Analysis of DNA Repair Proteins and Chromatin Remodeling Factors Using iPOND

Caroline Danielsen Søgaard

Acknowledgements

Summary

Most chemotherapeutic treatments rely on induction of severe DNA damage to kill the cancer cells. However, DNA repair pathways can repair the induced lesions, leading to survival of the tumor cells. Thus, inhibition of DNA repair pathways may increase the effect of chemotherapeutic drugs. Proliferating cell nuclear antigen (PCNA) functions as a binding platform for many proteins and has an essential role in co-ordination of DNA replication, DNA repair and other crucial processes for cell survival. In 2009, Gilljam and co-workers identified the peptide sequence AlkB homologue 2 PCNA-interacting motif (APIM), important for binding of proteins to PCNA. APIM is found in proteins involved in epigenetics, genome maintenance and cell cycle control, many of which are important after DNA damage. Studies have shown that APIM peptides sensitize cells to DNA damaging agents; hence APIM has a potential in cancer therapy. One hypothesis is that overexpressed APIM blocks the binding sites on PCNA and impairs the binding of APIM-containing proteins to PCNA, thus prevent optimal response to DNA damage.

Isolation of proteins on nascent DNA (iPOND) is a newly developed method for analysis of proteins involved in replication-related processes. The method relies on incorporation of the thymidine analog 5-ethynyl-2'-deoxyuridine (EdU) to nascent DNA. Crosslinking of proteins to DNA and biotin-conjugation to EdU results in biotin-tagged fragments of nascent DNA with bound proteins. The DNA-protein complexes are purified by exploiting the strong binding of biotin to streptavidin, before the proteins are eluted and analyzed by Western blotting.

The aim of this Master thesis has been to optimize iPOND to function with the Flp-IN^{TM} T$^-$Rex^{TM}-293 APIM-YFP cell line and for the detection of APIM-containing proteins with low abundance close to replication forks. Furthermore, the purpose of this study was to analyze how close proteins involved in epigenetics and DNA repair are to the replisome, and to evaluate if overexpression of APIM affects the presence of these proteins on nascent DNA, both before and after inducing DNA damage.

During optimization of iPOND, it was found that 3×10^7 cells/dish gives optimal EdU-incorporation and that 2×10^8 cells/sample is necessary for detection of proteins with low abundance close to replication forks. To maintain the proliferation rate, the cells need to be passaged the day before adding tetracycline, at a concentration of 0,02 µg/mL to induce and sustain APIM-expression. Finally, it was found that 0.5-1 mM methyl methanesulphonate (MMS) introduces DNA damage without excessive stalling of the replication machinery.

iPOND detected proteins involved in nucleotide excision repair (NER) (XPA and XPF)) and direct repair (hABH2) at newly replicated DNA, suggesting a function of these proteins in post-replicative repair. iPOND also verified the chromatin remodeling factors UHRF1 and hSNF5 as replisome proteins and trimethylated H3K9 (H3K9me3) and acetylated H4K16 (H4K16ac) as chromatin-bound proteins. Furthermore, a slightly reduced presence of XPA, hABH2 and hSNF5 (APIM-containing proteins) and of H3K9me3 and H4K16ac on nascent DNA was observed in MMS-treated APIM-expressing cells compared to cells not expressing APIM. The APIM-containing proteins EHMT1 and MRG15 are found in protein complexes that participate in trimethylation of H3K9 and acetylation of H4K16, respectively. Thus, overexpressed APIM seems to perturb the binding of APIM-containing proteins to nascent DNA, and to affect the function of APIM-containing protein complexes responsible for certain histone modifications.

Sammendrag

De fleste cellegiftbehandlinger induser alvorlige DNA skader for å drepe kreftcellene. DNA repareringsspor kan imidlertid reparere de introduserte skadene og føre til at kreftcellene overlever. Inhibering av DNA repareringsspor kan derfor øke effekten av kjemoterapi. «Proliferating nuclear cell antigen» (PCNA) fungerer som en bindingsplattform for mange proteiner og har en sentral rolle i koordinering av DNA replikasjon, DNA reparasjon og andre viktige prosesser for cellers overlevelse. I 2009 identifiserte Gilljam og medarbeidere peptidsekvensen «AlkB homologue 2 PCNA interacting-motif» (APIM) som er viktig for binding av proteiner til PCNA. APIM er funnet i proteiner involvert i epigenetikk, opprettholdelse av genomet og cellesykluskontroll, og mange av disse er viktige etter DNA skade. Studier har vist at APIM-peptider gjør cellene sensitive mot DNA skadende agenter, og dermed har APIM et potensiale i kreftterapi. En hypotese er at overuttrykt APIM blokker bindingssetene på PCNA og hemmer bindingen mellom APIM-innholdende proteiner og PCNA, og dermed hindrer optimal respons på DNA skade.

«Isolation of proteins on nascent DNA» (iPOND) er en nyutviklet metode for å analysere proteiner involvert i replikasjonsrelaterte prosesser. Metoden tar utgangspunkt i inkorporering av tymidinanalogen «5-ethynyl-2'-deoxyuridine» (EdU) til nyreplikert DNA. Kryssbinding av proteiner til DNA og biotin-konjugering til EdU gir biotin-merkede fragmenter av nyreplikert DNA med bundne proteiner. DNA-protein kompleksene renses ved å utnytte at biotin binder sterkt til streptavidin, før proteinene elueres og analyseres ved western-blotting.

Målet med denne masteroppgaven har vært å optimalisere iPOND i forhold til Flp-IN^{TM} T$^-$Rex^{TM}-293 APIM-YFP cellelinjen og til å kunne detektere APIM-innholdende proteiner som er tilstede ved små mengder nær replikasjonsgafler. Videre var det ønskelig å analysere hvor nære proteiner involvert i epigenetikk og DNA reparasjon er replisomet, og evaluere om overuttrykk av APIM påvirker tilstedeværelsen av slike proteiner på nylig syntetisert DNA, både før og etter DNA skade er indusert.

Under optimaliseringen av iPOND ble det funnet at 3×10^7 celler/skål gir optimal EdU-inkorporering og at 2×10^8 celler/prøve er nødvending for å detektere proteiner som kun er tilstede ved små mengder nær replikasjonsgafler. For å opprettholde celledelingshastigheten må cellene splittes dagen før tilsats av tetrasyklin, med en konsentrasjon på 0,02 µg/mL for å indusere og opprettholde APIM-uttrykkelse. Til slutt ble det funnet at 0,5-1 mM «methyl methanesulphonate» (MMS) introduserer DNA skade uten å stanse replikasjonsgaffelen betraktelig.

iPOND detekterte proteiner involvert i 'nucleotide excision repair' (NER) (XPA og XPF) og direkte reparasjon (hABH2) på nyreplikert DNA, som indikerer en funksjon i post-replikativ reparasjon. iPOND bekreftet også at kromatin remodeleringsfaktorene UHRF1 og hSNF5 er replisomproteiner og at trimetylert H3K9 (H3K9me3) og acetylert H4K16 (H4K16ac) er kromatinbundne proteiner. Det ble observert en noe redusert tilstedeværelse av XPA, hABH2 og hSNF5 (APIM-innholdende proteiner) og av H3K9me3 og H4K16ac på nyreplikert DNA i MMS-behandlede celler som uttrykte APIM i forhold til celler som ikke uttrykte APIM. De APIM-innholdende proteinene EHMT1 og MRG15 er funnet i proteinkomplekser som bidrar til henholdsvis trimetylering av H3K9 og acetylering av H4K16. Overuttrykt APIM ser derfor ut til å kunne forstyrre bindingen av APIM-innholdende proteiner til nylig syntetisert DNA, og påvirke APIM-innholdende proteinkomplekser som bidrar til visse histon modifikasjoner.

Abbreviations

A	Adenine
A	Alanine
ac	Acetylated
APIM	AlkB homologue 2 PCNA-interacting motif
ATP	Adenosine triphosphate
BER	Base excision repair
C	Cytosine
CFP	Cyan fluorescent protein
CPD	Cyclobutane pyrimidine dimers
D	Aspartate
DNA	Deoxyribonucleic acid
DNMT1	DNA methyltransferase 1
DSB	Double-strand breaks
EdU	5-ethynyl-2'-deoxyuridine
EHMT1	Euchromatic histone methylase 1
E.coli	*Escherichia coli*
F	Phenylalanine
FBS	Fetal bovine serum
G	Guanine
H	Histone
hABH2	Human AlkB homologue 2
HAT	Histon acetyl transferase
HDAC	Histone deacetylase complex
HEK293	Human embryonic kidney 293
HKDM	Histone lysine demetylase
HKMT	Histone lysine methyl transferase
hSNF5	Human sucrose non fermentable 5
I	Isoleucine
iPOND	Isolation of proteins on nascent DNA
K	Lysine
L	Leucine
LIG	DNA ligase
M	Methionine
me3	Trimethylated

meA	Methyladenine
meC	Methylcytosine
MDM2	Murine double minute 2
MMS	Methyl methanesulphonate
MRG15	MORF-related gene on chromosome 15
NER	Nucleotide excision repair
O_2	Oxygen
p53	Protein 53
PCNA	Proliferating cell nuclear antigen
PIP	PCNA-interacting peptide
Pol	DNA polymerase
PPs	Pyrimidine-pyrimidine photoproducts
PTM	Posttranslational modification
PVDF	Polyvinylidene fluoride
Q	Glutamine
R	Arginine
S-phase	Synthesis-phase
SAM	S-adenosyl-L-methionine
SENP2	SUMO specific protease 2
SNF	Sucrose non fermentable
SUV39H1	Suppressor of variegation 39H1
SWI	Switch
T	Thymine
TFII-I	Transcription factor II-I
TFIIS-L	Transcription factor II S-Like
TLS	Translesion synthesis
Topo	Topoisomerase
UHRF1	Ubiquitin-like PHD RING finger 1
UV	Ultraviolet
V	Valine
W	Tryptophan
XPA-G	Xeroderma pigmentosum group A-G
Y	Tyrosine
YFP	Yellow fluorescent protein

Table of Contents

1 Introduction

1.1 DNA Replication

The ability to reproduce itself is a fundamental property of all living organisms. Proliferating cells duplicate their DNA in a process called DNA replication that must be strictly regulated and happen accurately in the synthesis-phase (S-phase) of the cell cycle. Each of the two DNA strands serve as a templates for the formation of new DNA strands with bases complementary to the template; adenine (A) complements with thymine (T) and guanine (G) with cytosine (C) (Becker et al., 2009), as shown in Figure 1.1.
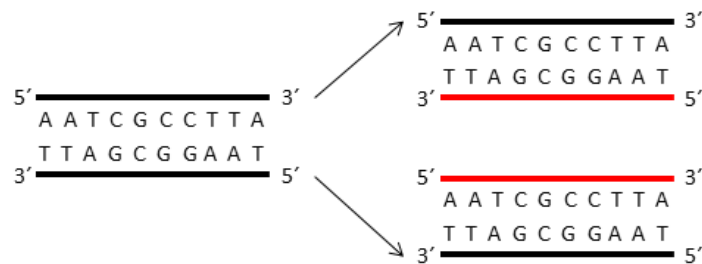
Figure 1.1 Each of the original DNA strands (black) serve as templates for the formation of new DNA strands (red) with bases complementary to the template during DNA replication.

The replication fork is an actively replicating region that moves along the parental DNA (Alberts et al., 2008). The DNA double helix is unwound and opened up in front of the replication fork to allow efficient duplication of the genetic material. The DNA polymerases (pol) δ and Ɛ are mainly responsible for DNA synthesis, and through co-operation with other proteins in the replication machinery, an error-free and efficient replication process is ensured (Burgers, 2009).

Even though DNA replication in eukaryotes can proceed as fast as 2,900 nucleotides per minute, only about one mistake occurs per 100 million nucleotides copied (Moldovan et al., 2007). The high fidelity during replication is partly due to the proofreading activity of the DNA polymerases itself, which strongly reduces the risk of misincorporation of bases. Despite of the high accuracy of the replication machinery, DNA repair systems are important both before and after replication to maintain

genome stability and prevent mutations causing diseases such as cancer. DNA replication is coupled both to DNA repair and to the assembly of DNA into chromatin, thus proteins at the replication fork are important for the fidelity of DNA replication, the co-ordination with progression in the cell cycle and for the inheritance of chromatin complexes (Waga and Stillman, 1998).

1.2 Proliferating Cell Nuclear Antigen

Proliferating cell nuclear antigen (PCNA) was first identified as an antigen found only in the nucleus of dividing cells, and later as an indispensable factor in DNA replication (Warbrick, 2000). PCNA is made up by three similar monomers that are joined firmly into a ring (Moldovan et al., 2007). The overall charge of PCNA is negative, but the inner surface is positively charged due to many arginine and lysine residues. This allows PCNA to encircle the negatively charged DNA without electrostatic repulsions and slide freely in both directions (Moldovan et al., 2007). The most known and characterized function of PCNA is as a sliding clamp for Pol δ and Pol Ɛ to ensure efficient DNA replication. PCNA has therefore become a marker for cell proliferation (Naryzhny, 2008). It is suggested that PCNA in mammalian cells is present as a double homotrimer formed by a back-to-back complex that works as both a moving platform and a binding station (Naryzhny, 2008). PCNA can thus be involved in DNA synthesis and other processes simultaneously, and has rightfully been called 'the maestro of the replication fork' (Moldovan et al., 2007).

Several proteins involved in DNA replication, DNA repair, cell cycle control, transcription, chromatin assembly and chromatin remodeling are found to interact with PCNA. As PCNA is a homotrimer, it can in theory bind three proteins simultaneously, but it is still believed that the binding of proteins to PCNA happens in a competitive manner (Naryzhny, 2008). Several proteins interact with PCNA through the PCNA-interacting peptide (PIP) box with the amino acid sequence Q-X-X-(L/M/I)-X-X-(F/Y)-/F/Y) identified by (Warbrick et al., 1997). Another PCNA-interacting motif has been identified, namely the AlkB homologue 2 PCNA-interacting motif (APIM) (Gilljam et al., 2009). Preliminary studies indicate that the PIP-box and the APIM sequence share the same binding site on PCNA. However, pull-down with a PIP-box and an APIM peptide suggests that posttranslational modifications (PTMs) on PCNA

regulate the interaction with and affinity towards PIP- or APIM-containing proteins (unpublished), as illustrated in Figure 1.2. Moreover, proteins containing the PIP-box are involved in DNA replication and are so called "housekeeping" proteins, while proteins containing the APIM motif are involved in DNA repair, apoptosis, epigenetics and other processes important after cellular stress. APIM is discussed more thoroughly in section 1.4.
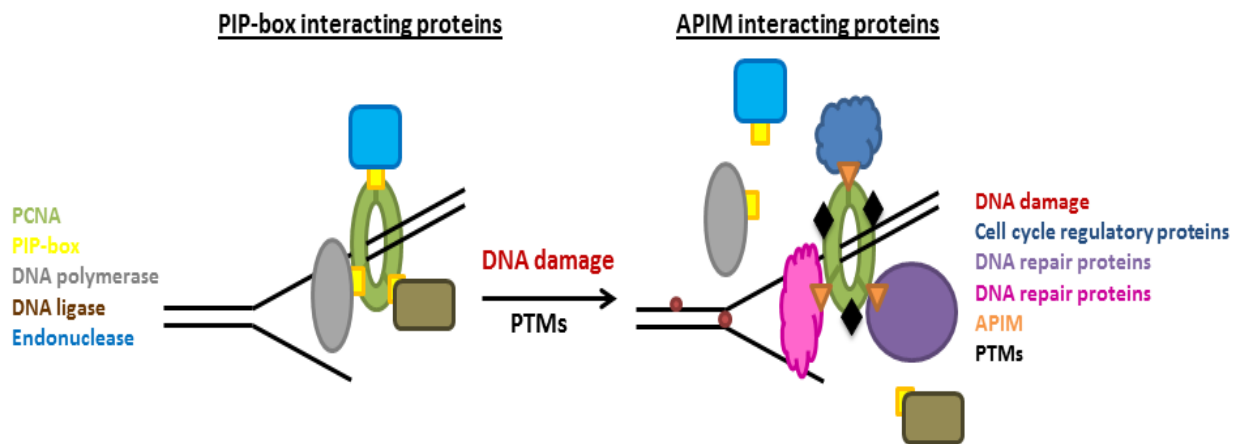
Figure 1.2 Hypothesized relationships between PIP- (yellow) and APIM-containing (orange) proteins and their interaction with PCNA (green). Proteins with the PIP-box, like DNA polymerase (grey), DNA ligase (brown) and endonucleases (blue) are involved in DNA replication. Proteins with APIM as the interacting motif possibly bind to PCNA after DNA damage (red) and PTMs (black) on PCNA. APIM-containing proteins include cell cycle regulatory proteins (dark blue) and DNA repair proteins (purple and pink), among others.

1.3 DNA Repair

Every day, thousands of lesions arise in the DNA caused by reactive metabolites and exposure to chemicals in the environment, among others. It is important that the cells can handle these lesions to ensure genetic stability and survival. Mammalian cells have conserved DNA damage sensor mechanisms to respond to such threats. Initiation of DNA repair pathways that removes the lesions and tolerance of damage are two such possible responses (Madhusudan and Middleton, 2005). As the range of lesions are wide, the cells have several DNA repair pathways; direct repair, base excision repair (BER) and nucleotide excision repair (NER) are a few examples (Houtgraaf et al., 2006). The repair processes, both before and after DNA replication, correct the mistakes and ensure that most DNA lesions and spontaneous changes

are temporary and do not manifest as mutations. DNA damage in front of the replication fork may block the replication progression if not repaired, and bypassed damage must be repaired after replication (Moldovan et al., 2007). Furthermore, the chromatin is in an open state during replication, allowing access for DNA repair proteins at the damage site (Luijsterburg and Van Attikum, 2011).

1.3.1 Nucleotide Excision Repair

One of the major and most versatile DNA repair pathways is NER. NER removes large DNA adducts or base modifications that cause distortions in the DNA double helix, and uses the undamaged strand as template for complete repair (Helleday et al., 2008). Perhaps the most known feature of NER is in the repair of ultraviolet (UV)-induced lesions. UV light may result in cyclobutane pyrimidine dimers (CPDs) and (6-4) pyrimidine-pyrimidine photoproducts ((6-4) PPs) (Nouspikel, 2009). The UV-induced DNA lesions repaired by NER are shown in Figure 1.3.
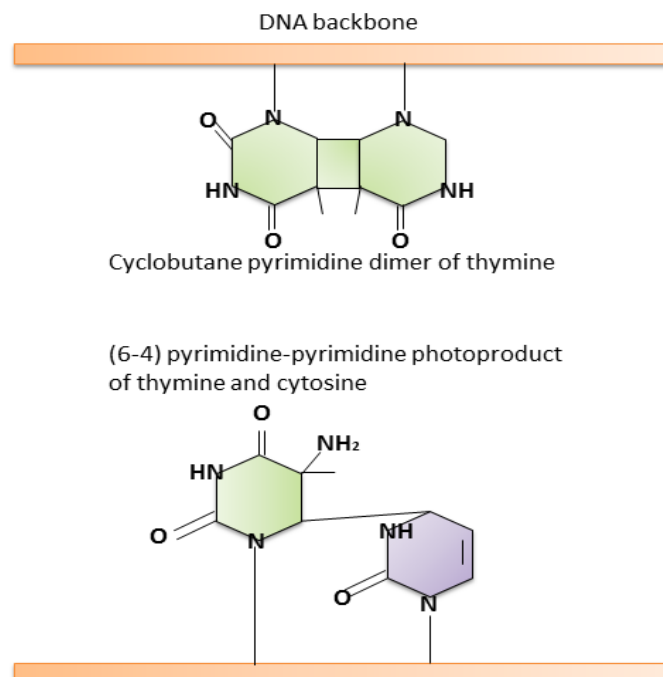
Figure 1.3 UV-induced DNA damage recognized by NER. **Top:** Cyclobutane pyrimidine dimer of thymine (green). **Bottom:** (6-4) pyrimidine-pyrimidine photoproduct of thymine and cytosine (purple).

Once the DNA damage has been recognized and NER activated, the DNA is denaturated to form a bubble around the lesion by a helicase (Nouspikel, 2009). The phosphodiester backbone of the abnormal strand is cleaved on both sides of the lesion by xeroderma pigmentosum group G and F (XPG and XPF), resulting in the excision of approximately 24-30 base pairs (Helleday et al., 2008, Nouspikel, 2009). The resulting gap is afterwards filled by DNA polymerases and sealed by DNA ligases, both involving PCNA. A simplified mechanism of NER is shown in Figure 1.4.
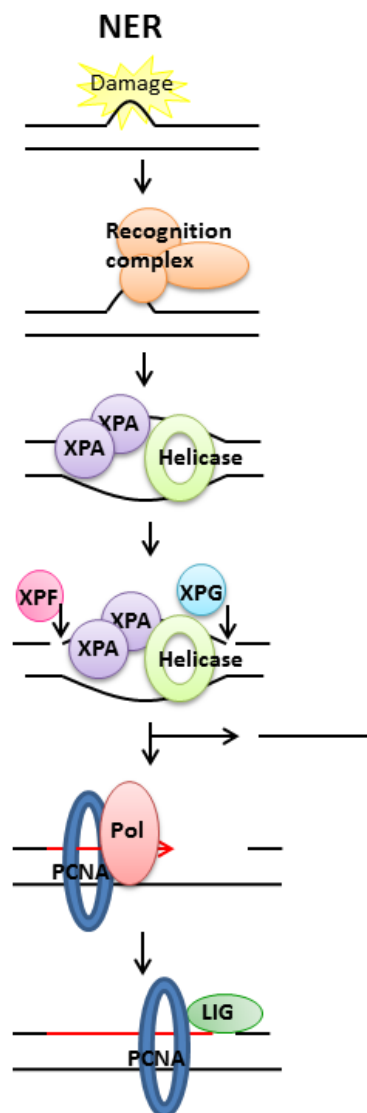
Figure 1.4 Simplified mechanism of NER. DNA damage (yellow) is recognized by a recognition complex (orange). The DNA is unwound by a helicase (light green) to form a bubble around the lesion. The XPA complex (purple) is recruited and the endonuclease activity of XPF (pink) and XPG (light blue) excises the DNA on each side of the lesion. DNA polymerases (Pol, red) fill the gap and the remaining nick is sealed by DNA ligases (LIG, green), both involving PCNA (blue).

Xeroderma pigmentosum group A (XPA) is believed to be recruited to the damage site after the lesion is recognized. XPA is part of the core preincision complex of NER and is associated with several other NER proteins (Köberle et al., 2006), including itself to form dimers (Yang et al., 2002). The role of XPA has not yet been fully identified. However, it is clear that XPA is absolutely necessary for NER to occur (Köberle et al., 2006). It has been suggested that the XPA complex may be involved in verification of the lesion or in identification of the damaged DNA strand, which is important to ensure that the correct strand is excised (Nouspikel, 2009). XPA is one of the many proteins that contain the APIM motif (K-F-I-V-K (in humans)). XPA has been found to interact with PCNA in replication foci via its APIM sequence, although a connection between replication and NER has not previously been reported (Gilljam et al., 2012, submitted).

1.3.2 Direct Repair

Direct repair is one mechanism for reversal of alkylation lesions. In contrast to other DNA repair pathways, direct repair does neither acquire excision of damaged nucleotides nor partly resynthesis of DNA (Hakem, 2008). The toxic lesions 1-methyladenine (1meA) and 3-methylcytosine (3meC) can be directly reverted by the oxidative demethylase human AlkB homologue 2 (hABH2) (Duncan et al., 2002). In the proposed mechanism, shown in Figure 1.5, hABH2 consumes oxygen (O_2) to demethylate 1meA and 3meC, and regenerates adenine and cytosine while releasing the methyl group as free formaldehyde (Sedgwick et al., 2007). hABH2 has been found to co-localize with PCNA in replication foci, suggesting that it performs direct repair in the proximity of the replication fork (Aas et al., 2003). Furthermore, hABH2 was the first protein where APIM (R-F-L-V-K (in humans)) was functionally verified as a PCNA-interacting motif (Gilljam et al., 2009).
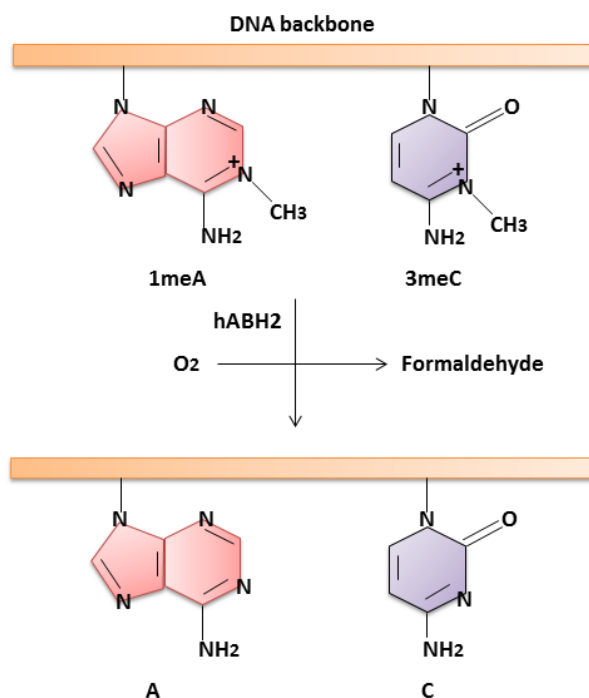
Figure 1.5 Simplified mechanism of direct reversal of alkylation damage by hABH2. hABH2 consumes oxygen to release the methyl group of 1meA (red) and 3meC (purple) as formaldehyde and regenerates adenine (A) and cytosine (C).

1.3.3 Translesion Synthesis

DNA damage that is not repaired before the cell replicates represents obstacles for the replication machinery and may lead to stalled or collapsed replication forks (Moldovan et al., 2007). Double-strand breaks (DSBs) and chromosomal rearrangements are possible outcomes of prolonged stalling, which might further result in cell cycle arrest or cell death. To avoid this potential threat, the cells have evolved a mechanism called translesion synthesis (TLS), that enhances further replication by bypassing the lesions (Moldovan et al., 2007). Normally, Pol δ and Pol ε carries out the DNA synthesis. These polymerases are switched to specialized polymerases that are able to bypass different types of damaged bases, but they are often more error prone (Helleday et al., 2008). A simplified mechanism of TLS is shown in Figure 1.6.

Currently fifteen different polymerases have been identified in mammals, where at least seven of these have TLS activity; Pol ζ (zeta), REV1, Pol η (eta), Pol ι (iota), Pol κ (kappa), Pol ν (nu) and Pol θ (theta) (Lange et al., 2011). It is believed that TLS occurs in a two-polymerase reaction, where the first polymerase acts to insert

nucleotides opposite the lesion, while the second polymerase functions to extend the polymerization. The second polymerase is reported to often be Pol ζ (Hendel et al., 2011). Pol η has an essential role in bypassing UV-induced CPDs with high efficiency and fidelity, as it most often inserts the correct A bases opposite thymine-thymine-CPDs (Lehmann, 2006). Alkylation damage in DNA poses a potential block for the replication machinery. In *Escherichia coli* (*E.coli*) it has been found that 1meA can be bypassed with low mutagenicity (Falnes et al., 2007) by PolV (Nieminuszczy et al., 2009) that most often inserts the correct T base above the lesion. Notably, the bypassed lesions are not removed and need to be repaired post-replicatively.
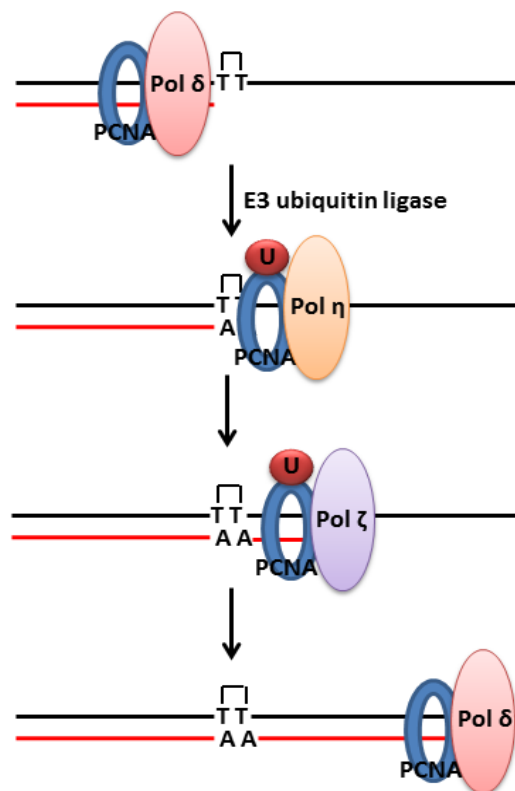
Figure 1.6 Proposed and simplified mechanism of translesion synthesis. The replication fork stalls when it reaches DNA damage, such as a thymine-thymine (TT)-dimer. The ubiquitylation (red) of PCNA (blue) may activate translesion synthesis polymerases, such as Pol η (orange) that inserts nucleotides opposite the lesion and Pol ζ (purple) that extends the polymerization before the normal replicative polymerases (light red) take over again (Lehmann, 2006).

1.4 Epigenetics

1.4.1 Chromatin Structure

The human genome is packaged into 23 pairs of chromosomes. Each of the chromosomes is made up of a single, long and linear DNA molecule that, together with proteins, is folded and packed into a more compact structure. The complex of DNA and its associated proteins is called chromatin. Histones are one class of proteins that bind to and participate in the folding of DNA. The histones have an overall positive charge that allows them to interact with DNA without electrostatic repulsion (Alberts et al., 2008). The primary level of chromatin packaging forms the nucleosome. The nucleosome looks like beads on a string, with the DNA as the string and nucleosome core particles as the beads. These particles consist of double stranded DNA wound 1.7 times around a protein core. The protein core consists of an octamer of histones, two of each of the histones H2A, H2B, H3 and H4 (Groth et al., 2007). The DNA is further condensed and packed until the mature chromosomes are formed. Chromatin is often divided into two types; heterochromatin and euchromatin. Heterochromatin is a highly condensed form of the chromatin associated with gene silencing, while euchromatin is less compact and associated with gene expression (Alberts et al., 2008). The organization of DNA is crucial to maintain genome stability, but needs to be dynamic to control processes such as DNA repair, replication and gene expression.

1.4.2 Epigenetics and Chromatin Replication

The structure of chromatin with the right DNA patterns and histone codes needs to be carefully duplicated during the cell cycle. This is called epigenetic inheritance and is essential to maintain gene expression patterns (McNairn and Gilbert, 2003). Epigenetics thus defines heritable changes in gene expression that are not encoded by the DNA sequence itself (Egger et al., 2004).

DNA replication is a crucial process for epigenetic inheritance. In order for replication to take place, the chromatin needs to be disrupted in front of the replication fork, which involves enormous alterations of the chromatin structure (Alabert and Groth, 2012). The accessibility to DNA and the disruption of the nucleosomes are probably

facilitated by chromatin remodeling complexes (Groth et al., 2007). The switch/sucrose non fermentable (SWI/SNF) chromatin remodeling complex uses energy from ATP to disrupt nucleosomes. The complex is composed of at least nine proteins, but only four are required for nucleosome remodeling; an ATPase and the three common core subunits (Roberts and Orkin, 2004). The complex is known to be involved in transcriptional activation of gene expression, but has also been implicated to have a role in DNA replication (Flanagan and Peterson, 1999, Lee et al., 1999). Human sucrose non fermentable 5 (hSNF5), one of the common core subunits, has been found to be required for efficient replication, suggesting that hSNF5 alone or as part of the complex has a function in DNA replication (Lee et al., 1999). Furthermore, hSNF5 has been found to co-localize and interact with XPC, which is a part of the recognition complex in NER, suggesting that the SWI/SNF complex has a function in increasing DNA accessibility for NER proteins (Ray et al., 2009). hSNF5 is among the proteins where APIM is found, with the sequence K-F-A-L-K (in humans), but it is not yet functionally verified.

Behind the replication fork the chromatin needs to be correctly restored. New histones are therefore synthesized and added to the daughter strand with the same modifications as the old histones, to maintain the epigenetic code and the density of the nucleosome (Alabert and Groth, 2012).

1.4.3 Histone Modifications and DNA Methylation

Each of the core histones has an N-terminal amino acid tail extending out of the DNA-histone core. The histone tails are subjected to different covalent modifications such as acetylation, methylation, ubiquitination and phosphorylation that alter the higher order chromatin structure. These types of modifications can thus form domains in the genome that affects the accessibility to transcription, repair and replication (McNairn and Gilbert, 2003).

Histone acetylation:

Histone acetyl transferases (HATs) use acetyl-CoA as a substrate to transfer an acetyl group to lysine residues on histone tails. The acetylation is reversible by the action of histone deacetylase complexes (HDACs). Acetylation of histones leads to a looser chromatin structure. It is believed that this is partly due to the neutralization of the positive charge on lysine when an acetyl group is added, which leads to a reduced interaction between DNA and histones in the nucleosomes (Bannister and Kouzarides, 2011). Hyperacetylation of the core histone tails is therefore a hallmark of transcriptionally active chromatin (McNairn and Gilbert, 2003). Acetylation of H4K16 (H4K16ac) directly unfolds the chromatin structure and has been reported to be important in the regulation of DNA repair (Luijsterburg and Van Attikum, 2011), transcriptional activation and the maintenance of euchromatin (Shogren-Knaak et al., 2006). The MORF-related gene on chromosome 15 (MRG15)-dependent acetyltransferase complex, among others, is involved in the acetylation of H4K16 (Wu et al., 2011). Several proteins in this complex contains the APIM motif, including MRG15 (K-Y-L-A-K (in humans)), although it has not yet been functionally verified.

Histone methylation:

Lysines on the tail of histones are also subjected to methylation. Methylation of histones is both related to active and silenced gene expression, dependent on which lysine residues that are methylated. Furthermore, each methylated lysine can exist in a mono-, di- or trimethylated state. Histone lysine methyl transferases (HKMTs) catalyze the transfer of one, two or three methyl groups from S-adenosyl-L-methionine (SAM) to lysine residues (Bannister and Kouzarides, 2011). The enzymatic activity of histone lysine demethylases (HKDMs) can reverse the methylation (Zhang et al., 2012). Euchromatic histone methyltransferase 1 (EHMT1) participates in mono- and dimethylation of H3K9 in euchromatin (Chen et al., 2010). EHMT1 is one of the proteins where APIM has been found (K-Y-L-I-K (in humans)), but not yet functionally verified. Suppressor of variegation 39H1 (SUV39H1) is reported to be the enzyme responsible for trimethylation of H3K9 (H3K9me3), either directly (Rice et al., 2003) or by further modifying mono- and dimethylation (Chen et al., 2010). H3K9me3 is a hallmark of heterocromatin, and its presence inhibits DNA repair proteins to gain access to the damage site (Luijsterburg and Van Attikum,

2011). Protein 53 (p53) is a tumor suppressor important in control and regulation of the cell cycle. Upon DNA damage, stress signals lead to the activation and stabilization of p53 that furthermore causes cell arrest until the damage is repaired or induced apoptosis if the damage is too severe to be repaired. Murine double minute 2 (MDM2) is a key regulator of p53. MDM2 can bind to p53 and its ubiquitin ligase activity prompts degradation of p53 (Chen et al., 2010). The ubiquitin ligase activity of MDM2 is stimulated by SUMO specific protease 2 (SENP2) (Chiu et al., 2008). SENP2 is an APIM-containing protein with the sequence R-W-L-V-R in humans, which is not yet functionally verified. When p53 is needed, DNA damage and other signals that induce p53 activation inhibit the p53 degradation by MDM2. Interestingly, MDM2 has been found to interact with both SUV39H1 and EHMT1 and to promote the formation of a p53-MDM2-SUV39H1/EHMT1 complex. The complex has been found to participate in the regulation of p53 by methylate H3K9 at p53 target promotors, leading to repressed p53 mediated transcription, hence inhibited p53 activity (Chen et al., 2010).

DNA methylation:

DNA methylation is another epigenetic mark, and generally serves to repress transcription, thus it is recognized as an epigenetic silencing mechanism. The activity of DNA methyltransferase 1 (DNMT1) is responsible for the correct inheritance of the DNA methylation patterns during DNA replication and repair (McNairn and Gilbert, 2003). Ubiquitin-like PHD RING Finger 1 (UHRF1) can interact with DNMT1, and it is believed that the affinity of UHRF1 towards hemi-methylated DNA helps DNMT1 to be recruited to the right place at the right time, though the mechanism is not clear (Bronner et al., 2010). Furthermore, it is proposed that UHRF1 and DNMT1 slide along with the replication fork by interactions with PCNA (Cedar and Bergman, 2009). APIM is found in UHRF1 with the sequence R-Y-L-L-R in humans, but it is not yet functionally verified.

1.5 APIM and Cancer Therapy

APIM is a PCNA-interacting motif found in several proteins involved in epigenetics, genome maintenance and cell cycle control (Gilljam et al., 2009). The motif and its functionality were first identified in hABH2. Through examination of the five conserved N-amino acids of hABH2, the APIM peptide was defined as [KR]-[FYW]-[LIVA]-[LIVA]-[KR] (Gilljam et al., 2009). All together more than 200 proteins have been found to contain the conserved APIM sequence, all though its functionality in most of these proteins has not yet been verified (Gilljam et al., 2009). At this point, APIM has been proven functional in six proteins; hABH2, Transcription factor II S-Like (TFIIS-L), Transcription factor II-I (TFII-I), DNA topoisomerase (Topo) II α, RAD51B (Gilljam et al., 2009), and recently XPA (Gilljam et al., 2012, submitted).

Most chemotherapeutic treatments rely on induction of severe DNA damage to the cells. As the attempt to replicate damaged DNA can increase cellular death and a known feature of cancer cells is their rapid proliferation compared to normal cells, the DNA damaging drugs mainly affect the cancer cells. DNA repair pathways can repair the lesions induced by chemotherapeutic agents, leading to survival of the tumor cells. The effect of chemotherapeutic drugs may therefore be increased by inhibitors of DNA repair pathways, especially since some cancer cells rely on a reduced number of DNA repair pathways to survive (Helleday et al., 2008). PCNA has an essential role in the co-ordination of events during DNA repair, DNA replication and other processes that are crucial for the survival of the cell. Many proteins involved in these processes bind to PCNA via their APIM sequence, thus peptides containing APIM can disrupt protein-PCNA interactions, hence improve the treatment of cancer (Gilljam et al., 2009). The hypothesis explaining how APIM can contribute to reduced tolerance to chemotherapy is illustrated for in Figure 1.7.
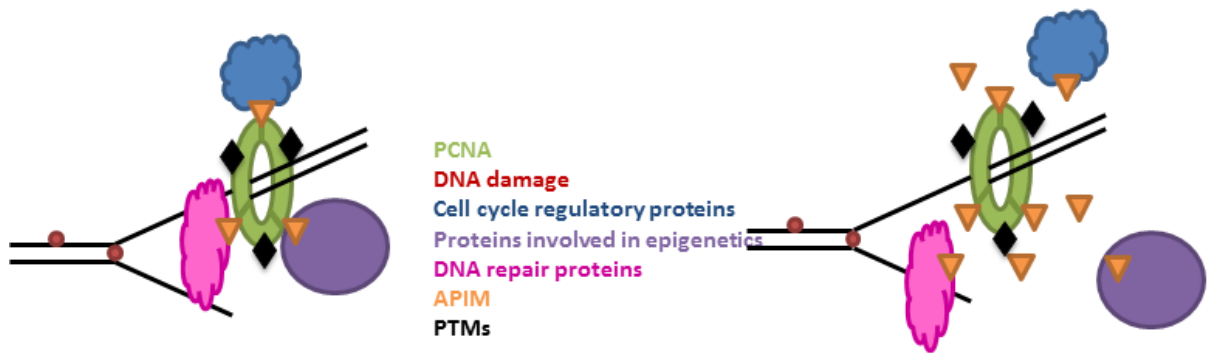
Figure 1.7 Left: After DNA damage (red), DNA repair proteins (pink), proteins involved in epigenetics (purple) and cell cycle regulatory proteins (blue) are recruited to PCNA (green); some of them through APIM (orange). **Right:** Overexpression of APIM may block the binding site on PCNA and impair the binding of APIM-containing proteins to PCNA, thus prevent optimal function of the proteins.

1.6 iPOND

Isolation of proteins on nascent DNA (iPOND) is a newly developed method for analysis of proteins present close to active or damaged replication forks in cultured mammalian cells. The method relies on incorporation of the thymidine analog 5-ethynyl-2'-deoxyuridine (EdU), see Figure 1.8, to label fragments of newly replicated DNA. After incubating cells with EdU for a short period of time, the proteins are fixed to DNA by formaldehyde, which also serves to stop replication (Sirbu et al., 2012).
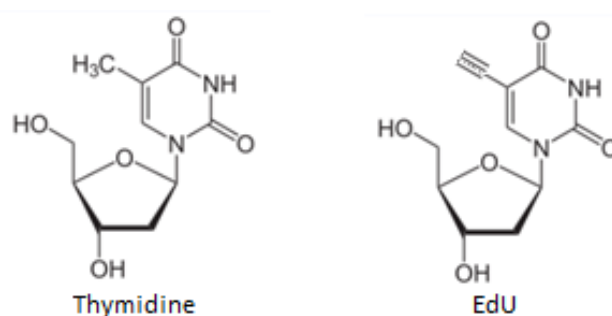
Figure 1.8 Left: Thymidine **Right:** The thymidine analog 5-ethynyl-2'-deoxyuridine (EdU) (Sirbu et al., 2012).

Click chemistry is an approach to assemble new and pure molecular entities rapidly and selectively in a reliable matter (Kolb et al., 2001). The alkyne group of EdU makes it possible to perform a copper-catalyzed click reaction to bind biotin azide and EdU covalently, resulting in biotin-tagged fragments of nascent DNA. Subsequently, DNA fragmentation is accomplished by cell lysis and sonication. Purified EdU-labeled DNA-protein complexes are achieved by exploiting the strong binding of biotin to streptavidin beads, before the proteins are eluted and analyzed by Western blotting (Sirbu et al., 2011). A simplified outline of the iPOND procedure is shown in Figure 1.9.
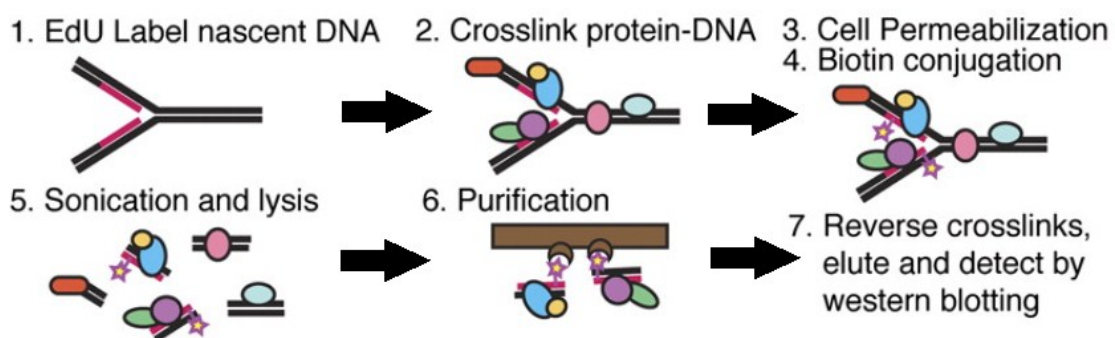
Figure 1.9 General outline of the iPOND procedure (Sirbu et al., 2011). The cells are cultured with EdU to label nascent DNA. Subsequently, the cells are treated with formaldehyde to crosslink DNA-protein complexes. Copper catalyzes a click reaction to bind biotin covalently to DNA, before the cells are lysed by sonication. The DNA-protein complexes are purified by exploiting the strong binding of biotin to streptavidin. Proteins are then eluted and analyzed by Western blotting.

iPOND has been successfully applied to identify proteins at active replisomes and to detect protein recruitment and post-translational modifications after introducing DNA damage (Sirbu et al., 2011). These applications are done with a pulse-chase experiment, as illustrated in Figure 1.10. For pulse samples, cells are added EdU for a short period of time to label nascent DNA. In a chase sample, the EdU-pulse is followed by a thymidine-chase, where EdU is replaced with thymidine for a longer period of time to move the EdU-labeled DNA segment away from the replication fork. After either the pulse or the chase period, the cells are fixed and harvested. If a protein is a true replisome protein, it should be identified only in pulse samples and not in chase samples, while proteins bound to the chromatin may be found in both

types of samples. An incubation time of 2.5 minutes with EdU is suggested to be sufficient for capturing proteins at the replisome, but longer incubations are necessary to isolate newly deposited chromatin factors (Sirbu et al., 2011).
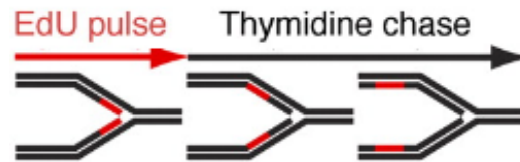
Figure 1.10 EdU-pulse and thymidine-chase (Sirbu et al., 2011). EdU is incorporated into the DNA during a pulse period, while EdU is replaced with thymidine for a longer period of time in the chase period to move the labeled fragment away from the replication fork.

1.7 Aim of Study

The aim of this Master thesis has been to optimize iPOND to function with the Flp-IN^{TM} T$^-$Rex^{TM}-293 APIM-YFP cell line used in the experiments and for the detection of proteins with low abundance close to replication forks. The purpose of using iPOND was to analyze APIM-containing proteins or consequence of APIM-containing proteins involved in epigenetics and DNA repair at various distances away from the fork, and to evaluate whether overexpression of APIM could perturb their presence on nascent DNA.

2 Materials and Methods

2.1 Cell Line and Culture

The Flp-IN^{TM} T$^-$Rex^{TM}-293 APIM-YFP cell line (Invitrogen) derived from human embryonic kidney 293 (HEK293) cells was used in the experimental work. The cell line was stably transfected with an APIM gene fused to yellow fluorescent protein (YFP) under the control of a tetracycline-inducible promotor. The APIM peptide expressed has the amino acid sequence M-D-R-W-L-V-K, the same as in hABH2 (amino acid 1-7) but with F replaced by W, as this has been shown to bind stronger to PCNA and result in the best effect in cell experiments. A simplified presentation of the cell line is shown in Figure 2.1.
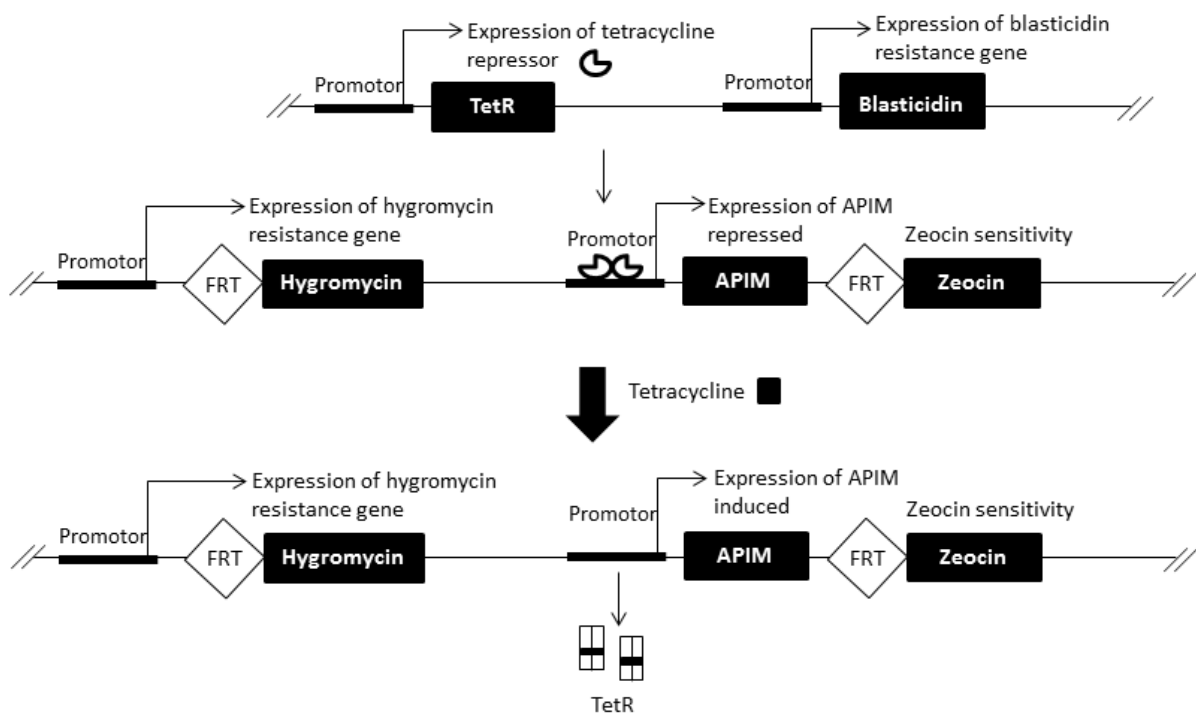
Figure 2.1 Simplified presentation of the Flp-IN^{TM} T$^-$Rex^{TM}-293 APIM-YFP cell line used in the experimental work. The cells exhibit blasticidin and hygromycin resistance and zeocin sensitivity. The FRT sites were the sites of recombination in order to stably transfect the cells. Expression of the APIM gene is repressed by the Tet repressor, but inducible upon tetracycline addition.

The cells were maintained in Dulbecco's modified Eagle's medium (Sigma-Aldrich) supplemented with fetal bovine serum (FBS) (10%, Sigma-Aldrich), L-glutamine (2 mM, Sigma-Aldrich), gentamicin (0.1 mg/mL, Invitrogen), amphotericin B (2.5 µg/mL, Sigma-Aldrich), hygromycin B (50 µg/mL, Invitrogen) and blasticidin (15 µg/mL, Invitrogen). The medium used for cell expansion was supplemented with 20% FBS and was without hygromycin and blasticidin. The cells were cultivated at 37°C in a 5% carbon dioxide-humidified atmosphere.

The cells were split when they reached 80-90% confluence, either for general maintenance or for expansion prior to the experiments. The medium was then aspirated off and the cells were washed once with 10 mL PBS (1x, Oxoid) to remove excess medium and serum. 2 mL trypsin-EDTA (1x, Lonza) was added to detach the adherent cells, and the cells were incubated (5 min) before re-suspension in the appropriate medium. The cell suspension was either split to dishes (15 cm) for expansion, or kept in the flask (75 cm^2) for maintenance.

2.2   Tetracycline Concentration

To determine the most appropriate tetracycline concentrations, Flp-IN^{TM} T$^-$Rex^{TM}-293 APIM-YFP cells were split (1:3) to nine dishes (15 cm) when they were 90% confluent, and incubated overnight (37°C, 5% CO$_2$). Subsequently, the number of cells in one dish was counted in a Bürker chamber (day 0). The remaining eight dishes were added tetracycline at different concentrations according to previous experience by others. Two and two dishes were added 0.00 µg/mL, 0.01 µg/mL, 0.02 µg/mL and 0.03 µg/mL tetracycline. The cells were incubated for approximately 24 and 48 hours.

Day one; The number of cells in one dish of each tetracycline concentration was counted in Bürker chambers, and each dish split to a confocal dish by adding 1.5 mL of a 20 mL cell suspension together with 0.5 mL medium to further dilute the suspension. The confocal dishes were incubated for approximately 6 hours before viewed at in the confocal fluorescence microscope. Fluorescent images were acquired using a laser-scanning confocal fluorescence microscope (LSM 510 Meta, Carl Zeiss). The YFP-tag fused to APIM was excited at λ=514 nm and detected at

λ=530-600 nm. The process was repeated at day two, but with 0.5 mL of a 30 mL cell suspension added to the confocal dishes together with 1.5 mL medium and an incubation time of approximately 4.5 hours.

2.3 Transfection

One confocal dish of Flp-IN™ T⁻Rex™-293 APIM-YFP cells was transient transfected by adding a solution of 95 µL OptiMEM medium (GIBCO, Invitrogen), 1500 ng CFP-PCNA expression construct and 3 µL Xtreme Gene DNA Transfection Reagent (Version 05, Roche Diagnostics). The solution was mixed and incubated for 20 minutes before added to the cells. Tetracycline (0.02 µg/mL) was added to the cells to express APIM-YFP, and incubated overnight (37°C, 5% CO_2). The cells were examined in the laser-scanning confocal fluorescence microscope, with settings allowing detection of two color images. The YFP-tag of APIM was excited and detected as described in the previous section, and the cyan fluorescent protein (CFP)-tag fused to PCNA was excited at λ=458 nm and detected at λ=470-500 nm using consecutive scans.

2.4 iPOND

The iPOND procedure was performed based on the protocol described by Sirbu and coworkers (Sirbu et al., 2011, Sirbu et al., 2012). A flow sheet of the procedure is shown in Figure 2.2.

2.4.1 Sample Preparation and EdU Labeling

Sample preparation: Cell cultures were expanded to six dishes (15 cm) per sample two days before EdU incubation, in order to achieve $2x10^8$ cells per sample with a cell confluence of $3x10^7$ at the day of iPOND. One additional dish was included to count the number of cells. For experiments with APIM-expression, tetracycline (Sigma-Aldrich) was added to the medium (0.02 µg/mL) the day before performing iPOND. The media needed to pulse and chase the samples was equilibrated in the incubator (37°C, 5% CO_2) overnight to ensure proper temperature and pH during the experiment.
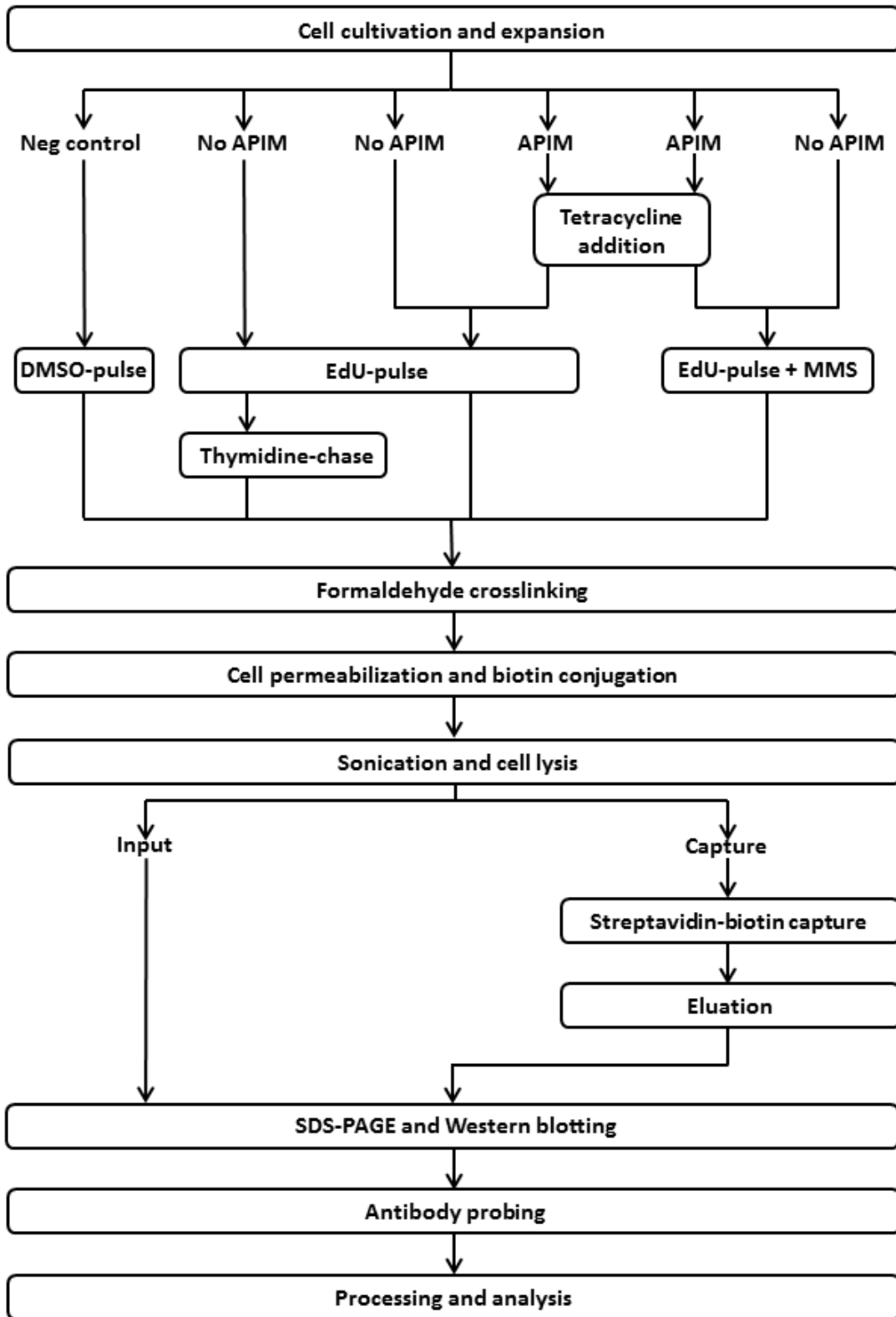
Figure 2.2 Flow sheet of the iPOND procedure.

Pulse: The medium was aspirated off each dish and replaced with 23 mL cell culture medium with EdU (10 µM, Invitrogen). The cells were then incubated for the desired pulse time to incorporate EdU. For experiments with DNA damage, methyl methanesulphonate (MMS) (0.5 and 1 mM, Sigma-Aldrich) was added to the EdU-medium.

Chase: For pulse-chase samples, the EdU-medium was discarded and the cells were washed once with 5 mL preheated cell culture medium containing thymidine (10 µM, Sigma-Aldrich). 20 mL of the thymidine-medium was added to each dish, and the cells were left in the incubator for the desired chase time.

Negative control: The medium was aspirated off each dish and replaced with 23 mL cell culture medium with DMSO (0.1%, Sigma-Aldrich). The negative control sample was incubated for the same length of time as the pulse sample with the longest incubation time.

Crosslinking: After performing pulse or pulse-chase, the media was decanted and the cells fixed in 10 mL formaldehyde in PBS (1%, Roth) for 20 minutes at room temperature. The crosslinking was quenched by adding 1 mL glycine (1.25 M, Merck). The samples were collected in centrifugation tubes (50 mL) by scraping with a cell lifter, and subsequently centrifuged (5 min, 900g, 4°C). The cell pellets were washed three times in cold PBS with the same volume as observed when collecting the samples. The samples were snap-frozen in liquid nitrogen and stored at -80°C overnight.

2.4.2 Click Reaction

Permeabilization: The cell pellets were resuspended in Triton-X-100 in PBS (0.25%, Sigma-Aldrich) at a concentration 1×10^7 cells/mL, and incubated for 30 minutes at room temperature while rotating. The samples were centrifuged (5 min, 900g, 4°C) and the pellets were washed once with cold BSA in PBS (0.5%, Sigma-Aldrich) and once with PBS.

Click reaction: The cell pellets were resuspended in a click reaction cocktail that contained 3.9 mL PBS, 0.05 mL biotin-azide (1mM, Invitrogen), 0.5 mL sodium ascorbate (100 mM, Sigma-Aldrich) and $CuSO_4$ (100 mM, Merck) per $1x10^8$ cells. The cell suspensions were rotated for 1-2 h at room temperature. The samples were centrifuged (5 min, 900g, 4°C) and washed once with cold BSA in PBS (0.5%) and once with PBS with the same volumes as used in the click reaction.

2.4.3 Cell Lysis

Cell lysis: The cell pellets were resuspended in lysis buffer (1% SDS (Sigma-Aldrich), 50 mM Tris (pH 8.0, Merck)) containing the protease inhibitors aprotinin (1 µg/mL, Sigma-Aldrich) and leupeptin (1 µg/mL, Sigma-Aldrich), at a concentration of $1.5x10^7$ cells per 100 µL of lysis buffer and transferred to eppendorf tubes (1.5 mL). The cell lysates were sonicated for 40 seconds using a microtip sonicator with 20% duty cycle and 2.5 in output control (Branson Sonifier 250, Branson Ultrasonics Corporation), followed by 40 seconds pause. The sonication pulse was repeated one time for every 200 µL of cell lysate. The samples were kept on ice slurry during sonication to avoid excessive heat generation. The samples were centrifuged (10 min, 16,100g, 18°C) and the supernatants filtered through an 80 micron nylon mesh into new tubes (15 mL). The filtrates were diluted with 1:1 (V/V) PBS containing aprotinin (1 µg/mL, Sigma-Aldrich) and leupeptin (1 µg/mL, Sigma-Aldrich).

Measuring protein concentration: Protein concentrations in the cell extracts were measured in a spectrophotometer (NanoDrop ND-1000 Spectrophotometer, Software V3.1.2, Life Science). The extracts were adjusted by calculating the volume needed to have the same protein content as the sample with the lowest, and replacing the remaining volume with cold PBS with protease inhibitors.

Input samples: 15 µL of the lysates were transferred to new tubes (1.5 mL) and added 1:1 (V/V) 2xSDS Laemmli sample buffer (2xSB) that contained 0.4 g SDS, 2 mL glycerol (100%, Merck), 1.25 mL Tris (1M, pH 6.8), 0.01 g bromophenol blue (Sigma-Aldrich) and dithiothreitol (DTT, 0.2 M, Sigma-Aldrich) in 8 mL H_2O. The input samples were stored at -80°C overnight.

2.4.4 Purification

Biotin capture: Streptavidin agarose beads (Novagen) were prepared at a concentration of 100 µl bead slurry (corresponds to 50µL packed beads) per $2x10^8$ cells to capture EdU-labeled and biotin-tagged DNA fragments with associated proteins. The bead slurry was centrifuged (1 min, 1,800g, 18°C) and the storage buffer was carefully aspirated off. The beads were washed twice with 1:1 (V/V) lysis buffer that contained aprotinin and leupeptin, and once in PBS, also containing the protease inhibitors. The beads were re-suspended in 1:1 (V/V) PBS with inhibitors and 100 µl was added to each sample. The samples were rotated overnight (16-20 h, 4°C).

Washing and elution of proteins: The samples were centrifuged (3 min, 1,800g, 18°C). The pellets, which consisted of beads with captured DNA and crosslinked proteins, were washed once with 1 mL cold lysis buffer, transferred to eppendorf tubes (1.5 mL), and once with 1 mL NaCl (1M, Merck) for each $2x10^8$ cells/100 µL biotin slurry. These two washing steps were repeated before washing twice with 1 mL cold lysis buffer. Between each washing step, the samples were rotated (5 min, 18°C) and centrifuged (1 min, 1,800g, 18°C). The supernatants were discarded after the last wash. The protein-DNA complexes isolated on the beads represented the capture samples. Captured proteins were eluted and crosslinks were reversed by adding 1:1 (V/V) 2xSB to the packed beads and incubating for 30 min at 95°C together with the input samples. The samples were centrifuged (1 min, 1,800g, 18°C) and the supernatants were kept for SDS-PAGE and Western blotting. The capture samples were concentrated (Concentrator 5301, Eppendorf) to a volume of 20 µL (≈15 minutes).

2.5   Western Blotting

The samples were loaded on a NuPAGE Novex 4-12% Bis-Tris gel (1.5 mm, 10 well, Invitrogen). The input samples were added 1 µL OmniCleave Endonnuclease (200 U/µL, Epicentre) 5 minutes prior to loading to reduce viscosity caused by DNA content. The MagicMark XP Western Protein Standard (20-220 kDa, Invitrogen) and SeeBlue Plus2 Pre-Stained Standard (4-250 kDa, Invitrogen) were used as standards for molecular weight. The electrophoresis was carried out in NuPAGE MOPS SDS running buffer (1x, Invitrogen) added 500 µL of NuPAGE antioxidant (Invitrogen) at 200 V as the limiting factor for approximately 1 hour.

After gel electrophoresis, the proteins were transferred from the gel onto a polyvinylidene fluoride (PVDF) membrane (Immobilon, Millipore) by blotting in NuPAGE transfer buffer (1x Invitrogen) with 10% (V/V) methanol. The transfer of proteins from gel to membrane was run for 2 hours at 30 V as the limiting factor. The membrane was blocked in 5% (w/V) low fat dry milk in PBST (PBS with 0.1% Tween 20)) for 1 hour. The primary and secondary antibodies were diluted in 5% low fat dry milk in PBST. The primary antibodies used are listed in Table 2.1. The secondary antibodies polyclonal swine anti-rabbit immunoglobulins/HRP (Dako) and polyclonal rabbit anti-mouse immunoglobulins/HRO (Dako) were used for polyclonal rabbit and monoclonal mouse primary antibodies respectively, diluted 1:5,000. The membrane was treated with a chemiluminescence reagent (SuperSignal West Femto Maximum Sensitivity Substrate, Thermo Scientific) and the proteins visualized in the KODAK Image Station (4000R, KODAK) and analyzed in the software program KODAK MI molecular imaging software (version 4.0, KODAK).

Table 2.1 Primary antibodies used to probe for proteins after iPOND.

Protein	Antibody	Clone	Dilution	Supplier
PCNA	SC-56	Mouse monoclonal	1:2,000	Santa Cruz Biotechnology
H3	AB1791	Rabbit-polyclonal	1:1,000	Abcam
XPA	AB65963	Mouse monoclonal	1:500	Abcam
XPF	AB17798-500	Mouse monoclonal	1:100	Abcam
hSNF5	HPA019127	Rabbit-polyclonal	1:250	Sigma-Aldrich
HIF3A	HPA041141	Rabbit-polyclonal	1:250	Sigma-Aldrich
hABH2	A8228	Mouse monoclonal	1:500	Sigma-Aldrich
UHRF1	AB57083	Mouse monoclonal	1:500	Abcam
H4K16ac	17-10101	Rabbit-polyclonal	1:500	Millipore
H3K9me3	AB8898	Rabbit-polyclonal	1:1000	Abcam

3 Results

All data from quantification presented in this chapter is given in the Appendix.

3.1 Optimization the Experimental Procedure

The iPOND protocol presented in chapter 2.4 is an optimized version, adjusted through experience to the Flp-IN^{TM} T$^-$Rex^{TM}-293 APIM-YFP cell line and for the detection of low abundant proteins close to replication forks.

3.1.1 APIM-Expression is Induced and Sustained by 0.02 µg Tetracycline/mL

In the pilot study where APIM-expression was induced, tetracycline (0.03 µg/mL) was added approximately four hours after passaging the cells. The tetracycline concentration was chosen according to experience by others, but resulted in lower proliferation rate compared to cells not added tetracycline. Tetracycline may be toxic to the cells, implicating that the concentration used were too high. Therefore, the effect of different tetracycline concentrations had to be determined. The most appropriate concentration should be the lowest concentration still inducing APIM-YFP. The cells were split to nine dishes and incubated overnight, instead of four hours as in the pilot study. At day 0, the number of cells in one dish was counted in a Bürker chamber. Two and two of the remaining dishes were added tetracycline (0.00-0.03 µg/mL) and incubated for 24 and 48 hours before being counted. The cell proliferation did not seem to be affected by the tetracycline doses tested, as shown in Figure 3.1.
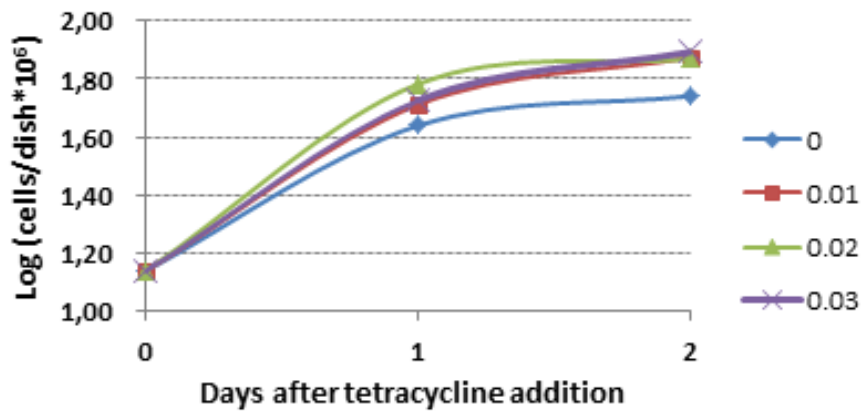
Figure 3.1 Growth curves of cells added different tetracycline concentrations; 0.00 µg/mL (blue), 0.01 µg/mL (red), 0.02 µg/mL (green) and 0.03 µg/mL (purple). The number of cells was counted in Bürker chambers.

The cell line is stably expressing APIM fused to YFP. The fluorescent light emitted by the YFP-tag was detected in a confocal fluorescence microscope. Thus, the expression of APIM can be seen directly with confocal fluorescence imaging, and give a quantitative measurement of protein expression. The confocal fluorescent images of APIM-YFP are shown in Figure 3.2. The first two and the last two rows show images from day one and day two after tetracycline addition, respectively. At day one, the cells seemed to express APIM-YFP equally. At day two, APIM-YFP was still expressed at all concentrations. The expression was possibly increasing slightly with increasing concentrations, but the difference was not pronounced. A tetracycline concentration of 0.02 µg/mL was chosen to induce APIM-expression to ensure both unaffected cell proliferation and sustained APIM-expression.

A negative control of cells not added tetracycline was also included (first column, Figure 3.2). Blasticidin prevents APIM-expression in the absence of tetracycline. It was therefore of interest to evaluate both cells grown in media for maintenance with blasticidin (first two images) and for expansion without blasticidin (last image). Ideally, no fluorescence should be detected from these cells. However, some leakage was observed from cells grown in both media. It should be emphasized that the detector gain in these images are higher than in the others to more clearly detect potential leakage. Despite of the leakage, the difference was evident between the negative control and the cells with induced APIM-expression.

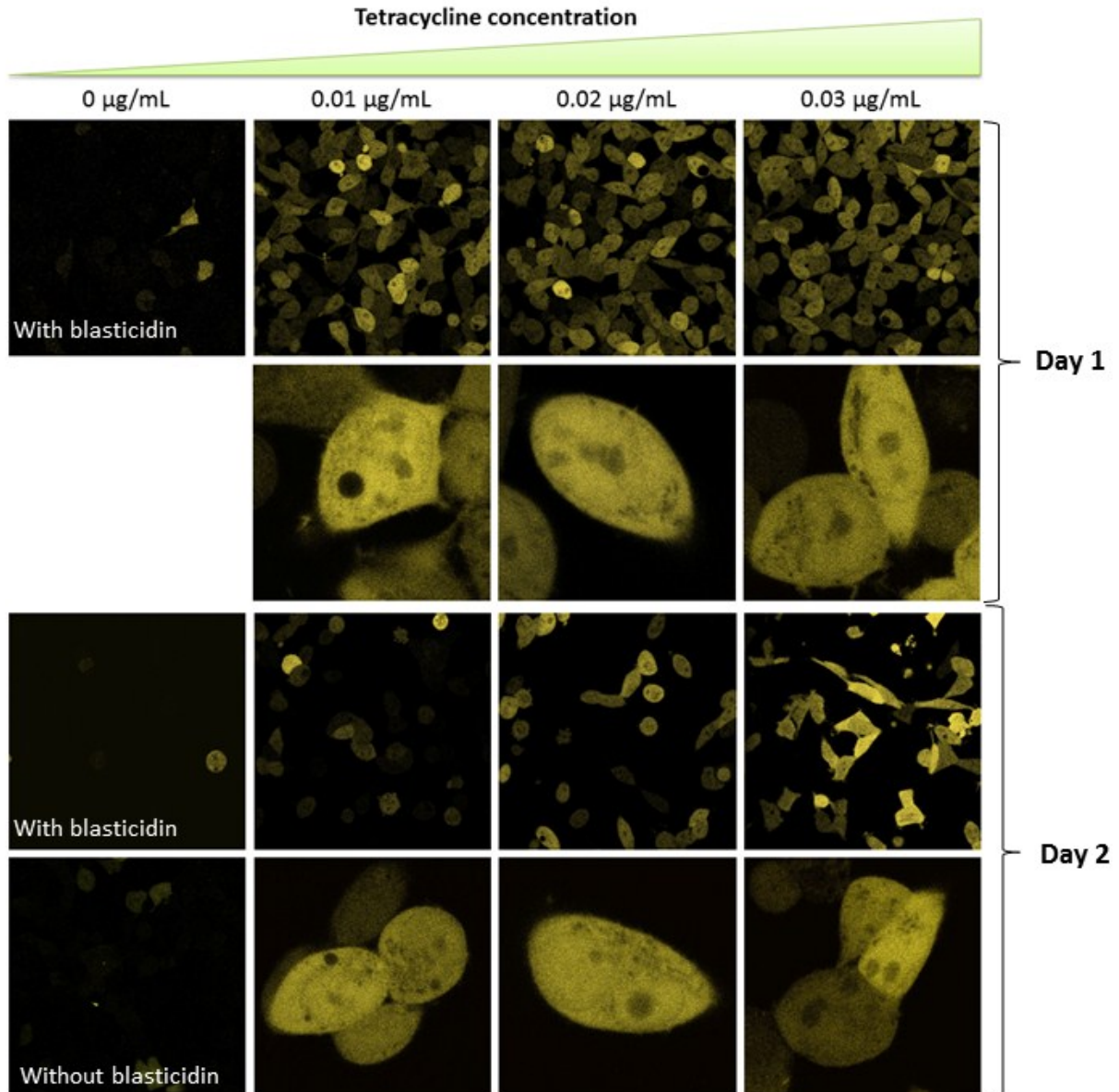**Figure 3.2** Expression of APIM-YFP induced by tetracycline at various concentrations. The first two and the last two rows show the expression at day one and two after tetracycline addition, respectively. Blasticidin should prevent APIM-expression in the absence of tetracycline. The first column shows the leakage of APIM-YFP from cells not added tetracycline, both for cells grown in media with and without blasticidin.

3.1.2 The APIM-YFP Construct Expressed Show Correct Intranuclear Localization

APIM co-localizes with PCNA in replication foci (Gilljam et al., 2009). To get a qualitative control of the cells, which stably expresses APIM-YFP, the cells were transient transfected with CFP-tagged PCNA. The co-localization of these proteins was visualized by confocal fluorescence imaging, as shown in Figure 3.3. The last column shows the merged images of APIM-YFP and CFP-PCNA, where a co-localization (shown as yellow dots) is observed at replication foci. The co-localization verifies that the cells express a correct construct.
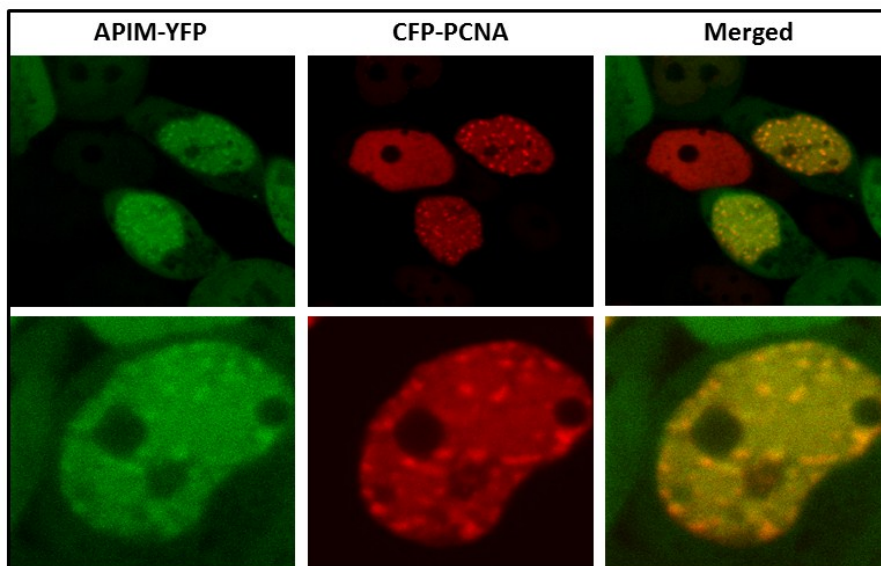
Figure 3.3 Confocal fluorescent images of cells stably expressing APIM-YFP and transient transfected with CFP-PCNA. The first and second columns show the localization of APIM-YFP and CFP-PCNA, respectively. When these images are merged, a co-localization (yellow dots) between APIM and PCNA is observed.

3.1.3 Cell Confluence of 3×10^7 Cells per Dish gives Optimal EdU-Incorporation

The cells must be in the log phase of growth to ensure optimal EdU-incorporation. While a cell confluence of $4\text{-}6 \times 10^7$ cells/dish is recommended for HEK293T cells (Sirbu et al., 2012), a cell confluence of 3×10^7 cells/dish gave the best results for the Flp-IN^{TM} T$^-$Rex^{TM}-293 APIM-YFP cell line.

3.1.4 2x10^8 Cells per Sample are needed to Detect Proteins with Low Abundance at the Replication Fork

After performing two experiments with 1x10^8 cells/sample, as suggested by (Sirbu et al., 2012), it became clear that 1x10^8 cells was sufficient for the detection of proteins highly abundant close to the replication fork, such as PCNA, but not for the detection of proteins with lower abundance, such as XPA. In these experiments, the signals of lower abundant proteins were not above the background signal. By increasing the number of cells to 2x10^8 cells/sample, the signal became higher than the background and XPA and other low abundant proteins could be detected.

3.1.5 Protein Detection is not an Artifact of Comprehensive Crosslinking

HIF3A is a transcription factor and should not be detected in the proximity of the replication fork. To examine whether the detection of proteins after iPOND were an artifact of too extensive crosslinking resulting in pull-down of non-replicative proteins, it was probed for HIF3A. HIF3A could be detected in input samples but could not be pulled down in iPOND captures, as shown in Figure 3.4. PCNA is here a positive control for replicative proteins, and H3 is a control for chromatin-bound proteins.
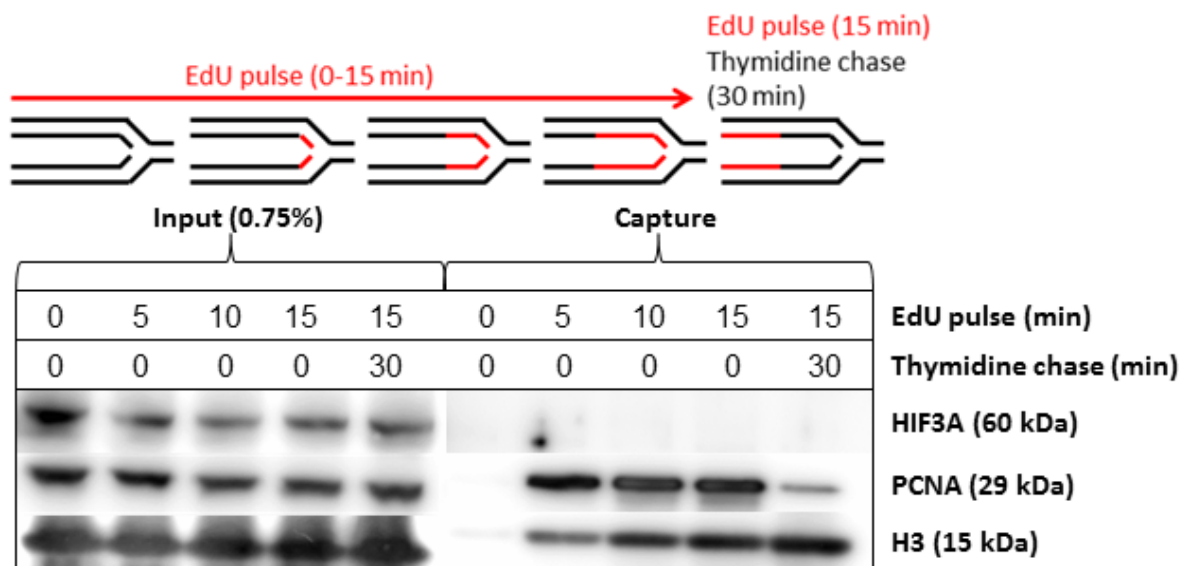

Figure 3.4 Pull-down of the non-replicative protein HIF3A after iPOND with different times of EdU-pulse (0-15 min). One additional sample was thymidine-chased (30 min) after the pulse (15 min). PCNA and H3 are here controls of replicative proteins and chromatin-bound proteins, respectively.

3.1.6 MMS Concentrations of 0.5-1 mM Introduce DNA damage without Excessive Stalling of the Replication Machinery

One hypothesis is that proteins interacting with PCNA through their APIM motif are recruited after DNA damage or other cellular stress that cause modifications of PCNA. It was therefore of interest to evaluate which MMS concentrations that would introduce DNA damage without causing a total block of replication. Different doses of MMS (0-4 mM) were added to the medium together with EdU, and the samples were pulsed (30 min). A pulse time of 30 minutes was chosen to take into account that damaged cells replicate slower and to ensure that DNA damage was introduced to the labeled fragments. In addition, one sample was not EdU-labeled and one sample was followed by a thymidine-chase (45 min) after being pulsed (30 min). The results are shown in Figure 3.5.

PCNA is an indicator of replication. More PCNA could be pulled down in the sample pulsed with 0.5 mM MMS than in the sample pulsed without MMS, which could be caused by more cells arrested in S-phase. Less PCNA could be pulled down after inducing 1 mM MMS compared to the sample pulsed without MMS. Higher MMS concentrations gave even lower PCNA detection, while not as low as the chase sample, suggesting reduced replication rate. hABH2 is known to repair damage introduced by MMS and to interact directly with PCNA in replication foci, indicating that it repairs damage close to the replication fork (Aas et al., 2003, Gilljam et al., 2009). hABH2 was found to be more concentrated in pulse samples compared to the chase sample, thus its presence near the replication machinery was verified, also in the absence of DNA damage. Furthermore, a pull-down of hABH2 seemed to result in a slightly higher protein detection in samples pulsed with 0.5-3 mM MMS compared to the sample pulsed without MMS, indicating hABH2 recruitment. Together these results indicate that 0.5-1 mM MMS can introduce DNA damage without reducing the replication rate excessively.
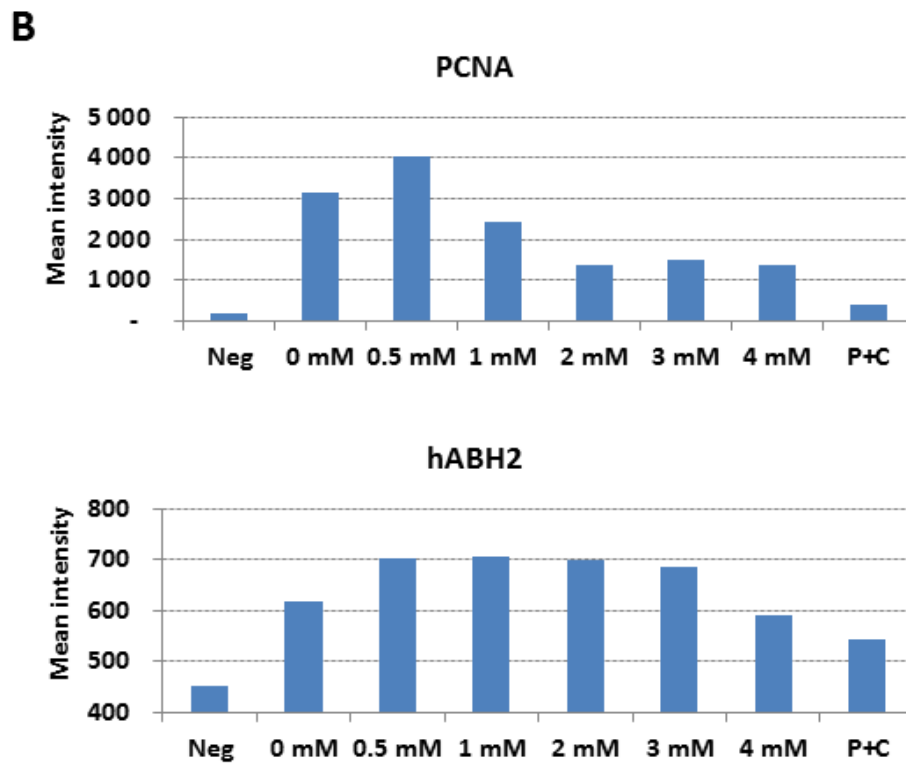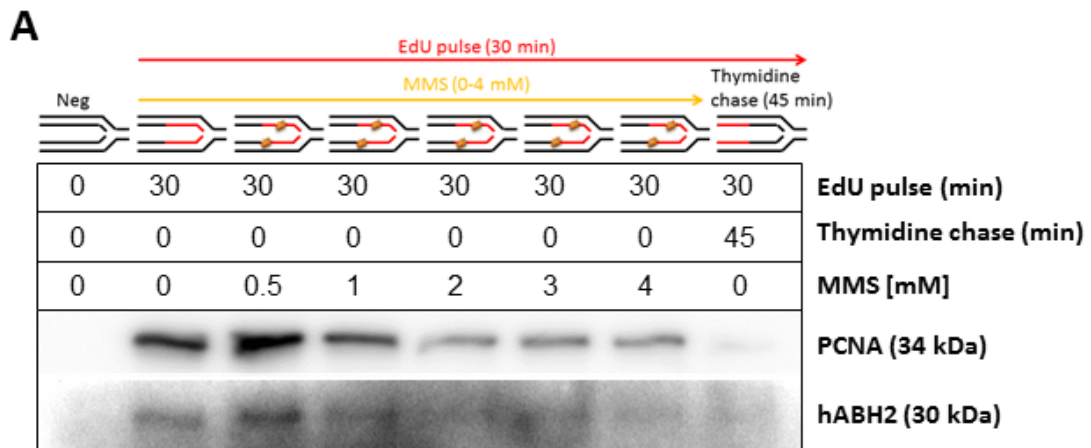
Figure 3.5 iPOND from cells pulsed with EdU (30 min) together with different concentrations of MMS (0-4 mM). In addition, one sample was not EdU-labeled (neg) and one sample was followed by a thymidine-chase (45 min) after the pulse (30 min) (P+C). **A)** Pull-down of PCNA and hABH2. **B)** Quantification of the bond intensities in A).

3.2 NER Proteins Travel with Active Replication Forks

XPA has recently been found to co-localize and interact with PCNA in replication foci, implying a connection between the replication machinery and NER (Gilljam et al., 2012, submitted). It was therefore of interest to further study how close XPA is to the replication fork by performing iPOND with different times of EdU-pulse (0-15 min). One additional sample was followed by a thymidine-chase (30 min) after the pulse (15 min). The results are shown in Figure 3.6. PCNA is here a positive control for replisome proteins. XPA and XPF could be detected in capture samples after only five minutes, as shown in Figure 3.6. Slightly increased protein detection was observed after 10 minutes, followed by a slight decrease after 15 minutes.
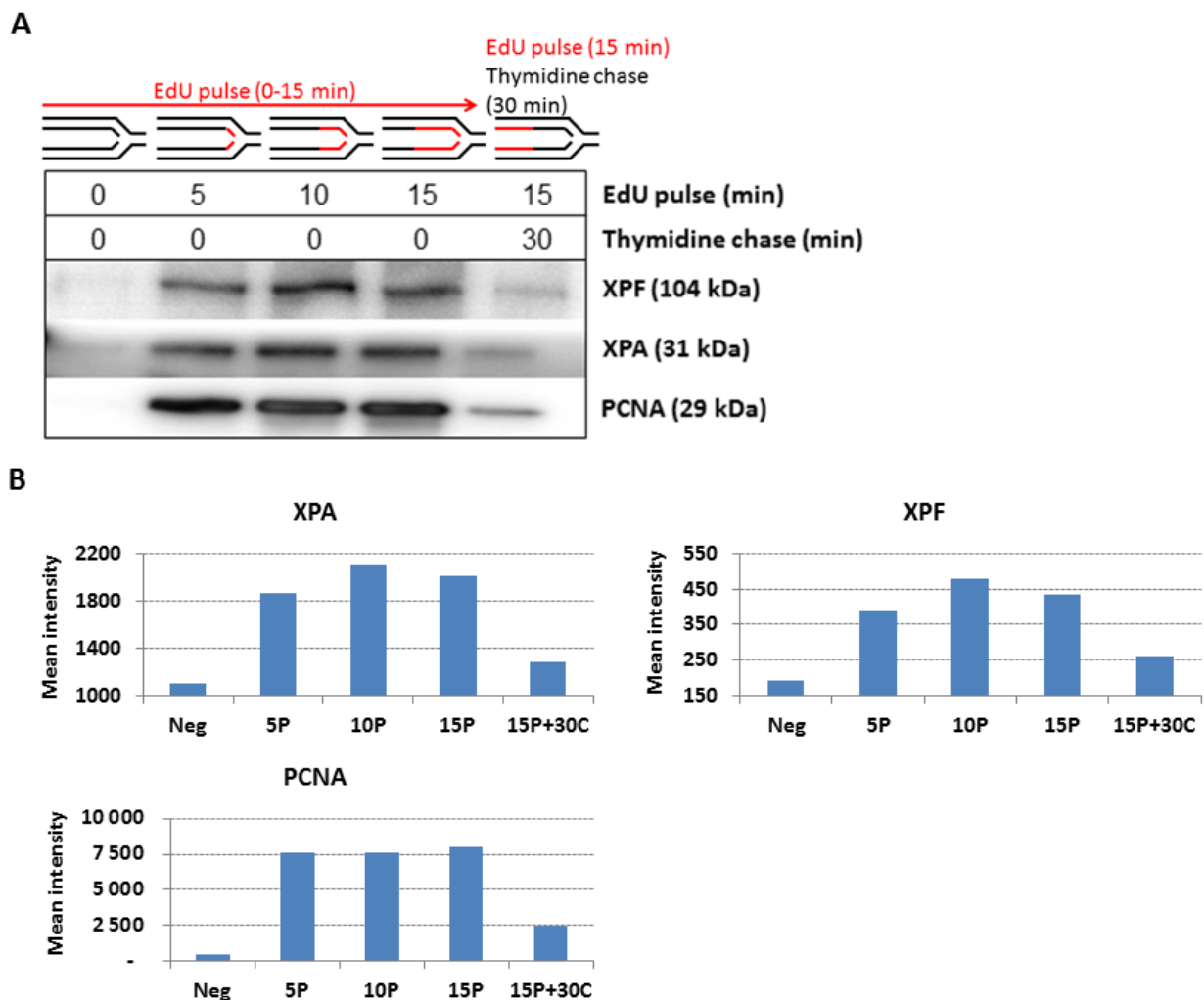
Figure 3.6 iPOND from cells pulsed with EdU (0-15 min). One additional sample was followed by a thymidine-chase (30 min) after the pulse (15 min) **A)** Pull-down of XPF, XPA and PCNA. PCNA bonds are the same as shown in Figure 3.4. **B)** Quantification of bond intensities in A). P=pulse, C=chase.

The detection of both XPA and XPF were in agreement with what was observed for the positive PCNA control, with higher pull-down in pulse samples than in the chase sample, which has been reproduced in other experiments. Consequently, the results indicate that XPA, XPF and probably other NER proteins are in the proximity of the replication fork in the absence of DNA damage.

3.3 iPOND Detects Epigenetic and Chromatin Remodeling Factors

hSNF5, as part of the SWI/SNF complex, and UHRF1 have both been reported to interact with PCNA and be involved during DNA replication (Cedar and Bergman, 2009, Euskirchen et al., 2011). Histones are modified to regulate the chromatin structure (Bannister and Kouzarides, 2011). H3K9me3 and H4K16ac are examples of such modified histones. Pulse-chase experiments were therefore performed to verify the presence of these proteins on nascent DNA with one sample pulsed with EdU and one sample thymidine-chased after the pulse. One additional sample was not EdU-labeled (negative control). Both UHRF1 and hSNF5 were confirmed to be replisome proteins by iPOND, as they were more concentrated in pulse than chase samples, as shown in Figure 3.7A. On the other hand, H3K9me3 and H4K16ac were detected as chromatin-bound proteins as they were more concentrated in chase than pulse samples, as shown in Figure 3.7B. PCNA is here a positive control for replisome proteins.
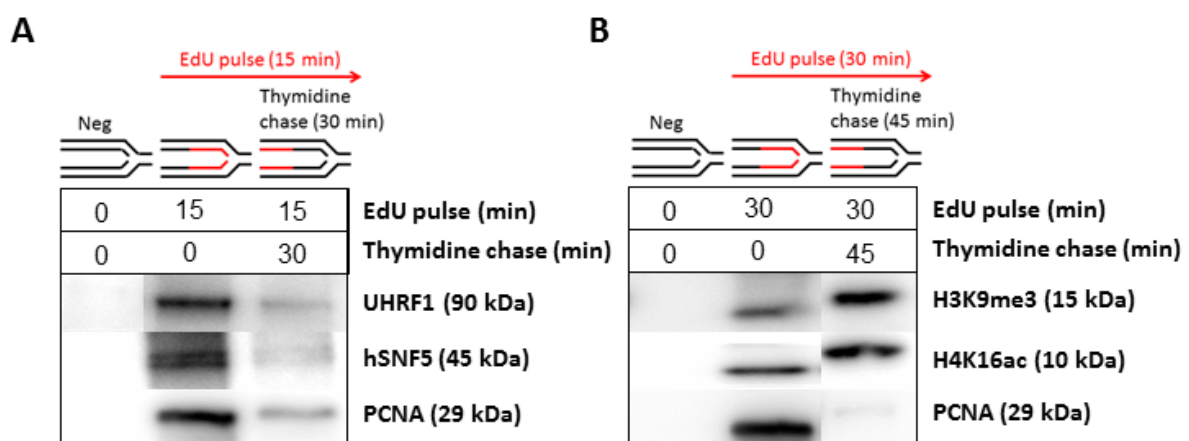
Figure 3.7 A) Capture of UHRF1, hSNF5 and PCNA from cells EdU-pulsed (15 min) or thymidine-chased (30 min) after the pulse (15 min). One additional sample was not EdU-labeled (neg). **B)** Capture of H3K9me3, H4K16ac and PCNA from cells EdU-pulsed (30 min) or thymidine-chased (45 min) after the pulse (30 min). One additional sample was not EdU-labeled (neg).

3.4 APIM-Expression Affects the Capture of Several APIM-Containing Proteins

APIM has been found in several DNA repair proteins and chromatin remodeling factors. It was therefore of interest to examine if iPOND could detect a perturbed protein presence on nascent DNA when overexpressing APIM, both of APIM-containing proteins and of histones modified by such proteins. It has been speculated that APIM-containing proteins bind more strongly to PCNA with PTMs caused by cellular stress. Therefore, it was also of interest to evaluate whether a potential effect of overexpressed APIM was more pronounced after DNA damage. Both cells that did and did not express APIM were EdU-pulsed (30 min) together with MMS (0-1 mM). In addition, one sample was not EdU-labeled (negative control) and one sample was followed by a thymidine-chase (45 min) after the pulse (30 min).

3.4.1 APIM Affects the Binding of XPA to Newly Synthesized DNA

Mutated APIM in XPA has recently been found to result in reduced NER efficiency and cell survival (Gilljam et al., 2012, submitted). Therefore, it was particularly interesting to examine whether overexpression of APIM could affect the binding of XPA to nascent DNA. Pull-down of XPA was slightly reduced in APIM-expressing cells treated with MMS compared to cells not expressing APIM, while no such reduction could be observed in untreated cells, as shown in Figure 3.8. However, a slightly decreased pull-down in untreated APIM-expressing cells has been observed previously (data not shown). XPF, which does not contain APIM, is here a negative control and was not pulled down in reduced amount in APIM-expressing cells.

Figure 3.8 iPOND from cells labeled with EdU (30 min) and MMS (0-1mM). In addition, one sample was not EdU-labeled (neg) and one sample was followed by a chase in thymidine-containing medium (45 min) (P+C). **A)** Pull-down of XPF, XPA and PCNA. PCNA bonds are the same as partly shown in Figure 3.7B. **B)** Quantification of bond intensities in A). Blue bars: Cells not expressing APIM (control). Red bars: Cells expressing APIM.

3.4.2 APIM Reduces hABH2 Binding to Nascent DNA

APIM has been functionally verified in hABH2 as a PCNA-interacting motif. Furthermore, overexpressed APIM has been found to make cells more sensitive towards DNA alkylation damage and to result in slower repair of 1meA generated by MMS (Gilljam et al., 2009). For these reasons, the effect of overexpressed APIM on binding of hABH2 to nascent DNA was explored. As shown in Figure 3.9, less hABH2 could be detected in all capture samples from APIM-expressing cells compared to cells not expressing APIM. PCNA was used as a positive control for replisome proteins, but did also show a reduced pull-down in APIM-expressing cells treated with 1 mM MMS.
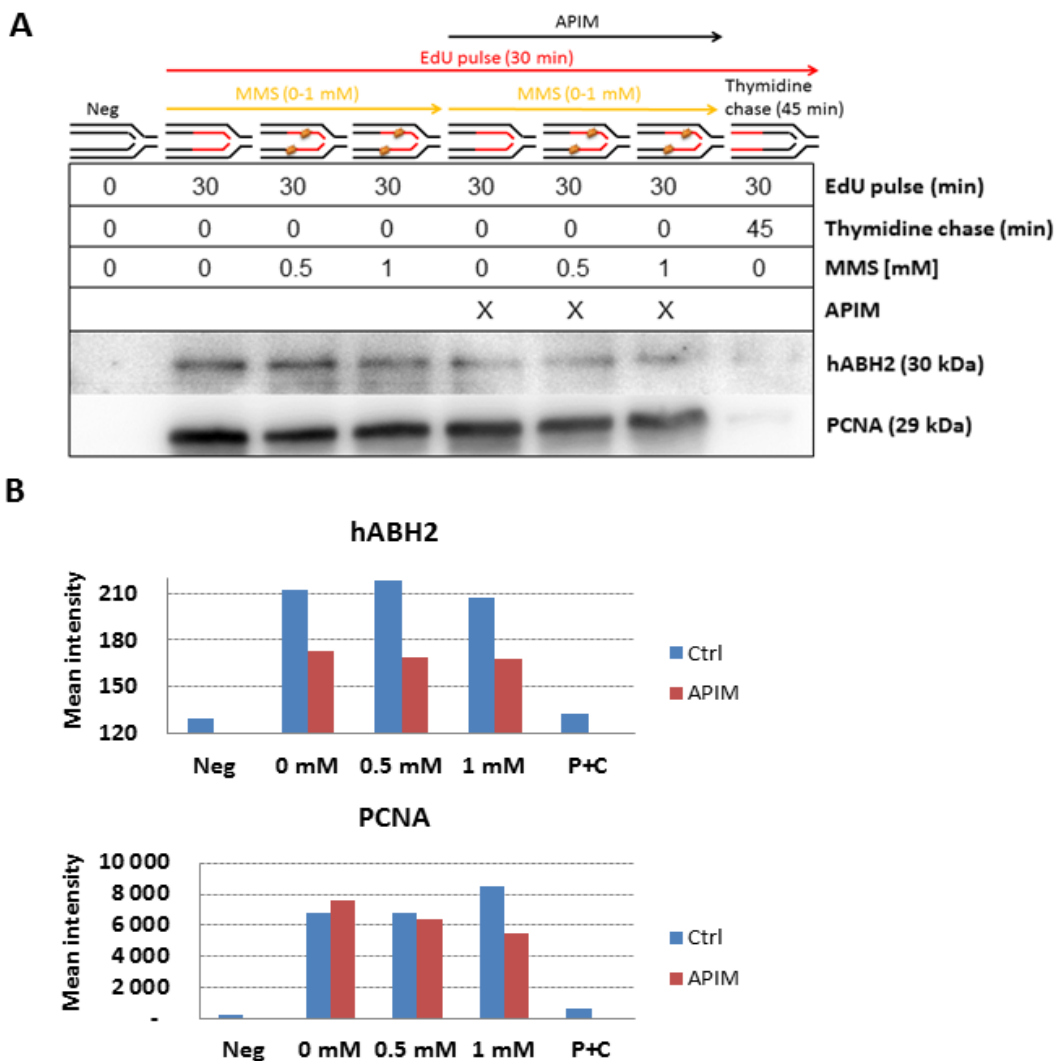
Figure 3.9 iPOND from cells labeled with EdU (30 min) and MMS (0-1mM). In addition, one sample was not EdU-labeled (neg) and one sample was followed by a chase in thymidine-containing medium (45 min) (P+C). **A)** Pull-down of hABH2 and PCNA. PCNA bonds are the same as shown in Figure 3.8. **B)** Quantification of bond intensities in A). Blue bars: Cells not expressing APIM (control). Red bars: Cells expressing APIM.

3.4.3 APIM Affects Chromatin Remodeling and Epigenetics

APIM has been found in UHRF1, hSNF5, EHMT1 and MRG15, but not functionally verified. EHMT1 is involved in both di- and trimethylation of H3K9 (Chen et al., 2010) and MRG15 is suggested to participate in the acetylation of H4K16 ((Wu et al., 2011). Therefore, it was of interest to investigate whether overexpressed APIM could affect the presence of the chromatin remodeling factors UHRF1 and hSNF5 and the modified histones H3K9me3 and H4K16ac on nascent DNA. The results from overexpressing APIM are shown in Figure 3.10. UHRF1 was pulled down in approximately equal amounts from both cells expressing and not expressing APIM. On the other hand, reduced amount of hSNF5 binding to nascent DNA in APIM-expressing cells was seen, and has also been observed in another experiment. The reduction was more pronounced in cells treated with MMS. The presence of both H3K9me3 and H4K16ac on newly replicated DNA seemed to be slightly reduced by the expression of APIM in MMS-treated cells.

	Neg								
EdU pulse (min)	0	30	30	30	30	30	30	30	
Thymidine chase (min)	0	0	0	0	0	0	0	45	
MMS [mM]	0	0	0.5	1	0	0.5	1	0	
APIM					X	X	X		

UHRF1 (90 kDa)
hSNF5 (45 kDa)
H3K9me3 (15 kDa)
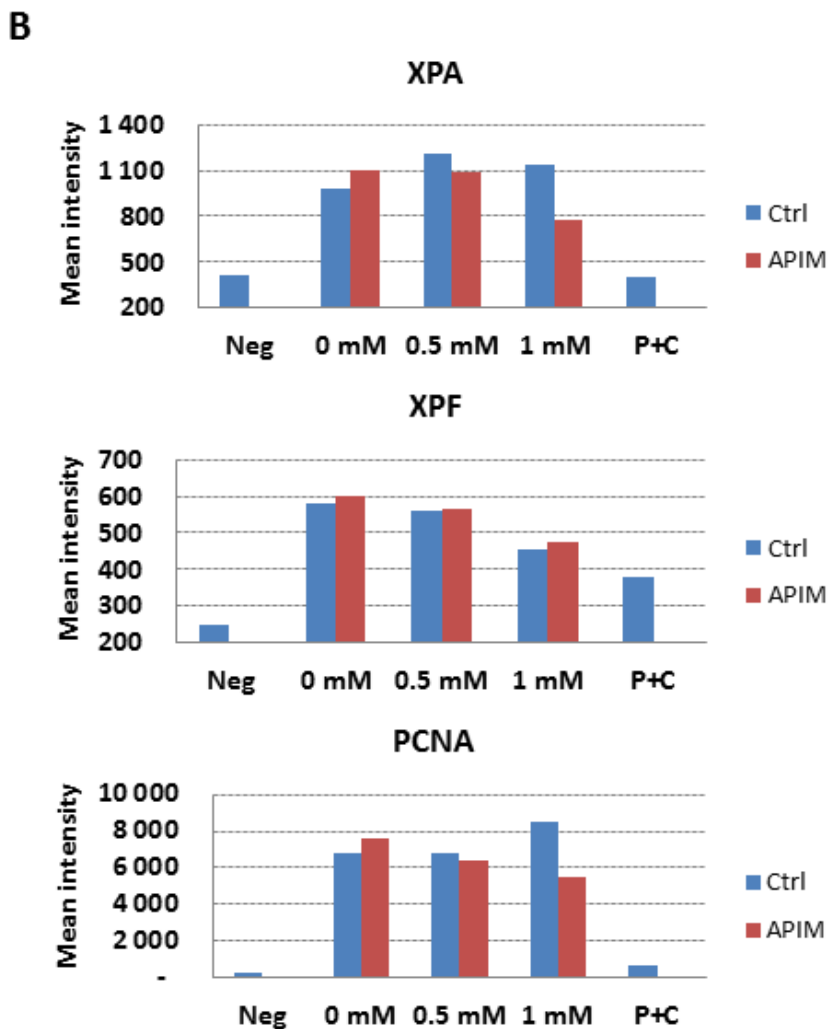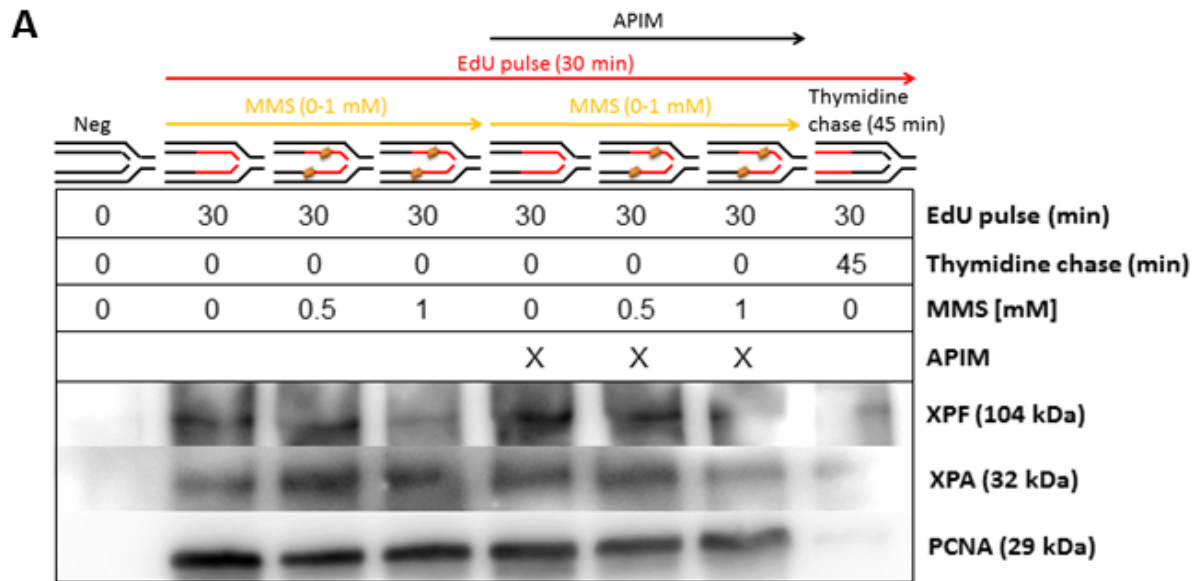H4K16ac (10 kDa)
PCNA (29 kDa)
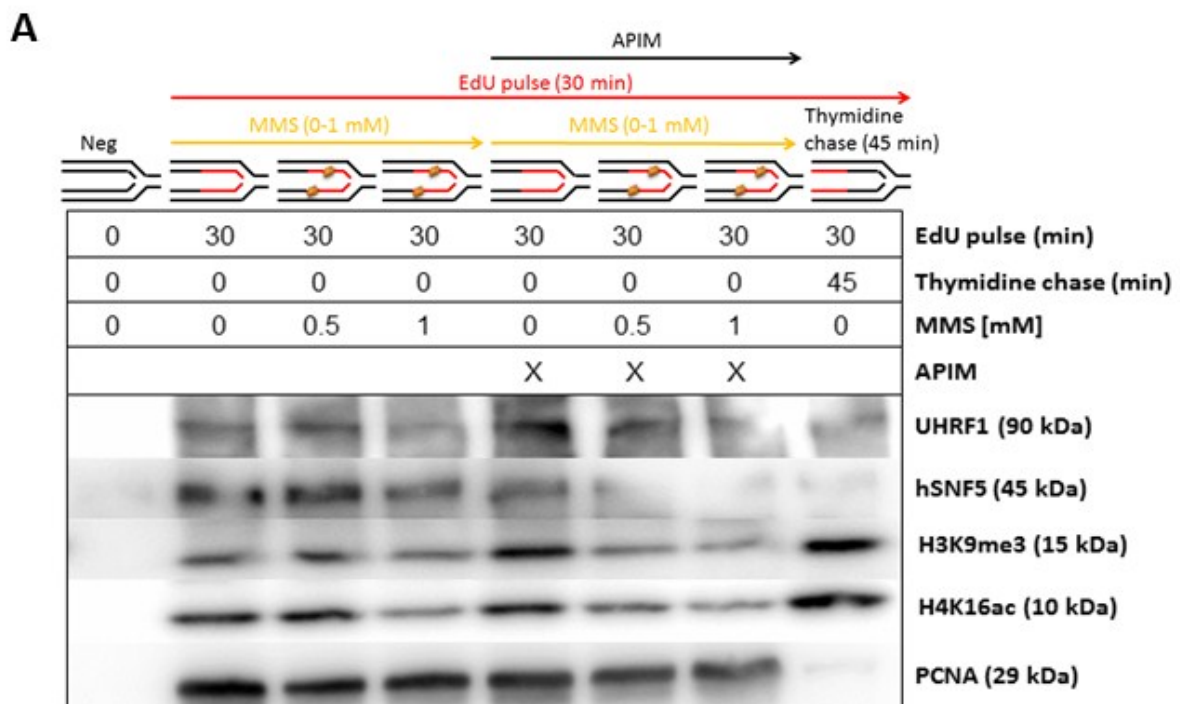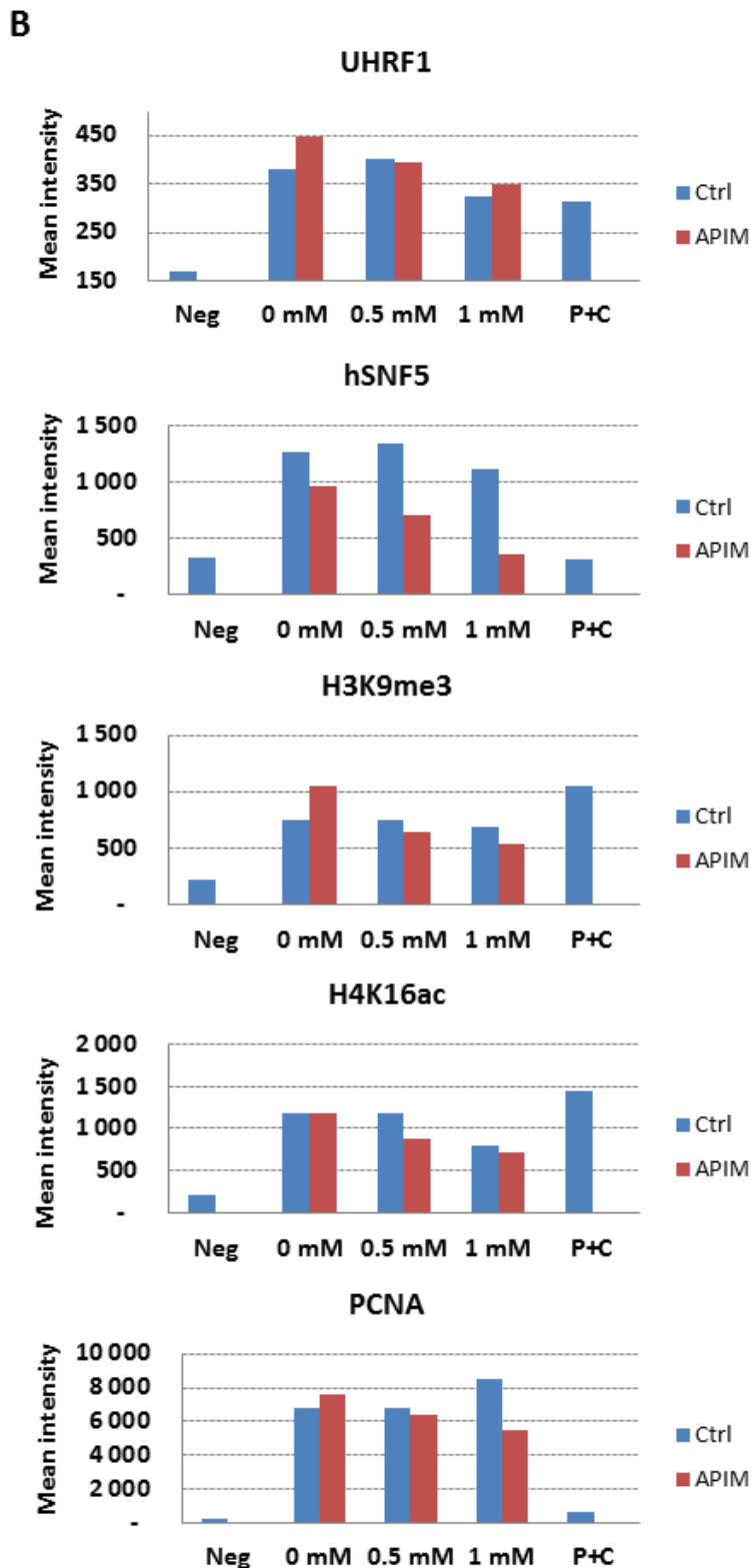
Figure 3.10 iPOND from cells labeled with EdU (30 min) and MMS (0-1mM). In addition, one sample was not EdU-labeled (neg) and one sample was followed by a chase in thymidine-containing medium (45 min) (P+C). **A)** Pull-down of UHRF1, hSNF5, H3K9me3, H4K16ac and PCNA. PCNA bonds are the same as shown in Figure 3.8. The bonds of H3K9me3 and H4K16ac are also partly shown in Figure 3.7B. **B)** Quantification of bond intensities in A). Blue bars: Cells not expressing APIM (control). Red bars: Cells expressing APIM.

4 Discussion

4.1 Optimization of iPOND

The pilot study for APIM-expression implied that a tetracycline concentration of 0.03 µg/mL was toxic to the cells. A decline in cell proliferation with increasing tetracycline concentrations was therefore expected. However, tetracycline did not show a pronounced toxicity at the concentrations tested when added the day after passaging the cells. The Flp-IN^{TM} T$^-$Rex^{TM}-293 APIM-YFP cells probably need more than four hours to attach properly, and the cells could be stressed right after attachment. Tetracycline might be more toxic to stressed cells, explaining the lower proliferation rate observed in the pilot study. Consequently, the cells should be split one day prior to tetracycline addition.

A cell confluence of $3x10^7$ cells/dish gave the best iPOND results. Higher confluence seemed to result in lower EdU-incorporation, probably due to less replication. The cell proliferation of Flp-IN^{TM} T$^-$Rex^{TM}-293 APIM-YFP cells appears to be reduced when the cells are more confluent. Furthermore, $2x10^8$ cells per sample were needed to detect signals of proteins with low abundance close to replication forks. High background signal could be caused by unspecific protein binding to the streptavidin beads. Therefore, the concentration of beads were adjusted to 100 µL bead slurry per $2x10^8$ cells and the number of washing steps prior to elution of proteins were increased with two more steps compared to the original protocol. In addition, the time of elution was increased with 5 minutes to ensure complete crosslink reversal, as formaldehyde crosslinking may interfere with epitope detection (Sirbu et al., 2012).

It was concluded that MMS concentrations of 0.5-1 mM introduce DNA damage without stalling the replication machinery excessively. Higher doses of MMS would probably give a more distinct difference when evaluating pull-down of DNA repair proteins before and after damage, but it would be at the sacrifice of slower replication rate.

4.2   Variations in Pull-Down of Proteins

The results in chapter 3.2 showed variations in pull-down of both XPA and XPF among the pulse samples. More of the protein captures could be detected after 10 and 15 minutes than after 5 minutes of EdU-incorporation, and less was detected after 15 minutes than after 10 minutes. An explanation for the variations in protein detection could be uneven loading of the sample material. Unfortunately, there is no appropriate loading control to standardize the results. Another explanation for the observed trend is presented by the model in Figure 4.1.
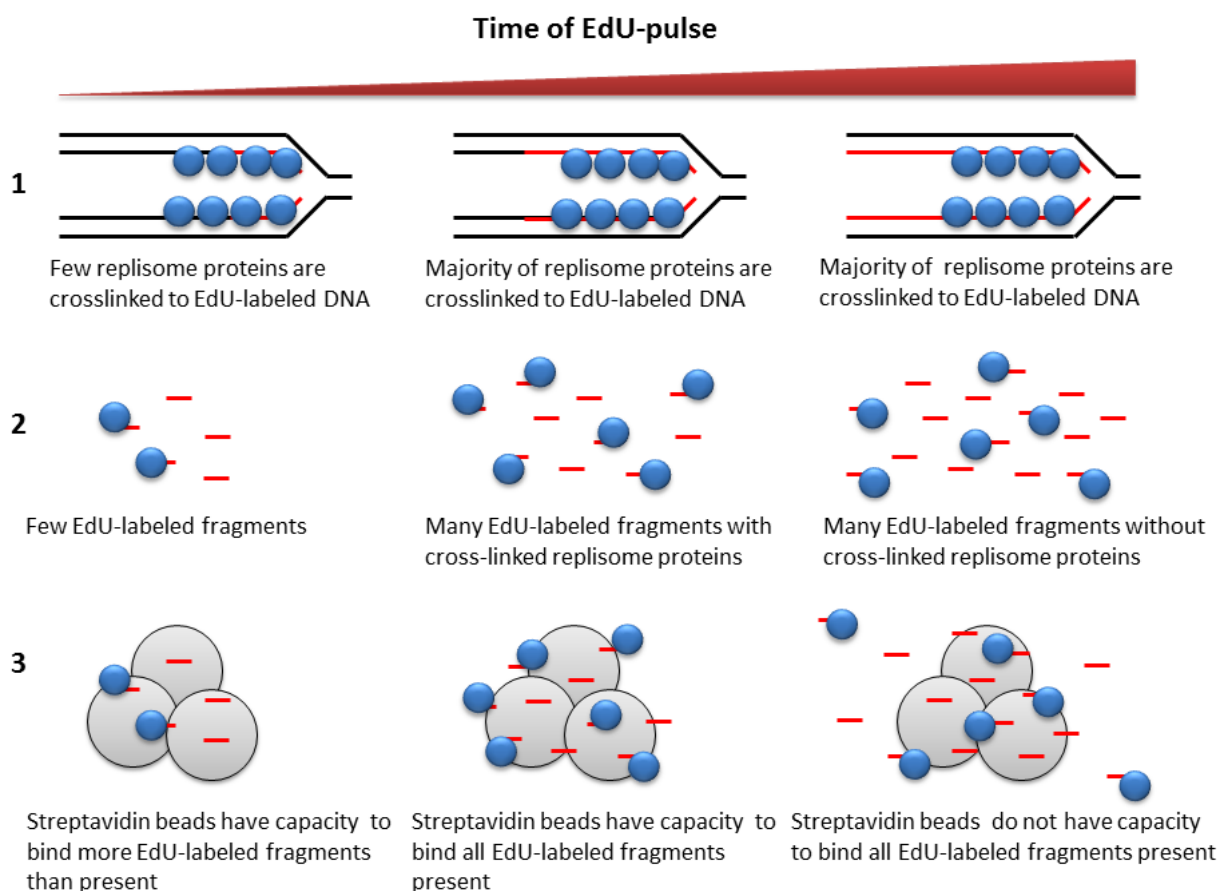
Figure 4.1 Model for various detection of replisome proteins (blue). 1) Short time of EdU-pulse might not be sufficient for all replisome proteins to be crosslinked to the labeled DNA, while longer time of EdU-incorporation can detect proteins further from the replication fork. 2) After cell lysis and sonication, the size of DNA fragments should be independent on length of EdU-pulse, but longer pulses will yield higher amount of labeled DNA fragments. For replisome proteins, long labeling times could give high amount of EdU-labeled fragments without cross-linked replisome proteins. 3) EdU-labeled fragments are pulled down by streptavidin-beads. If the capacity of the beads is exceeded, lower protein detection could be the result.

iPOND resolution is dependent on length of EdU-pulse, rate of DNA synthesis and size of DNA fragments after sonication (Sirbu et al., 2011). Higher protein detection with increasing length of EdU-pulse suggests that more proteins can be found further from the replication fork, when assuming equal replication rate and size of fragments in all samples. The capacity of the streptavidin beads is not known. Long pulses with EdU might give high concentration of DNA fragments without crosslinked proteins of interest, if these proteins are close to the replication fork. Thus, if the beads do not have capacity to bind all fragments, a decrease in protein detection could be the result. However, pull-down of PCNA from the same experiment did not decrease with increasing length of EdU-pulse, suggesting that unevenly loading of the sample material is the most plausible explanation.

4.3   Post-Replicative Repair by NER and hABH2

The findings of XPA and XPF in close proximity to the replication fork suggest that NER is associated with the replisome performing post-replicative repair. The results support the model suggested by (Gilljam et al., 2012, submitted) as presented in Figure 4.2. The model hypothesizes that NER in S-phase tightly follows damage bypass by TLS polymerases. Pol η has the ability to replicate past UV-induced CPDs with high fidelity, and 6-4 PPs has been reported to be bypassed by Pol ζ or Rev1 (Lehmann, 2006). Thus, damage recognized by NER is often bypassed and needs to be repaired after replication. An association between NER and the replication machinery is somewhat controversial, as NER is one of the most well-documented repair pathways and no one has previously reported a link to replication. However, it is reasonable for NER to repair bypassed DNA lesions subsequent to replication as the chromatin is in an uncondensed state; hence the NER proteins have access to the damage site.
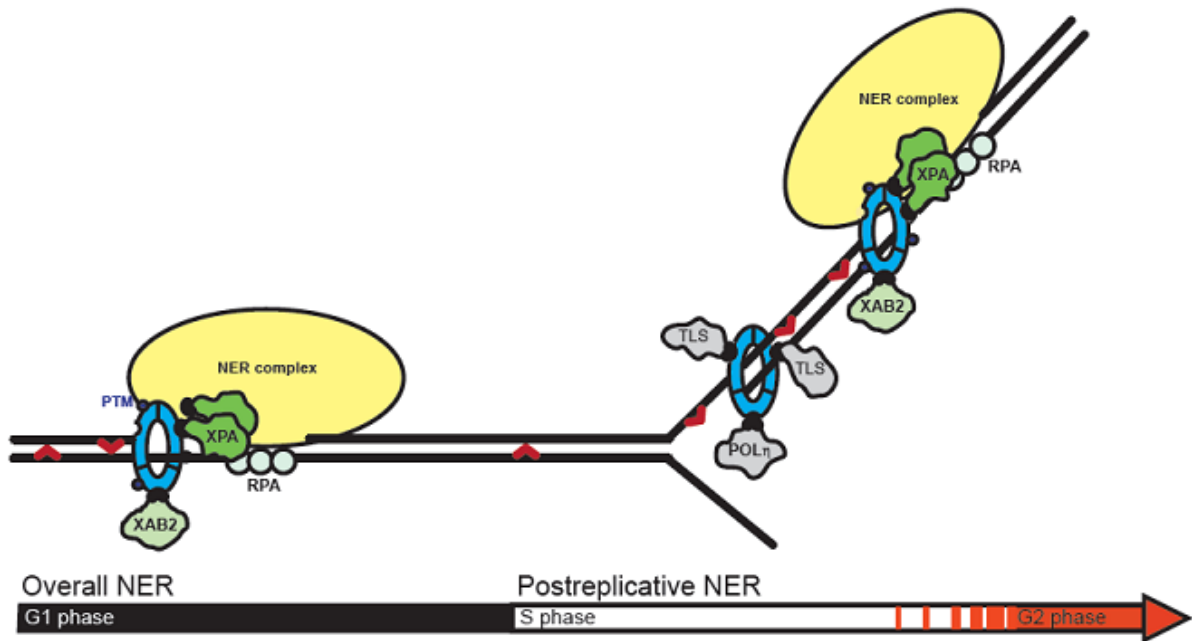
Figure 4.2 Model presenting the role of NER in S-phase, where it tightly follows damage bypass by TLS polymerases (Gilljam et al., 2012, submitted).

hABH2's presence close to the replication fork was verified by iPOND. Its binding to nascent DNA suggests that hABH2 has, at least partly, a function in post-replicative repair. Thus, the model presented for NER proteins probably also applies for hABH2. hABH2 is known to repair 1meA and 3meC lesions caused by alkylating agents, but works most effectively on 1meA (Nieminuszczy and Grzesiuk, 2007). Methyl lesions have been reported to be bypassed by TLS polymerases, at least in *E.coli*. Bypass of 1meA has been shown to result in low mutagenicity, as the correct nucleotide (T) is inserted above 99% of the lesions (Falnes et al., 2007). This supports that hABH2 might remove alkylation damage after they are bypassed. hABH2 preferentially work on double stranded DNA, and can remove alkylation lesions directly without excision, hence does not cause breaks in the DNA. Therefore, it is also possible for hABH2 to function in pre-replicative repair, as its action in front of the replication fork would not pose a potential threat for replication forks collapse.

4.4 Effects of APIM Expression

The binding of UHRF1 to nascent DNA did not seem to be reduced by overexpressed APIM, suggesting that APIM in UHRF1 is not responsible for its PCNA-interaction. However, DNMT1 can interact directly with both PCNA and UHRF1. If APIM is a functional PCNA-interacting motif in UHRF1, overexpressed APIM might block its binding site on PCNA and lead to binding of UHRF1 to DNMT1 instead. Hence, even if overexpressed APIM disturbs the direct interaction between UHRF1 and PCNA, UHRF1 still travels with active replication forks, as illustrated in Figure 4.3. This could be another possible explanation for why overexpression of APIM did not reduce pull-down of UHRF1.
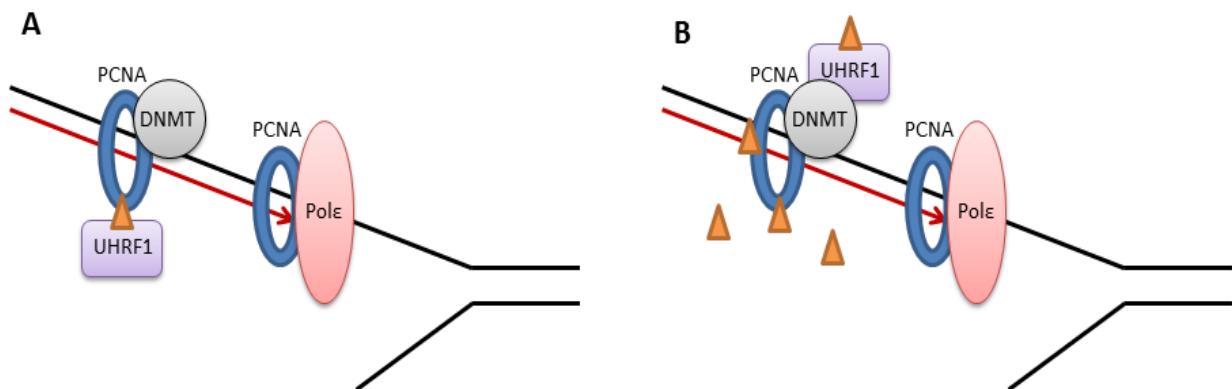
Figure 4.3 Model proposing how overexpression of APIM could affect binding of UHRF1 to PCNA. **A)** In normal cells, UHRF1 (purple) might bind to PCNA (blue) through its APIM motif (orange). **B)** In cells where APIM is overexpressed, APIM may block the binding site of UHRF1 on PCNA and lead to UHRF1 binding to DNMT (grey) instead.

Reduced capture of XPA, hABH2, hSNF5, H3K9me3 and H4K16ac was observed in APIM-expressing cells. In general, the reduced level of proteins present on nascent DNA when overexpressing APIM was slightly more pronounced in cells treated with MMS. As PCNA is believed to be modified after DNA damage (Ulrich, 2009), these results support the hypothesis that APIM-containing proteins bind more strongly to PCNA with PTMs. The reason for reduced pull-down of proteins from APIM-expressing cells is probably that overexpressed APIM blocks the binding site on PCNA, leading to impaired interaction between PCNA and proteins that bind to PCNA through their APIM motif. This leads to reduced presence of APIM-containing

proteins or proteins modified by APIM-containing proteins at newly synthesized DNA. However, reduced capture of PCNA was also observed in APIM-expressing cells treated with 1 mM MMS, suggesting that DNA damage together with overexpressed APIM cause more replication arrest. This could explain why a reduced capture of XPA, hABH2, hSNF5, H3K9me3 and H4K16ac was observed in this sample. Nevertheless, capture of UHRF1 and XPF did not show reduced signal in MMS-treated APIM-expressing cells compared to cells not expressing APIM. Furthermore, XPF, which do not have APIM, was pulled down in slightly increased amount in APIM expressing cells, and thus appeared to have an opposite trend compared to APIM-containing proteins that showed reduced pull-downs. For these reasons, the results support that overexpression of APIM perturbs the binding of APIM-containing proteins to nascent DNA and interfere with the function of APIM-containing protein complexes responsible for certain histone modifications. It should, however, be emphasized that the results need to be reproduced before they can be given too much value.

Inhibition of central DNA repair proteins may lead to increased efficiency of chemotherapeutic agents. XPA does not appear to have any functions outside NER, as the rest of the NER proteins do (Köberle et al., 2006), and may thus be a target for inhibition of the NER pathway in cancer therapy. Reduced binding of XPA and hABH2 to DNA by overexpression of APIM supports the findings that APIM leads to decreased NER efficiency (Gilljam et al., 2012, submitted) and decreased efficiency of hABH2 in removal of 1meA (Gilljam et al., 2009) and can partly explain the potential of APIM in cancer therapy.

Interestingly, a connection between NER and hSNF5 has been identified. hSNF5 has been found to co-localize with XPC and to contribute to increased access of NER proteins at the site of damage (Ray et al., 2009). MRG15, which contain APIM, is suggested to participate in the acetylation of H4K16 (Wu et al., 2011) that leads to unfolding of the chromatin (Luijsterburg and Van Attikum, 2011). Both reduced presence of hSNF5 and H4K16ac on nascent DNA by overexpression of APIM can thus result in more compact chromatin. Furthermore, this could inhibit DNA repair proteins to gain access to the damage site, and thus promote apoptosis of damaged cells. On the other hand, overexpressed APIM also seemed to reduce the capture of H3K9me3. H3K9me3 condenses chromatin, hence reduced amount of this

modification possible leads to a more open chromatin landscape and better access for DNA repair proteins. SUV39H1, that catalyzes trimethylation of H3K9, has been found in a p53-MDM2-SUV39H1/EHMT1 complex that inhibits the activity of p53, where EHMT1 I this complex contains APIM. p53 cause cell arrest or apoptosis upon DNA damage (Chen et al., 2010). If the reduced amount of H3K9me3 found on nascent DNA is caused by overexpressed APIM interfering with the function of the p53-MDM2-SUV39H1/EHMT1 complex, increased apoptosis could be the result. Overexpressed APIM could also possibly interfere with the function of the APIM-containing SENP2. SENP2 stimulates MDM2 to degrade p53, thus reduced function of SENP2 could also lead to increased apoptosis, and partly explain the potential use of APIM in cancer therapy.

Although many APIM-containing proteins have defined roles in histone modifications, it is difficult to point out the exact reasons why overexpressed APIM affects the histone modifications. The field of epigenetics and histone modifications is relatively new and still remains to be totally resolved. However in this context, the observation that APIM-expression effects histone modifications is of importance.

4.5 Future Work

First and foremost, the observations of reduced protein presence on nascent DNA when overexpressing APIM are subtle; hence they need to be reproduced before they can be given too much value. Furthermore, it could be interesting to probe for other proteins with APIM and for other histones that are modified by APIM-containing proteins to examine whether their presence on newly synthesized DNA are reduced when APIM is overexpressed. Antibodies for other APIM-containing proteins than those presented in this thesis have been tested, but their protein targets were not detected neither in input nor capture. Thus, it is a challenge to find antibodies that are sensitive enough for the detection of low abundant proteins. Moreover, it could be interesting to perform iPOND in combination with another form of DNA damage that is more relevant for epigenetics. Most of what is reported regarding epigenetics and response to DNA damage is after introducing DSBs. Bleomycin is one DNA damaging agent that introduces DSBs. An iPOND experiment could be performed with an EdU-pulse followed by a long thymidine-chase and subsequently a bleomycin-chase to investigate epigenetics and evaluate the effects of overexpressed APIM outside replication. Reduced pull-down of hSNF5, H3K9me3 and H4K16ac in APIM-expressing cells, suggest that APIM might be a functional motif in hSNF5, EHMT1 and MRG15. Therefore, experiments to confirm whether APIM actually is a functional PCNA interacting motif in these proteins should be determined.

5 Conclusion

During optimization of iPOND it was found that 3×10^7 cells/dish gives optimal EdU-incorporation in Flp-IN^{TM} T$^-$Rex^{TM}-293 APIM-YFP cells. Furthermore, it was found that 2×10^8 cells/sample is necessary to detect proteins with low abundance close to replication forks. To avoid reduced proliferation rate, the cells must be passaged the day before adding tetracycline at a concentration of 0.02 µg/mL to induce and sustain APIM-expression. Finally, it was found that 0.5-1 mM MMS are the most appropriate concentrations to introduce DNA damage without excessive stalling of the replication machinery.

NER proteins and hABH2, performing direct repair, were found in close proximity to active replication forks, suggesting that they have a function in post-replicative repair. iPOND also verified the chromatin remodeling factors UHRF1 and hSNF5 as replisome proteins and the modified histones H3K9me3 and H4K16ac as chromatin-bound proteins.

XPA, hABH2, hSNF5, H3K9me3 and H4K16ac were all pulled down in slightly reduced amounts in APIM-expressing cells treated with MMS compared to cells not expressing APIM. The reason for the reduced pull-down is probably that overexpressed APIM blocks the binding site on PCNA where PCNA-interacting proteins bind through their APIM motif. This leads to reduced presence of APIM-containing proteins or proteins modified by APIM-containing proteins at newly synthesized DNA. In conclusion, overexpressed APIM seems to perturb the binding of APIM-containing proteins involved in DNA repair and chromatin remodeling to nascent DNA, in addition to interfere with the function of APIM-containing protein complexes responsible for certain histone modifications. Thus, APIM seems to affect several cell processes, partly explaining its potential use in cancer therapy.

Literature and References

AAS, P. A., OTTERLEI, M., FALNES, P. Ø., VÅGBØ, C. B., SKORPEN, F., AKBARI, M., SUNDHEIM, O., BJØRÅS, M., SLUPPHAUG, G. & SEEBERG, E. 2003. Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature,* 421**,** 859-863.

ALABERT, C. & GROTH, A. 2012. Chromatin replication and epigenome maintenance. *Nature Reviews Molecular Cell Biology,* 13**,** 153-167.

ALBERTS, B., JOHNSON, A., LEWIS, J., RAFF, M., ROBERTS, K. & WALTER, P. 2008. *Molecular Biology of the Cell*, Garland Science.

BANNISTER, A. J. & KOUZARIDES, T. 2011. Regulation of chromatin by histone modifications. *Cell research,* 21**,** 381-395.

BECKER, W. M., KLEINSMITH, L. J., HARDIN, J. & BERTONI, G. P. 2009. *The World of the Cell*, Pearson Education.

BRONNER, C., FUHRMANN, G., CHÉDIN, F. L., MACALUSO, M. & DHE-PAGANON, S. 2010. UHRF1 links the histone code and DNA methylation to ensure faithful epigenetic memory inheritance. *Genetics & epigenetics,* 2009**,** 29.

BURGERS, P. M. J. 2009. Polymerase dynamics at the eukaryotic DNA replication fork. *Journal of Biological Chemistry,* 284**,** 4041.

CEDAR, H. & BERGMAN, Y. 2009. Linking DNA methylation and histone modification: patterns and paradigms. *Nature Reviews Genetics,* 10**,** 295-304.

CHEN, L., LI, Z., ZWOLINSKA, A. K., SMITH, M. A., CROSS, B., KOOMEN, J., YUAN, Z. M., JENUWEIN, T., MARINE, J. C. & WRIGHT, K. L. 2010. MDM2 recruitment of lysine methyltransferases regulates p53 transcriptional output. *The EMBO journal,* 29**,** 2538-2552.

CHIU, S. Y., ASAI, N., COSTANTINI, F. & HSU, W. 2008. SUMO-specific protease 2 is essential for modulating p53-Mdm2 in development of trophoblast stem cell niches and lineages. *PLoS biology,* 6**,** e310.

DUNCAN, T., TREWICK, S. C., KOIVISTO, P., BATES, P. A., LINDAHL, T. & SEDGWICK, B. 2002. Reversal of DNA alkylation damage by two human dioxygenases. *Proceedings of the National Academy of Sciences,* 99**,** 16660.

EGGER, G., LIANG, G., APARICIO, A. & JONES, P. A. 2004. Epigenetics in human disease and prospects for epigenetic therapy. *Nature,* 429**,** 457-463.

EUSKIRCHEN, G. M., AUERBACH, R. K., DAVIDOV, E., GIANOULIS, T. A., ZHONG, G., ROZOWSKY, J., BHARDWAJ, N., GERSTEIN, M. B. & SNYDER, M. 2011. Diverse roles and interactions of the SWI/SNF chromatin remodeling complex revealed using global approaches. *PLoS genetics,* 7**,** e1002008.

FALNES, P. Ø., KLUNGLAND, A. & ALSETH, I. 2007. Repair of methyl lesions in DNA and RNA by oxidative demethylation. *Neuroscience,* 145**,** 1222-1232.

FLANAGAN, J. F. & PETERSON, C. L. 1999. A role for the yeast SWI/SNF complex in DNA replication. *Nucleic Acids Research,* 27**,** 2022-2028.

GILLJAM, K. M., FEYZI, E., AAS, P. A., SOUSA, M. M. L., MÜLLER, R., VÅGBØ, C. B., CATTERALL, T. C., LIABAKK, N. B., SLUPPHAUG, G. & DRABLØS, F. 2009. Identification of a novel, widespread, and functionally important PCNA-binding motif. *The Journal of cell biology,* 186**,** 645.

GILLJAM, K. M., MÜLLER, R., LIABAKK, N. B., OTTERLEI, M. 2012. Nucleotide excision repair is associated with the replisome and its efficiency depends on a direct interaction between XPA and PCNA. *Nucleic acids research.*

GROTH, A., ROCHA, W., VERREAULT, A. & ALMOUZNI, G. 2007. Chromatin challenges during DNA replication and repair. *Cell,* 128**,** 721-733.

HAKEM, R. 2008. DNA-damage repair; the good, the bad, and the ugly. *The EMBO journal,* 27**,** 589-605.

HELLEDAY, T., PETERMANN, E., LUNDIN, C., HODGSON, B. & SHARMA, R. A. 2008. DNA repair pathways as targets for cancer therapy. *Nature Reviews Cancer,* 8**,** 193-204.

HENDEL, A., KRIJGER, P. H. L., DIAMANT, N., GOREN, Z., LANGERAK, P., KIM, J., REIßNER, T., LEE, K., GEACINTOV, N. E. & CARELL, T. 2011. PCNA Ubiquitination Is Important, But Not Essential for Translesion DNA Synthesis in Mammalian Cells. *PLoS genetics,* 7**,** e1002262.

HOUTGRAAF, J. H., VERSMISSEN, J. & VAN DER GIESSEN, W. J. 2006. A concise review of DNA damage checkpoints and repair in mammalian cells. *Cardiovascular Revascularization Medicine,* 7**,** 165-172.

KOLB, H. C., FINN, M. & SHARPLESS, K. B. 2001. Click chemistry: diverse chemical function from a few good reactions. *Angewandte Chemie International Edition,* 40**,** 2004-2021.

KÖBERLE, B., ROGINSKAYA, V. & WOOD, R. D. 2006. XPA protein as a limiting factor for nucleotide excision repair and UV sensitivity in human cells. *DNA repair,* 5**,** 641-648.

LANGE, S. S., TAKATA, K. & WOOD, R. D. 2011. DNA polymerases and cancer. *Nature Reviews Cancer,* 11**,** 96-110.

LEE, D., SOHN, H., KALPANA, G. V. & CHOE, J. 1999. Interaction of E1 and hSNF5 proteins stimulates replication of human papillomavirus DNA. *Nature,* 399**,** 487-491.

LEHMANN, A. R. 2006. Translesion synthesis in mammalian cells. *Experimental cell research,* 312**,** 2673-2676.

LUIJSTERBURG, M. S. & VAN ATTIKUM, H. 2011. Chromatin and the DNA damage response: The cancer connection. *Molecular oncology.*

MADHUSUDAN, S. & MIDDLETON, M. R. 2005. The emerging role of DNA repair proteins as predictive, prognostic and therapeutic targets in cancer. *Cancer treatment reviews,* 31**,** 603-617.

MCNAIRN, A. J. & GILBERT, D. M. 2003. Epigenomic replication: linking epigenetics to DNA replication. *Bioessays,* 25**,** 647-656.

MOLDOVAN, G. L., PFANDER, B. & JENTSCH, S. 2007. PCNA, the maestro of the replication fork. *Cell,* 129**,** 665-679.

NARYZHNY, S. 2008. Proliferating cell nuclear antigen: a proteomics view. *Cellular and molecular life sciences,* 65**,** 3789-3808.

NIEMINUSZCZY, J. & GRZESIUK, E. 2007. Bacterial DNA repair genes and their eukaryotic homologues: 3. AlkB dioxygenase and Ada methyltransferase in the direct repair of alkylated DNA. *ACTA BIOCHIMICA POLONICA-ENGLISH EDITION-,* 54**,** 459.

NIEMINUSZCZY, J., MIELECKI, D., SIKORA, A., WRZESIŃSKI, M., CHOJNACKA, A., KRWAWICZ, J., JANION, C. & GRZESIUK, E. 2009. Mutagenic potency of MMS-induced 1meA/3meC lesions in E. coli. *Environmental and molecular mutagenesis,* 50**,** 791-799.

NOUSPIKEL, T. 2009. DNA repair in mammalian cells. *Cellular and molecular life sciences,* 66**,** 994-1009.

RAY, A., MIR, S. N., WANI, G., ZHAO, Q., BATTU, A., ZHU, Q., WANG, Q. E. & WANI, A. A. 2009. Human SNF5/INI1, a component of the human SWI/SNF chromatin remodeling complex, promotes nucleotide excision repair by influencing ATM recruitment and downstream H2AX phosphorylation. *Molecular and cellular biology,* 29**,** 6206-6219.

RICE, J. C., BRIGGS, S. D., UEBERHEIDE, B., BARBER, C. M., SHABANOWITZ, J., HUNT, D. F., SHINKAI, Y. & ALLIS, C. D. 2003. Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Molecular cell,* 12**,** 1591-1598.

ROBERTS, C. W. M. & ORKIN, S. H. 2004. The SWI/SNF complex—chromatin and cancer. *Nature Reviews Cancer,* 4**,** 133-142.

SEDGWICK, B., BATES, P. A., PAIK, J., JACOBS, S. C. & LINDAHL, T. 2007. Repair of alkylated DNA: recent advances. *DNA repair,* 6**,** 429-442.

SHOGREN-KNAAK, M., ISHII, H., SUN, J. M., PAZIN, M. J., DAVIE, J. R. & PETERSON, C. L. 2006. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science,* 311**,** 844-847.

SIRBU, B. M., COUCH, F. B. & CORTEZ, D. 2012. Monitoring the spatiotemporal dynamics of proteins at replication forks and in assembled chromatin using isolation of proteins on nascent DNA. *Nature Protocols,* 7**,** 594-605.

SIRBU, B. M., COUCH, F. B., FEIGERLE, J. T., BHASKARA, S., HIEBERT, S. W. & CORTEZ, D. 2011. Analysis of protein dynamics at active, stalled, and collapsed replication forks. *Genes & Development,* 25**,** 1320-1327.

ULRICH, H. D. 2009. Regulating post-translational modifications of the eukaryotic replication clamp PCNA. *DNA repair,* 8**,** 461-469.

WAGA, S. & STILLMAN, B. 1998. The DNA replication fork in eukaryotic cells. *Annual review of biochemistry,* 67**,** 721-751.

WARBRICK, E. 2000. The puzzle of PCNA's many partners. *Bioessays,* 22**,** 997-1006.

WARBRICK, E., LANE, D. P., GLOVER, D. M. & COX, L. S. 1997. Homologous regions of Fen1 and p21Cip1 compete for binding to the same site on PCNA: a potential mechanism to co-ordinate DNA replication and repair. *Oncogene,* 14**,** 2313.

WU, J., CHEN, Y., LU, L. Y., WU, Y., PAULSEN, M. T., LJUNGMAN, M., FERGUSON, D. O. & YU, X. 2011. Chfr and RNF8 synergistically regulate ATM activation. *Nature structural & molecular biology,* 18**,** 761-768.

YANG, Z., LIU, Y., MAO, L. Y., ZHANG, J. T. & ZOU, Y. 2002. Dimerization of human XPA and formation of XPA2-RPA protein complex. *Biochemistry,* 41**,** 13012-13020.

ZHANG, X., WEN, H. & SHI, X. 2012. Lysine methylation: beyond histones. *Acta biochimica et biophysica Sinica,* 44**,** 14-27.

Appendix – Quantification of bonds after Western blotting

To quantify the bond intensities after Western blotting, the manual ROI tools in the KODAK MI software were used. The bonds were marked and the mean intensities of the marked areas were calculated by the software program.

The data of mean intensities are shown in Table A.1-A.3. The same table number indicates that the data are from the same experiment. The quantifications are presented in chapter 3.

Table A.1.1 Mean intensity of PCNA bonds from capture samples.

PCNA	Mean Intensity
Negative	178
Pulse (30 min)	3 144
Pulse (30 min)+MMS (0.5 mM)	4 036
Pulse (30 min)+MMS (1 mM)	2 412
Pulse (30 min)+MMS (2 mM)	1 363
Pulse (30 min)+MMS (3 mM)	1 508
Pulse (30 min)+MMS (4 mM)	1 381
Pulse (30 min)+Chase (45 min)	409

Table A.1.2 Mean intensity of hABH2 bonds from capture samples.

hABH2	Mean Intensity
Negative	453
Pulse (30 min)	619
Pulse (30 min)+MMS (0.5 mM)	702
Pulse (30 min)+MMS (1 mM)	705
Pulse (30 min)+MMS (2 mM)	698
Pulse (30 min)+MMS (3 mM)	684
Pulse (30 min)+MMS (4 mM)	592
Pulse (30 min)+Chase (45 min)	543

Table A.2.1 Mean intensity of XPF bonds from capture samples.

XPF	Mean Intensity
Negative	194
Pulse (5 min)	390
Pulse (10 min)	480
Pulse (15 min)	434
Pulse (15 min)+Chase (30 min)	262

Table A.2.2 Mean intensity of XPA bonds from capture samples.

XPA	Mean Intensity
Negative	1 100
Pulse (5 min)	1 865
Pulse (10 min)	2 105
Pulse (15 min)	2 011
Pulse (15 min)+Chase (30 min)	1 279

Table A.2.3 Mean intensity of PCNA bonds from capture samples.

PCNA	Mean Intensity
Negative	482
Pulse (5 min)	7 608
Pulse (10 min)	7 567
Pulse (15 min)	8 050
Pulse (15 min)+Chase (30 min)	2 489

Table A.2.4 Mean intensity of H3 bonds from capture samples.

H3	Mean Intensity
Negative	1 240
Pulse (5 min)	7 345
Pulse (10 min)	10 558
Pulse (15 min)	12 372
Pulse (15 min)+Chase (30 min)	13 649

Table A.3.1 Mean intensity of XPF bonds from capture samples.

XPF	Mean Intensity
Negative	246
Pulse (30 min)	581
Pulse (30 min)+MMS (0.5 mM)	562
Pulse (30 min)+MMS (1 mM)	454
Pulse (30 min)+APIM	602
Pulse (30 min)+MMS (0.5 mM)+APIM	567
Pulse (30 min)+MMS (1 mM)+APIM	473
Pulse (30 min)+Chase (45 min)	378

Table A.3.2 Mean intensity of UHRF1 bonds from capture samples.

UHRF1	Mean Intensity
Negative	171
Pulse (30 min)	383
Pulse (30 min)+MMS (0.5 mM)	401
Pulse (30 min)+MMS (1 mM)	326
Pulse (30 min)+APIM	447
Pulse (30 min)+MMS (0.5 mM)+APIM	394
Pulse (30 min)+MMS (1 mM)+APIM	350
Pulse (30 min)+Chase (45 min)	313

Table A.3.3 Mean intensity of hSNF5 bonds from capture samples.

hSNF5	Mean Intensity
Negative	331
Pulse (30 min)	1 262
Pulse (30 min)+MMS (0.5 mM)	1 331
Pulse (30 min)+MMS (1 mM)	1 107
Pulse (30 min)+APIM	966
Pulse (30 min)+MMS (0.5 mM)+APIM	696
Pulse (30 min)+MMS (1 mM)+APIM	357
Pulse (30 min)+Chase (45 min)	306

Table A.3.4 Mean intensity of XPA bonds from capture samples.

XPA	Mean Intensity
Negative	413
Pulse (30 min)	986
Pulse (30 min)+MMS (0.5 mM)	1 211
Pulse (30 min)+MMS (1 mM)	1 139
Pulse (30 min)+APIM	1 103
Pulse (30 min)+MMS (0.5 mM)+APIM	1 086
Pulse (30 min)+MMS (1 mM)+APIM	770
Pulse (30 min)+Chase (45 min)	401

Table A.3.5 Mean intensity of hABH2 bonds from capture samples.

hABH2	Mean Intensity
Negative	129
Pulse (30 min)	212
Pulse (30 min)+MMS (0.5 mM)	218
Pulse (30 min)+MMS (1 mM)	207
Pulse (30 min)+APIM	173
Pulse (30 min)+MMS (0.5 mM)+APIM	169
Pulse (30 min)+MMS (1 mM)+APIM	167
Pulse (30 min)+Chase (45 min)	132

Table A.3.6 Mean intensity of PCNA bonds from capture samples.

PCNA	Mean Intensity
Negative	249
Pulse (30 min)	6 775
Pulse (30 min)+MMS (0.5 mM)	6 770
Pulse (30 min)+MMS (1 mM)	8 539
Pulse (30 min)+APIM	7 625
Pulse (30 min)+MMS (0.5 mM)+APIM	6 397
Pulse (30 min)+MMS (1 mM)+APIM	5 542
Pulse (30 min)+Chase (45 min)	678

Table A.3.7 Mean intensity of H3K9me3 bonds from capture samples.

H3K9me3	Mean Intensity
Negative	216
Pulse (30 min)	745
Pulse (30 min)+MMS (0.5 mM)	756
Pulse (30 min)+MMS (1 mM)	693
Pulse (30 min)+APIM	1 051
Pulse (30 min)+MMS (0.5 mM)+APIM	639
Pulse (30 min)+MMS (1 mM)+APIM	540
Pulse (30 min)+Chase (45 min)	1 048

Table A.3.8 Mean intensity of H4K16ac bonds from capture samples.

H4K16ac	Mean Intensity
Negative	212
Pulse (30 min)	1 173
Pulse (30 min)+MMS (0.5 mM)	1 178
Pulse (30 min)+MMS (1 mM)	797
Pulse (30 min)+APIM	1 185
Pulse (30 min)+MMS (0.5 mM)+APIM	878
Pulse (30 min)+MMS (1 mM)+APIM	717
Pulse (30 min)+Chase (45 min)	1 443