj.emboj.7600993

Saugar, I., Parker, J. L., Zhao, S., & Ulrich, H. D. (2012). The genome maintenance factor Mgs1 is targeted to sites of replication stress by ubiquitylated PCNA. *Nucleic Acids Research*, *40*(1), 245–257. https://doi.org/10.1093/nar/gkr738

Schwab, R. A., Nieminuszczy, J., Shah, F., Langton, J., Lopez Martinez, D., Liang, C. C., Cohn, M. A., Gibbons, R. J., Deans, A. J.,Niedzwiedz, W. (2015). The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication and Transcription. *Molecular Cell*, *60*(3), 351–361. https://doi.org/10.1016/j.molcel.2015.09.012

Shibata, A., Moiani, D., Arvai, A. S., Perry, J., Harding, S. M., Genois, M. M., Maity, R., van Rossum-Fikkert, S., Kertokalio, A., Romoli, F., Ismail, A., Ismalaj, E., Petricci, E., Neale, M. J., Bristow, R. G., Masson, J. Y., Wyman, C., Jeggo, P. A., Tainer, J. A. (2014). DNA Double-Strand Break Repair Pathway Choice Is Directed by Distinct MRE11 Nuclease Activities. *Molecular Cell*, *53*(1), 7–18. https://doi.org/10.1016/j.molcel.2013.11.003

Smeaton, M. B., Hlavin, E. M., Mason, T. M., Noronha, A. M., Wilds, C. J., & Miller, P. S. (2008). Distortion-dependent unhooking of interstrand cross-links in mammalian cell extracts. *Biochemistry*, *47*(37), 9920–9930. https://doi.org/10.1021/bi800925e

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., Elledge, S. J. (2010). A Genetic Screen Identifies FAN1, a Fanconi Anemia-Associated Nuclease Necessary for DNA Interstrand Crosslink Repair. *Molecular Cell*, *39*(1), 36–47. https://doi.org/10.1016/j.molcel.2010.06.023

Suhasini, A. N., & Brosh, R. M. (2012). Fanconi anemia and Bloom's syndrome crosstalk through FANCJ-BLM helicase interaction. *Trends in Genetics*, *28*(1), 7–13. https://doi.org/10.1016/j.tig.2011.09.003

Suhasini, A. N., Rawtani, N. A., Wu, Y., Sommers, J. A., Sharma, S., Mosedale, G., North, P. S., Cantor, S. B., Hickson, I. D., Brosh, R. M. (2011). Interaction between the helicases genetically linked to Fanconi anemia group J and Bloom's syndrome. *The EMBO Journal*, *30*(4), 692–705. https://doi.org/10.1038/emboj.2010.362

Sung, P., & Klein, H. (2006). Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nature Reviews. Molecular Cell Biology*, *7*(10), 739–750. https://doi.org/10.1038/nrm2008

Suzuki, N., Rohaim, A., Kato, R., Dikic, I., Wakatsuki, S. and Kawasaki, M., 2016. A novel mode of ubiquitin recognition by the ubiquitin-binding zinc finger domain of WRNIP1. *The FEBS journal*, *283*(11), pp.2004-2017.

Taniguchi, T., Garcia-Higuera, I., Andreassen, P. R., Gregory, R. C., Grompe, M., & D'Andrea, A. D. (2002). S-phase-specific interaction of the Fanconi anemia protein, FANCD2, with BRCA1 and RAD51. *Blood*, *100*(7), 2414–2420. https://doi.org/10.1182/blood-2002-01-0278

Tian, Y., Paramasivam, M., Ghosal, G., Chen, D., Shen, X., Huang, Y., Akhter, S., Legerski, R., Chen, J., Seidman, M. M., Qin, J. Li, L. (2015). UHRF1 contributes to DNA damage repair as a lesion recognition factor and nuclease scaffold. *Cell Reports*, *10*(12), 1957–1966. https://doi.org/10.1016/j.celrep.2015.03.038

Tomasz, M. (1995). Mitomycin C: small, fast and deadly (but very selective). *Chemistry and Biology*, *2*(9), 575–579. https://doi.org/10.1016/1074-5521(95)90120-5

Trujillo, K. M., & Sung, P. (2001). DNA Structure-specific Nuclease Activities in the Saccharomyces cerevisiae Rad50-Mre11 Complex. *Journal of Biological Chemistry*, *276*(38), 35458–35464. https://doi.org/10.1074/jbc.M105482200

Tsurimoto, T., Shinozaki, A., Yano, M., Seki, M., & Enomoto, T. (2005). Human Werner helicase interacting protein 1 (WRNIP1) functions as novel modulator for DNA polymerase delta *Genes to Cells*, *10*(1), 13–22. https://doi.org/10.1111/j.1365-2443.2004.00812.x

Turnbull, C., & Rahman, N. (2008). Genetic predisposition to breast cancer: past, present, and future. *Annual Review of Genomics and Human Genetics*, *9*, 321–45. https://doi.org/10.1146/annurev.genom.9.081307.164339

Ulrich, H. D., & Walden, H. (2010). Ubiquitin signalling in DNA replication and repair. *Nature Reviews. Molecular Cell Biology*, *11*(7), 479–489. https://doi.org/10.1038/nrm2921

Unno, J., Itaya, A., Taoka, M., Sato, K., Tomida, J., Sakai, W., Sugasawa, K., Ishiai, M., Ikura, T., Isobe, T., Kurumizaka, H., Takata, M. (2014). FANCD2 binds CtIP and regulates DNA-end resection during DNA interstrand crosslink repair. *Cell Reports*, *7*(4), 1039–1047. https://doi.org/10.1016/j.celrep.2014.04.005

Vijeh Motlagh, N. D., Seki, M., Branzei, D., & Enomoto, T. (2006). Mgs1 and Rad18/Rad5/Mms2 are required for survival of Saccharomyces cerevisiae mutants with novel temperature/cold sensitive alleles of the DNA polymerase delta subunit, Pol31. *DNA Repair*, *5*(12), 1459–1474. https://doi.org/10.1016/j.dnarep.2006.07.006

Wang, A. T., Sengerová, B., Cattell, E., Inagawa, T., Hartley, J. M., Kiakos, K., Burgess-Brown, N. A., Swift, L. P., Enzlin, J. H., Schofield, C. J, Gileadi, O., McHugh, P. J. (2011). Human SNM1A and XPF – ERCC1 collaborate to initiate DNA interstrand cross-link repair. *Genes and Development*, 1859–1870. https://doi.org/10.1101/gad.15699211

Wang, H., Yang, H., Shivalila, C. S., Dawlaty, M. M., Cheng, A. W., Zhang, F., & Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/cas-mediated genome engineering. *Cell*, *153*(4), 910–918. https://doi.org/10.1016/j.cell.2013.04.025

Warren, A. J., Maccubbin, A. E., & Hamilton, J. W. (1998). Detection of mitomycin c-DNA adducts in vivo by 32P-postlabeling: Time course for formation and removal of adducts and biochemical modulation. *Cancer Research*, *58*(3), 453–461.

Watanabe,K., Tateishi,S., Kawasuji,M., Tsurimoto, T., Inoue, H., Yamaizumi M., (2004), Rad18 guides polη to replication stalling sites through physical interaction and PCNA monoubiquitination. *EMBO*, Sep 29; 23(19): 3886–3896. https://doi.org/10.1038/sj.emboj.7600383

Williams, H. L., Gottesman, M. E., & Gautier, J. (2012). Replication-Independent Repair of DNA Interstrand Crosslinks. *Molecular Cell*, *47*(1), 140–147. https://doi.org/10.1016/j.molcel.2012.05.001

Williams, H.L., Gottesman, M.E. and Gautier, J., 2013. The differences between ICL repair during and outside of S phase. *Trends in biochemical sciences*, *38*(8), pp.386-393.

Williams, R. S., Dodson, G. E., Limbo, O., Yamada, Y., Williams, J. S., Guenther, G., Classen, S.,

Glover, J. N., Iwasaki, H., Russell, P., Tainer, J. A. (2009). Nbs1 Flexibly Tethers Ctp1 and Mre11-Rad50 to Coordinate DNA Double-Strand Break Processing and Repair. *Cell*, *139*(1), 87–99. https://doi.org/10.1016/j.cell.2009.07.033

Yamada, K., Miyata, T., Tsuchiya, D., Oyama, T., Fujiwara, Y., Ohnishi, T., Iwasaki, H., Shinagawa, H., Ariyoshi, M., Mayanagi, K., Morikawa, K. (2002). Crystal structure of the RuvA-RuvB complex: A structural basis for the holliday junction migrating motor machinery. *Molecular Cell*, *10*(3), 671–681. https://doi.org/10.1016/S1097-2765(02)00641-X

Yoshikiyo, K., Kratz, K., Hirota, K., Nishihara, K., Takata, M., Kurumizaka, H., Horimoto, S., Takeda, S., 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*(50), 21553–7. https://doi.org/10.1073/pnas.1011081107

Yoshimura, A., Seki, M. and Enomoto, T., 2017. The role of WRNIP1 in genome maintenance. *Cell Cycle*, *16*(6), pp.515-521.

Yoshimura, A., Kobayashi, Y., Tada, S., Seki, M., & Enomoto, T. (2014). WRNIP1 functions upstream of DNA polymerase eta in the UV-induced DNA damage response. *Biochemical and Biophysical Research Communications*, *452*(1), 48–52. https://doi.org/10.1016/j.bbrc.2014.08.043

Yoshimura, A., Seki, M., Kanamori, M., Tateishi, S., Tsurimoto, T., Tada, S., & Enomoto, T. (2009). Physical and functional interaction between WRNIP1 and RAD18. *Genes & Genetic Systems*, *84*(2), 171–178. https://doi.org/10.1266/ggs.84.171

Zeman, M. K., & Cimprich, K. A. (2014). Causes and consequences of replication stress. *Nature Cell Biology*, *16*(1), 2–9. https://doi.org/10.1038/ncb2897

Zhu, Z., Chung, W. H., Shim, E. Y., Lee, S. E., Ira, G. (2008). Sgs1 Helicase and Two Nucleases Dna2 and Exo1 Resect DNA Double–Strand Break Ends. Cell, 134(6), 981–994. https://doi.org/10.1016/j.cell.2008.08.037), 981–994. https://doi.org/10.1016/j.cell.2008.08.037

APPENDIX

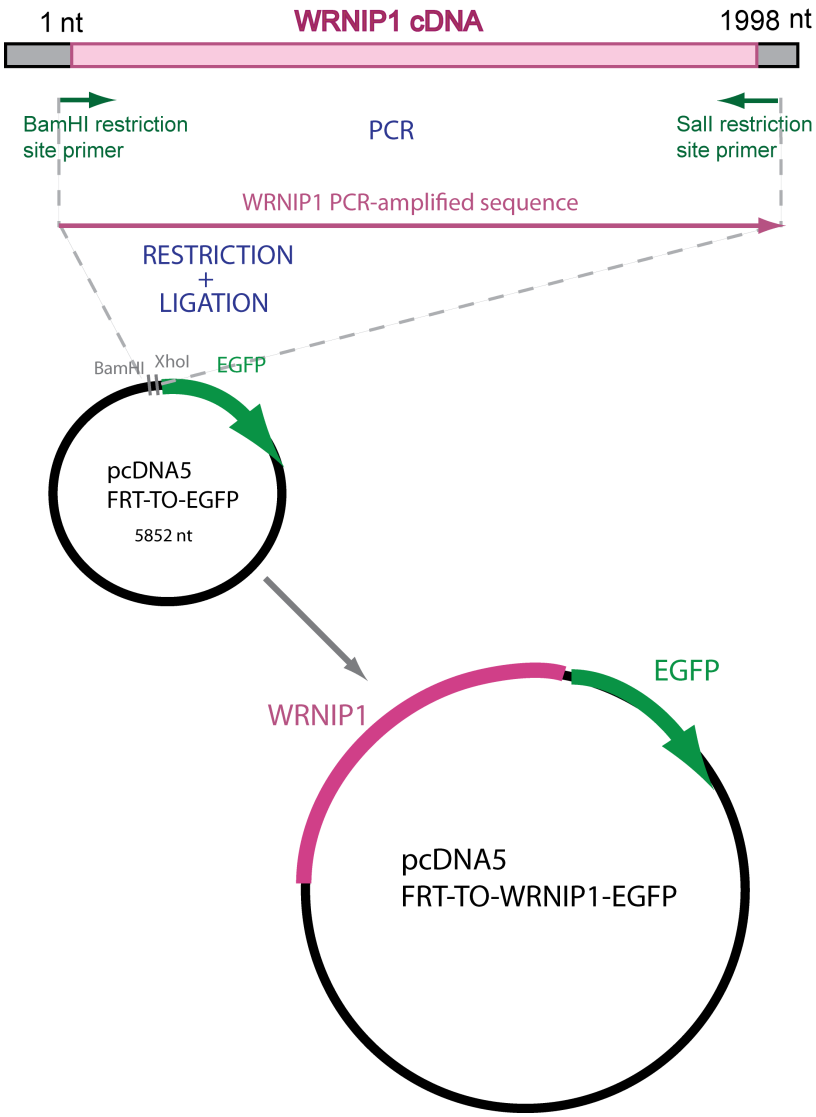


Figure A1: pcDNA5-FRT-TO(EGFP) plasmid.

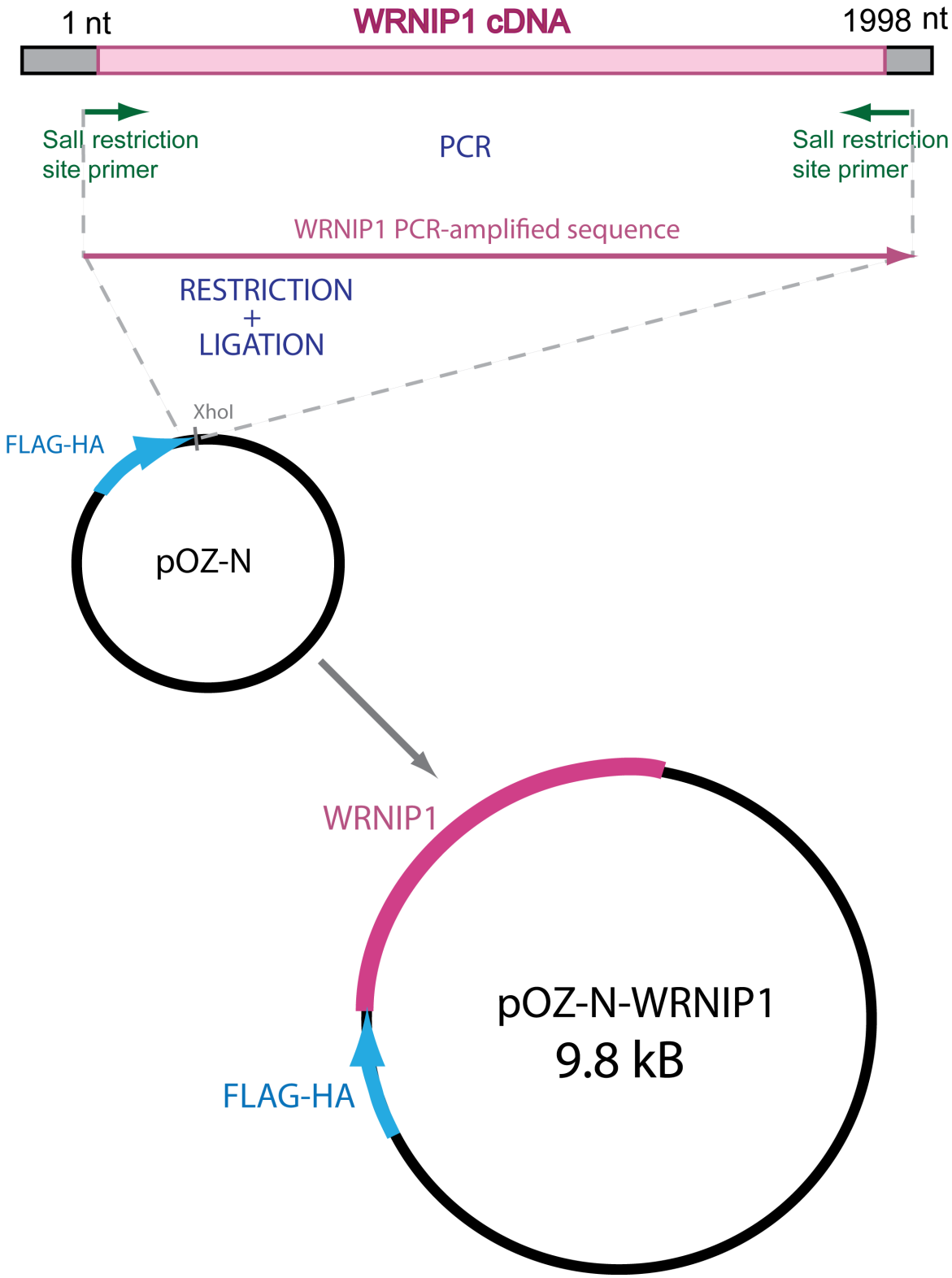


Figure A2: pOZ-N(FLAG-HA-WRNIP1) construct.

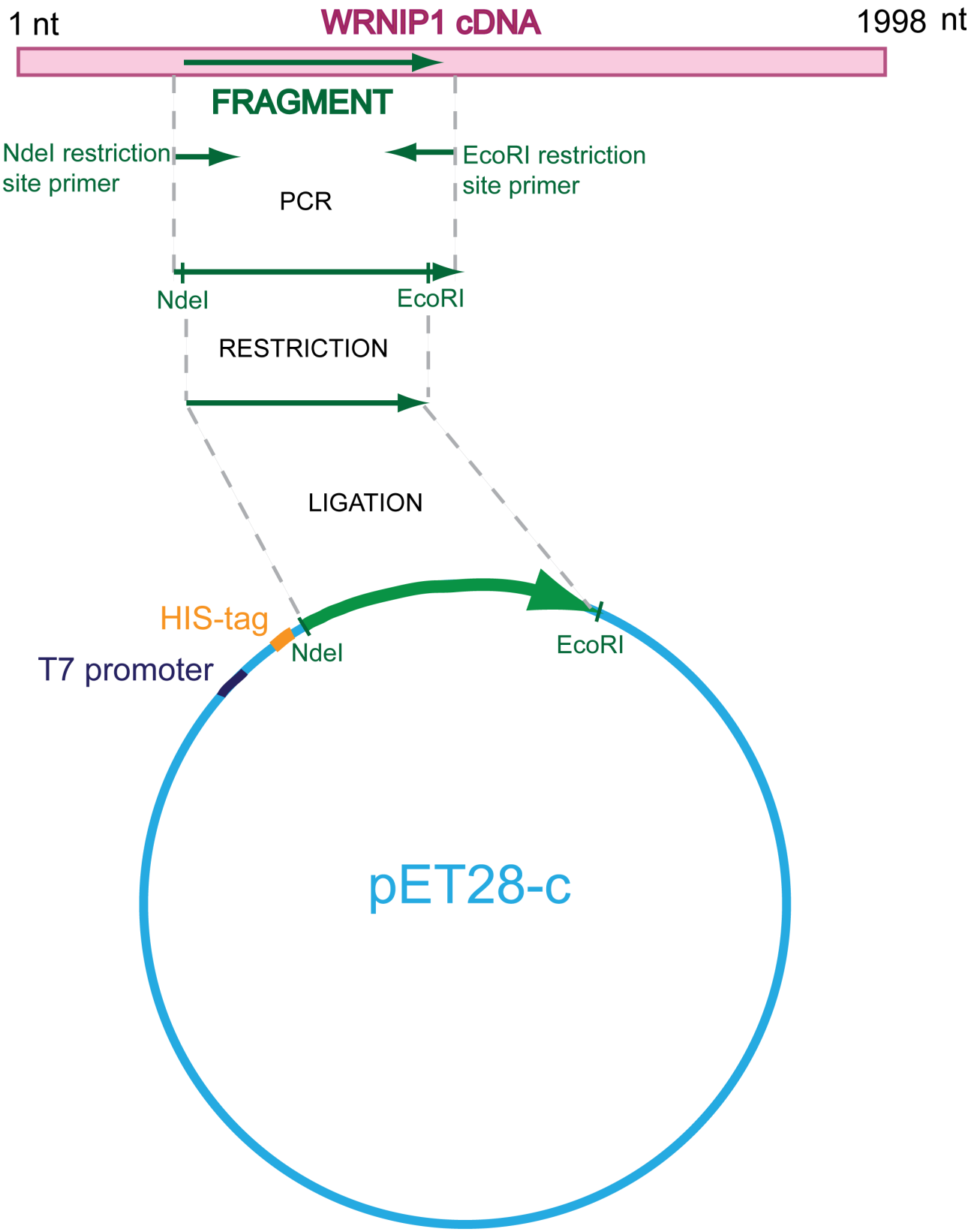


Figure A3: pET(His-WRNIP1_fragment) construct.

Figure A4: pX459 plasmid.

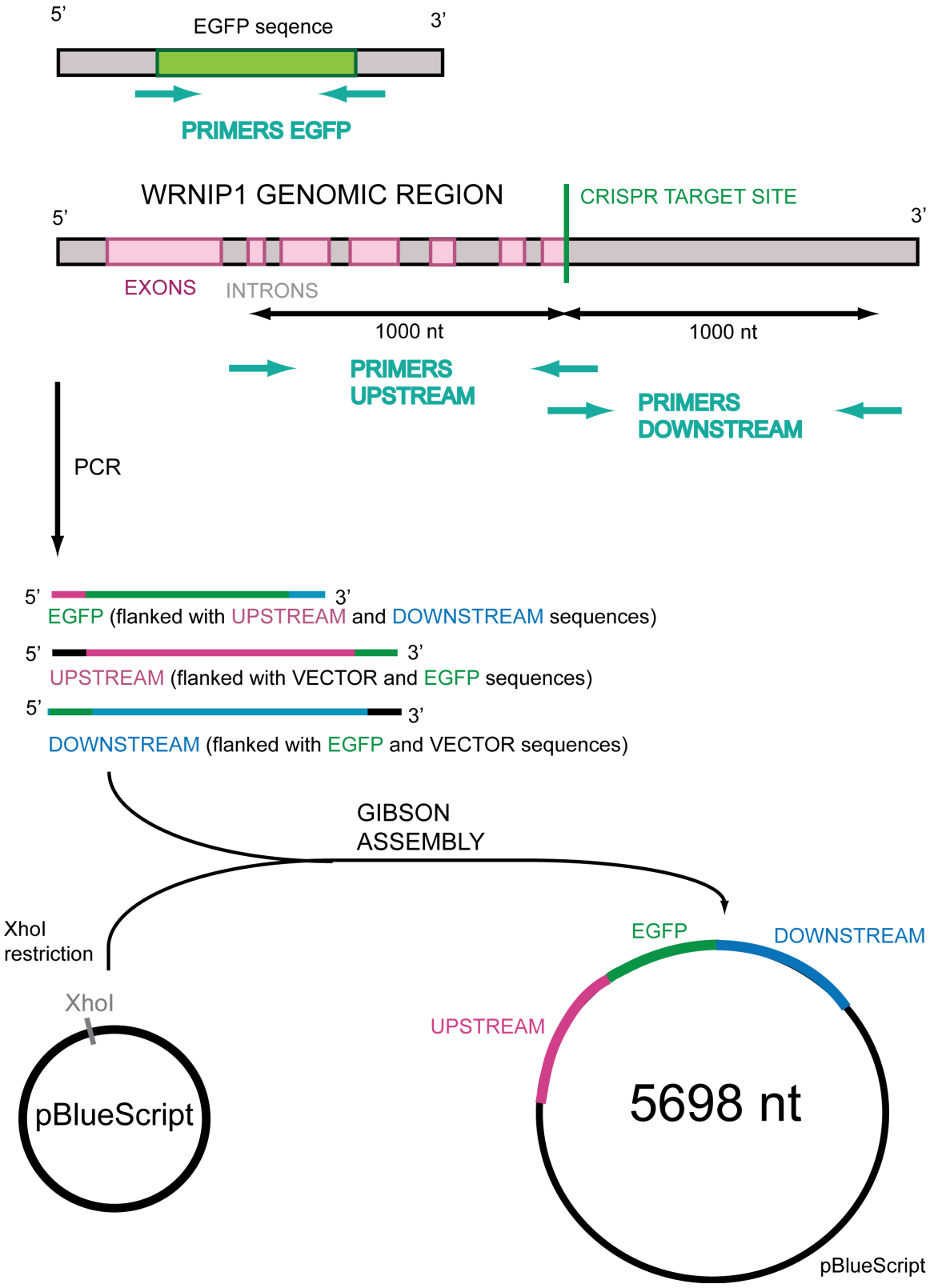


Figure A5: Gibson assembly.