Understanding the Genetic Relationship between Key Effectors of DNA Interstrand Crosslink Repair

Sanja Brolih
Green Templeton College
University of Oxford

A thesis submitted to the University of Oxford for the degree of

***MSc by Research* in Oncology**

Michaelmas 2016

ABSTRACT

Understanding the Genetic Relationship between Key Effectors of DNA Interstrand Crosslink Repair

A thesis submitted to the University of Oxford for the degree of *MSc by Research* in Oncology

With each cell division, our whole genome is duplicated in an error-free and controlled manner. Failure to do so leads to genomic instability. DNA interstrand crosslinks (ICLs) are particularly toxic DNA lesions as when left unrepaired, they are able to block replication producing devastating health conditions such as the genetic syndrome Fanconi Anemia (FA) that predisposes its patients to cancer.

The repair of ICLs is thought to be initiated by the concerted action of multiple nucleases, some of which are defective in FA patients. In particular, the 5'-3' exonuclease SNM1A works in a common pathway with the heterodimeric endonuclease XPF (ERCC4/FANCQ)–ERCC1 during replication-associated ICL repair. We wish to investigate the role of the less well-characterised nuclease FAN1. It has previously been shown that cells deficient in key nucleases XPF-ERCC1, SNM1A and FAN1 all show increased sensitivity to widely used crosslinking anticancer drugs, underlining the clinical importance of exemplifying the exact mechanism of ICL repair.

The project both focuses on a biochemical and cellular characterization of the nuclease FAN1. Through nuclease assays on radioactively end-labelled synthetic DNA substrates mimicking replication fork intermediates, FAN1 was characterised as a 5' flap endonuclease as well as a 5' to 3' exonuclease. On a cellular level, a functional redundancy between the two repair nucleases FAN1 and SNM1A was identified. In order to better understand the genetic relationship between FAN1 and the key exonuclease XPF-ERCC1, I generated a XPF -/- human cell line using the CRISPR/Cas9 gene-editing technology. In future, this tool will enable us to better understand the genetic relationship between XPF-ERCC1 and FAN1 leading to a more complete understanding of ICL repair.

Altogether, the project offers an overview of the biochemical and genetic relationship between key nucleases of the DNA ICL repair and can offer interesting mechanistic insight in ICL incision, unhooking and repair, potentially identifying new therapeutic targets.

ACKNOWLEDGEMENTS

My first thank you goes to my supervisor Prof. Peter McHugh and his invaluable advice and support he has given me throughout my project. I am also very grateful to all members of the McHugh lab, current and past: Anna, Lianne, Ghadeer, Sal and Erkin for their support and help.

Thank you to Lonnie for being a great teacher all year long, to Susie for being such a great pal and the daily top quality banter, and Hannah for being my life pal and being the only one enjoying our endless rowing chat. Thank you Rami for showing me the world is a tough place and keeping me entertained everyday. Special thanks to Sook and Ummi for helping me integrate into the lab and get the project started.

I would also like to thank my Green Templeton family and rowing team, for always being there and providing support outside of the lab.

Final thanks go to my family and parents, because nothing could ever be possible without the support of such great individuals.

TABLE OF CONTENT

INTRODUCTION

With each cell division, our whole genome is duplicated in an error-free and controlled manner. Failure to do so leads to genomic instability, which constitutes a key step in the process of cell transformation and causes a variety of other human diseases (Hanahan et al. 2011). A variety of DNA lesions affecting the cell's viability have been reported (Schärer 2003; Iyama & Wilson 2013). In order to prevent accumulation of DNA damage which is detrimental to the cell's survival and function, a wide array of DNA repair mechanisms have evolved (Iyama & Wilson 2013).

The importance of understanding the details of DNA repair mechanisms becomes clear when looking at several rare genetic disorders that are based on defects in DNA damage repair proteins causing devastating health consequences. Conditions such as Fanconi Anemia (FA), Xeroderma Pigmentosa (XP) or Cockayne Syndrome (CS) all have deficient DNA repair pathways leading to symptoms such as premature aging, neurological disorders or predisposition to cancer (Knoch et al. 2011). The cell's inability to survive with increased DNA damage levels is also explored in cancer treatment (O 'Connor 2015). In fact, a variety of anti-cancer treatments' cytotoxicity is based on DNA damage induction. Unfortunately, the use of DNA damaging agents as an anti cancer therapy is limited by dose limiting side effects as well as the development of drug resistance (Cheung-Ong, Giaever & Nislow 2013). Understanding the molecular details of DNA repair pathways is therefore crucial for both elucidating the origins of a variety of genetic syndromes as well as improving anti-cancer treatments (Helleday et al. 2008).

Endogenous and exogenous DNA damage sources

Constant genomic maintenance is essential for the viability and longevity of a healthy organism (Iyama & Wilson 2013). Our DNA is under constant threat from both endogenous and exogenous sources that can cause a variety of DNA lesions. If left unrepaired, accumulation of DNA damage is a major trigger of mutagenesis, ageing and eventually the development of cancer (Hanahan et al. 2011).

Most endogenous sources of DNA damage involve reactive molecules, usually by-products of essential cellular metabolism reactions. Most of these reactions are based on the chemical instability of DNA bases (hydrolysis or depurination), lipid oxidation or the effect of reactive oxygen species (ROS) and other metabolites (Jackson & Loeb 2001). For instance, the N-glycosidic bond between a DNA base and the deoxyribose duplex is subject to spontaneous hydrolysis cleaving to an abasic or apurinic (AP) site. Hydrolytic deamination can lead to the conversion of a cytosine to a uracil base (Lindahl 1993). Oxidative damage of DNA is cause by reactive oxygen species that are generated as a by-product of cellular metabolism. Examples of ROS include superoxide (O_2-.) generated during mitochondrial respiration that can be converted to hydrogen peroxide (H_2O_2). Other reactive species include hydroxyl radicals (OH^-) and singlet oxygen ($^1O_2$) (Frenkel 1992; Cadet et al. 2011). All together, these lead to a variety of DNA damage such as base modifications, DNA single-strand and double-strand breaks as well as DNA-protein crosslinks and intra-and interstrand crosslinks (ICLs) (Helleday et al. 2014). Polyunsaturated fatty acid residues of phospholipids are also sensitive to oxidation. Lipid hydroperoxides can react with metals to produce major aldehydes products that can damage the DNA by the formation of exocyclic adducts and DNA crosslinks (Iyama & Wilson 2013).

DNA damage from exogenous sources can come from a variety of different environmental factors, including non-ionising and ionising radiation (Friedberg et al. 2004). One of the major origins of DNA lesions is exposure to ultraviolet (UV) light from the sun (Nikitaki et al. 2015). DNA has the ability to absorb photons from the UV light that excite DNA bases leading to the formation of bi-pyrimidine sequences. Those large DNA adducts can be divided into two main classes, both covalently linking two adjacent pyrimidine bases: Cyclobutane bi-pyrimidine dimers (CPD) and 6-4 pyrimidone photoproducts (6-4PP) (Mouret et al. 2006).

Ionising radiation (IR) can have both a direct and indirect effect on DNA causing a variety of complex DNA lesions. IR can be both naturally occurring (such as in gamma radiation) or from an artificial source (such as X-rays or radiotherapy) (Santivasi & Xia 2014). The consequent DNA lesions are caused by either the direct one-electron oxidation of DNA or the indirect water radiolysis leading to the formation of toxic ROS (Cadet & Wagner 2013). Repair of DNA double-strand breaks (DSBs) generated by IR is crucial, as they constitute one of the most harmful forms of damage (Santivasi & Xia 2014).

Another major source of exogenous DNA damage can be identified when looking at chemotherapeutic agents. Some of the most widely used anti-cancer therapies rely on inducing DNA damage to selectively kill cancer cells (Fojo 2001). One of the predominant mechanisms is based on introducing DNA interstrand crosslinks (ICLs) through alkylating agents (Cheung-Ong, Giaever & Nislow 2013).

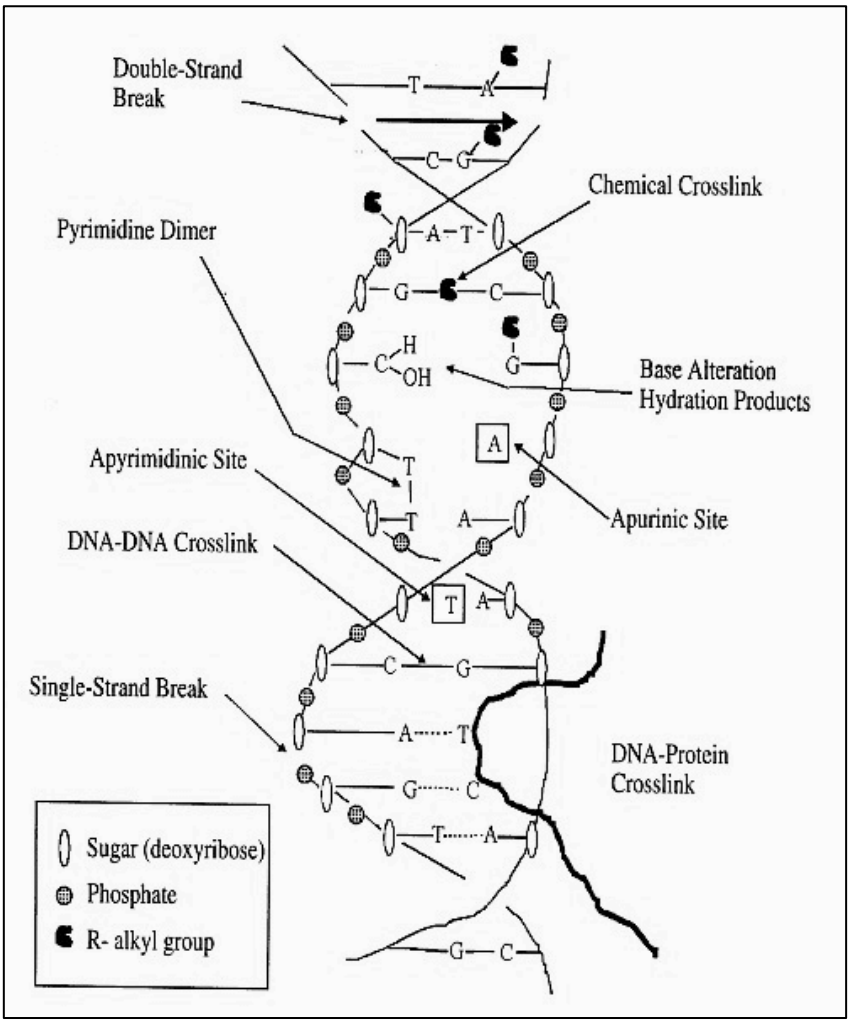


Figure 1- Overview of the different types of DNA damage that can occur

Drawing of an overview of the different types of DNA damage that can occur. DNA damage can arise from both endogenous and exogenous sources causing different lesions varying from bulky additions to single or double strand DNA breaks. From *The Molecular Basis of Aging*, Alvaro Macieira-Coelho; 1995.

The DNA Damage Response (DDR)

Cells have developed an array of signalling cascades to detect and promote the repair of the various types of DNA lesions. The so-called 'DNA Damage response' (DDR) comprises the different transduction cascades responsible for a number of cellular fates depending on the severity of the DNA damage. Cells will either enter cell cycle arrest in order to allow time to repair the DNA damage and activate DNA damage repair pathways, or will engage in cellular senescence or cell death responses if DNA damage is non repairable (Jackson & Bartek 2009). Although each DNA lesion will lead to a different repair response, it is important to notice that all pathways have a certain degree of overlap and share common mechanistic features (Maréchal & Zou 2013; Jackson & Bartek 2009). The DNA damage response is a crucial mechanism for maintaining genome integrity and any defects constitute one of the initiators of cancer (Pearl et al. 2015).

To describe DDR in short, recognition of DNA damage or replication stress will cause the recruitment of DNA damage sensors that initiate DDR by activating the serine/threonine-protein kinases ATM and ATR, phosphorylating downstream effectors (Maréchal & Zou 2013). Depending on the degree of damage in the cell, activation of either the survival pathway or apoptotic pathway (programmed cell death) will occur. The survival pathway will trigger cell cycle arrest to allow time for DNA repair (Shiloh 2003). Errors in DNA repair or unrepaired lesions can lead to the accumulation of mutations and consequent malignant transformation of the cell (O 'Connor 2015).

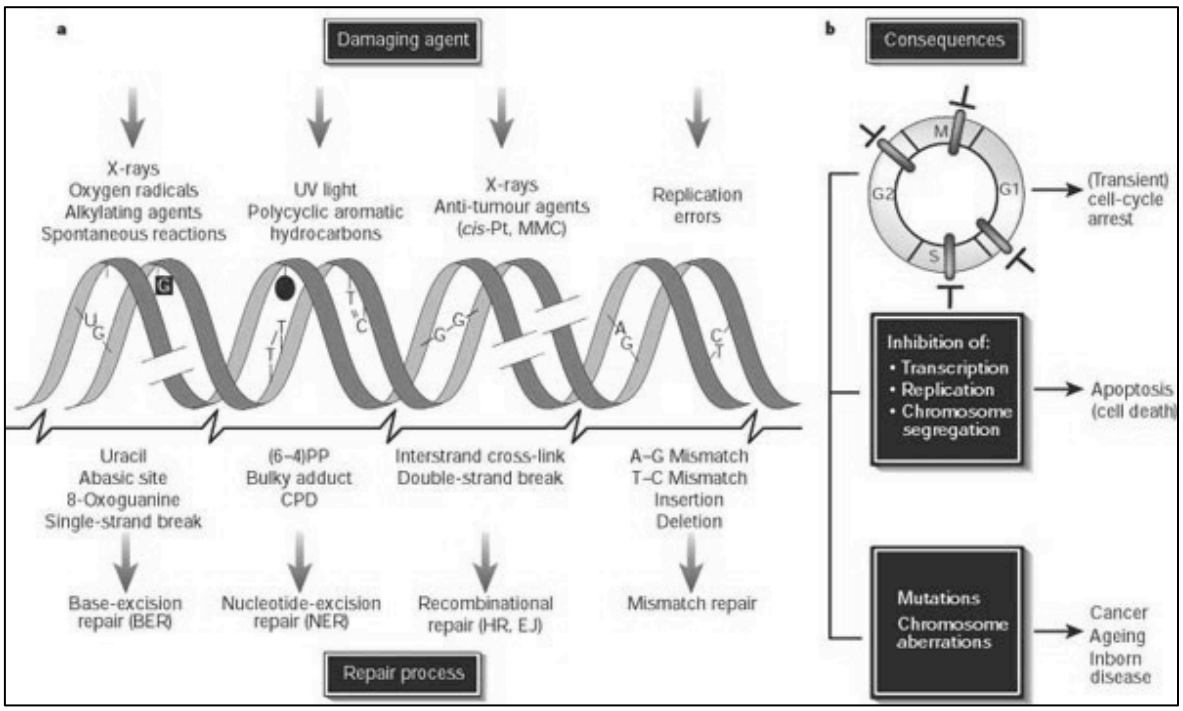


Figure 2- The DNA Damage Response and cellular responses to DNA Damage

Schematic summarising the DNA damage response and the main cellular responses. The DNA damage response is a crucial mechanism for maintaining genome integrity. Different damaging agents trigger the activation of specific repair pathways in order to prevent accumulation of DNA damage. Failure to repair the DNA lesion can have detrimental effects on the cell's survival, leading to the activation of cellular death or mutations or chromosomal aberrations that can initiate disease progression (Hoeijmakers 2001).

Types of DNA Repair Pathways

A large number of DNA repair mechanisms exist to match the variety of DNA lesions that can occur through both the endogenous and exogenous sources (Iyama & Wilson 2013). Although some repair mechanism can occur directly, most of them involve the concerted action of multiple repair proteins through a series of catalytic events (Jackson & Bartek 2009). The main repair pathways include mismatch repair

(MMR), base excision repair (BER), translesion DNA synthesis (TLS), nucleotide-excision repair (NER) as well as non-homologous end joining (NHEJ) and homologous recombination (HR) that are more specific to resolution of double strand breaks.

BER is the main pathway for the repair of chemically modified or damaged DNA bases resulting from deamination, alkylation or oxidation. Once the damaged base is recognized, it is excised by DNA glycosylases hydrolysing the N-glycosidic bond between the base and deoxyribose. This creates an apurinic/apyrimidinic (AP) site that is subsequently removed by the AP endonuclease APE1 as the correct nucleotide is incorporated and the resulting nick ligated (Schärer 2003).

MMR mainly eliminates base-base mismatches or nucleotide insertions/deletions resulting from DNA polymerase errors. The damage is recognised by the MutSα-MutLα complex that recruits the exonuclease EXO1 to the damage site. This triggers a single strand incision by EXO1 upstream of the mismatch that then leads to degradation of the DNA strand past the mismatch base generating a single stranded DNA gap. The repair process is then completed by filling in the DNA gap by DNA Pol δ and nick ligation by DNA ligase I (Jiricny 2006).

Translesion DNA synthesis (TLS) is a DNA damage tolerance mechanism that allows the DNA replication machinery to bypass a DNA lesion. TLS is a unique process in which specialised, low-fidelity DNA polymerases (Pol η, ι, κ and Rev1 in the Y-family and Polζ in the B family) are recruited to the site of damage by ubiquitinated PCNA and are able to replicate the DNA across from a DNA lesion. Contrary to most other DNA repair mechanisms, TLS is error-prone and predisposes the cell to

mutagenesis. This process is crucial for stalled replication fork restart thereby avoiding the cell's death if the replication block was to persist (Waters 2009).

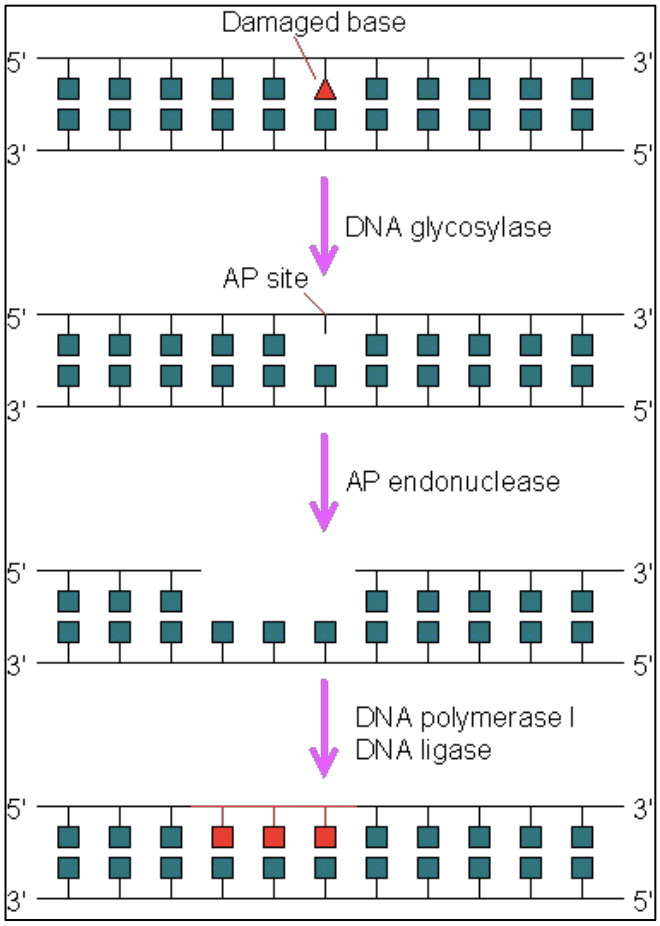


Figure 3- Base Excision repair (BER) schematic in mammalian cells.

When a damaged base is recognized, it is removed by DNA glycosylases creating an AP site. This AP site then in turn serves as a signal for AP endonuclease APE1 that creates a 5' incision to the AP site. DNA polymerase I subsequently is able to incorporate undamaged nucleotides and DNA ligase seals the nicks to complete DNA repair (Schärer 2003).

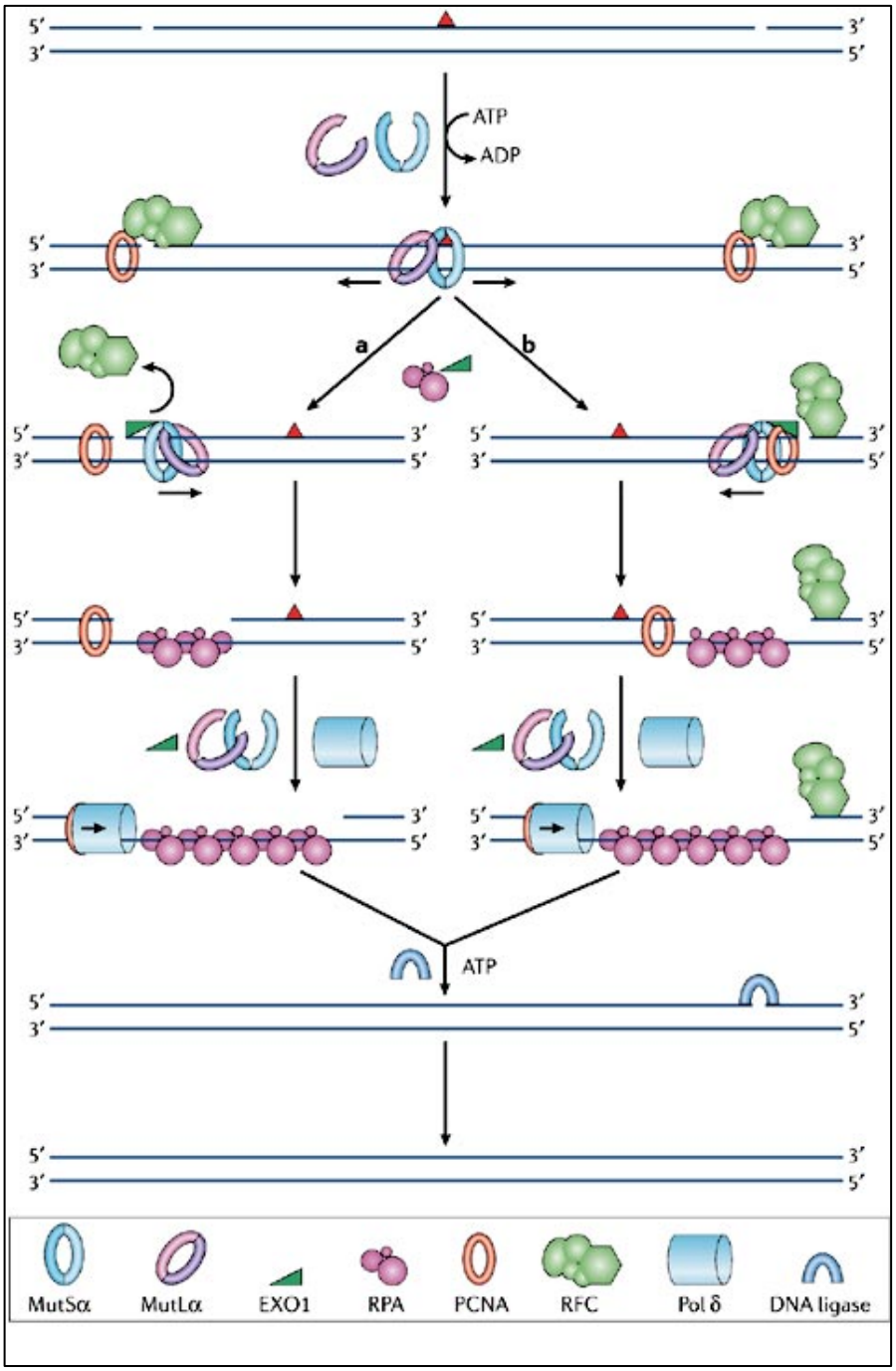


Figure 4- Schematic of the mismatch repair pathway in mammalian cells

The DNA lesion represented in the red square is recognised by the MutSα-MutLα complex that recruits the exonuclease EXO1 to the damage site. The complex undergoes conformation change in an ATP-driven process that enables it to load onto the DNA strand and translocate away from the damage. In both a) and b), Exo1 is the key nuclease involved in the incision of the lesion and degradation of the damaged base. The gap is filled by DNA Polδ and the nick ligated to complete repair (Jiricny 2006).

Helix distorting lesions, as such caused by UV light and chemotherapeutics are processed by nucleotide excision repair (NER) in mammalian cells (Schärer 2013). NER can be divided into two sub-pathways: global-genome NER (GG-NER) that targets the entire genome and transcription-coupled NER (TC-NER) that is specific to lesions that inhibit RNA polymerase II progression and can cause transcription fork collapse or transcription block (Jackson & Bartek 2009). Although both pathways differ in the DNA damage recognition stage, they share the same downstream process.
In GG-NER, DNA damage recognition is mediated by the XPC-RAD23B complex that recognises structural changes in the DNA and binds to the undamaged strand opposite the lesion. In TC-NER, arrested RNA polymerase at the lesion serves as a signal for CS proteins, CSA and CSB that initiate the removal of the lesion (Iyama & Wilson 2013). Once a recognition signal has been initiated, the multi-subunit transcription factor II H (TFIIH) complex is recruited to the damage site. Two of the subunits of this complex, the helicases XPB and XPD unwind the DNA and promote the creation of a 'bubble' structure. All together, this enables the assembly of the pre-incision complex composed of the recognition complexes, the already present helicases as well as associated proteins XPA and RPA. TFIIH (and the XPC-RAD23B complex in the case of GG-NER) are released from the damage site recruiting the endonucleases XPF-ERCC1 and XPG. Both endonucleases are required for the 5' and 3' incisions around the damage site, respectively. The DNA fragment containing the lesion is then removed and repair is initiated. DNA polymerase δ, κ and ε perform gap filling repair synthesis in cooperation with associated factors RFC and PCNA. The nick is sealed by DNA ligases and repair is complete. In TC-NER, ubiquitination and degradation of CSB restores transcription after completion of DNA repair (Iyama & Wilson 2013; Schärer 2013).

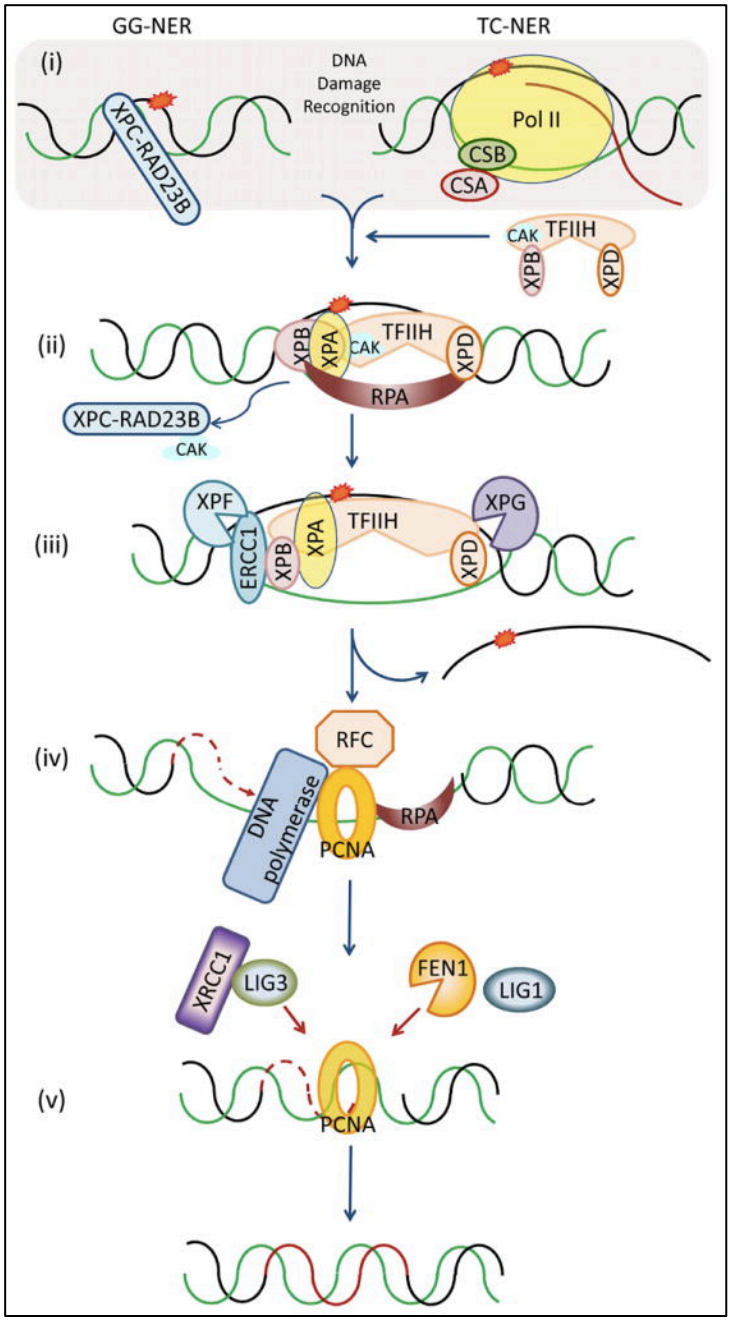


Figure 5- Schematic representation of nucleotide excision repair (NER) in mammalian cells

NER can be divided into two sub-pathways. In GG-NER, DNA damage is recognised by the XPC/RAD23B complex. In TC-NER, arrested RNA polymerase at the lesion serves as a signal for CS proteins, CSA and CSB that initiate the removal of the lesion. Both pathways share the same downstream reactions: together with the helicases XPB and XPD, transcription factor II H (TFIIH) unwinds the DNA creating a bubble structure that enables the endonucleases XPF-ERCC1 and XPG to incise on both ends of the lesion. The resulting gap is filled by DNA Polymerases and re-ligated (Iyama & Wilson 2013).

There are two distinct mechanisms involved in double-strand break repair: Non-homologous end joining (NHEJ) and homologous recombination (HR) (Li & Heyer 2008; Lieber 2010). Due to the lack of involvement of a homologous DNA template, NHEJ is a more error-prone pathway but holds the advantage of being able to operate in any phase of the cell cycle. In contrast, HR mediates faithful repair but is restricted to S and G2 phase. Ineffective repair of DSB can result in chromosomal rearrangements and subsequent loss of genetic information (Lieber 2010).

NHEJ involves the direct ligation of the broken DNA ends for repair. Because of the various sources of DSB formation, the repair factors involved in NHEJ have evolved a certain degree of mechanistic adaptability and structural tolerance enabling them to detect a large variety of DNA end chemistries (Lieber 2010).
In NHEJ, DNA damage recognition is mediated by the Ku heterodimer that binds around the exposed ends and triggers recruitment of repair proteins. In fact, Ku serves as a scaffold for DNA-PKcs kinases that activate a cascade of end-processing enzymes (such as nucleases) and downstream NHEJ factors bringing the exposed DNA ends together. During this process, small deletions of DNA sequences can occur making this pathway more error-prone. DNA ligase IV complex completes repair by sealing the ends (Jackson & Bartek 2009).

HR is characterized by the use of homologous DNA of the undamaged sister chromatid as a template for high fidelity repair of DSBs. HR involves two major steps: the creation of single-stranded DNA (ssDNA) overhangs and strand invasion (Li & Heyer 2008). The MRN complex (which includes Mre11-RAD50-NBS1) mediates recognition of DSBs recruiting processing factors to the site of damage via activation of the ATM

signalling cascade (Jackson & Bartek 2009). One of these processing factors, CtIP, associates with the exonuclease Mre11 to create 3' ssDNA overhangs that are required for strand exchange. The ssDNA serves as a substrate for RPA recruiting the repair protein Rad51 to the damage site. Rad51-dependent strand invasion causes the formation of a heteroduplex DNA structure containing the invading 3' overhang strand and the template strand together with the displaced parental strand known as a displacement loop (D-loop) structure. The 3' end of the invading strand is used as a primer for synthesis of the new DNA strand using the sister chromatid as a template (Li & Heyer 2008). Repair synthesis leads to the formation of a Holliday Junction (HJ) structure that can be resolved via one of the independent HR sub-pathways restoring two intact double-stranded DNA molecules. One of theses HR sub-pathways is mediated by structure selective endonucleases, including GEN1, SLX1-SLX4 and Mus81-EME1 (Moynahan & Jasin 2010). HR has been implicated in rescue of stalled and collapsed replication forks and ICL repair, where a number of these endonucleases play key roles (Helleday et al. 2014; Saleh-Gohari et al. 2005).

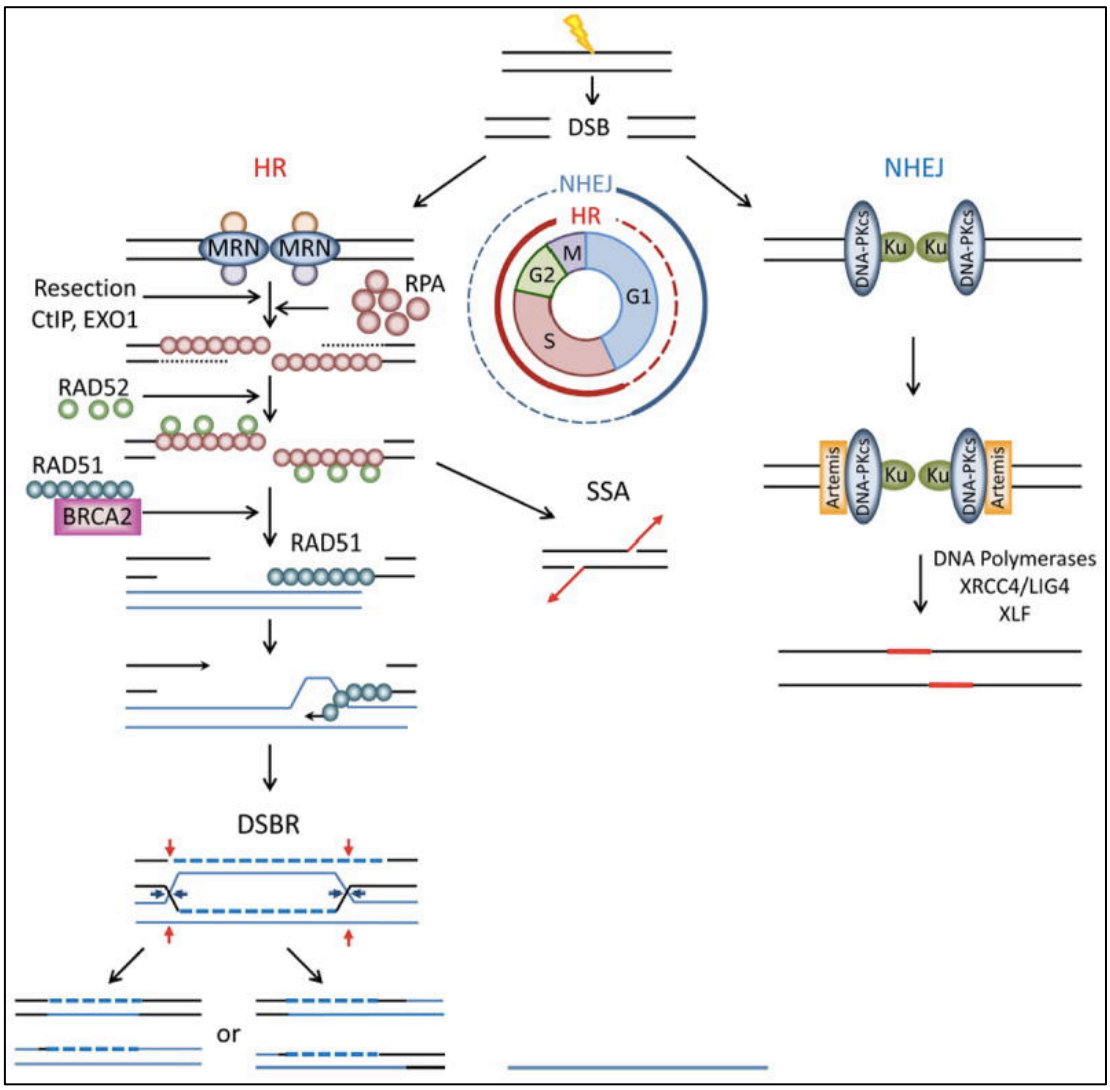


Figure 6- Schematic representation of double-strand break repair (DSBR) in mammalian cells

Double strand break repair occurs through non-homologous end joining (NHEJ) or homologous recombination (HR). Both pathways differ in when they are activated in the cell cycle and in their repair fidelity. NHEJ is a more error-prone pathway that involves the direct ligation of the broken DNA ends for repair mediated by the Ku heterodimer that allows DNA-PKcs to contact DNA ends. HR is characterized by the use of homologous DNA of the undamaged sister chromatid as a template for high fidelity repair of DSBs. HR involves two major steps: the creation of ssDNA overhangs and strand invasion mediated by the MRN recognition complex, structure specific nucleases and further processing factors. Repair synthesis leads to the formation of a Holliday Junction (HJ) structure that can be resolved via one of the independent HR sub-pathways restoring two intact double-stranded DNA molecules (Iyama & Wilson 2013).

DNA Interstrand cross-links (ICLs): Sources and Repair Mechanisms

DNA interstrand crosslinks (ICLs) are extremely toxic lesions that covalently link the complementary DNA strands thereby blocking essential cellular processes such as DNA replication and transcription (Deans & West 2011). Their repair is believed to be primarily replication dependent, occurring when a stalled replication fork is detected in S-phase (Räschle et al. 2008). Although there is evidence for a replication-independent ICL repair pathway, it will not be discussed in this report (Enoiu et al. 2012).

Their recognition and repair is initiated and promoted by the Fanconi Anemia pathway proteins (Patel & Joenje 2007). Because the lesion involves a covalent link between the DNA strands, its repair process is unique in the sense that incisions on both strands are required to remove the lesion. This makes ICL repair a complex multi-step process involving the concerted action of several DNA repair proteins and structure-specific nucleases (Kottemann & Smogorzewska 2013). Failure to repair an ICL can have detrimental effects on cell survival, in particular rapidly dividing cells, thereby explaining the efficacy of crosslinking agents in anti-cancer therapy (Cheung-Ong, Giaever, Nislow, et al. 2013). Individuals suffering from the rare genetic syndrome Fanconi Anemia (FA), display heightened sensitivity to common crosslinking agents underlining the importance of FA proteins in ICL resolution (Cheryl Clauson et al. 2013).

ICLs are both caused by reactive molecules, in particular bi-functional alkylating agents, that arise in the natural processes of cellular metabolism, exposures to exogenous factors and are clinically more relevant, by their use as chemotherapeutic

agents (Iyama & Wilson 2013). The major source of endogenous ICL formation is due to reactive aldehydes that can arise as a by-product of cellular metabolism or lipid peroxidation (Muniandy et al. 2010). Mutations or defects in the ICL repair pathway can have detrimental health consequences and are believed to be the cause of the rare genetic syndrome FA. FA patients are characterised by genetic instabilities and are predisposed to cancer and display severe growth delay and bone marrow failure (Hashimoto et al. 2016).

ICL-inducing agents are widely used in anti-cancer therapy due to their extreme cytotoxicity. Common crosslinking reagents include Nitrogen Mustard (HN2), Mytomycin C (MMC) as well as Cisplatin and its common derivatives (Cheung-Ong, Giaever & Nislow 2013). They mainly differ in the location and chemistry of ICL formation in the DNA as well as the percentage of ICL lesions that occur as a function of the total lesions induced and their degree of structural distortion of the DNA. Nitrogen mustards are small, bi-functional compounds that contain a reactive amine group that crosslinks with the nitrogen group of guanine. This leads to formation of ICLs in the major groove of the DNA. In contrast, MMC creates ICLs in the minor groove of the DNA due to its quinone ring that reacts with the amine group of guanine (Muniandy et al. 2010). The most clinically widely used compounds are Cisplatin and its derivatives. Cisplatin is one of the first members of platinum containing anti-cancer drugs. Cisplatin reacts with the guanine bases of DNA, forming ICLs at GpC sites in the DNA. Cisplatin is reported to create the largest helical distortion of the DNA, forming 96% of ICLs, 2% mono-adducts, 1% intrastrand crosslinks and 1% DNA-protein crosslinks (Hashimoto et al. 2016).

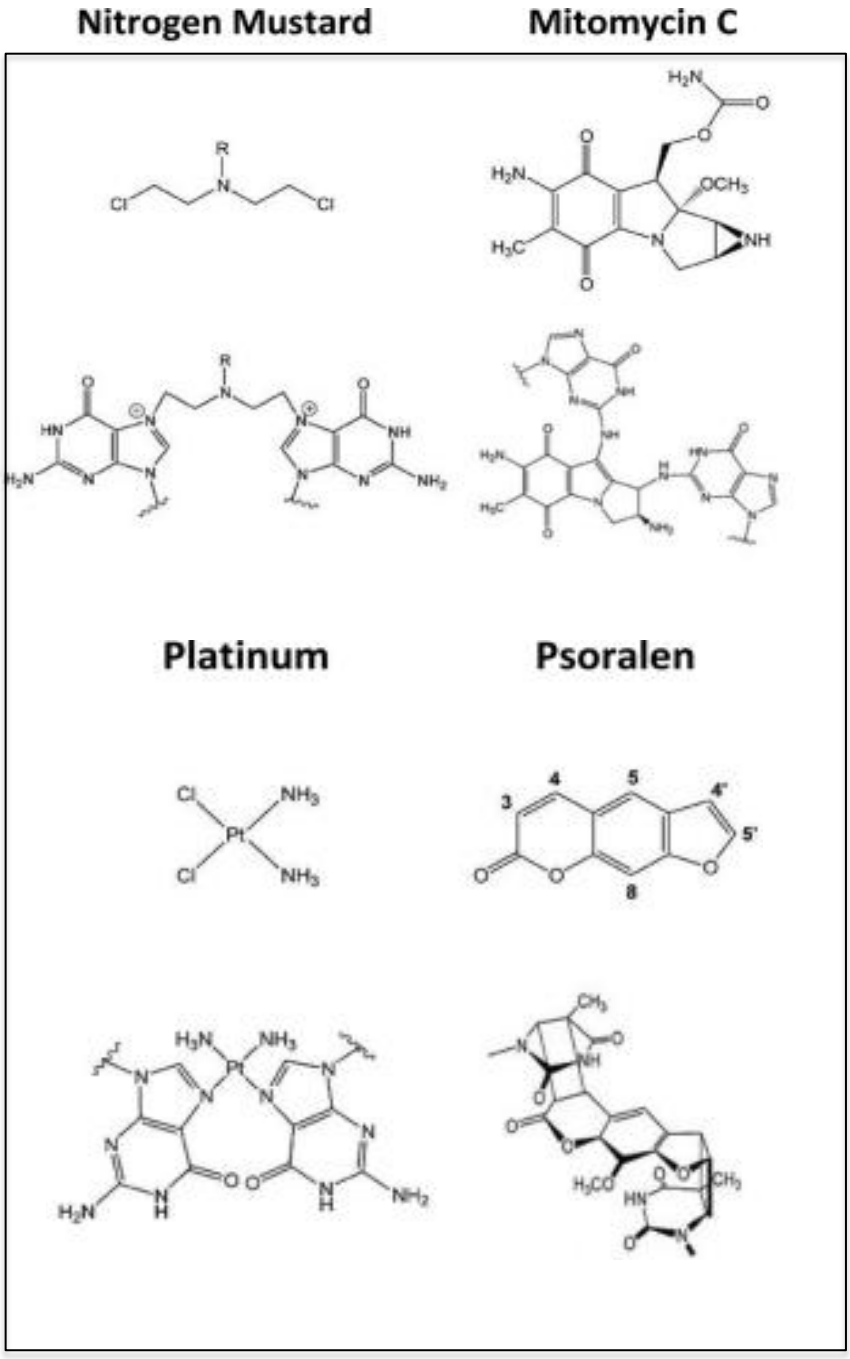


Figure 7- Common DNA damaging agents and their lesion formations

ICL-inducing agents are widely spread in anti-cancer therapy due to their extreme cytotoxicity. Common crosslinking reagents include Nitrogen Mustard (HN2), Mytomycin C (MMC) as well as Cisplatin and its common derivatives. In this report, we use Cisplatin. Cisplatin reacts with the guanine bases of DNA, forming ICLs at GpC sites in the DNA. It is reported to create the largest helical distortion of the DNA (Wilson & Seidman 2010).

Replication-dependent ICL repair pathway: The Fanconi Anemia Pathway

ICLs are complex lesions that require the coordination of multiple repair pathways for removal of the DNA damage and the successful maintenance of genomic integrity (Kottemann & Smogorzewska 2013). The current model of ICL repair is thought to involve the concerted action of several repair systems, including the Fanconi Anemia (FA) pathway, homologous recombination (HR), Nucleotide excision repair (NER) and translesion synthesis (TLS) (Deans & West 2011).

FA is a rare, clinically heterogeneous disease characterized by congenital abnormalities, growth retardation, bone marrow failure and genome instability that predisposes its patients to cancer (Bhagwat et al. 2009). 17 proteins, known as FA factors, have been identified to be mutated in FA patients. Those proteins are assigned as FA complementation groups (FANCA to FANCS) and together with FA associated proteins; they play key roles in a common pathway implicated in the repair of ICLs (Wang & Smogorzewska 2015).

Replication dependent ICL repair is a multistep process that requires the interaction of the FA factors and its associated proteins. As described in Wang & Smogorzewska, fork stalling at an ICL site during S phase serves as a signal for the recruitment of the heterodimer FANCM-FAAP24, FAAP16/MHF1 and FAAP10/MHF2. This is followed by the assembly of the FA core complex (composed of FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, FAAP100 and FAAP20). Some of the components such as FANCM also have additional functions such as phosphorylation of the DNA damage response kinase ATR, which in response activates essential repair

downstream factors. Activation of ICL repair is initiated by the ubiquitination of the FANCD2-FANCI complex by the E3 ubiquitin ligase FANCL, one of the proteins of the FA core complex (Ghosal & Chen 2013; Smogorzewska et al. 2007). Ubiquitinated FANCD2-FANCI complex serves as a recruitment signal for a variety of structure-specific nucleases implicated in ICL repair (Ishiai et al. 2008). In fact, multiple repair factors and in particular Fan1 (Fanconi-associated nuclease 1) as well as the scaffold protein SLX4 are suggested to be recruited to the site of damage through their interaction with the ubiquitin-binding zinc finger (UBZ) domains of FANCD2-FANCI (MacKay et al. 2010). SLX4 acts as a regulator for the recruitment of repair factors XPF-ERCC1, SLX1 and Mus81-EME1 to the damage site (Sengerová et al. 2011). Once the lesion has been recognised and repair initiated, incisions on either side of the ICL occur in order to convert the lesion into a double strand break. This step, termed ICL unhooking, requires the spatial and temporal coordination of various candidate structure-specific nucleases: XPF-ERCC1, SLX1 and Mus81-EME1 recruited by SLX4, Fan1 as well as SNM1A and SNM1B (Sengerová et al. 2011; Wang et al. 2011). The exact molecular mechanisms by which ICL unhooking occurs are not clear. A more detailed outline will be discussed in the next section.

After the ICL has been unhooked, the bypass of the cross-linked base pair occurs via translesion synthesis (TLS). TLS polymerases Pol z and Rev1 are implicated in ICL repair. In fact Rev1 has the ability to introduce a nucleotide opposite the lesion enabling nascent DNA strand extension past the ICL (Ho & Schärer 2010). ICL repair is then completed via homologous recombination (HR). HR is suggested to promote error-free DSB repair in S phase, in a Rad51-dependent manner (Deans & West 2011; Hinz 2010). Rad51 is able to facilitate strand invasion into the template DNA of the sister chromatid thereby repairing and re-establishing the replication fork (Bhagwat et al. 2009).

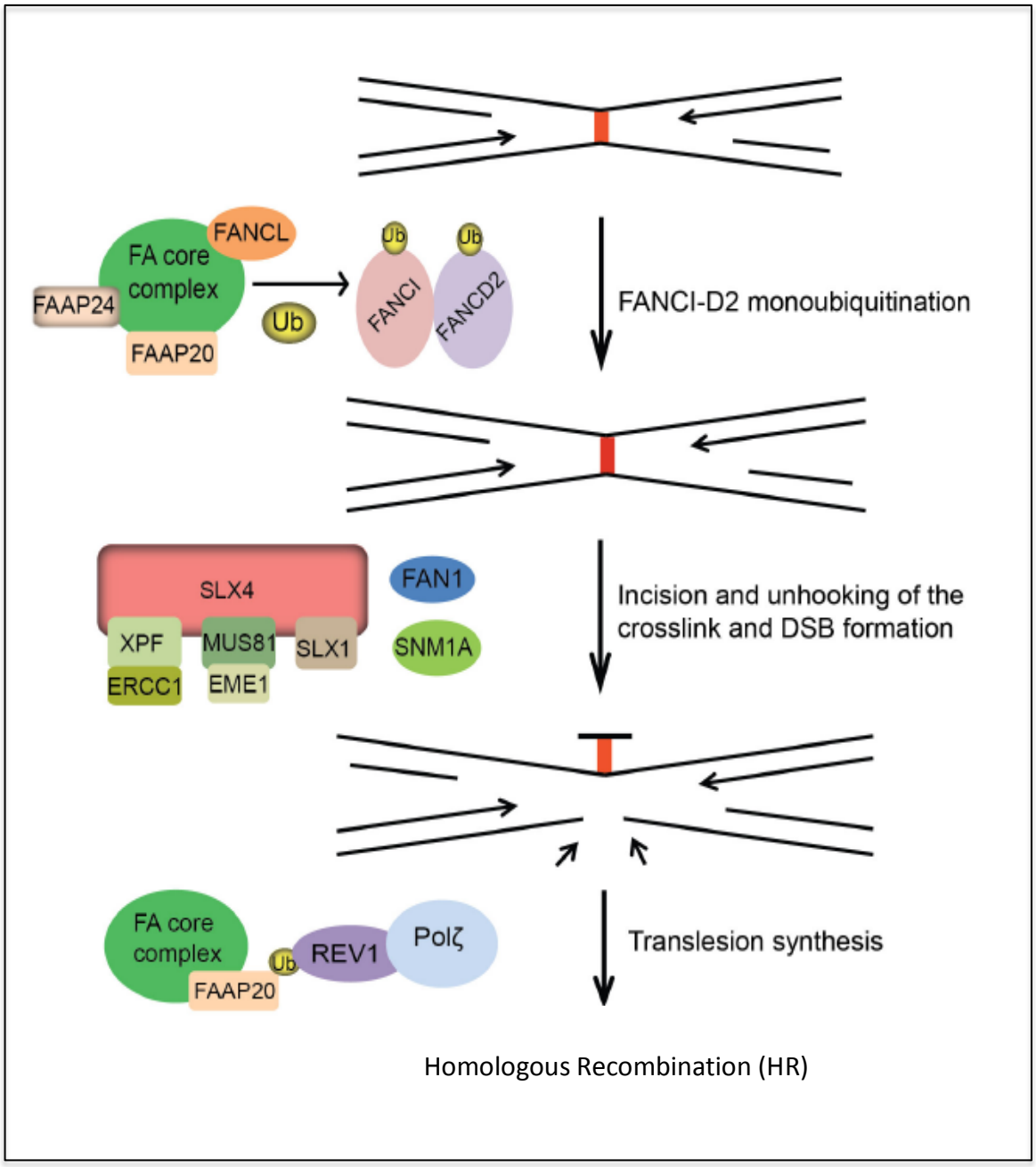


Figure 8- Simplified current model of the Fanconi Anemia Pathway in DNA ICL repair in mammalian cells

In the dual fork convergence model, collision of replication forks with the DNA ICL (represented here in red) leads to the assembly of the FA core complex that subsequently mono-ubiquitinates the FANCI-FANCD2 complex. This serves as an activation signal for several structure specific nucleases, including XPF-ERCC1, MUS 81-EME1, SLX1, FAN1 and SNM1A, that incise and unhook the DNA crosslink. The translesion synthesis machinery then further processes the damage site and repair is completed by homologous recombination (Adapted from Ghosal & Chen 2013).

ICL Unhooking: A more precise molecular description

The key step in ICL repair is the unhooking of the ICL performed by the concerted action of several structure-specific nucleases (Cheryl Clauson et al. 2013). ICL unhooking refers to the dual incisions on the 5' and 3' side of the ICL, thereby creating an intermediate structure containing a double-strand break (Wang & Smogorzewska 2015). The exact mechanisms as well as structural requirements of the DNA to trigger this step remain elusive. So far, numerous repair nucleases have been implicated in this process, namely XPF-ERCC1, Mus81-EME1 and SLX1 that are recruited by the scaffold SLX4, the nuclease Fan1 as well as repair factors SNM1A. All of these nucleases play a key role in ICL repair, with deficiency in any of these factors causing hypersensitivity to ICL-inducing agents (Sengerová et al. 2011; Hashimoto et al. 2016).

A proposed model for ICL repair suggests that the 3' flap endonuclease Mus81-EME1 is implicated in the 3' incision of the ICL. It has been suggested that Mus81-EME1 creates the first cut around the ICL thereby creating a gapped intermediate that can be further processed by other nucleases. Mus81-EME1 has also been demonstrated to have roles in HR and the resolution of HR intermediates indicating it could also play a further downstream role in ICL repair (Bhagwat et al. 2009; Wang et al. 2011).

Fanconi-associated nuclease 1 (FAN1) has been suggested to have a role in ICL repair as FAN1 deficiency causes hypersensitivity to ICL inducing agents such as Mytomycin C or Cisplatin (Kratz, Schöpf, Kaden, Sendoel, Eberhard, et al. 2010). Studies reporting the structural organisation of FAN1 also implied its involvement in ICL repair. FAN1 has been reported to possess a RAD18-like zinc finger domain, a UBZ domain at its N-

terminus and a PD-D/E(X)K type endonuclease motif within a Virus-type replication repair nuclease domain at its C terminus (MacKay et al. 2010). FAN1 is believed to interact with mono-ubiquitinated FANCD2 through its UBZ domain recruiting it to the site of ICL damage (Kratz, Schöpf, Kaden, Sendoel, Eberhard, et al. 2010). Although both FAN1 -/- mice and FA patients show microcephaly and neurodevelopmental retardations, no anaemia or bone marrow failure were observed in FAN1 deficient patients (Thongthip et al. 2016). Previously published studies demonstrate that FAN1 is not epistatic with other FA factors and deficiency in FAN1 does not lead to FA progression suggesting an alternative pathway FAN1 might be involved in (Sengerová et al. 2011; MacKay et al. 2010; Kratz, Schöpf, Kaden, Sendoel, Eberhard, et al. 2010). FAN1 has been reported to have both an endo- and exonuclease activity suggesting multiple roles for the protein (Pizzolato et al. 2015; MacKay et al. 2010). Studies suggest both a 5' flap endo- and a 5' to 3' exonuclease activity. Its exonuclease activity has been reported to be a series of sequential cuts leading to 3nt products, with a maximum of 4 consecutive cuts (Wang et al. 2011; Pizzolato et al. 2015). All together, current studies underline FAN1's structural and functional flexibility, making its precise role unclear.

XPF-ERCC1 has been implicated in several DNA repair responses including NER, and ICL repair (Bhagwat et al. 2009). XPF-ERCC1 is one of the key nucleases involved in the repair of bulky DNA lesions such as UV photoproducts, incising 5' to the lesion on the damaged strand (Wang & Gautier 2010). XPF has recently been characterized as a Fanconi Anemia complementation group (FANCQ) underlining its importance in ICL repair. XPF-ERCC1 is demonstrated to be the key protein involved in ICL repair (Sengerová et al. 2011). Immunodepletion experiments in *Xenopus laevis* cell-free egg

extracts reveal that only XPF depletion abolishes ICL repair whereas depletion of other nucleases only has minor effects. XPF-ERCC1 depleted cells also display accumulation of DSBs as a result of impaired DNA repair. Altogether, this suggests XPF-ERCC1 is a pre-required factor to create an incision around the ICL (Zhang et al. 2015; Bhagwat et al. 2009; Douwel et al. 2014). The repair protein XPF-ERCC1 exists in a heterodimeric form, with dimerization occurring through the C terminal tandem helix-hairpin-helix domain $(HhH)_2$ of both proteins . The catalytic activity of the heterodimer is harboured in the XPF subunit that contains a highly conserved central PD-D/E(X)K type endonuclease domain (Enzlin & Schärer 2002). XPF-ERCC1 has been reported to have a 5' endonuclease activity, incising on the 5' of the ICL (Bhagwat et al. 2009). Work in our lab has also reported that the heterotrimeric single-stranded DNA binding protein RPA is required to stimulate and restore XPF-ERCC1 activity on a DNA substrate modelling a nascent leading strand on a native fork as well as establishing a duplex DNA region as a pre-requisite for XPF-ERCC1 activity. Defects in XPF-ERCC1 can lead to rare autosomal recessive DNA disorders including Xeroderma pigmentosum (XP) and Fanconi Anemia (FA) that both predispose its patients to cancer. While XP is mainly resulting in defects in the NER capacity, FA patients exhibit defects in ICL repair (Lehmann 2003; Iyama & Wilson 2013).

Another key factor that has been implicated in ICL repair is the exonuclease SNM1A. Previous work in our lab has shown that SNM1A -/- human cells display hypersensitivity to ICL-inducing agents (Sengerová et al. 2012). Interestingly, SNM1A-/- cells are particularly sensitive to ICL-inducing agents that target the minor groove of the DNA. Contrary to cells depleted on other key factors of ICL repair such as XPF or ERCC1, SNM1A-/- cells display minor sensitivity to agents reacting with the major

groove of the DNA (Wang et al. 2011). Studies suggest SNM1A is recruited to the DNA lesion by monoubiquitinated PCNA in a RAD18-dependent manner. The crystal structure of SNM1A reveals a wide DNA-binding groove that further confirms its role in DNA repair (Allerston et al. 2015). Experiments have also suggested SNM1A and XPF-ERCC1 cooperate in a common ICL repair pathway. It is therefore hypothesised that SNM1A is able to digest the oligos tethered to the ICL in a 5' to 3' direction on a synthetic DNA structure after incisions by nucleases such as XPF-ERCC1 to enable further downstream processes to finish ICL repair (Wang et al. 2011; Sengerová et al. 2011).

Models of replication-dependent ICL repair

Current studies in ICL repair propose two potential repair models that suggest different DNA strand organizations for repair initiation. One model requires dual fork convergence onto an ICL site (as shown in Figure 8) while the other assumes a single fork colliding with an ICL lesion is enough to start ICL repair.

The dual fork convergence model is supported by recent studies on *Xenopus laevis* cell-free egg extracts that enables experiments to be carried out on a single site-specific ICL located on a 6kb plasmid. Using this system, Zhang et al. (2015) established some key requirements for ICL repair to occur. They propose the dual replication fork ICL repair model by which the convergence of two replication forks onto an ICL is essential for ICL repair activation. This is supported by their findings that an X-shaped structure is an absolute requirement for the initiation of ICL repair as well as the presence of a nascent leading strand up to 1nt from the ICL. The authors suggest that when a replication fork encounters an ICL, stalling of both nascent leading strands occurs at -20

nucleotides (nt) from the lesion. Only one leading strand then progresses to 1nt from the ICL stimulating unhooking of the ICL through incisions on either side of the lesion on the opposite template strand. Repair is then completed by TLS and subsequent HR before fork restart (Douwel et al. 2014; Zhang et al. 2015). In the single fork collision model, collision of a single replication fork with an ICL is able to trigger ICL repair. This model is supported by the idea that a single fork is more likely to strike a randomly formed ICL than a convergent fork. No data suggest whether the leading or lagging strand template is chosen for incisions to occur (C. Clauson et al. 2013; Patel & Joenje 2007).

Work in our lab performed by Dr. U. Abdullah proposed the following ICL repair model: In the event of a single replication fork encountering the ICL, progression of a nascent leading strand up to 1-nt to the ICL induces an XPF-ERCC1-RPA incision that occurs six nucleotides 5' to the junction in the double-stranded DNA region. The 5' to 3' exonuclease SNM1A is able to load onto the nick and digest past the ICL thereby unhooking the ICL. The gapped intermediate created is then processed by the TLS machinery and the broken DNA repaired via HR. In the event of dual replication fork convergence on an ICL, nascent leading strand stalling occurs and only one strand further progressed as previously described to initiate ICL unhooking. It has also been suggested that both the dual fork and single fork model could co-exist in a cell cycle dependent manner. In fact, the single replication fork model could occur in early S phase while it would be more likely for two converging forks to trigger ICL repair in late S phase (Wang et al. 2011; Dr. U. Abdullah's thesis).

Nonetheless, the exact mechanisms and requirements for repair initiation still remain elusive and more experiments will need to be conducted to obtain a better understanding of ICL repair.

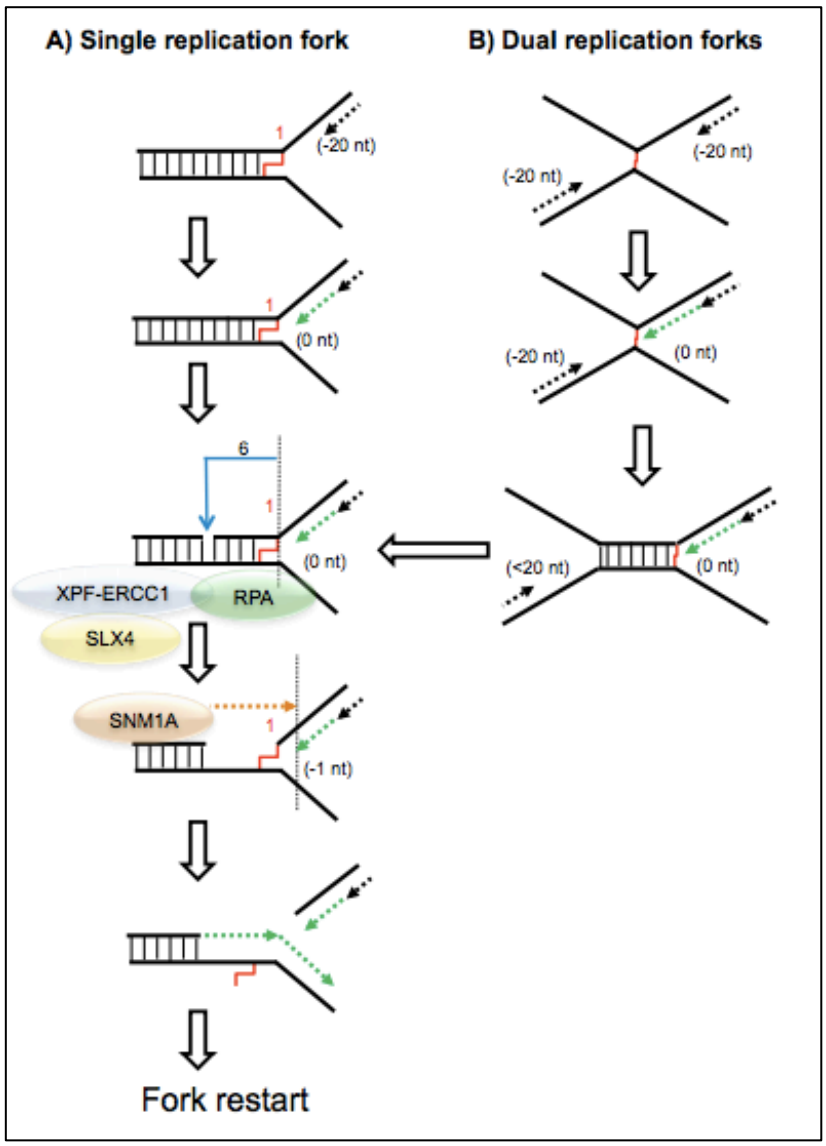


Figure 9- Model of DNA Interstrand Crosslink repair
As suggested by Dr U. Abdullah from our lab

(A) When a single replication fork encounters an ICL, the nascent leading strand stalls at 20 nt from the ICL. It then progresses to 1nt triggering an XPF-ERCC1-RPA incision in the duplex region. SNM1A can load onto the DNA and digest past the ICL thereby unhooking the ICL. The gapped intermediate is then processed by TLS proteins and the broken DNA strand is repaired by HR.

(B) In the event of dual replication fork convergence into an ICL, both nascent leading strands initially stall. Only one nascent leading strand gradually progresses to 1nt from the ICL initiating XPF-ERCC1-RPA action. After incision, SNM1A is able to load onto the DNA. Repair is terminated by the TLS machinery and subsequent HR of the DNA strand.

AIMS AND OBJECTIVES

In a cell-free *Xenopus laevis* egg model system, depletion of XPF has been proven to abolish ICL repair affirming XPF-ERCC1 as the key effector of ICL repair. Depletion of other nucleases involved in ICL repair, namely MUS81 or FAN1 did not have such a major effect (Zhang et al. 2015; Douwel et al. 2014). Previous work in our lab performed by Dr. U. Abdullah has also shown that XPF-ERCC1 plays a pivotal role in the initiation of ICL repair. Biochemical reconstitution assays of various synthetic DNA structures mimicking fork substrates demonstrated that XPF-ERCC1 is able to incise 5' to the ICL but not 3' to the ICL. All together, previous findings demonstrate that XPF-ERCC1 alone is not able to unhook an ICL alone suggesting ICL unhooking is a collaborative action between several nucleases.

Previous genetic analysis in our lab put the scaffold protein SLX4, the 5' to 3' exonuclease SNM1A and the endonuclease XPF-ERCC1 in the same pathway. Furthermore, the characterization of SNM1A and in particular the evaluation of its collaborative activity with XPF-ERCC1 has been studied (Sengerová et al. 2011; Wang et al. 2011). Wang, et al. (2011) demonstrated that SNM1A is able to load on and digest an ICL-containing double-stranded DNA substrate after incision by XPF-ERCC1 on the 5' side of the ICL.

This thesis aims to analyse the role of the less studied nuclease FAN1. Recently, FAN1 has become an interesting candidate to be involved in ICL repair as FAN1 -/- cells display hypersensitivity to ICL-inducing agents (MacKay et al. 2010). Through both biochemical and cellular characterisation of FAN1, we aimed to establish a genetic relationship between the key effectors of ICL repair, namely XPF-ERCC1, SNM1A and FAN1. Cellular studies concentrate on investigating a possible functional redundancy between FAN1 and the exonuclease SNM1A. We further wish to characterise the nuclease activity of FAN1 in order to be able to evaluate its role in ICL repair. All together, this will enable us to assess the possible role of FAN1 in an XPF-ERCC1 dependent pathway.

MATERIALS AND METHODS

I) Biochemical Methods

Protein purification of human XPF-ERCC1

Human XPF-ERCC1 protein was purified from insect cells at the Research Complex in Harwell with the help of Dr Denis Ptchelkine. The protocol was adapted from Dr Ummi Abdullah's thesis (Department of Oncology, Oxford) based on already published protocols (Enzlin & Schärer 2002).

Firstly, competent DH10Bac E.coli cells were transformed with pFastBac1 vectors containing either the cDNA encoding for full length XPF or full length ERCC1-His respectively. Using the BAC-to-BAC system (from Invitrogen), bacmid DNA was isolated and transfected into Sf-21 insect cells for baculovirus amplification. Next, 500 ml of Hi-5 insect cell culture at $2.5x10^7$ cells/ml were co infected with P2 XPF Wild-type (WT) and P2 ERCC1-His WT viruses and cells harvested after 72 hours. After centrifuging cells at 400 xg for 10 minutes at room temperature, cells were re-suspended in 200 mL 100 mM NaCl and centrifuged again at 400 xg for 10 minutes at RT. Cells were lysed in 100 mL lysis buffer (50 mM Tris-HCl pH 8.0, 800 mM NaCl, 20 mM Imidazole, 5 mM TCEP, 10% glycerol, 0.1% NP-40 and one EDTA-free protease inhibitor cocktail table (Roche)). After sonication (Soniprep 150 Plus, MSE), the cell lysate was centrifuged for 1 hour at 20 000 rpm at 4^oC. The supernatant was then collected and incubated on a rotator overnight with 2.5 mL Ni-NTA agarose beads (Qiagen) at 4^oC. Protein levels were checked throughout the process by SDS-PAGE.

The beads were centrifuged at 400xg for 5 minutes at 4^{o}C and re-suspended in 20 mL Ni-buffer (50 mM Tris-HCl pH 8.0, 10% glycerol, 5 mM TCEP) containing 800 mM NaCl and 4 mM imidazole. The sample was then loaded onto a column and the column washed twice via gravity flow with 50 mL Ni-buffer with increasing Imidazole concentrations. (500 mM NaCl with 5 mM and 20 mM imidazole respectively). Purified XPF-ERCC1 protein was eluted in 1 mL fractions with two times 12 mL Ni-Buffer with increasing imidazole concentrations (300 mM NaCl with 50 mM and 250 mM imidazole respectively). Protein levels of chosen fractions were determined by SDS-PAGE electrophoresis. Fractions with the highest levels of XPF-ERCC1 protein were pulled together and dialysed in 2 L dialysis buffer (20 mM Tris-HCl pH 8.0, 10% glycerol, 500 mM NaCl, 5 mM DTT) at 4^{o}C overnight using a 3.5K molecular weight cut-off dialysis cassette (Thermo Scientific). Gel filtration was then performed as a final purification step using a HiLoad 16/60 Superdex 200 column. (Pharmacia) equilibrated with gel filtration buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% glycerol and 5 mM TCEP). Fractions containing the XPF–ERCC1 heterodimer, eluting from 13 to 15 mL, were pooled and concentrated. The protein concentration was measured directly at 280 nm and the purified protein was flash-frozen in aliquots and stored at -80^{o}C. Typically, yields between 1.56-1.95 mg of the complex were obtained.

In this thesis, nuclease assays using human XPF-ERCC1 were performed using already purified protein from one of the members of our lab, Dr Ummi Abdullah. Full length human Fan1 was a kind gift from Dr John Rouse (University of Dundee). Recombinant human RPA complex purified from *E. coli* was purchased from Enzymax. (Catalog #61).

Generation of radiolabelled DNA substrates

In this thesis, both 3'-end and 5'-end radiolabelled DNA substrates were used. 3'-end radiolabelling was performed by incubating 50 pmoles of DNA oligonucleotide with 1 unit of terminal deoxytransferase (TdT) (NEB) in the presence of 3.3 pmoles [α-^{32}P] dATP for 1 hour at 37^{o}C. For 5'-end radiolabelling, 10 pmoles of DNA oligonucleotide were incubated with 1 unit T4 polynucleotide kinase (T4 PNK) (NEB) in the presence of 6.8 pmoles [γ-^{32}P]dATP for 1 hour at 37^{o}C. The reaction mixture was spun in a Bio-Gel P-6 spin column (Bio-Rad) at 1,000xg for 4 minutes and diluted with water to a concentration of 500 nM. In order to create the different structures, the radiolabelled oligonucleotide was annealed with unlabelled ('cold') complementary oligonucleotide(s) using annealing buffer (10 mM Tris pH 7.5, 50 mM NaCl, 1 mM EDTA). The mixture was boiled at 100^{o}C for 3-5 minutes and gradually cooled to below 30^{o}C before use. Finally, radiolabelled oligonucleotides are diluted to 100 nM and run on a 10% non-denaturing polyacrylamide (PAGE) gel to check for correct annealing.

Structures in Figure 10 were created by end labelling single synthetic DNA oligonucleotide sequences and annealing structures mimicking intermediate structures involved in DNA replication.

Oligo- sequences used to create DNA structures were:
TOP STRAND:
5'ATAAATATTTTTATTAATAATAGATCACCTTTCTTTCTCTTCTCCCCT3'

BOTTOM STRAND:
5'TTCCCCTCCTCTCCTTCCTTCCTGATCTATTATTAATAAAAAATATTTA3'

LEADING STRAND:
5'AAGGGGAGAAGAGAAAGAAAGG3'

LAGGING STRAND:
5'GGAAGGAAAGGAGAGGAGGGGAA3'

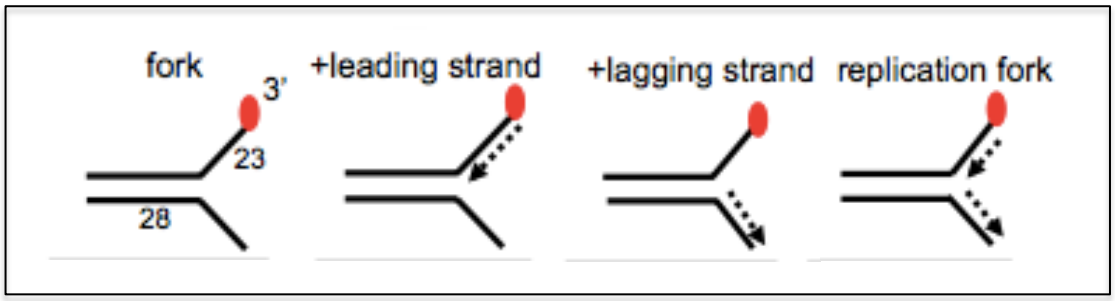


Figure 10- Synthetic 3' top strand labeled DNA structures used for nuclease assays

Structures were annealed using the above mentioned oligonucleotide sequences to create synthetic DNA structures mimicking common DNA replication structure intermediates. Structures are used to assess the activity and specificity of the nucleases involved in ICL.
Strands can also be bottom labelled and either 3' or 5' end- labelled. Red dot indicated the 3' radioactive label. Numbers (23 and 28 nucleotides) refer to the length of the DNA section.

Nuclease assay

For the experiments performed in this thesis, reactions were carried out with 10 nM of DNA substrate and 40 nM of human XPF-ERCC1 in nuclease buffer (25 mM HEPES pH 8.0, 40 mM NaCl, 10% glycerol, 0.5 mM ß-mercaptoethanol, 0.1 mg/ml BSA) containing 1 mM $MnCl_2$. For reactions with FAN1, 10 nM of DNA substrate were reacted with 1 µM of FAN1 protein in the same nuclease buffer. (Concentrations of protein used have previously been described in the literature and tested in our lab.) For experiments analysing the cooperative activity of XPF-ERCC1 and FAN1, substrates and protein were incubated for 1 hour at 30°C for XPF-ERCC1 reactions and 30 min at 37°C for FAN1 reactions. Assays were quenched by adding 3 µl STOP solution (95 % formamide/5 % EDTA) and heated at 95°C for 5 minutes.

To analyse the activity of FAN1 post-incision by XPF-ERCC1, DNA substrates were reacted with 40 nM of XPF-ERCC1 first, followed by the addition of 1 μM FAN1. The reactions were then further incubated at 37^{o}C for 30 minutes. All nuclease assays were then quenched by adding 3 μL STOP solution (95 % formamide/5 % EDTA) and heated at 95^{o}C for 5 minutes.

Analysis was performed by loading the samples on a 20% denaturing (8 M urea) PAGE gel (19:1 Acrylamide/Bis) gel containing 1x TBE. Gels were run for 2 hours at 525 V. Gels were fixed in fixing solution (40% methanol, 20% Acetic Acid and 5% glycerol) for 1 hour and dried in a Gel Dryer at 50^{o}C for 4 hours. Images of the gels were visualised by phosphorimager.

II) Cell Biology Methods

Tissue Culture

Human 293-FT and U2OS cells used in this report were a kind gift from Lonnie Swift, a member of our group. Cells were grown in DMEM (Gibco, Thermo Scientific) media supplemented with 10% (v/v) fetal calf serum and L-glutamine and no antibiotics and incubated at 37°C at 5% CO_2. When appropriate, cells were harvested using Trypsin solution.(Sigma Aldrich). Cells were regularly tested for mycoplasma and general tissue culture good practice respected.

siRNA knockdown experiments

For siRNA experiments, U2OS cells were transfected with FAN1 siRNA using HiPerfect reagent (Qiagen) according to the manufacturer's protocol.

Following siRNA duplexes at increasing concentrations (20 nM, 50 nM, 100 nM) were used:

FAN1-1 : GUAAGGCUCUUUCAACGUA

FAN1-2 : GCAGGAAGGCAGAGUGGCU

siRNA were pooled together for the experiments performed in this thesis.

A reverse transfection protocol was performed: cells were harvested and immediately treated with the transfection complexes. A second round of transfections was performed 24 hours after. Cells were then harvested 48 hours after first transfection and used for subsequent experiments. Control sample was transfected with AllStars siRNA negative control duplex (Qiagen).

Reverse transcription was performed using the High Capacity cDNA reverse transcription Kit (Applied Biosystems) according to the manufacturer's manual. A total of 2 μg of RNA was converted to single stranded cDNA in a single 20 μL reaction using the following reaction mix. cDNA was quantified and used for RT-qPCR to examine inhibition levels of siRNA.

RT-qPCR was used to confirm siRNA gene silencing. SYBR Green RT-PCR Reagent kit (Applied Biosystems) was used to measure expression levels of SNM1A and FAN1. GAPDH housekeeping gene levels were measured as a control. 10 ng of cDNA

template were used in each reaction. Primers were designed according to the general guidelines outlined in the manufacturer's protocol.

Clonogenic Cell Survival

Transfected cells were used to perform clonogenic survival experiments. Cells were seeded at increasing cell numbers in 10 cm dishes and allowed to adhere for a minimum of 4 hours before drug treatment. Cells were treated with increasing concentrations (0 μM to 10 μM) of Cisplatin for 4 hours before fresh media was added. Colonies were allowed to grow for 12 days before plates were stained with Commassie Brilliant Blue solution.

Clustered regularly interspaced short palindromic repeats (CRISPR) /Cas9 knock out cell line generation

A XPF -/- human cell line was generated in U2OS cells using the CRISPR/Cas9 system. CRISPR-Cas9 plasmid design and protocol were adapted as published from the Zhang Laboratory (Ran et al. 2013). Plasmids were a gift from Dr P. Hublitz (University of Oxford). Following plasmids were used: pX330-GFP and pX458-Ruby both containing a puromycin resistance cassette and a GFP or Ruby reporter gene respectively.

An RNA guide sequence (sgRNA) was designed to target exon 1 of human XPF protein. Two RNA guide sequences were used to create a XPF knock out cell line: one

at the beginning and one at the end of exon 1. Guide sequences were chosen using http://crispr.mit.edu/ from Zhang's laboratory.

Following guide sequences were used:

BEGINNING OF EXON 1:
5' CACC GCCGGCTCGACGGATTGCCA 3'
END OF EXON 1:
5' CACC GCTCAACACGCAGCCGGCCG 3'

Guide sequences were cloned into the backbone vectors according to Ran et al. 2013 protocol. Plasmids were digested with Bbsl for 30 min at 37°C using FastDigest kit (Fermentas). sgRNA oligos were phosphorylated and annealed using T4 PNK (NEB) according to the manufacturer's protocol. A ligation reaction was performed using 50 ng of digested plasmid for 1:200 diluted oligo duplexes with Quick Ligase (NEB). Finally, competent DH5a cells were transformed with the ligation product and after heat shock and incubation, cells were plated onto LB agar with 100 μg/ml of ampicillin and incubated at 37^{o}C overnight.

Colonies were picked from agar plates and overnight cultures set up in LB broth with 1:1000 ampicillin at 37^{o}C. Plasmid DNA extraction was performed using QIAprep Miniprep Kit (Qiagen) according to manufacturer's protocol. Plasmids were sent for sequencing to confirm correct product formation before being used for further experiments.

U2OS and 293-FT human cells were transfected with both plasmids using Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's protocol. After 24 hours, cells were single-cell sorted for both GFP and Ruby positive cells by FACS. Single colonies were grown in 96 well plates before expanding surviving

colonies into 6 well plates. Cells were harvested and tested for XPF expression by Western Blot. Positive clones were sent for sequencing to confirm deletion in the XPF gene.

Western Blotting

To observe levels of XPF protein, whole cell extracts were prepared. Cells were centrifuged for 5 min at 2.4 rpm, washed with PBS solution and lysed using 2x Laemmli buffer (buffer (2% SDS, 10% glycerol, 62.5 mM Tris-HCl at pH 6.8) at 95^{o}C for 5 min. Samples were then mixed with loading buffer and reducing agent (Thermo Scientific NuPAGE Sample buffer Kit) and loaded onto a 4-12% SDS-Page gel (Thermo Scientific). Gel was run at 130 V for 1 hour 30 min in Tris-Bis buffer and transferred overnight at 4°C on PVDF membrane. The membrane was then blotted with appropriate antibodies (monoclonal XPF antibody from ThermoFisher) for visualisation of protein levels. Actin was used as a loading control.

RESULTS

Experiments using cell-free replication coupled repair system in *Xenopus laevis* egg cell-free extracts have established the endonuclease XPF-ERCC1 as the key factors for DNA interstrand crosslink initiation (Zhang et al. 2015). In fact, immunodepletion experiments in *Xenopus leavis* egg extracts reveal that only XPF depletion abolishes ICL repair whereas depletion of other nucleases, namely FAN1 and MUS81 have only minor effects (Zhang et al. 2015; Räschle et al. 2008; Douwel et al. 2014).

Work in our lab confirmed that XPF-ERCC1, a structure-selective heterodimeric endonuclease, is able to incise 5' to the ICL cutting at several sites in the duplex region, between 2 and 6 nt from the junction. However, no incisions were observed 3' to the lesion, which would be required for complete ICL unhooking. This suggests that although XPF-ERCC1 is essential for ICL repair initiation, it is not sufficient on its own and requires the action of other nucleases to successfully unhook the ICL (Bhagwat et al. 2009; Douwel et al. 2014; Wang & Smogorzewska 2015). SNM1A has already been demonstrated to participate in ICL repair in the same pathway as XPF-ERCC1 (Sengerová et al. 2012). In fact, the 5'-3' ICL repair exonuclease SNM1A can load from XPF-ERCC1 induced incisions and digest past the ICL to complete ICL unhooking (Sengerová et al. 2012; Wang et al. 2011).

Nonetheless, some observations have been made that suggest that other factors might collaborate with XPF-ERCC1, probably in a redundant way. It has been demonstrated that SNM1A -/- cells display a milder phenotype when exposed to ICL-inducing agents as opposed to cells disrupted in XPF (Wang et al. 2011). Furthermore, some Fanconi Anemia patients display symptoms of the disease but do not have any mutations in the known FA genes (Deans & West 2011).

All together, this suggests another factor is involved in the unhooking step of ICL. FAN1 has been identified as a good candidate for this role as FAN1 deficiency causes hypersensitivity to ICL inducing agents such as Mytomycin C or Cisplatin (Thongthip et al. 2016; Kratz, Schöpf, Kaden, Sendoel, Legerski, et al. 2010). FAN1 has also been demonstrated to interact with mono-ubiquitinated FANCD2, recruiting it to the site of ICL damage (Kratz, Schöpf, Kaden, Sendoel, Legerski, et al. 2010). Finally, FAN1 has been reported to both have an endo- and exonuclease activity further promoting the hypothesis of FAN1 being involved in ICL repair in an XPF-ERCC1 dependent pathway (MacKay et al. 2010; Pizzolato et al. 2015).

In this thesis, we wish to further investigate its biochemical role as well as establish its genetic relationship with SNM1A and XPF-ERCC1.

Biochemical Characterisation of ICL repair factors

In order to characterise the cooperative role of nucleases XPF-ERCC1 and FAN1 in ICL unhooking, we used an in vitro approach using nuclease assays on synthetic DNA oligonucleotide structures that mimic replication-associated intermediates in ICL repair.

Human XPF-ERCC1 Protein Purification

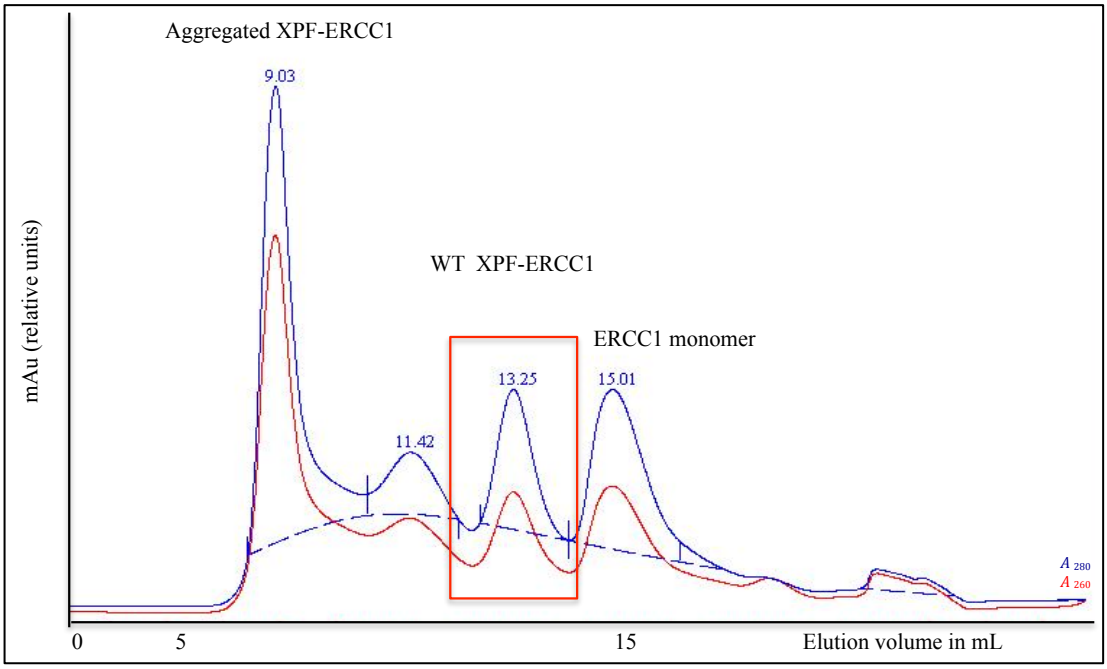
Human XPF-ERCC1 protein was purified from insect cells at the Research Complex in Harwell with Dr Denis Ptchelkine's help. Recombinant full length WT XPF and His-tagged ERCC1 were co-purified from insect cells. Purity of the heterodimer was

assured by performing nickel agarose affinity chromatography followed by gel filtration.

The gel filtration elution profile (Figure 11-A) below was obtained. As indicated on the trace, four main peaks were obtained. The peak around 9 mL corresponds to aggregated protein and cannot be used for any biochemical assays. Elution from peaks at 11 mL and 13 mL both contain purified XPF and ERCC1-His although the peak at 13 mL seems to have a more significant amount of protein as indicated by the stronger signal. The peak at 15 mL represents mostly unbound ERCC1-His monomers. Fractions corresponding to the labelled peaks were pooled together and tested by SDS-PAGE gel in order to prove the presence of purified XPF-ERCC1-His (Figure 11-A).

The activity of the purified protein was tested against a DNA structure mimicking a simple fork, an established XPF-ERCC1 substrate (Confirmed in Dr. U Abdullah's thesis) (Figure 11-B). A nuclease assay was performed by adding increasing concentrations of XPF-ERCC1 complex (0, 20, 40 and 80 nM) on a 10 nM DNA substrate for 1 hour at 30°C (Figure 12). Purified XPF-ERCC1-His already present in our lab (courtesy of Dr Ummi Abdullah) was used as a reference control. Nuclease assay products were analysed on a 20% denaturing acrylamide gel. Unfortunately, a relatively weak signal was obtained with the batch of newly purified protein. This may be due to the purification conditions and could be improved in future by adding an extra metal affinity chromatography purification step or designing a single virus containing both XPF and ERCC1 rather than co-infecting the insect cells. We therefore decided to use the already available purified human XPF-ERCC1-His from our lab for all experiments performed in this thesis (purified by Dr Ummi Abdullah in a previous project).

A



B

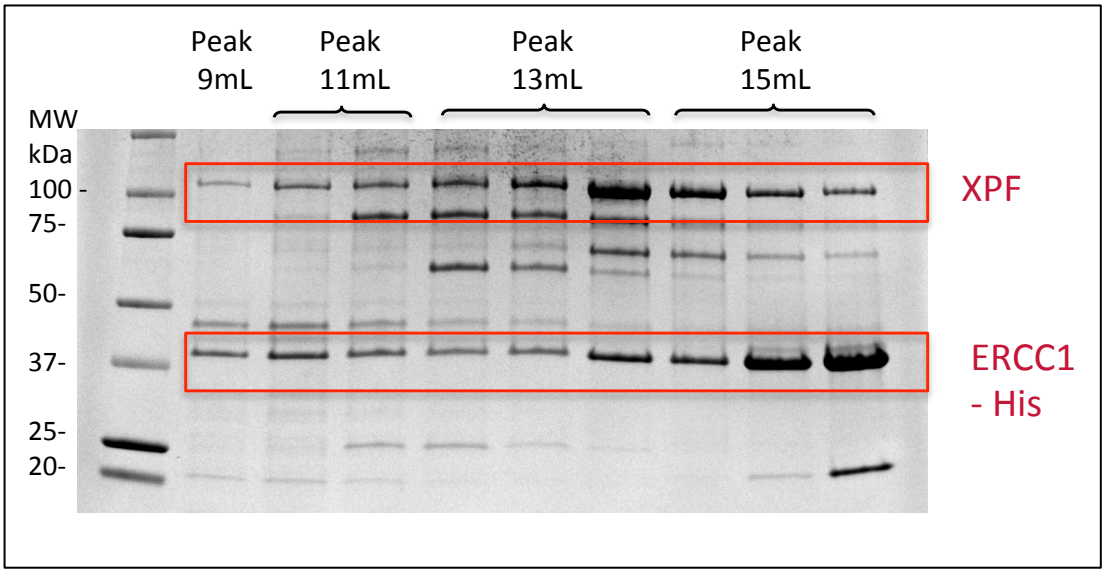
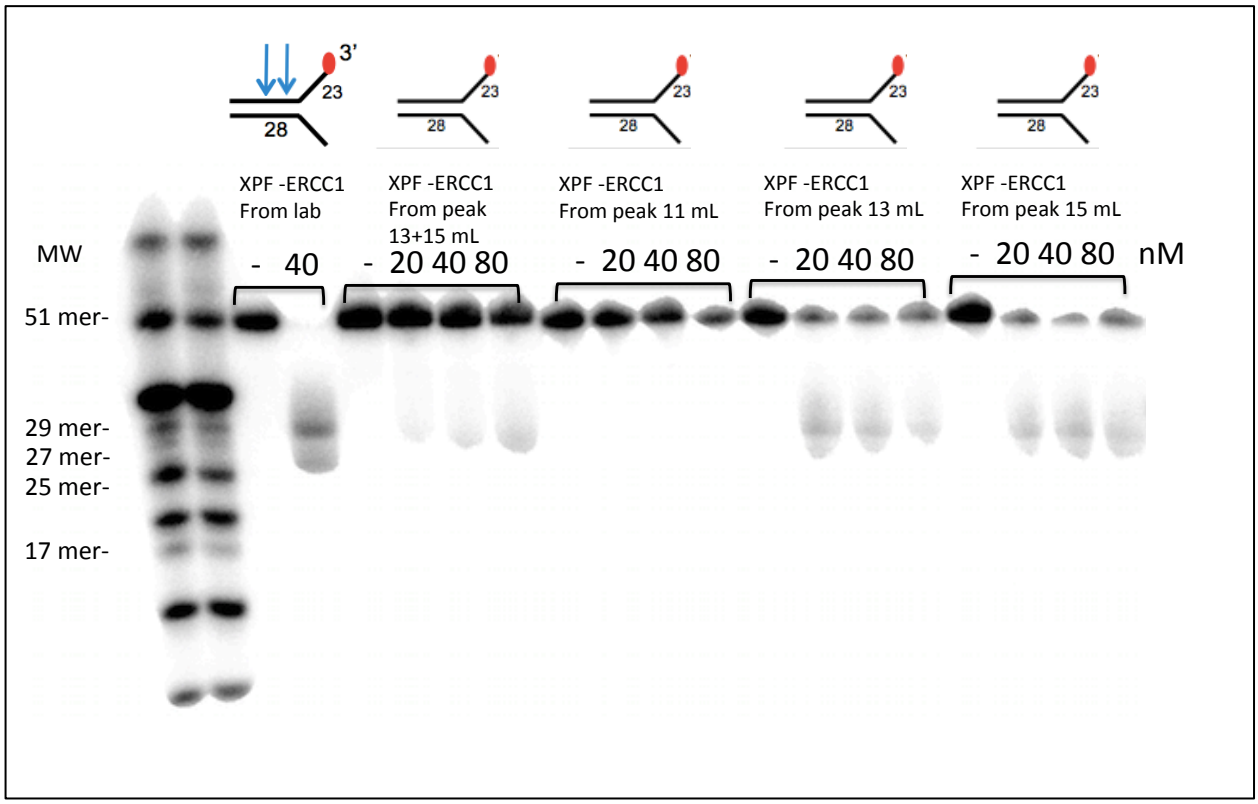


Figure 11- Purification of human XPF-ERCC1-His from insect cells

(A) Elution profile of gel filtration for XPF-ERCC1-His. Peaks at 11 mL and 13 mL contain purified XPF-ERCC1 protein.

(B) 4-12% SDS PAGE of samples taken from each peak as indicated. MW indicates the molecular weight ladder. Gel was stained with Instant Blue to visualize bands. Peak at 13 mL contains the most purified protein whereas peak at 9 mL only has aggregated protein and peak at 15 mL displays mainly unbound ERCC1.

A



B

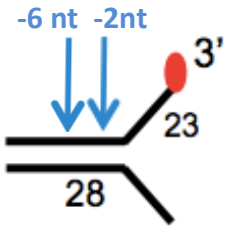


Figure 12- Activity test of purified XPF-ERCC1

(A) A nuclease assay was performed by adding increasing concentrations of the XPF-ERCC1 complex (0, 20, 40 and 80 nM) on a 10 nM 3'[^{32}P]-radiolabelled simple fork substrate for 1 hour at 30^{o}C and samples ran on a 20% denaturing acrylamide gel. DNA structures were labelled on the top strand at the 3' end. (red dot). MW shows the molecular weight. Activity of the newly purified protein was low compared to the protein already available in the lab. **(B)** Cartoon indicating the cleavage pattern of XPF-ERCC1 on a 3'[^{32}P]-radiolabelled simple fork substrate. The two incisions observed in the nuclease assay are indicated by the arrows confirming XPF-ERCC1's function as a 5' endonuclease. XPF-ERCC1 incises at -2 nt and -6 nt from the fork creating products of 29 mer and 25 mer. (Identified in previous work in our lab).

Creation of radioactively end-labelled synthetic DNA structures

In order to characterise the function of the different nucleases involved in ICL repair, we use *in vitro* nuclease assays. Nuclease assays enable us to analyse the effect of one or more structure-specific nucleases on an array of DNA substrates that represent the different intermediates that can be formed during DNA repair at a stalled fork. Two types of labelling exist: 3' labelling that adds a radioactive α- [^{32}P] phosphate to the 3' end of a DNA oligonucleotide strand and 5'labelling that adds a radioactive γ-[^{32}P] to the 5' end of the nucleotide sequence. Once a radioactively labelled oligonucleotide has been generated, it is annealed with an unlabelled complementary DNA oligonucleotide in order to create an array of different DNA substrate structures. Substrates can either be labelled on the top or bottom strand of the structure, on either end. In this report, we predominantly use four different types of DNA substrates, mimicking a simple replication fork (Y-structure), a fork with either a leading strand or a lagging strand and a fork with both a leading and lagging strand (named replication fork). All DNA substrates were created according to the protocol described in the Materials and Methods section. The purity of the structures was assessed on a 10% non-denaturing polyacrylamide electrophoresis gel (PAGE). Only one band per structure is visible confirming that the annealing reaction was successful (Figure 13).

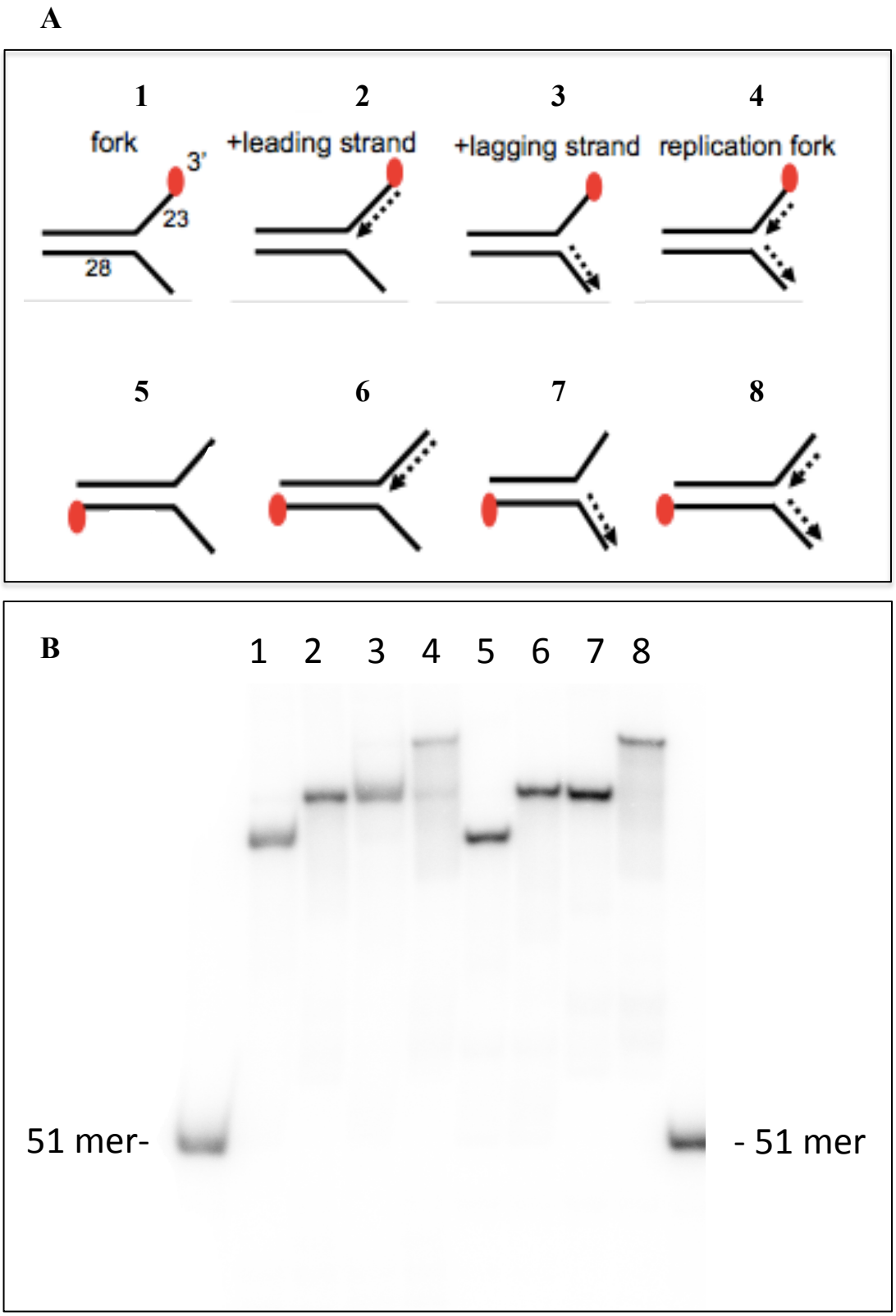


Figure 13- Synthetic radioactively labeled DNA substrates on a 10% Native Gel

(A) Schematics of the radiolabelled DNA substrates created. In this case, structures are 3' labeled on the top strand (1-4) and on the bottom strand (5-8). 5' labeled structures have also been created and tested on a 10% non-denaturing gel. 23 and 28 indicate nucleotide length of the DNA.

(B) 10 % non-denaturing PAGE gel showing the quality and purity of the annealed DNA substrates. Lanes correspond to the above schematic. 51-mer strands were used as molecular weight markers.

Functional Characterisation of the activity of Fanconi-associated nuclease 1

Fanconi-associated nuclease 1 (FAN1) has recently been discussed as a factor involved in ICL repair and in particular in the ICL unhooking step. FAN1 has been reported to both have an intrinsic endo- and exonuclease activity (Pizzolato et al. 2015; Kratz, Schöpf, Kaden, Sendoel, Eberhard, et al. 2010; MacKay et al. 2010). Our aim was to characterise the basic ionic and structural requirements for FAN1 digestion and evaluate its role in combination with the key effector of ICL repair, XPF-ERCC1.

In order to validate the activity of our wild-type FAN1 protein preparation, we firstly performed a comparative 20% PAGE gel with a nuclease dead mutant form of FAN1 (both courtesy of Prof. J.Rouse). Nuclease assay showed no products for the nuclease dead mutant whereas incisions were observed for our WT protein thereby confirming its activity. We also sought to determine the ideal protein:DNA ratio for best signal. The nuclease assays were performed on a 3' bottom strand labelled fork structure with a leading strand. Results confirm 10:1 protein:DNA ratio that is consistent with published data (MacKay et al. 2010) (Figure 14-A).

We then identified the cationic preferences of FAN1 for optimal activity. FAN1 contains a catalytic viral replication and repair (VRR) nuclease domain that has two metal binding sites predicted to interact with metal ions stabilised by aspartic and glutamic acid residues (Gwon et al. 2014). The nuclease assays were performed on a 3' bottom strand labelled fork structure with a leading strand with a 10:1 protein-DNA ratio. We tested the activity of FAN1 on increasing concentrations of Mn^{2+}, Mg^{2+} and Zn^{2+} ions. The reactions were incubated at 37^{o}C for 30 min and run on a 20% denaturing PAGE gel. FAN1 has an optimal activity for Mn^{2+} ions at a final concentration of 1 mM of Mn^{2+} (Figure 14-B).

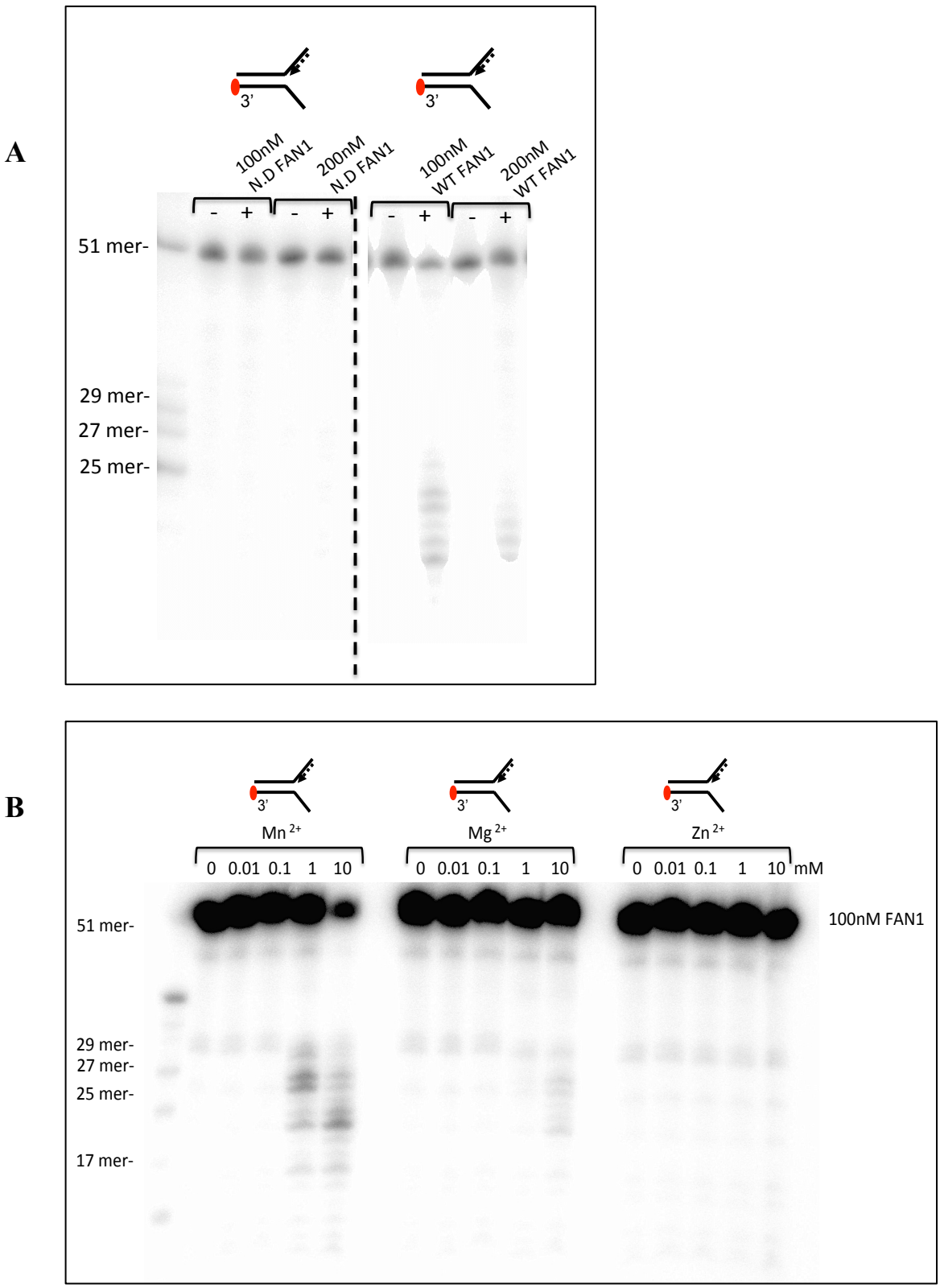
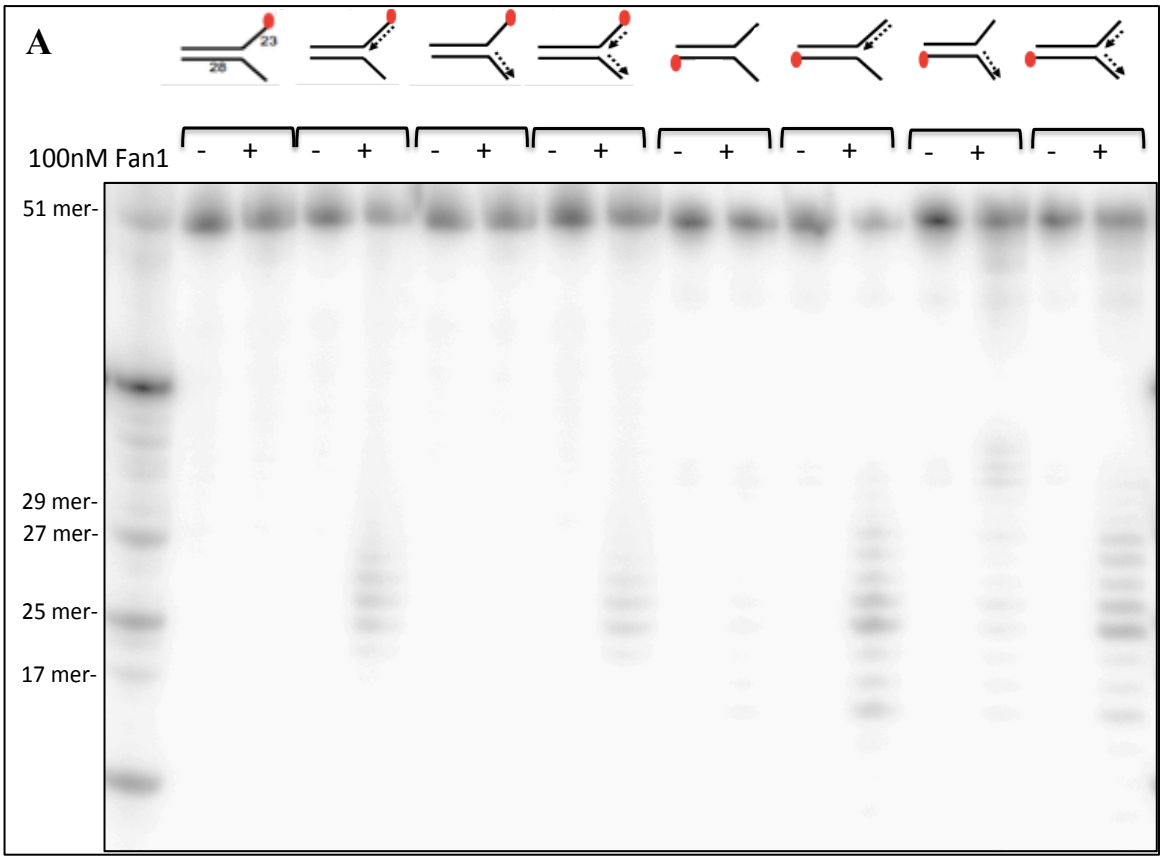
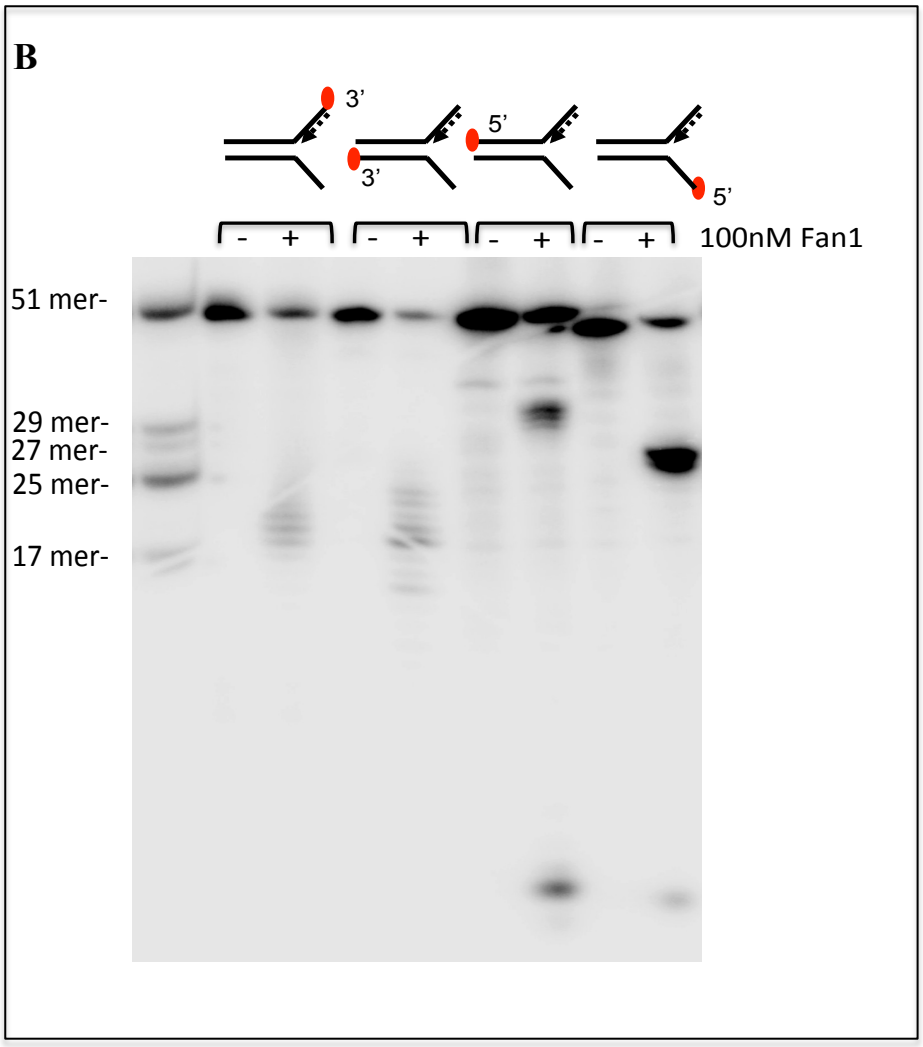


Figure 14- Preliminary characterization of FAN1 by nuclease assays

(A) Comparative nuclease assay between wild type (WT) FAN1 and nuclease dead (N.D) FAN1. Nuclease assays on a 10 nM 3'[^{32}P]-radiolabelled fork substrate with a leading strand revealed no signal for the nuclease dead protein but show incision for the WT confirming its activity. Ideal protein:DNA ratio appears to be 10:1. **(B)** FAN1 shows a preference for Mn^{2+} ions for optimal activity at a 1 mM concentration. Nuclease assays were performed with increasing ion concentrations (0 to 10 mM) at 37°C for 30 min on a 10 nM 3'[^{32}P]-radiolabelled simple fork substrate with a leading strand.

We then performed nuclease assays on FAN1 and an array of DNA structures in order to identify the DNA structural requirements for FAN1 activity, always using a 10:1 protein:DNA ratio and nuclease buffer (25 mM HEPES pH 8.0, 40 mM NaCl, 10% glycerol, 0.5 mM ß-mercaptoethanol, 0.1 mg/ml BSA) containing 1 mM $MnCl_2$ unless stated otherwise. 100nM of FAN1 was incubated with different α-[^{32}P] and γ-[^{32}P] labelled DNA structures at a concentration of 10 nM at 37^{o}C for 30 min with 1 mM of Mn^{2+} as a cofactor (Figure 15). Our results identify FAN1 as a 5' flap endonuclease that cuts the DNA in the duplex region, which is in agreement with the current publications. We believe FAN1 has a preferential endonuclease activity that creates a cut in the double stranded region thereby creating intermediate DNA structures. FAN1 also displays a 5'-3' exonuclease activity, cutting at every 3 nt: it is able to load onto the 5' end of the DNA and create consecutive cuts. FAN1 is also able to load onto the intermediate structures created by its endonuclease activity thereby giving rise to complex product formation making it hard to differentiate the contribution of its endo or exonuclease activity on the results obtained (MacKay et al. 2010; Wang et al. 2014; Pizzolato et al. 2015).


A
100nM Fan1
- + - + - + - + - + - + - + - +
51 mer-
29 mer-
27 mer-
25 mer-
17 mer-


B
3'
3'
5'
5'
- + - + - + - +
100nM Fan1
51 mer-
29 mer-
27 mer-
25 mer-
17 mer-

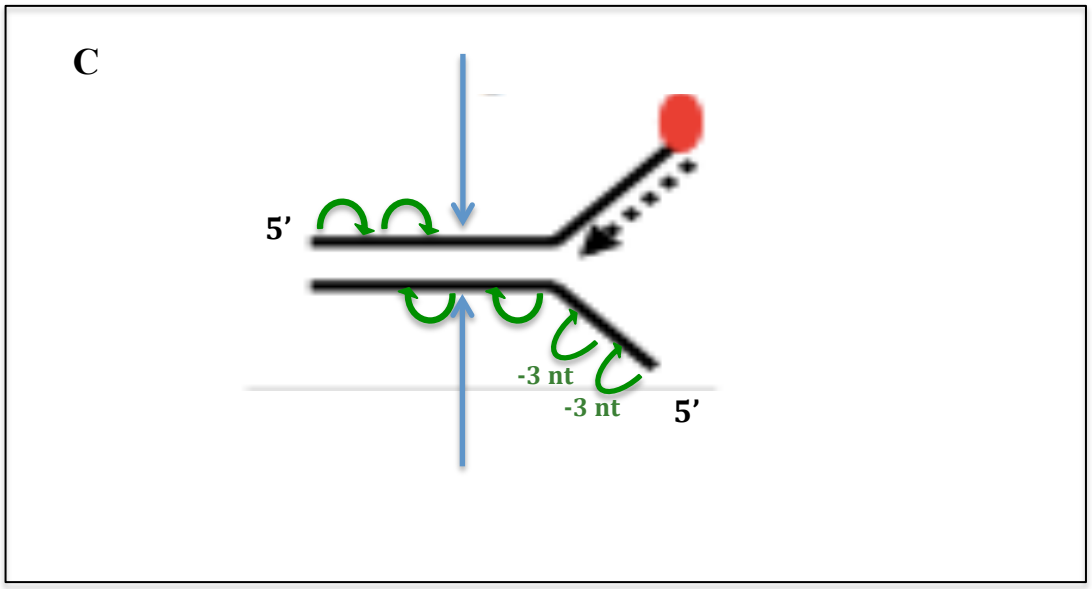


Figure 15- Identification of FAN1's 5' flap endonuclease activity

(A) Nuclease assays with 100 nM FAN1 were performed on a series of 10 nM 3'[^{32}P]-radiolabelled DNA structures. FAN1 only incises on structures that display a 5' flap as well as the presence of a leading strand. FAN1's preferential substrate seems to be the fork structure with a leading strand.

(B) Further nuclease assays were performed on both 10 nM 5'[^{32}P]-radiolabelled and 3' [^{32}P]-radiolabelled fork substrate with a leading strand (preferred substrate) with 100 nM of FAN1. FAN1 cleaves the substrate in the duplex region. Labels are represented by the red dot. Molecular weight is labeled on the side.

(C) Cartoon of the cleavage pattern of FAN1 on a 3'[^{32}P]-radiolabelled DNA leading strand structure. FAN1 possesses both an endo- and exo-activity leading to a complex digestion pattern. The blue arrow indicates the site of the main cut performed in the double stranded region of DNA by the endonuclease activity of FAN1. The green arrows show the consecutive digestion of the DNA in a 5' to 3' direction performed by the exonuclease activity of FAN1, cleaving every 3 nt. FAN1 both has an endo- and exonuclease activity leading to the formation of complex digestion products.

Previous research has implicated the 5'-3' ICL repair exonuclease SNM1A in ICL repair in an XPF-ERCC1 dependent pathway. In fact, SNM1A is able to load from XPF-ERCC1 induced incisions and digest past the ICL to complete ICL unhooking (Sengerová et al. 2012; Wang et al. 2011). We wish to investigate whether FAN1 is able to digest the DNA after an incision by XPF-ERCC1. Nuclease assays were performed on a 10 nM 3'[^{32}P]-radiolabelled fork substrate with a leading strand using both XPF-ERCC1 and FAN1. DNA substrates were incubated with 40 nM of XPF-ERCC1 at 30^{o}C for 1 hour with nuclease buffer containing 1 mM $MnCl_2$ before adding 100 nM of FAN1 and further incubating the mixture at 37^{o}C for 15 to 30 min. Results reveal that FAN1 is indeed able to digest the DNA after XPF-ERCC1 incision confirming its previously reported 5'-3' exonuclease activity (Pizzolato et al. 2015; MacKay et al. 2010). At first, only the two main incisions created by XPF-ERCC1 (products at 25 and 29mer) are visible. As incubation with FAN1 occurs, FAN1 digests the DNA in a 5'-3' direction, leaving a regular footprint that corresponds to successive cuts at every 3rd nucleotide (Figure 16). Although more experiments will need to be conducted, results suggest FAN1 could be implicated in ICL repair in an XPF-ERCC1 dependent pathway. FAN1 could load onto the DNA after an initial nick by XPF-ERCC1 and digest in a 5' to 3' direction. It still remains to be understood whether FAN1 plays a redundant role to other nucleases such as SNM1A.

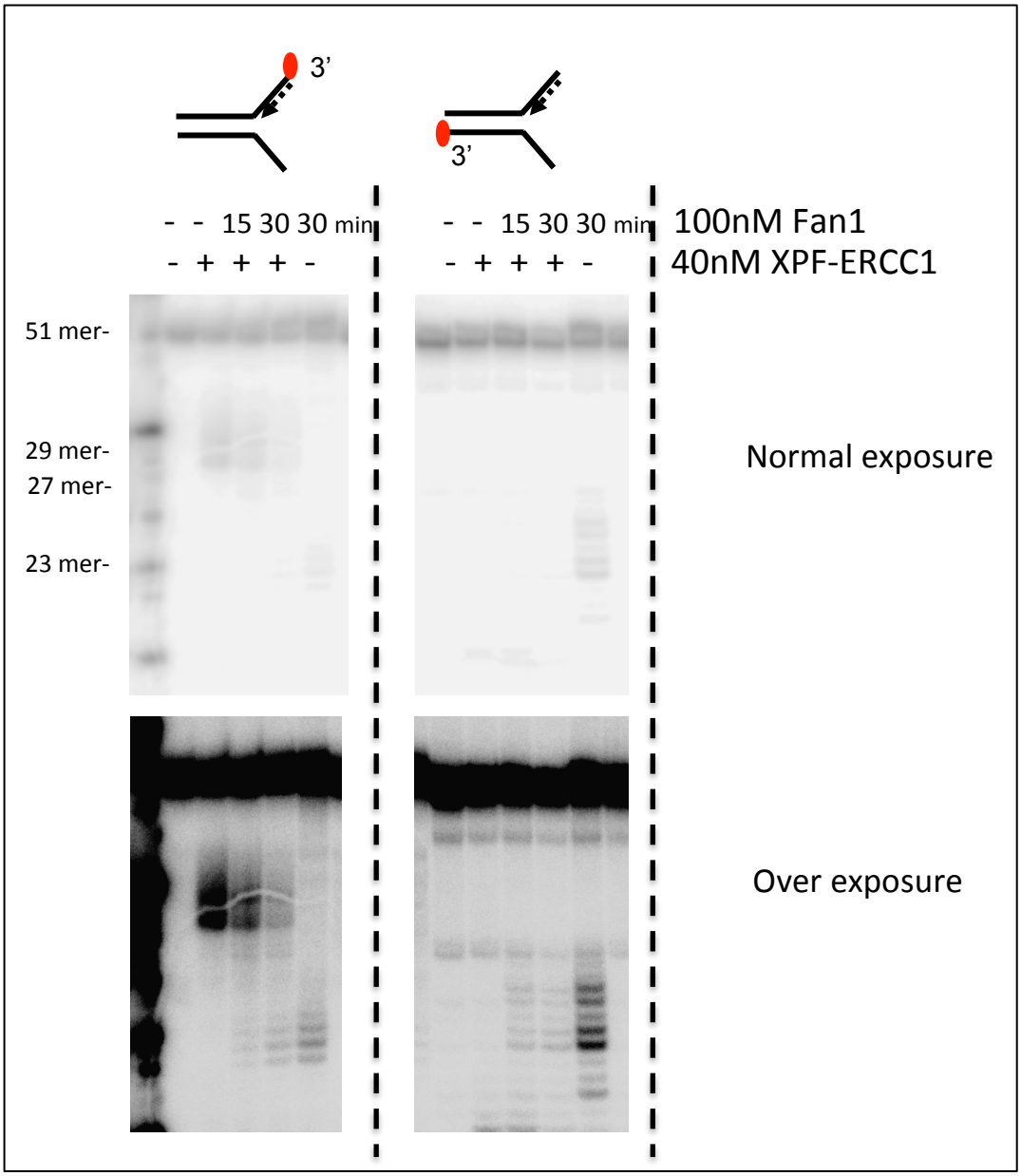


Figure 16- FAN1 is able to digest past an XPF-ERCC1 incision in a 5'-3' direction

Nuclease assays were performed with both XPF-ERCC1 and FAN1. Firstly, 3' [^{32}P]-radiolabelled DNA fork substrates with a leading strand were incubated with 40 nM XPF-ERCC1 for 1 hour at 30°C. Then, 100 nM FAN1 was added and the reaction further incubated for 15 to 30 min at 37°C. Gels confirm FAN1's 5' to 3' exonuclease activity. In fact, FAN1 cuts the DNA successively every 3rd /4th nucleotide.

Cellular and Genetic Characterisation of FAN1

Previous work in our lab confirmed that the 5'-3' exonuclease SNM1A participates in the same pathway as the endonuclease XPF-ERCC1, the essential factor in ICL unhooking (Sengerová et al. 2012; Wang et al. 2011). We wish to investigate the role of the nuclease FAN1 and in particular, its genetic relationship between SNM1A revealing a possible functional redundancy between the two factors.

Establishing the genetic relationship between SNM1A and FAN1

In order to establish the genetic relationship between SNM1A and FAN1, we decided to knock-down FAN1 using small-interfering RNA in human U2OS SNM1A -/- cells. (Courtesy of Dr Lonnie Swift from our lab) SNM1A -/- cells (named U2OS E2 SNM1A-/-) were transfected with FAN1 siRNA using HiPerfect reagent (Qiagen) according to the manufacturer's protocol. We first decided to optimise the knock-down experiment, using increasing concentrations of siRNA in order to establish the optimal concentration of siRNA needed. RT-qPCR was used to confirm siRNA gene silencing. SYBR Green RT-PCR Reagent kit (Applied Biosystems) was used to measure expression levels of SNM1A and FAN1. GAPDH housekeeping gene levels were measured as a control (Figure 17-A). Results determine 20 nM as the optimal concentration of siRNA to use to knock down FAN1 in U2OS derived cells. SNM1A knock out was also confirmed using RT-qPCR validating the cell line used in this experiment (Figure 17-B).

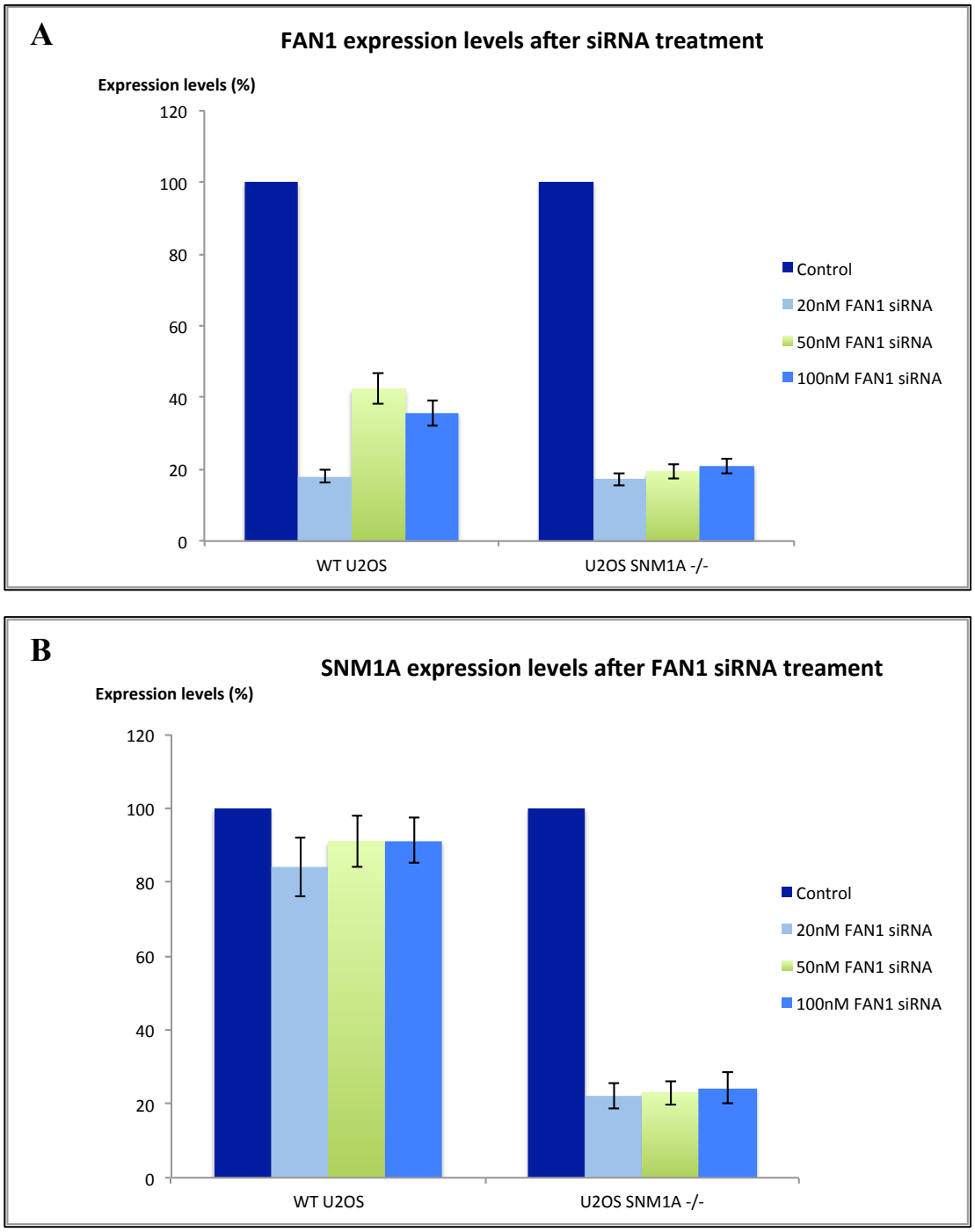


Figure 17- qPCR results for FAN1 and SNM1A expression levels in U2OS Wild type (WT) and SNM1A -/- cells

(A) FAN1 expression levels after siRNA treatment. U2OS WT and SNM1A -/- cells were treated with increasing concentrations (20 nM, 50 nM and 100 nM) of siRNA and expression levels tested by qPCR. Optimal siRNA concentration seems to be 20 nM.

(B) SNM1A expression levels in U2OS WT and SNM1A -/- U2OS E2 cells. In the same set up, we tested SNM1A expression in order to validate the SNM1A knock out cell line. WT U2OS were used as a positive control.

Clonogenic survival experiments

Successfully transfected U2OS WT and SNM1A -/- U2OS E2 cells with confirmed FAN1 knock down were used to perform clonogenic survival experiments. Cells were seeded at increasing cell numbers in 10 cm dishes and allowed to adhere for a minimum of 4 hours before drug treatment. Cells were treated with increasing concentrations (0 μM to 10 μM) of Cisplatin for 4 hours before fresh media was added. Cisplatin is a common ICL-inducing platinum-based reagent that has been demonstrated to create highly distorting DNA lesions (Wilson & Seidman 2010).

U2OS WT and U2OS SNM1A -/- cells were transfected with 20 nM siRNA against FAN1. Knock down was tested and confirmed by qRT-PCR (Figure 18-A). Survival clonogenics were then set up and treated with Cisplatin. After 12 days, the number of colonies that survived was counted and survival fractions determined (Figure 18-B). We observe increased sensitivity to Cisplatin for either SNM1A knock out or FAN1 knock down conditions individually, agreeing with previously published data (Wang et al. 2011). Interestingly, FAN1 knock down in SNM1A -/- cells hypersensitises the cells to the damaging agent, causing an additive response compared to a single condition. Altogether, this data demonstrates the functional redundancy of the exonuclease SNM1A and the repair factor FAN1. This could be interesting when thinking about the exact mechanism of FAN1 in ICL repair and how it interplays with the other nucleases in ICL unhooking.

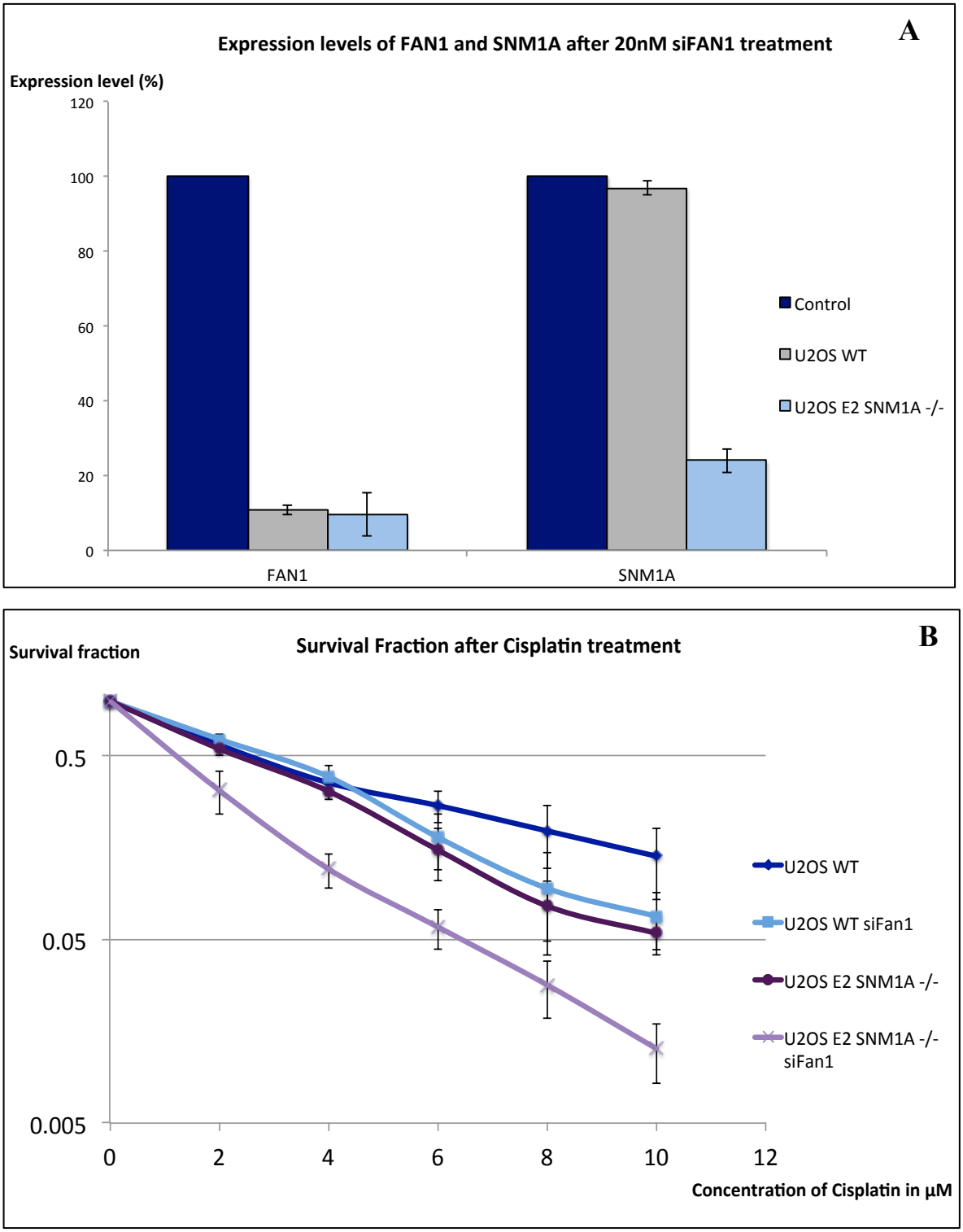


Figure 18- The nucleases SNM1A and FAN1 have redundant roles

(A) Validation of FAN1 knock down by qPCR in U2OS WT and U2OS SNM1A -/- cells.

(B) Survival fraction of colonies counted after Cisplatin treatment. A clonogenics experiment was set up with four different conditions: U2OS WT, U2OS WT with siRNA against FAN1, SNM1A -/- (E2) cells and SNM1A -/- (E2) cells with siRNA against FAN1. An additive, hypersensitive response can be seen in the SNM1A -/- cells with siRNA against FAN1. This suggest a functional redundancy between SNM1A and FAN1.

Establishing the genetic relationship between XPF-ERCC1 and FAN1

We have just established a functional and genetic redundancy between the key repair nucleases SNM1A and FAN1. Wang et al. (2010) have conducted previous genetic analysis confirming SNM1A operates in the same pathway as the key effector of ICL repair, the heterodimer endonuclease XPF-ERCC1. It would therefore be of interest to determine the genetic relationship between FAN1 and XPF-ERCC1 itself, revealing whether FAN1 is able to operate in an XPF-ERCC1 dependent pathway.

In order to establish the genetic relationship between FAN1 and XPF-ERCC1, we decided to generate an XPF disrupted human cell line. This goal was achieved using the CRISPR/Cas9 nuclease genome editing system. CRISPR/Cas9 relies on a site-specific RNA guide sequence cloned into the vector to cause targeted gene disruption by the Cas9 nuclease (Ran et al. 2013).

We decided to create an XPF -/- human cell line. Since previous attempts in our lab has failed to generate a knock out cell line using a single RNA guide sequence, we decided to design two guide sequences, targeting either end of exon 1 of XPF hoping to generate a deletion of around 200 base pairs that would disrupt gene function and expression (Figure 19).

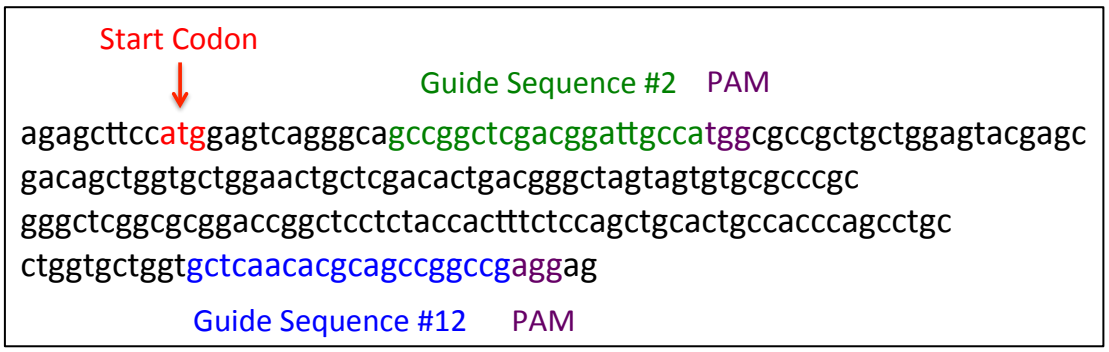


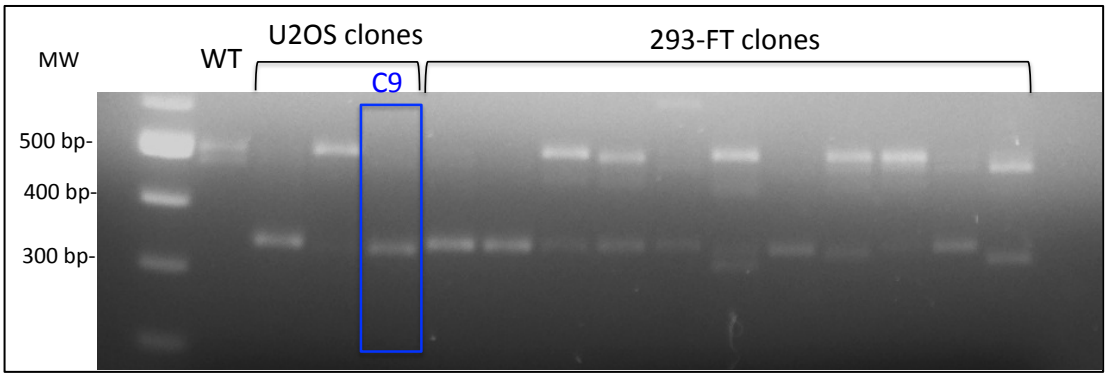
Figure 19- Sequence of Exon 1 of XPF

We decided to create a 200 base pair deletion in Exon 1 of the XPF gene. Guide sequences #2 and #12 were designed according the Zhang et al. (2013) guidelines. PAM sites serve a signal for Cas9 recruitment and cuts. Both guide sequences had a quality score of > 80% and minimal number of off-target sites.

After successful cloning of the guide sequences into pX330-GFP and pX458-Ruby CRISPR/Cas9 vectors, U2OS and 293-FT human cells were transfected with both plasmids using Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's protocol. After FACS sorting, single colonies were grown in 96 well plates before expanding surviving colonies into a 6 well plates and testing for XPF expression levels. Firstly, we decided to perform a PCR on the surviving clones. Genomic DNA was extracted from each sample and a PCR reaction completed according to the manufacturer's guidelines. Wild type exon 1 yields a product of around 500 base pairs (bp). We expect XPF -/- clones to give a PCR product of around 300 bp if our anticipated 200 bp deletion has been successful. In fact, a number of clones from both cell lines seem to have successfully been disrupted. In order to further confirm XPF disruption on a protein level, we performed Western Blot analysis using monoclonal XPF antibody to confirm our PCR results (Figure 20 B).
Western blots further confirmed our successful generation of XPF -/- human cell lines. XPF -/- clones were obtained in both U2OS and 293-FT cell lines.

Final validation of the knock out was done through comparative sequencing analysis of all the clones obtained and wild type U2OS cells.

A



B

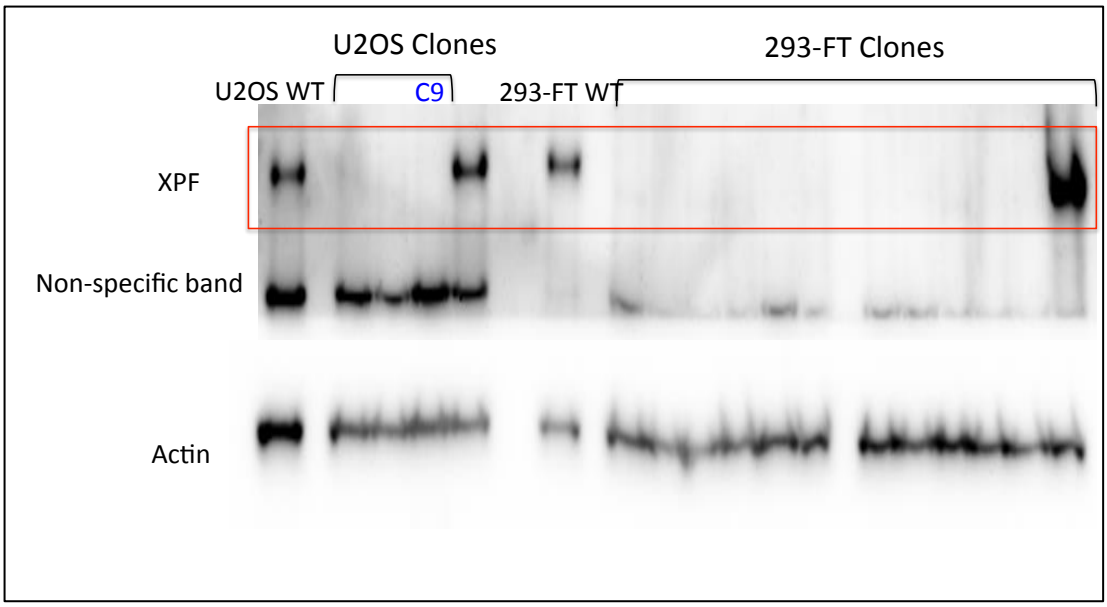


Figure 20- Validation of XPF -/- human cell lines generated through CRISPR/Cas 9 genome editing method

(A) Genomic PCR of XPF exon 1 of obtained clones. Wild type exon 1 has a size of around 500 bp. Expected XPF -/- clones should have a 200bp deletion in Exon 1, yielding a 300 bp product. MW= molecular weight.
(B) Western Blot of U2OS and 293-FT clones. Results show no detectable XPF expression in most of the clones tested. Wild Type U2OS and 293-FT cells were used as positive controls. Actin was used as a loading control.

The U2OS C9 XPF -/- clone was chosen to perform further experiments.

We decided to use the XPF -/- C9 U2OS clone for further experiments. Since we wish to perform clonogenic experiments, U2OS cells seemed the better choice as 293-FT cells hold the properties of being very easily detachable when in culture making them less suitable for clonogenic survival experiments. The C9 XPF -/- clone was chosen based on its sequencing analysis: during CRISPR/Cas9 treatment, an out of frame mutation occurred introducing a premature STOP codon in Exon 2 of XPF, making it a likely complete loss-of-function cell line.

In order to validate fully our XPF -/- cell line (U2OS C9 clone), we performed a clonogenic survival experiment with increasing doses of the ICL-inducing agent Cisplatin. Strikingly, human XPF -/- cells exhibit extreme levels of sensitivity to the drug (Figure 21). Clonogenic survival experiments were also performed with higher concentrations of Cisplatin. In fact, concentration above as little as 2 μM of Cisplatin yielded no cell survival making data analysis impossible. Altogether, the data further confirms XPF's crucial role in ICL repair initiation and establishes the role of XPF-ERCC1 as the key mediator ICL unhooking.

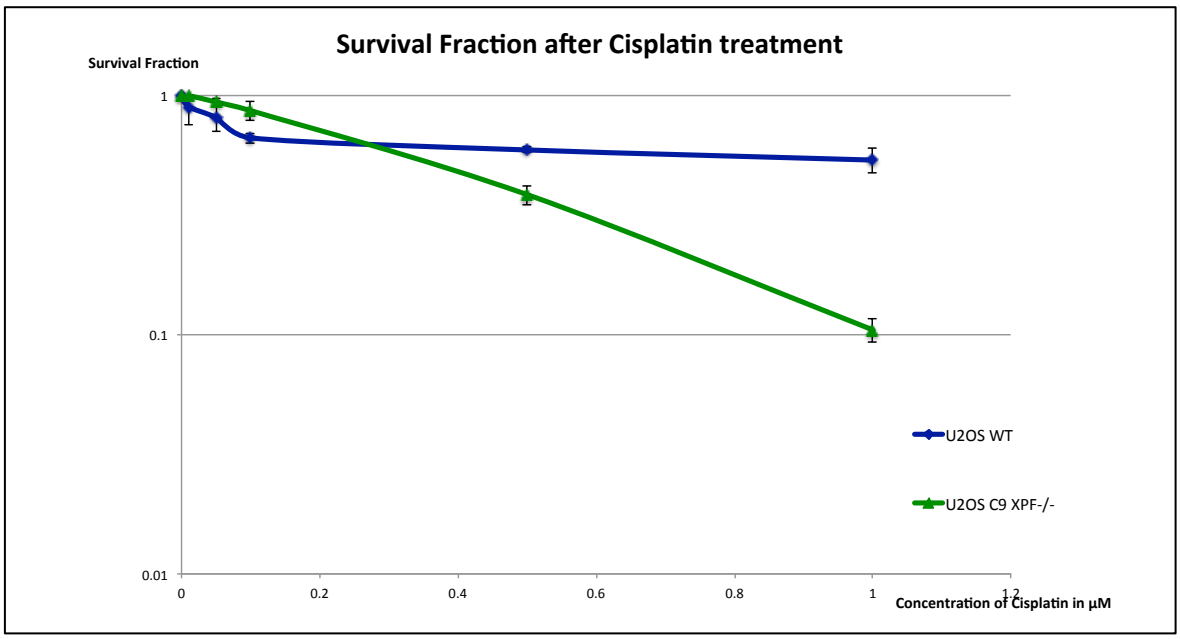


Figure 21- Validation of XPF -/- human cell lines generated through CRISPR/Cas 9 genome editing method by measuring Cisplatin sensitivity in a clonogenic survival experiment

The surviving fraction of the colonies was calculated after U2OS Wild type and U2OS XPF -/- cells were treated with increasing doses of Cisplatin. XPF -/- cells exhibited hypersensitivity to cisplatin compared to the Wild Type, further validating our knock out.

Clonogenic survival experiments

With this new genetic tool generated, we wished to further understand the genetic relationship between XPF-ERCC1 and FAN1. Biochemistry results suggested FAN1 might be involved in ICL repair in an XPF-ERCC1 dependent pathway. In order to establish whether FAN1 and XPF-ERCC1 operate in the same pathway, we knocked down FAN1 in U2OS WT and XPF -/- cells using small-interfering RNA. Cells were transfected with FAN1 siRNA using HiPerfect reagent (Qiagen) according to the manufacturer's protocol. SYBR Green RT-PCR Reagent kit (Applied Biosystems) was used to measure expression levels of SNM1A and FAN1. GAPDH housekeeping gene

levels were measured as a control. Successful knock down of FAN1 was obtained (Figure 22-A).

U2OS WT and XPF -/- C9 cells with confirmed FAN1 knock down were then used to perform clonogenic survival experiments. Cells were seeded at increasing cell numbers in 10 cm dishes and allowed to adhere for a minimum of 4 hours before drug treatment. Cells were treated with increasing concentrations (0 μM to 1 μM) of Cisplatin for 4 hours before fresh media was added. After 12 days, the number of colonies that survived was counted and surviving fractions determined (Figure 22-B). We observed a very pronounced increased sensitivity to Cisplatin for XPF knock out cells. Although we know U2OS WT cells with siRNA against FAN1 are sensitive to Cisplatin, the concentration range of 0 μM to 1 μM used in this specific experiment is too low for it to be observed. Strikingly, FAN1 knock down in XPF -/- cells further sensitises the cells to the damaging agent, causing an additive response compared to the knock down alone. This suggests FAN1 and XPF-ERCC1 do not operate in the same pathway, a surprising conclusion considering FAN1 and SNM1A have been suggested to have redundant roles and SNM1A has been shown to operate in an XPF-ERCC1 dependent pathway (Wang et al. 2011). This result offers new, interesting research questions when thinking about the mechanism of FAN1 in ICL repair in vivo.

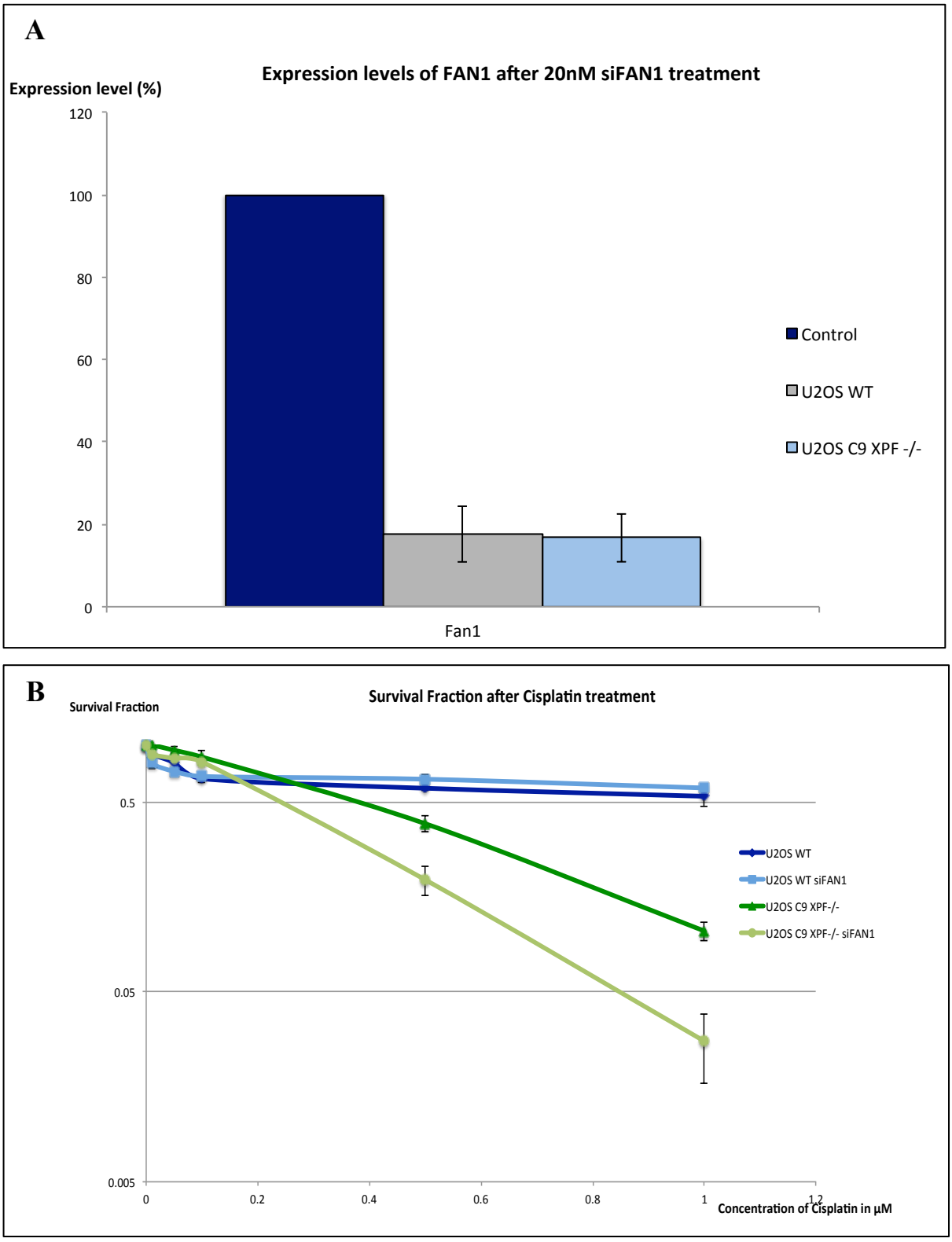


Figure 22- The nucleases FAN1 and XPF-ERCC1 do not operate in the same pathway

(A) Validation of FAN1 knock down by qPCR in U2OS WT and U2OS XPF -/- cells. **(B)** Survival fraction of colonies counted after Cisplatin treatment. A clonogenics experiment was set up with four different conditions: U2OS WT, U2OS WT with siRNA against FAN1, XPF -/- (C9) cells and XPF -/- (C9) cells with siRNA against FAN1. An additive, extremely sensitive response can be seen in the XPF -/- cells with siRNA against FAN1. This suggest FAN1 does not operate in an XPF-ERCC1 dependent pathway.

DISCUSSION

DNA interstrand crosslinks (ICLs) are particularly toxic DNA lesions as when left unrepaired, they are able to block replication producing devastating health conditions such as the genetic syndrome Fanconi Anemia (FA) that predisposes its patients to cancer (Hashimoto et al. 2016; Wang & Gautier 2010).

The repair of ICLs is thought to be initiated by the concerted action of multiple nucleases, some of which are defective in FA patients (Sengerová et al. 2011; Wang et al. 2011). In this thesis, we aim to understand the genetic relationship between the key endonuclease XPF-ERCC1 and two repair factors SNM1A and FAN1 during replication-associated ICL repair. XPF-ERCC1, SNM1A and FAN1 all show increased sensitivity to widely used crosslinking anticancer drugs, underlining the clinical importance of exemplifying the exact mechanism of ICL repair (Wang et al. 2011; Kratz, Schöpf, Kaden, Sendoel, Eberhard, et al. 2010; Douwel et al. 2014).

XPF-ERCC1 has been demonstrated to be a key protein involved in ICL unhooking. Immunodepletion experiments in *Xenopus laevis* cell free egg extracts reveal that only XPF depletion abolishes ICL repair whereas depletion of other nucleases only has minor effects (Hashimoto et al. 2016; Zhang et al. 2015). Work from our own laboratory suggests, however, that XPF-ERCC1 is able to incise 5' to the ICL cutting at two distinct sites in the duplex region. Since no incisions 3' to the ICL were observed, the action of one or more other nucleases might be required to complete ICL unhooking. The 5' to 3' exonuclease SNM1A has already been demonstrated to participate in ICL repair in the same pathway as XPF-ERCC1 (Sengerová et al. 2012).

In fact, SNM1A can load from XPF-ERCC1 induced incisions and digest past the ICL to complete ICL unhooking (Sengerová et al. 2012; Wang et al. 2011).

Nonetheless, some observations have been made that other repair factors might collaborate with XPF-ERCC1, possibly in a redundant way. It has been demonstrated that SNM1A -/- cells display a milder phenotype when exposed to ICL-inducing agents as opposed to cells disrupted in XPF (Wang et al. 2011). Furthermore, some Fanconi Anemia patients display symptoms of the disease but do not have any mutations in the FA genes (Deans & West 2011). We therefore decided to investigate the role of Fanconi-associated protein 1 (FAN1) in ICL unhooking. FAN1 has been reported to possess a UBZ domain required for its interaction with monoubiquitylated FANCD2 recruiting it to the site of ICL damage. It also contains a PD-D/E(X)K type endonuclease motif within a Virus-type replication repair nuclease domain at its C terminus, present in other repair nucleases (MacKay et al. 2010; Allerston et al. 2015).

Through *in vitro* nuclease assays, we sought to understand the contribution of FAN1 on an array of synthetic DNA structures that mimic ICL processing intermediates. Our results identify FAN1 as a 5' flap endonuclease that cuts the DNA in the duplex region, which is in agreement with previous publications. It is important to note that FAN1 has also been identified to have an intrinsic 5'-3' exonuclease-like activity giving rise to complex product formation making it hard to differentiate the contribution of its endo or exonuclease activity on the results obtained (MacKay et al. 2010; Wang et al. 2014; Pizzolato et al. 2015). Collaborative studies of FAN1 with XPF-ERCC1 on a fork substrate with a leading strand suggest FAN1 could be implicated in ICL repair in an XPF-ERCC1 dependent pathway. It is hypothesized that FAN1 is able to load onto the DNA after an initial nick by XPF-ERCC1 and digest in a

5’ to 3’ direction. It still remains to be understood whether FAN1 plays a redundant role to other nucleases such as SNM1A or not.

In order to obtain a better idea of the genetic relationship between FAN1 and SNM1A and establish a possible functional redundancy, we performed clonogenic survival experiments on WT U2OS and U2OS SNM1A -/- cells treated with siRNA against FAN1. Plates were exposed to increasing concentration of Cisplatin, a platinum based ICL-inducing chemotherapeutic drug. Results confirmed a functional redundancy between FAN1 and SNM1A. Since SNM1A is proven to operate in the same pathway as the endonuclease XPF-ERCC1, we decided to establish the genetic relationship between FAN1 and XPF. In order to do so, we created a XPF disrupted human U2OS cell line using the genome editing CRISPR-Cas9 method. Once a validated cell line was confirmed, clonogenic survival experiments were performed on WT U2OS and U2OS XPF -/- with siRNA against FAN1. Strikingly, FAN1 knock down in XPF -/- cells extremely sensitises the cells to the damaging agent, causing an additive response compared to the knock down only. This suggests FAN1 and XPF-ERCC1 do not operate in the same pathway, a surprising conclusion considering FAN1 and SNM1A have redundant roles and SNM1A has been implicate to operate in an XPF-ERCC1 dependent pathway. This result offers new, interesting research questions when thinking about the mechanism of FAN1 in ICL repair *in vivo*.

Altogether, the results offer new insight into the interplay of several nucleases involved in ICL repair and in particular, ICL unhooking. In order to further understand the biochemical and genetic image of ICL repair, further experiments need to be conducted. In particular, a better understanding of the FAN1 digestion footprint is

necessary. Because of both its endo- and exo-nucleolytic activities, FAN1 reactions lead to complex digestion products. It is therefore necessary to establish an alternative method to analyse FAN1's action with more precision. It would also be of interest to determine the action of FAN1 in collaboration with XPF-ERCC1 on synthetic DNA structures containing ICLs. It can be hypothesized that FAN1 is able to load onto a nick previously made by XPF-ERCC1 and digest past a DNA ICL, comparable to the activity of SNM1A. On a genetic level, confirming the epistasis of the nucleases XPF-ERCC1 and SNM1A could provide a more complete image of their genetic background. With the newly established human U2OS XPF -/- cell lines, clonogenic survival experiments with treatment with siRNA for either or both FAN1 and SNM1A could offer a more comprehensive map of the genetic relationships between all the nucleases involved in ICL unhooking. New insight into the mechanistic and genetic details of ICL unhooking could provide an important source for potential therapies in a variety of rare genetic syndromes and cancer chemotherapeutics.

FUTURE IMPLICATIONS

In this thesis, we identify the role of FAN1 on a biochemical and cellular level. Our biochemical study reveals FAN1's complex nuclease activity. In fact, we establish FAN1 as a 5' flap endonuclease that cuts the DNA in the duplex region as well as a 5'-3' exonuclease. Because of both the endo- and exonuclease activity of FAN1, our experiments give rise to a complex digestion pattern. It is of interest to further characterise the exact cleavage pattern of FAN1 and aim to differentiate the contribution of each of these different activities of FAN1 on DNA substrates. This would eventually be achievable by carefully designing various DNA substrates with specific modifications on either end (for example terminal phosphate or biotin moieties). In addition, having different lengths of DNA oligos to create various intermediate structures could help understand FAN1's footprint and its structural DNA preferences. In the context of DNA ICL unhooking, further nuclease assays with XPF-ERCC1 and SNM1A should be conducted to clarify a possible functional redundancy between SNM1A and FAN1.

On a cellular level, our results prove that FAN1 does not seem to be operating in an XPF-dependent pathway. In contrast, previous studies in our lab established that SNM1A and XPF operate in the same pathway. As biochemical studies suggest that FAN1 and SNM1A have redundant roles, it is of interest to further understand the exact genetic relationship between the two nucleases and the ICL repair machinery. A clonogenic survival assay with increasing Cisplatin concentrations in XPF -/- cells with both SNM1A and FAN1 siRNA knock downs will help to confirm the non-epistatic relationship between the repair factors. On a broader level, further clonogenic survival

assays where other key protein of the ICL FA pathway are silenced will shed more light on the genetic and temporal relationship between them. Further, using different types of ICL-inducing drugs, such as Nitrogen Mustard or Mitomycin C would be a further validation of our findings and could help clarify the contribution of DNA distortion and type of DNA damage to DNA damage response. We hope that by better understanding how the specific factors interact with each other, we can reveal the precise ICL unhooking mechanism.

In conclusion, we believe that further work in this field will eventually lead to a better understand of ICL unhooking, the key step in ICL repair. This will offer new therapeutic potential, especially when looking at chemotherapeutics.

REFERENCES

Allerston, C.K. et al., 2015. The structures of the SNM1A and SNM1B/Apollo nuclease domains reveal a potential basis for their distinct DNA processing activities. *Nucleic Acids Research*, 43(10), pp.11047–11060.

Bhagwat, N. et al., 2009. XPF-ERCC1 participates in the Fanconi anemia pathway of cross-link repair. *Molecular and cellular biology*, 29(24), pp.6427–37. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19805513 [Accessed October 13, 2016].

Cadet, J., Douki, T. & Ravanat, J.-L., 2011. Measurement of oxidatively generated base damage in cellular DNA. *Mutation research*, 711(1–2), pp.3–12. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21329709 [Accessed October 28, 2016].

Cadet, J. & Wagner, J.R., 2013. DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation. *Cold Spring Harbor perspectives in biology*, 5(2). Available at: http://www.ncbi.nlm.nih.gov/pubmed/23378590

Cheung-Ong, K., Giaever, G., Nislow, C., et al., 2013. DNA-Damaging Agents in Cancer Chemotherapy: Serendipity and Chemical Biology. *Chemistry & Biology*, 20(5), pp.648–659. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1074552113001312 [Accessed October 28, 2016].

Cheung-Ong, K., Giaever, G. & Nislow, C., 2013. DNA-Damaging Agents in Cancer Chemotherapy: Serendipity and Chemical Biology. *Chemistry & Biology*, 20(5), pp.648–659.

Clauson, C., Scharer, O.D. & Niedernhofer, L., 2013. Advances in Understanding the Complex Mechanisms of DNA Interstrand Cross-Link Repair. *Cold Spring Harbor Perspectives in Biology*, 5(10), pp.a012732–a012732. Available at: http://cshperspectives.cshlp.org/lookup/doi/10.1101/cshperspect.a012732 [Accessed November 2, 2016].

Clauson, C., Schärer, O.D. & Niedernhofer, L., 2013. Advances in understanding the complex mechanisms of DNA interstrand cross-link repair. *Cold Spring Harbor perspectives in biology*, 5(10), p.a012732. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24086043 [Accessed November 2, 2016].

Deans, A.J. & West, S.C., 2011. DNA interstrand crosslink repair and cancer. *Nature Reviews Cancer*, 11(7), pp.467–480. Available at: http://www.nature.com/doifinder/10.1038/nrc3088 [Accessed October 26, 2016].

Douwel, D.K. et al., 2014. XPF-ERCC1 Acts in Unhooking DNA Interstrand Crosslinks in Cooperation with FANCD2 and FANCP/SLX4. *Molecular Cell*, 54, pp.460–471. Available at: http://dx.doi.org/10.1016/j.molcel.2014.03.015 [Accessed October 13, 2016].

Enoiu, M., Jiricny, J. & Schärer, O.D., 2012. Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. *Nucleic acids research*, 40(18), pp.8953–64. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22810206 [Accessed October 27, 2016].

Enzlin, J.H. & Schärer, O.D., 2002. The active site of the DNA repair endonuclease XPF-ERCC1 forms a highly conserved nuclease motif. *The EMBO journal*, 21(8), pp.2045–53. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11953324 [Accessed October 23, 2016].

Fojo, T., 2001. Cancer, DNA repair mechanisms, and resistance to chemotherapy. *Journal of the National Cancer Institute*, 93(19), pp.1434–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11584051 [Accessed October 28, 2016].

Frenkel, K., 1992. Carcinogen-mediated oxidant formation and oxidative DNA damage. *Pharmacology & Therapeutics*, 53(1), pp.127–166. Available at: http://linkinghub.elsevier.com/retrieve/pii/0163725892900474 [Accessed October 27, 2016].

Friedberg, E.C., Mcdaniel, L.D. & Schultz, R.A., 2004. The role of endogenous and exogenous DNA damage and mutagenesis. *Current Opinion in Genetics & Development*, 14, pp.5–10.

Ghosal, G. & Chen, J., 2013. DNA damage tolerance: a double-edged sword guarding the genome. *Translational Cancer Research*, 2(3), pp.107–129.

Gwon, G.H. et al., 2014. Crystal structure of a fanconi anemiaassociated nuclease homolog bound to 5??? flap DNA: Basis of interstrand cross-link repair by FAN1. *Genes and Development*, 28(20), pp.2276–2290.

Hanahan, D. et al., 2011. Hallmarks of cancer: the next generation. *Cell*, 144(5), pp.646–74. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21376230 [Accessed October 27, 2016].

Hashimoto, S. et al., 2016. Mechanisms of interstrand DNA crosslink repair and human disorders. *Genes and Environment*, 38(1), p.9. Available at: http://www.genesenvironment.com/content/38/1/9 [Accessed October 26, 2016].

Helleday, T. et al., 2008. DNA repair pathways as targets for cancer therapy. *Nature Reviews Cancer*, 8(3), pp.193–204. Available at: http://www.nature.com/doifinder/10.1038/nrc2342 [Accessed October 26, 2016].

Helleday, T., Eshtad, S. & Nik-Zainal, S., 2014. Mechanisms underlying mutational signatures in human cancers. *Nature Reviews Genetics*, 15(9), pp.585–598. Available at: http://www.nature.com/doifinder/10.1038/nrg3729 [Accessed October 27, 2016].

Hinz, J.M., 2010. Role of homologous recombination in DNA interstrand crosslink repair. *Environmental and Molecular Mutagenesis*, 51(6), pp.582–603.

Ho, T.V. & Schärer, O.D., 2010. Translesion DNA synthesis polymerases in DNA interstrand crosslink repair. *Environmental and Molecular Mutagenesis*, 51(6), pp.552–566.

Hoeijmakers, J.H., 2001. Genome maintenance mechanisms for preventing cancer. *Nature*, 411(6835), pp.366–74. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11357144 [Accessed October 27, 2016].

Ishiai, M. et al., 2008. FANCI phosphorylation functions as a molecular switch to turn on the Fanconi anemia pathway. *Nature structural & molecular biology*, 15(11), pp.1138–46. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18931676 [Accessed October 27, 2016].

Iyama, T. & Wilson, D.M., 2013. DNA repair mechanisms in dividing and non-dividing cells. *DNA Repair*, 12(8), pp.620–636.

Jackson, A.L. & Loeb, L.A., 2001. The contribution of endogenous sources of DNA damage to the multiple mutations in cancer. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 477(1–2), pp.7–21. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0027510701000914 [Accessed October 27, 2016].

Jackson, S.P. & Bartek, J., 2009. The DNA-damage response in human biology and disease. *Nature*, 461(7267), pp.1071–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19847258 [Accessed October 28, 2016].

Jiricny, J., 2006. The multifaceted mismatch-repair system. *Nature reviews. Molecular cell biology*, 7(5), pp.335–46. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16612326 [Accessed October 28, 2016].

Knoch, J. et al., Rare hereditary diseases with defects in DNA-repair. *European journal of dermatology : EJD*, 22(4), pp.443–55. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22436139 [Accessed October 26, 2016].
Kottemann, M.C. & Smogorzewska, A., 2013. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature*, 493(7432), pp.356–363. Available at: http://www.nature.com/doifinder/10.1038/nature11863 [Accessed October 27, 2016].
Kratz, K., Schöpf, B., Kaden, S., Sendoel, A., Eberhard, R., et al., 2010. Deficiency of FANCD2-Associated Nuclease KIAA1018/FAN1 Sensitizes Cells to Interstrand Crosslinking Agents. *Cell*, 142(1), pp.77–88. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092867410006768 [Accessed September 18, 2016].
Kratz, K., Schöpf, B., Kaden, S., Sendoel, A., Eberhard, R., et al., 2010. Deficiency of FANCD2-associated nuclease KIAA1018/FAN1 sensitizes cells to interstrand crosslinking agents. *Cell*, 142(1), pp.77–88. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20603016 [Accessed October 27, 2016].
Lehmann, A.R., 2003. DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *Biochimie*, 85(11), pp.1101–11. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14726016 [Accessed October 26, 2016].
Li, X. & Heyer, W.-D., 2008. Homologous recombination in DNA repair and DNA damage tolerance. *Cell research*, 18(1), pp.99–113. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18166982 [Accessed November 2, 2016].
Lieber, M.R., 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual review of biochemistry*, 79, pp.181–211. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20192759 [Accessed November 2, 2016].
Lindahl, T., 1993. Instability and decay of the primary structure of DNA. *Nature*, 362(6422), pp.709–15. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8469282 [Accessed October 27, 2016].
MacKay, C. et al., 2010. Identification of KIAA1018/FAN1, a DNA repair nuclease recruited to DNA damage by monoubiquitinated FANCD2. *Cell*, 142(1), pp.65–76. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20603015 [Accessed October 27, 2016].
Maréchal, A. & Zou, L., 2013. DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harbor perspectives in biology*, 5(9). Available at: http://www.ncbi.nlm.nih.gov/pubmed/24003211 [Accessed October 28, 2016].
Mouret, S. et al., 2006. Cyclobutane pyrimidine dimers are predominant DNA lesions in whole human skin exposed to UVA radiation. *Proceedings of the National Academy of Sciences of the United States of America*, 103(37), pp.13765–70. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16954188 [Accessed October 28, 2016].
Moynahan, M.E. & Jasin, M., 2010. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nature Reviews Molecular Cell Biology*, 11(3), pp.196–207. Available at: http://www.nature.com/doifinder/10.1038/nrm2851 [Accessed November 2, 2016].
Muniandy, P.A. et al., 2010. DNA interstrand crosslink repair in mammalian cells: step by step. *Critical Reviews in Biochemistry and Molecular Biology*, 45(1), pp.23–49. Available at: http://www.tandfonline.com/doi/full/10.3109/10409230903501819 [Accessed November 2, 2016].

Nikitaki, Z. et al., 2015. Stress-induced DNA damage biomarkers: applications and limitations. *Frontiers in chemistry*, 3, p.35. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26082923 [Accessed October 28, 2016].
O 'connor, M.J., 2015. Targeting the DNA Damage Response in Cancer. *Molecular Cell*, 60, pp.547–560. Available at: http://dx.doi.org/10.1016/j.molcel.2015.10.040 [Accessed October 28, 2016].
Patel, K.J. & Joenje, H., 2007. Fanconi anemia and DNA replication repair. *DNA Repair*, 6(7), pp.885–890.
Pearl, L.H. et al., 2015. Therapeutic opportunities within the DNA damage response. *Nature Reviews Cancer*, 15(3), pp.166–180. Available at: http://www.nature.com/doifinder/10.1038/nrc3891 [Accessed October 28, 2016].
Pizzolato, J. et al., 2015. FANCD2-associated nuclease 1, but not exonuclease 1 or flap endonuclease 1, is able to unhook DNA interstrand cross-links in vitro. *Journal of Biological Chemistry*, 290(37), pp.22602–22611.
Ran, F.A. et al., 2013. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11), pp.2281–2308. Available at: http://www.nature.com/doifinder/10.1038/nprot.2013.143 [Accessed October 24, 2016].
Räschle, M. et al., 2008. Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell*, 134(6), pp.969–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18805090 [Accessed October 27, 2016].
Saleh-Gohari, N. et al., 2005. Spontaneous homologous recombination is induced by collapsed replication forks that are caused by endogenous DNA single-strand breaks. *Molecular and cellular biology*, 25(16), pp.7158–69. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16055725 [Accessed November 2, 2016].
Santivasi, W.L. & Xia, F., 2014. Ionizing radiation-induced DNA damage, response, and repair. *Antioxidants & redox signaling*, 21(2), pp.251–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24180216 [Accessed October 28, 2016].
Schärer, O.D., 2003. Chemistry and Biology of DNA Repair. *Angewandte Chemie International Edition*, 42(26), pp.2946–2974. Available at: http://doi.wiley.com/10.1002/anie.200200523 [Accessed October 28, 2016].
Schärer, O.D., 2013. Nucleotide excision repair in eukaryotes. *Cold Spring Harbor perspectives in biology*, 5(10), p.a012609. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24086042 [Accessed October 28, 2016].
Sengerová, B. et al., 2012. Characterization of the human SNM1A and SNM1B/Apollo DNA repair exonucleases. *Journal of Biological Chemistry*, 287(31), pp.26254–26267.
Sengerová, B., Wang, A.T. & McHugh, P.J., 2011. Orchestrating the nucleases involved in DNA interstrand cross-link (ICL) repair. *Cell Cycle*, 10(23), pp.3999–4008.
Shiloh, Y., 2003. ATM and related protein kinases: safeguarding genome integrity. *Nature reviews. Cancer*, 3(3), pp.155–68. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12612651 [Accessed October 28, 2016].
Smogorzewska, A. et al., 2007. Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. *Cell*, 129(2), pp.289–301. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17412408 [Accessed October 27, 2016].
Thongthip, S. et al., 2016. Fan1 deficiency results in DNA interstrand cross-link repair defects, enhanced tissue karyomegaly, and organ dysfunction. *Genes & development*, 30(6), pp.645–59. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26980189 [Accessed October 28, 2016].

Waters, L.S. et al., 2009. Eukaryotic Translesion Polymerases and Their Roles and Regulation in DNA Damage Tolerance. *Microbiology and Molecular Biology Reviews : MMBR*, 73(1), pp.134–154. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2650891/.

Wang, A.T. et al., 2011. Human SNM1A and XPF-ERCC1 collaborate to initiate DNA interstrand cross-link repair. *Genes & development*, 25(17), pp.1859–70. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21896658 [Accessed October 27, 2016].

Wang, A.T. & Smogorzewska, A., 2015. SnapShot: Fanconi Anemia and Associated Proteins. *Cell*, 160(1–2), p.354–354.e1. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092867414016377 [Accessed November 2, 2016].

Wang, L.C. & Gautier, J., 2010. The Fanconi anemia pathway and ICL repair: implications for cancer therapy. *Critical reviews in biochemistry and molecular biology*, 45(5), pp.424–39. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20807115 [Accessed November 2, 2016].

Wang, R. et al., 2014. Mechanism of DNA interstrand cross-link processing by repair nuclease FAN1. *Science*, 346(6213), pp.1127–1130. Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.1258973 [Accessed September 18, 2016].

Wilson, D.M. & Seidman, M.M., 2010. A novel link to base excision repair? *Trends in Biochemical Sciences*, 35(5), pp.247–252.

Zhang, J. et al., 2015. DNA interstrand cross-link repair requires replication-fork convergence. *Nature Structural & Molecular Biology*, 22(3), pp.242–247. Available at: http://www.nature.com/doifinder/10.1038/nsmb.2956 [Accessed October 5, 2016].