DNA-Protein Cross-links: Formation in Cells and Tissues, Repair, and Inhibition of DNA Transcription

A THESIS

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF THE

UNIVERSITY OF MINNESOTA

BY

Daeyoon Park

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

Dr. Natalia Y. Tretyakova, Advisor

April 2019

Acknowledgement

First and foremost, I would like to thank my advisor, Dr. Natalia Tretyakova, for her great support and guidance throughout my degree. She was a great role model for me as a scientist, and all the opportunities in her lab encouraged me to grow as a scientist. I feel very fortunate to have worked with such a great mentor who inspires her students.

I thank Dr. Colin Campbell for his advice and supports during my graduate career. I also want to acknowledge him for the collaboration throughout my projects and valuable feedback he has provided on my work. I would like to express my thanks to my thesis committee chair, Dr. Kate Adamala, and members for their advice, support, and feedbacks throughout my thesis.

I wish to thank Dr. Mark Greenberg and Dr. Kun Yang at John Hopkins University for the opportunities, allowing me to contribute to their project focused on DNA-protein cross-links induced by monofunctional alkylating agents. Furthermore, I would like to thank all the collaborators who contributed to my projects, especially Dr. Lei Li (The University of Texas MD Anderson Cancer Center), Dr. Yuichi Machida (Mayo Clinic), Dr. Deborah Ferrington (University of Minnesota), and Dr. Ashis Basu (University of Connecticut).

I would like to thank to the former and current lab members in the Tretyakova lab who helped me with my projects and provided great friendship and valuable discussions.

I wish to thank my parents, Jindal Park and Youngsun Lee, for making me see the greater picture, and my sister, Jeeyoon, for encouraging me. And lastly, I would like to thank my wife, Hyojin, for her perpetual love and support.

Abstract

DNA is constantly damaged by exogenous and endogenous agents, generating a range of nucleobase lesions. It is important to understand the biological consequences and repair mechanisms of DNA adducts. Cellular proteins can become covalently trapped on DNA to generate DNA-protein crosslinks (DPCs). Because of their unusually bulky nature, DPCs are anticipated to block many cellular processes including replication, transcription, and repair. However, cellular effects of DPCs have not been fully elucidated. **Chapter 1** of this thesis provides background information on the formation, biological consequences, and repair pathways of DPCs studied in previous studies.

In **Chapter 2,** we employed a quantitative nanoLC-ESI$^+$-MS/MS assay to investigate the formation of free radical-induced DPCs between thymidine in DNA and tyrosine sidechains of proteins. This methodology was used to examine the role of SPRTN protease and immunoproteasome in DPC repair in human cells and mouse models.

In **Chapter 3,** a mass spectrometry based CTAB assay was used to study the effects of DNA-peptide crosslinks on transcription in human cells. We constructed plasmid molecules containing DPCs between C5 of dC and lysine sidechains of polypeptides in order to mimic conjugates that form endogenously at DNA epigenetic marks (5-formylc-dC). Lesion bearing and control plasmids were transfected into human cells, and the amounts of RNA transcripts were determined using a mass spectrometry based approach. Moreover, DNA lesion bearing plasmid models were used to determine the importance of NER pathway in DPC repair. In **Chapter 4**, we investigated *in vivo* formation of DPCs in cells exposed to monofunctional alkylating agent, methyl methanesulfonate (MMS). A mass spectrometry-based TMT proteomics approach was used to characterize MMS-

induced DNA-protein cross-linking in Chinese hamster lung fibroblasts (V79). utilizing Our results revealed that DPCs can be produced via nucleophilic attack of proteins at the C8 position of N7-methylguanine (MdG). Our results revealed novel DPC formation mechanisms and the toxicities of monofunctional agent induced DPCs.

In summary, mass spectrometry-based quantification was used to the amounts of free radical induced DPCs in cells, providing evidence for the role of DPC proteolysis in repair, while CTAB assay demonstrated the effect of endogenously formed DPCs on transcription. Moreover, a mass spectrometry-based methodology was applied to examine a novel DPC formation mechanism following treatment with monofunctional alkylating agents.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF SCHEMES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACN	acetonitrile
AP	abasic site
APE1	apurinic endonuclease 1
BER	base excision repair
dA	Deoxyadenosine
dC	Deoxycytidine
dG	Deoxyguanosine
dL	Deoxyribonolactone
dT	Thymidine
DNMT	DNA methyltransferase
DPC	DNA-protein Crosslinks
dR	2'-deoxyribose sugar
DSB	double strand breaks
dT-Tyr	2-amino-3-(4-hydroxy-3-((1-((2R,4R,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)methyl)phenyl)propanoic acid
FA	Formaldehyde
Fapy-Gua	2,6-diamino-4-hydroxy-5-formamidopyrimidine

FPG	formamidopyrimidine glycosylase
H_2O_2	hydrogen peroxide
HR	homologous recombination
IR	ionizing radiation
MMS	Methyl methanesulfonate
MdG	N7-methylguanine
N3-MeA	3-methlyladenine
$NaBH_4$	sodium borohydride
NCP	nucleosome core particle
NER	nucleotide excision repair
NHEJ	non-homologous end joining
NM	nitrogen mustards
$\bullet O_2^-$	superoxide
•OH	hydroxyl radical
Ogg	8-oxo-guanine DNA glycosylase
PM	phosphoramide mustard
ROS	reactive Oxygen Species
SSB	single-strand break
SOD	superoxide dismutase

SPRTN	Spartan
TC-NER	transcription-coupled NER
TDP	tyrosyl-DNA phosphodiesterase
TEAA	Triethylamine Acetate
TET	ten-eleven translocation
TOP	topoisomerase
TLS	trans-lesion synthesis
Thy-Gly	5,6-dihydroxy-5,6-dihydrothymine
UV	ultraviolet light
5fC	5-formylcytosine
5caC	5-carboxylcytosine
5′-dRP	abasic 2-deoxyribose-5-phosphate
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
8-oxo-dG	8-oxo-7,8-dihydroguanine

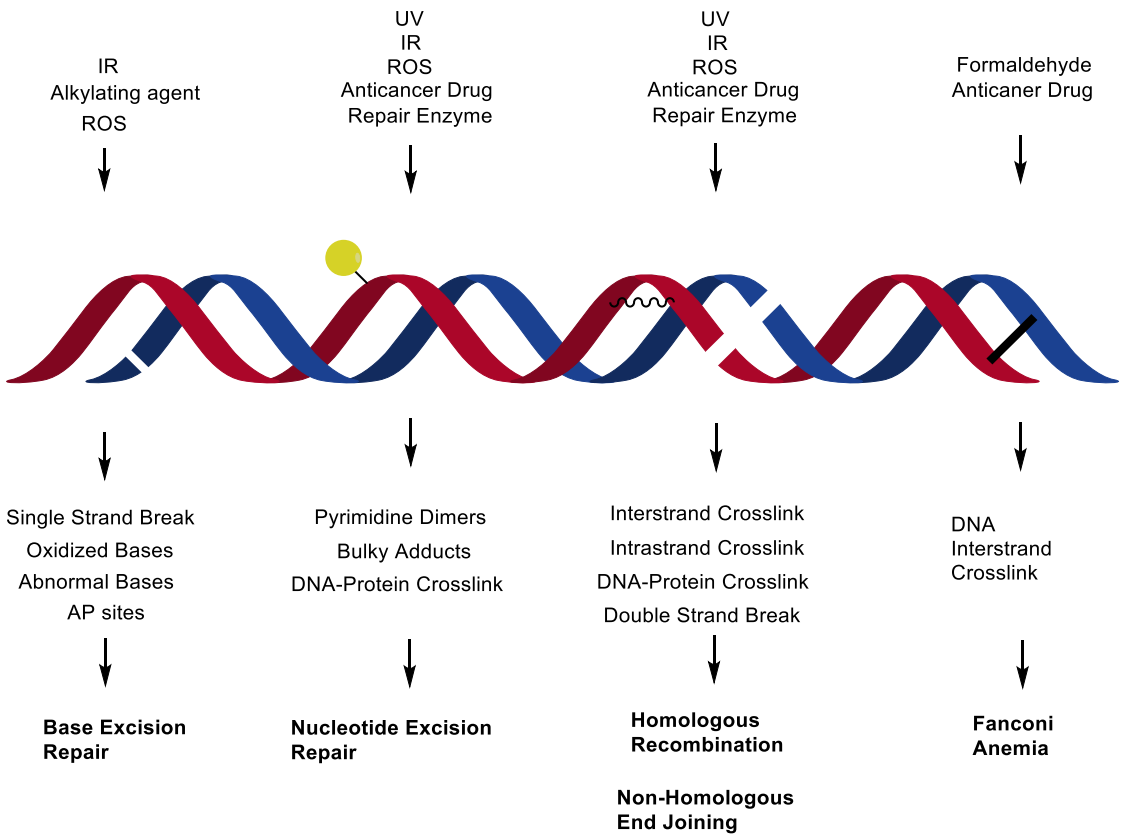
I. Introduction

1.1. DNA Damage and Its Biological Influence

Living cells store genetic information in the form of DNA. However, DNA is continuously damaged by various endogenous and exogenous agents. DNA damage can be caused by many exogenous agents, including ultraviolet light (UV)[1], ionizing radiation (IR)[2], reactive metabolites[3–5], and anti-cancer drugs[6], as well as many endogenous agents, such as aldehydes formed from lipid peroxidation[5,7] and reactive oxygen species (ROS)[8]. When these agents react with DNA, it can induce multiple types of DNA damages such as nucleobase monoadducts, DNA-intra/inter Crosslinks, DNA-protein Crosslinks (DPC), and strand breaks[9,10] **(Figure 1.1.)**.

DNA damage can lead to a loss of genetic integrity by inducing errors during DNA replication, resulting in mutagenesis, cancer, or apoptosis[11]. Thus, DNA adducts should be recognized and repaired by cellular pathways[11–14]. Because these different types of DNA lesions are generated in cells[15–18], it is important to understand the biological consequences that can be caused by these lesions[8].

Figure 1.1. Type of DNA Damage and the Known DNA Repair Pathways



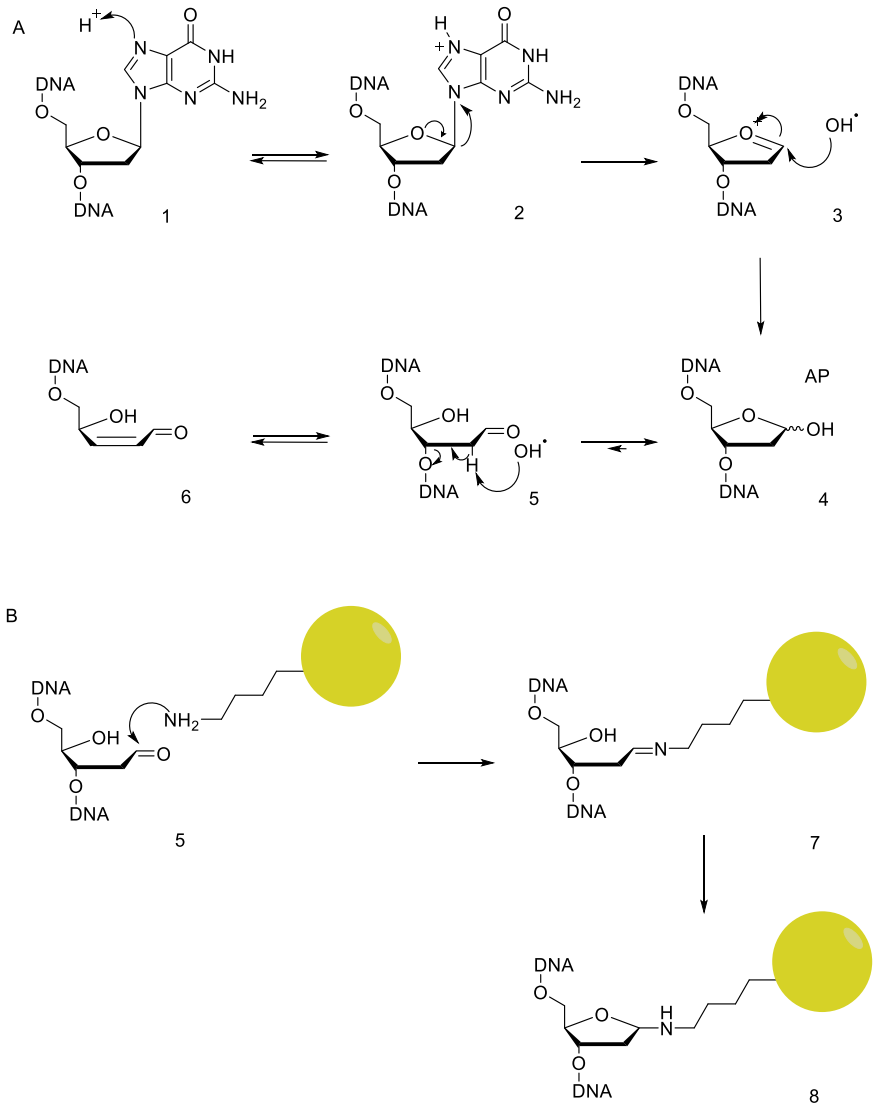
1.2. DNA-Protein Crosslinks

Irreversible (covalent) trapping of proteins on DNA generates unusually bulky DNA lesions called DNA-Protein Crosslinks (DPCs). DPCs can be created by many different mechanisms and involve hundreds of protein and many positions of DNA, leading to a wide range of structurally diverse lesions. For example, DPCs can involve abasis sites in DNA. When protonation of the N7-positions of purines stimulates the cleavage of the glycosidic bond to yield a free base and an oxycarbenium ion, the latter undergoes hydrolysis to yield an abasic site (AP site)[19–21] **(Scheme 1.1.).** More than 10,000 AP sites are generated in a cell per day via depurination[19], and the AP sites engage in secondary reactions with nearby bases in the DNA or amino acids in the protein, resulting additional DNA damages. Abstraction of a proton adjacent to the aldehyde group leads to β-elimination, inducing single strand breaks[21,22]. Alternatively, the open form of the abasic site reacts lysine residues of histones to form DPCs via Schiff base formation **(Scheme 1.1.)**.

Due to their bulky nature, DPCs are known to block replication, transcription, and repair[23–26]. Nakano *et al*. demonstrated that DPCs with a size over 14.1 kDa completely block helicases, while smaller DPCs can be bypassed with low efficiency[24]. DPCs also interfere with key DNA-protein interactions required for replication and transcription[27]. Recent studies suggest that DPC formation may contribute to human diseases, such as cancer, heart disease, and aging problems[28–30]. Therefore, several methods, including mass spectrometry, were used to investigate DPC formation in cells, revealing that hundreds of proteins are capable of forming crosslinks with DNA[31–33]. However, because different cross-linking agents induce DPC through distinct mechanisms[17], the structures

and biological effects of DPCs has not been fully elucidated. Moreover, there are controversial opinions on the DPC repair mechanisms[34]. An improved understanding of the mechanisms of DPCs formation and their repair pathways is needed to comprehend the biological consequences of DPCs.

Scheme 1.1. Depurination of N7-Alkylated Guanine and Subsequent Reactions of Abasic Sites Leading to DPC formation A) Protonation of a purine cleavages the glycosidic bond to yield a free base and an oxycarbenium ion (3), the resulting ion undergoes hydrolysis to yield an abasic site (4, AP site). (B) a lysine residue of histone (yellow circle) reacts with the aldehyde group (5) to form DPC (8) via Schiff base formation.



1.2.1. Types of DPC Formation in Cells.

DPCs can be classified into two groups; enzymatic and non-enzymatic DPCs. Enzymatic DPCs are generated when intermediates of enzyme activities are trapped on their DNA substrate. Topoisomerases and DNA polymerases are the good examples of proteins that commonly form enzymatic DPCs [35,36]. Non-enzymatic DPCs are formed when endogenous and exogenous agents induce non-specific crosslinking of any protein in the vicinity of DNA. For example, cisplatin, a common anticancer agent, induces non-enzymatic DPCs[37].

1.2.2. DPC Formation in Cells.

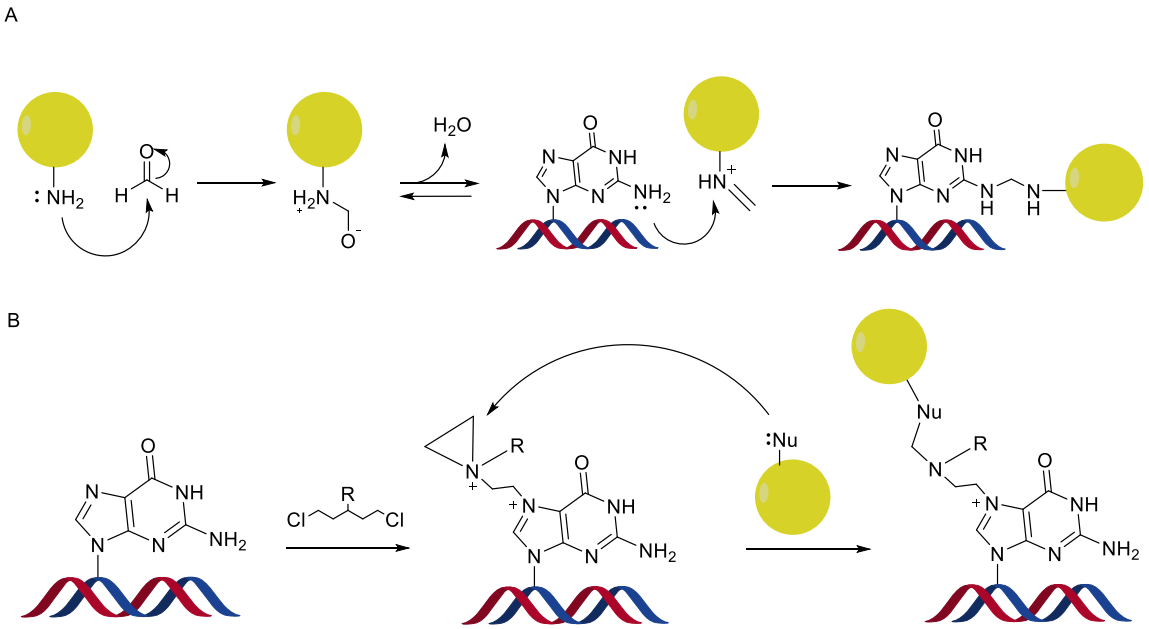
Formaldehyde (FA) induced DPCs are one of the most extensively studied DPC inducing agents **(Figure 1.2.)**. Humans are exposed to FA in industrial and household environments due to consumer products[38,39], and it is also produced in cells endogenously as a product of methylamine deamination[40] and histone demethylation[41,42]. It has been shown that FA induces several types of DNA lesions resulting point mutations and gene deletions[38]. Interestingly, FA-mediated DPCs are formed between lysine, histidine, tryptophan, and cysteine residues of proteins and dG, dA, and dC nucleosides of DNA[43,44,45]. FA induced histone-DNA crosslinks may play a role FA in tumorigenesis and carcinogenesis[30,46].

DPCs can also be induced by exogenous bifunctional alkylating agents such as nitrogen mustards (NM). NM, a type of anticancer drugs, are used in chemotherapy regimens for lymphoma, leukemia, myeloma, breast and brain cancers[47]. NMs can form reactive aziridinium ions which alkylate nucleophilic positions on DNA and proteins to

form DPCs **(Figure 1.2.)**. A dose-dependent increase of DPCs has been detected in human HT1080 cells treated with 0 - 500 μM of phosphoramide mustard (PM)[48]. Interestingly, Hansson *et al*. showed that 60 – 70% of total cross-linked lesions were identified as DPCs via alkaline elution analyses[49]. Work in our laboratory has revealed that NM-induced DPCs can form between the N7 position of guanine and cysteine residues of proteins[50–52], whereas mass spectrometry-based proteomic studies revealed that the crosslinked proteins play a role in chromatin organization, DNA repair, replication and transcription[48].

In summary, DPC lesions are structurally complex and may occur at multiple sites of proteins and DNA. Because of their unusually large size and their ability to block replication and transcription, cellular DPCs must be efficiently removed to avoid mutagenicity and cell death.

Figure 1.2. Mechanism of DPC formation by FA (A) and NM (B)



1.3. DPC-inducing Agents Studied in This Thesis

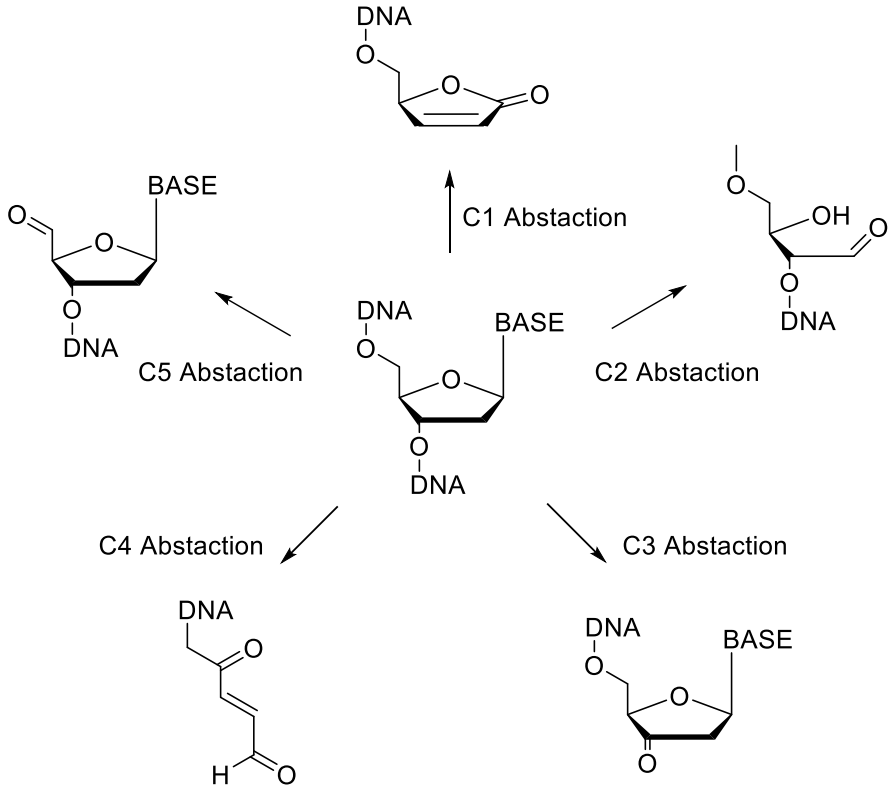
1.3.1. Reactive Oxygen Species (ROS)

Reactive oxygen species (ROS) are radical and non-radical oxygen species formed from the partial reduction of oxygen. Superoxide anion ($O_2^{\bullet -}$), hydroxyl radical (•OH), hydroperoxyl radial (HO_2•), alkoxyl radicals (RO•), and peroxyl radicals (ROO•) are known as oxygen radicals, whereas singlet oxygen ($^1O_2$), hydrogen peroxide (H_2O_2), peroxynitrous acid (ONOOH), hypochlorous acid (HOCl), hypobromous acid (HOBr), hypothiocyanous acid (HOSCN) and organic peroxide (ROOH) are classified as non-radical forms **(Table 1.1.)**. Due to their intrinsic instability, ROS can react with DNA, affecting the integrity of DNA in the cell[53]. For example, all five carbon positions of the 2'-deoxyribose sugar (dR) can undergo proton abstraction by •OH to induce direct strand scission and electrophilic moieties[54–56] **(Figure 1.3.)**. Moreover, hydroxyl radical can be added to the double bonds of DNA bases[57]. •OH addition to C8 of the guanine bases generates 8-hydroxy-7,8-dihydroguan-8-yl radicalswhich undergo further reactions to form 8-oxo-7,8-dihydroguanine (8-oxo-dG) and 2,6-diamino-4-hydroxy-5-formamidopyrimidine (Fapy-Gua) **(Scheme 1.2.)**. Similarly, •OH can be added to C5 and C6 of thymine, making C5-yl and C6-yl radicals, respectively[58]. O_2 addition to the radical sites yields thymine hydroperoxides, which will undergo subsequent reactions to form the 5,6-dihydroxy-5,6-dihydrothymine (Thy-Gly)[59–61] **(Scheme 1.2.)**.

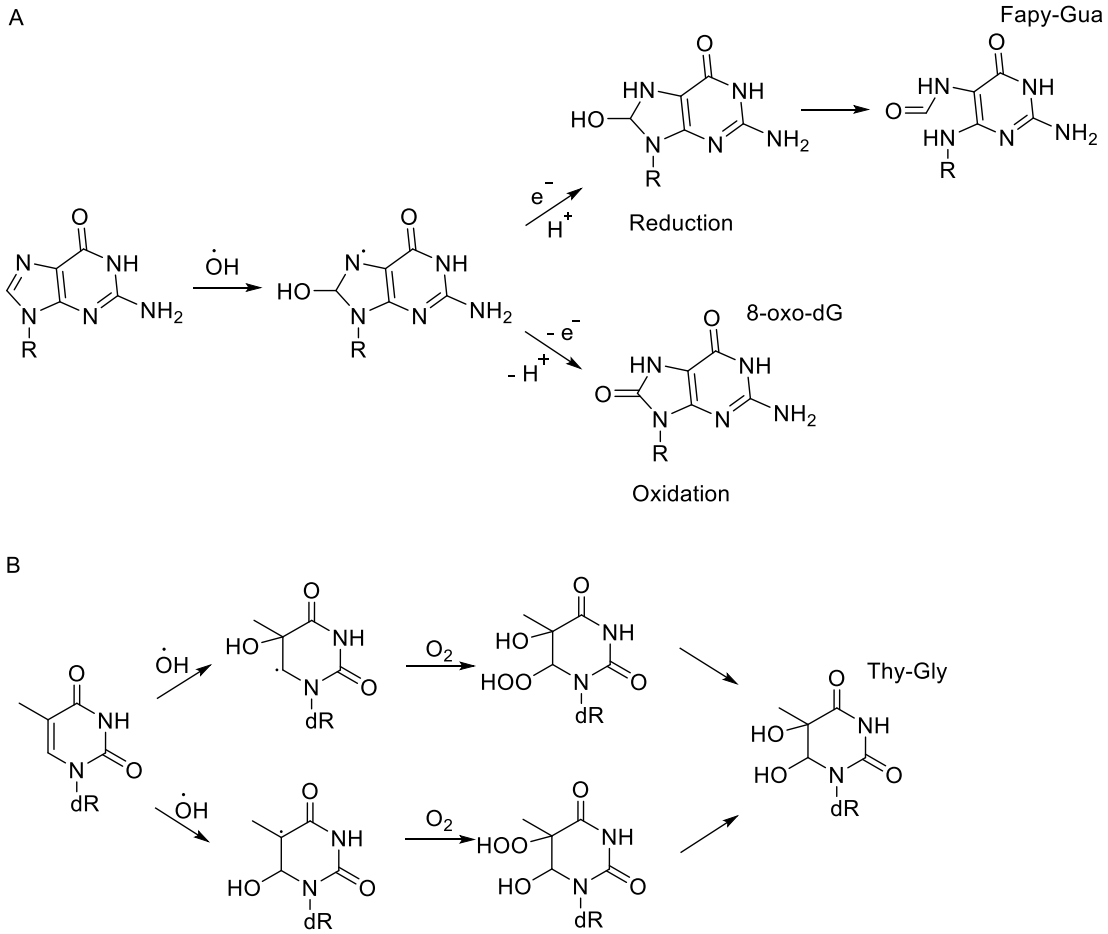
Table 1.1. Reactive Oxygen Species

Reactive Oxygen Species			
Oxygen Free radicals		**Organic Free Radicals**	
$^1O_2$	Singlet oxygen	ROO•	Peroxyl radical
$•O_2–$	Superoxide anion	RO•	Alkoxy radical
•OH	Hydroxyl radical		
$•HO_2$	Hydroperoxyl radical		
Non-radical Oxidant		**Others**	
H_2O_2	Hydrogen peroxide	HOCl	Hypochlorous acid
ONOOH	Peroxynitrous acid	HOBr	Hypobromous acid
ROOH	Peroxide	HOSCN	Hypothiocyanous acid

Figure 1.3. DNA Adducts Produced by Hydrogen Abstraction from 2'-Deoxyribose Sugar of DNA. Hydroxyl radical can abstraction hydrogen from the different position of 2-deoxyribose sugar to yield different electrophilic lesions, inducing strand scission and crosslinks.



Scheme 1.2. Mechanism of ROS-induced Imidazole Ring Opening in Purine. A) Hydroxyl radical can attack the C8 position of guanine to produce 8-hydroxy-7,8-dihydroguan-8-yl radicals, which give rise to 8-oxo-7,8-dihydroguanine (8-oxo-dG) via oxidation and 2,6-diamino-4-hydroxy-5-formamidopyrimidine (Fapy-Gua) via reduction. B) Hydroxyl radical can attack C5 and C6 of thymine to yield radicals, which can generate thymine hydroperoxides and thymine glycol (Thy-Gly).



1.3.1.1. Endogenous Formation of Reactive Oxygen Species

Mitochondrial respiration is known as the major source of ROS in cells[62,63]. Mitochondria reduce oxygen to form water molecules through the Electron Transport Chain (ETC). The ETC consists of four complexes - Complex I (NADH ubiquinone oxidoreductase), Complex II (succinate dehydrogenase), Complex III (ubiquinol cytochrome c oxidoreductase) and Complex IV (cytochrome c oxidase), as well as additional proteins including coenzyme Q and cytochrome c. Complex I and II accept electrons from NADH and succinate respectively, and these electrons are transferred to Complex III via coenzyme Q. Later, Complex III transfers electron to Complex IV via cytochrome c to reduce oxygen and to produce water molecules.

However, during the ETC cycle, some portion of oxygen can be partially reduced to superoxide ($O_2^{\cdot -}$), leading to unexpected reactions. For example, when coenzyme Q receives electrons from Complex I and II, coenzyme Q is reduced into the QH_2 form via the formation of an unstable intermediate semiquinone anion ($^{\cdot}Q^-$). This semiquinone anion transfers electrons to oxygens, and $O_2^{\cdot -}$ is released into both the mitochondrial matrix and the intermembrane space. The resulting superoxide anion is converted to hydrogen peroxide (H_2O_2) by the mitochondrial superoxide dismutase (SOD) and produces hydroxyl radicals which can damage DNA molecules and hinder normal functions of the cell [56,64].

Peroxisome also has ability to transfer the electrons from various metabolite to the oxygen, whereas beta-oxidation of fatty acids is used as the source to produce H_2O_2 [65].

Several enzymes, including amino acid oxidase and xanthine oxidase, has been shown to produce different ROS[66]. Moreover, ROS can be produced during immune response and inflammation[57,67,68]. For example, macrophage activation increases O_2 uptake, which can lead to a rise of various ROS[69].

1.3.1.2. Exogenously Induced Formation of Reactive Oxygen Species

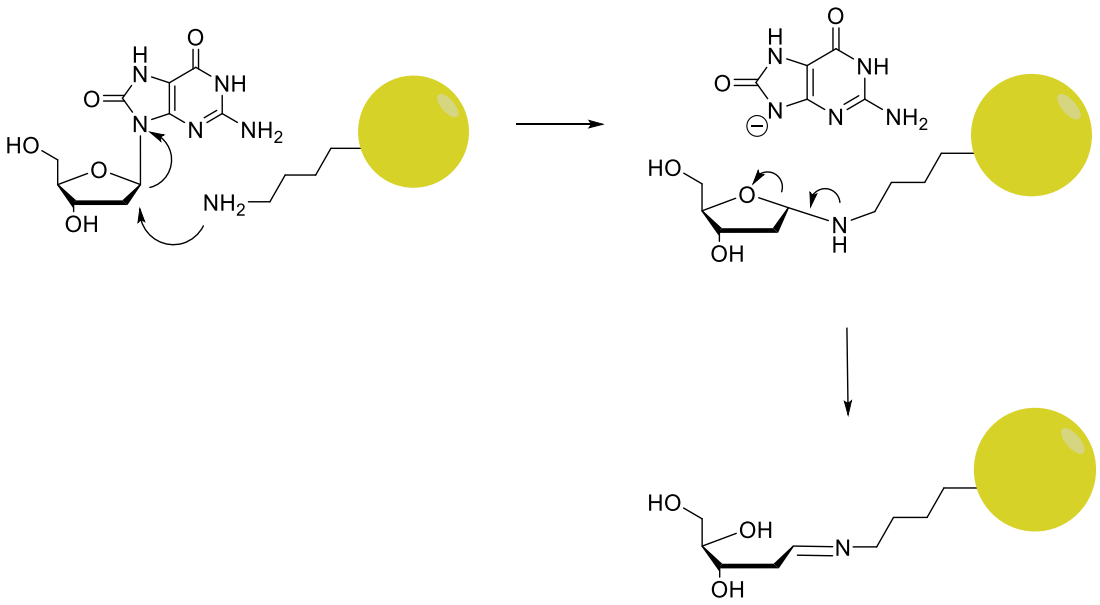
ROS can also be produced by exogenous agents, such as IR and UV irradiation. Gamma radiation can excite cellular water molecules, generating ROS[70]; the water molecules are sequentially converted to hydroxyl radicals, superoxide radicals, and finally hydrogen peroxide. Moreover, UV exposure can indirectly induce ROS generation through photosensitization. When photosensitizers (nucleic acids, aromatic amino acids, NADH, NADPH etc.) absorb energy from UV, the excited photosensitizer interacts with a target biomolecule, such as oxygen, to yield singlet oxygen and superoxide anions[71]. In addition, recent studies have shown that ROS regeneration can be initiated by cigarette smoke and urban air[72,73], as well as during alcohol consumption[74,75].

1.3.1.3. ROS induced DNA-Protein Crosslinks

When a nucleophilic site of a protein reacts with a ROS-induced DNA lesion, a DPC can be generated though secondary reactions[76,77]. For example, when a hydroxyl radical is added to the C8' position of guanine, it forms 8-oxo-dG and Fapy-Gua[60,61]. The one-electron oxidation of the radical creates 8-oxo-dG, while the reduction generates Fapy-Gua by opening the imidazole ring[18,78]. Interestingly, the C1' position of 8-oxo-Gua can make a covalent bond with the N-terminal proline of *E. coli* formamidopyrimidine

glycosylase (FPG) during base excision repair[79]. Similarly, 8-oxo-guanine DNA glycosylase (Ogg) can be crosslinked to DNA through its lysine residues **(Scheme 1.3.)** [80]. Furthermore, when a free radical attacks the 5'-methyl group of thymidine to yield 5-(uracilyl)methyl radical, the C3' position of a neighboring tyrosine can react with the radical, generating thymidine-tyrosine (dT-Tyr) cross-links[81,82]. This structure was used to investigate the role of proteolysis in the repair of DPCs in mouse and human cells in Chapter 2.

Scheme 1.3. Mechanism of 8-oxo-dG mediated DPC Formation. 8-oxo-guanine DNA glycosylase (Ogg) displaces 8-oxo-dG and forms a reversible crosslink with DNA through its lysine residues.



1.3.2. Ionizing Radiation

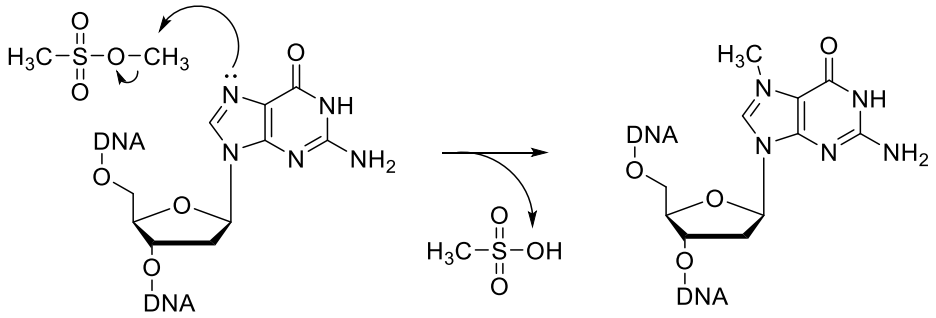
DNA can be damaged by IR via direct and indirect mechanisms. When DNA is exposed to IR, electrons can be directly removed from DNA bases to form radical cations, which undergo decomposition. Alternatively, water molecules can be excited to produce reactive oxygen species such as •OH, which attack DNA and generate DNA radials, followed by further reactions and decomposition. DNA lesions produced by IR are structurally identical to those formed by ROS[58,83,84]. Similar to DNA, radiation energy induces protein radicals, which decay by themselves. However, these DNA and protein radicals undergo secondary reactions alternatively with the proximal constituents of proteins and DNA, inducing DPC[85].

1.3.3. Methyl Methanesulfonate

Methyl methanesulfonate (MMS), a model compound for methylating agents, is a powerful genotoxic agent which produces a wide spectrum of DNA methyl adducts. MMS methylates nitrogen atoms of DNA to form a range of methylated lesions. For example, MMS modifies both guanine and adenine to produce N7-methylguanine (N7-MeG, or MdG) and 3-methlyladenine (N3-MeA), comprising 82% and 11% of all DNA adducts, respectively **(Figure 1.4.)** [86,87]. These lesions are known to cause base mispairing and replication blocks[86]. Cells use the base excision repair (BER) pathway to repair alkylating agent-induced DNA damage[88]. Specifically, *N*-methylpurine DNA glycosylase removes N7-MeG, N3-MeA, and N3-MeG [89], and apurinic endonuclease 1 (APE1) removes abasic (AP) sites, leading to the formation of single strand breaks[90,91]. Moreover, the cytotoxicity of MMS-induced damage increases in cells where other DNA

repair pathways are compromised. Specifically, DNA double strand breaks (DSB) were increased in homologous recombination (HR) deficient cells[92]. Therefore, MMS was known as DNA double-strand break (DSB) agent for many years, including in some recent publications. However, it is now believed that MMS stalls replication forks, and HR is important for repair of MMS damaged replication forks[93]. Interestingly, Yang *et al.* showed that monofunctional agents also have abilities to induce DPC[94]. Therefore, MMS induced DPC formations *in vivo* have been investigated in the Chapter 4 of the thesis.

Figure 1.4. Mechanism of MdG Formation by MMS MMS attacks the N7 of guanine to produce a methyl adduct.

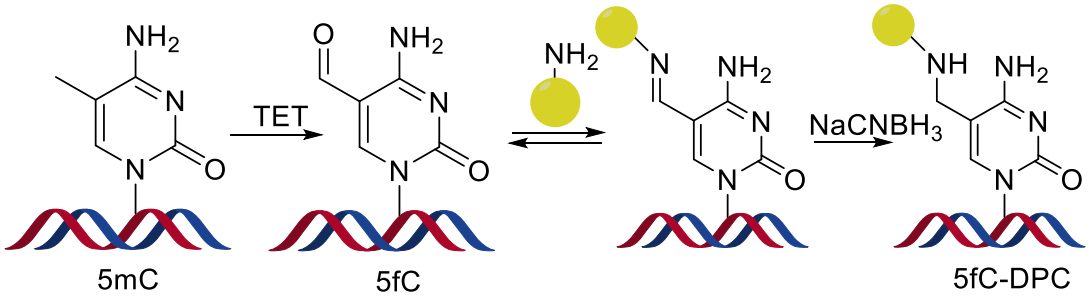


1.3.4. 5-Formylcytosine Epigenetic Marks

DNA methylation occur at the C5' position of cytosine in DNA. DNA methyltransferases (DNMT) methylate the C-5 position of cytosine residues that are adjacent to a guanine nucleotide (CG sites) to produce two methylated cytosine residues (5-methylcytosine, 5mC) positioning diagonally on opposite DNA strands[95,96]. Furthermore, that it has been reported that different members of the DNMT family have distinct functions. For example, DNMT3a and DNMT3b are responsible for de novo methylation of DNA, whereas DNMT1 maintains DNA methylation during semiconservative replication[97,98]. DNA methylation has been shown to play an important role in gene expression by blocking transcription factors from binding to the promoter region of a gene and recruiting chromatin remodeling machinery. Suzuki *et al.* showed that higher level of methylation in a promotor region causes lower transcription[99]. However, the actual mechanism of how the methylation regulates gene expression is still unknown. Recent studies have demonstrated that 5mC can be converted to other epigenetic marks. Ten eleven translocation dioxygenases (TET1-3) mediate the iterative oxidation of 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC)[100]. Because the concentration of 5hmC is dramatically decreased in tumors [101,102], it has been suggested that oxidized forms of 5mC play an important role in gene expression. Among the oxidized forms, 5fC may have the ability to regulate gene expression due to the presence of an aldehyde group in its structure. This aldehyde group can react with cellular nucleophiles such as amino acids, polypeptides, and proteins. Moreover, significantly greater numbers of nuclear proteins have higher affinity for 5fC than 5mC, 5hmC and 5caC[103,104]. Interestingly, the Tretyakova group found that histone

proteins, such as H2A and H4, can make a reversible conjugate with 5fC-containing DNA via Schiff base; proteins containing Lys and Arg residues can form crosslinks to the formyl group of 5fC **(Scheme 1.4.)**. Therefore, reversible conjugation between regulatory proteins and 5fC in DNA may play a role in chromatin dynamics and epigenetic regulation. Conversely, DNA replication, transcription, and repair can be blocked due to 5fC mediated DPCs[9,105,106].

Scheme 1.4. Schematic of 5fC mediated DPC Formation Once 5mC is modified to 5fC, it reacts with a nucleophilic residue of protein to make a DPC via Schiff base formation which can be stabilized after reduction.



1.4. DPC Repair Pathways

1.4.1. Significance of DNA-Protein Crosslink Repair

The study of DPC repair will improve the understanding of the role that endogenous DPCs play in human disease. An increased level of DPC lesions was detected in cancer[107–109] and appears to contribute to aging problems, such as Alzheimer's and Parkinson's diseases.[28,110]. Furthermore, the Tretyakova group has identified an increase in the levels of DPCs in cardiomyocytes from a rat model of ischemia/reperfusion injury[111]. Therefore, it is possible that deficient DPC repair can play a role in these common human diseases.

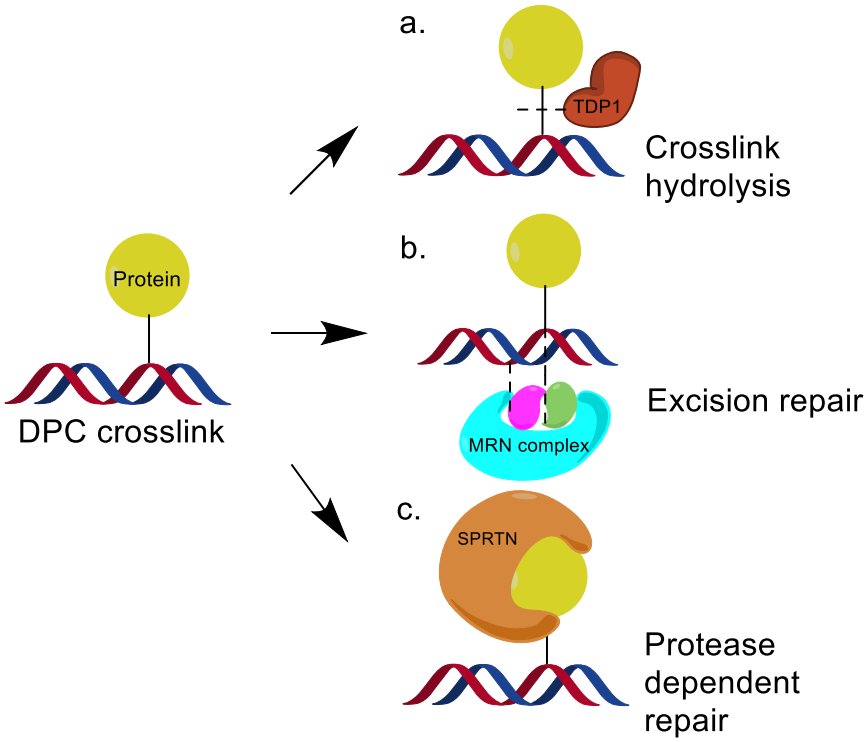
Most classical anticancer chemotherapeutics induce the formation of bulky DNA adducts, such as DNA-DNA and DNA-Protein crosslinks (DPC). These types of DNA damage are hypothesized to block major DNA-dependent cellular pathways, including transcription and replication[9]. In order to increase the therapeutic efficiency of antitumor drugs and to understand inter-individual differences in response to therapy, it is crucial to elucidate cellular mechanisms triggered in response to these bulky DNA lesions.

1.4.2. DPC Repair by Canonical Repair Pathways

There are three general tactics that can be applied in order to remove DPCs **(Scheme 1.5.)**: 1) breaking the DNA-protein covalent bond, 2) removing and replacing DNA regions containing DPCs and 3) degrading crosslinked proteins to peptide fragments that will allow bypass at the replication fork. Topoisomerase I and II (TOP1 and TOP2) and their Tyrosyl-DNA phosphodiesterases (TDP1 and TDP2) are clear examples of direct

crosslink hydrolysis activities; the recruitment of TDP1 and TDP2 hydrolyzes the phosphodiesterase bond between DNA and tyrosine residues of TOP1 and TOP2, respectively[112]. However, targeting the DNA-protein bond has limitations in the interface accessibility and the requirement of specialized repair enzymes to remove each specific type of DPC. Also, this mechanism appears to be limited to DPCs involving topoisomerase-DNA intermediates. To overcome these limitations, cell must utilize more general repair mechanisms to deal with other types of DPC lesions.

Scheme 1.5. Three strategies to Remove DPC A) if DNA crosslinked protein is topoisomerase, tyrosyl-DNA phosphodiesterase (TDP) can remove DPC by breaking covalent bond through hydrolysis. B) MRN complex cleave DPC by removing DNA sequence nearby through endonuclease activity. The resulting shorter DNA without DPC is repaired by HR or NHEJ. C) SPRTN function as protease to degrade crosslinked proteins into short peptides that bypass the replication fork.



Recent studies suggested that nucleotide excision repair (NER) and homologous recombination (HR) orchestrate to repair DPCs and maintain genomic integrity[113]. Work by De Graaf *et al.* demonstrated that NER can repair FA-induced DPCs[114], while Stingele *et al.* showed that NER is dominant in DPC repair in yeast[115]. Moreover, several studies have demonstrated that mammalian NER can remove DPCs *in vitro* and *in vivo*[116–119]. Because NER only excises a region of damaged DNA strand containing a bulky lesion, DNA polymerases and ligases must be recruited to the complementary stand to repair the excised regions. However, NER is known to repair DPC sites with a protein size of 10 kDa or less[116], and is not available during replication. Therefore, proteases are required to degrade the crosslinked proteins into smaller fragments when the replication forks encounter a DPC site. Interestingly, another lesion tolerance mechanism called trans-lesion synthesis (TLS), which recruits TLS polymerases that tolerate DNA lesions, is available in cells [120,121]. This process permits replicative bypass of DNA lesion so that canonical DNA repair pathways can repair them at a later stage[122]. However, TLS polymerases cannot accommodate large DPCs in their active site and require DPC proteolysis.

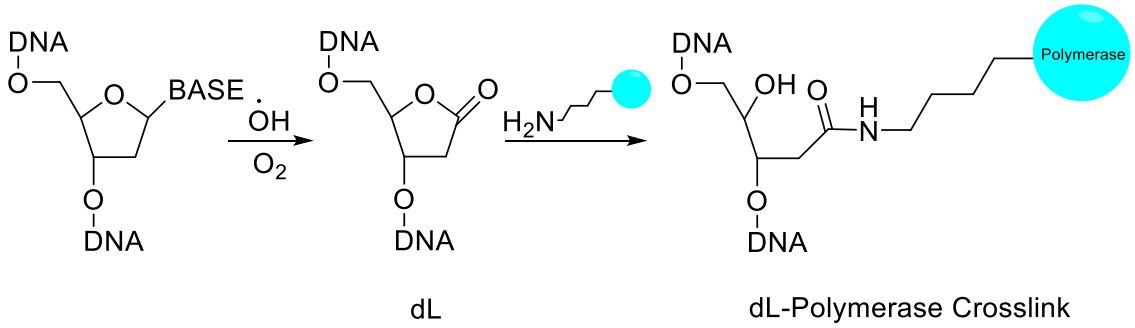
Double strand break (DSB) can be generated by DPCs, leaving the DPC at the end of a DSB strand. When a DPC is present at the 5' end or 3' end of dsDNA, the MRE11-Rad50-NBS1(MRN) complex activates its intrinsic endonucleolytic incision mechanisms to remove the DPC by cleaving several base pairs at the end of dsDNA[123,124]. For example, the MRN complex cleaves 15-20bp downstream of TOP2 cleavage complex (TOP2cc)[123,125]. The resulting 'shorter' dsDNA without a DPC can be repaired by HR. However, undamaged complementary strands must be recruited to recover the lost

sequence information. Furthermore, the involvement of HR in DPC repair has not been fully elucidated. Although FA-induced damages were more cytotoxic in HR-deficient E.coli[117], other pathways including Fanconi anemia can be activated to response to the various types of FA-induced DNA damage, not DPCs alone. Interestingly, Chesner *et al*. showed that HR is crucial to repair DPCs in mammalian cells through a quantitative PCR-based assay, by transfecting DPC containing plasmids[126]. Therefore, further studies are required to determine whether double-strand break repair pathways such as the Fanconi anemia pathway, non-homologous end-joining, and microhomology mediated end-joining are involved in DPC repair pathways.

1.4.3. Endogenous DNA Damage by Attempted DNA repair

The eleven DNA glycosylases in mammalian cells recognize DNA lesions and initiate BER[127], leaving an AP site in the DNA[128]. The resulting AP sites are incised by an AP endonuclease[129,130] to generate a normal 3′ OH that can be used as a primer by DNA polymerase, and a 5′ end bearing the abasic 2-deoxyribose-5-phosphate (5′-dRP). During Short Patch BER, Polβ is recruited to 5′-dRP[131] and form a covalent linkage to the DNA during its lyase reaction. However, enzyme attack on some lesions form a stable covalent link to the DNA[132]. For example, oxidation at the C1’ position of the 2’-deoxyribose sugar can lead to the formation of deoxyribonolactone (dL), thus leading to DPC formation with endonuclease III (Endo III) and DNA polymerase β (Pol β)[133,134] **(Figure 1.5.).** Moreover, this has also been shown to obstruct BER steps[132]. Therefore, attempted DNA repair on some lesions often results in further damage[135–137].

Figure 1.5. DNA Damage by Attempted DNA repair Hydroxyl radicals induce the formation of dL in DNA. DNA polymerase form a stable amide linkage to the lesion during its enzymatic activity.



dL

dL-Polymerase Crosslink

1.4.4. Novel DPC-Proteolysis Repair

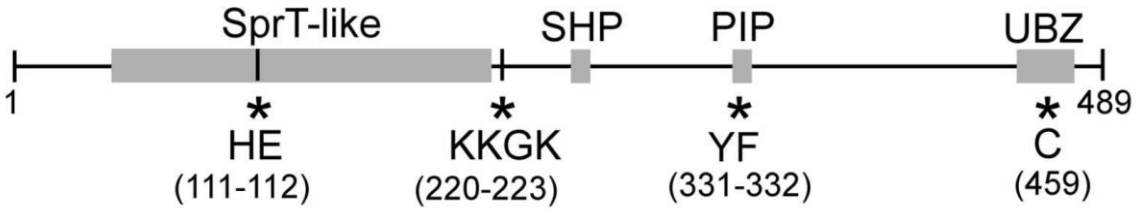
1.4.4.1. Spartan DPC Protease in DPC Repair

Recent studies suggest that Spartan (SPRTN) protease plays a critical role in DPC repair pathways. The novel proteolytic DPC repair pathway was found in yeast initially as Wss1 replaced the role of TDP1 in formaldehyde-induced DPC repair in tdp1 deficient cells[138]. This finding provided a clue which helped identify the role of SPRTN-mediated proteolysis during DPC repair, since Wss1 and SPRTN show similarities in their structures[139]. Several studies demonstrated that SPRTN knockout is embryonic lethal and showed incomplete DNA replication, formation of micronuclei and chromatin bridges, leading to the cell death[109,140]. Moreover, it has been reported that SPRTN mutations cause Ruijs-Aalfs syndrome (RJALS)[109,141], and patients with SPRTN mutation suffer hepatocellular carcinoma at young age due to the high level of DPCs in their liver[108,109,140,142]. Therefore, by elucidating the role of SPRTN, the role of DPC repair in common human diseases can be better understood.

SPRTN contains several distinct motifs; the putative zinc metalloprotease domain (SprT), DNA binding domain (KKGK box), p97 interaction domain (SHP), PCNA interacting domain (PIP), and C terminal ubiquitin-interacting domain (UBZ)[143] **(Scheme 1.6.)**. SPRTN is known to regulate DNA damage response by assisting p97 to stalled replication forks[144] and is recruited to damaged DNA by ubiquitinated PCNA and RAD18[145]. Interestingly, the N-terminal SprT domain of SPRTN, which contains a HexxH motif, negatively regulates TLS by changing the interaction of POLD3, a subunit of replicative polymerase with Rev1 and pol ζ[146]. This finding implies that SPRTN

interacts with replicative polymerase and that this regulation may affect mutagenesis. Moreover, since Rev1 interacts with Rev7 directly to regulate pol ζ[147], it is suggested that SPRTN may play a critical role in the negative regulation of pol δ / pol ζ switching. The importance of SPRTN-TLS polymerase interactions has also been reported by showing that SPRTN can help pol η target DNA damage sites[140,145,148–150]. However, the connection between DNA binding and Ub-PCNA binding by SPRTN-involved-pathways, and the role of SPRTN during polymerase switching mechanisms are unclear and thus should be investigated further[109,140,145].

Scheme 1.6. Domain Structure of Homosapien SPRTN SPRTN contain SprT, SHP, PIP, and UBZ domain[144].



Domain in SPRTN	
SprT	Zinc metalloprotease domain
HE	HExxH motif
KKGK	KKGK box for DNA binding
SHP	p97 interaction domain
PIP	PCNA binding domain
UBZ	Ubiquitin binding domain

SPRTN has high binding affinity to DNA, with a K_d value of 100 nM[151], and the KKGK box in SPRTN N-terminal regulates DNA-binding activity[148]. A mutation study on this site displayed a decrease in the amount of chromatin-bound SPRTN and an increase in the amount of non-chromatin-bound SPRTN, indicating its contribution to DNA anchoring. Stingele *et al.* and Vaz *et al*. showed that SPRTN removes DPCs as a protease and its proteolysis function is activated when DNA or DNA crosslinked proteins are present[151,152]. Furthermore, *Haracska et al.* showed that SPRTN has a DNA-dependent protease and self-degrading activities by comparing WT to a KKGK box mutant[143]. To be more specific, the study found that SPRTN cleaves several DNA binding proteins such as Fan1, HLTF, and yRad5, while RFC and RPA bind to single-stranded DNA to inhibit the protease activities of SPRTN; these results suggest that SPRTN-single-stranded DNA binding may function as a proteolytic DPC repair rate-limiting step. Moreover, several assays have been developed to measure the degree of DPC removal in replication *in vivo*[143]. In brief, BrdU comet assay demonstrated that SRPTN knock-down HEK293 cells have a deficiency in DPC repair and that the addition of exogenous SRPTN could rescue the DPC repair. These results highlight the importance of not only DNA-binding but also protease-domain of SPRTN against the DPCs. However, it is still unknown whether DPC proteolysis occurs only in ssDNA or in both ssDNA and dsDNA.

1.4.4.2. Proteasome in DPC Repair

The involvement of proteasome in DPC repair has been previously investigated. However, the exact role of proteasome in DPC pathway is unclear due to contradictory results in recent studies. For example, inhibition of ATP-dependent proteases showed no effect on the toxicity of FA-induced DPCs in bacteria[117], while proteasome inhibition

increased the amount of DPCs[39,153,154]. Moreover, the repair of DPCs on plasmid decreased with MG132 proteasome inhibitor treatment; however, Ide *et al*. demonstrated no difference in the amount of ubiquitinated DPCs in presence or absence of proteasome inhibitor MG132[119,155]. Moreover, Nakano *et al.* showed that DPCs are not marked by polyubiquitination for proteasome recognition, while the crosslinked topoisomerase I were polyubiquitylated[155,156]. Taken together, further studies are required to elucidate the mechanism of DPC proteolysis activity. Recently, Larsen et al. showed that both SPRTN and proteasome can function as two replication-coupled DPC protease[157]. In brief, proteasome recognize ubiquitylated DPC, while SPRTN mediated DPC degradation occur with DPC ubiquitylation during replication.

1.4.4.3. Mechanism of Novel DPC Repair Pathway

The following novel proteolytic DPC repair mechanism can be suggested (Discussed in chapter 2). When a replication fork is stalled due to a DPC, PCNA is ubiquitylated by Rad6-Rad18 complex, followed by the recruitment of SPRTN to that site. Here, the mono-ubiquitylation and exposed single stranded DNA attracts SPRTN through high affinities toward UBZ and PIP domains, and DNA-binding domain (KKGK box), respectively. Once SPRTN is recruited to the site, the conformational change activates DNA-dependent protease function to cleave the crosslinked protein to small peptide or nucleotide scale to bypass the lesion. Moreover, since SPRTN interacts with both pol δ and pol η and regulates the dissociation of pol η through p97, it is believed that the recruitment of TLS polymerase and the exchange of TLS and replicative polymerases are induced by SPRTN to continue replication after bypass. In addition, some subunits of

replicative polymerase may leave the complex with SPRTN so that remaining subunits can interact with newly recruited polymerase such as pol η. Moreover, more research is needed to identify the exact mechanisms which regulate SPRTN and how they prevent unspecific substrate cleavage. Recent findings have shown that DPCs are removed during the S phase of the cell cycle[151], however the involvement of HR and NER during S phase is not fully understood, and thus this, too, should be further investigated. In addition, it would be interesting to explore the SPRTN - pol I, pol k or Rev1 interaction to understand the bypass and the DPC repair because SPRTN is interacting with pol η, a member of TLS polymerase.

II. Role of Proteolysis in Repair of Free Radical-induced DNA-Protein Crosslinks in Mouse Cells and Tissues

Adapted from:

Arnie Groehler IV, Daeyoon Park, Reeja Maskey, Lei Li, Yuichi Machida, and Natalia Y. Tretyakova; Roles of Spartan Metalloprotease and Proteasomal Degradation in Repair of ROS-induced Chromosomal DNA-Protein Cross-links

Manuscript in preparation

Daeyoon Park participated in all experiments described in this chapter.

Arnie Groehler IV performed cytotoxicity of DPC-inducing agents in Spartan proficient and Spartan deficient cell lines and examined the impact of pharmacological inhibition of proteasome. Arnie Groehler IV and Daeyoon Park performed quantitation of dT-Tyr obtained from wild type or SPRTN$^{H/H}$ tissues. Daeyoon Park performed quantitative analysis of radical-induced DPC formation in IR-treated MEF cells, IR-treated mouse samples and Immunoproteasome KO mouse samples. Daeyoon Park also performed clonogenic assay and SPE method development for nanoLC-ESI^{+}-MS/MS.

2.1. Introduction

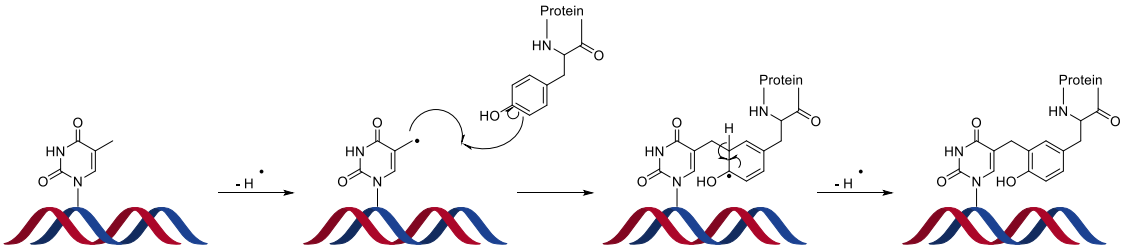
Endogenous and exogenous agents induce various types of DNA damage, including oxidative damage, alkylation, bulky lesions, inter and intrastrand DNA crosslinks, and DNA-protein crosslinks. However, the characteristics and biological outcomes of DNA-protein cross-links (DPCs) have not been fully investigated. DPCs, super-bulky DNA lesions which are formed by the covalent bond between chromosomal DNA and a protein, inhibit DNA replication[23–26], transcription and gene expression[158]. Considering these interferences, it is important to elucidate the mechanisms of DPC recognition and repair. Several studies showed that tyrosyl-DNA phosphodiesterase-1 or -2 (TDP1 and TDP2) can repair DPCs first by proteolysis of the crosslinked topoisomerase to smaller polypeptide fragments and then hydrolysis of the phosphotyrosyl bond[159,160]. However, specialized repair enzymes were needed to repair these topoisomerase DPCs. Thus, a more general repair mechanism must be required to deal with varied types of DPC lesions.

Canonical DNA repair pathways have been implicated in DPC repair. The MRN complex can cleave DNA on both sides of the DPC to complete nuclease dependent repair of DPCs[123,161], and the resulting double-stranded break is repaired by either HR or NHEJ[162–164]. Moreover, NER can cleave the DNA on both sides of the DPC to remove the damaged DNA strand, and the complimentary strand can then be used to repair the excised region[165]. Interestingly, recent studies showed that DNA-peptide crosslinks can be bypassed in error-prone manner by recruiting TLS polymerases during DNA replication[121,166], suggesting that DPC repair is orchestrated by proteolysis performed by

SPRTN[34]. SPRTN degrades crosslinked protein, such as topoisomerases and histones, in the presence of ssDNA[143,151,152], while bypass of a FA-induced DPC was inhibited in SPRTN deficient cells[34]. Due to these evidences, Stingele *et al.*[139,152] and Vaz *et al.*[34,151] have suggested a replication-coupled DPC repair model. In this model, SPRTN is recruited once the replication complex is stalled and then degrades the crosslinked proteins so the resulting DNA-peptide crosslink can be bypassed by TLS polymerases[34,143,167] and repaired by canonical repair pathways[39,119,122]. However, it has not been demonstrated whether ROS-induced DPCs can be repaired by SPRTN in a similar manner. So, the purpose of this study was to determine if SPRTN can cleave ROS-induced DPCs.

When the 5-methyl group of thymidine obtains free radicals from hydroxyl radicals, it forms a stable methylene linkage by reacting with the 3-position of tyrosine[168]. Once hydrogen is released to make an aromatic ring in the compound, a stable DNA-protein cross-link is generated **(Scheme 2.1.)**[81–83].

Scheme 2.1. Mechanism of Free-Radical Induced DPC Formation DPC can be formed between thymidine in DNA and tyrosine in protein. When hydroxyl radical abstract a hydrogen from the methyl group of thymidine, thymidine can be converted to a reactive thymidine radical. The resulting radicals are subjected to a one-electron addition to the 3-position of tyrosine to generate a stable methylene bond. The 3-position of tyrosine undergoes a hydrogen abstraction to re-aromatize the phenol ring to make the stable dT-Tyr crosslink structure.



By adapting this knowledge, our lab has developed a nanoLC-ESI$^+$-MS/MS based absolute quantitation method which measures the amount of free radical-induced DPC formation by detecting 2-amino-3-(4-hydroxy-3-((1-((2R,4R,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)methyl)phenyl)propanoic acid (dT-Tyr)[82,83]. Because hydroxyl radicals can also be induced by ionizing radiation, ROS-induced DPCs are identical to IR induced DPCs[169,170]. Therefore, in this chapter, ROS-induced DPCs were measured in IR-treated SPRTN WT and hypomorphic cells as well as mouse models to investigate the role of SPRTN. In addition, an SPE method has been developed to improve the nanoLC-ESI$^+$-MS/MS assay for dT-Tyr detection.

Interestingly, DPC degradation can also be accomplished by the proteasome[39,119,153,156,171]. Top1ccs and Top2cc are degraded by the 26S proteasome[153,156,171,172] prior to repair by TPD1 and TPD2[173–175]. Moreover, the involvement of the proteasome in repairing FA induced DPCs was also observed in a previous study[39]. In contrast, the cytotoxicity of hydrogen peroxide, 1,2,3,4-diepoxybutane, cisplatin, phosphoramide mustard, and DEB induced DPCs was not affected by proteasome inhibition[176]. However, to our knowledge, none of the previous studies examined whether ROS induced DPCs can be recognized by the 26S proteasome. Therefore, by examining the involvement of the proteasome and immunoproteasome in DPC repair in this chapter, we aimed to elucidate a novel DPC repair mechanism which is required to maintain DNA integrity.

2.2. Materials and Methods

Materials

Cell Lysis Solution, Protein Precipitation Solution, and RNase A were purchased from Qiagen (Hilden, Germany). UltraPure buffer-saturated phenol was obtained from Invitrogen (Carlsbad, CA). Pronase, alkaline phosphatase, PMSF, leupeptin, pepstatin, aprotinin, Nuclease P1 from *Penicillium citrinum*, crystal violet were purchased from Sigma (St. Louis, MO). DNase I, Phosphodiesterase I and phosphodiesterase II were obtained from Worthington Biochemical Corporation (Lakewood, NJ). Amicon 3K filters were purchased from Millipore (Darmstadt, Germany). Omega Nanosep 10K filters were obtained from PALL Life Science (Port Washington, NY). Oasis MCX were purchased from Waters Corporation (Milford, MA).

Cell Culture

Wild type Mouse embryonic fibroblast (MEF5, *SPRTN* +/+) and *SPRTN*-deficient (MEF7, *SPRTN* f/-) cells were obtained from Dr. Yuichi Machida's laboratory (Mayo Clinic, MN)[140]. The cells were maintained as exponentially growing monolayer cultures in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin streptomycin, in a humidified incubator at 37 °C with 5% CO2.

Clonogenic Assay

MEF5 and MEF7 cells (2.5×10^7) were seeded in 15 cm cell culture dishes and incubated until confluency of 50% or higher was obtained. The cells were aliquoted into 15 mL

tubes and treated with different doses of ionizing radiation (0, 2.5, 5, 7.5 and 10 Gy) on a RadSource RS-2000 biological irradiator (Buford, GA). The dose rate was 1.9 Gy/Min. After irradiation, cells were seeded in 6 well plates (1×10^3 cells per well, in triplicate), and incubated for 14 days to form colonies. The number of colonies were counted manually after crystal violet staining.

MTT Cell Viability Test

MEF5 and MEF7 cells were plated in 48 well plates (1.0×10^5, in triplicate) and incubated for 24 hours. The plates were irradiated with different doses of ionizing radiation (0, 2.5, 5, 7.5 and 10 Gy) on a RadSource RS-2000 biological irradiator (Buford, GA). and incubated for an additional 48 h at 37 °C. Cell viability was measured using an Alamar Blue assay[177] and a Synergy HI Microplate reader (BioTek, Winooski, VT).

Cell Treatment with IR and Isolation of DNA-Protein Crosslinks

MEF5 and MEF7 cells (2.5×10^7, in triplicate) were seeded in 15cm cell culture dishes and incubated until a confluency of 80% or higher was obtained. The cells were then treated with an increasing dose of ionizing radiation (0, 2.5, 5, 7.5, 10, 20 and 40 Gy) on a RadSource RS-2000 biological irradiator (Buford, GA). To quantify DPCs, MEF5 and MEF7 cells were incubated for an additional 3 h at 37C before harvested for DPC repair. The cells were lysed in Qiagen cell lysis solution (2 mL for 10 million cells) and incubated with 10 µL of RNase A, followed by proteinase K (8U) incubation at 37 °C overnight. 750 µL of Qiagen protein precipitation solution was added to sample, and

100% ethanol (approximately 2.5mL) was added to the supernatant to collect DPC containing DNA pellets. The resulting pellets were washed with 70% ethanol, dried for 10 min, and reconstituted in 10mM Tris-HCl pH 7.5 (100 – 500 µL). The amount of DNA was measured using UV spectrophotometry (A260) and subsequently determined by quantitation of dG in enzymatic hydrolysates.

DNA Quantitation by dG Analysis

To quantify the amount of DPC-containing DNA extracted from cells and to check for potential RNA contamination in the extracts, 5 µg aliquots of DNA were digested with phosphodiesterase I (210 mU), phosphodiesterase II (105 mU), DNase I (35 U) and alkaline phosphatase (22 U) in 20 µL of 10 mM Tris-HCl/15 mM $ZnCl_2$ (pH 7.0) overnight at 37 °C. The digested samples were analyzed by HPLC-UV on an Agilent Technologies 1100 HPLC system equipped with a diode array UV detector and an autosampler to measure the amount of dG by calculating the corresponding peak area. Atlantis T3 C18 column (2.1 × 150 mm, 5 µm, from Waters Corporation, Milford, MA) and 5 mM ammonium formate, pH 4.0 (Solvent A) and methanol (Solvent B) were used for the system. Solvent composition was changed linearly from 3 to 30% B over 15 min, increased further to 80% B over 3 min, held at 80% B for 1 min, and returned to 3% B over 2 min, where it was kept for the final 8 min of the HPLC run. UV absorbance was monitored at 260 nm. dG was eluted at 12.7 min. UV absorbance was monitored at 260 nm. dG amounts were determined by comparing HPLC peak areas to a calibration curve constructed by injecting known dG amounts.

Quantitation of dT-Tyr Conjugates Using nanoLC-ESI+-MS/MS

Once DPC-containing DNA (50 - 100 μg) was isolated from cells as described above, the samples were reconstituted in 100 μL of 10 mM Tris pH 7.5 and digested with 4 Units proteinase K and 50 μg of pronase (Sigma, St. Louis, MO) overnight at 37 °C. Next, samples were subjected to Nuclease P1 (2.5U) incubation overnight at 37 °C in 15 mM Tris pH 5.5 with 5 mM $ZnCl_2$. Following this, the sample buffer was adjusted to pH 7 by adding 1M ammonium bicarbonate for PDEI incubation. PDEI (210mU) was spiked in each sample with 10 mM $MgCl_2$ for incubation overnight at 37 °C. Finally, this was followed by an alkaline phosphatase (22U) incubation overnight at 37 °C. 200 fmol of $^{15}N_1,^{13}C_9$ isotope labeled-dT-Tyr was spiked in each sample, and the enzymes were removed by Nanosep 10K filters at 5,000 x g for 10 minutes. The filtrates were dried completely and re-suspended in 200 μL of 0.1% formic acid in H_2O for SPE purification. The dT-Tyr internal standard and analyte were captured on a MCX column and eluted by ammonium hydroxide in 30% methanol. The dT-Tyr internal standard and analyte were concentrated to dryness again, re-suspended in 100 μL 0.1% acetic acid in water, and enriched by offline HPLC purification. HPLC purification was done using an Agilent Technologies 1100 HPLC system; Supelcosil LC-18DB (4.6 × 250 mm, 5 μm, from Thermo Scientific, Waltham, MA) and of 0.1% acetic acid in water (Solvent A) and 0.1% acetic acid in acetonitrile (Solvent B) were used for the system. Solvent was held at 2% B for 5 min, increased to 30% B over 7 min, increased to 75% B over 1 min, held at 75% B for 2 min, decreased to 2% B over 1 min and re-equilibrated for the next run. Fractions containing dT-Tyr (around 14.4 min) were collected, dried and reconstituted in 0.1% formic acid (16 μL) for nanoHPLC-ESI+-MS/MS analysis. The samples were analyzed

using a Quantiva mass spectrometer (Thermo Scientific, Waltham, MA) equipped with Dionex UltiMate 3000 RSLC nanoHPLC system (Thermo Scientific, Waltham, MA), and Synergi 4 μm Hydro RP 80A resin (Phenomenex, Torrence, Ca). The neutral loss of dR and the loss of dR and CO_2 from protonated molecules of the analyte (m/z 422.2 [M + H]+ → 306.1 [M + H − dR]+, m/z 422.2 [M + H]+ → 260.1) and the $^{15}N_{14}$,$^{13}C9$- labeled internal standard (m/z 432.1 [M + H]+ → 316.1 [M + H − dR]+, m/z 432.2 [M + H]+ → 269.1) were detected to confirm and quantify the dT-Tyr, and the amounts were calculated using the Xcaliber software program.

nanoLC-ESI+-MS/MS Method Validation

To check The HPLC-ESI$^+$-MS/MS limit of detection, various amounts of dT-Tyr analyte (0.1 - 50 fmol) and 200 fmol of internal standard was diluted in Solvent A(20ul) and injected for the scan. dT-Tyr were quantified by nanoHPLC-ESI+-MS/MS analysis as described above, to observe the analyte/internal standard ratio. To confirm the limit of quantification, 50 μg of calf thymus DNA (in triplicate) or 50 μg of MEF7 cell DNA was enzymatically digested as described above. The digests were spiked with various amounts of dT-Tyr analyte (10 - 500 fmol) and 200 fmol of internal standard, followed by cation exchange and offline HPLC purification as described above. dT-Tyr were quantified by nanoHPLC-ESI$^+$-MS/MS analysis as described above, to observe the analyte/internal standard ratio.

Quantitation of dT-Tyr in Wild Type and SPRTN Deficient Mouse Tissues

To investigate the role of SPRTN, tissues from 10-month old wild type ($SPRTN^{+/+}$) and Spartan hypomorphic mice ($SPRTN^{H/H}$) were provided by Machida's lab as described previously[140]. 200 mg of tissue (liver, spleen, kidney, heart, thymus, and brain, 3 per group) were extracted, diced into small pieces and homogenized in 250 mM sucrose/ 20 mM sodium phosphate buffer (pH 7.5). DPC-containing DNA was extracted and digested for dT-Tyr conjugates quantitation as described above. dT-Tyr conjugate quantitation was performed as described above.

Quantitation of dT-Tyr in IR-treated Mouse Tissues

To investigate IR-induced DPC formation, C57BL/6 wild type mice were irradiated in a 137Cs Mark I-68 irradiator (JL Shepherd & Associates, San Francisco, CA) to yield doses of 0, 8 and 12 Gy (3 mice per group). After irradiation, the animals were euthanized and the liver, spleen, kidney, heart, thymus, and brain were excised for DPC-containing DNA extraction. DPC-containing DNA extraction and dT-Tyr Conjugate quantitation were performed as described above.

Quantitation of dT-Tyr in Immunoproteasome Knockout Mouse Tissues

To investigate the role of the immunoproteasome in DPC repair, 10-month old wild type and LMP7/MECL-1 double knockout mice (3 per group) were prepared by the Ferrington group as described in their previous studies[178]. 200 mg of tissues (liver, spleen, kidney, heart, thymus, and brain, 3 per group) were homogenized and used for DPC-containing

DNA extraction as described above. dT-Tyr Conjugate quantitation was performed as described above.

2.3. Results

2.3.1. Cytotoxicity and Quantification of DPCs in the IR treated WT and SPRTN Deficient Cells

Our lab has developed an isotope dilution tandem mass spectrometry assay for dT-Tyr conjugates[111] to quantify IR-induced DPCs in cells. In brief, DPCs were extracted from IR treated cells, and the resulting DPCs were subjected to enzymatic digestion to generate dT-Tyr, which can be analyzed by isotope dilution HPLC-ESI^+-MS/MS. dT-[$^{15}N_1$,$^{13}C_9$]Tyr was synthesized as in previous studies[179], and a known amount of internal standard was spiked in enzymatically digested DPC extracts to allow for quantification the analyte. The simultaneous loss of dR and CO_2 from protonated molecules of the analyte and internal standard were scanned to measure the amount of IR-induced DPCs **(Figure 2.1)**.

In a recent study, the Tretyakova group demonstrated that radical-induced DPCs were increased in an IR-dose dependent manner[176]. Therefore, SPRTN proficient and deficient mouse embryonic fibroblasts (MEF5 and MEF7) were used to investigate if the amount of radical-induced DPC could be affected by the presence or absence of this DPC repair pathway. Once MEF cells were treated with 40 Gy of ionizing radiation, the amounts of dT-Tyr in MEF7 were slightly increased than that in MEF5 cells **(Figure 2.2.)**. In addition, the cytotoxicity of radical-induced DPCs was examined by clonogenic assay **(Figure 2.3.) (Table 2.1.)**. The assay demonstrated that MEF7 cells are more sensitive to

IR exposure. In brief, MEF7 cell were 1.591-fold more sensitive (IC_{50} =4.613) to IR compared to MEF5 cells (IC_{50} =2.899). These findings support our hypothesis that SPRTN plays important role in ROS-induced DPC repair.

Figure 2.1 Representative Trace of dT-Tyr Detection

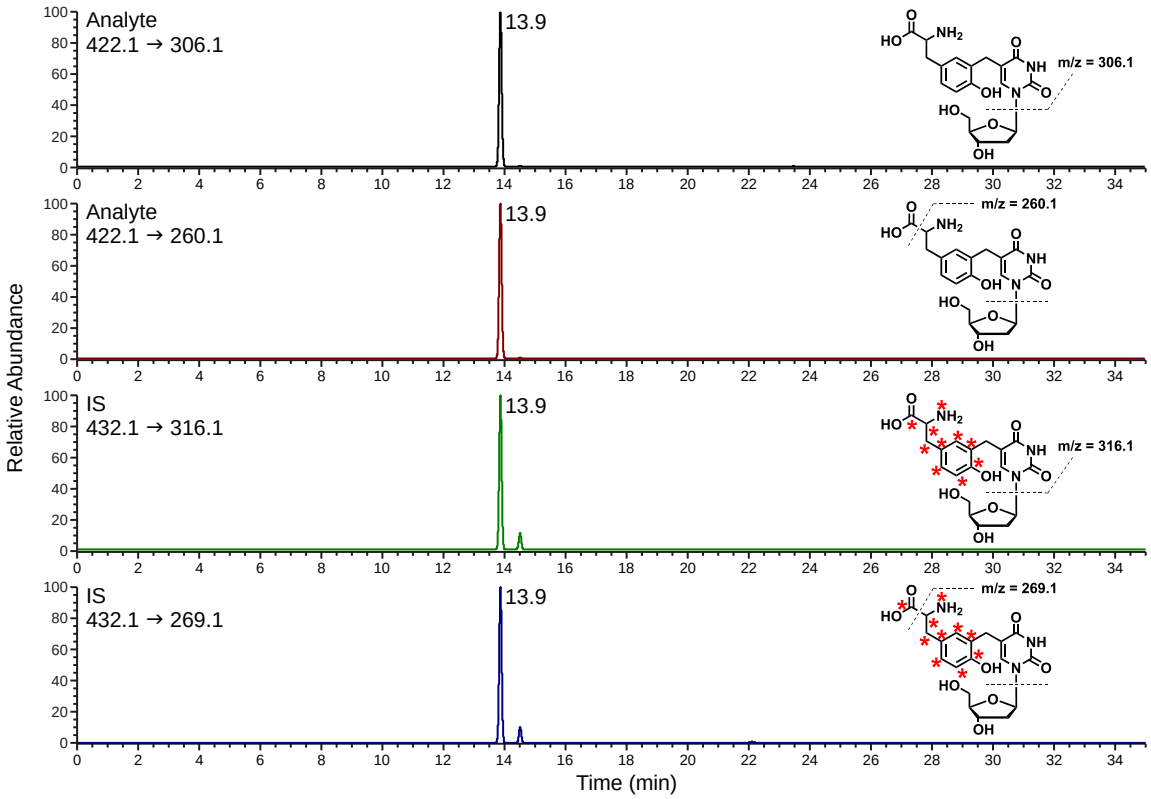


Figure 2.2. The amounts of dT-Tyr in IR treated SPRTN proficient and deficient MEF cells. The levels of dT-Tyr were increased IR dose dependently in MEF cells. When MEF cells were exposed to high dose of IR (40Gy), the adduct levels were higher in SPRTN deficient cells than that in proficient cells.

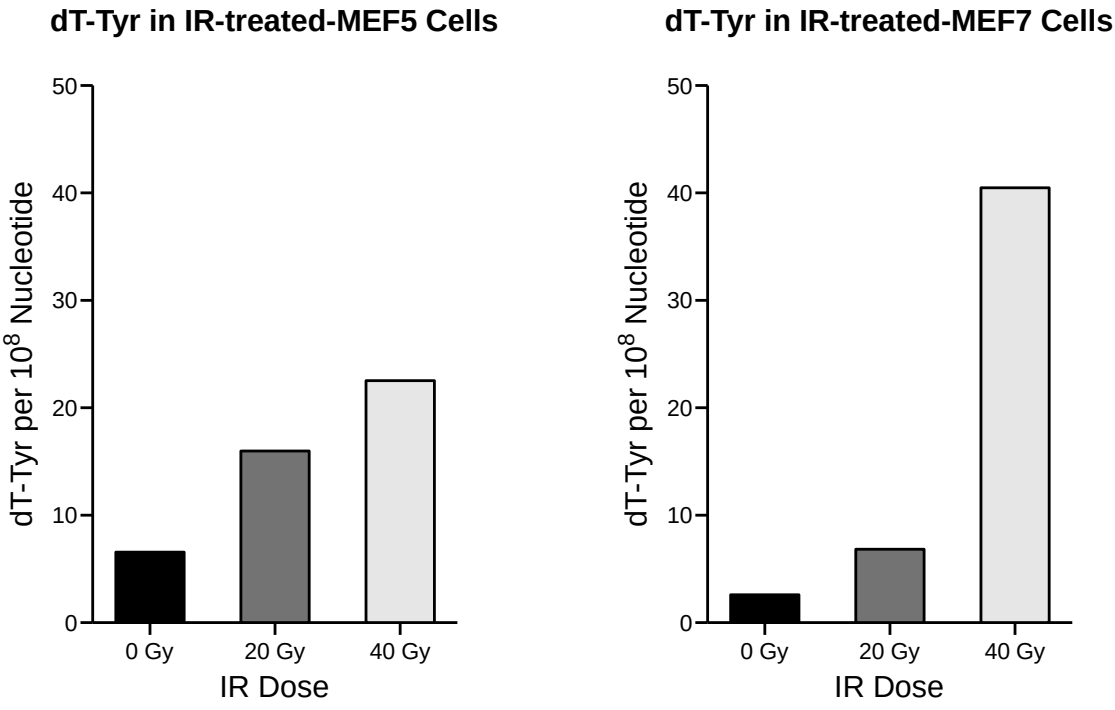


Figure 2.3. Cytotoxicity of IR Induced DPC in Spartan Proficient (MEF5) and Deficient Cells (MEF7) Calculated IC_{50} values of MEF5 and MEF7 cells treated with 0 – 10 Gy IR were 4.613 and 2.899, respectively.

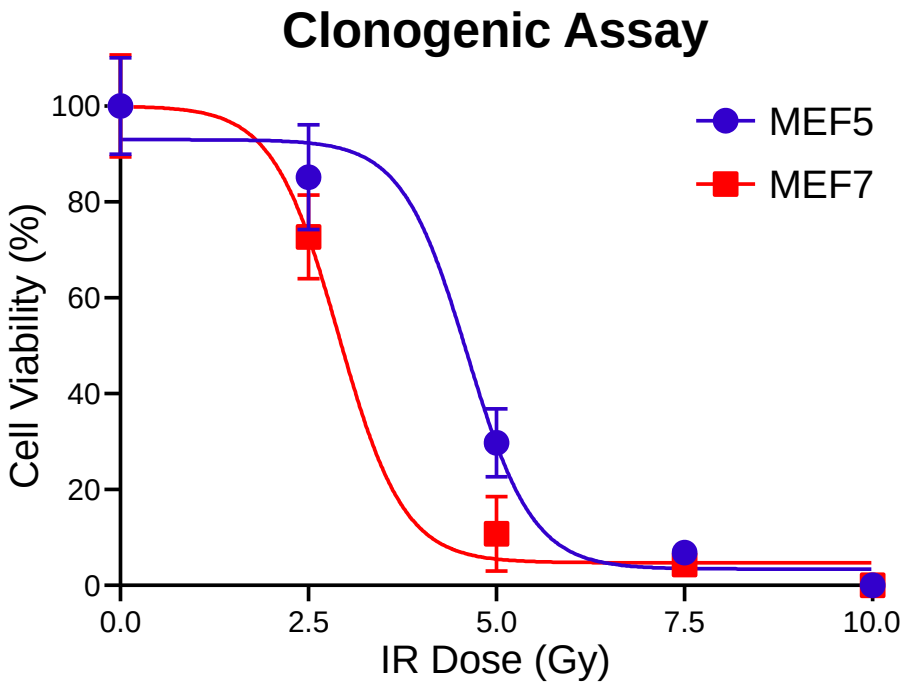


Table 2.1. Clonogenic Assay: Surviving Fraction of Irradiated cells. The surviving fraction of IR treated MEF5 and MEF7 cells was calculated through clonogenic assay. The average and SD were calculated based on three experimental results.

IR Dose (Gy)	MEF5		MEF7	
	Average (%)	S.D.	Average (%)	S.D.
0	100.00	10.04	100.00	10.65
2.5	85.14	10.89	72.66	8.72
5	29.71	7.14	10.79	7.78
7.5	6.86	1.71	4.32	2.16
10	0.00	0.00	0.00	0.00

2.3.2. Quantitation of DPCs in Wild Type and SPRTN Deficient Mouse Tissues

To examine the role of SPRTN in radical-induced DPC repair, DPCs were extracted from liver and brain tissues of wild type and $SPRTN^{H/H}$ (hypomorphic) mice[140] and were quantified through the isotope dilution tandem mass spectrometry assay **(Table 2.2.)**. Interestingly, this analysis demonstrated that $SPRTN^{H/H}$ mice had higher levels of dT-Tyr (41.6 ± 2.64 adducts per 10^8 nucleotides, Liver) compared to that of the wild type (22.05 ± 3.05 adducts per 10^8 nucleotides, Liver) **(Figure 2.4.)**. Moreover, the amount of dT-Tyr was higher in the liver compared to the brain. However, the difference was less than 2-fold. Therefore, these results suggests that another protease other than Spartan must be involved in DPC removal.

Figure 2.4. dT-Tyr Quantitation in Wild Type and SPRTN Deficient Mouse Tissues

dT-Tyr conjugates from liver and brain of 10-month old wild type and SPRTN $^{H/H}$ mice were quantified via isotope dilution tandem mass spectrometry assay. nanoLC-ESI-MS/MS detected significant amounts of dT-Tyr analyte (10 - 50 per 10^8 nucleotides) in the liver and brain.

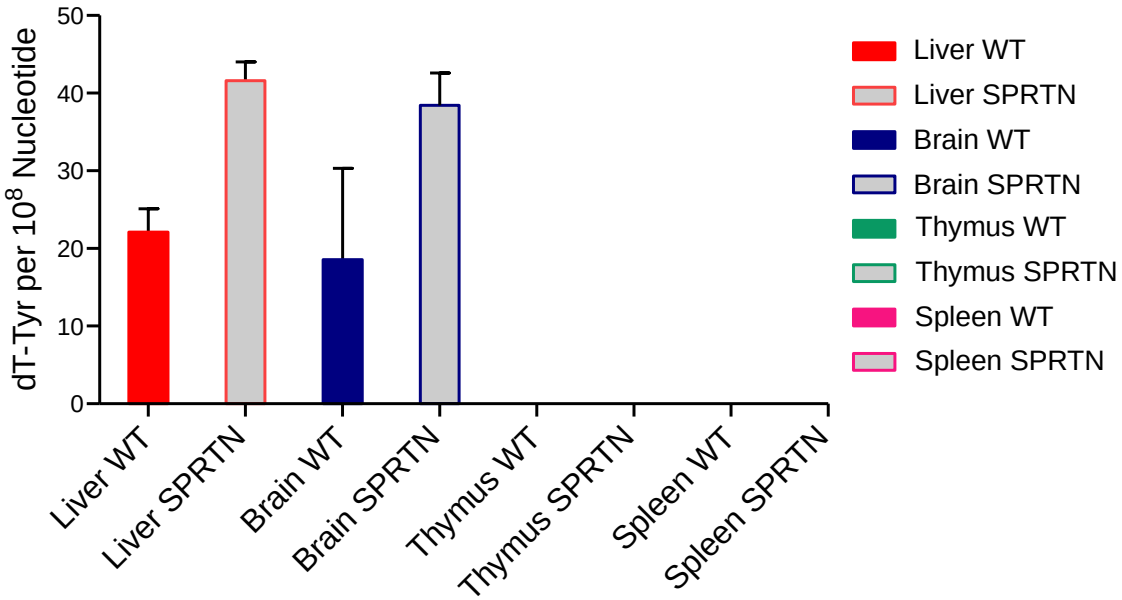


Table 2.2. dT-Tyr Quantitation in Wild Type and SPRTN Deficient Mouse Tissues

nanoLC-ESI-MS/MS detected significant amounts of dT-Tyr (10 - 50 per 10^8 nucleotides) in the liver and brain of 10-month old wild type and SPRTN $^{H/H}$ mice. The average and SD were calculated based on three experimental results.

Sample	dT-Tyr per 10^8 Nucleotide	
	Average	S.D.
Liver WT	22.05	3.05
Liver SPRTN	41.6	2.43
Brain WT	18.54	11.79
Brain SPRTN	38.42	4.36
Thymus WT	0	0
Thymus SPRTN	0	0
Spleen WT	0	0
Spleen SPRTN	0	0

2.3.3. Method Development for Isotope Dilution Tandem Mass Spectrometry Assay

MS-based absolute quantification for dT-Tyr requires multiple preparation steps including enzyme digestion, enzyme removal, and multiple offline HPLC purification steps to enrich the analyte and its internal standard prior to MS injection. However, offline HPLC purification steps are laborious and contaminations in the samples are removed by selecting smaller fractions through sequential HPLC purification. Since repeating same offline HPLC purification steps is not an efficient way to eliminate contamination, a single offline HPLC purification step can be replaced with a solid phase ion-exchange column to improve the existing protocol and to purify the analyte and its internal standard faster and more efficiently. Since dT-Tyr becomes a cation in low pH (**Scheme 2.2.)**, the dT-Tyr ion can be captured by a cation ion exchange column. Therefore, 200 fmol of internal standard and various amounts (10 - 500 fmol) of analyte were spiked into 50µg of DNA from MEF7 cells and used to validate the dT-Tyr enrichment protocol. The limit of quantification was 10 fmol on column **(Figure 2.5.)**, while the HPLC-ESI^+-MS/MS limit of detection was calculated to be 500 amol of dT-Tyr. This indicates that the cation exchange column could help to enrich the analyte and its internal standard. Therefore, the cation exchange and the single offline HPLC purification can be performed instead of multiple offline HPLC purification steps.

Scheme 2.2. Structure of dT-Tyr Cation dT-Tyr becomes a cation in low pH (pH 2.2) value

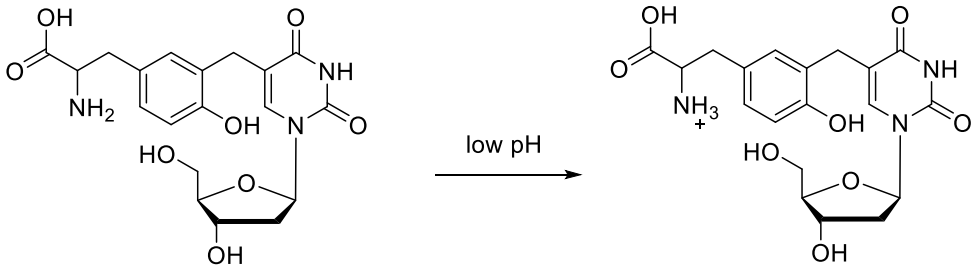
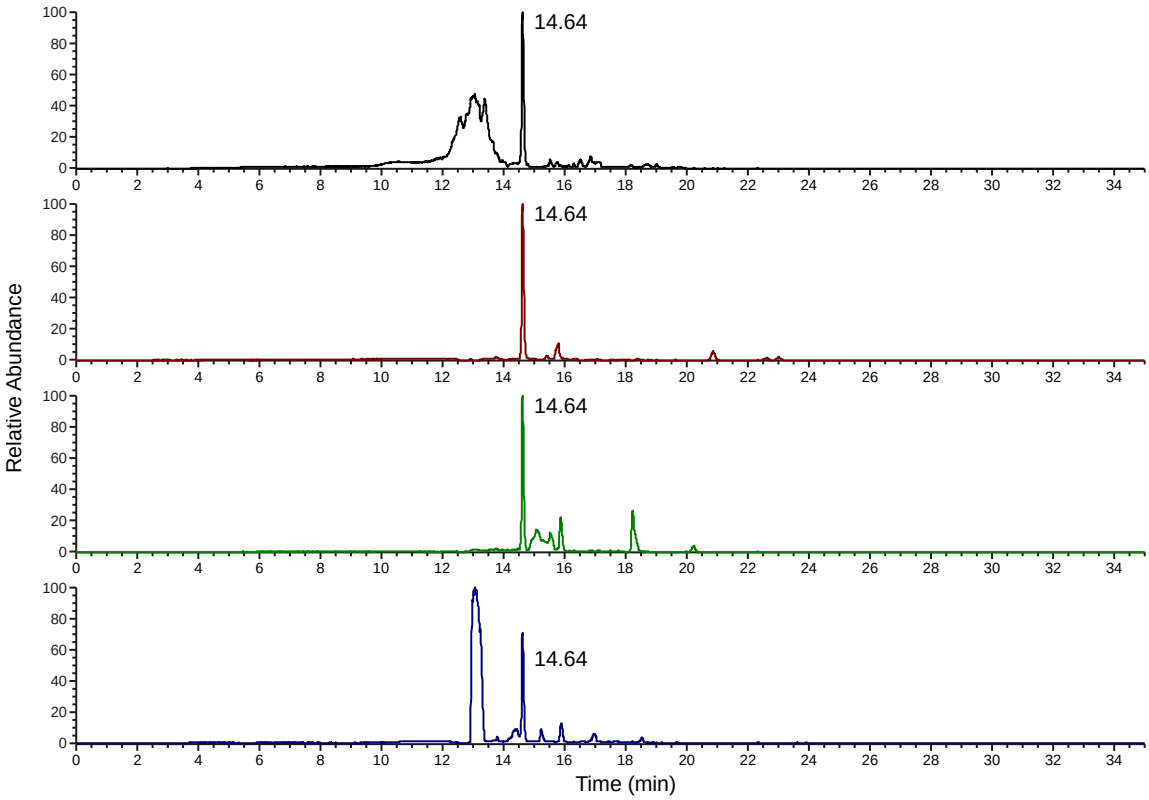


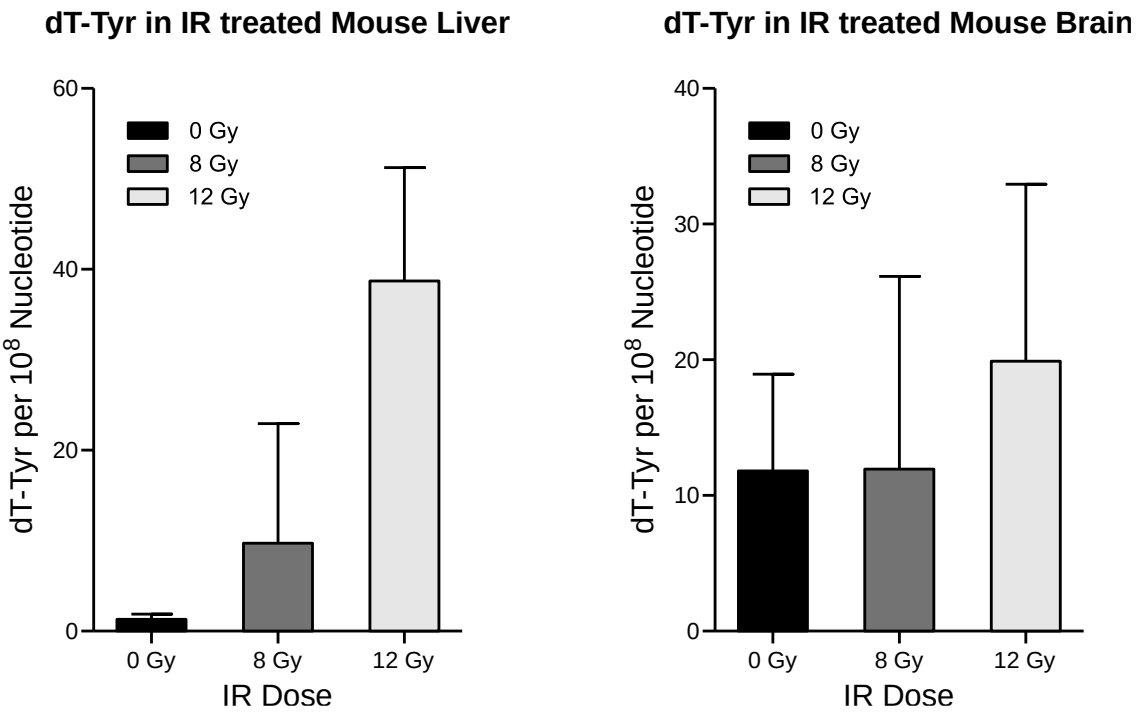
Figure 2.5. Representative Trace of dT-Tyr Detection in MEF7 cells



2.3.4. Quantitation of DPCs in IR treated Mouse Brain and Liver samples

Since our nanoLC-ESI^{+}-MS/MS assay was improved by adapting the cation exchange method described above, the improved assay was utilized to detect IR-induced DPC formation in mouse models. C57BL/6 wild type mice were irradiated with 0, 8 and 12 Gy of gamma-ray, and the liver and brain of the IR-treated mouse were collected for DPC analysis. We found a linear increase of dT-Tyr conjugate in IR-treated mice **(Figure 2.6.)**. 9.712 ± 13.22 and 38.709 ± 12.53 adducts per 10^8 nucleotides were observed in mouse liver treated with 8 and 12 Gy IR, respectively, while the brain samples showed 11.952 ± 14.190 and 19.898 ± 13.05 adducts per 10^8 nucleotides with 8 and 12 Gy. Interestingly, as we observed in wild type and SPRTN $^{H/H}$ mice tissues **(Figure 2.4.)**, the amount of dT-Tyr was slightly higher in the liver compared to the brain. In short, we were successful in detecting dT-Tyr in both the liver and brain, and confirmed that our MS-based absolute quantification method could be used to determine IR-induced DPC formation *in vivo.*

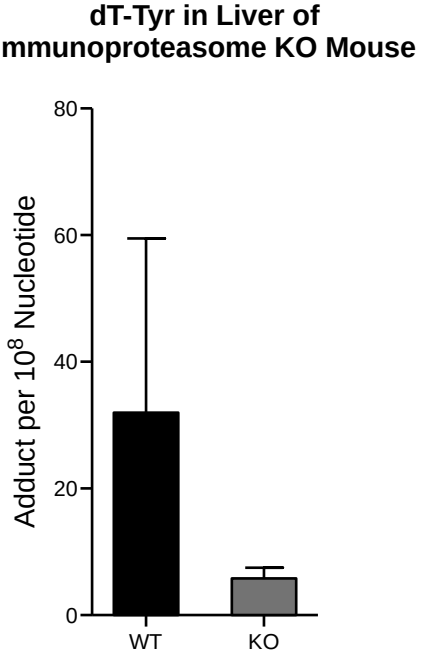
Figure 2.6. dT-Tyr Quantitation in IR-treated mouse tissues The amounts of dT-Tyr were increased in brain and liver tissues after IR treatment. The data are the average ± SD of three experiments.



2.3.5. Quantitation of DPCs In Tissue of Immunoproteasome Knockout Mouse

Because a significant difference in dT-Tyr amounts was not observed between the WT and $SPRTN^{H/H}$ group **(Figure 2.4.)**, the ability of the proteasome to repair DPCs was examined with radical-induced DPCs to elucidate an alternative pathway for DPC degradation. Briefly, our lab has tested the cytotoxicity of IR-induced DPCs in MEF cells with lactacystin to inhibit the 20S subunit of the proteasome. An increased level of dT-Tyr was expected when MEF7 cells were treated with lactacystin, whereas MEF5 cells were expected to show a slight increase of dT-Tyr. Unfortunately, proteasome inhibition had no effect on dT-Tyr levels **(data not shown)**. Interestingly, the absence of dT-Tyr in mouse spleen and thymus tissues suggest the involvement of immunoproteasome in DPC repairs[176]. Therefore, dT-Tyr was also quantified in WT mice and immunoproteasome subunit knockouts mice (LMP7 and MECL genes) to examine the role of the immunoproteasome in DPC repair **(Figure 2.7.)**. Interestingly, dT-Tyr levels in liver tissues were reduced to one-fifth when comparing the KO group to control group. This suggest that proteases, such as SPRTN, can be recruited to ubiquitinylated crosslinked proteins when immunoproteasome could not participate in ROS-induced DPCs repair. This also support that immunoproteasome may not be the major mechanism in DPC repair.

Figure 2.7. dT-Tyr Quantitation in Wild Type and Immunoproteasome Knockout Mouse dT-Tyr conjugates from liver of 10-month old wild type and immunoproteasome knockout (LMP7 and MECL KO) mice were quantified via isotope dilution tandem mass spectrometry assay. nanoLC-ESI-MS/MS detected significant amounts of dT-Tyr analyte (4 - 60 per 10^8 nucleotides) in the liver tissues. The data are the average ± SD of three experiments.



2.4. Discussion

As described in **(Scheme 2.1.)**, DPC can be formed between the methyl group of thymidine and tyrosine residue upon IR exposure[82,83]. Since hydroxyl radicals can abstract hydrogen atoms from Thy[16,57,180], the same thymidine-tyrosine crosslinks can be generated endogenously in cells not exposed to IR[81,83]. Due to the structural complexity of DPC lesions and the difficulty of generating relevant DPC structures, quantitation of radical induced DPCs was difficult to be conducted. However, the Tretyakova group recently succeeded in quantifying free radical-induced DPCs via a sensitive nanoLC-ESI+-MS/MS assay which accurately measure the amount of dT-Tyr in cells. Using this assay, our lab was able to demonstrate an increased level of ROS-induced DPCs in the cardiomyocytes of rats that underwent LAD ligation/reperfusion[111]. Quantitative proteomic analysis of the same samples showed that 80 proteins were increased by at least 1.2-fold after LAD ligation/reperfusion.

Although our nanoLC-ESI^+-MS/MS assay is sensitive enough to detect the dT-Tyr conjugate, the assay requires extensive enzyme incubation and purification steps to be analyzed by the mass spectrometer. To make the method faster and more accurate, multiple HPLC offline purifications were modified to ion-exchange and reverse phase HPLC offline purification combined protocols. Several SPE columns, including a reverse phase column, cation exchange column, and anion exchange column, were tested to purify dT-Tyr conjugates **(data not shown)**. However, only the cation exchange column (Oasis MCX) was able to purify dT-Tyr efficiently. By using newly developed method, more efficient dT-Tyr enrichment was accomplished, leading to a successful

quantification of radical-induced DPCs from the tissues of IR-exposed C57BL/6 mice **(Figure 2.6.)** and from the tissue of hypomorphic ($SPRTN^{H/H}$) mice **(Figure 2.4.)**.

By showing a linear increase of dT-Tyr in HeLa cells treated with 0, 8, and 16 Gy of IR through our preliminary experiments[176], nanoLC-ESI^+-MS/MS assay was confirmed as an accurate method to quantify IR-induced DPC. In this chapter, we applied the improved assay to elucidate the roles of SPRTN in the repair of radical-induced DPCs. After treating MEF5 and MEF7 cells with 40Gy dose of ionizing radiation, we showed that the amount of dT-Tyr conjugate was increased in MEF7 than that in MEF5 **(Figure 2.2.)**, indicating the involvement of SPRTN in DPC repair as Vaz *et al.* suggested in their earlier studies[34]. In addition, a clonogenic assay revealed that SPRTN deficient cells were more sensitive to IR **(Figure 2.3.)**, highlighting the importance of SPRTN protease. Unfortunately, it was difficult to confirm the role of SPRTN with low dose of IR (0 – 10 Gy) due to similar amounts of dT-Tyr in MEF5 and MEF7 cells.

nanoLC-ESI^+-MS/MS assay is sensitive enough to detect dT-Tyr conjugates *in vivo* as well. Our lab has demonstrated that hypomorphic ($SPRTN^{H/H}$) mice showed higher dT-Tyr levels than wild type mice, suggesting that SPRTN plays a role in ROS-induced DPC repair **(Figure 2.4.)**. The amount of dT-Tyr was increased by 1.5 - 2 fold in SPRTN deficient mice, and similar results were obtained when comparing endogenous DPC levels in wildtype and Spartan-depleted HeLa cells through a SDS/K+ precipitation assay[151]. Although $SPRTN^{H/H}$ mice express only 10% of Spartan compared to wild type,

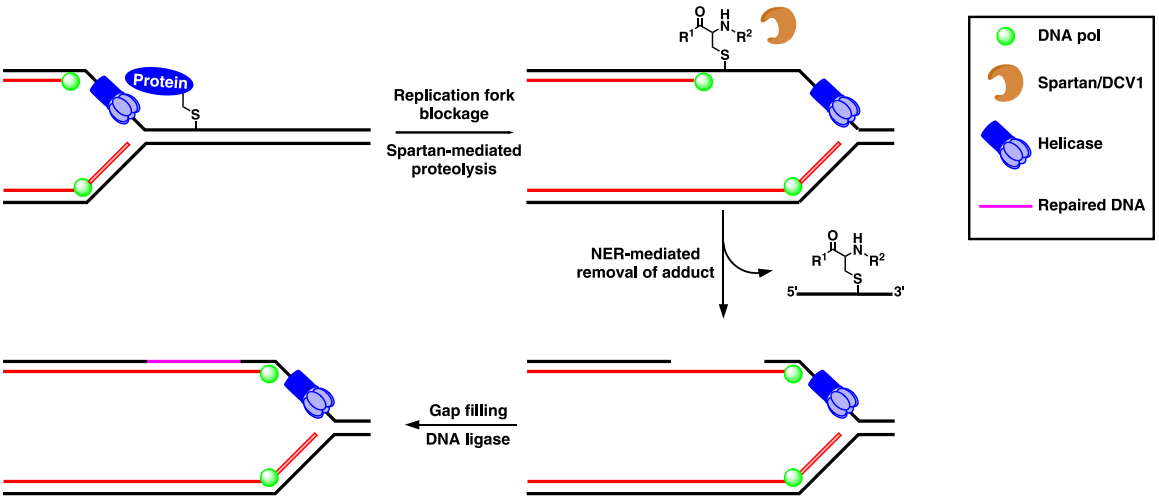
no dramatic difference in the amount of dT-Tyr was observed between two groups. Therefore, we suggested that other proteases may repair endogenous DPCs when SPRTN is not available. Interestingly, recent studies have found that other proteases which play a similar role as SPRTN, supporting our hypothesis **(Machida's unpublished data).** Taken together, IR-induced and ROS-induced DPCs can be examined to investigate novel DPC repairs through our nanoLC-ESI^{+}-MS/MS assay.

Previously, the repair pathway involving proteasomal DPC degradation was tested in MEF5 and MEF7 cells. DPCs in MEF cells were induced by several different DPC inducing agents, such as hydrogen peroxide, 1,2,3,4-diepoxybutane, cisplatin, phosphoramide mustard, and DEB, but pharmacological proteasome inhibition by MG-132 or lactacystin had no effects on cell viability[176]. Therefore, it was suggested that the immunoproteasome may play a role in DPC repair due to the low amount of dT-Tyr in mouse spleen and thymus tissues where lymphatic immunoproteasomes are present. Because immunoproteasomes contain a replacement for the three subunit, including low-molecular-weight proteins (LMP2 and LMP7) and the multicatalytic endopeptidase complex subunit 1 (MECL-1), compared to 20S proteasome core, some unique peptides are produced through its distinct proteolysis activity[181,182]. Moreover, immunoproteasome knockout (LMP2 KO) mice have been shown to have more oxidative damaged proteins in their tissues[183]. Therefore, we quantified dT-Tyr in tissues of the wild type and double knockout (LMP7 and MECL-1 KO) mice to investigate whether the immunoproteasome can repair endogenous ROS-induced DPCs. However, our nanoLC-ESI^{+}-MS/MS assay showed the level of dT-Tyr was decreased in KO group **(Figure 2.7.)**, supporting our

hypothesis that DPC proteases, including SPRTN, might compete with proteasome to target ubiquitinated DPCs and play major roles in ROS-induce DPC repair. Interestingly, recent experiments in our lab studying DPC ubiquitylation showed similar patterns[184]. Briefly, plasmid containing hOGG1 protein conjugated to an abasic site of DNA were subjected to transfection, enzyme digestion, pulldown against ubiquitin, and qPCR to check DPC ubiquitylation. PCR products from the region where the DPC was located were enriched, while those from the control region showed no change. This indicates that crosslinked proteins can be ubiquitinylated to be recognized by SPRTN or other proteases with high affinities, but immunoproteasome may not be the major mechanism in DPC repair.

To accomplish replication coupled DPC repair, DPCs should be degraded into smaller peptides on both the leading and lagging DNA strand[167] and SPRTN prevents these DPCs from causing replication fork stalling as a part of the replisome[143,151]. As demonstrated in **(Figure 2.8.),** when the replisome/CMG helicase is blocked by a DPC[23,24,108,119,138,152,167], SPRTN will be recruited to ssDNA to initiate its proteolysis activity. The resulting peptide conjugates can be bypassed by TLS polymerases[152,167] and repaired by NER[39,122]. In addition, the activation of SPRTN in presence of ssDNA [135,144,145] also supports this suggested pathway. In conclusion, multiple proteases and proteasomes may be required to repair endogenously and exogenously formed DPCs and their functions in novel DPC repair pathway should be further investigated.

Figure 2.8. Replication Coupled DPC Repair When helicase reaches a DPC site, replication fork is stalled and Rad6-Rad18 activates monoubiquitylation. Monoubiquitylation and ssDNA attract SPRTN to the site, resulting DPC degradation. Upon degradation, SPRTN dissociates and TLS polymerase bypass the lesion site. After bypass, TLS polymerase is replaced by replicative polymerase and the remaining peptide chain will be excised via NER.



III. Effects of 5-Formylcytosine Mediated DNA-Peptide Cross-links on Transcription in Human Cells

Adapted from:

Shaofei Ji, Daeyoon Park, Konstantin Kropachev, Marina Kolbanovskii, Nicholas Geacintov and Natalia Y. Tretyakova; Effects of 5-formylcytosine mediated DNA-peptide cross-link on transcription and its nucleotide excision repair in human cells

Manuscript in preparation

Daeyoon Park performed all experiments described in this chapter. Shaofei Ji performed oligodeoxynucleotides containing DNA-peptide crosslink substrate synthesis.

3.1 Introduction

5-Methylcytosine (5mC) is an important DNA epigenetic mark that plays a central role in regulating the levels of gene expression[185]. 5mC residues are primarily found at CpG dinucleotides, and their formation from cytosine is catalyzed by DNA methyltransferases (DNMTs) including DNMT1, DNMT3A, DNMT3B and DNMT3C[186,187]. DNA methylation can regulate the transcriptional machinery by inducing chromatin remodeling and modifying the access of transcription factors to promoter regions of genes. DNA methylation patterns are affected in many human diseases including asthma, autism, and cancer[101]. Furthermore, changes in DNA methylation patterns are observed across the genome in aging process, giving rise to the concept of 'epigenetic clock'[188,189].

Ten-eleven translocation (TET) family of proteins oxidizes 5mC in DNA to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxycytosine (5caC)[98,100]. Although these oxidized forms of 5mC may have the ability to regulate gene expression, their biological functions are incompletely understood, especially given their relatively low abundance in cells. The abundances of 5fC, 5caC, and 5hmC are 0.002% ~ 0.2%, 0.002% ~ 0.2%, and 0.02 ~ 0.1% of all cytosines, respectively, while mC constitutes 4-6% of all cytosines[98,190,191]. Among the oxidized forms of 5mC, 5fC is unique due to the presence of inherently reactive aldehyde group in its structure. The Tretyakova group recently reported that 5fC generates reversible DNA-protein crosslinks by forming Schiff base conjugates with lysine and arginine side chains of histone proteins[106]. Furthermore, a recent study suggested that 5fC-mediated covalent DNA-

histone interactions is important in the regulation of gene expression by controlling nucleosome position and establishing distinct regulatory regions[192].

Because of their size, 5fC-mediated DNA–protein conjugates may interfere with DNA replication, transcription, and repair[9,105,106]. Recent work conducted in our laboratory revealed that full size 5fC-histone conjugates can block DNA replication[105,106], while the corresponding DNA-peptide lesions can be bypassed by TLS polymerases[26,193]. It has been reported that DNA-protein crosslinks are subject to proteolysis via SPRTN proteases[34,108,151,167] and/or the proteasome[39,153,154], which degrades DPCs into smaller DNA-peptide conjugates (DpCs) which can be removed by NER [116,119]. The effects of DPC on transcription are not completely understood. Lippard *et al.* first reported that cisplatin-induced DPCs block transcription[194]. Moreover, the mass spectrometry based proteomic study revealed that the proteins trapped on DNA in the presence of cisplatin are involved in DNA damage response and repair, transcriptional regulation, RNA processing, and cell cycle regulation[195]. Our group utilized an *in vitro* assay to examine T7 RNA polymerase bypass of 5fC-mediated DPC and DpC conjugates **(Ji and Tretyakova, submitted)**. We found that DPCs present in the transcribed strand were a strong block to RNA polymerases, while DPCs present in noncoding strand were less blocking **(Ji and Tretyakova, submitted)**. Smaller DpC lesions were bypassed by T7 RNA polymerase is either error free or error-prone manner, depending on their structure. However, mammalian RNA polymerases are significantly more complex, multi-subunit proteins and may have different interactions with bulky DPC lesions.

In the present work, we examined the effects of 5fC-mediated conjugates on transcription in human cells. A plasmid substrate containing a structurally defined DpC lesion, such as 5fC conjugated 11 mer peptide (RPKPQQFFGLM-$CONH_2$) or a single Lys, was synthesized and transfected into cells, and their effects on transcription were studied using the competitive transcription and adduct bypass (CTAB) assay developed in the Wang group[196]. In this approach, RNA transcripts are reverse transcribed to cDNA, amplified by PCR, and sequenced by HPLC-ESI-MS/MS.

3.2. Materials and Methods

Materials

Sodium cyanoborohydride, and sodium phosphate were purchased from Sigma (St. Louis, MO). T4 PNK, NcoI, SfaNI, and Nt.BstNBI were purchased from New England Biolabs (Beverly, MA). Phenol:chloroform:isoamyl alcohol solution (25:24:1), HiPure plasmid Mini-prep Kit and HiPure plasmid Maxi-prep Kit were purchased from Invitrogen (Carlsbad, CA). Sep-Pak C18 solid phase extraction column was obtained from Waters Corp. (Millford, MA). 100K Amicon filters were purchased from Millipore (Darmstadt, Germany). E.Z.N.A. Total RNA Kit I and E.Z.N.A. Cycle Pure kit were purchased from Omega Bio-Tek (Norcross, GA). Ambion DNA-free kit was purchased from Life Technologies (Carlsbad, CA). M-MLV reverse transcriptase was purchased from Promega (Madison, WI).

Methods

Synthesis of DNA Oligodeoxynucleotides

5fC-containing oligodeoxynucleotides (5'-ATGGCGGGXTAT-3', where X= 5-formylcytosine) were synthesized using an ABI 394 DNA synthesizer from Applied Biosystems (Foster City, CA). 5-formyl dC nucleoside phosphoramidite was purchased from Glen Research (Sterling, VA). Solid phase synthesis was performed under standard coupling conditions on 1umole scale. Newly synthesized oligodeoxynucleotides were cleaved from the resin by incubation with 30% ammonium hydroxide at room temperature overnight and deprotected by incubation with 80% acetic acid at room temperature for 6 h. The resulting synthetic DNA strands were purified by Agilent Technologies 1100 HPLC system equipped with a diode array UV detector. A Jupiter 4u proteo 90 Å (250 x 10.0mm, Phenomenex, Torrence, CA) was eluted with a gradient of 100mM TEAA pH 7.0 (Solvent A) and ACN (Solvent B) were used for the system at flow rate of 2.5ml/min. Solvent composition was linearly changed from 3% solvent B to 11% in 10min, further to 15% in 25min and to 60% in 3min, kept at 60% for 5 min, then brought back to 3% in 2min. The DNA oligodeoxynucleotides were eluted at 27min. HPLC purified oligodeoxynucleotides were desalted using Illustra NAP-10 columns (GE Healthcare, Pittsburgh, PA) and characterized by HPLC-ESI$^-$-MS on an Agilent ion trap mass spectrometer.

Construction of Oligodeoxynucleotides Containing DNA-peptide and DNA-lysine Crosslinks

DNA oligodeoxynucleotides (5'-ATGGCGGGXTAT-3', where X =5fC) (300pmol) were incubated with 20 fold molar excess of 11-mer peptide (RPKPQQFFGLM-$CONH_2$) in 16 μl of 4.5 mM sodium phosphate (pH 7.5) at 37 °C for 3 h and then incubated with 4 μl of

100 mM $NaCNBH_4$ overnight. The samples were heated at 90 °C for 15min and separated by 20% PAGE. Gel bands containing DNA-peptide conjugates were excised and eluted via classic freeze-thaw method using ethanol. The resulting DNA-peptide conjugates were desalted using Sep-Pak C18 solid phase extraction column from Waters Corporation (Milford, MA). Crosslink formation was confirmed by MALDI-TOF-MS under 3-HPA matrix (50:50 acetonitrile: H2O containing 50 mg/mL 3-HPA and 10 mg/mL diammonium hydrogen citrate). The analyses were performed at of 800-15,000 m/z on an AB-Sciex 5800 MALDI-TOF/TOF-MS (AB SCIEX, Foster City, CA) operated in the positive ion and 2,000 laser shots accumulation.

DNA oligodeoxynucleotides (5'-ATGGCGGGXTAT-3', where X =5fC) (300pmol) were incubated 25 mM lysine in 16 μl of 4.5 mM sodium phosphate (pH 7.5) at 37 °C for 3 h and then incubated with 4 μl of 100 mM $NaCNBH_4$ overnight. The DNA-lysine conjugates reaction mixtures were injected to HPLC using an Agilent Technologies 1100 HPLC system for purification. An Xbridge BEH C18 column (4.6mm × 150 mm, 2.5 μm, from Waters Corporation, Milford, MA) was eluted with a gradient of 100 mM TEAA (Solvent A) and ACN (Solvent B) with flow rate of 0.9ml/min. Solvent composition was linearly changed from 6% solvent B to 15% in 40 min, further to 40% in 20min and to 75% in 1min, kept at 75% for 2 min, then brought back to 6% in 2 min. The DNA oligodeoxynucleotides were eluted at 15.4 min. The purified samples were examined by HPLC-ESI-MS on an Agilent ion trap mass spectrometer in ESI positive mode with the mass range m/z 300-1800. Zorbex 300 SB-c18(150 × 0.5 mm, 5μm, from Agilent Technologies, Santa Clara, CA) was eluted with a gradient of 100 mM TEAA (Solvent

A) and ACN (Solvent B) were used at flow rate of 15ul/min. Solvent composition was linearly changed from 2% solvent B to 75% in 35min, then brought back to 2% in 6min. The oligodeoxynucleotides containing DNA-lysine crosslinks oligodeoxynucleotides were eluted at 11min.

Construction of Plasmid Substrates Containing Site Specific DNA-peptide and DNA-Lys crosslinks

Two plasmids, control and competitor pTGFP-H7 Hha10 plasmids, were obtained from Dr. Yinsheng Wang's Lab at University of California at Riverside. The two plasmids were introduced into TOP10 competent cells to be amplified through transformation. When the colonies are formed on LB plates with kanamycin, a single colony was amplified in 4ml of LB broth with kanamycin (50ug/mL) for 16 h. The LB broth are aliquoted to be purified with Mini-prep Kit from Invitrogen (Carlsbad, CA) for Sanger sequencing and to be stored at -80 °C for later use. Once the sequence had been confirmed, the plasmid (1ml of LB broth) was amplified in one-liter LB broth and resulting pellet was purified using a Maxi-prep Kit from Invitrogen (Carlsbad, CA). The purified control plasmids used to produce plasmids containing site specific DNA-peptide conjugate as previously described[196]. In brief, approximately 300 μg of control plasmid were incubated with Nt.BstNBI from New England Biolabs (Beverly, MA) to introduce nicks. An excessive amount(5nmol) of 25-mer oligonucleotides (5'-CATCGACTC CCGAATAGCCCGCCAT-3') complementary to the nicked area were introduced to detach the resulting oligonucleotides. The 25 mer dsDNA and the gapped plasmid was separated using an 100k Amicon filter from Millipore (Darmstadt, Germany). The

purified gapped plasmids were ligated with 5fC-containing oligodeoxynucleotides (5'-ATGGCGGGXTAT-3') and 13mer oligodeoxynucleotides (5'-pTCGGGAGTCGATG-3'). Lesion bearing supercoiled plasmids were then subjected to agarose gel purification, followed by restriction enzyme treatments. 40 ng of plasmids (control and lesion bearing plasmid) were incubated with NcoI (2 U) and alkaline phosphatase. The 5' end of the DNA was labeled with γ-^{32}P ATP and T4 PNK (5 U), and treated with SfaNI (2 U) to produce the radiolabeled 13-mer oligonucleotides containing DNA-peptide crosslink. The samples were loaded into 20% denaturing PAGE gel, and DNA fragments were visualized by phosphorimaging.

Cell Culture

Human Xeroderma Pigmentosum Complementation Group A (XPA, XPA deficient) cells were obtained from NIGMS Human Genetic Cell Repository (Camden, NJ). Human embryonic kidney cells (HEK293T) were obtained from Dr. Basu's Lab (University of Connecticut, CT). CSB cell were obtained from the Coriell Institute for Medical Research (Camden, NJ). Cells were maintained as exponentially growing monolayer cultures in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin streptomycin, in a humidified incubator at 37 °C with 5% CO2.

Competitive Transcription and Adduct Bypass (CTAB) Assay

Competitive transcription and adduct bypass assays were performed as described in a previous study[196] to examine the effect of DNA lesions on transcription. In brief, lesion

bearing or lesion free control plasmids were mixed with non-lesion competitor vector at a 3:1 ratio. The mixed plasmids (50 ng) were then transfected into XPA-deficient cells (NER deficient), XPA-complemented cells (NER proficient) or HEK293T cells in 24 well plates (1.25×10^5 cell per well) using Lipofectamine 2000. Following 24 h incubation, cells were subjected to RNA extraction using the E.Z.N.A. Total RNA Kit I from Omega Bio-Tek (Norcross, GA), DNA removal using Ambion DNA-free kit from Life Technologies (Carlsbad, CA), and cDNA synthesis. M-MLV reverse transcriptase and a primer (5'- TCGGTGTTGCTGTGAT-3') were used for cDNA synthesis. The resulting cDNA was amplified using PCR and the following two primers, 5'-CTAGCGGATGCATCGACTC-3' and 5'-TGCTGCGGATGATCTTGTCG-3'. The PCR product was purified using E.Z.N.A. Cycle Pure kit from Omega Bio-Tek (Norcross, GA) and digested with SfaNI and NcoI to generate 13mer or 16mer fragments, which were purified via phenol/chloroform extraction. Specifically, an equal volume of phenol:chloroform:isoamyl alcohol solution (25:24:1) was added to each sample, and the DNA fragments were collected from the aqueous layer.

Time-course Competitive Transcription and Adduct Bypass (CTAB) Assay

To examine time dependence for competitive transcription and adduct bypass assays were performed as described above, but cells were allowed to grow for different amounts of time prior to RNA extraction. After transfection of the DNA-peptide crosslink containing plasmid, cells were grown for 24, 48, or 72 h and then harvested for RNA extraction. RNA was extracted using Ambion DNA-free kit and immediately converted into cDNA on the same day to minimize RNA degradation. The amplified cDNA

underwent PCR purification, restriction enzyme digestion, and phenol/chloroform extraction prior to HPLC-ESI-MS/MS analysis as described below.

Mass Spectrometry based Transcriptional Product Analysis

DNA samples were reconstituted in 0.1% formic acid in water(8ul) and analyzed using an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific) interfaced with a Dionex Ultimate 3000 RS autosampler and Dionex UltiMate 3000 RSLCnano system. Solvent system consisted of 15 mM ammonium acetate in water (Solvent A) and 0.1% formic acid in acetonitrile (Solvent B). A Zorbax 300SB-C18 column (150 × 0.5 mm, 5 μm from Agilent Technologies, Santa Clara, CA) was eluted at a 15ul/min flow rate using the following gradient. Solvent composition was linearly changed from 2% solvent B to 20% in 23min, further to 90% in 7min, kept at 90% for 2 min, then brought back to 2% in 3min. cDNA fragments from the control or lesion-bearing site (5'-CCGAATAXCCCGC-3', where X=A, T, C or G) and that from competitor plasmids (5'-CACAATAGCATATCGC-3') were observed by mass spectrometry. The sequences and their relative amounts were compared by MS/MS analysis and calculating the intensities of the corresponding peaks. Mass spectrometry analyses were performed at a resolution of 60,000 and a scan range of *m/z* 300-2000. 974.6650 m/z , 970.6670 m/z, 986.4140 m/z, 964.6640 m/z, 892.4020 m/z, and 1207.4450 *m/z* were scanned to detect monitor 13-mer oligonucleotides containing different transcription products from the control or lesion-bearing plasmids and 16-mer oligonucleotides from competitor plasmids.

3.3. Results

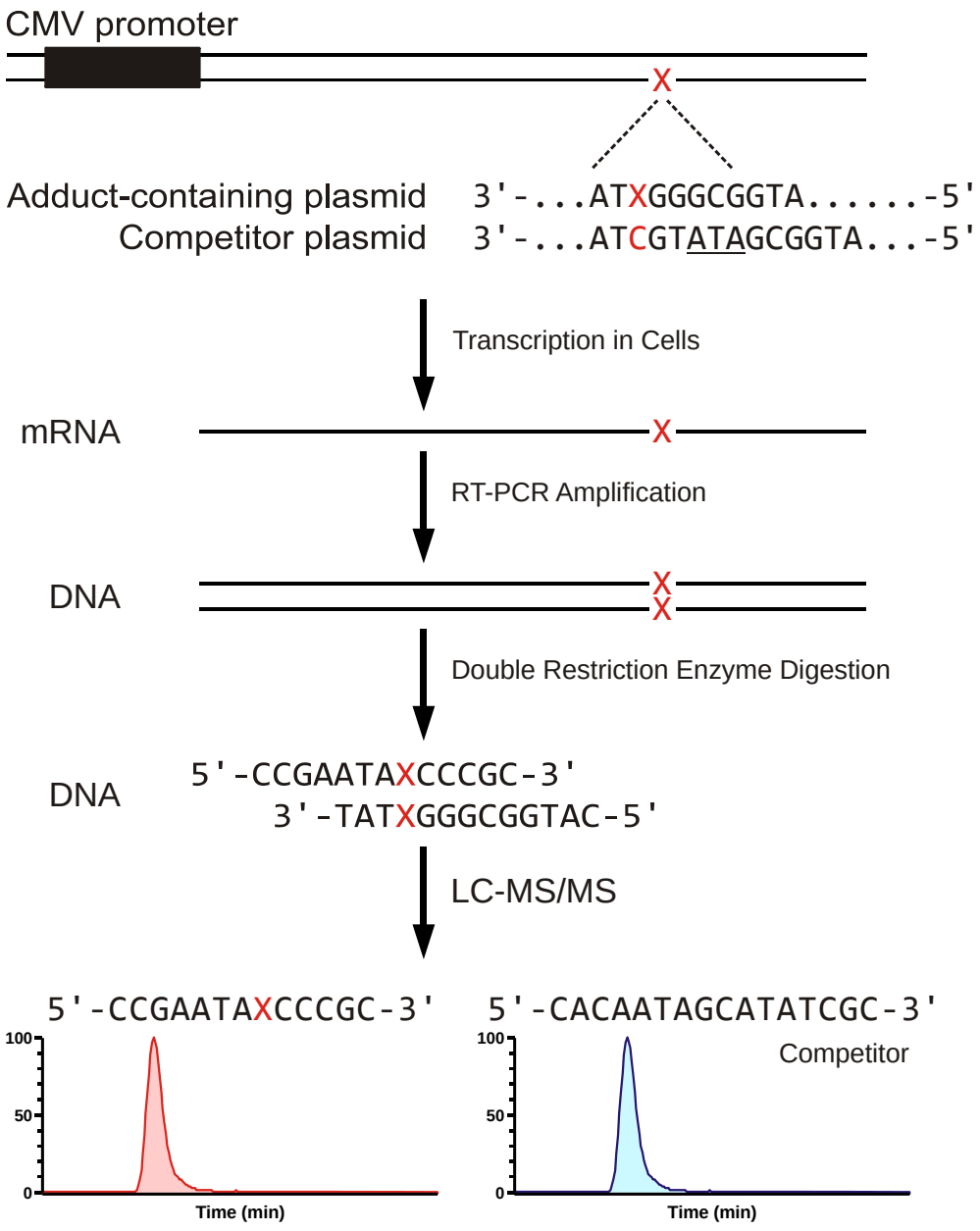
3.3.1. Generation of DPC containing Plasmid Substrate

Competitive transcription and adduct bypass (CTAB) assay is a powerful approach that makes it possible to examine the effects of DNA lesions on the efficiency and fidelity of transcription in cells[196]. To account for transfection efficiency, known amounts of competitor plasmids are transfected with lesion bearing plasmids. Because this competitor plasmids have three additional base pairs near the lesion position, RNA transcripts from two different plasmid can be distinguished with LC-MS/MS analysis. Following transfection of lesion bearing plasmids in cells, RNA transcripts are isolated and amplified through RT-PCR, and their sequences and amounts are examined by HPLC-ESI-MS/MS **(Scheme 3.1.)**. The assay requires a plasmid substrate where a lesion is incorporated at a specified site within the coding sequence of gene. The sequence of gapped region is located 56 nucleotides downstream of the CMV protomer region in our plasmid.

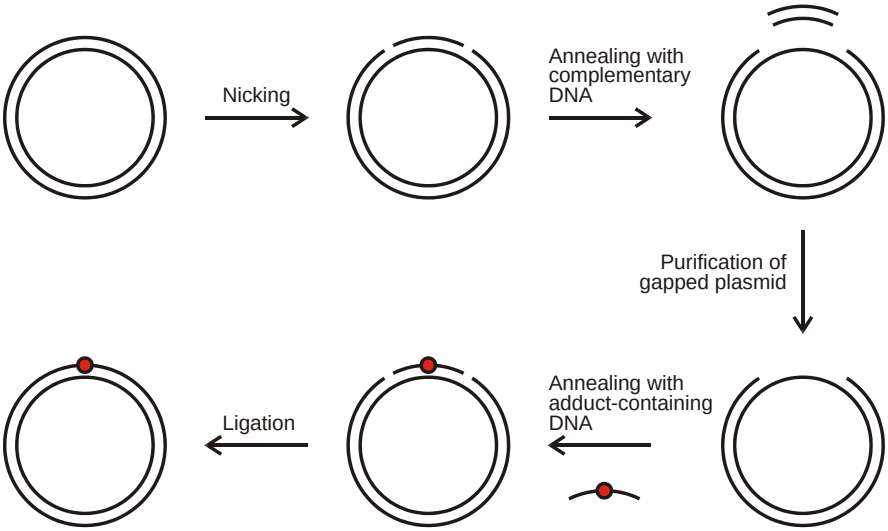
We employed a gapped plasmid approach to generate plasmids containing DNA-peptide and DNA-Lys conjugates conjugated to the C5 position of cytosine[106]. As described in our earlier publications, oligonucleotide substrates containing C5-dC DpCs (RPKPQQFFGLM-$CONH_2$) were generated by reductive amination **(Scheme 1.4.)**. To incorporate DpCs, pTGFP-H7 Hha10 plasmids were digested with Nt.BstNBI enzyme to generate a gapped plasmid **(Scheme 3.2.)**. The sequence of the gapped region is complementary to our synthesized DpC – containing 12-mers. Therefore, the DNA-peptide/lysine crosslink can be inserted via removal of the cleaved fragment, annealing the engineered 12-mer, and ligation.

The resulting lesion containing plasmids were purified with 100k Amicon filter, and the presence of DpC in the specified region of the plasmid was confirmed by gel shift. In brief, gel purified plasmids were subjected to NcoI and SfaNI restriction enzyme digestion, radiolabeling, and denaturing PAGE. The presence of DpC was confirmed by a characteristic gel shift (slower mobility) due to the presence of a DpC lesion **(Figure 3.1.)**. Because the competitor plasmid has three additional base pairs as compared to control plasmids between two restriction enzyme sites, lesion free control plasmid and competitor plasmid produced 13-mer and 16-mer oligonucleotides, respectively **(Figure 3.1.)**.

Scheme 3.1. Schematic of CTAB Assay System to Investigate the Effect of DNA-peptide Crosslink on Transcription The CTAB assay, adapted from You et al.[196], quantifies RNA transcripts from the plasmid to measure how a site-specific DNA lesion effect the efficiency and the fidelity of transcription in mammalian cell[197]. DNA lesion is marked as 'X' in the figure. A DNA lesion locate at the coding sequence of TurboGFP, which is the 56 nucleotides downstream of CMV promotor, to be transcribed.



Scheme 3.2. Construction of Lesion Bearing Plasmid . DNA lesion-containing plasmids were constructed by using gapped strategy. pTGFP-T7-Hha10T pTGFP-H7 Hha10 plasmids were digested with Nt.BstNBI enzyme to incorporate DpCs into the gapped region of plasmid. Nt.BstNBI enzyme digestion site is highlighted with bold. DpCs (5fC-11mer peptide conjugate, and 5fC-lys conjugate) were marked as 'X' in the figure. The DpC bearing strands were ligated to the plasmid by using T4 DNA ligase.



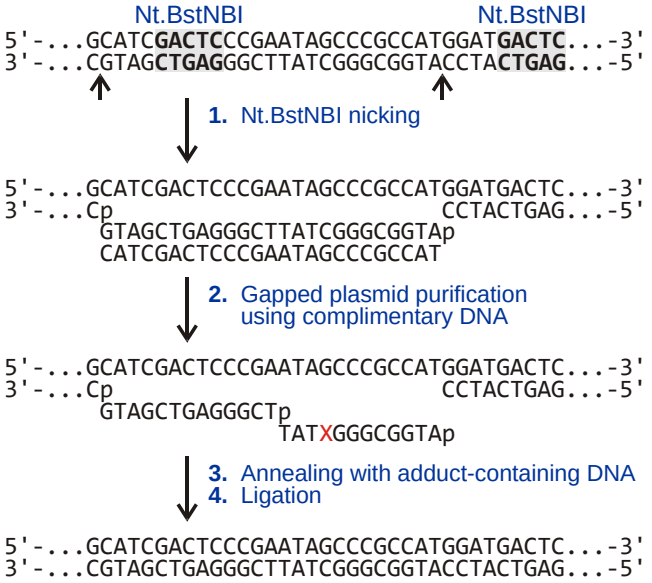
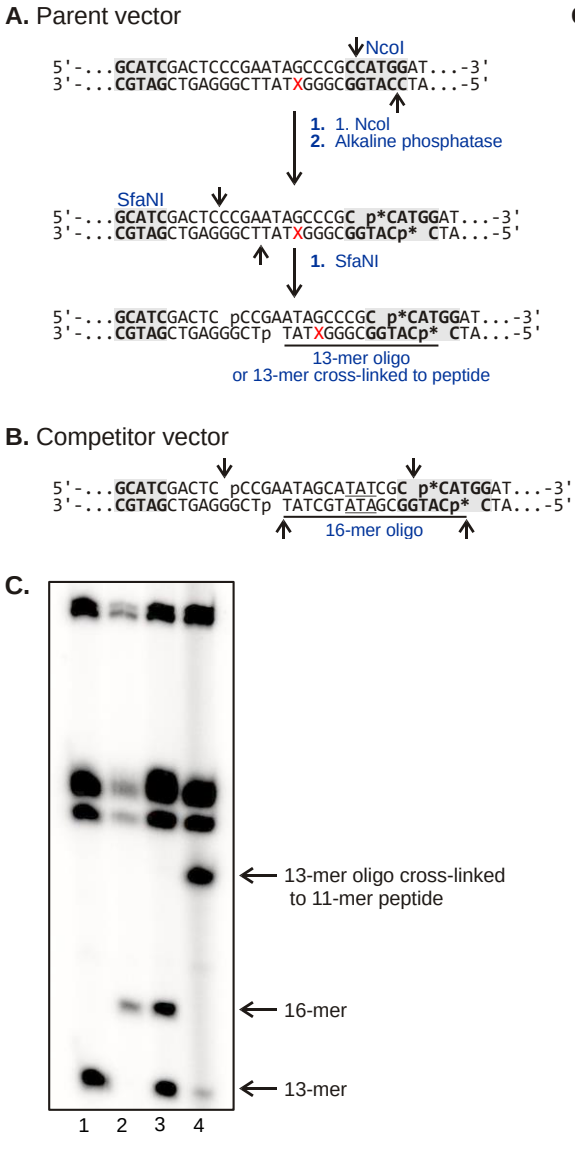


Figure 3.1. PAGE Analysis to Measure the Effect of DpC on Transcription RNA transcripts from A) lesion containing plasmids and B) competitor plasmids are reverse transcribed to cDNA, amplified by PCR. The resulting PCR products were subjected to enzyme digestion using NcoI and SfaNI, and p-32 labeling. The enzyme digestion sites are highlighted with bold and the ^{32}P-labeled phosphate group is marked as P*. C) The samples were loaded into 20% denaturing PAGE gel, and visualized by phosphoimaging. The 13mer-oligo with peptide conjugate showed slower mobility.



3.3.2. Influence of DNA-peptide and DNA-Lys cross-links on transcription in human cells

Lesion-bearing or lesion-free control plasmids were co-transfected into HEK293T cells along with the competitor plasmids, which are used to normalize for transfection efficiency. Because our control and competitor plasmids do not contain a replication origin, they only produce RNA transcripts with a different sequence with the same efficiency in transcription, allowing us to distinguish between RNA transcripts originating from the two plasmids. In order to compare transcription efficiency, RNA transcripts were isolated, amplified through RT-PCR, and subjected to restriction enzyme digestion to generate 13-mer and 16-mer oligodeoxynucleotides, respectively **(Figure 3.2.)**. These oligonucleotides were sequenced and quantified by using HPLC-ESI-MS/MS methodology **(Scheme 3.1)**.

We found that the presence of DNA-peptide/lysine crosslink had no influence on fidelity of transcription **(Figure 3.3.)**. No deletions or mis-insertions were detected on the position opposite to the DpC conjugate site. Only error-free products (5'-CCGAATAGCCCGC-3') (974.6650 m/z, **(Table 3.1.)**) were detected by HPLC-ESI-MS/MS. Similar results were observed for DNA-Lys conjugates. However, the presence of DNA-peptide/lysine crosslinks greatly reduced the amount of RNA transcripts **(Figure 3.3.)**. The bypass efficiency was decreased approximately 10-fold when the DNA-peptide crosslink was present in the plasmid. In addition, the DNA-lysine crosslink decreased bypass efficiency by 60%. Overall, this data suggests that RNA polymerase bypass rate is

affected by the DNA lesions and supports the importance of a DPC crosslink repair pathway.

Figure 3.2. Enzyme Digestion of Amplified PCR Product to Produce DNA Fragment for LC/MS-MS cDNA is amplified with a pair of primers to produce 527 bp size PCR product. The PCR products are enzymatically digested with NcoI and SfaNI to produce 13mer strand and 16mer strand from DNA-lesion bearing and competitor, respectively. The enzyme digestion was confirmed by detecting expected bands on 1% agarose gel (234bp, 147bp, and 66bp).

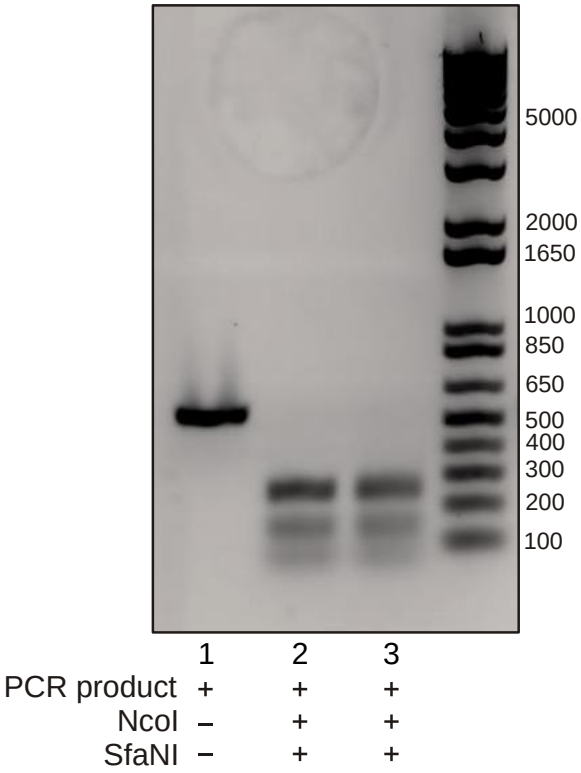
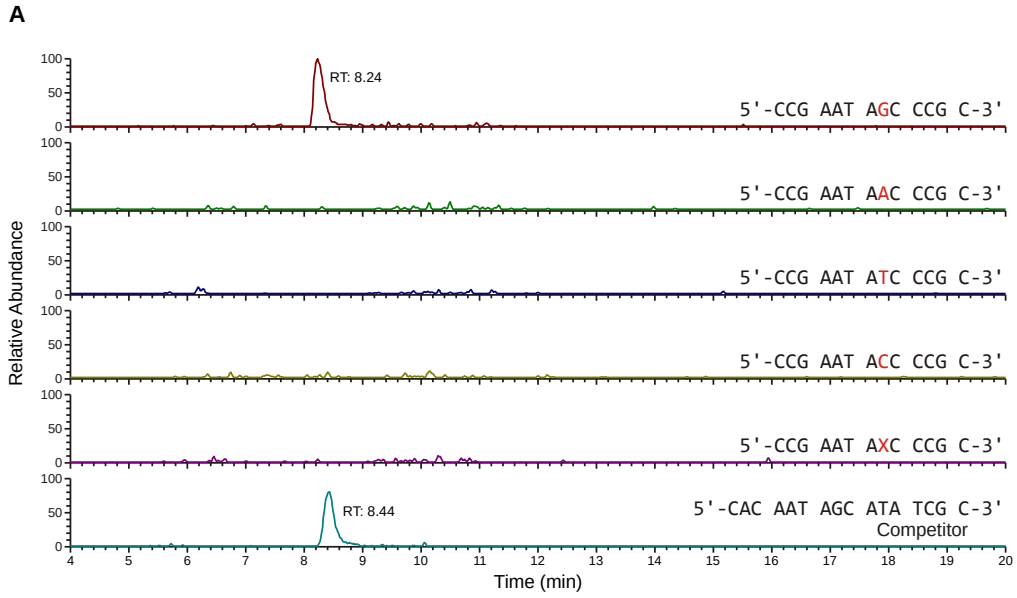
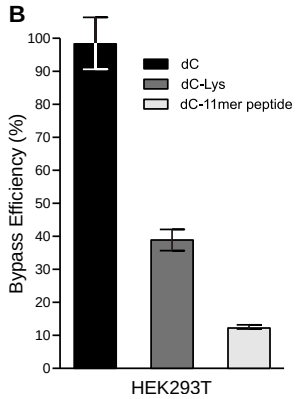


Figure 3.3. Influence of DNA-peptide and DNA-Lys Cross-links on Transcription in HET293T A) DpC bearing plasmid (5fC-11mer or 5fc-lys) and competitor plasmid are transfected to HET293T cells to produce cDNA from RNA transcripts. cDNA fragments from the control or lesion-bearing site (5'-CCGAATAXCCCGC-3', where X=A, T, C or G) and that from competitor plasmids (5'-CACAATAGCATATCGC-3') were observed to measure the bypass efficiency and the mutations. B) The size of conjugate is crucial for the bypass tolerance during transcription. The bypass efficiencies were dropped to ~39% and ~12% by 5fC-lys and 5fC-11mer peptide conjugates, respectively. The data are the average ± SD of three experiments.





DNA lesion	Control	5fC-lys	5fC-pep
Bypass Efficiency(%)	100	39.33	12.35
Std	7.9681	3.1179	0.7248

Table 3.1. Nucleobase Sequence and Mass Spectrometry Characterization of Strand from cDNA Fragment

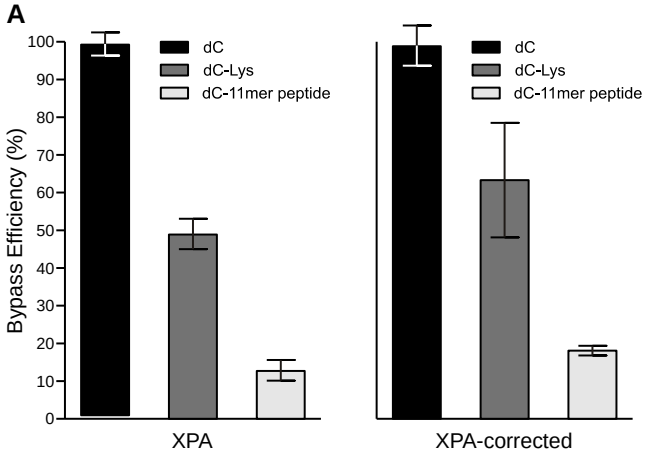
Sequence of Strand	Mutation	Observed Mass [m/z]
5'-CCGAATAGCCCGC-3'	None	$[M-4H]^{4-}$=974.665
5'-CCGAATAACCCGC-3'	C to T	$[M-4H]^{4-}$=970.667
5'-CCGAATATCCCGC-3'	C to A	$[M-4H]^{4-}$=968.414
5'-CCGAATACCCCGC-3'	C to G	$[M-4H]^{4-}$=964.664
5'-CCGAATACCCGC-3'	Deletion	$[M-4H]^{4-}$=892.402
5'-CACAATAGCATATCGC-3'	Competitor	$[M-4H]^{4-}$=1206.957

3.3.3. Influence of NER on Transcription of DpCs

In previous studies, it has been suggested that nucleotide excision repair can remove DpC lesions in cells. Nakano *et al.* later showed that DPC smaller than 10kDa can be excised by NER in mammalian cells, while baker *et al*. showed a 38kDa size DPC on plasmid can be repaired by NER directly[119,155]. Interestingly, Chesner *et al*. recently revealed that a DNA-crosslinked 39 amino acid polypeptide is repaired more efficiently than a 345 amino acid protein crosslink[126]. Therefore, CTAB assays was performed in XPA (NER-deficient) and XPA corrected (NER-proficient) cells to investigate the roles of nucleotide excision repair in the removal of these DNA lesions. Although XPA-corrected cells showed slightly higher bypass efficiency than that of XPA cells, the difference was not statistically significant **(Figure 3.4.)**. The presence of 5fC-lys and 5fC-11mer conjugates dropped the bypass efficiency to ~49% and ~12% respectively in XPA, decreased to ~64% and ~18% respectively in XPA corrected cells. However, no transcriptional mutations were observed.

Figure 3.4. Influence of NER on Transcription The effects on DpC formation are examined in NER deficient and proficient condition via CTAB assay. The amount of RNA transcripts in XPA were compared to that in XPA corrected. The data are the average ± SD of three experiments.

A



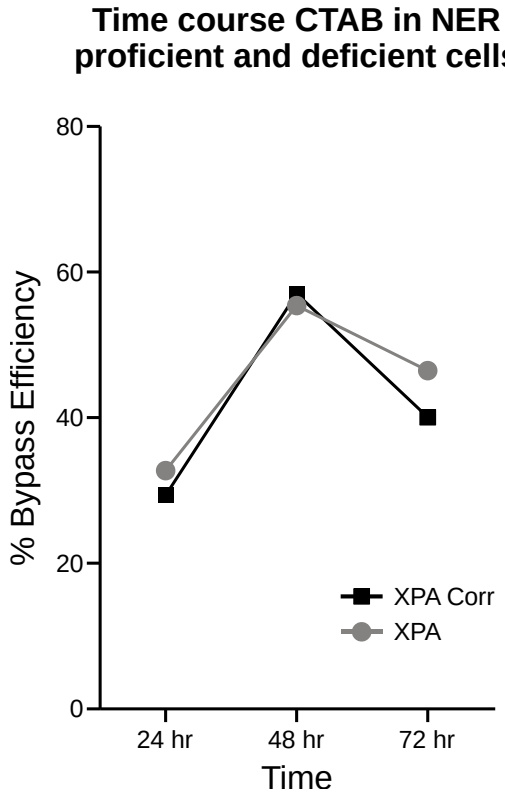
B

XPA Cells			
DNA lesion	**Control**	**5fC-lys**	**5fC-pep**
Bypass rate (%)	100	48.91	11.9
S.D.	3.1091	4.0075	2.288
XPA Corrected Cells			
DNA lesion	**Control**	**5fC-lys**	**5fC-pep**
Bypass rate (%)	100	63.99	18.05
S.D.	5.5552	15.4034	1.3781

3.3.4. Transcriptional Mutagenesis at Longer Time Points.

Because previous studies have shown that nucleotide excision repair can remove small DPCs, we hypothesized that DNA-peptide/lysine conjugates can be recognized by DNA damage repair pathways, including NER, to recover the RNA polymerase bypass efficiency. Since DNA-peptide/lysine conjugates decrease the bypass, the repair of the crosslink can be examined by the amount of transcripts observed. Therefore, we performed CTAB assay and RNA transcripts from lesion-bearing plasmid were collected at different time points to evaluate the repair. Because cells have a greater opportunity to repair conjugate sites with a longer incubation time, a higher bypass efficiency was observed at 48-hour timepoint compared to 24-hour points as we expected **(Figure 3.5.)**. Unexpectedly, the amount of RNA transcript decreased again at 72-hour point. We hypothesize that the decrease is caused by RNA degradation. In addition, no significant difference in bypass efficiency was detected at any timepoints between XPA and XPA-corrected cell groups. This data indicates that DNA-peptide/lysine conjugate could be repaired by nucleotide excision repair, but 5fC mediated conjugates were not active NER substrates. Therefore, there should be another pathway which can remove crosslinking damages when this repair pathway is not possible.

Figure 3.5. Time course CTAB assay in XPA and XPA Corrected Cells Cells have a greater opportunity to repair conjugate sites with a longer incubation time. After transfection of the DNA-peptide crosslink containing plasmid, cells were grown for 24, 48, or 72 h and subjected to RNA product collection to examine DpC repair via NER pathway.



3.4. Discussion

Recent papers have revealed that 5fC-histone DPCs form in living cells[106]. However, DPCs lesions are likely to block DNA replication, transcription, and repair[9]. Therefore, it is important to know how the presence of DPCs affect the efficiency and accuracy of DNA dependent pathways, and how cells recognize and repair the crosslinks to overcome the blockage or mutagenesis. To investigate this research question, our group has previously conducted primer extension assays and kinetic measurements with TLS polymerase and a DNA substrate containing 5fC site conjugates *in vitro*; the conjugate reduced the bypass efficiency during replication, and the bypass rate was decreased as the length of peptide conjugate increased[105]. However, it was still unknown whether 5fC mediated conjugates could affect the efficiency of transcription. Through the experiments in this chapter, such as the CTAB assay, it was shown that DNA-peptide/lysine conjugates can decrease the bypass of RNA polymerase during transcription. The-5fC-mediated peptide cross-link conjugate significantly reduced the RNA transcripts which indicates that the size of conjugate is a critical factor which determines the bypass tolerance.

Our results revealed that DpC lesions did not induce transcriptional errors. This in contrast to our in vitro primer extension assays that demonstrated that 5fC site conjugates can induce C to T transition mutations or deletions during replication[105]. Moreover, it was reported that RNA polymerase could bypass drug-induced DPCs or 5fC-peptide conjugates in an error-prone manner and induce mutations and deletions[184]. However, our CTAB assay data showed that the presence of 5fC site conjugates have no effect on the

sequence of RNA transcripts **(Figure 3.3.)**. I hypothesize that the structure of DNA lesions and conjugates has a strong impact on DNA and RNA polymerase bypass activity. Because an 11mer peptide that is crosslinked to a 5fC site is similar to 5fC-lysine crosslink formed in histones[105,106] and formed endogenously in the cell, cells may have an ability to tolerate this crosslink. Therefore, several different DNA-peptide crosslink structures should be investigated to elucidate the effect of DNA lesions on transcription.

Previous research has suggested that NER and HR orchestrate repair of DPCs and NER plays a role in removing DpCs (< 16kD) [34,116,119]. Our earlier *in vitro* NER assays show that DNA-peptide/lysine can be repaired by NER. In brief, DNA substrates containing 5fC site conjugates (11mer-peptide or lysine crosslink) were incubated with cell extract, and the resulting 24 – 30 mer oligonucleotide NER products were examined to detect the repair activities. These assays demonstrated that the regions containing DNA damage were excised by NER enzymes and that the 5fC-site conjugates can be repaired by NER, although the repair was not significant[184]. Our CTAB assay results also revealed that DNA-peptide/lysine conjugate could be repaired by nucleotide excision repair, but 5fC mediated conjugates were not active NER substrates **(Figure 3.4.)**. Therefore, the repair of 5fC site conjugates by a global NER pathway or transcription-coupled NER should be examined to elucidate the repair pathways. Interestingly, transcription-coupled NER (TC-NER) removes DNA lesion faster from the transcribed strand of gene than from the non-transcribed stand[198,199], while cell removes the RNAP-blocking DNA lesions by TC-NER to overcome prolonged blockage of transcription[200]. Moreover, CSB is recruited

to DNA lesion faster than XPC, whereas OGG1 binding appeared to precede CSB[201]. This suggest that TC-NER can repair DPCs and CTAB assay should be performed at different timepoints. Therefore, I hypothesize that this is the reason why 5fC site conjugates induced C to T transition mutation during replication, but not during transcription.

IV. Quantification of DNA-Protein Crosslinks induced by Monofunctional Alkylating Agents *in vivo*

Adapted from published article:

Yang, K., Park, D., Tretyakova, N. Y., and Greenberg, M. M. (2018) Histone tails decrease N7-methyl-2′-deoxyguanosine depurination and yield DNA–protein cross-links in nucleosome core particles and cells. *Proc. Natl. Acad. Sci. 115*, E11212–E11220.

This work was performed in collaboration with Dr. Mark Greenberg and Dr. Kun Yang. Kun Yang measured hydrolysis rate of MdG induced DPC *in vitro*, and Daeyoon Park performed MMS-induced-DPC quantification and identification *in vivo.*

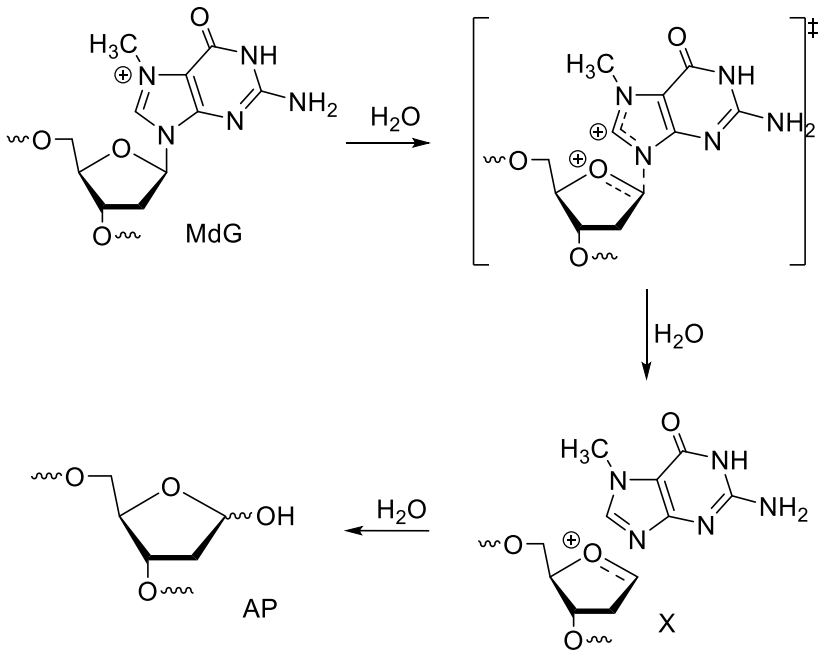
4.1. Introduction

In general, DPCs are produced by bifunctional alkylating agents; bifunctional alkylating agents target nucleophilic sites within proteins and DNA and after sequential alkylation, result in DPCs **(Figure 1.2.)**. Monofunctional alkylating agents can induce alkylated DNA structures which can undergo secondary reactions to target a site within proteins, thus also producing DPCs. However monofunctional agent induced DPC formations have not been fully investigated.

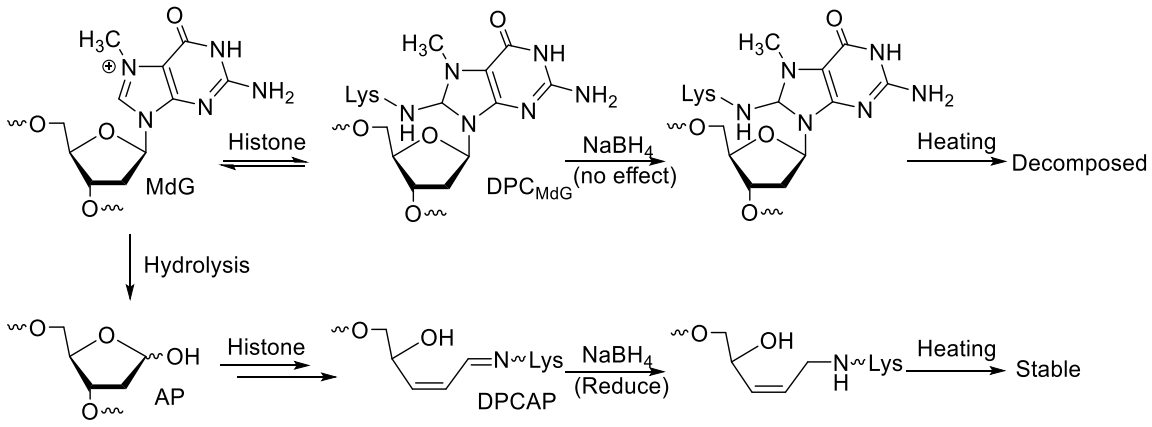
DNA alkylation has been investigated in many previous studies due to its important biological effects. For example, various endogenous and exogenous agents such as S-adenosylmethionine and methyl methanesulfonate (MMS) have been reported to induce various genotoxic methylation products, with N7-methyl-2′-deoxyguanosine (MdG) comprising 60 - 80% of the total methylated lesions[11,202–204]. The positive charge from the N7 position of MdG weakens the glycosidic bond and induces an abasic site through depurination. The resulting abasic site can then cause G to T transversion mutations through replication[205]. Moreover, MMS induced MdG cause base mispairing or block replication[86]. Although the steric effects of N7-methylation in the major groove only minimally perturb the duplex structure so that MdG can retain the positioning of dG, it has been shown that MdG encourages blockage of DNA polymerase during replication through tautomerization[206]. However, there is no clear evidence that tautomerization leads to a decrease in fidelity during bypass by a DNA polymerase. Therefore, its cytotoxicity has been ascribed to DPC formation because DPCs are known to block DNA replication and transcription.

N7-Alkylation encourages abasic sites via hydrolysis of the glycosidic bond[207]. Although the hydrolysis rate is slow with MdG, many studies have shown that DPCs can be generated by abasic sites (DPC_{AP}). Therefore, abasic sites are suggested as the main source of cytotoxicity from MdG lesions as shown in **Scheme 4.1.**. Although human cellular DNA is organized in chromatin and other DNA lesions interact closely with histone octameric nucleosome core particles (NCPs), most studies on MdG have been performed in free DNA condition[208–212]. Interestingly, Yang *et al*. suggested that the DPC can be formed by nucleophilic attack at the C8 position of MdG as shown in S**cheme 4.2.**[94]. Therefore, this study has demonstrated the possibility that MdG production is more than a source of abasic sites, but can also induce a secondary lesion including DPCs.

Scheme 4.1. Mechanism of MdG Hydrolysis The protonated base is detached from MdG and form an oxocarbenium ion (marked as X) via a unimolecular (S_N1) mechanism, resulting an AP site.



Scheme 4.2. MdG Mediated DPC DPC can be formed in two different ways at MdG. MdG can induced AP site, resulting DPC (DPC_{AP}), whereas the nucleophilic site of a protein, such as lysine, react with the C8 position of MdG site to form DPC on MdG (DPC_{MdG}). $NaBH_4$ reduces the Schiff-base within DPC_{AP} and stabilizes the structure, while DPC_{MdG} are not affected by the treatment. Therefore, DPC_{MdG} are released after heating due to their reversibility of the process.



MdG reactivity in nucleosome core particles (NCPs) is significantly different from that in free DNA[94]. By constructing synthetic 145-nt-long oligonucleotides with a specific MdG site, they demonstrated that hydrolysis is significantly retarded in NCPs due to the proximity and magnitude of positive charge from the lysine or arginine rich histone tail. In brief, MdG123 position, where the lesion was most proximal to the histone tail of H2A and H2B, showed the highest decrease in the hydrolysis rate. Moreover, hydrolysis rate was changed when the sequence of histone tail was mutated from lysine to alanine. This suggests that the octameric core of histone proteins play an important role in stabilizing the DNA structures against hydrolysis. Importantly, this study has shown that MdG directly reacts with histone proteins in nucleosome core particles (NCPs) to form DNA–protein cross-links (DPCs), which was not known to be induced by simple monoalkylation adducts. Unfortunately, monofunctional agent induced DPC formations have only been examined *in vitro*. Therefore, the experiments described below have been done to gain evidence of monoalkylation agent induced DPC formation *in vivo*.

4.2. Materials and Methods

Materials

Ammonium bicarbonate, ammonium acetate, phenylmethanesulfonyl fluoride (PMSF), leupeptin, pepstatin, aprotinin, methoxyamine, dithiothreitol (DTT), iodoacetamide, chloroform, ribonuclease A, deoxyribonuclease I, alkaline phosphatase, methanesulfonate, and sodium borohydride were purchased from Sigma (St. Louis, MO). Methyl UltraPure buffer-saturated phenol was obtained from Invitrogen (Carlsbad, CA). Mass spectrometry grade trypsin was purchased from Promega (Madison, WI). Pierce

C18 column and TMT 6-plex reagents were purchased from Thermo Fischer Scientific (Waltham, MA).

Methods

Cell Culture

Chinese hamster lung fibroblast cell lines V79 (GM16136) were obtained from the Coriell Institute for Medical Research (Camden, NJ) and cultured in Ham's F-12 modified essential Eagle's media supplemented with 10% fetal bovine serum and 1% penicillin streptomycin. Human fibrosarcoma (HT1080) cells were obtained from the American Type Culture Collection (Camden, NJ) and maintained as exponentially growing monolayer cultures in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin streptomycin, in a humidified incubator at 37 °C with 5% CO2.

Cell Treatment with MMS and Isolation of DNA-Protein Crosslinks

HT1080 cells (~1 x 10^7, in triplicate) and V79 cells (~2 x 10^7) in culture were treated with increasing concentrations of MMS (0, 10mM, and 25mM) for 3 h at 37 °C. The cells were washed and resuspended with phosphate-buffered saline (PBS) and DPC-containing DNA was extracted by a modified phenol-chloroform procedure as described previously[50,213]. In brief, the cells were lysed by adding an equal volume of 2X cell lysis buffer (20 mM Tris-HCl/10 mM MgCl2/2% Triton-X100/0.65 M Sucrose), incubated on ice for 5 min, and centrifuged at 2000 g for 10 min at 4 °C. The resulting pellets were re-suspended in saline-EDTA solution (2ml, 75 mM NaCl/24 mM EDTA/1% sodium

dodecyl sulfate) with RNase A (10 μg/mL) and a protease inhibitor cocktail (1 mM PMSF; 1 μg/mL pepstatin; 0.5 μg/mL leupeptin; 1.5 μg/mL aprotinin) for 2hr incubation at 37°C. After the incubation, an equal amount of tris-saturated phenol was added, and the solutions were mixed for 5 minutes. Following centrifugation at 4000 RPM for 10 minutes, the aqueous layer containing DNA and the DPC-containing white interface layer were collected. Two additional phenol/chloroform extractions were performed, and the DPC containing DNA was precipitated with cold ethanol at -20 °C. The samples are dissolved in 400ul of 10mM HEPES, pH 7.5. The amount of DNA was measured using UV spectrophotometry (A260) and subsequently determined by dG analysis.

$NaBH_4$ Treatment to Distinguish Between Types of DPC in HT1080

To distinguish AP site DPCs (DPC_{AP}) from MdG site DPCs (DPC_{MdG}), a reduction step was performed with $NaBH_4$. After DPC extraction from MMS treated or untreated cells, 20μg of DNA (18ul, 10mM HEPES pH 7.5) was mixed with fresh ice-cold 1M $NaBH_4$ (2μl), and incubated for 1hr at 4 °C. The samples were quenched with 1M acetic acid (1.5μl) and a pH of 7 - 7.5 was confirmed with pH paper. Samples were heated to 70 °C for 1 hr and resolved by NuPAGE Novex 4 - 12% Bis-Tris gels (Invitrogen, Carlsbad, CA), and visualized by staining with SimplyBlue SafeStain (Invitrogen, Carlsbad, CA).

Gel Based DPC Detection in MMS-treated HT1080 cells

HT1080 cells (~1 x 10^7, in triplicate) were treated with MMS (0, 10mM, 25 mM) for 3 h at 37 °C. DNA was extracted and quantified in a manner to preserve DPCs, as described above. 20 ug of DNA was resolved by NuPAGE Novex 4 - 12% Bis-Tris gels

(Invitrogen, Carlsbad, CA), and visualized by staining with SimplyBlue SafeStain (Invitrogen, Carlsbad, CA).

Mass Spectrometric Identification of Cross-Linked Proteins

Gel lanes (0 - 30kDa area) were excised and further diced into 1 mm pieces. The standard in-gel tryptic digestion method was performed with the gel pieces. In brief, gel pieces were washed with 1:1 100 mM ammonium bicarbonate:acetonitrile, and rehydrated with 10 mM DTT in 50 mM ammonium bicarbonate (1hr, 56°C). After removing DTT through centrifugation, the samples were alkylated with saturated iodoacetamide. Gel pieces were dehydrated with digestion buffer (50 mM NH_4HCO_3, 5 mM $CaCl_2$, 5 ng/µl trypsin) and incubated on ice for 15 min. The supernatant was removed and replaced with 50 mM NH_4HCO_3, 5 mM $CaCl_2$. for an overnight incubation at 37 °C. The resulting tryptic peptides were extracted with 60% acetonitrile containing 0.3% aqueous formic acid, evaporated to dryness, and desalted using Pierce C18 column. Samples were reconstituted in 0.1% formic acid in water (10µl) for HPLC−ESI^+−MS/MS analysis and analyzed using an LTQ Orbitrap Fusion mass spectrometer (Thermo Scientific) interfaced to a Dionex Ultimate 3000 RS autosampler and Dionex UltiMate 3000 RSLCnano system. 0.1% formic acid in water (Solvent A) and 0.1% formic acid in acetonitrile (Solvent B) with flow rate of 0.3 µL/min were used for the system and the samples were loaded onto 5 µm Luna-C18 resin (Phenomenex, Torrence, CA). Solvent composition was held at 5% solvent B for 6min, followed by an increase to 7% over 2 min, 25% over 57 min, 60% over 12 min, and finally to 95% over 1 min. The solvent composition was held at 95% for 4 min, and returned to 5% over 0.5 minute and re-

equilibration for 7 min. The mass spectrometer was operated in a data dependent mode with dynamic exclusion enabled (repeat count: 1, exclusion duration: 15 sec). Every scan cycle, one full MS scan (*m/z* 320 to 2000) was collected at a resolution of 120,000 at an AGC target value of $2x10^5$, followed by MS2 scans of as many dependent scans as possible within a cycle time of 3 seconds with HCD (normalized collision energy = 38%, isolation width = 0.7 *m/z*, resolution = 30,000) at an AGC target value of $1x10^5$. Ions with a charge state of +1 were excluded from the analysis.

Mass Spectrometric Identification and Quantification of DPC Formation in Chinese Hamster Lung Cells.

DPCs were extracted from MMS (0 or 25mM) treated V79 cells as described above. Based on dG analysis, aliquots containing 30μg of DNA were prepared in 200μl of 100 mM HEPES buffer (pH 7.5). To distinguish between cross-links involving MdG and AP, 30μg of DNA was incubated with 22μl of 1M $NaBH_4$ (1h, 4 °C) to stabilize the DPCs involving AP and quenched with an appropriate amount of 1M acetic acid until the sample reach pH 7.5. DPC-containing DNA was subjected to neutral thermal hydrolysis (1 h, 70 °C), incubated with 3 volume equivalents of cold ethanol (1h, -20 °C), and centrifuged at 14,000g for 30 min. The supernatants were dried under vacuum and dissolved in 100mM HEPES (pH 8.0) and subsequently washed to remove salts using Amicon 3K filters (14,000g, 10 min). The HEPES buffer was exchanged with ammonium bicarbonate buffer (25 mM, pH 8.0) using Amicon 3K filters and concentrated to ~200 μL. After resuspension in 200μl of 25mM ammonium bicarbonate, each sample was incubated with 1μg trypsin (overnight, 37 °C). The resulting tryptic peptides were

recovered by centrifugation (14,000g, 15 min) and concentrated to dryness in glass MS vials. The dried samples were re-suspended in 35µl of 100mM HEPES, followed by incubation with 5µl of ACN and 10µl of TMT 6-plex reagent (19.5µg/µL) for 2 hours at RT. The reactions were quenched with 4µl of 5% hydroxylamine (15min, RT). The peptides from each condition (control, $NaBH_4$, MMS, MMS + $NaBH_4$) were combined and dried, and then underwent desalting through a Pierce C-18 column prior to nanoLC-MS analysis (Thermo Orbitrap Fusion). TMT-labeled peptides were reconstituted in 8µL 0.1% formic acid and loaded onto a pulled-tip fused silica column with a 100 µm inner diameter packed in-house 5µm Luna-C18 resin (Phenomenex, Torrence, Ca). Same setting is used for LTQ Orbitrap Fusion mass spectrometer as described above. The raw data was searched against UniProt Chinese Hamster database with Thermo Scientific Proteome Discoverer 2.1 software. The database searches were carried out using the following settings in the Sequest HT algorithm: precursor ion mass tolerance, 10 ppm; Fragment mass tolerance 0.02 Da; variable modification, TMT6plex, Methylation, Dimethylation, Trimethylation, and Acetylation; allowed number of mis-cleavages, 3. To elucidate the crosslinked histone family, a smaller UniProt database which contained all proteins that were detected in analyzed samples was established.

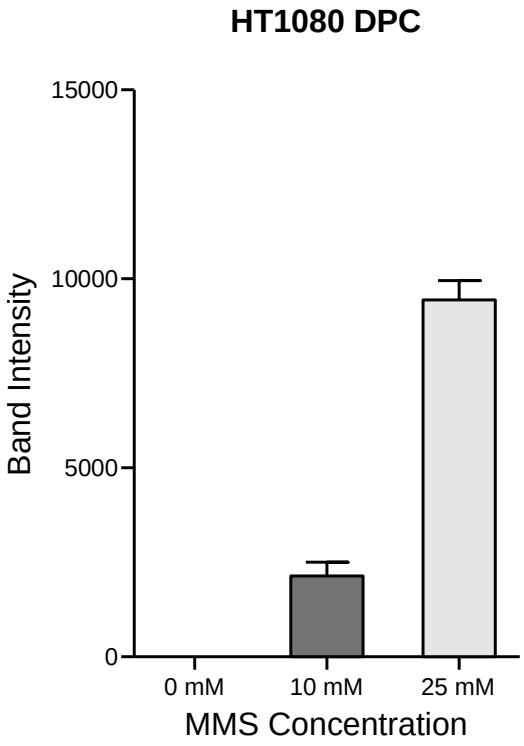
4.3. Results

4.3.1. Concentration-dependent Formation of DPCs in MMS-treated Human Cells

DNA was extracted under conditions that preserve the DNA-protein conjugate from human fibrosarcoma cells (HT1080) that were treated with increasing concentrations of MMS (0, 10, and 25 mM) for 3 h. The samples were hydrolyzed by heat to release

proteins conjugates so that they could be resolved by gel electrophoresis. Small amounts of endogenous DPCs were present in untreated cells, whereas the amounts of cross-linked proteins increased in a concentration dependent manner with MMS treatment **(Figure 4.1.)**. However, compared to the dose of anticancer drugs, such as cisplatin, used in previous studies[195], a relatively higher concentration of MMS was required to detect the DPCs. This suggests that MMS can induce DPCs, but it has a lesser ability to induce DPCs than cisplatin.

Figure 4.1. DPC Band Intensities from MMS treated HT1080 The DPCs from MMS treated HT1080 cells were resolved by SDS-PAGE gel (NuPAGE Novex 4 -12% Bis-Tris). The intensities of bands were digitized by ImageJ program. The data are the average ± SD of three experiments.



4.3.2. Identification of MMS Induced Cross-linked Proteins in HT1080 cells

To identify MMS induced DPCs in HT1080 cells, cells were treated with 25 mM MMS for 3 h and the resulting DPCs were compared to that from untreated cells. In addition, HT1080 cells were treated with both MMS and $NaBH_4$ to identify the AP site DPCs (DPC_{AP}) and MdG site DPCs (DPC_{MdG}). DPC containing DNA was extracted as described above, and the samples were loaded onto a SDS-PAGE gel (4 - 12%, NuPage) for gel electrophoresis. Because a lysine residue of NCP histone can be covalently linked to DNA via AP site (DPC_{AP}) or the C8 position of MdG (DPC_{MdG}) **(Scheme 4.2.),** a molecular weight range of 0 - 30kDa region was excised from the gel lanes. The gel pieces from each lane were subjected to in-gel tryptic digestion to obtain peptides, and the resulting peptides were extracted, desalted, and analyzed with HPLC-ESI^+-MS/MS methodology for protein identification. The amino acid sequence of tryptic peptides from cross-linked proteins were characterized by performing b and y series ions test, and at least two unique peptides were used to identify the crosslinked proteins. A total of 30 proteins were identified **(Table 4.1.)**. Interestingly, Histone H4, H2B, and H2A were detected from this region. A large portion of the cross-linked proteins were known to participating in translation, transcription and chromatin organization. Moreover, the amounts of crosslinked Histone H4, H2B, and H2A were increased dramatically in MMS treated groups, whereas MMS and $NaBH_4$ treated groups showed only a slight increase **(Figure 4.2.)**. Unfortunately, Histone H3 didn't show a similar pattern as the other histone groups. Because the amount of crosslinked histone was estimated based on the MS peak area from the raw data through MaxQuant program, the accuracy can be increased by performing other measurement. Therefore, another method, such as TMT-

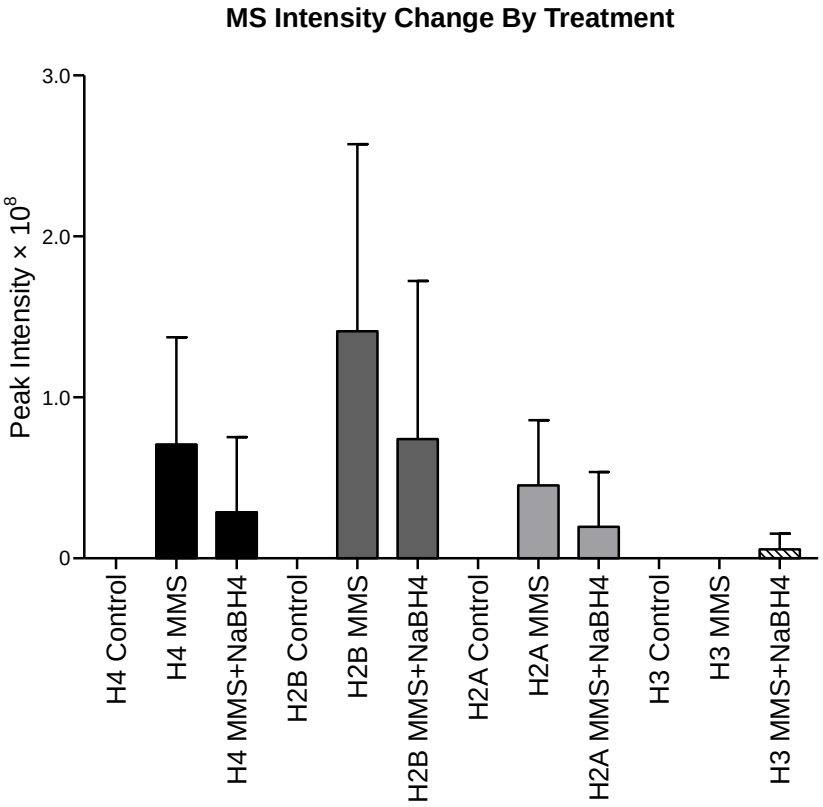
tag, should be performed to determine the accurate amount of histone in each experimental group.

Table 4.1. Proteins Forming Covalent Crosslink to DNA in MMS treated HT1080

Gene Name	Accession	Identified Proteins	% Coverage	# Peptides	No. of Unique Peptides	Primary biological Function	MW [kDa]
HMGA1	P17096	High mobility group protein HMG-I/HMG-Y	48.6	5	2	transcription	11.669
HMGA1	P17096-2	Isoform HMG-Y of High mobility group protein HMG-I/HMG-Y	42.7	4	1		10.673
HMGA2	F5H2A4	High mobility group protein HMGI-C	38.1	3	3		12.707
HMGA1	P17096-3	Isoform HMG-R of High mobility group protein HMG-I/HMG-Y	14.0	2	2		19.682
RPS25	P62851	40S ribosomal protein S25	20.8	3	3	translation	13.734
RPL22	P35268	60S ribosomal protein L22	18.8	2	2		14.778
RPL22L1	H0Y8C2	60S ribosomal protein L22-like 1 (Fragment)	8.5	1	1		16.89
RPS20	P60866-2	Isoform 2 of 40S ribosomal protein S20	7.7	1	1		15.995
RPS9	P46781	40S ribosomal protein S9	8.8	2	2		22.578
RPS16	Q6IPX4	40S ribosomal protein S16	13.8	2	2		17.096
RPL28	H0YKD8	60S ribosomal protein L28	6.5	1	1		19.06
RPL19	P84098	60S ribosomal protein L19	4.6	1	1		23.451
RPS15A	P62244	40S ribosomal protein S15a	20.0	3	3		14.83
HIST1H4A	P62805	Histone H4	50.5	5	5	chromatin organization	11.36
HIST1H2BN	U3KQK0	Histone H2B	17.5	3	3		18.792
HMGN1	P05114	Non-histone chromosomal protein HMG-14	34.0	3	3		10.653
HIST1H3D	A0A0U1RRH7	Histone H2A	9.4	2	2		18.541
H3F3B	K7ES00	Histone H3.3 (Fragment)	8.6	2	2		16.611
HIST1H2BJ	P06899	Histone H2B type 1-J	28.6	1	1		13.896
LYZ	P61626	Lysozyme C	43.9	8	8	immune system	16.526
DCD	P81605-2	Isoform 2 of Dermcidin	34.7	4	4		12.406
S100A8	P05109	Protein S100-A8	39.8	4	4		10.828
S100A9	P06702	Protein S100-A9	31.6	3	3		13.234
DEFA1	P59665	Neutrophil defensin 1	19.1	2	2		10.194

FABP5	Q01469	Fatty acid-binding protein, epidermal	13.3	2	2		15.155
S100A7	P31151	Protein S100-A7	21.8	2	2		11.464
IGLC2	A0A075B6K9	Ig lambda-2 chain C regions (Fragment)	9.4	1	1		11.34
PIP	P12273	Prolactin-inducible protein	13.7	2	2		16.562
PRDX1	Q06830	Peroxiredoxin-1	27.6	6	5	oxidative stress regulation	22.096
PRDX2	P32119	Peroxiredoxin-2	13.6	3	2		21.878
TXN	P10599	Thioredoxin	21.0	2	2		11.73
HBA1	P69905	Hemoglobin subunit alpha	16.9	2	2	oxygen transport	15.248
RPS27A	P62979	Ubiquitin-40S ribosomal protein S27a	30.1	4	4	cellular regulation	17.953
LCN1	P31025	Lipocalin-1	17.0	3	3		19.238
CALML5	Q9NZT1	Calmodulin-like protein 5	5.5	1	1		15.883
LGALS7	P47929	Galectin-7	19.1	2	2		15.066
CSTA	P01040	Cystatin-A	15.3	2	2		11
PPIA	P62937	Peptidyl-prolyl cis-trans isomerase A	9.1	2	2	protein folding	18.001

Figure 4.2. MMS and $NaBH_4$ Treatment to Distinguish the Types of DPC The MMS treatment induced the histone crosslink to chromosomal DNA in HT1080. Additional $NaBH_4$ treatment, which reduces the Schiff-base within DPC_{AP}, were used to estimate the amount of $DPC_{AP\ and}$ $DPC_{MdG.}$ The intensities of histone peaks were measured by HPLC−ESI^+−MS/MS. The data are the average ± SD of three experiments.



4.3.3. Quantification of MMS Induced Cross-linked Proteins in V79 cells

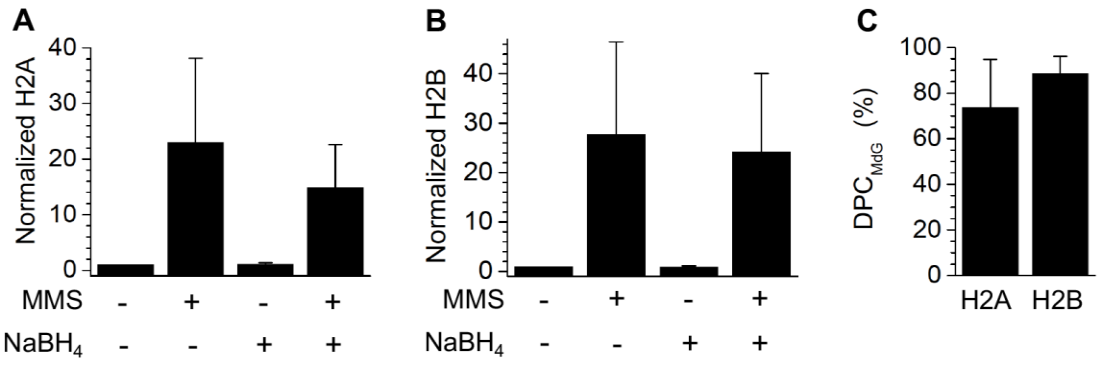
DPC formation between histone proteins and MdG was also observed in MMS-treated V79 Chinese hamster lung cells. In this experiment, V79 cells were used instead of HT1080 cells, because DPC repair is less efficient in this cell line compared to that in human cells[214]. To quantify the MMS induced DPCs and to distinguish between AP site DPCs (DPC_{AP}) and MdG site DPCs (DPC_{MdG}), V79 cells were incubated either with or without MMS, and the DNA extracted from the two groups was incubated either with or without $NaBH_4$. All four samples (control, $NaBH_4$, MMS treated, MMS + $NaBH_4$ treated) were subjected to tryptic digestion, desalting, and quantitative proteomics analysis incorporating TMT isotope tags as described above. This experiment clearly demonstrated the formation of DPCs with several proteins in MMS treated cells **(Table 4.2)**. A total of 17 proteins, including H2B, and H2A, were identified. Interestingly, a large portion of the cross-linked proteins were known to participating in translation, transcription and chromatin organization. Moreover, mass spectrometry-based TMT proteomics showed that the amounts of crosslinked Histone were increased dramatically in MMS treated groups, whereas MMS and $NaBH_4$ treated groups showed only a slight increase. Specifically, crosslinked histones H2A and H2B levels were increased 23-fold and 28-fold respectively when comparing the MMS treated group to control group **(Figure 4.3.)**. DPC_{MdG} and DPC_{AP} were distinguished by their different thermal stabilities **(Scheme 4.2.)**. Yang *et al.* showed that after $NaBH_4$ treatment (0.1 M, 4 °C, 1 h) and heating (70 °C, 1 h), no decomposition was observed in DPC_{AP123} while DPC_{MdG123} was completely decomposed *in vitro*[94]. Similar to *in vitro* data, we found that following $NaBH_4$ treatment, ~75% of H2A and ~90% of H2B DPCs were thermally released. This

indicates that the majority of the DPCs in MMS treated cells were from MdG, with the remaining conjugates forming at AP sites. Therefore, we conclude that a monofunctional alkylating agent, such as MMS, can induce DPCs by forming the DNA lesions which will undergo secondary reactions.

Table 4.2. Heat map analysis of the TMT abundance data for MMS-induced DPCs in V79 cells All four samples (control, $NaBH_4$, MMS treated, MMS + $NaBH_4$ treated) were TMT tagged to quantify the MMS induced DPC_{MdG} and DPC_{AP}. The amounts of crosslinked protein were increased dramatically in MMS treated groups, while NaBH4 treatment had no effect on DPC formation.

	Abundance Ratio			
	Control	MMS	NaBH4	MMS+NaBH4
High mobility group protein HMG-I/HMG-Y	0.17	1.986	0.034	1.411
Nucleolin	0.09	2.157	0.048	1.251
High mobility group protein HMGI-C	0.129	1.874	0.1	1.247
Vimentin	0.138	1.706	0.082	1.782
Nucleophosmin	0.101	1.293	0.11	2.183
PC4 and SFRS1-interacting protein-like isoform 2	0.048	1.765	0.072	1.55
Histone H2B	0.028	1.968	0.027	1.922
Histone H2A type 1	0.069	2.119	0.06	1.254
Sp110 nuclear body protein	0.169	1.074	0.061	2.998
Heterogeneous nuclear ribonucleoprotein A/B	0.059	2.145	0.061	1.291
Chloride intracellular channel protein	0.01	0.553	0.113	3.287
Clathrin light chain	0.11	1.8	0.142	2.324
Heat shock cognate protein	0.928	2.067	0.153	2.449
ATP-binding cassette sub-family F member 1-like protein	0.024	1.435	0.021	2.355
Ribosome-binding protein 1	0.01	2.003	0.01	1.395
Actin, cytoplasmic 1	0.08	0.99	0.102	1.769
ucleolar and coiled-body phosphoprotein 1-like isoform 1	0.095	1.667	0.049	2.13

Figure 4.3. Histone DPC Formation in MMS-treated V79 cells The normalized abundance of thermally released **A)** H2A and **B)** H2B within MMS induced DPCs from V79 cells. The DPC abundance under each condition was normalized to that without MMS and NaBH4 treatment. **C)** The percentage of the thermally released histone proteins after NaBH4 treatment were compared with those released without NaBH4 treatment were used to calculate the ratio of DPC_{MdG} in MMS-treated V79 cell. ~75% of H2A and ~90% of H2B DPCs were thermally released. The data are the average ± SD of three experiments.



4.4. Discussion

MdG is known as a non-toxic and non-mutagenic DNA lesion although it comprises 60 - 80% of the total methylated lesions[11,215]. However, the hydrolysis of MdG induces abasic site[216], resulting DPC, while a nucleophilic site of histone proteins makes a crosslink directly at the C8 position of MdG **(Scheme 4.2.)**. Yang *et al.* compared the rate constants for hydrolysis in NCPs to those in free DNA for the first time[94]. The hydrolysis rate was decreased one-half when NCPs are present, suggesting that the histone proteins stabilized the DNA against abasic site formation. Moreover, MdG positions close to positively charged histone tails showed greater decrease in hydrolysis rate than those further from the tails. As shown in **Scheme 4.1.**, MdG hydrolysis can occur via the S_N1 mechanism. Therefore, the presence of a positive histone tail stabilizes the ground state and decreases the amount of transition state MdG, and thus abasic site formation is suppressed in a cellular environment where cellular DNA interacts closely with NCPs.

DPCs induced by monofunctional alkylating agents have not been well documented. Interestingly, this study showed for the first time that DPCs are also formed in cells treated with a monofunctional alkylating agent. The relatively long half-life of MdG (~200 hour) allows histone proteins to be crosslinked before depurination occurs. Our LC-MS/MS analysis of MMS-treated cells also demonstrated that more than 75% of histone crosslinks are produced at the C8' position of MdG whereas abasic sites induce the remaining less than 25% of DPC_{AP}. These findings suggest that an alkylating agent can induce DPCs not only via abasic site formation, but also via DNA alkylation which can undergo secondary reactions. Because DPCs, which block DNA dependent pathways

such as replication and transcription, are extremely toxic, the cytotoxicity of monofunctional alkylating agent should be examined more carefully.

MMS showed higher cytotoxicity in a NER deficient cells[217,218]. However, the repair of MMS induced DPCs has not been investigated because the possibility of DPC formation by a monofunctional agent had been dismissed. In short, the possible lesions produced by monofunctional alkylating agents should be studied further and their repair pathways, biological effects and consequences should be investigated in future studies. Moreover, this also raises the question of whether chemists could increase DPC yields in cellular DNA and/or exploit their formation to improve the efficacy of drugs such as temozolomide.

V. Summary and Conclusion

In Chapter 2 of this thesis, we quantified free radical induced DPC *in vivo* with sensitive nanoLC-ESI$^+$-MS/MS methodologies. Since DPCs produced by IR exposure are identical to those formed by ROS, MEF cells are treated with IR to examine whether IR can generate radical induced DPCs. By using an isotope dilution tandem mass spectrometry assay for dT-Tyr conjugates, the correlation between IR exposure and DPC formation has been demonstrated **(Figure 2.2.)**. Furthermore, dT-Tyr quantification in IR treated C57BL/6 wild type mouse showed that ROS-induced-DPC can be generated by IR in mouse models as well **(Figure 2.6.)**. Therefore, it has been shown that nanoLC-ESI+-MS/MS assay can be applied to analyze DPC and to observe direct evidence of DPC formations both in cells and mouse models. Moreover, the sensitivity of nanoLC-ESI+-MS/MS assay can be improved further by adapting several methods, such as SPE **(Figure 2.5.)**. We also utilized our nanoLC-ESI$^+$-MS/MS assay to examine DPC repair pathways. For example, the role of SPRTN in DPC repair was revealed by comparing the amount of dT-Tyr in SPRTN proficient and deficient cells **(Figure 2.2.)**. The increased amount of dT-Tyr in SPRTN deficient cell lines is direct evidence for the role of SPRTN as a protease functioning in the repair of DPCs. In addition, a clonogenic assay successfully showed not only the cytotoxicity of radical induced DPCs but also the importance of SPRTN in DPC repair; SPRTN deficient cells were more sensitive to IR exposure than SPRTN proficient cells **(Figure 2.5.)**. Furthermore, the amounts of endogenously formed ROS induced DPCs were increased by 1.5~2 fold in 10-month old wild type (SPRTN$^{+/+}$) compared to that in Spartan hypomorphic mouse (SPRTN$^{H/H}$) (**Figure 2.6.)**. Therefore, the role of other proteases which can play similar role as SPRTN should be examined in

future studies. The role of immunoproteasome in DPC repair was tested in chapter 2 because the amount of dT-Tyr was lower in tissue where lymphatic immunoproteasomes are present. To our knowledge, this is the first report investigating the contribution of the immunoproteasome to DPC repair. DPCs were extracted from the tissues from the wildtype and immunoproteasome knockout mice and the amount of dT-Tyr was quantified in each group **(Figure 2.7.)**. The amount of dT-Tyr was higher in the wildtype, indicating that immunoproteasome mediated proteolysis is not a major pathway for DPC repair. This supports our previous hypothesis that when SPRTN repair of DPCs is not possible, other proteases recognize ubiquitinated crosslinked proteins to degrade them or other canonical pathways will remove crosslinking damages.

Because 5fC epigenetic mark is hypothesized to regulate gene expression and to form DPC using its aldehyde group which block DNA replication, transcription, and repair[9,105,106], the effect of 5fC-DpC on transcription has been examined *in vivo* through CTAB assay. Control and competitor pTGFP-H7 Hha10 plasmids were enzymatically digested to produce a gapped plasmid, and synthesized 5fC-mediated peptide/lysine conjugates were introduced to the gapped region to establish a specific DNA damage containing plasmids **(Scheme 3.3.)**. We confirmed the establishment of 5fC site conjugates with gel electrophoresis by observing slower mobility phase on a PAGE gel **(Figure 3.1.)**. Interestingly, the RNA transcripts from the transfected plasmids showed that RNA polymerase bypass efficiency is decreased by the DNA lesions and highlighted the importance of a DPC crosslink repair pathway. Compared to the control, 5fC-lys and 5fC-peptide conjugate decreased the amounts of RNA transcripts by ~61% and ~88%,

respectively **(Figure 3.3.)**. This revealed that the size of conjugate is a critical factor which determines the bypass tolerance.

Our CTAB assay showed that DNA-peptide/lysine conjugate could be repaired by NER pathway. Although the difference was not significant, the bypass efficiency was slightly higher (18% or 6%) in XPA corrected cells compared to that in XPA **(Figure 3.4.)**. This data indicate that DNA-peptide/lysine conjugate could be repaired by nucleotide excision repair, but 5fC mediated conjugates were not active NER substrates. Therefore, several different DNA-peptide crosslink structures should be investigated with CTAB assay to elucidate the effect of DNA lesions on transcription. Moreover, no significant difference in bypass efficiency was detected at any timepoints between XPA and XPA-corrected cell groups **(Figure 3.5.)**, while the presence of 5fC site conjugates was not mutagenic for the sequence of RNA transcripts. Because RNA transcripts were produced during CTAB assay, it is possible that 5fC-site conjugates are repaired by TC-NER rather than global NER. Since 5fC site conjugate was mutagenic during the replication[105] and the transcription *in vitro* **(Ji and Tretyakova, submitted)**, the involvement TC-NER in repairing 5fC-site conjugates should be examined. In conclusion, several different DNA-peptide crosslink structures and different repair pathways should be investigated with CTAB assay to elucidate the effect of DNA lesions on transcription, and to elucidate the mechanism of DPC repair pathways.

In Chapter 4 of this thesis, we investigated monofunctional alkylating agent induced DPC formation *in vivo*. Typically, DPCs are produced by bifunctional alkylating agents.

However, monofunctional alkylating agents also generate DPCs. When DNA is treated with MMS, MdG is produced as a major methylation product. This DNA lesion can induce DPCs directly through nucleophilic attack at the C8 position and AP sites via hydrolysis of the glycosidic bond generating DPCs through secondary reactions **(Scheme 4.2.)**. Therefore, we examined whether monofunctional alkylating agents can induce DPCs and their cytotoxicity is ascribed to DPC formation for the first time. Yang *et al.* showed that MdG directly reacts with lysine residue of histone nucleosome core particles (NCPs) to form DPCs *in vitro*[94]. Therefore, DPC-containing DNA was extracted from MMS treated cells and their untreated controls to quantify MMS induced DPCs *in vivo* via gel electrophoresis, and the dose dependent increase of DPCs upon treatment with MMS has been reported **(Figure 4.1.).** In addition, proteomic studies were utilized to investigate the biological properties and cytotoxicity of MMS induced DPCs **(Table 4.1.)**. The proteomic result showed that the crosslinked proteins participate in translation, transcription, and chromatin organization. Interestingly, the amounts of crosslinked histones are increased significantly in MMS treated cells **(Figure 4.2.)**, indicating MMS induced DPC formation in vivo. Therefore, MMS induced DPCs in V79 cells were TMT-tagged to determine the rate of DPC_{AP} and DPC_{MdG} through mass spectrometry quantification (**Table 4.2.**). Our LC-MS/MS analysis of MMS-treated cells demonstrated that more than 75% of DPCs are produced by MdG site DPCs **(Figure 4.3.)**. To our knowledge, this is the first report which provides evidence of monofunctional alkylating agent induced DPC formation *in vivo*. In short, monofunctional alkylating agents induced DPCs should be considered carefully to understand the DPC toxicity and to elucidate their repair pathways in future studies.

In summary, mass spectrometry based analytical methodologies can be utilized to detect the direct evidence of DPC formation *in vitro* and *in vivo*, to examine novel mechanism of DPC formation, and to investigate novel DPC repair pathways.

VI. Future Direction

6.1.Absolute Quantification and Proteomic Analysis of Radical Induced Chromosomal DPCs and Mitochondrial DPCs

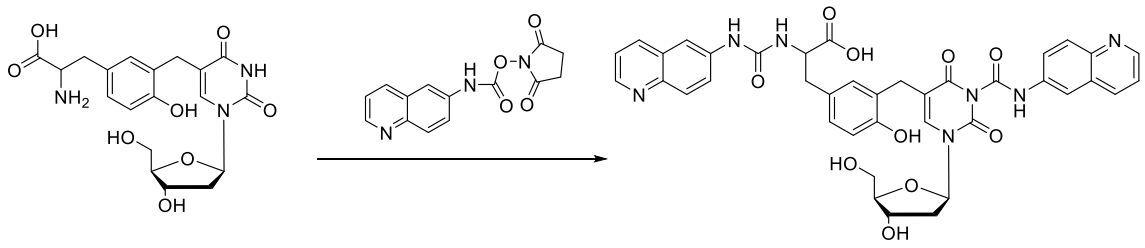
We quantified IR-induced DPC in cells and mouse models using nanoLC-ESI$^+$-MS/MS assay as described in Chapter 2. Because ROS react with guanine, cytosine, or thymine of DNA and lysine or tyrosine residues in proteins to generate crosslinks through free radical or electrophilic lesion[17,219], mitochondrial DNA may have a higher frequency of DPCs than chromosomal DNA[220,221]. However, in our experiments, DPCs were analyzed without consideration of their cellular location. We lysed both the nuclei and the mitochondrial membranes to extract DPCs from the cultured cells or mouse tissue samples. Because DPCs can inhibit replication and transcription[23–26] and mitochondrial apoptosis can be regulated by ROS generation[222], it is important to distinguish chromosomal and mitochondrial DPCs and their toxicities. Therefore, cells can be treated with a mild cell lysis solution that only breaks the cell membrane to separate nuclei and mitochondria via ultracentrifugation[48,50,223,224]. The resulting nuclei and mitochondria can be subjected to separate DPC extraction as described in Chapter 2. The radical induced DPCs in nuclei and mitochondria can be quantified using nanoLC-ESI$^+$-MS/MS assay. In addition, mass spectrometry-based quantification and identification of cross-linked proteins can be performed to characterize the cross-linked proteins in each organelle.

6.2.Derivatization of dT-Tyr to Improve Sensitivity of nanoLC-ESI+-MS/MS assay

As described in chapter 2, we have developed and optimized a nanoLC-ESI$^+$-MS/MS assay to detect dT-Tyr *in vivo* and *in vitro*. With this method, we revealed direct evidence

that ROS-induced DPC can be repaired by SPRTN, while the importance of other proteases and proteasomes which play similar role to SPRTN has been suggested. The limit of quantification of this assay was around 10 fmol. The sensitivity of the nanoLC-ESI^+-MS/MS assay should be improved further to detect ROS-induced DPC more accurately. For example, AccQ-tag from Waters Corp. can be utilized to derivatize the amine group of an amino acid **(Scheme 6.1.)**. Since the derivatized dT-Tyr is less polar than dT-Tyr, SPE columns, such as Sep-Pak C18 or Discovery C18, can be investigated for sample purification rather than the currently used cation ion exchange column. Furthermore, this derivatization will confirm the effectiveness of the enzymatic DPC digestion for the assay as well. Once the derivatization reaction is completed, nanoLC-ESI+-MS/MS can be performed to quantify dT-Tyr derivates and new isotopically labeled internal standard derivates in mammalian tissues.

Scheme 6.1. Schematic of dT-Tyr derivatization with AccQ-Tag AccQ-tag from Waters Corp. derivatize the amine group of dT-Tyr.



dT-Tyr

dT-Tyr Derivate

6.3. Role of Transcription Coupled NER in DNA-peptide Cross-link Repair

As described in Chapter 3, we have investigated the effect of 5fC mediated conjugates on transcription in various cell lines to elucidate the mechanism of DNA-peptide crosslink repair. 5fC conjugates containing plasmids were transfected into XPA deficient cells and XPA corrected cells to determine the importance of NER pathways in DNA-peptide crosslink repair. We showed that XPA-corrected cells have slightly higher bypass efficiency than that of XPA cells, but the difference was not significant **(Figure 3.4.)**. Interestingly, recent studies have demonstrated that cells can remove RNAP-blocking DNA lesions by transcription-coupled repair (TC-NER) to overcome prolonged blockage of transcription[200]. Specifically, CSB play important role during TC-NER by recruiting NER proteins, chromatin remodelers and CSA- E3-ubiquitin ligase complex to the stalled RNAP. Therefore, CTAB assay can be performed with CSB deficient cells and the wild-type to elucidate the role of TC-NER in DNA-peptide crosslink repair. We anticipate observing 5fC site conjugates induced transition mutations or deletions, and lower RNA polymerase bypass efficiency in CSB deficient cells.

6.4. Effects of 5-formylcytosine Mediated DNA-Peptide Cross-link on Replication

As described in Chapter 3, we have investigated the effect of 5fC mediated conjugates on transcription using CTAB assay. CTAB assay successfully demonstrated that DpC conjugates decrease the bypass of RNA polymerase during transcription and their sizes are critical for the bypass tolerance. Although it is known that DPCs block replication fork[105,106], the effect of 5fC mediated conjugates on replication has not been revealed *in vivo* yet. Therefore, MS-based assay for replication studies in cells[225], which is similar to

CTAB, can be performed to examine the impact of 5fC mediated conjugate on replication. Briefly, when 5fC mediated 12mer peptide/lysine crosslink is introduced into a plasmid with replication origin, the resulting lesion bearing and lesion free plasmids can be transfected and replicated in host mammalian cells. Therefore, the progeny plasmids can be isolated from the host cells and the region where the crosslink was initially located can be PCR amplified with a pair of primers. The PCR products can be treated with restriction enzymes to generate short oligo-deoxyribonucleotides fragment for LC-MS/MS analysis as described in Chapter 3. We anticipate that the size of conjugate would be a critical factor for bypass tolerance during replication as transcription.

6.5. Impact of Crosslink Site Neighboring Sequence on DNA-peptide Cross-link Repair

As described in chapter 3, 12mer-5fC-containing oligodeoxynucleotides (5'-ATGGCGGGXTAT-3') were crosslinked to 11mer peptide (RPKPQQFFGLM-$CONH_2$) and this structure was used to determine the RNA polymerase bypass rate using CTAB assay in the thesis. However, it is still unknown whether neighboring sequence could affect the bypass rate during transcription and replication. Therefore, DNA substrates with different sequence can be synthesized to perform primer extension assays and kinetic analysis as described in a previous study[105]. For example, 5'-ATGGCGGGXTAT-3' can be modified to 5'-ATGGCGGNXNAT-3' to reveal the sequence induced difference in TLS polymerase bypass rate and their mutagenic effect during replication *in vitro*. Furthermore, the sequence of crosslinked peptide or the sequence of DNA of lesion bearing plasmid can be modified to any sequence (ex. RPPQQKFFGLM-$CONH_2$ or 5'-

ATGGCGGNXNAT-3') to demonstrate the sequence effects on transcription through CTAB assay *in vivo*.

6.6. DPC Repair Pathways in Mitochondria

Mitochondrial DNA is constantly damaged by endogenous and exogenous agents, inducing damage, hydrolytic damage, DNA adducts, mismatched bases, strand breaks, and oxidative damage[5,226–229]. Therefore, a higher frequency of mtDNA lesion generation may disrupt the function of mitochondria. Although the intrinsic base excision repair(BER) pathways is known to remove oxidative damage in mitochondrial DNA[222,230,231], attempted DNA repair on some lesions often results in further damage[135–137]. For example, Demple *et al.* demonstrated that mitochondrial pol ɣ is crosslinked to mtDNA during repair[132,232]. Recent studies have shown that DPC proteolysis repair involving SPRTN protease activity helps to maintain chromosomal DNA integrity[151]. Moreover, Chesner *et al.* showed that both NER and HR play role in DPC repair using a quantitative PCR-based assay[126]. Despite a large number of studies on DPCs, the influence of mitochondrial DPC and their repair pathways have not been elucidated yet. Specifically, it is still unknown whether DPCs undergo proteolytic processing in the mitochondria as in the nucleus and if the resulting smaller DNA-peptide crosslink can be bypassed through TLS machinery. To elucidate this, DPC containing plasmid can be introduced into purified mitochondria to examine the repair. In addition, any proteases can be knocked-down with siRNA to investigate their role in mitochondrial DPC repair. Although the recruitment of bypass DNA polymerases to overcome replisome stalling on oxidative DNA lesions or DPC sites has been investigated, it is still unknown how the

mitochondrial replisome complex bypasses mtDNA lesions. Interestingly, Wisnovsky et al. employed an siRNA screen, demonstrating the recruitment of polymerases, such as pol θ or REV3, to mitochondria upon oxidative DNA damage[233]. This indicates that several polymerases may play a role in TLS to maintain DNA replication timing in mitochondria. Therefore, the development of a DNA substrate containing a site-specific DPC lesion will help to examine mitochondrial TLS mechanisms *in vitro* and to demonstrate the influence of mtDPCs on replication efficiency and accuracy through primer extension assay. Moreover, a novel affinity capture methodology can be used to identify the necessary proteins for TLS and mtDPC recognition. Because SPRTN is known to degrade nuclear DNA-crosslinked proteins into peptide chains, I hypothesize that a mitochondrial analog could degrade mtDPCs prior to canonical mtDNA repair pathways.

BIBLIOGRAPHY

(1) Sinha, R.P., and Häder, D.P. (2002). UV-induced DNA damage and repair: a review. *Photochem. Photobiol. Sci. 1,* 225–236.

(2) Rastogi, R.P., Kumar, A., Tyagi, M.B., and Sinha, R.P. (2010). Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. *J. Nucleic Acids 2010*. 16, 592980.

(3) Farmer, P.B., Singh, R., Kaur, B., Sram, R.J., Binkova, B., Kalina, I., Popov, T.A., Garte, S., Taioli, E., and Gabelova-Wasilewska, A. (2003). Molecular epidemiology studies of carcinogenic environmental pollutants: effects of polycyclic aromatic hydrocarbons (PAHs) in environmental pollution on exogenous and oxidative DNA damage. *Mutat. Res. 544,* 397–402.

(4) Xue, W., and Warshawsky, D. (2005). Metabolic activation of polycyclic and heterocyclic aromatic hydrocarbons and DNA damage: a review. *Toxicol. Appl. Pharmacol. 206,* 73–93.

(5) De Bont, R., and Van Larebeke, N. (2004). Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis 19,* 169–185.

(6) Gates, K.S. (2009). An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. *Chem. Res. Toxicol. 22,* 1747–1760.

(7) Tudek, B., Zdżalik-Bielecka, D., Tudek, A., Kosicki, K., Fabisiewicz, A., and Speina, E. (2017). Lipid peroxidation in face of DNA damage, DNA repair and other cellular processes. *Free Radic. Biol. Med. 107,* 77–89.

(8) Cooke, M.S., Evans, M.D., Dizdaroglu, M., and Lunec, J. (2003). Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB J. 17,* 1195–1214.

(9) Barker, S., Weinfeld, M., and Murray, D. (2005). DNA–protein crosslinks: their induction, repair, and biological consequences. *Mutat. Res. 589,* 111–135.

(10) Barker, S., Weinfeld, M., Zheng, J., Li, L., and Murray, D. (2005). The identification of mammalian proteins crosslinked to DNA by ionizing radiation. *J. Biol. Chem. 40,* 33826-33838

(11) Fu, D., Calvo, J.A., and Samson, L.D. (2012). Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat. Rev. Cancer 12*, 104.
(12) Drabløs, F., Feyzi, E., Aas, P.A., Vaagbø, C.B., Kavli, B., Bratlie, M.S., Peña-Diaz, J., Otterlei, M., Slupphaug, G., and Krokan, H.E. (2004). Alkylation damage in DNA and RNA—repair mechanisms and medical significance. *DNA Repair 3*, 1389–1407.
(13) Hoeijmakers, J.H. (2001). Genome maintenance mechanisms for preventing cancer. *nature 411*, 366.
(14) Ghosal, G., and Chen, J. (2013). DNA damage tolerance: a double-edged sword guarding the genome. *Transl. Cancer Res. 2*, 107.
(15) Cadet, J., Douki, T., Gasparutto, D., and Ravanat, J.L. (2003). Oxidative damage to DNA: formation, measurement and biochemical features. *Mutat. Res. 531*, 5–23.
(16) Cadet, J., and Wagner, J.R. (2013). DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation. *Cold Spring Harb. Perspect. Biol. 5*, a012559.
(17) Tretyakova, N.Y., Groehler IV, A., and Ji, S. (2015). DNA–protein cross-links: formation, structural identities, and biological outcomes. *Acc. Chem. Res. 48*, 1631–1644.
(18) Cadet, J., Douki, T., and Ravanat, J.L. (2010). Oxidatively generated base damage to cellular DNA. *Free Radic. Biol. Med. 49*, 9–21.
(19) Lindahl, T., and Nyberg, B. (1972). Rate of depurination of native deoxyribonucleic acid. *Biochemistry 11*, 3610–3618.
(20) Zoltewicz, J.A., Clark, D.F., Sharpless, T.W., and Grahe, G. (1970). Kinetics and mechanism of the acid-catalyzed hydrolysis of some purine nucleosides. *J. Am. Chem. Soc. 92*, 1741–1750.
(21) Lindahl, T., and Andersson, A. (1972). Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid. *Biochemistry 11*, 3618–3623.
(22) Crine, P., and Verly, W.G. (1976). A study of DNA spontaneous degradation. *Biochim. Biophys. Acta. 442*, 50–57.
(23) Yeo, J.E., Wickramaratne, S., Khatwani, S., Wang, Y.C., Vervacke, J., Distefano, M.D., and Tretyakova, N.Y. (2014). Synthesis of site-specific DNA–protein conjugates and their effects on DNA replication. *ACS Chem. Biol. 9*, 1860–1868.

(24) Nakano, T., Miyamoto-Matsubara, M., Shoulkamy, M.I., Salem, A.M., Pack, S.P., Ishimi, Y., and Ide, H. (2013). Translocation and stability of replicative DNA helicases upon encountering DNA-protein cross-links. *J. Biol. Chem. 288*, 4649-4658.
(25) Wickramaratne, S., Boldry, E.J., Buehler, C., Wang, Y.C., Distefano, M.D., and Tretyakova, N.Y. (2015). Error-prone translesion synthesis past DNA-peptide cross-links conjugated to the major groove of DNA via C5 of thymidine. *J. Biol. Chem. 290*, 775–787.
(26) Wickramaratne, S., Ji, S., Mukherjee, S., Su, Y., Pence, M.G., Lior-Hoffmann, L., Fu, I., Broyde, S., Guengerich, F.P., Distefano, M., Schärer, O.D., Sham Y.Y., and Tretyakova N.Y. (2016). Bypass of DNA-protein cross-links conjugated to the 7-deazaguanine position of DNA by translesion synthesis polymerases. *J. Biol. Chem. 291*, 23589-23603.
(27) Varshavsky, A.J., Sundin, O., and Bohn, M. (1979). A stretch of "late" SV40 viral DNA about 400 bp long which includes the origin of replication is specifically exposed in SV40 minichromosomes. *Cell 16*, 453–466.
(28) Zahn, R.K., Zahn-Daimler, G., Ax, S., Hosokawa, M., and Takeda, T. (1999). Assessment of DNA-protein crosslinks in the course of aging in two mouse strains by use of a modified alkaline filter elution applied to whole tissue samples. *Mech. Ageing Dev. 108*, 99–112.
(29) Izzotti, A., Cartiglia, C., Taningher, M., De Flora, S., and Balansky, R. (1999). Age-related increases of 8-hydroxy-2′-deoxyguanosine and DNA–protein crosslinks in mouse organs. *Mutat. Res. 446*, 215–223.
(30) Hubal, E.A., Schlosser, P.M., Conolly, R.B., and Kimbell, J.S. (1997). Comparison of inhaled formaldehyde dosimetry predictions with DNA–protein cross-link measurements in the rat nasal passages. *Toxicol. Appl. Pharmacol. 143*, 47–55.
(31) Groehler IV, A., Degner, A., and Tretyakova, N.Y. (2017). Mass Spectrometry-based Tools to Characterize Dna–protein Cross-linking by bis-electrophiles. *Basic Clin. Pharmacol. Toxicol. 121*, 63–77.
(32) Loeber, R.L., Michaelson-Richie, E.D., Codreanu, S.G., Liebler, D.C., Campbell, C. R., and Tretyakova, N.Y. (2009). Proteomic analysis of DNA- protein cross-linking by antitumor nitrogen mustards. *Chem. Res. Toxicol. 22*, 1151–1162.

(33) Qiu, H., and Wang, Y. (2009). Exploring DNA-binding proteins with in vivo chemical cross-linking and mass spectrometry. *J. Proteome Res. 8*, 1983–1991.
(34) Vaz, B., Popovic, M., and Ramadan, K. (2017). DNA–Protein Crosslink Proteolysis Repair. *Trends Biochem. Sci. 42*, 483–495.
(35) Kiianitsa, K., and Maizels, N. (2013). A rapid and sensitive assay for DNA–protein covalent complexes in living cells. *Nucleic Acids Res. 41*, e104–e104.
(36) Chen, S.H., Chan, N.L., and Hsieh, T.S. (2013). New mechanistic and functional insights into DNA topoisomerases. *Annu. Rev. Biochem. 82*, 139–170.
(37) Chválová, K., Brabec, V., and Kašpárková, J. (2007). Mechanism of the formation of DNA–protein cross-links by antitumor cisplatin. *Nucleic Acids Res. 35*, 1812–1821.
(38) Ma, T.H., and Harris, M.M. (1988). Review of the genotoxicity of formaldehyde. *Mutat. Res. 196*, 37–59.
(39) Quievryn, G., and Zhitkovich, A. (2000). Loss of DNA–protein crosslinks from formaldehyde-exposed cells occurs through spontaneous hydrolysis and an active repair process linked to proteosome function. *Carcinogenesis 21*, 1573–1580.
(40) O'Sullivan, J., Unzeta, M., Healy, J., O'Sullivan, M.I., Davey, G., and Tipton, K.F. (2004). Semicarbazide-sensitive amine oxidases: enzymes with quite a lot to do. *Neurotoxicology 25*, 303–315.
(41) Cloos, P.A., Christensen, J., Agger, K., and Helin, K. (2008). Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. *Genes Dev. 22*, 1115–1140.
(42) Hou, H., and Yu, H. (2010). Structural insights into histone lysine demethylation. *Curr. Opin. Struct. Biol. 20*, 739–748.
(43) Lu, K., Ye, W., Zhou, L., Collins, L.B., Chen, X., Gold, A., Ball, L.M., and Swenberg, J.A. (2010). Structural characterization of formaldehyde-induced cross-links between amino acids and deoxynucleosides and their oligomers. *J. Am. Chem. Soc. 132*, 3388–3399.
(44) Heck, H.D., Casanova, M., and Starr, T.B. (1990). Formaldehyde toxicity—new understanding. *Crit. Rev. Toxicol. 20*, 397–426.
(45) Merk, O., and Speit, G. (1998). Significance of formaldehyde-induced DNA–protein crosslinks for mutagenesis. *Environ. Mol. Mutagen. 32*, 260–268.

(46) O'Connor, P.M., and Fox, B.W. (1989). Isolation and characterization of proteins cross-linked to DNA by the antitumor agent methylene dimethanesulfonate and its hydrolytic product formaldehyde. *J. Biol. Chem. 264*, 6391–6397.
(47) Emadi, A., Jones, R.J., and Brodsky, R.A. (2009). Cyclophosphamide and cancer: golden anniversary. *Nat. Rev. Clin. Oncol. 6*, 638-647.
(48) Groehler IV, A., Villalta, P.W., Campbell, C., and Tretyakova, N.Y. (2016). Covalent DNA–Protein Cross-Linking by Phosphoramide Mustard and Nornitrogen Mustard in Human Cells. *Chem. Res. Toxicol. 29*, 190–202.
(49) Hansson, J., Lewensohn, R., Ringborg, U., and Nilsson, B. (1987). Formation and removal of DNA cross-links induced by melphalan and nitrogen mustard in relation to drug-induced cytotoxicity in human melanoma cells. *Cancer Res. 47*, 2631–2637.
(50) Gherezghiher, T.B., Ming, X., Villalta, P.W., Campbell, C., and Tretyakova, N.Y. (2013). 1, 2, 3, 4-Diepoxybutane-induced DNA–protein cross-linking in human fibrosarcoma (HT1080) cells. *J. Proteome Res. 12*, 2151–2164.
(51) Loeber, R., Rajesh, M., Fang, Q., Pegg, A.E., and Tretyakova, N.Y. (2006). Cross-linking of the human DNA repair protein O 6-alkylguanine DNA alkyltransferase to DNA in the presence of 1, 2, 3, 4-diepoxybutane. *Chem. Res. Toxicol. 19*, 645–654.
(52) Loeber, R., Michaelson, E., Fang, Q., Campbell, C., Pegg, A.E., and Tretyakova, N.Y. (2008). Cross-linking of the DNA repair protein O 6-alkylguanine DNA alkyltransferase to DNA in the presence of antitumor nitrogen mustards. *Chem. Res. Toxicol. 21*, 787–795.
(53) Phaniendra, A., Jestadi, D.B., and Periyasamy, L. (2015). Free radicals: properties, sources, targets, and their implication in various diseases. *Indian J. Clin. Biochem. 30*, 11–26.
(54) Greenberg, M.M. (1998). Investigating nucleic acid damage processes via independent generation of reactive intermediates. *Chem. Res. Toxicol. 11*, 1235–1248.
(55) Pogozelski, W.K., and Tullius, T.D. (1998). Oxidative strand scission of nucleic acids: routes initiated by hydrogen abstraction from the sugar moiety. *Chem. Rev. 98*, 1089–1108.
(56) Breen, A.P., and Murphy, J.A. (1995). Reactions of oxyl radicals with DNA. *Free Radic. Biol. Med. 18*, 1033–1077.

(57) Halliwell, B., and Aruoma, O.I. (1991). DNA damage by oxygen-derived species Its mechanism and measurement in mammalian systems. *FEBS Lett. 281*, 9–19.
(58) Von Sonntag, C. (2006). Free-radical-induced DNA damage and its repair. Springer.
(59) Wagner, J.R., Van Lier, J.E., Berger, M., and Cadet, J. (1994). Thymidine hydroperoxides: structural assignment, conformational features, and thermal decomposition in water. *J. Am. Chem. Soc. 116*, 2235–2242.
(60) Pouget, J.P., Frelon, S., Ravanat, J.L., Testard, I., Odin, F., and Cadet, J. (2002). Formation of modified DNA bases in cells exposed either to gamma radiation or to high-LET particles. *Radiat. Res. 157*, 589–595.
(61) Douki, T., Ravanat, J.L., Pouget, J.P., Testard, I., and Cadet, J. (2006). Minor contribution of direct ionization to DNA base damage inducedby heavy ions. *Int. J. Radiat. Biol. 82*, 119–127.
(62) Nohl, H., Kozlov, A.V., Gille, L., and Staniek, K. (2003). Cell respiration and formation of reactive oxygen species: facts and artefacts. *Biochem. Soc. Trans. 6*, 1308-1311.
(63) Circu, M.L., and Aw, T.Y. (2010). Reactive oxygen species, cellular redox systems, and apoptosis. *Free Radic. Biol. Med. 48*, 749–762.
(64) Valko, M., Rhodes, C.J., Moncol, J., Izakovic, M., and Mazur, M. (2006). Free radicals, metals and antioxidants in oxidative stress-induced cancer. *Chem. Biol. Interact. 160*, 1–40.
(65) De Duve, C., and Baudhuin, P. (1966). Peroxisomes (microbodies and related particles). *Physiol. Rev. 46*, 323–357.
(66) Schrader, M., and Fahimi, H.D. (2006). Peroxisomes and oxidative stress. *Biochim. Biophys. Acta. 1763*, 1755–1766.
(67) Forman, H.J., and Torres, M. (2002). Reactive oxygen species and cell signaling: respiratory burst in macrophage signaling. *Am. J. Respir. Crit. Care Med. 166*, S4–S8.
(68) Torres, M.A., Jones, J.D., and Dangl, J.L. (2006). Reactive oxygen species signaling in response to pathogens. *Plant Physiol. 141*, 373–378.
(69) Conner, E.M., and Grisham, M.B. (1996). Inflammation, free radicals, and antioxidants. *Nutrition 12*, 274–277.

(70) Singh, A., and Singh, H. (1982). Time-scale and nature of radiation-biological damage: approaches to radiation protection and post-irradiation therapy. *Prog. Biophys. Mol. Biol. 39*, 69–107.
(71) Heck, D.E., Vetrano, A.M., Mariano, T.M., and Laskin, J.D. (2003). UVB light stimulates production of reactive oxygen species: unexpected role for catalase. *J. Biol. Chem 25*, 22432-22436.
(72) Valavanidis, A., Vlachogianni, T., and Fiotakis, K. (2009). Tobacco smoke: involvement of reactive oxygen species and stable free radicals in mechanisms of oxidative damage, carcinogenesis and synergistic effects with other respirable particles. *Int. J. Environ. Res. Public. Health 6*, 445–462.
(73) Church, D.F., and Pryor, W.A. (1985). Free-radical chemistry of cigarette smoke and its toxicological implications. *Environ. Health Perspect. 64*, 111-126.
(74) Montoliu, C., Vallés, S., Renau-Piqueras, J., and Guerri, C. (1994). Ethanol-induced oxygen radical formation and lipid peroxidation in rat brain: effect of chronic alcohol consumption. *J. Neurochem. 63*, 1855–1862.
(75) Bailey, S.M., Pietsch, E.C., and Cunningham, C.C. (1999). Ethanol stimulates the production of reactive oxygen species at mitochondrial complexes I and III. *Free Radic. Biol. Med. 27*, 891–900.
(76) Mozaffarian, D., Benjamin, E.J., Go, A.S., Arnett, D.K., Blaha, M.J., Cushman, M., Das, S.R., de Ferranti, S., Després, J.P., and Fullerton, H.J. (2015). Heart disease and stroke statistics—2016 update: a report from the American Heart Association. *Circulation 4*, e29-322.
(77) O'Connor, T.R., Boiteux, S., and Laval, J. (1988). Ring-opened 7-methylguanine residues in DNA are a block to in vitro DNA synthesis. *Nucleic Acid Res*. *13*, 5879-5894.
(78) Cadet, J., Douki, T., and Ravanat, J.L. (2008). Oxidatively generated damage to the guanine moiety of DNA: mechanistic aspects and formation in cells. *Acc. Chem. Res. 41*, 1075–1083.
(79) Gilboa, R., Zharkov, D.O., Golan, G., Fernandes, A.S., Gerchman, S.E., Matz, E., Kycia, J.H., Grollman, A.P., and Shoham, G. (2002). Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA. *J. Biol. Chem. 277*, 19811–19816.

(80) Banerjee, A., Yang, W., Karplus, M., and Verdine, G.L. (2005). Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. *nature 434*, 612.
(81) Dizdaroglu, M., and Gajewski, E. (1989). Structure and mechanism of hydroxyl radical-induced formation of a DNA-protein cross-link involving thymine and lysine in nucleohistone. *Cancer Res. 49*, 3463–3467.
(82) Gajewski, E., Fuciarelli, A.F., and Dizdaroglu, M. (1988). Structure of hydroxyl radical-induced DNA-protein crosslinks in calf thymus nucleohistone in vitro. *Int. J. Radiat. Biol. 54*, 445–459.
(83) Dizdaroglu, M., Gajewski, E., Reddy, P., and Margolis, S.A. (1989). Structure of a hydroxyl radical-induced DNA-protein crosslink involving thymine and tyrosine in nucleohistone. *Biochemistry 28*, 3625–3628.
(84) Dizdaroglu, M., and Jaruga, P. (2012). Mechanisms of free radical-induced damage to DNA. *Free Radic. Res. 46*, 382–419.
(85) Cress, A.E., and Bowden, G.T. (1983). Covalent DNA-protein crosslinking occurs after hyperthermia and radiation. *Radiat. Res. 95*, 610–619.
(86) Beranek, D.T. (1990). Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutat. Res. 231*, 11–30.
(87) Wyatt, M.D., and Pittman, D.L. (2006). Methylating agents and DNA repair responses: Methylated bases and sources of strand breaks. *Chem. Res. Toxicol. 19*, 1580–1594.
(88) Lindahl, T., and Wood, R.D. (1999). Quality control by DNA repair. *Science 286*, 1897–1905.
(89) O'Brien, P.J., and Ellenberger, T. (2004). Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase. *J. Biol. Chem. 279*, 9750–9757.
(90) Bennett, R.A., Wilson, D.M., Wong, D., and Demple, B. (1997). Interaction of human apurinic endonuclease and DNA polymerase β in the base excision repair pathway. *Proc. Natl. Acad. Sci. 94*, 7166–7169.
(91) Erzberger, J.P., Barsky, D., Schärer, O.D., Colvin, M.E., and Wilson III, D.M. (1998). Elements in abasic site recognition by the major human and Escherichia coli apurinic/apyrimidinic endonucleases. *Nucleic Acids Res. 26*, 2771–2778.

(92) Krogh, B.O., and Symington, L.S. (2004). Recombination proteins in yeast. *Annu Rev Genet 38*, 233–271.
(93) Lundin, C., North, M., Erixon, K., Walters, K., Jenssen, D., Goldman, A.S., and Helleday, T. (2005). Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable in vivo DNA double-strand breaks. *Nucleic Acids Res. 33*, 3799–3811.
(94) Yang, K., Park, D., Tretyakova, N.Y., and Greenberg, M.M. (2018). Histone tails decrease N7-methyl-2′-deoxyguanosine depurination and yield DNA–protein cross-links in nucleosome core particles and cells. *Proc. Natl. Acad. Sci. 115*, E11212–E11220.
(95) Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. *Annu Rev Biochem 74*, 481–514.
(96) Turek-Plewa, J., and Jagodzinski, P.P. (2005). The role of mammalian DNA methyltransferases in the regulation of gene expression. *Cell. Mol. Biol. Lett. 10*, 631-647.
(97) Jones, P.A., and Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. *Science 293*, 1068–1070.
(98) Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science 333*, 1300–1303.
(99) Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet. 9*, 465.
(100) Hu, L., Lu, J., Cheng, J., Rao, Q., Li, Z., Hou, H., Lou, Z., Zhang, L., Li, W., and Gong, W. (2015). Structural insight into substrate preference for TET-mediated oxidation. *Nature 527*, 118.
(101) Jin, S.G., Jiang, Y., Qiu, R., Rauch, T.A., Wang, Y., Schackert, G., Krex, D., Lu, Q., and Pfeifer, G.P. (2011). 5-Hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. *Cancer Res.* canres–2023.
(102) Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science 324*, 929–930.

(103) Spruijt, C.G., Gnerlich, F., Smits, A.H., Pfaffeneder, T., Jansen, P.W., Bauer, C., Münzel, M., Wagner, M., Müller, M., and Khan, F. (2013). Dynamic readers for 5-(hydroxy) methylcytosine and its oxidized derivatives. *Cell 152*, 1146–1159.
(104) Iurlaro, M., Ficz, G., Oxley, D., Raiber, E.A., Bachman, M., Booth, M.J., Andrews, S., Balasubramanian, S., and Reik, W. (2013). A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. *Genome Biol. 14*, R119.
(105) Ji, S., Fu, I., Naldiga, S., Shao, H., Basu, A.K., Broyde, S., and Tretyakova, N.Y. (2018). 5-Formylcytosine mediated DNA–protein cross-links block DNA replication and induce mutations in human cells. *Nucleic Acids Res. 46,* 6455-6469.
(106) Ji, S., Shao, H., Han, Q., Seiler, C.L., and Tretyakova, N.Y. (2017). Reversible DNA–Protein Cross-Linking at Epigenetic DNA Marks. *Angew. Chem. Int. Ed. 56*, 14130–14134.
(107) Wu, F.Y., Lee, Y.J., Chen, D.R., and Kuo, H.W. (2002). Association of DNA-protein crosslinks and breast cancer. *Mutat. Res. 501*, 69–78.
(108) Maskey, R.S., Flatten, K.S., Sieben, C.J., Peterson, K.L., Baker, D.J., Nam, H.J., Kim, M.S., Smyrk, T.C., Kojima, Y., Machida, Y., Santiago, A., van Deursen, J.M., Kaufmann, S.H., and Machida, Y.J. (2017). Spartan deficiency causes accumulation of Topoisomerase 1 cleavage complexes and tumorigenesis. *Nucleic Acids Res. 45*, 4564–4576.
(109) Lessel, D., Vaz, B., Halder, S., Lockhart, P.J., Marinovic-Terzic, I., Lopez-Mosqueda, J., Philipp, M., Sim, J.C., Smith, K.R., Oehler, J., Cabrera, E., Freire, R., Pope, K., Nahid, A., Norris, F., Leventer, R.J., Delatycki, M.B., Barbi, G., von Ameln, S., Högel, J., Degoricija, M., Fertig, R., Burkhalter, M.D., Hofmann, K., Thiele, H., Altmüller, J., Nürnberg, G., Nürnberg, P., Bahlo, M., Martin, GM., Aalfs, CM., Oshima, J., Terzic, J., Amor, D.J., Dikic, I., Ramadan, K., Kubisch, C. (2014). Mutations in SPRTN cause early onset hepatocellular carcinoma, genomic instability and progeroid features. *Nat. Genet. 46*, 1239–1244.
(110) Ross, C.A., and Poirier, M.A. (2004). Protein aggregation and neurodegenerative disease. *Nat. Med. 10*, S10-17

(111) Groehler IV, A., Kren, S., Li, Q., Robledo-Villafane, M., Schmidt, J., Garry, M., and Tretyakova, N.Y. (2018). Oxidative cross-linking of proteins to DNA following ischemia-reperfusion injury. *Free Radic. Biol. Med. 120*, 89–101.
(112) Pommier, Y., Shar-yin, N.H., Gao, R., Das, B.B., Murai, J., and Marchand, C. (2014). Tyrosyl-DNA-phosphodiesterases (tdp1 and tdp2). *DNA Repair 19*, 114–129.
(113) Ide, H., Shoulkamy, M.I., Nakano, T., Miyamoto-Matsubara, M., and Salem, A.M. (2011). Repair and biochemical effects of DNA–protein crosslinks. *Mutat. Res. Mol. Mech. Mutagen. 711*, 113–122.
(114) de Graaf, B., Clore, A., and McCullough, A.K. (2009). Cellular pathways for DNA repair and damage tolerance of formaldehyde-induced DNA-protein crosslinks. *DNA Repair 8*, 1207–1214.
(115) Stingele, J., Schwarz, M.S., Bloemeke, N., Wolf, P.G., and Jentsch, S. (2014). A DNA-dependent protease involved in DNA-protein crosslink repair. *Cell 158*, 327–338.
(116) Minko, I.G., Zou, Y., and Lloyd, R.S. (2002). Incision of DNA–protein crosslinks by UvrABC nuclease suggests a potential repair pathway involving nucleotide excision repair. *Proc. Natl. Acad. Sci. 99*, 1905–1909.
(117) Nakano, T., Morishita, S., Katafuchi, A., Matsubara, M., Horikawa, Y., Terato, H., Salem, A.M., Izumi, S., Pack, S.P., and Makino, K. (2007). Nucleotide excision repair and homologous recombination systems commit differentially to the repair of DNA-protein crosslinks. *Mol. Cell 28*, 147–158.
(118) Reardon, J.T., and Sancar, A. (2006). Repair of DNA–polypeptide crosslinks by human excision nuclease. *Proc. Natl. Acad. Sci. 103*, 4056–4061.
(119) Baker, D.J., Wuenschell, G., Xia, L., Termini, J., Bates, S.E., Riggs, A.D., and O'Connor, T. R. (2007). Nucleotide excision repair eliminates unique DNA-protein cross-links from mammalian cells. *J. Biol. Chem. 282*, 22592–22604.
(120) Sale, J. E. (2013). Translesion DNA synthesis and mutagenesis in eukaryotes. *Cold Spring Harb. Perspect. Biol. 5*, a012708.
(121) Prakash, S., Johnson, R.E., and Prakash, L. (2005). Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annu Rev Biochem 74*, 317–353.
(122) Reardon, J.T., Cheng, Y., and Sancar, A. (2006). Repair of DNA–protein cross-links in mammalian cells. *Cell Cycle 5*, 1366–1370.

(123) Deshpande, R. A., Lee, J.H., Arora, S., and Paull, T.T. (2016). Nbs1 converts the human Mre11/Rad50 nuclease complex into an endo/exonuclease machine specific for protein-DNA adducts. *Mol. Cell 64*, 593–606.
(124) Cannavo, E., and Cejka, P. (2014). Sae2 promotes dsDNA endonuclease activity within Mre11–Rad50–Xrs2 to resect DNA breaks. *Nature 514*, 122-125.
(125) Hoa, N.N., Shimizu, T., Zhou, Z.W., Wang, Z.Q., Deshpande, R.A., Paull, T.T., Akter, S., Tsuda, M., Furuta, R., and Tsutsui, K. (2016). Mre11 is essential for the removal of lethal topoisomerase 2 covalent cleavage complexes. *Mol. Cell 64*, 580–592.
(126) Chesner, L. (2018). Repair of DNA-protein crosslinks in mammalian cells. *J Vis Exp. 133.*
(127) Sidorenko, V.S., and Zharkov, D.O. (2008). Role of base excision repair DNA glycosylases in hereditary and infectious human diseases. *Mol. Biol. 42*, 794–805.
(128) Dianov, G.L., and Hübscher, U. (2013). Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Res. 41*, 3483–3490.
(129) Demple, B., Herman, T., and Chen, D.S. (1991). Cloning and expression of APE, the cDNA encoding the major human apurinic endonuclease: definition of a family of DNA repair enzymes. *Proc. Natl. Acad. Sci. 88*, 11450–11454.
(130) Demple, B., and Sung, J.S. (2005). Molecular and biological roles of Ape1 protein in mammalian base excision repair. *DNA Repair 4*, 1442–1449.
(131) Caston, R.A., and Demple, B. (2017). Risky repair: DNA-protein crosslinks formed by mitochondrial base excision DNA repair enzymes acting on free radical lesions. *Free Radic. Biol. Med. 107*, 146–150.
(132) DeMott, M.S., Beyret, E., Wong, D., Bales, B.C., Hwang, J.T., Greenberg, M.M., and Demple, B. (2002). Covalent trapping of human DNA polymerase β by the oxidative DNA lesion 2-deoxyribonolactone. *J. Biol. Chem. 277*, 7637–7640.
(133) Guan, L., and Greenberg, M.M. (2010). Irreversible inhibition of DNA polymerase β by an oxidized abasic lesion. *J. Am. Chem. Soc. 132*, 5004–5005.
(134) Kroeger, K.M., Hashimoto, M., Kow, Y.W., and Greenberg, M.M. (2003). Cross-linking of 2-deoxyribonolactone and its β-elimination product by base excision repair enzymes. *Biochemistry 42*, 2449–2455.

(135) Quiñones, J.L., Thapar, U., Yu, K., Fang, Q., Sobol, R.W., and Demple, B. (2015). Enzyme mechanism-based, oxidative DNA–protein cross-links formed with DNA polymerase β in vivo. *Proc. Natl. Acad. Sci. 112*, 8602–8607.
(136) Quinones, J.L., and Demple, B. (2016). When DNA repair goes wrong: BER-generated DNA-protein crosslinks to oxidative lesions. *DNA Repair 44*, 103–109.
(137) Nakano, T., Terato, H., Asagoshi, K., Masaoka, A., Mukuta, M., Ohyama, Y., Suzuki, T., Makino, K., and Ide, H. (2003). DNA-Protein Cross-link Formation Mediated by Oxanine A NOVEL GENOTOXIC MECHANISM OF NITRIC OXIDE-INDUCED DNA DAMAGE. *J. Biol. Chem. 278*, 25264–25272.
(138) Stingele, J., Schwarz, M.S., Bloemeke, N., Wolf, P.G., and Jentsch, S. (2014). A DNA-dependent protease involved in DNA-protein crosslink repair. *Cell 158*, 327–338.
(139) Stingele, J., Habermann, B., and Jentsch, S. (2015). DNA–protein crosslink repair: proteases as DNA repair enzymes. *Trends Biochem. Sci. 40*, 67–71.
(140) Maskey, R.S., Kim, M.S., Baker, D.J., Childs, B., Malureanu, L.A., Jeganathan, K. B., Machida, Y., Van Deursen, J.M., and Machida, Y.J. (2014). Spartan deficiency causes genomic instability and progeroid phenotypes. *Nat. Commun. 5*, 5744.
(141) Ramadan, K., Halder, S., Wiseman, K., and Vaz, B. (2017). Strategic role of the ubiquitin-dependent segregase p97 (VCP or Cdc48) in DNA replication. *Chromosoma 126*, 17–32.
(142) Ruijs, M.W., van Andel, R.N., Oshima, J., Madan, K., Nieuwint, A.W., and Aalfs, C.M. (2003). Atypical progeroid syndrome: an unknown helicase gene defect? *Am. J. Med. Genet. A. 116*, 295–299.
(143) Mórocz, M., Zsigmond, E., Tóth, R., Enyedi, M.Z., Pintér, L., and Haracska, L. (2017). DNA-dependent protease activity of human Spartan facilitates replication of DNA–protein crosslink-containing DNA. *Nucleic Acids Res. 45*, 3172–3188.
(144) Mosbech, A., Gibbs-Seymour, I., Kagias, K., Thorslund, T., Beli, P., Povlsen, L., Nielsen, S.V., Smedegaard, S., Sedgwick, G., Lukas, C., Hartmann-Petersen, R., Lukas, J., Choudhary, C., Pocock, R., Bekker-Jensen, S., and Mailand, N. (2012). DVC1 (C1orf124) is a DNA damage–targeting p97 adaptor that promotes ubiquitin-dependent responses to replication blocks. *Nat. Struct. Mol. Biol. 19*, 1084–1092.

(145) Juhasz, S., Balogh, D., Hajdu, I., Burkovics, P., Villamil, M.A., Zhuang, Z., and Haracska, L. (2012). Characterization of human Spartan/C1orf124, an ubiquitin-PCNA interacting regulator of DNA damage tolerance. *Nucleic Acids Res. 40*, 10795–10808.
(146) Kim, M.S., Machida, Y., Vashisht, A.A., Wohlschlegel, J.A., Pang, Y.P., and Machida, Y.J. (2012). Regulation of error-prone translesion synthesis by Spartan/C1orf124. *Nucleic Acids Res. 41*, 1661–1668.
(147) Murakumo, Y., Ogura, Y., Ishii, H., Numata, S., Ichihara, M., Croce, C.M., Fishel, R., and Takahashi, M. (2001). Interactions in the error-prone postreplication repair proteins hREV1, hREV3, and hREV7. *J. Biol. Chem. 276*, 35644–35651.
(148) Toth, A., Hegedus, L., Juhasz, S., Haracska, L., and Burkovics, P. (2017). The DNA-binding box of human SPARTAN contributes to the targeting of Polη to DNA damage sites. *DNA Repair 49*, 33–42.
(149) Centore, R.C., Yazinski, S.A., Tse, A., and Zou, L. (2012). Spartan/C1orf124, a reader of PCNA ubiquitylation and a regulator of UV-induced DNA damage response. *Mol. Cell 46*, 625–635.
(150) Ghosal, G., Leung, J. W.C., Nair, B.C., Fong, K.W., and Chen, J. (2012). PCNA-binding protein C1orf124 is a regulator of translesion synthesis. *J. Biol. Chem. 287*, 34225-24233.
(151) Vaz, B., Popovic, M., Newman, J.A., Fielden, J., Aitkenhead, H., Halder, S., Singh, A.N., Vendrell, I., Fischer, R., Torrecilla, I., Drobnitzky, N., Freire, R., Amor, D.J., Lockhart, P.J., Kessler, B.M., McKenna, G.W., Gileadi, O., and Ramadan, K. (2016). Metalloprotease SPRTN/DVC1 orchestrates replication-coupled DNA-Protein crosslink repair. *Mol. Cell 64*, 704–719.
(152) Stingele, J., Bellelli, R., Alte, F., Hewitt, G., Sarek, G., Maslen, S.L., Tsutakawa, S. E., Borg, A., Kjaer, S., Tainer, J.A., Skehel, J.M., Groll, M., and Boulton, S.J. (2016). Mechanism and regulation of DNA-protein crosslink repair by the DNA-dependent metalloprotease SPRTN. *Mol. Cell 64*, 688–703.
(153) Mao, Y., Desai, S.D., Ting, C.Y., Hwang, J., and Liu, L.F. (2001). 26 S proteasome-mediated degradation of topoisomerase II cleavable complexes. *J. Biol. Chem. 276*, 40652–40658.

(154) Lin, C.P., Ban, Y., Lyu, Y.L., Desai, S.D., and Liu, L.F. (2008). A ubiquitin-proteasome pathway for the repair of Top1-DNA covalent complexes. *J. Biol. Chem 283*, 21074-21083.
(155) Nakano, T., Katafuchi, A., Matsubara, M., Terato, H., Tsuboi, T., Masuda, T., Tatsumoto, T., Pack, S.P., Makino, K., and Croteau, D.L. (2009). Homologous recombination but not nucleotide excision repair plays a pivotal role in tolerance to DNA-protein crosslinks in mammalian cells. *J. Biol. Chem. 284*, 27065-27076.
(156) Desai, S.D., Liu, L.F., Vazquez-Abad, D., and D'Arpa, P. (1997). Ubiquitin-dependent destruction of topoisomerase I is stimulated by the antitumor drug camptothecin. *J. Biol. Chem. 272*, 24159–24164.
(157) Larsen, N.B., Gao, A.O., Sparks, J.L., Gallina, I., Wu, R.A., Mann, M., Räschle, M., Walter, J.C., and Duxin, J.P. (2019). Replication-coupled DNA-protein crosslink repair by SPRTN and the proteasome in Xenopus egg extracts. *Mol. Cell 73*, 574–588.
(158) Nakano, T., Ouchi, R., Kawazoe, J., Pack, S.P., Makino, K., and Ide, H. (2012). T7 RNA polymerases backed up by covalently trapped proteins catalyze highly error prone transcription. *J. Biol. Chem. 287*, 6562–6572.
(159) Debéthune, L., Kohlhagen, G., Grandas, A., and Pommier, Y. (2002). Processing of nucleopeptides mimicking the topoisomerase I–DNA covalent complex by tyrosyl-DNA phosphodiesterase. *Nucleic Acids Res. 30*, 1198–1204.
(160) Zeng, Z., Cortés-Ledesma, F., El Khamisy, S.F., and Caldecott, K.W. (2011). TDP2/TTRAP is the major 5′-tyrosyl DNA phosphodiesterase activity in vertebrate cells and is critical for cellular resistance to topoisomerase II-induced DNA damage. *J. Biol. Chem. 286*, 403–409.
(161) Aparicio, T., Baer, R., Gottesman, M., and Gautier, J. (2016). MRN, CtIP, and BRCA1 mediate repair of topoisomerase II–DNA adducts. *J Cell Biol 212*, 399–408.
(162) Davis, A.J., and Chen, D.J. (2013). DNA double strand break repair via non-homologous end-joining. *Transl. Cancer Res. 2*, 130-143.
(163) Li, X., and Heyer, W.D. (2008). Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res. 18*, 99-113.
(164) Wright, W.D., Shah, S.S., and Heyer, W.D. (2018). Homologous recombination and the repair of DNA Double-Strand Breaks. *J. Biol. Chem. 293*, 10524-10535.

(165) Spivak, G. (2015). Nucleotide excision repair in humans. *DNA Repair 36*, 13–18.
(166) Vaisman, A., and Woodgate, R. (2017). Translesion DNA polymerases in eukaryotes: what makes them tick? *Crit. Rev. Biochem. Mol. Biol. 52*, 274–303.
(167) Duxin, J.P., Dewar, J.M., Yardimci, H., and Walter, J.C. (2014). Repair of a DNA-protein crosslink by replication-coupled proteolysis. *Cell 159*, 346–357.
(168) Weir-Lipton, M.S., Fuciarelli, A.F., Springer, D.L., and Edmonds, C.G. (1996). Characterization of radiation-induced thymine-tyrosine crosslinks by electrospray ionization mass spectrometry. *Radiat. Res. 145*, 681–686.
(169) Ward, J. F. (1988). DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability. *Prog. Nucleic Acid Research and Molecular Biology*, pp 95–125.
(170) Karu, T. (1999). Primary and secondary mechanisms of action of visible to near-IR radiation on cells. *J. Photochem. Photobiol. B 49*, 1–17.
(171) Desai, S.D., Li, T.K., Rodriguez-Bauman, A., Rubin, E.H., and Liu, L.F. (2001). Ubiquitin/26S proteasome-mediated degradation of topoisomerase I as a resistance mechanism to camptothecin in tumor cells. *Cancer Res. 61*, 5926–5932.
(172) Zhang, A., Lyu, Y.L., Lin, C.P., Zhou, N., Azarova, A.M., Wood, L.M., and Liu, L.F. (2006). A protease pathway for the repair of topoisomerase II-DNA covalent complexes. *J. Biol. Chem. 281*, 35997–36003.
(173) Interthal, H., and Champoux, J.J. (2011). Effects of DNA and protein size on substrate cleavage by human tyrosyl-DNA phosphodiesterase 1. *Biochem. J. 436*, 559–566.
(174) Interthal, H., Chen, H.J., and Champoux, J.J. (2005). Human Tdp1 cleaves a broad spectrum of substrates including phosphoamide linkages. *J. Biol. Chem. 280*, 36518-36528.
(175) Yang, S., Burgin, A.B., Huizenga, B.N., Robertson, C.A., Yao, K.C., and Nash, H. A. (1996). A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases. *Proc. Natl. Acad. Sci. 93*, 11534–11539.
(176) Groehler IV, A. (2018). Mass Spectrometry-Based Characterization, Quantitation, And Repair Investigations Of Complex DNA Lesions. Retrieved from the University of Minnesota Digital Conservancy,.

(177) O'brien, J., Wilson, I., Orton, T., and Pognan, F. (2000). Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity. *Eur. J. Biochem. 267*, 5421–5426.
(178) Ferrington, D.A., Hussong, S.A., Roehrich, H., Kapphahn, R.J., Kavanaugh, S.M., Heuss, N.D., and Gregerson, D.S. (2008). Immunoproteasome responds to injury in the retina and brain. *J. Neurochem. 106*, 158–169.
(179) Charlton, T.S., Ingelse, B.A., Black, D.S., Craig, D.C., Mason, K.E., and Duncan, M.W. (1999). A covalent thymine-tyrosine adduct involved in DNA-protein crosslinks: synthesis, characterization, and quantification. *Free Radic. Biol. Med. 27*, 254–261.
(180) Cadet, J., Delatour, T., Douki, T., Gasparutto, D., Pouget, J.P., Ravanat, J.L., and Sauvaigo, S. (1999). Hydroxyl radicals and DNA base damage. *Mutat. Res. 424*, 9–21.
(181) Drews, O., Wildgruber, R., Zong, C., Sukop, U., Nissum, M., Weber, G., Gomes, A.V., and Ping, P. (2007). Mammalian proteasome subpopulations with distinct molecular compositions and proteolytic activities. *Mol. Cell. Proteomics 6*, 2021–2031.
(182) Dahlmann, B., Ruppert, T., Kuehn, L., Merforth, S., and Kloetzel, P.M. (2000). Different proteasome subtypes in a single tissue exhibit different enzymatic properties1. *J. Mol. Biol. 303*, 643–653.
(183) Ding, Q., Martin, S., Dimayuga, E., Bruce-Keller, A.J., and Keller, J.N. (2006). LMP2 knock-out mice have reduced proteasome activities and increased levels of oxidatively damaged proteins. *Antioxid. Redox Signal. 8*, 130–135.
(184) Ji, S. (2019). Chemistry and Biology of DNA-Protein Cross-links. Retrieved from the University of Minnesota Digital Conservancy.
(185) Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nat. Rev. Genet. 14*, 204-220.
(186) Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell 99*, 247–257.
(187) Barau, J., Teissandier, A., Zamudio, N., Roy, S., Nalesso, V., Hérault, Y., Guillou, F., and Bourc'his, D. (2016). The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science 354*, 909–912.

(188) Niculescu, M.D., and Lupu, D.S. (2011). Nutritional influence on epigenetics and effects on longevity. *Curr. Opin. Clin. Nutr. Metab. Care 14*, 35–40.
(189) Jones, M.J., Goodman, S.J., and Kobor, M.S. (2015). DNA methylation and healthy human aging. *Aging Cell 14*, 924–932.
(190) Pfaffeneder, T., Hackner, B., Truss, M., Münzel, M., Müller, M., Deiml, C. A., Hagemeier, C., and Carell, T. (2011). The discovery of 5-formylcytosine in embryonic stem cell DNA. *Angew. Chem. Int. Ed. 50*, 7008–7012.
(191) Globisch, D., Münzel, M., Müller, M., Michalakis, S., Wagner, M., Koch, S., Brückl, T., Biel, M., and Carell, T. (2010). Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PloS One 5*, e15367.
(192) Raiber, E.A., Portella, G., Cuesta, S.M., Hardisty, R., Murat, P., Li, Z., Iurlaro, M., Dean, W., Spindel, J., and Beraldi, D. (2017). 5-Formylcytosine controls nucleosome positioning through covalent histone-DNA interaction. *BioRxiv* 224444.
(193) Pande, P., Ji, S., Mukherjee, S., Schärer, O.D., Tretyakova, N.Y., and Basu, A.K. (2016). Mutagenicity of a model DNA-peptide cross-link in human cells: roles of translesion synthesis DNA polymerases. *Chem. Res. Toxicol. 30*, 669–677.
(194) Todd, R.C., and Lippard, S.J. (2009). Inhibition of transcription by platinum antitumor compounds. *Metallomics 1*, 280–291.
(195) Ming, X., Groehler IV, A., Michaelson-Richie, E.D., Villalta, P.W., Campbell, C., and Tretyakova, N.Y. (2017). Mass Spectrometry Based Proteomics Study of Cisplatin-Induced DNA–Protein Cross-Linking in Human Fibrosarcoma (HT1080) Cells. *Chem. Res. Toxicol. 30*, 980–995.
(196) You, C., and Wang, Y. (2015). Quantitative measurement of transcriptional inhibition and mutagenesis induced by site-specifically incorporated DNA lesions in vitro and in vivo. *Nat. Protoc. 10*, 1389.
(197) You, C., Dai, X., Yuan, B., Wang, J., Wang, J., Brooks, P.J., Niedernhofer, L.J., and Wang, Y. (2012). A quantitative assay for assessing the effects of DNA lesions on transcription. *Nat. Chem. Biol. 8*, 817-822.
(198) Hanawalt, P.C., and Spivak, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. *Nat. Rev. Mol. Cell Biol. 9*, 958-970.

(199) Mellon, I., Spivak, G., and Hanawalt, P.C. (1987). Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell 51*, 241–249.
(200) Fousteri, M., and Mullenders, L.H. (2008). Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects. *Cell Res. 18*, 73-84.
(201) Menoni, H., Hoeijmakers, J.H., and Vermeulen, W. (2012). Nucleotide excision repair–initiating proteins bind to oxidative DNA lesions in vivo. *J Cell Biol 199*, 1037–1046.
(202) Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *nature 362*, 709-715.
(203) Sedgwick, B. (2004). Repairing DNA-methylation damage. *Nat. Rev. Mol. Cell Biol. 5*, 148-157.
(204) Reiner, B., and Zamenhof, S. (1957). Studies on the chemically reactive groups of deoxyribonucleic acids. *J. Biol. Chem. 228*, 475–486.
(205) Strauss, B.S. (1991). The 'A rule'of mutagen specificity: A consequence of DNA polymerase bypass of non-instructional lesions? *Bioessays 13*, 79–84.
(206) Kou, Y., Koag, M.C., and Lee, S. (2015). N7 methylation alters hydrogen-bonding patterns of guanine in duplex DNA. *J. Am. Chem. Soc. 137*, 14067–14070.
(207) Mattes, W.B., Hartley, J.A., and Kohn, K.W. (1986). Mechanism of DNA strand breakage by piperidine at sites of N7-alkylguanines. *Biochim. Biophys. Acta. 868*, 71–76.
(208) Sczepanski, J.T., Wong, R.S., McKnight, J.N., Bowman, G.D., and Greenberg, M.M. (2010). Rapid DNA–protein cross-linking and strand scission by an abasic site in a nucleosome core particle. *Proc. Natl. Acad. Sci. 107*, 22475–22480.
(209) Zhou, C., Sczepanski, J.T., and Greenberg, M.M. (2013). Histone modification via rapid cleavage of C4′-oxidized abasic sites in nucleosome core particles. *J. Am. Chem. Soc. 135*, 5274–5277.
(210) Sczepanski, J.T., Zhou, C., and Greenberg, M.M. (2013). Nucleosome core particle-catalyzed strand scission at abasic sites. *Biochemistry 52*, 2157–2164.

(211) Zhou, C., Sczepanski, J.T., and Greenberg, M.M. (2012). Mechanistic studies on histone catalyzed cleavage of apyrimidinic/apurinic sites in nucleosome core particles. *J. Am. Chem. Soc. 134*, 16734–16741.
(212) Zhou, C., and Greenberg, M.M. (2012). Histone-catalyzed cleavage of nucleosomal DNA containing 2-deoxyribonolactone. *J. Am. Chem. Soc. 134*, 8090–8093.
(213) Michaelson-Richie, E.D., Ming, X., Codreanu, S.G., Loeber, R.L., Liebler, D.C., Campbell, C., and Tretyakova, N.Y. (2011). Mechlorethamine-induced DNA–protein cross-linking in human fibrosarcoma (HT1080) cells. *J. Proteome Res. 10*, 2785–2796.
(214) Chesner, L.N., and Campbell, C. (2018). A quantitative PCR-based assay reveals that nucleotide excision repair plays a predominant role in the removal of DNA-protein crosslinks from plasmids transfected into mammalian cells. *DNA Repair 62*, 18–27.
(215) Chang, M., Bellaoui, M., Boone, C., and Brown, G.W. (2002). A genome-wide screen for methyl methanesulfonate-sensitive mutants reveals genes required for S phase progression in the presence of DNA damage. *Proc. Natl. Acad. Sci. 99*, 16934–16939.
(216) Gates, K.S., Nooner, T., and Dutta, S. (2004). Biologically relevant chemical reactions of N7-alkylguanine residues in DNA. *Chem. Res. Toxicol. 17*, 839–856.
(217) Torres-Ramos, C.A., Johnson, R.E., Prakash, L., and Prakash, S. (2000). Evidence for the involvement of nucleotide excision repair in the removal of abasic sites in yeast. *Mol. Cell. Biol. 20*, 3522–3528.
(218) Memisoglu, A., and Samson, L. (2000). Contribution of base excision repair, nucleotide excision repair, and DNA recombination to alkylation resistance of the fission yeast Schizosaccharomyces pombe. *J. Bacteriol. 182*, 2104–2112.
(219) Xu, X., Muller, J. G., Ye, Y., and Burrows, C.J. (2008). DNA- protein cross-links between guanine and lysine depend on the mechanism of oxidation for formation of C5 vs C8 guanosine adducts. *J. Am. Chem. Soc. 130*, 703–709.
(220) Bogenhagen, D.F. (2012). Mitochondrial DNA nucleoid structure. *Biochim. Biophys. Acta. 1819*, 914–920.
(221) Alexeyev, M., Shokolenko, I., Wilson, G., and LeDoux, S. (2013). The maintenance of mitochondrial DNA integrity—critical analysis and update. *Cold Spring Harb. Perspect. Biol. 5*, a012641.

(222) Murphy, M.P. (2009). How mitochondria produce reactive oxygen species. *Biochem. J. 417*, 1–13.
(223) Gray, G.M., and Yardley, H.J. (1975). Mitochondria and nuclei of pig and human epidermis: isolation and lipid composition. *J. Invest. Dermatol. 64*. 423-430.
(224) Dimauro, I., Pearson, T., Caporossi, D., and Jackson, M. J. (2012). A simple protocol for the subcellular fractionation of skeletal muscle cells and tissue. *BMC Res. Notes 5*, 513.
(225) You, C., and Wang, Y. (2016). Mass spectrometry-based quantitative strategies for assessing the biological consequences and repair of DNA adducts. *Acc. Chem. Res. 49*, 205–213.
(226) Rydberg, B., and Lindahl, T. (1982). Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J. 1*, 211–216.
(227) Wunderlich, V., Schütt, M., Böttger, M., and Graffi, A. (1970). Preferential alkylation of mitochondrial deoxyribonucleic acid by N-methyl-N-nitrosourea. *Biochem. J. 118*, 99–109.
(228) Tomasi, A., Albano, E., Banni, S., Botti, B., Corongiu, F., Dessi, M.A., Iannone, A., Vannini, V., and Dianzani, M.U. (1987). Free-radical metabolism of carbon tetrachloride in rat liver mitochondria. A study of the mechanism of activation. *Biochem. J. 246*, 313–317.
(229) Jung, D., Cho, Y., Collins, L.B., Swenberg, J.A., and Di Giulio, R.T. (2009). Effects of benzo [a] pyrene on mitochondrial and nuclear DNA damage in Atlantic killifish (Fundulus heteroclitus) from a creosote-contaminated and reference site. *Aquat. Toxicol. 95*, 44–51.
(230) Croteau, D.L., Stierum, R.H., and Bohr, V.A. (1999). Mitochondrial DNA repair pathways. *Mutat. Res. Repair 434*, 137–148.
(231) Thorslund, T., Sunesen, M., Bohr, V.A., and Stevnsner, T. (2002). Repair of 8-oxoG is slower in endogenous nuclear genes than in mitochondrial DNA and is without strand bias. *DNA Repair 1*, 261–273.
(232) Liu, P., Qian, L., Sung, J.S., de Souza-Pinto, N.C., Zheng, L., Bogenhagen, D.F., Bohr, V.A., Wilson, D.M., Shen, B., and Demple, B. (2008). Removal of oxidative DNA

damage via FEN1-dependent long-patch base excision repair in human cell mitochondria. *Mol. Cell. Biol. 28*, 4975–4987.

(233) Wisnovsky, S., Jean, S.R., and Kelley, S.O. (2016). Mitochondrial DNA repair and replication proteins revealed by targeted chemical probes. *Nat. Chem. Biol. 12*, 567.