Formation, Degradation, and Bypass of DNA-Protein Crosslinks

By

Carl A. Sedgeman

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

September 30th, 2018

Nashville, Tennessee

Approved By:

F. Peter Guengerich, PhD

Martin Egli, PhD

Brandt Eichman, PhD

Neil Osheroff, PhD

Carmelo Rizzo, PhD

For my brother, Dan, for always pushing me throughout my life and for setting the bar high.

I cannot wait to see you again.

ACKNOWLEDGEMENTS

I would like to thank Dr. Fred Guengerich for his constant support for me during my tenure in his lab. He was a great role model for me as a scientist, and I will never forget the skills and lessons that he taught me.

I thank Kathy Trisler for her help and support throughout my time in lab. I also need to thank the rest of the Dogs of the Guengerich Lab for their guidance, support, and company throughout these years. There have been too many of them to mention personally, but I am grateful to have shared my time in the lab with so many good scientists, great people, and everlasting friends.

I wish to thank my committee members, Dr. Neil Osheroff, Dr. Martin Egli, Dr. Carmelo Rizzo, and Dr. Brandt Eichman. I appreciated you volunteering your time in guiding my research and keeping me on track.

I also want to thank the Biochemistry department, especially Dr. John York and Jen Smith. They have worked tirelessly to make this department one of the finest on campus by bringing in excellent speakers, hosting great department retreats, and giving me a wonderful social network of scientists to work with.

I wish to acknowledge the Chemical Biology Interface training grant (T32 GM065086) for funding my research as well as giving me another layer of support on campus.

I would like to thank my family. My parents, Tom and Joyce, have been instrumental in both my education and personal well-being. I would like to thank my sister, Lynnette, for always making me see the greater picture and reminding me to look at things from a different prospective.

And lastly, I would be remorse if I didn't thank my wife, Leslie, for her perpetual love and support for me. She makes me be a better person and my life more enjoyable.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AGT	O^6-alkylguanine DNA alkyltransferase
BER	base excision repair
BOC	*N*-(*tert*-butyloxycarbonyl)
BSA	bovine serum albumin
CPD	cyclobutane pyrimidine dimer
DBE	1,2-dibromoethane
DIPEA	*N,N*-diisopropylethylamine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
dNTP	deoxynucleotide triphosphate
DPC	DNA-protein crosslink
ds-DNA	double-stranded DNA
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionization
FAM	6-carboxyfluorescein
FPLC	fast protein liquid chromatography
GSH	glutathione
GST	glutathione transferase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	high-performance liquid chromatography
hPol	human DNA polymerase
HR	homologous recombination
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl-β-D-1-thiogalactopyranoside
LB	lysogeny broth
LC	liquid chromatography
LDS	lithium dodecyl sulfate
MRE11	meiotic recombination 11
MS	mass spectrometry
MSH	*O*-(mesitylsulfonyl)hydroxylamine
NBS1	Nijmegan breakage syndrome protein 1
NHEJ	nonhomologous end joining

NMR	nuclear magnetic resonance
NTA	nitrilotriacetic acid
PCNA	proliferating cell nuclear antigen
RJALS	Ruijs-Aalfs syndrome
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
ss-DNA	single-stranded DNA
TCEP	tris(2-carboxyethyl)phosphine
TDP	tyrosyl-DNA phosphodiesterase
TDT	terminal deoxynucleotidyl transferase
TEV	tobacco etch virus
TLS	translesion DNA synthesis
UDG	uracil-DNA glycosylase
UPLC	ultra-performance liquid chromatography

CHAPTER 1

Introduction

1.1 DNA Polymerases.

DNA polymerases are enzymes that replicate DNA and are vital to the overall function of all living organisms. The enzymes work to create two new identical strands of DNA from two parent strands of DNA. The DNA polymerase 'reads' an existing template strand of DNA and creates a new strand that complements the first strand. This works in a semi-conservative fashion where the two 'parent' strands of DNA are unwound and the two newly synthesized strands form separately with the complementary pair for each, giving two copies of double-stranded DNA with one 'parent' and one 'daughter' strand of DNA (1). This family of enzymes is highly proficient in respect to both its rate of catalysis, at least for the processive DNA polymerases, as well as its astonishing fidelity in preventing misinsertions in the replication of DNA (2).

The catalytic function of DNA polymerases is to synthesize DNA from deoxyribonucleotides (Figure 1.1). The polymerase works by binding to the template strand of DNA that the new primer strand is synthesized against. After binding to the template strand, the polymerase binds the nucleotide that complements the base in the template strand. The base pairing combines cytosine with guanine and thymine with adenine through Watson-Crick base interactions. Once the correct deoxyribonucleotide is in the active site of the polymerase with the template strand, the enzyme adds the nucleotide to the primer strand by forming a new phosphodiester bond between the terminal 3´-hydroxyl group on the primer strand and the 5´-α-phosphate on the deoxyribonucleotide. The deoxyribonucleotide is inserted onto the primer strand and pyrophosphate is released. The polymerase then shifts down to the next free template base and the catalytic cycle is repeated. All DNA polymerases work in the 5´ to 3´ direction for the insertion of nucleotides into the primer strand of DNA.

The one commonality of all DNA polymerases is their DNA synthesis catalytic domain. In addition to the catalytic domain, the sequence of each DNA polymerase codes for secondary functions such as exonuclease activity, binding motifs (PCNA binding), or regulatory purposes (ubiquitylation sites). The structure of the catalytic domain of a DNA polymerase is often compared to that of a right hand, complete with subdomains called the thumb, palm, and finger (3-6). The thumb subdomain is responsible for binding to the DNA substrate. The finger subdomain interacts with the incoming deoxynucleotide and plays a role in positioning it near the template strand of DNA. The finger domain is important for monitoring the interactions between the incoming nucleotide and the DNA template to insure correct nucleotide incorporation and the preservation of the sequence of DNA. The finger domain is also responsible for the major conformational change that the polymerase complex undergoes prior to the catalytic step (7). The palm subdomain contains the two magnesium ions that are required for catalytic function of the enzyme. The palm domain contains the catalytic residues of the polymerases that perform the chemistry of the phosphoryl transfer (8).

1.2 Families of DNA Polymerases.

There are multiple families of DNA polymerases in both prokaryotic and eukaryotic cells, including families A, B, C, D, X, Y, and RT (Table 1). Of these, Families A, B, X, Y, and RT are found in eukaryotic organisms and in humans. Humans have 17 known DNA polymerases, many with overlapping functions (9).

Family B DNA polymerases include α, δ, ε, and ζ. Pols δ and ε are the two crucial DNA polymerases found in humans. These are responsible for the majority of all replication. Pol δ performs DNA replication in the lagging strand of DNA synthesis, whereas Pol ε advances replication in the leading strand of DNA synthesis. Both of these polymerases have high processivity and insert several nucleotides per second. Although these polymerases have similar

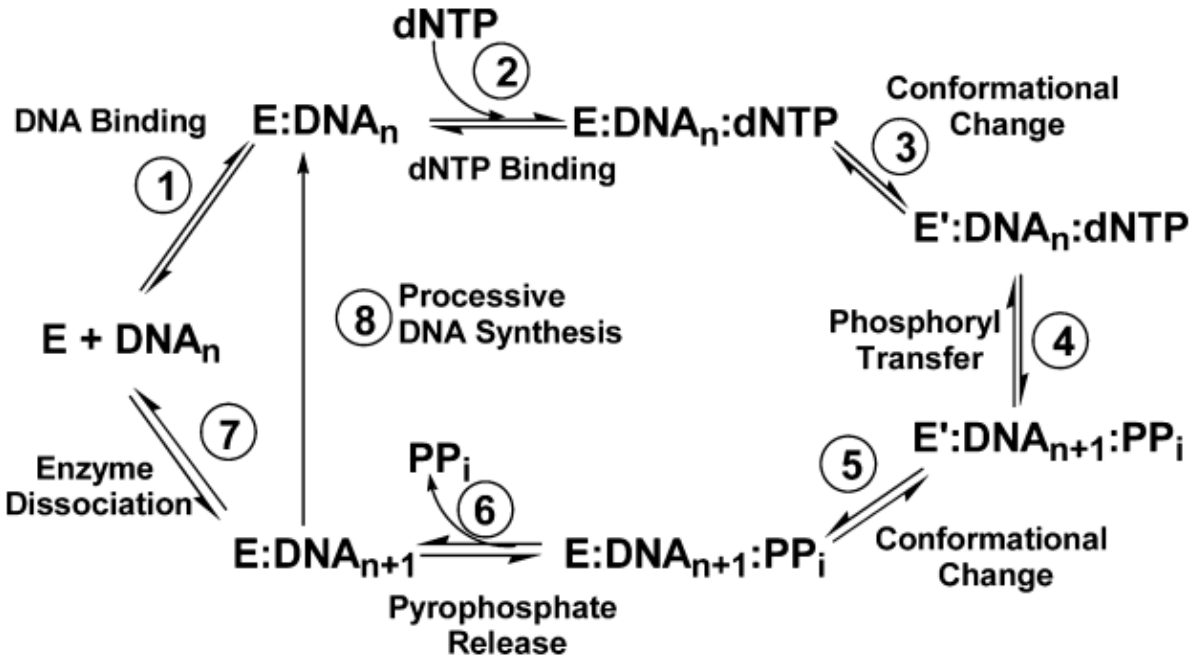


Figure 1.1 Catalytic Cycle for DNA Polymerases. Adapted from Choi *et al.* (10).

functions, however on different strands of DNA, loss of either polymerase δ or ε is embryonic lethal. While these polymerases are naturally very selective in incorporating the correct nucleotide during replication, they also have 3´-5´ exonuclease activity, which allows them to proofread its previous insertions in order to replace mis-inserted nucleotides from the newly synthesized DNA strand when the wrong nucleotide is inserted on the primer strand. This ability to replace the misinserted nucleotide alleviates the errors caused by these replicative polymerases and gives these polymerases quite low levels of misincorporation (~1 x 10^{-9}) (11). As well, the loss of this exonuclease function has been shown to result in tumorigenesis in mice (12). Pol α is responsible for initiating DNA replication on both strands of DNA, where it forms a complex with DNA primase. DNA primase inserts a short RNA primer onto the DNA template to start the replication process. Pol α then switches to the DNA template and uses an RNA primer to replicate approximately twenty nucleotides before switching off to the Pols δ and ε (13). Pol ζ is a translesion DNA polymerase and is involved in the inserting deoxyribonucleotides opposite stalled DNA replication forks when Pols δ and ε are stalled. It is also one of two known DNA polymerases to localize and perform mitochondrial DNA replication and repair (14). Pol ζ is believed to have a role in recruiting other translesion DNA polymerases to the active site of stalled replication forks (15).

DNA polymerases that belong to family A include polymerases γ, θ, and ν. Polymerase γ is the major mitochondrial DNA polymerase and is responsible for mitochondrial DNA replication (16,17). It also has a DNA repair function in the mitochondria alongside pol ζ and contains a 3´-5´ exonuclease function similar to the B family polymerases. Pol θ is a nuclear DNA polymerase involved in DNA repair and translesion DNA synthesis (18). One of its major functions is for the repair of double-strand DNA breaks through non-homologous end-joining (NHEJ) (19). Pol θ is capable of bypassing various DNA lesions such as thymidine glycol DNA adducts and abasic sites (20). Its amino acid sequence is similar to other A family polymerases except that it does not have any exonuclease function, resulting in higher error rates that are similar to Y-family

translesion DNA polymerases (see below). Pol ν is the last of the A family of polymerases. This enzyme also is capable of performing translesion DNA synthesis opposite thymidine glycol (21). As well, it has a role in DNA repair through homologous recombination and is suspected of being involved in the repair of DNA cross-link repair through the Fanconi anemia pathway (22).

Family X DNA polymerases β, λ, μ, and TdT are mainly responsible for DNA repair but also have some translesion replication activity (23). The most recognized polymerase in this family is pol β, which is well characterized for its role in base excision repair (BER) in the removal of alkylated and oxidized bases as well as abasic sites from DNA (24). Loss of pol β in animal studies was shown to be lethal and resulted in death immediately after birth (25). Both pols β and λ have been shown to bypass the DNA adducts 8-oxoguanine and 2-hydroxyadenine. DNA polymerase λ is distinct from most other DNA polymerases for its preferential utilization of manganese as its divalent metal ion compared to magnesium for other polymerases (26). Pol λ also is considered to be involved in non-homologous end-joining (27,28). Pol μ associates with XRCC4 and DNA ligase IV, suggesting that it is involved in V(D)J recombination (29). Terminal deoxynucleotidyl transferase (TdT) is a specialized DNA polymerase that adds nucleotides in the variable regions of immunoglobulins during V(D)J recombination, and it does not have the traditional ability of a DNA polymerase to incorporate nucleotides opposite a template strand of DNA (30).

The Y-family DNA polymerases are involved in performing translesion DNA synthesis (TLS). This family includes polymerases η, ι, κ, and REV1. Y-family DNA polymerases are capable of bypassing DNA lesions when the replicative DNA polymerases are stalled. The importance of this family in performing TLS is signified by the evolutionary history of these polymerases, which are found in budding yeast (REV1) and even have archaeal and bacterial ancestry (pol κ orthologues Dpo4 and Pol IV). These polymerases characteristically have a significantly larger active site compared to the replicative DNA polymerases δ and ε, which allows

Table 1.1 The 17 Human DNA Polymerases and Their Functions. Adapted from van Loon *et al.* (9).

DNA pol	HUGO name	Family	Proposed functions
α	POLA	B	Nuclear DNA replication/Initiation and priming
δ	POLD1	B	Nuclear DNA replication; DNA repair
ε	POLE1	B	Nuclear DNA replication; DNA repair
ζ	POLZ	B	TLS
Y	POLG	A	Mitochondrial DNA replication and repair
θ	POLQ	A	TLS, DNA repair (end-joining)
v	POLN	A	TLS
λ	POLL	X	NHEJ; base excision repair; TLS
β	POLB	X	Base excision repair
TdT	TDT	X	NHEJ; diversity
μ	POLM	X	NHEJ; diversity
η	POLH	Y	TLS
ι	POLI	Y	TLS
κ	POLK	Y	TLS
REV1	REV1	Y	TLS
Telomerase	TERT	RTd	Telomere maintenance
PrimPol	PRIMPOL	AEP	DNA replication, replication fork repriming; TLS

the Y-family polymerases to tolerate the additional size of the DNA adduct and continue on with the DNA replication. Along with the X-family of polymerases, the Y-family enzymes do not contain exonuclease functionality.

1.3 Features of Translesion DNA Polymerases.

The general structural features of the catalytic domain for Y-family polymerases are similar for the replicative DNA polymerases (31). Although they have poor sequence conservation, both feature similar catalytic domains featuring the 'right hand' topology. The active site for the Y-family polymerases is ample to accommodate the adduct in the template strand of DNA, leaving the active site more solvent exposed compared to the replicative polymerases (32). The finger and thumb domains are shorter, causing fewer DNA connections with both the DNA and the incoming dNTP. These modifications are crucial for allowing TLS to occur but lead to a decrease in processivity and fidelity of the enzyme. Y-family DNA polymerases also contain a little finger domain that replicative polymerases lack (31,33). This domain is responsible for mediating the contact with the DNA lesion in the active site of the polymerase. Orientation of this little finger domain also is selective for the type of DNA lesion that the polymerase is capable of bypassing (33).

Y-family DNA polymerases are activated during DNA replication when the replicative DNA synthesis machinery pauses. There are two known pathways by which TLS can become activated. The first canonical pathway is through the ubiquitination of proliferating cell nuclear antigen (PCNA) (Figure 1.2) (34). PCNA ubiquitination signals the replicative polymerase (δ or ε) to dissociate from the DNA and the subsequent recruitment of Y-family polymerases to bind to the replication fork. The polymerases are capable of bypassing the lesion on the template strand and incorporating a free nucleoside on the newly synthesized primer. After this incorporation, the TLS polymerase will extend the replication fork along for several bases before PCNA is deubiquitinated, thus ending TLS replication. The Y-family polymerase dissociates from the

replication fork, allowing the replicative polymerase to switch back onto the fork. The second pathway is independent of PCNA ubiquitination, utilizing interactions within the polymerase REV1 to be recruited to stalled replication forks as described below (35,36).

Humans contain four DNA polymerases specific for translesion DNA synthesis, and each of these enzymes performs replication opposite different types of DNA lesions (32). Pol η is best known for its bypass of cyclobutane pyrimidine dimer (CPD) lesions (37). These lesions are commonly the result of UV radiation from sunlight. One of the variants of the disease xeroderma pigmentosum is caused by the deficiency of DNA polymerase η, resulting in skin discoloration and melanoma (38). It has the ability to bypass various other bulky lesions, primarily in the major groove of DNA. Pol η is also known for causing chemotherapeutic resistance in patients treated with cisplatin and other related drugs (39-41). Pol κ has been shown to be important in the bypass of the major cellular oxidation DNA lesion, 8-oxo-(7,8-dihydro)guanine, as well as other common DNA lesions such as thymidine glycol and benzo[*a*]pyrene (42-46). The active site configuration of pol κ allows it to perform TLS opposite bulky minor groove adducts, namely N^2-guanine and N^3-adenine adducts. Pol κ has also been implicated for being involved in a number of DNA repair mechanisms including UV irradiation and the repair of DNA interstrand crosslinks through an involvement in the nucleotide excision repair (NER) pathway (47-49). Pol ι is a unique DNA polymerase for a number of reasons. First, it uses Hoogsteen base pairing for inserting nucleotides into the DNA primer instead of normal Watson-Crick base pairing (50). This causes pol ι to have a lower fidelity in incorporating the correct nucleotide compared to other Y-family and other DNA polymerases, especially for incorporation opposite dT. Magnesium is the canonical metal in the active site of DNA polymerases, but pol ι has also been shown to be more active in the presence of manganese (51,52). DNA replication performed by pol ι is distributive; the enzyme will incorporate one nucleotide onto the primer and then dissociate from the DNA (53,54). Longer incubations with the enzyme with replication forks will result the incorporation of

only a few nucleotides, while the other Y-family polymerases will continue on and incorporate several nucleotides. REV1 is not particularly known for its activity as a DNA polymerase. Its ability is restricted to the insertion of deoxycytidine opposite guanines in undamaged DNA as well in the bypass of bulky N^2-G adducts (55-58). REV1 is mainly used as a scaffolding protein to recruit other Y-family DNA polymerases to the stalled replication fork (35,36). The other human Y-family DNA polymerases each contain a REV1-interacting region (RIR) motif that binds to REV1 (59-61). This interaction is believed to crucial for the non-canonical mechanism for the recruitment of translesion polymerases for TLS. REV1 is also known to associate with pol ζ and is believed to be involved in the scaffolding for other DNA repair pathways (62).

While the ability to resume stalled replication forks is advantageous to the cell to prevent long strands of single-stranded DNA, one major disadvantage of TLS DNA polymerases is that they are prone to causing misinsertions in DNA. This defect is caused in part by the lack of exonuclease function of the Y-family DNA polymerases, which replicative polymerases use to remove misinserted nucleotides and drastically reduce the mutation frequency for the replication cycle. The larger active sites of the Y-family DNA polymerases are vital to allow the enzymes to replicate past DNA lesions, but the larger active site also infringes on the proofreading ability of the enzymes. Y-family DNA polymerases perform TLS with greater fidelity compared to the replicative polymerases, but this misincorporation rate is considerably higher compared to replicative polymerases opposite undamaged DNA, and the Y-family DNA polymerases are also less stringent for the normal incorporation of nucleotides opposite undamaged DNA. The Y-family DNA polymerases can also be detrimental during certain chemotherapeutic treatments, permitting the cell to survive the damage caused by the drug instead of undergoing programed cell death. This phenomenon has been widely studied in the case of cisplatin, and the resistance that is caused by pol η. This information has led to the development of potential chemotherapeutics to also target these polymerases.

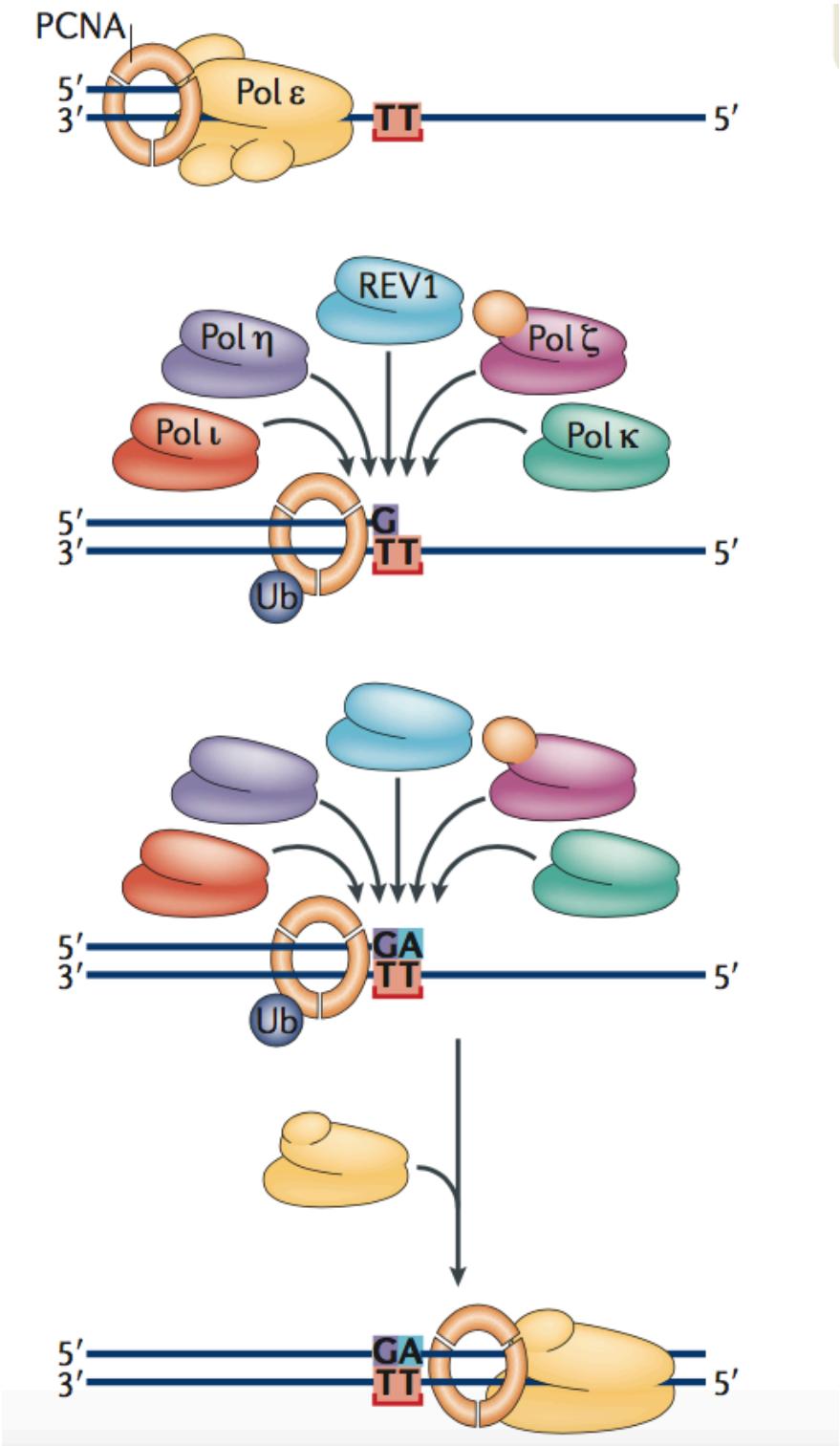


Figure 1.2 Mechanism for Translesion DNA synthesis. Adapted from Sale *et al* (63).

1.4 Chemical Carcinogens.

Chemical carcinogens are substances that can act as genotoxins and damage DNA, creating alterations to the genetic code (Figure 1.3). While causing dangerous effects in cells, carcinogens are not necessarily immediately toxic to the cell. Therefore, their effects can manifest for extended periods of time before afflicting the cell. These lesions can lead to a variety of severe diseases such as neurodegenerative disorders and cancer. Chemical carcinogens come from a wide array of sources. They can be formed naturally in various species such as aflatoxin B_1 from *Aspergillus flavus* as well as be found in various environments (arsenic). Carcinogens can also be man-made products, such as the nitrogen mustards and cisplatin that are used as chemotherapeutic agents. Carcinogens are also commonly found in the byproducts of combustion reactions such as industrial and tobacco smoke, byproducts of cooking of food, and fuel exhaust (formaldehyde, benzo[*a*]pyrene, nitrosamines, and 1,3-butadiene).

Many chemical carcinogens work by adducting to the DNA (Figure 1.4). This adduction, or formation of DNA adducts, can cause negative effects numerous ways, based on the nature and characteristics of the DNA adduct. Bulky DNA adducts can alter the overall structure of DNA, thus changing the pattern of DNA binding proteins and disrupting signaling events on nearby DNA (64,65). Bulky DNA adducts can also block nuclear machinery from progressing alongside DNA (64). This disruption can affect processes such as DNA replication and RNA transcription, triggering long stretches of single strand DNA that is prone to breakages. Other DNA adducts localize inside of the DNA double helix and act as DNA intercalators (66). Intercalators position between the stacked DNA bases and often cause a kink or disruption in the DNA stacking interactions of the double helix. The disruption of the stacking interactions can also interfere with in the overall tertiary structure of DNA and block nuclear processes similarly to bulky DNA adducts. These types of adducts normally directly cause DNA damage due to the formation of DNA breaks, which can lead to DNA mutations and cancer. Smaller DNA adducts do not disrupt the cell until they are encountered by replication and transcription machinery. Depending on the

location of the modification, an adduct can disrupt the normal Watson-Crick base pairing of the DNA and cause a misinsertion to occur. If DNA repair proteins do not remove the adduct or the mispaired base, the misinsertion will remain in the DNA strand and become a DNA mutation.

Although there is a wide array of DNA adducts, our cells have numerous ways to identify and correct DNA damage. Cellular antioxidants, such as glutathione, are utilized to sequester various reactive oxygen species and prevent their reaction with DNA (67). Direct repair proteins such as O^6-alkylguanine DNA alkyltransferase (AGT) and AlkB directly remove DNA adducts to repair the original DNA base (68). Other repair mechanisms such as base-excision repair (BER), nucleotide excision repair (NER), and mismatch repair use enzymes to remove the DNA surrounding the damage before replacing it with new correctly synthesized DNA (69).

1.5 Halogenated Hydrocarbons.

One class of chemical carcinogens consists of halogenated hydrocarbons. These are carbon molecules linked to one or more halogens (fluorine, chlorine, bromine, or iodine). Halogenated hydrocarbons are broadly used commercially and industrially as organic solvents, refrigerants, and propellants and in the generation of plastics and other polymers. Although these are widely used and are commercially important, there are some toxic effects of this class of molecules, dating back one hundred years to the use of mustard gases as chemical warfare agents in World War I. Chlorofluorocarbons were extensively used until their role in the depletion of the ozone was discovered. Many halogenated hydrocarbons are also alkylating agents, causing the transfer of an alkyl group from the halogenated hydrocarbon to nucleophilic groups such as thiols and amines on other compounds (70). This can have severe medical ramifications in that these alkylation reactions can occur with various biological molecules, and their roles in alkylating DNA defines some as chemical carcinogens (71,72). The alkylation of DNA can cause the disruption of nuclear processes such as DNA replication and RNA transcription, which can initiate tumors in

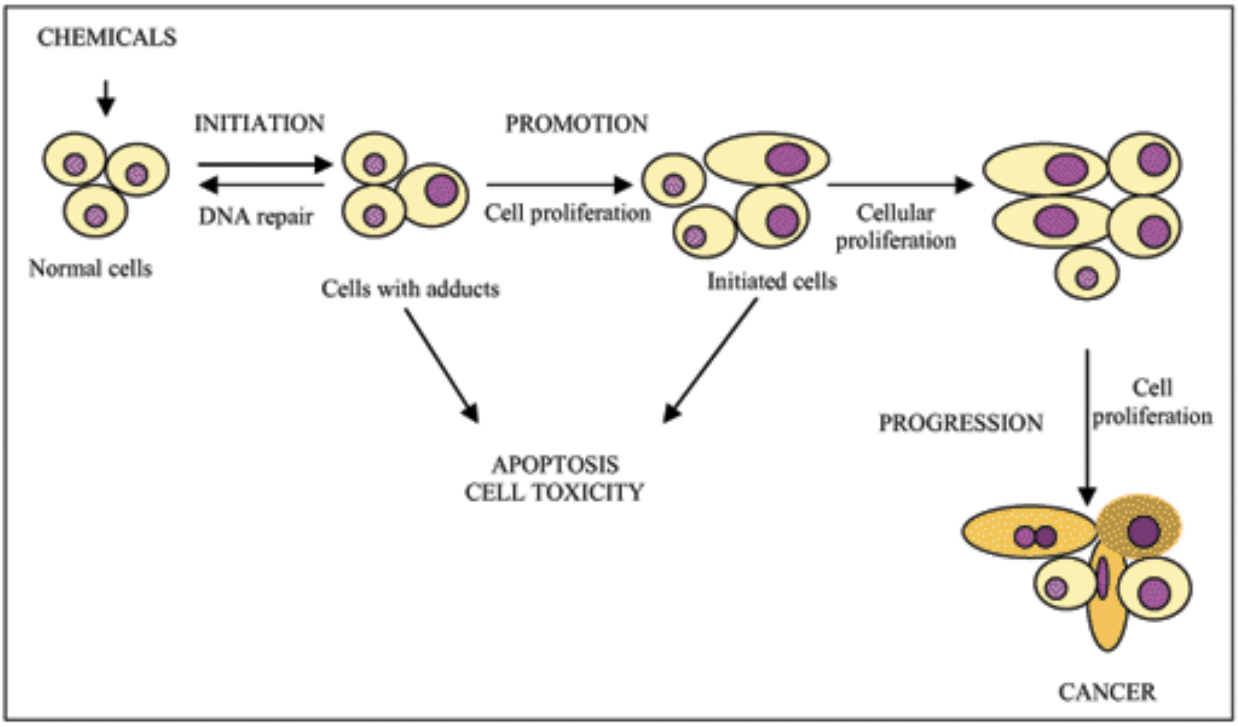


Figure 1.3 Mechanism of Action for Chemical Carcinogens. Adapted from Oliveira *et al.* (73).

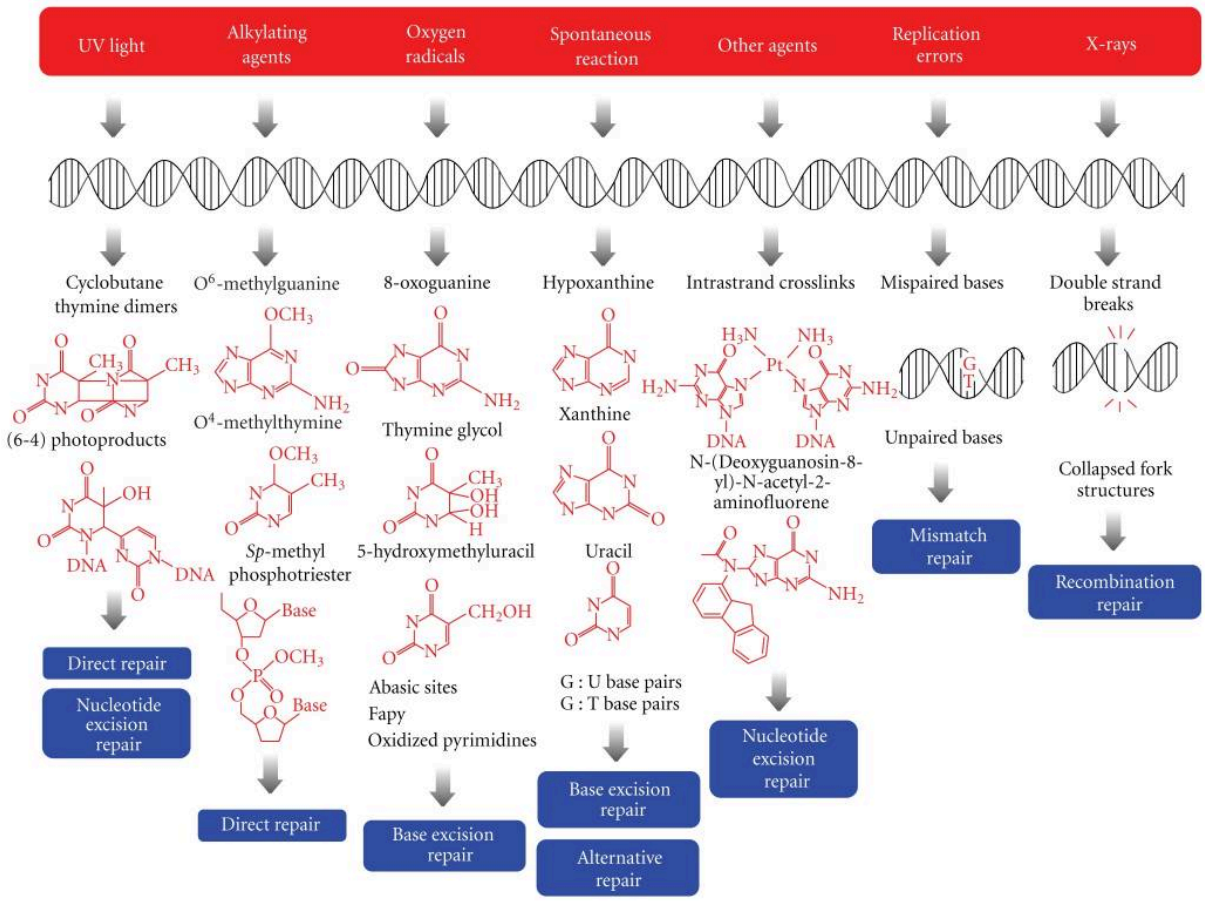


Figure 1.4 Types of DNA Damage and Repair. Adapted from Morita *et al.* (74).

an organism. Their ability to alkylate DNA also gave rise to their use as chemotherapeutic agents to treat cancers as well (75).

Historically, 1,2-dibromoethane (DBE, ethylene dibromide) has been used as an anti-knock additive in leaded gasoline and also has had uses as a pesticide in soil and in the organic synthesis of various compounds, e.g. vinyl bromide. Annual production of DBE was >330 million pounds in the United States during the 1970s and has since rapidly declined, due to its health risks and ensuing regulations, to 1-10 million pounds per year by the end of the 20th century. DBE is a known carcinogen and had the highest hazard score for carcinogens in the Human Exposure/Rodent Potency Index prior to changes in the restrictions on the compound (76). DBE causes tumors in rats and mice in several different tissues, including the nasal cavity, blood vessels, skin, lung, kidney, and liver (77-80).

DBE is detrimental because it is a *bis*-electrophile. The presence of two bromine atoms in the molecule allows it to undergo two nucleophilic substitution reactions, enabling the molecule to interact and become covalently bound to two different entities. This gives DBE the ability to crosslink two biomolecules to each other, including two proteins, two strands of nucleic acids, or a protein to a nucleic acid.

1.6 DNA-Protein Crosslinks.

When a protein becomes covalently bound to DNA through a crosslinking agent such as DBE, this process can cause severely detrimental effects to the cell including cytotoxicity and DNA mutations. DNA-protein crosslinks (DPCs) can comprise very bulky proteins that can alter the three-dimensional folding of DNA in the nucleus by affecting the ability of the DNA supercoiling as well as its interactions with the DNA histones, indirectly leading to alteration of DNA synthesis initiation and transcription initiation of RNA (81). In addition, DPCs can directly block the nuclear processes of DNA replication and RNA transcription (82-84). Both DNA and RNA polymerases contain narrow active sites to ensure high fidelity. Both of these enzyme classes are not

structurally designed to accommodate a protein covalently bound to the coding strand of DNA and consequently become blocked, which can cause detrimental effects to the cell by means of DNA replication stalling, single- and double-strand DNA breaks, and prevention of protein synthesis.

A few studies have looked at the peptides and proteins that can become bound to DNA in DPCs (85-87). The first peptide found to be crosslinked was glutathione (GSH), which as described above is crucial for the removal of reactive electrophiles from the cell (88,89). Its mechanism for action uses GSH *S*-transferase (GST) to conjugate the protein to an electrophile through its reactive cysteine thiol, thus sequestering it from reacting with DNA (Figure 1.5). However, in the presence of a *bis*-electrophile like DBE, the second electrophile is still present after the conjugation to the first bromine to GSH. In the case of DBE, the second bromine can then react with the sulfur moiety of GSH via anchimeric assistance and form an episulfonium ion (89). This three-member ring is quite unstable due to the steric hinderance of the cyclic bonds. This process allows the cyclic group to react rather easily and was found to react with DNA. This paradoxically makes GSH part of the DNA adduct instead of preventing the electrophile from reacting with DNA.

The same phenomenon has been characterized with the repair protein AGT. AGT utilizes its active site cysteine to find sites of DNA damage on the O^6 position of guanine. Once it finds an adduct on this site, the cysteine reacts with the alkyl lesion on the guanine base and adducts it to the protein, thus signaling the protein for degradation. However, the *bis*-electrophile DBE will react with the active site cysteine sulfur atom of AGT (Cys-145) and form a highly reactive episulfonium ion, based on the results of work done with GSH (90). This episulfonium (also called thiiranium) ion then reacts with DNA to form a DPC at various sites on the DNA bases (92). The lesion has been found to occur at numerous sites on DNA and can cause significant detrimental effects to the cell, including cytotoxicity and an increase in DNA mutations (91,93).

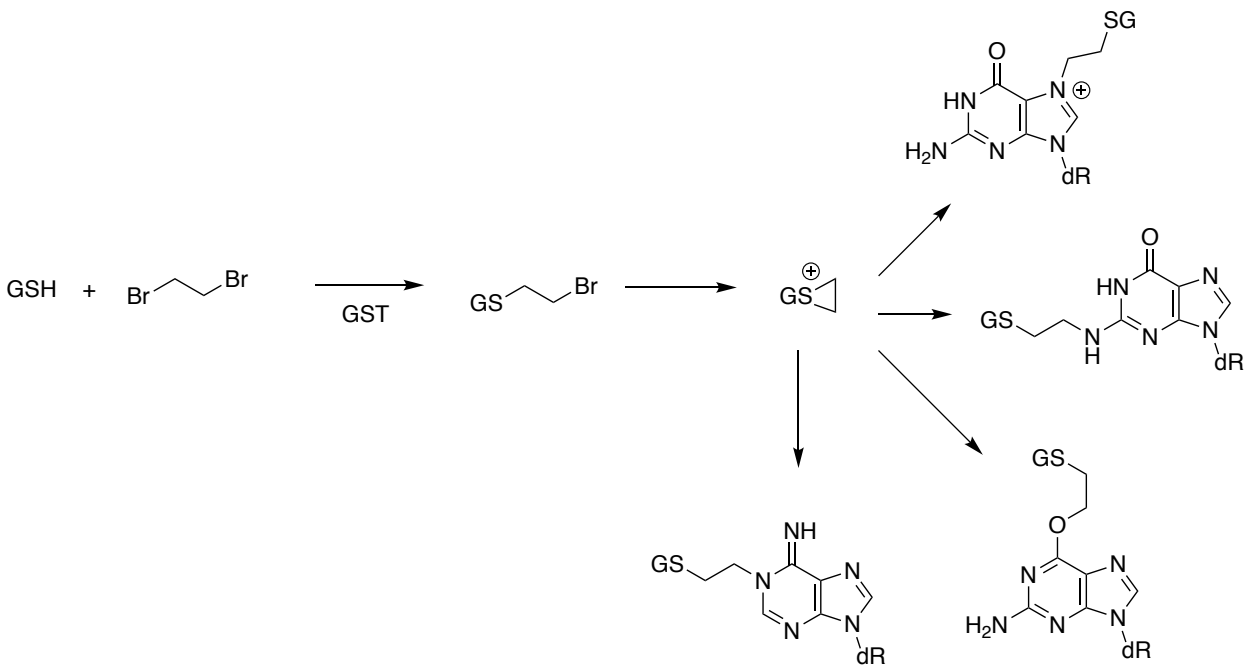


Figure 1.5 Mechanism of Formation of DNA-Protein Crosslink from Glutathione and Dibromoethane. Adapted from Cmarik *et al.* (99).

Broader studies have utilized proteomics to search for other proteins that can form DPCs via crosslinking agents such as BDE (86,87). These results show hundreds of proteins of varying cellular function and localization (Figure 1.6). Only two other proteins (histone H2b and glyceraldehyde 3-phosphate dehydrogenase) have been studied to test the effects of their crosslink (94,95). Peculiarly, neither of these proteins caused an increase in the number of cellular mutations in the cells treated (Figure 1.7). This leads to the hypothesis that only a select number of proteins will form DPCs with detrimental effects. The reason for why these proteins have detrimental effects and others do not is still unknown.

1.7 Repair of DNA-Protein Crosslinks.

DNA-protein crosslinks can be repaired in multiple ways. Select DPCs can be directly removed by the tyrosyl-DNA phosphodiesterase family of enzymes (TDP1 and TDP2). These enzymes can repair crosslinks that are formed through a linkage between the DNA phosphodiester backbone and tyrosine amino acids (96). This repair pathway is used for the removal of DNA topoisomerase crosslinks that can become trapped when cleaving either one or both strands of DNA to resolve torsion in the winding of the DNA helical structure. TDP1 resolves topoisomerase 1 crosslinks that form single-strand breaks. The result of the hydrolysis of the phosphotyrosyl bond formed by topoisomerase 1 is handled by single-stranded break repair enzymes PNPK, PARP1, XRCC1, and LIG3 to process and to fill in the gap on the DNA (97). TDP2 resolves crosslinks formed by topoisomerase 2, which normally forms double-stranded DNA breaks to relieve the helical stress on DNA (98). These double-stranded breaks are repaired by either of the canonical pathways, homologous recombination or non-homologous end joining. Unlike TDP1, TDP2 is unable to gain access to the phosphotyrosyl bond of topoisomerase 2 and requires proteolytic degradation of the topoisomerase 2 prior to repair.

Another mechanism for the removal of DPCs is through nuclease-dependent repair. This type of repair is dependent on the degradation of the DNA that surrounds the crosslinked protein. This system requires repair proteins to remove the damaged DNA. Endonucleases cleave both strands of DNA by the crosslink, leaving a double strand DNA break. This can result in the recruitment of repair proteins that are involved in homologous recombination (HR). HR utilizes the second copy of DNA to repair the double-stranded DNA break and prevent the loss of the genetic information that was coded in the lost sequence of DNA. This process can proceed in one of two pathways. The first pathway results in the crossing over of the second copy of DNA into the first copy prior to replication, resulting in Holiday junctions. The subsequent pathway uses the second copy of DNA only as a template to fill in one strand of the DNA, then using that strand as the template for the second strand of DNA. This repair pathway, while cumbersome, is a preferred pathway for cells to repair double-stranded breaks because it restores the DNA without losing any genetic information. The other process for the repair of double-stranded breaks is NHEJ. NHEJ does not utilize the second copy of DNA but instead cleans the termini of the DNA prior to ligating the broken ends together. This mechanism can be mutagenic due to the loss of the DNA that is lost during the double strand break or during the processing by the NHEJ proteins to remove DNA bases prior to ligation.

Double-stranded DNA breaks can also occur without removing the DPC but rather leaving it on the terminus of cleaved DNA. This type of DPC can be removed by the MRN complex, consisting of meiotic recombination 11 (MRE11), RAD50, and Nijmegan breakage syndrome protein 1 (NBS1). These enzymes will cut the DNA prior to the DPC break that is then repaired by either HR or NHEJ. A simpler repair mechanism is NER. This strategy recognizes bulky DNA lesions and excises only the damaged strand of DNA and leaves the undamaged strand alone, thus causing only a single-stranded DNA break that are significantly less harmful compared to double-stranded DNA breaks. Once the damaged strand is excised from DNA, polymerases are recruited to synthesize the complementary sequence from the undamaged single-stranded DNA.

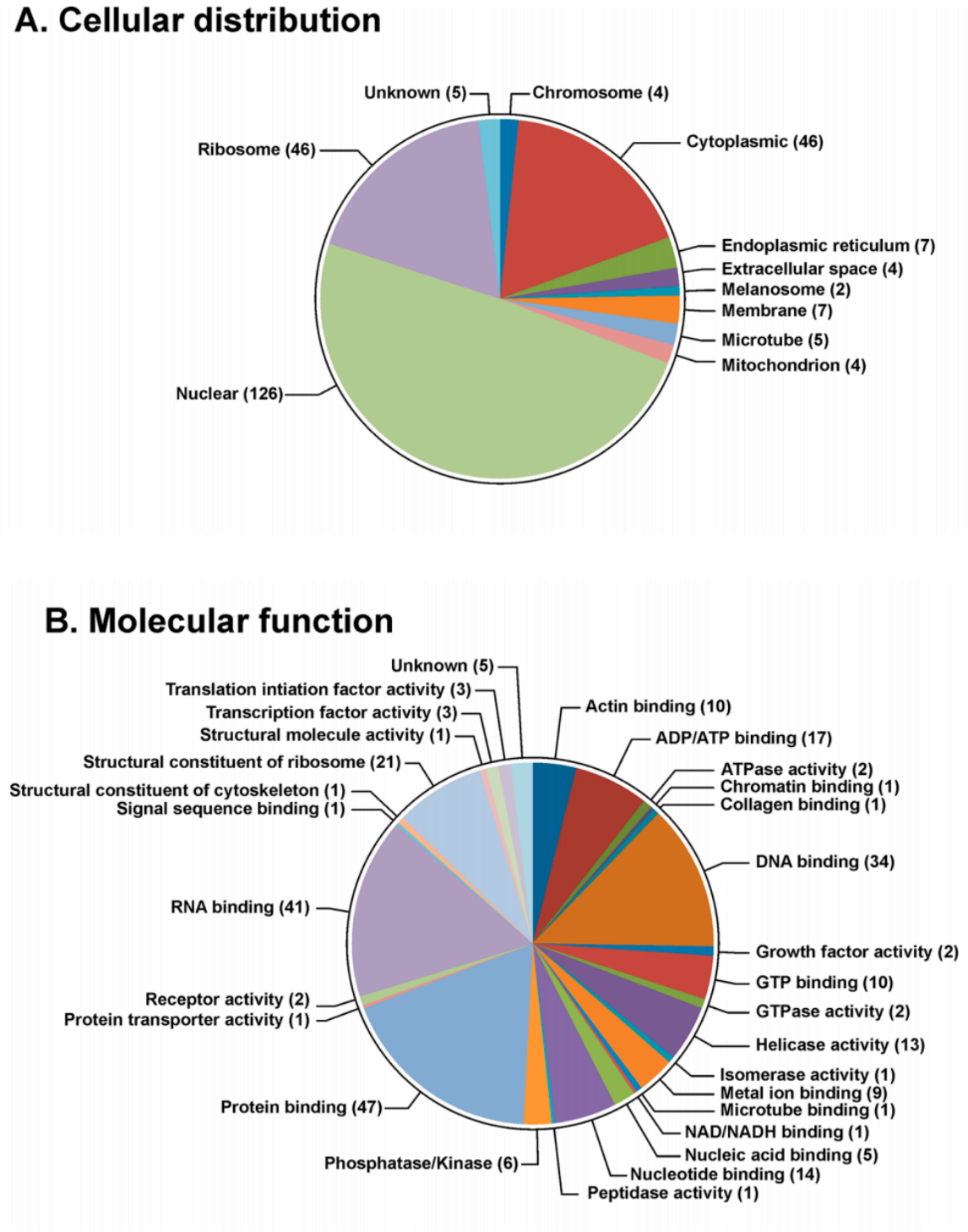


Figure 1.6 Distribution and Function of Proteins Found in DNA-Protein Crosslinks. Adapted from Ming *et al.* (87)**.**

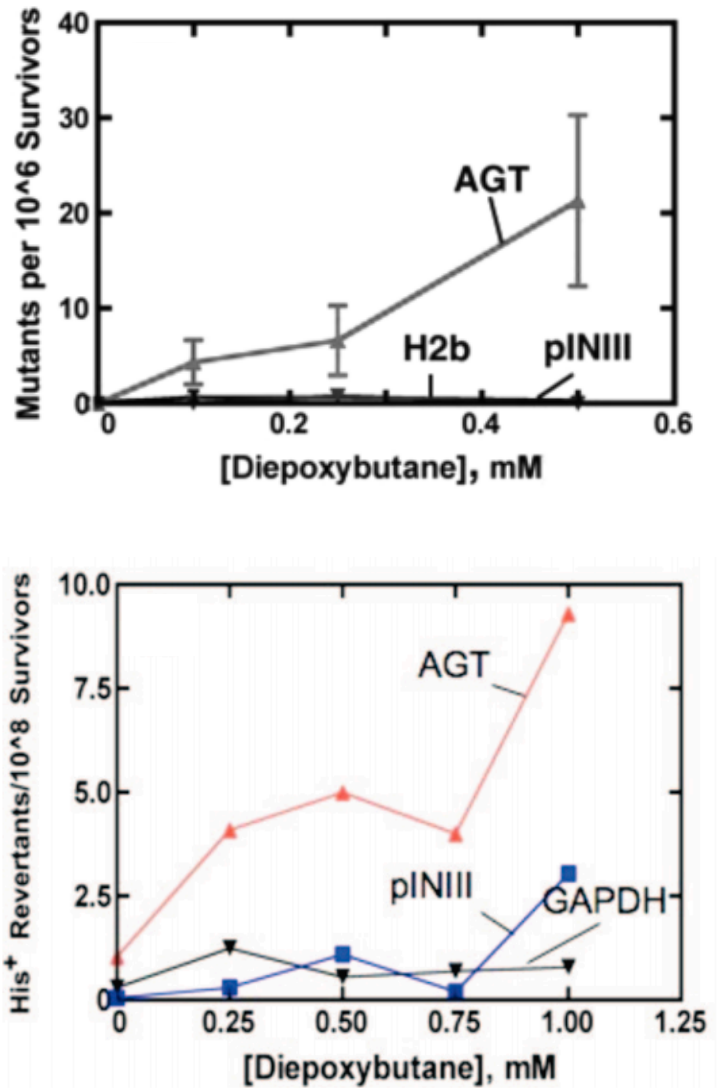


Figure 1.7 Co-treatment of the Crosslinking Agent Diexpoxybutane with Histone H2b and GAPDH Does Not Increase Mutation Frequency. Bacteria treated with the crosslinking agent, diepoxybutane, while expressing either histone H2b (top) or GADPH (bottom) do not incur the significant increase in DNA mutations as compared to when expressing AGT in the cells. Adapted from Loecken *et al.* and Loecken *et al.* (94,95).

One major limitation to the NER pathway is that it is unable to remove larger proteins (larger than 10 kDa) from DNA (100,101). Therefore, nuclear proteases are required to degrade the adducted protein on the DNA strand to a small enough size to allow NER to proceed.

One major challenge that can occur with DPCs is their effect when encountering a DNA replication fork. While one of the pathways outlined above could repair the lesion, these can be time-consuming or would not be accessible during replication. Therefore, it has been speculated that DPCs could be bypassed by the Y-family DNA polymerases during TLS. This would bypass the lesion, preventing a blocked replication fork and a stretch of single strand DNA. However, no TLS polymerase is capable of accommodating the great size of a crosslinked protein to allow direct bypass of the lesion. Therefore, for this pathway to work, a nuclear protease would be required to degrade the protein to a smaller (<10 kDa) peptide prior to TLS.

1.8 Degradation of DNA-Protein Crosslinks.

As described above, there has been evidence for the requirement of a nuclear protease to degrade DPCs for numerous DPC repair pathways as well as for potential TLS of the lesion (Figure 1.8). Without such a protease, DPCs would block cellular processes such as DNA synthesis and DNA repair, leading to severe DNA damage and even cell death. Because of this, there has been a recent interest in determining the identity and functionality of such a protease.

Recently the protease Wss1 was found in yeast and shown to degrade DPCs (102). Yeast cells that lacked the Wss1 protease accumulated DPCs after crosslinking treatment, leading to substantial chromosomal rearrangements and cell death. Purified Wss1 can cleave DNA binding proteins such as topoisomerases and histones but would not cleave non-DNA binding proteins, e.g. bovine serum albumin or GST, even in the presence of DNA. This was the first evidence of a nuclear protease that could degrade DPCs for either further repair processes or translesion DNA synthesis.

A human homologue of Wss1, SPRTN (Spartan, DVC1), was first discovered in 2012. The name SPRTN is derived from the putative zinc metalloprotease domain SprT located in the N-terminal part of the protein. However, it was first characterized by its C-terminal domain as a facilitator for p97 binding to stalled replication forks, and was thought to be involved in regulating DNA damage response (103). SPRTN is recruited to damaged DNA by ubiquitinated PCNA and RAD18, where it binds preferentially to ubiquitinated PCNA (104). It was also found to directly interact with DNA polymerases δ and η, leading to speculation about a function in either up- or down-regulating TLS (105,106). Further studies identified germline mutations in SPRTN as a cause of Ruijs-Aalfs syndrome (RJALS), which is characterized by DNA instability, premature aging, and early-onset hepatocellular carcinoma (107). Complete knockout of SPRTN is embryonic lethal in mice, and knockout studies in cells showed incomplete DNA replication, formation of micronuclei and chromatin bridges, and eventually cell death (108), as well as sensitivity to UV irradiation and various chemical crosslinking agents (109,110). SPRTN hypomorphic mice accumulated topoisomerase DPCs in their livers (111).

SPRTN contains 489 amino acids. The N-terminal side of the protein contains the putative zinc-dependent metalloprotease domain. Its C-terminal side contains regulatory and binding domains. The interaction domain for p97 (SHP domain) is located at residues 252-261. The PIP-box for its interaction with PCNA is located at residues 325-332. Located at the very end of the protein, SPRTN contains recognition module for ubiquitin (453-475). It has also been identified that there is a DNA binding pocket located between the Sprt domain and the SHP domain.

The role of SPRTN as a protease for degrading DPCs was reported in 2016 in by two different labs (110,112). Due to the effects of the cellular stress imposed when knocking down SPRTN and treating with crosslinking agents, it was speculated that SPRTN could have a role in the repair of crosslinks. Both research groups were able to show that SPRTN is a zinc-dependent metalloprotease and is dependent on DNA for its activity. They were able to show that SPRTN

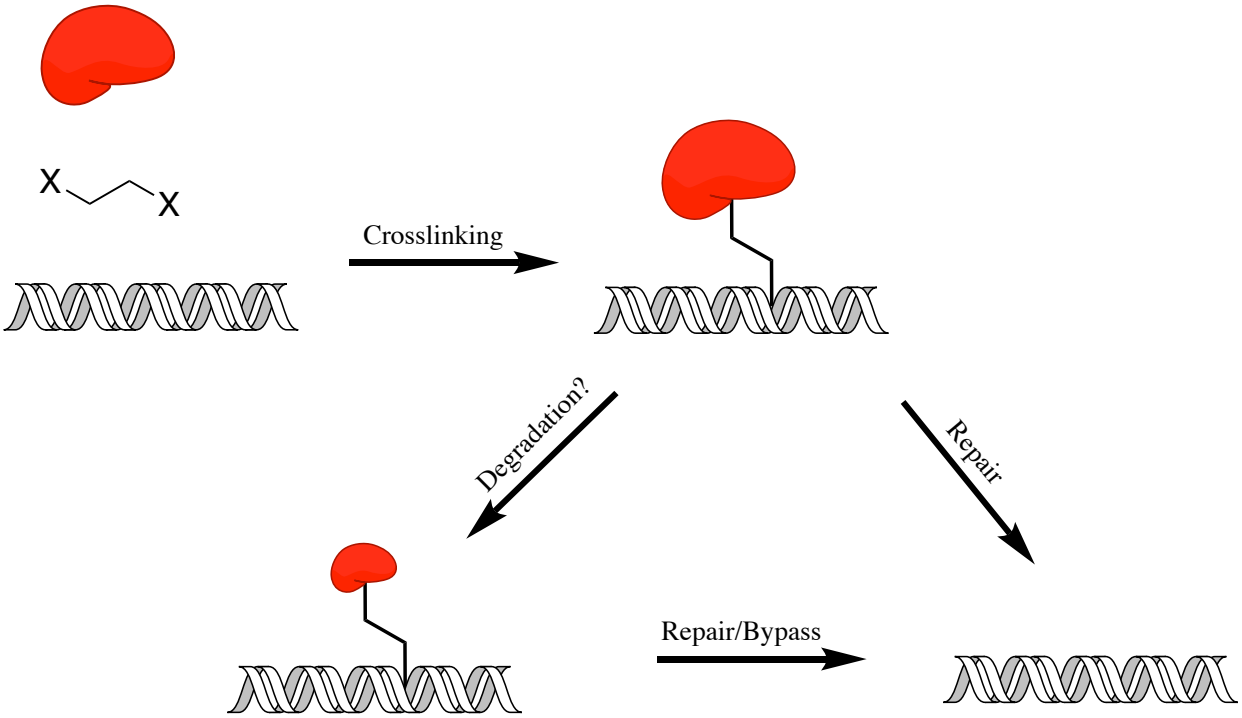


Figure 1.8 Pathways for the Formation, Degradation, Bypass, and Repair of DPCs

will cleave both itself and other DNA binding proteins in the presence of DNA but does not have any activity in the absence of DNA nor on non-DNA binding proteins (even in the presence of DNA). It was shown that SPRTN cleavage is dependent on the N-terminal SprT domain and that it is still active without any of the C-terminal binding motifs after its DNA binding area. SPRTN also has high binding affinity to DNA, with a K_d value of 100 nM. Its affinity for binding DNA is similar whether the DNA is single- or double-stranded or in a splayed duplex DNA. Its cleavage was also determined to be metal dependent as shown by its inhibition by 1,10-phenanthroline, a metal chelator. Point mutations of key residues in the SprT domain have also shown a loss in activity, notably E112Q and Y117C. The activity of SPRTN in its auto-cleavage was shown to occur in *trans*, where one molecule of SPRTN can only degrade another molecule of SPRTN and not itself.

Purified SPRTN can cleave topoisomerase, histones H2A, H2B, and H3, Fan1, HLTF, and Rad5. Interestingly, there have been conflicting results on the type of DNA required for this proteolysis. SPRTN autocleavage has been shown to occur in both ssDNA and dsDNA. However, the Boulton group concluded that SPRTN will only cleave other proteins in the presence of ssDNA but not dsDNA (112). Other groups including those of Ramadan and Haracska found that SPRTN will cleave these DNA binding proteins in either ssDNA or dsDNA (110,113).

1.9 Research Aims

The purpose of my dissertation is to understand the ability of Y-family DNA polymerases to bypass DPCs. Previous work had shown that crosslinks formed via adduction with GSH and AGT are both cytotoxic and mutagenic. My primary focus is to determine the mechanism for these adverse effects and to determine if the mutagenesis occurs due to TLS misinserting nucleotides opposite or nearby the site of the lesion. I am fascinated by the fact that these select adducts have harmful effects but that crosslinks formed with either GAPDH or histone H2b do not. Due to the fact that AGT is presumed to be too large a protein to be directly bypassed by any DNA

polymerase, another goal of this dissertation is to study the ability of the protease, SPRTN to degrade AGT protein crosslinks on DNA to allow for direct bypass by a translesion DNA polymerase.

Detailed materials and methods for the research I performed is described in Chapter II.

Chapter III describes the effects of GSH-crosslinks on the N^6 position of adenosine. This lesion was identified on DNA and its relative quantitation compared to the other sites of adduction was calculated. A DNA oligonucleotide was synthesized containing a site-specific GSH crosslink on the N^6 position of adenosine. Steady-state kinetics were performed to determine the changes in performance by various DNA polymerases as well as their misincorporation frequencies. LC-MS DNA sequencing of the bypass products was also tested to identify frameshift mutations.

Chapter IV describes the characterization of AGT-DNA crosslinks. DNA oligonucleotides were synthesized containing site-specific AGT peptide crosslinks. These oligonucleotides were then utilized for *in* vitro bypass assays and kinetic analysis. SPRTN protease was expressed and purified to degrade AGT-DNA crosslinks, which allowed translesion DNA synthesis to occur.

Chapter V contains my concluding remarks, discussion, and future directions for this research.

CHAPTER 2

Materials and Methods

2.1 Materials

2.1.1 Enzymes

Human DNA Pol η, ι, and κ (catalytic core—hPolη: 1-432 amino acids, hPolι: 1-420 amino acids, hPol κ: 19-526), *Sulfolobus solfataricus* DNA polymerase Dpo4, bacteriophage T7 DNA polymerase, and *Escherichia coli* Klenow fragment were expressed and purified as described previously (42,114-118). Uracil-DNA glycosylase (UDG) was purchased from New England Biolabs (Ipswich, MA). Alkaline phosphatase was purchased from Promega Corporation (Madison, WI). DNase I, Benzonase®, and phosphodiesterase I were purchased from Sigma Aldrich (St. Louis, MO). Tobacco etch virus (TEV) protease and the gene construct for SPRTN were purchased from Genscript (Piscataway, NJ).

2.1.2 DNA Substrates

6-Chloropurine deoxyribonucleoside was obtained from Alfa Aesar (Haverhill, MA). Unmodified oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA) and were HPLC-purified by the manufacturer. 6-Chloropurine phosphoramidite was obtained from Prof. Carmelo Rizzo (Vanderbilt University, Nashville, TN). Oligonucleotides containing a 6-chloropurine were synthesized on a Perspective Biosystems model 8909 DNA synthesizer on a 1-μmol scale using Expedite reagents (Glen Research, Sterling, VA), utilizing a standard synthetic protocol, and were purified by HPLC using a Phenomenex Alumina RP octadecylsilane (C_{18}) column (250 mm $\times$ 4.6 mm, 5 μm). Buffers consisted of Mobile Phase A (0.10 M NH_4HCO_2) and Mobile Phase B (CH_3CN). The following gradient program (v/v) was used with a flow rate of 1.5 mL/min: starting point 0% B, increased to 10% B over 15 minutes, then increased to 20% B at 20 min, held at 20% B for 5 min, increased to 80% B for 3 minutes, held at 80% B for 4 minutes,

and re-equilibrated for 5 min at 0% B (all v/v). The UV detector was set at 240 nm. All other nucleotides used for replication assays or SPRTN assays were purchased from either Integrated DNA Technologies (Skokie, IL) or Midland Certified Reagent Company (Midland, TX).

2.1.3 Other Materials

Isopropyl-*β*-D-1-thiogalactopyranoside (IPTG) was purchased from Anatrace (Maumee, OH). GSH, *N*-(*tert*-butyloxycarbonyl) (BOC)-protected 2-bromoethylamine, 1-bromo-2-chloroethane, Mesitylene sulfonyl chloride, and other chemical reagents were purchased from Sigma-Aldrich. Synthetic peptides from the AGT active site (PCH, LIPCHRV, and PVPILIPCHRVVSSS) were purchased from New England Peptides (Gardiner, MA), with the N-terminus protected using an acetyl group and the C-terminus protected using an amide group. The pNIC-ZB expression vector was purchased from Addgene (Cambridge, MA).

2.2 Methods

2.2.1 Chemical Synthesis

2.2.1.1 Synthesis of S-(2-Aminoethyl)GSH

GSH (50 mg, 163 *μ*mol) was dissolved in 10 mL of CH_3OH along with 20 mg (0.86 mmol) of Na°. A solution of *N*-BOC-protected 2-bromoethylamine (72.2 mg, 322 *μ*mol), in 5 mL CH_3OH, was added dropwise. The reaction mixture was stirred for 24 h before quenching with 0.7 mL glacial CH_3CO_2H. The reaction product was dried under a stream of N_2. The resulting white solid was redissolved in H_2O before purification by HPLC (mobile phase A: 0.01% CH_3CO_2H in H_2O, mobile phase B: 95% CH_3CN in H_2O with 0.01% CH_3CO_2H, all v/v) using a semi-preperative Beckman Ultrasphere RP octadecylsilane (C18) column (250 mm × 10 mm, 5 *μ*m). The following gradient program (v/v) was used with a flow rate of 3 mL/min: starting point 2% B, increased to 10% B over five minutes, then increased to 30% B at 10 min, held at 30% B for 5 min, and re-equilibrated for 5 min at 2% B (all v/v). The UV detector was set at 240 nm. The *N*-BOC-protected *S*-(2-aminoethyl)GSH was collected and its structure was confirmed by positive ion ESI-MS/MS.

The compound was then redissolved in a mixture of 2.5 mL of CH_2Cl_2 and 0.3 mL CF_3CO_2H and stirred for 1 h to remove the N-BOC. The reaction mixture was dried under a stream of N_2 before repurification by HPLC using the same protocol as above.

2.2.1.2 Synthesis of S-[2-(N^6-Deoxyadenosinyl)ethyl]GSH

S-(2-Aminoethyl)GSH (180 *μ*mol) was dissolved in a mixture of dimethyl sulfoxide (DMSO) (225 *μ*L) and *N,N*-diisopropylethylamine (DIPEA) (25 *μ*L) and added to dry 6-chloropurine deoxyribonucleoside (1 *μ*mol). Following the incubation at 60 °C for 24 h, the reaction mixture was purified by HPLC as shown above, identified by ^{1}H-NMR (Figure A.1), and quantified by UV absorbance at 260 nm (119).

2.2.1.3 Synthesis of S-[2-(N^6-Deoxyadenosinyl)ethyl]GSH-Containing Oligonucleotide

S-(2-Aminoethyl)GSH (750 nmol) was dissolved in a mixture of $(CH_3)_2SO$ (100 *μ*L) and DIPEA (25 *μ*L) and added to a dry 6-chloropurine-containing oligonucleotide (5′-TCTCXGTTTATGGACCACC-3′, where X is 6-chloropurine) (15 nmol) (Figure 2.1). Following incubation at 60 °C for 24 h, the reaction product was purified by 15% (w/v) polyacrylamide gel electrophoresis, visualized using a UV lamp (to locate bands), cut out of the gel, and extracted using a solution of 200 mM NaCl/1 mM EDTA (pH 8.0); the identity was confirmed by negative ion ESI-MS (Figure A.2).

2.2.1.4 Alkylation of Calf Thymus DNA

S-(2-Chloroethyl)GSH was synthesized by dissolving 0.5 g (1.6 mmol) of GSH in 12 mL of dry CH_3OH in which 0.12 g of Na° had been dissolved (120). A solution of 2.3 g (16 mmol) of 1-bromo-2-chloroethane in 12 mL of CH_3OH was then added dropwise to the GSH solution and stirred for 1 h before quenching with 0.25 mL glacial CH_3CO_2H. The resulting precipitate (*S*-(2-chloroethyl)GSH) was collected by centrifugation (2000 × g, 10 min). It was redissolved in 0.5 M Tris-HCl buffer (pH 7.7) containing 10% CH_3OH (v/v) to a final concentration of 10, 50, or 250 mM and immediately added to a solution of calf thymus DNA (2 mg/mL) in the same buffer (final

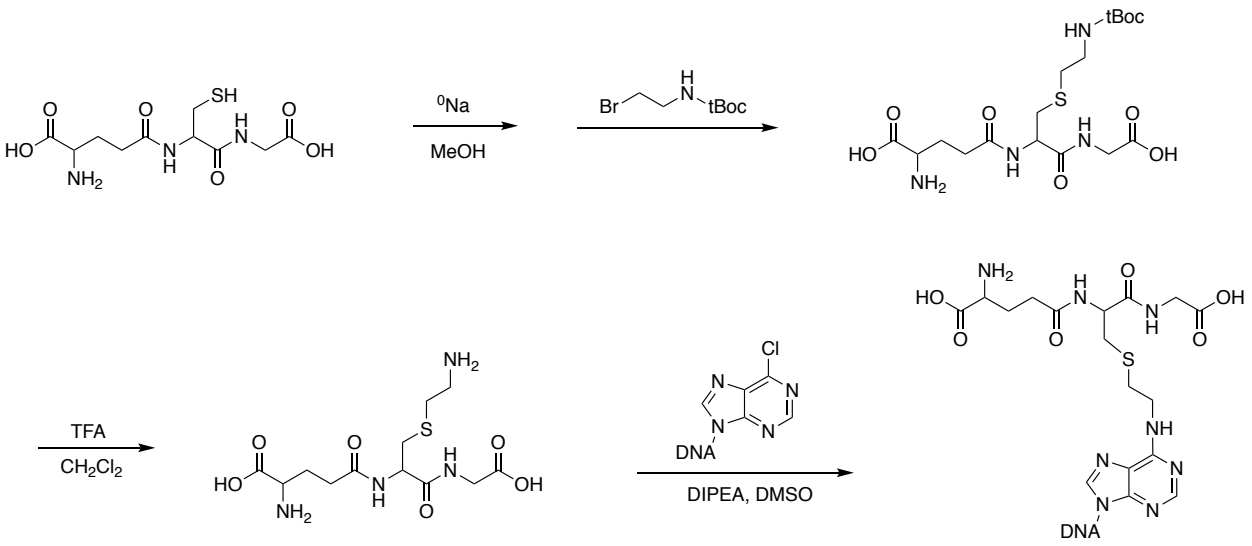


Figure 2.1 Reaction Scheme for the Synthesis of *S*-[2-(*N*^{6}-Deoxyadenosinyl)ethyl]GSH-Containing Oligonucleotide

volume for the reaction was 250 μL). The reaction was incubated at 37 °C for 30 minutes. After the incubation the DNA was precipitated by the addition of 2.5 volumes of cold C_2H_5OH and digested using a combination of nucleases consisting of phosphodiesterase I from *Crotalus adamanteus* venom (0.02 units), bovine pancreatic DNase 1 (20 units), and calf intestinal alkaline phosphatase (5 units) at 37 °C for 24 h. The resulting solution was filtered through a Microcon 10K Centrifugal filter before drying using centrifugal lyophilization. The samples were redissolved in 100 μL of running buffer (0.1% HCO_2H in H_2O, v/v) and analyzed by UPLC-MS. The solvents used were Mobile Phase A (0.1% HCO_2H in H_2O, v/v) and Mobile Phase B (0.1 % HCO_2H in 95% CH_3CN in H_2O, v/v). The following gradient was used at 0.3 mL/min: start at 0% B, hold for 5 min, increased to 30% B at 13 min, then increased to 95% B at 15 min, and maintained there for 2 min before re-equilibrating to 0% B at 19 min (all v/v). The temperature of the column was 40 °C. Product ion spectra were collected over the *m/z* range of 100-1000, and MS/MS fragmentation spectra were collected over the *m/z* range of 160-700. The MS conditions were as follows: capillary temperature, 350 °C; source voltage, 5 kV; capillary voltage, 13 V; tube lens voltage, 60 V. MS fragmentation analysis was performed with a collision energy of 35%, *m/z* isolation width of 2, activation *Q* of 0.25, and an activation time of 30 ms.

2.2.1.5 Synthesis of O-(Mesitylsulfonyl)hydroxylamine

Ethyl hydroxamate (23 mmol) was stirred in 9 mL of a 2:1 mixture of dimethylformamide (DMF) and triethylamine and cooled to 0 °C. Mesitylene sulfonyl chloride (23 mmol) was slowly added and stirred for an additional 15 min before diluting in 100 mL dichloromethane. The reaction product was washed ten times with H_2O, and the organic layer was collected and dried using brine and $MgSO_4$ before removing the solvent by vacuum. The reaction product was redissolved in 4 mL dioxane and cooled to 0 °C before adding 1.8 mL of 70% perchloric acid dropwise over 2 min. The solution was stirred for an additional 2 min and transferred into 200 mL of chilled H_2O prior to extracting with 100 mL diethyl either. The organic layer was neutralized and dried using KCO_3 and filtered. The filtrate was concentrated and added to 150 mL of chilled hexane to

crystalize the product overnight at -20 °C. The *O*-(mesitylsulfonyl)hydroxylamine (MSH) crystals were collected, dried under vacuum, and stored at -20 °C.

2.2.1.5 Synthesis of Dehydroalanine Peptides from AGT

Synthetic peptides from the active site of AGT (n = 3, 7, and 15 aa; 10 μmol) were dissolved in 200 μL H_2O and treated with 2.1 equivalents of MSH and 10 equivalents of potassium carbonate dissolved in 200 μL DMF at 0 °C for 20 min to form dehydroalanine-containing peptides. Peptides were purified by HPLC using the same protocol as above.

2.2.1.6 Synthesis of S-[2-(N^6-Deoxyadenosinyl)ethyl]AGT Peptide-Containing Oligonucleotides

The 6-chloropurine-containing oligonucleotide (characterized above) (15 nmol) was incubated with cystamine dihydrochloride (5 μmol) in 100 μL of 150 mM DIPEA in H_2O at 50 °C overnight. The reaction product was desalted and dried by lyophilization before resuspending it in 50 μL H_2O. The oligonucleotide was further reduced using 1 mM of either DTT or TCEP in a solution containing 50 mM Na_2PO_4 (pH 8.0) and 100 mM NaCl at 50 °C for 30 min. The oligonucleotide was reacted with 250 nmol of the dehydroalanine-containing peptide at 37 °C for 48 h. The product was desalted using a solid phase extraction column (Oasis HLB, Waters), and its identity was confirmed by negative ion ESI-MS.

2.2.1.7 Synthesis of AGT Protein DNA Crosslink

The 6-chloropurine oligonucleotide (characterized above) (15 nmol) was incubated with propargylamine (5 μmol) in a mixture of 1 M DIPEA in dioxane (25 μL) and H_2O (75 μL) at 55 °C for 24 h before desalting using a solid phase extraction column (Oasis HLB, Waters) and purifying by HPLC (Figure 2.2). The 12mer primer strand of DNA was then annealed to DNA in 50 mM HEPES (pH 7.4) with 50 mM NaCl at 95 °C for 5 min. The DNA duplex was modified using a copper-catalyzed azide-alkyne cycloaddition reaction for 4 h in the dark by the addition of 2 mM azide linker with an iodoacetamide attached to its terminus (Figure 2.2). The DNA duplex was then incubated with recombinant AGT for 3 h in the dark to allow the active site thiol of AGT to crosslink to the DNA duplex at the iodoacetamide position. The AGT crosslink was purified by

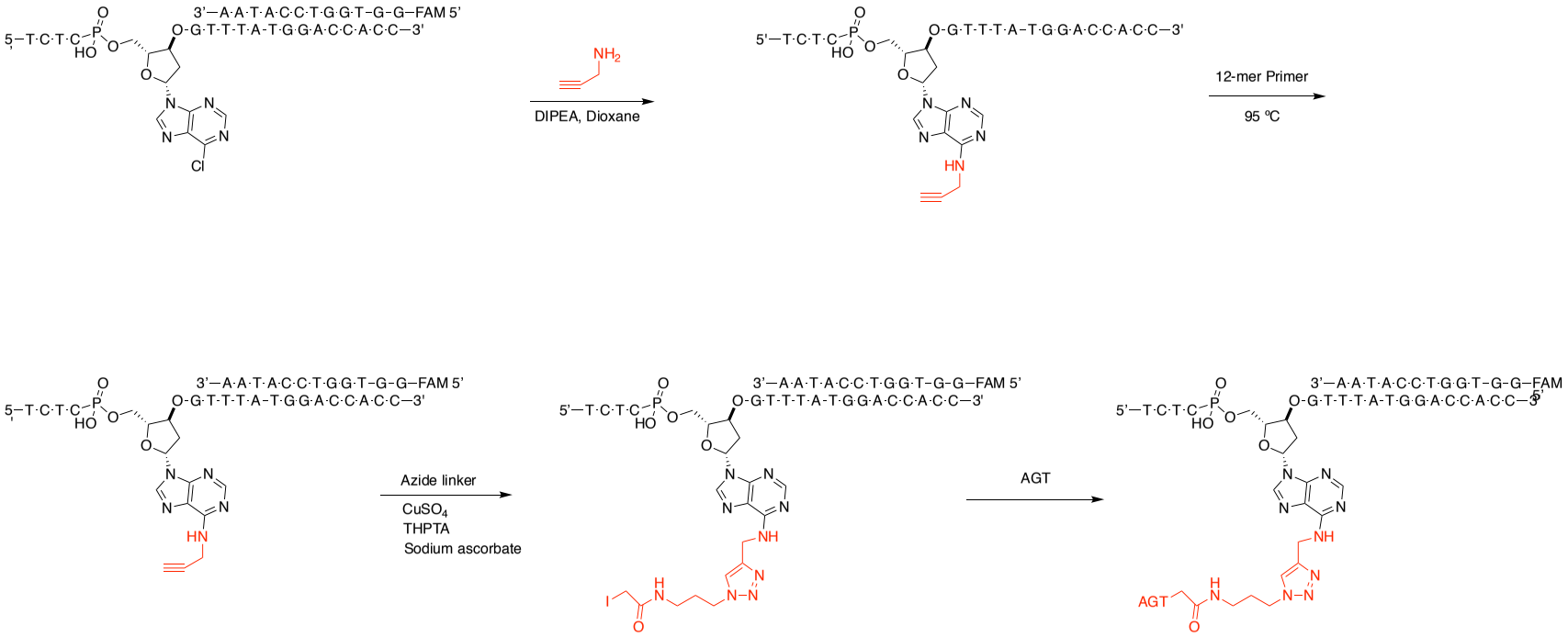

3'—A·A·T·A·C·C·T·G·G·T–G–G–FAM 5'
5'—T·C·T·C
O-G·T-T-T-A-T·G·G·A·C·C·A·C·C—3'
HO
Cl
NH2
DIPEA, Dioxane
HN
12-mer Primer
95 ºC
Azide linker
CuSO4
THPTA
Sodium ascorbate
NH
N–N
I
AGT

Figure 2.2 Synthesis scheme for AGT DNA crosslink. DNA oligonucleotide containing a 6-chloropurine was treated with propargyl amine before annealing the primer strand of DNA. The crosslinking agent was then attached to the oligonucleotide utilizling click chemistry before reacting with the AGT protein.

FPLC using a Ni^{2+}-nitrilotriacetic acid (NTA) or Co^{2+}-NTA based IMAC column (GE Healthcare) and analyzed by gel electrophoresis and proteomic analysis.

2.2.2 DNA Replication Studies

6-Carboxyfluorescein (FAM)-Labeled Primer Extension and Steady-State Kinetic Assays. A 12-mer oligomer (FAM-5´-GGTGGTCCATAA-3´, for full primer extension in the presence of four dNTPs) and a 14-mer oligomer (FAM-5´-GGTGGTCCATAAAC-3´, for single base primer extension and steady-state kinetics in the presence of single dNTPs) were annealed to the 19-mer oligomer template (5´-TCTCXGTTTATGGACCACC-3´, where X is dA or *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH). Primer extension was performed with 120 nM oligonucleotide complex in 50 mM Tris-HCl buffer (pH 7.5) containing 50 mM NaCl, 5 mM $MgCl_2$, 500 μM dNTPs, 2% (v/v) glycerol, 50 μg/mL bovine serum albumin (BSA), and 20 nM polymerase (with the exception of hPol ι, 40 nM) at 37 °C. Steady-state kinetics were performed under the same conditions except using polymerase concentrations from 1-40 nM, varying dNTP concentrations (0.1-500 μM), and incubation times from 5-30 min so that the maximum incorporation was < 20% of the substrate concentration. Reactions were quenched with 7 μL of 20 mM EDTA (pH 9.0) in 95% (v/v) formamide. Products were separated using 18% (w/v) polyacrylamide gel electrophoresis and visualized using a Typhoon imaging system (GE Healthcare).

LC-MS Primer Extension Assays. A 14-mer oligomer (FAM-5´-GGTGGTCCATAA(dU)C-3´) was annealed to the 19-mer oligomers as described above. Full-length primer extension was performed with 2.5 μM oligonucleotide complex in 50 mM Tris-HCl buffer (pH 7.5) containing 50 mM NaCl, 5 mM $MgCl_2$, 1 mM dNTPs, 2% (v/v) glycerol, 50 μg/mL BSA, and 5 μM hPol η at 37 °C for 1 h. The reactions were stopped by spin column filtration to remove the salts and buffer before redissolving in UDG stock buffer (1 mM DTT and 1 mM EDTA in 20 mM Tris-HCl, pH 8.0) and treating with 25 units of UDG for 37 °C for 4 h, then adding 0.25 M piperdine and heating at 95 °C for 1 h. To examine reaction products, samples were analyzed by LC-MS/MS, performed with a Waters Acquity UPLC system (Waters, Milford, MA) interfaced to a Thermo-Finnigan LTQ

mass spectrometer (Thermo Scientific Corp., San Jose, CA) equipped with an ESI source. Chromatographic separation was achieved using a Waters Acquity UPLC BEH octadecylsilane (C_{18}) column (1.0 mm × 10 mm, 1.7 μm) at a flow rate of 0.3 mL/min. Mobile Phase A consisted of 10 mM $NH_4CH_3CO_2$, and Mobile Phase B was 10 mM $NH_4CH_3CO_2$ in 95% CH_3CN in H_2O (v/v). The gradient started at 2% B, increased to 10% B at 5 min, 30% B at 9 min, and was held there for 2 min before returned to 2% B for 2 minutes (all v/v). The column was maintained at 50 °C. The MS conditions were as follows: capillary temperature, 350 °C; source voltage, 4kV; source current, 100 μA; capillary voltage, -35 V; tube lens voltage, -93 V. MS fragmentation analysis was performed using data dependent scanning with a collision energy of 35%, *m/z* isolation width of 2, activation *Q* of 0.25, and an activation time of 30 ms. Product ion spectra were taken over an *m/z* range of 300-2000, and the most abundant species (-2 charge) was used for collision-induced dissociation (CID).

2.2.3 Protein Expression

2.2.3.1 SPRTN

SPRTN was expressed and purified as described as previously reported (110). The SPRTN gene construct was assembled by Genscript, containing the full length SPRTN-WT codon-optimized for bacterial expression, an N-terminal Hisx6- and Z-basic- tag, and a C-terminal Hisx6-tag. The gene was inserted into pNIC-ZB expression vector, and 5 μL of the vector was transformed into BL21 ArcticExpress (DE3) competent *E. coli* cells. The transformed cells were plated on lysogeny broth (LB) agar plates supplemented with 50 μg/mL kanamycin to generate bacterial colonies. Isolated colonies were selected and subjected to sequence analysis to confirm the incorporation of the SPRTN gene and used to inoculate 100 mL of LB medium containing kanamycin (50 μg/mL) that was incubated at 37 °C with shaking at 250 rpm overnight. Expression cultures were grown by adding 10 mL of the overnight culture into 1 L of Terrific Broth medium containing kanamycin, and the protein was expressed by inducing the expression vector with 0.5 μM IPTG at 20 °C for 18 h. The expression culture was centrifuged at 5000 x g for 10 min to pellet

the bacteria. The supernatant was removed, and the bacterial pellet was resuspended in 250 mL of 100 mM HEPES buffer (pH 7.4) containing 500 mM NaCl, 1 mM $MgCl_2$, 10 mM imidazole, 10% glycerol (v/v), 1 mM DTT, 0.1% Triton X-100 (v/v), and a protease inhibitor mixture (cOmplete™, EDTA-free, Roche Applied Science). The cells were lysed by sonication until the solution was homologous, and 1 unit of Benzonase® was added to the cell lysate. The lysate was centrifuged to remove the cellular debris at 35,000 x g for 30 min. The supernatant was collected and loaded onto a 5-mL nickel-nitrilotriacetic acid (Ni-NTA) FPLC column (HisTrap HP, GE Healthcare) that has been equipped with both Buffer A (100 mM HEPES (pH 7.4) containing 500 mM NaCl, 10 mM imidazole, 10% glycerol (v/v), and 1 mM DTT) and Buffer B (100 mM HEPES (pH 7.4) containing 500 mM NaCl, 300 mM imidazole, 10% glycerol (v/v), and 1 mM DTT) and kept at 4 °C. The column was equilibrated with 7% of Buffer B at a flow rate of 0.5 mL/min prior to loading the protein. The bound protein was washed with two column volumes of 20% Buffer B before performing a linear gradient to 100% Buffer B over 10 column volumes (Figure A.4). The purified enzyme was then dialyzed against 100-fold volumes of 20 mM HEPES (pH 7.4) containing 500 mM NaCl and 10% glycerol (v/v). The Z-basic tag was removed from the protein by treating with 50 units of TEV protease at 4 °C for 72 h. The protein was then concentrated using a centrifugal filter with a 3 kDa molecular weight cut-off filter (Amicon® Ultra 4, Millipore Sigma) to a volume of ~500 μL. This solution was then loaded onto a size-exclusion column (Superdex 75 10/300, GE Healthcare) using FPLC to further purify the protein away from the TEV protease and the cleaved Z-basic tag. The column was equilibrated with 20 mM HEPES buffer (pH 7.4) containing 500 mM NaCl and 10% glycerol (v/v) and the protein was eluted with a flow rate of 0.5 mL/min (Figure A.5). The protein purity was confirmed by SDS-PAGE, and the sample was frozen at -80 °C in small aliquots until use.

2.2.3.2 AGT

Recombinant human AGT with a C-terminal Hisx6 tag was expressed in *E. coli* and purified by metal affinity chromatography using a protocol adapted from a previous report (93).

Briefly, an *E. coli* codon-optimized cDNA corresponding to the sequence reported by Pegg and associates (121) of the human wild type protein was transformed into XL-1 Blue cells and plated on LB agar containing 100 μg/mL ampicillin. Isolated colonies were used to inoculate overnight cultures consisting of 100 mL LB medium containing 100 μg/mL ampicillin, shaking at 250 rpm at 37 °C for 12-15 h. Four-liter flasks containing 1 L of Terrific Broth medium containing ampicillin were inoculated with 10 mL aliquots of the overnight culture and incubated at 37 °C for ~4 h until reaching an OD_{600} of ~0.6. Expression of the protein was induced by supplementing the cultures with 0.3 mM IPTG and incubated at 37 °C for 24 h with shaking at 220 rpm. The expression culture was centrifuged at 5000 x g for 10 min to pellet the bacteria. The supernatant was removed, and the bacterial pellet was resuspended in 250 mL of 20 mM Tris-HCl buffer (pH 7.4) containing 500 mM NaCl, 5 mM imidazole, 5% glycerol (v/v), 1 mM DTT, 0.1% Triton X-100 (v/v), and a protease inhibitor mixture (cOmplete™, EDTA-free, Roche Applied Science). The cells were lysed by sonication until the solution was homologous. The lysate was centrifuged to remove the cellular debris at 35,000 x g for 30 min. The supernatant was collected and loaded onto a 1-mL nickel-nitrilotriacetic acid (Ni-NTA) FPLC column (HisTrap HP, GE Healthcare) that had been equilibrated with both Buffer A (20 mM Tris-HCl (pH 7.4) containing 500 mM NaCl, 5 mM imidazole, 5% glycerol (v/v), and 1 mM DTT) and Buffer B (20 mM Tris-HCl (pH 7.4) containing 500 mM NaCl, 100 mM imidazole, 5% glycerol (v/v), and 1 mM DTT) and kept at 4 °C. The column was equilibrated with 7% Buffer B at 0.5 mL/min prior to loading the protein. The bound protein was washed with two column volumes of 20% Buffer B before performing a linear gradient to 100% Buffer B over 10 column volumes. The purified enzyme was then dialyzed against 100-fold volumes of 20 mM Tris-HCl (pH 7.4) containing 500 mM NaCl and 5% glycerol (v/v) prior to storing at -80 °C.

2.2.4 SPRTN Assays

A custom DNA oligonucleotide was designed for SPRTN protease activity assays (Figure A.3). The oligonucleotide was 50 nucleotides in length with a G:C content of 50%. Both ss- and ds-DNA were used in these assays.

2.2.4.1 SPRTN Protease Assays

Protease activity assays were performed with 1-6 μM SPRTN using 10 μM DNA as a modulator. Assays were performed either with other proteins (10 μM) or with SPRTN alone. Incubations were done in the presence of 25 mM Tris-HCl buffer (pH 7.4) containing 200 μM $ZnCl_2$. Incubations were performed with ss-DNA, ds-DNA, or no DNA. The reactions were done at 37 °C for up to 24 h prior to quenching 12 μL aliquots using 4 μL of NuPAGE™ LDS Sample Buffer (ThermoFisher Scientific) and heating to 95 °C for 5 min. Samples were loaded onto a denaturing SDS protein gel to resolve the reaction products using a 4-12% gradient gel and 2-(*N*-morpholino)ethanesulfonic acid (MES) buffer (pH 7.3). The proteins were separated at 150 V for 1 h before removing the gel and imaging using either Coomassie SimplyBlue™ or SYPRO™ Ruby Protein Gel Stains (ThermoFisher Scientific).

2.2.4.2 AGT-crosslink Degradation

AGT-DNA crosslinks were formed by incubating 20 μM DNA with 15 μM recombinant AGT and 10 mM DBE in 500 μL of 50 mM potassium phosphate buffer (pH 7.6) at 37 °C for 3 h. The excess DBE was removed by running the sample through a 3 kDa molecular weight cut-off filter (Amicon® Ultra 4, Millipore Sigma) and washing with excess amounts of 50 mM potassium phosphate buffer (pH 7.6). The reaction products were analyzed by SDS-PAGE to confirm the presence of a DPC. The crosslink was further incubated with 2 μM SPRTN to degrade the AGT protein crosslinks over a period of 24 h to monitor the degradation of the DPC band. Samples of the reaction mixture (25 μL) were removed at selected time points and treated at 55 °C for 5 min to heat inactivate SPRTN. The reaction was first monitored by loading 5 μL (from each timepoint) on a 4-12% denaturing protein gel and imaging using SYBR Safe™ (ThermoFisher Scientific) to

image the DNA and Coomassie SimplyBlue™ (ThermoFisher Scientific) to image the protein bands.

The remaining 20 μL from each sample was used for proteomic analysis, which was first reduced using 10 mM DTT at 37 °C for 30 min. The samples were then treated with 25 mM iodoacetamide and incubated in the dark on a shaker at room temperature for 30 min to alkylate all free thiols. The samples were further blocked by the addition of 20 mM DTT at 37 °C for 10 min. The samples were then split into two even aliquots. The first aliquot was used to purify the DNA from the sample using Fe^{3+}-IMAC columns. This was first diluted into 300 μL of 0.1 M acetic acid and then loaded onto Fe^{3+}-IMAC columns conditioned using PHOS-Select™ Fe^{3+} chelate matrix (Sigma Aldrich) in SigmaPrep™ Spin Columns (Sigma Aldrich). The columns were incubated for 30 min with mixing to bind the DNA to the matrix. After incubation, the columns were centrifuged at 8,200 x g for 30 s to remove unbound peptides/proteins. The columns were washed with 500 μL of 250 mM acetic acid with 30% CH_3CN (v/v) and 500 μL H_2O with centrifugation prior to adding 500 μL of 250 mM NH_4OH. The columns were incubated for 5 min, with shaking, prior to eluting the DNA from the columns by centrifugation for 1 min. Both groups of aliquots were then dried using vacuum centrifugation prior to suspending in 48% hydrofluoric acid (v/v) and incubating at 4 °C for 16 h. The samples were dried under a stream of nitrogen and resuspended twice in 50 μL of CH_3OH before drying again. The samples were redissolved in 20 μL of 0.1% HCO_2H in H_2O (v/v) and submitted for proteomic analysis.

CHAPTER 3†

Formation and Bypass of *S*-[2-(*N*^{6}-Deoxyadenosinyl)ethyl]glutathione in DNA by Translesion DNA Polymerases

3.1 Introduction

Halogenated hydrocarbons are widely used in commerce and industry, e.g. as solvents, pesticides, and propellants. 1,2-Dibromoethane (DBE, ethylene dibromide) has been historically used as an anti-knock additive in fuel and also has had uses as a pesticide in soil and in the organic synthesis of various compounds such as vinyl bromide. DBE is a known carcinogen (122,123) and had the highest hazard score for carcinogens in the Human Exposure/Rodent Potency Index prior to changes in the restrictions on the compound (76). DBE causes tumors in rats and mice in several different tissues, including nasal cavity, blood vessels, skin, lung, kidney, and liver (77-80).

DBE is not particularly reactive itself. It can be oxidized to 2-bromoacetaldehyde, but there is considerable evidence that the more important reaction in genotoxicity is conjugation with GSH, facilitated by GSH transferases (88,124-127). The resulting product *S*-(2-bromoethyl)GSH, a half-mustard, then reacts with DNA through an episulfonium ion intermediate (Figure 3.7) (89,90). *S*-(2-Chloroethyl)GSH can be synthesized and used as a substitute, generating the same products (120,128). Multiple DNA adducts arising from the GSH-DBE conjugate include those formed at the N7, N2, and O6 positions of 2´-deoxyguanosine (dG) as well as the N1 position of 2´-deoxyadenosine (dA) (99,129). Mutations of a transgene in mouse liver were attenuated by an inhibitor of GSH synthesis, indicating that these adducts strongly contribute to DBE genotoxicity *in vivo* (130).

† This chapter was adapted from Sedgeman, C. A., Su, Y., and Guengerich, F. P. (2017) Formation of S-[2-(N6-Deoxyadenosinyl)ethyl]glutathione in DNA and Replication Past the Adduct by Translesion DNA Polymerases. *Chem. Res. Toxicol.* 30, 1188-1196.

A similar chemical mechanism is involved in the DBE modification of the DNA repair protein O^6-alkylguanine DNA alkyltransferase (AGT, MGMT) (93,131-134) and other proteins that become crosslinked to DNA (94,95). After calf thymus DNA was treated with DBE in the presence of AGT, we identified adducts at the N7, N1, N2, and O6 positions of dG (92,134) and an N^6-dA adduct, which had not been seen in the earlier work with GSH. This adduct was reported at levels similar to those of the other non-labile adducts (the N^7-dG adduct is labile and can depurinate) (92,134).

In this work, we investigated the ability of DBE, and more specifically the GSH conjugate *S*-(2-chloroethyl)GSH [an analogue of *S*-(2-bromooethyl)GSH, with similar biological properties (120,128)] to react at the N6 position of deoxyadenosine to form *S*-[2-(N^6-deoxyadenosinyl)ethyl]GSH in calf thymus DNA. Extension studies were performed in duplex oligonucleotides containing the adduct by human translesion synthesis DNA polymerases (hPols) η, ι, and κ and with *S. solfataricus* DNA polymerase Dpo4, *E. coli* polymerase I Klenow fragment, and bacteriophage T7 polymerase.

3.2 Results

3.2.1 Synthesis of S-[2-(N^6-Deoxyadenosinyl)ethyl]GSH and an Oligonucleotide Containing the Lesion

S-[2-(N^6-Deoxyadenosinyl)ethyl]GSH was synthesized as a standard to quantitate the adduct in calf thymus DNA and characterized by positive ESI-MS/MS (Figure 3.2) and ^{1}H-NMR (Figure A.1). Additionally, an 19-mer oligonucleotide containing *S*-[2-(N^6- deoxyadenosinyl)ethyl]GSH was prepared for *in vitro* replication assays (Figure 3.1).

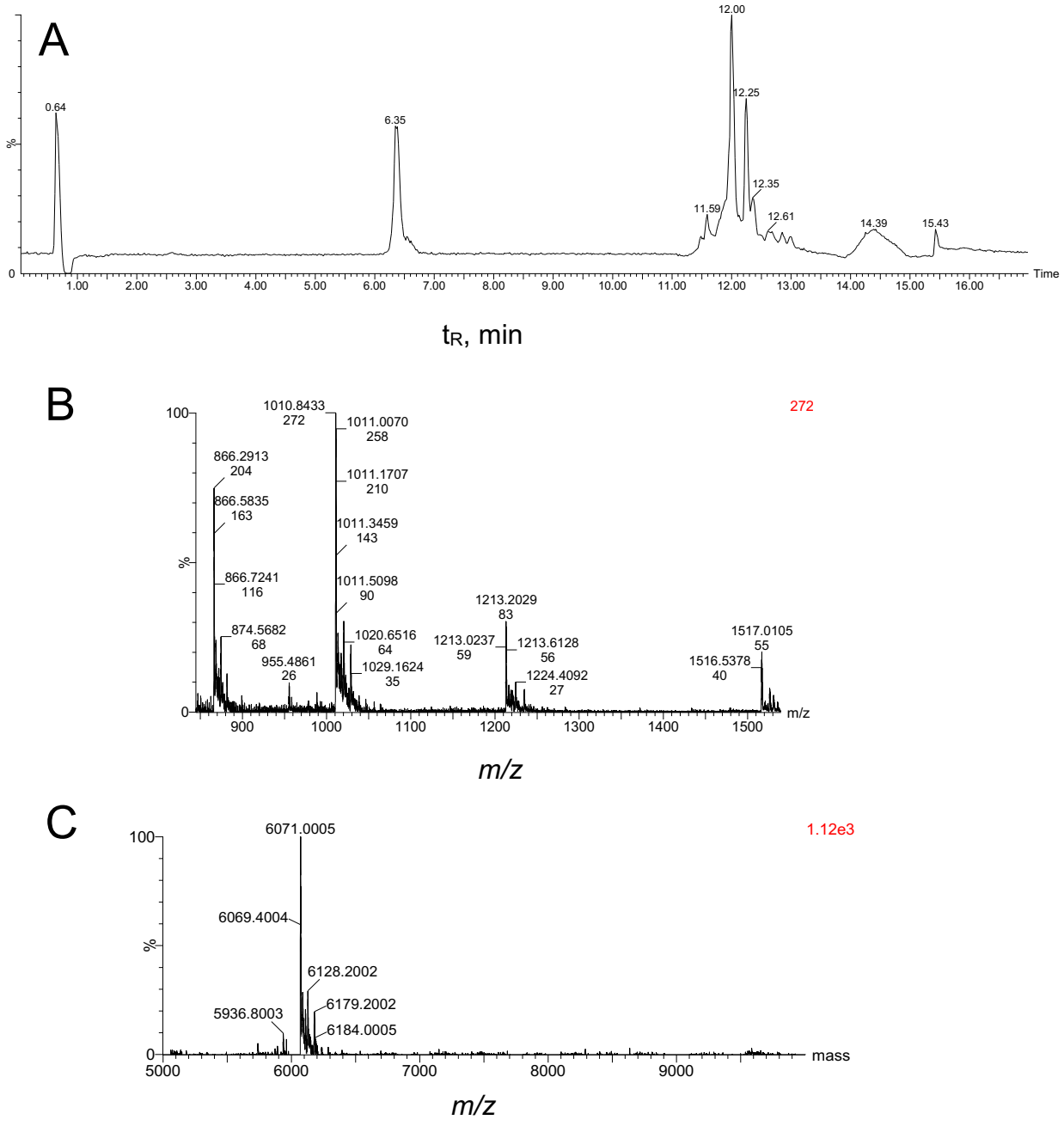


Figure 3.1 LC-MS Characterization of the *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH-containing Oligonucleotide. UPLC chromatogram (A), mass spectrum of product at 6.35 min (B), and deconvoluted mass spectrum of product (C).

GSH was treated with *N*-BOC-protected 2-bromoethylamine and deprotected to release the free amino group. The product was used for coupling with 6-chloropurine, either for the synthesis of a free nucleoside or a 19-mer oligonucleotide for the standard and replication studies, respectively. In the reaction, the amino group from 2-aminoethyl GSH displaced the 6-chloro group, resulting in the formation of dA with a thioethylene linkage to GSH. The site of attachment at the dA N6 atom is at the Cys-ethylamino nitrogen and not the Glu *α*-amino group, as revealed by the M-129 plus M-116 losses in the mass spectra; the former is indicative of a GSH adduct (Figure 3.1) (135). This is the same result seen with the synthesis of the N^2-dG adduct previously, established in a different way (136).

The *S*-[2-(N^6-deoxyadenosinyl)ethyl]GSH standard was purified by HPLC and characterized by positive ESI-MS. The oligonucleotide containing *S*-[2-(N^6-deoxyadenosinyl)ethyl]GSH was purified by gel electrophoresis and characterized by negative ESI-MS (Figure 3.2).

3.2.2 Formation of S-[2-(N^6-Deoxyadenosinyl)ethyl]GSH in DNA

Calf thymus DNA was treated with *S*-(2-chloroethyl)GSH at 37 °C for 30 min. The DNA was then extracted, digested with nucleases and alkaline phosphatase, and desalted prior to LC-MS analysis. A peak corresponding to the correct retention time and mass of the adduct and fragmenting to *m/z* 340 (loss of *γ*-glutamine plus loss of deoxyribose) was observed in the chromatogram (Figure 3.2A and C). Similar fragmentation patterns for the parent ion were found in both the standard and the reaction product (Figure 3.2B and D). The levels of the adduct were quantified in relation using the standard. Using a concentration of 250 mM *S*-(2-chloroethyl)GSH for the reaction, 2.5 *S*-[2-(N^6-deoxyadenosinyl)ethyl]GSH adducts/10^4 bases were detected.

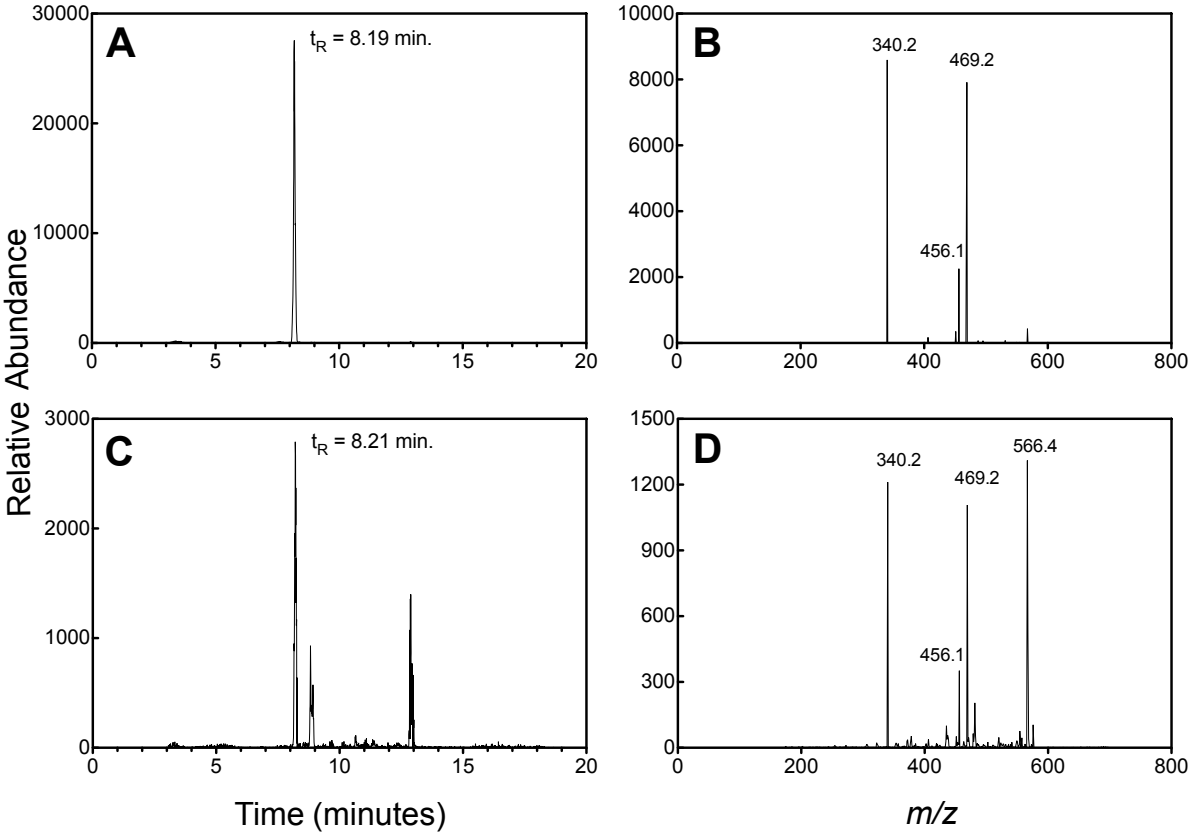


Figure 3.2 Detection of *S*-[2-(*N*^{6}-Deoxyadenosinyl)ethyl]GSH in Calf Thymus DNA treated with *S*-(2-choroethyl)GSH. UPLC chromatogram of synthetic *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH (A) and reaction product (C), observing precursor ion *m/z* 585 fragmenting to *m/z* 340. MS/MS data of precursor ion *m/z* 585 splitting into fragment ions for synthetic standard (t_R 8.19 min) (B) and reaction product (t_R 8.21 min) (D).

3.2.3 Primer Extension Assays

To investigate the ability of DNA polymerases to bypass the *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH adduct, FAM-labeled primer extension assays were performed in duplex oligonucleotide substrates containing the site-specifically incorporated *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH adduct in the template strand. Replication experiments were done with the human translesion polymerases hPol *κ*, *η*, and *ι*, as well as with *S. solfataricus* Dpo4, bacteriophage Pol T7, and *E. coli* DNA polymerase I Klenow fragment.

To determine the replication bypass of the GSH adduct, primer extension assays were performed in the presence of all four dNTPs (Figures 3.3 and 3.4). hPol *η*, Dpo4, and Klenow fragment were the most efficient of the four polymerases in terms of producing full-length primer products (P+7). hPol *ι* showed slight bypass of the adduct (P+2) but required higher enzyme concentrations (40 nM) and longer reaction incubations for extended bypass. However, hPol *κ* and Pol T7 were strongly blocked by the adduct. These polymerases were able to insert the two bases prior to the adduct site, but only low levels of bypass were observed after a 60 min incubation at 37 °C. Overall, the *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH adduct considerably blocked replication by each of the three human DNA polymerases.

Primer extension assays were also performed in the presence of single dNTPs to study misincorporation opposite the GSH adduct (Figure 3.5). hPol *κ* incorporated the correct dTTP in the control oligonucleotide but was completely blocked by the adduct. hPols *η* and *ι* each had similar or lower extents of misincorporation opposite the DNA adduct compared to the control template. Interestingly, Dpo4 showed less misincorporation opposite the *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH adduct compared to the control template.

To better understand the catalytic efficiency of the incorporation of each dNTP and the frequency of misinsertion, steady-state kinetic analysis was performed with hPols *κ*, *η*, and *ι* (Table 3.1). Compared to the unmodified substrate with dA, the *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH adduct decreased the catalytic efficiency of correct dTTP

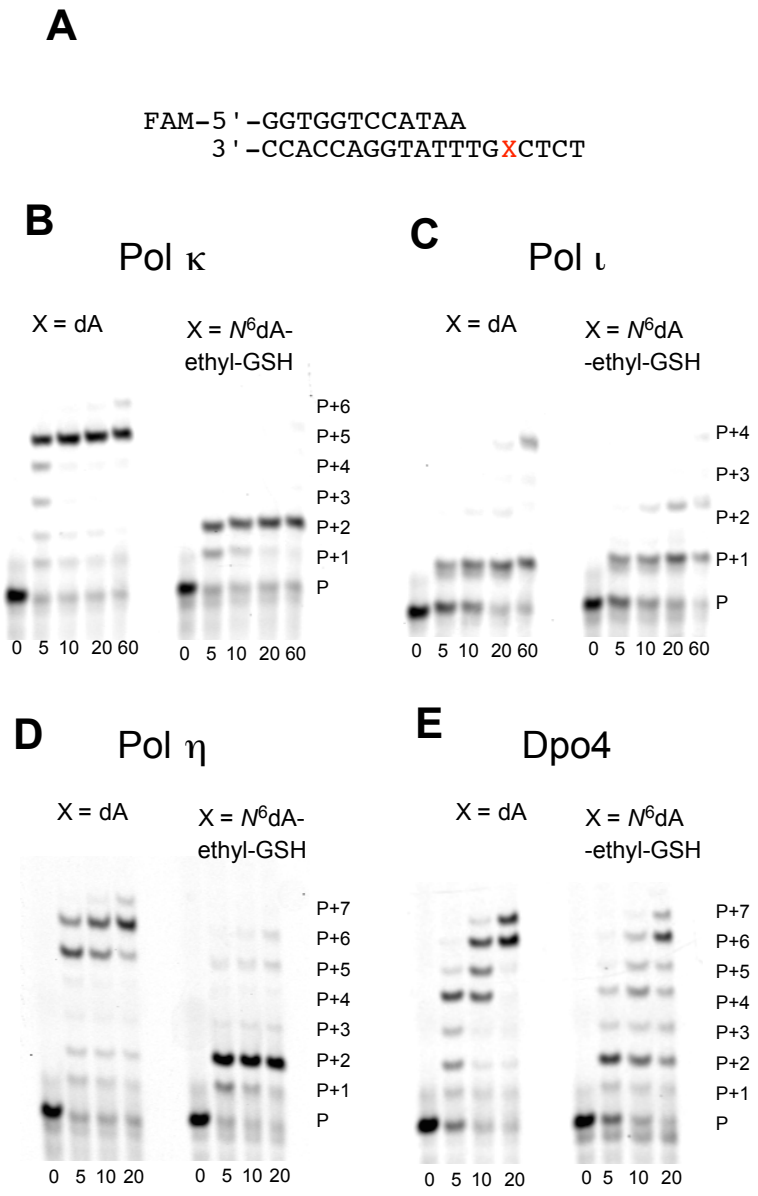


Figure 3.3 Primer Extension by DNA Polymerases in the Presence of All Four dNTPs. (A) Primer and template sequences where X is dA or *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH (*N*^{6}-dA-ethyl-GSH). Reactions were done with FAM-labeled primer in the presence of hPol *κ* (B), hPol *η* (C), hPol *ι* (D), and Dpo4 (E). Reactions were conducted for increasing times (indicated in minutes) in the presence of 20 nM enzyme with the exception of hPol *ι* (40 nM).

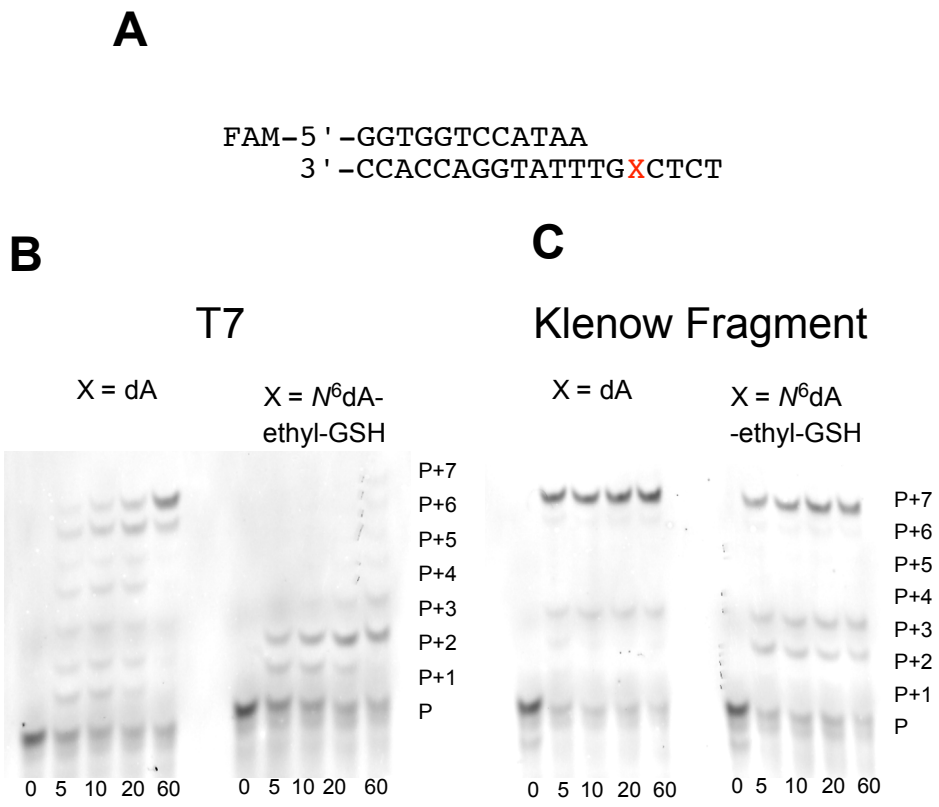


Figure 3.4 Additional Primer Extension by DNA Polymerases in the Presence of All Four dNTPs. (A) Primer and template sequences where X is dA or *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH (*N*^{6}-dA-ethyl-GSH). Reactions were done with FAM-labeled primer in the presence of bacteriophage T7 DNA polymerase (117) (B) or Klenow fragment (118) (C) at 20 nM for the indicated times (in min, at 37 °C).

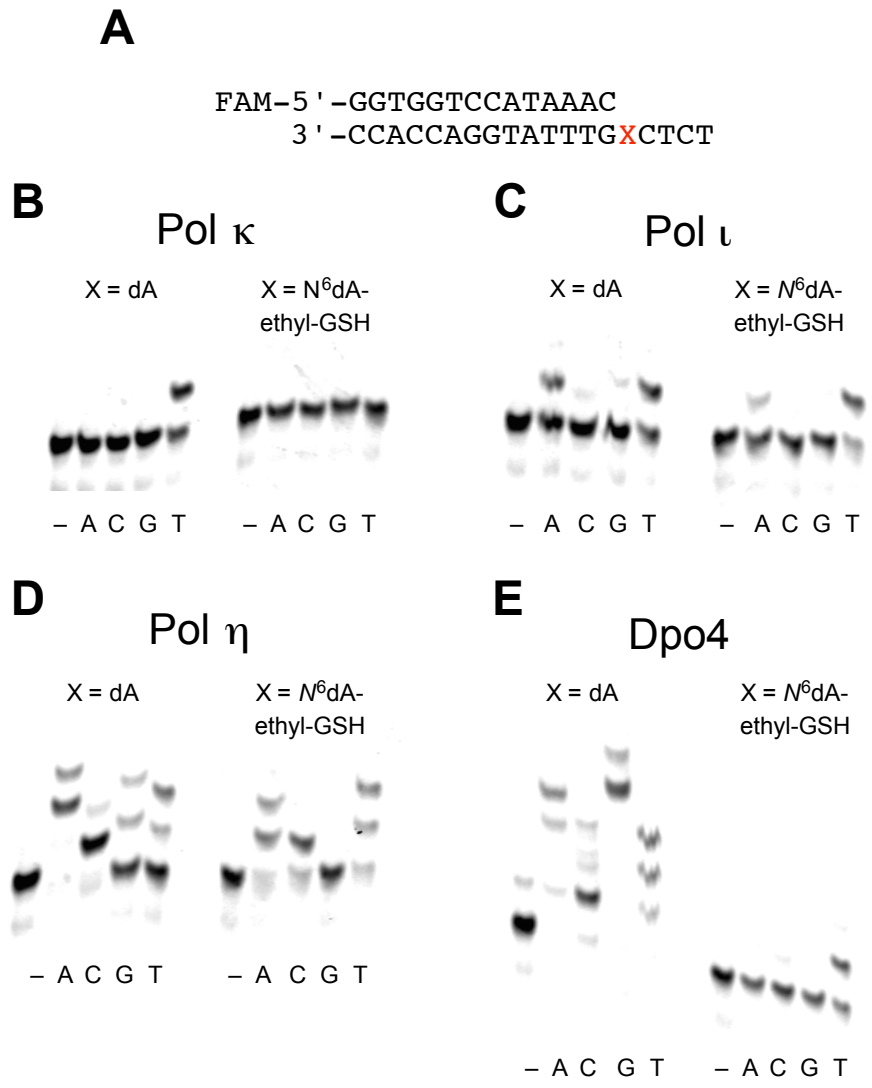


Figure 3.5 Primer Extension in the Presence of Single dNTPs by Y-family DNA Polymerases. (A) Primer and template sequences where X is dA or *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH (*N*^{6}-dA-ethyl-GSH). Reactions were done with FAM-labeled primers in the presence of hPol *κ* (B), hPol *η* (C), hPol *ι* (D), and Dpo4 (E). Reactions were conducted for 5 minutes in the presence of 20 nM enzyme with the exception of hPol *ι* (40 nM).

Table 3.1 Steady-State Kinetic Analysis of DNA Primer Single-Base Insertion Reaction

polymerase	*template*	*dNTP*	k_{cat} *(min^{-1})*	K_m *(μM)*	k_{cat}/K_m *(min^{-1} μM^{-1})*	*f*[a]
hPol *κ*	dA	T	5.0 ± 0.4	25 ± 7	0.20	1
		A	0.013 ± 0.002	170 ± 62	0.000079	0.00040
	N^6-dA-ethyl-GSH	T	0.044 ± 0.014	230 ± 140	0.00019	1
		A	*n/a*[b]	*n/a*[b]	*n/a*[b]	*n/a*
hPol *η*	dA	T	0.59 ± 0.02	0.67 ± 0.12	0.89	1
		A	0.49 ± 0.03	21 ± 5	0.023	0.026
		C	0.88 ± 0.07	91 ± 18	0.0097	0.011
		G	0.49 ± 0.04	20 ± 5	0.025	0.028
	N^6-dA-ethyl-GSH	T	0.89 ± 0.07	5.6 ± 1.2	0.16	1
		A	1.6 ± 0.2	120 ± 28	0.013	0.083
		C	0.89 ± 0.03	92 ± 7	0.0096	0.060
		G	0.12 ± 0.01	4.9 ± 1.0	0.024	0.15
hPol *ι*	dA	T	2.3 ± 0.2	33 ± 8	0.069	1
		A	0.076 ± 0.004	79 ± 10	0.00096	0.014
	N^6-dA-ethyl-GSH	T	1.9 ± 0.1	42 ± 9	0.046	1
		A	0.062 ± 0.003	36 ± 5	0.0017	0.038

[a] Misinsertion frequency $f = [k_{cat}/K_m]_{incorrect} / [k_{cat}/K_m]_{correct}$

[b] DNA incorporation was below limits of quantitation ($k_{cat} < 0.004$ min^{-1})

incorporation by each of the polymerases. hPol κ was strongly hindered by the adduct, with the catalytic efficiency for correct incorporation decreased >10^3-fold, while hPol η (5.6-fold) and pol ι (1.5-fold) had less attenuation. The catalytic efficiency of misincorporation was also low for the various polymerases. With hPol η, there was an increase in misinsertion frequency of dATP, dCTP, and dGTP of 3.2-, 5.5-, and 5.4-fold, respectively. The misinsertion frequency of dATP was also increased with hPol ι (2.7-fold) while the misinsertion of dATP by hPol κ was below the limit of quantitation (k_{cat} < 0.004 min^{-1}).

3.2.4 Primer Extension Analysis by LC-MS

While steady-state kinetic measurements provide useful information about the catalytic efficiency of the correct versus incorrect incorporation opposite the DNA adduct, this analysis does not provide any information about the patterns of extension nor other types of mistakes such as frame-shift events. Therefore, full primer extension assays were performed, in the presence of all four dNTPs, using LC-MS/MS analysis to sequence nucleotide incorporations opposite and after the *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH adduct. This analysis utilizes a uracil-containing primer (5´-GGTGGTCCATAA(dU)C-3´), and the product is cleaved using UDG followed by hot piperdine treatment to create a shorter DNA oligonucleotide for MS analysis (114,137). Analysis was performed with hPol η because its minsertion frequencies were the highest of the three DNA polymerases. The product species were isolated and fragmented using CID (Figure 3.6). These fragment ions were then compared to theoretical ions calculated using an Oligo Mass Calculator (The RNA Institute, University of Albany SUNY, Albany, NY; http://mods.rna.albany.edu/masspec/Mongo-Oligo) to identify the products. Analysis of the unmodified oligonucleotide showed only the correct full-length replication product, whereas the modified oligonucleotide revealed an n-1 frameshift (3%) and a dCTP misinsertion opposite the adduct (6%) in the full-length replication product (Table 3.2).

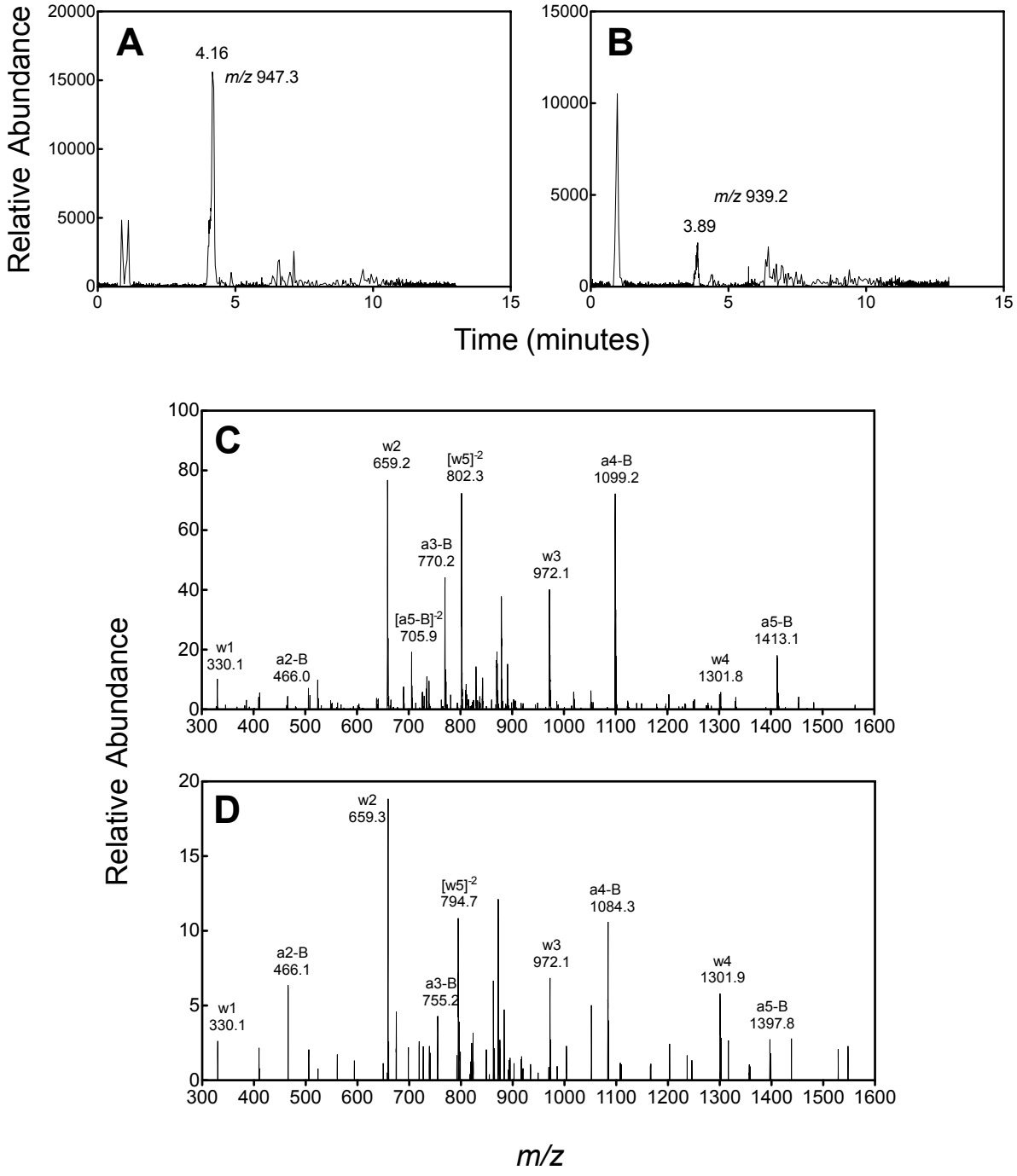


Figure 3.6 DNA Replication Sequencing by LC-MS/MS. Extracted ion chromatograms and MS/MS spectra of ions at *m/z* 947.3 (-2) (A and C) and 939.2 (-2) (B and D) from full-length primer extension products formed by hPol *η* in the presence of all four dNTPs.

Table 3.2 LC-MS/MS Analysis of DNA Primer Extension

5´6FAM–GGTGGTCCATAA**U**C
3´–CCACCAGGTATTTG**X**CTCT

product sequence	m/z (-2)	peak area	% frameshift	%C	%T
CTGAGA	947.1	60842			91
CTGAGAA	1103.3	125955			
CCGAGA	939.6	13057		6	
C_GAGA	795.0	5746	3		

3.3 Discussion

Mutagenic DNA-peptide crosslinks can be formed from DBE with the tripeptide GSH and the DNA repair protein AGT (138). We first reported a crosslink formed with GSH following the reaction of GSH and DBE with DNA in the presence of GSH transferase (89). The major adduct was identified at the N7 position of dG (89,139,140) with lesser amounts of adduction located at the N2 and O6 positions of dG (99) as well as the N1 position of dA (129). Our studies on the sites of adduction of AGT later identified the N6 position of dA as another adduction site that had not been seen using GSH (92). Therefore, we were interested in whether or not the GSH adduct was also formed at the N6 position of dA and, if so, what its biological properties are.

S-[2-(*N*^{6}-Deoxyadenosinyl)ethyl]GSH was detected in calf thymus DNA treated with *S*-(2-chloroethyl)GSH (Figure 3.2). Higher concentrations of *S*-(2-chloroethyl)GSH were used for DNA modification compared to our previous studies, which may explain why the adduct was not detected previously. However, the result can also be explained by advances in analytical chemistry, particularly in mass spectrometry, since 1992. Adjusting the concentration of *S*-(2-chloroethyl)GSH used to that in our earlier work would predict the extent of modification for the *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH adduct to be ~0.1% of the total adducts (Figure 3.7). The levels of the *N*^{6}-dA adduct formed with GSH are in the range of those formed at the dG *N*^{2} and *O*^{6} and the dA *N*^{1} positions (99). We had previously reported that *S*-[2-(*N*^{1}-deoxyadenosinyl)ethyl]GSH did not convert to *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH in a Dimroth rearrangement unless the sample was subjected to high pH (129).

Replication studies were done with the *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH adduct placed in a 19-mer oligonucleotide, utilizing a sequence that has been used previously in this laboratory for GSH crosslinking studies (141). Qualitative analysis of the bypass efficiency of various translesion polymerases (Figure 3.3) showed that Dpo4 was the most efficient polymerase for replicating past the adduct, followed by hPol η and hPol ι. hPol κ was strongly

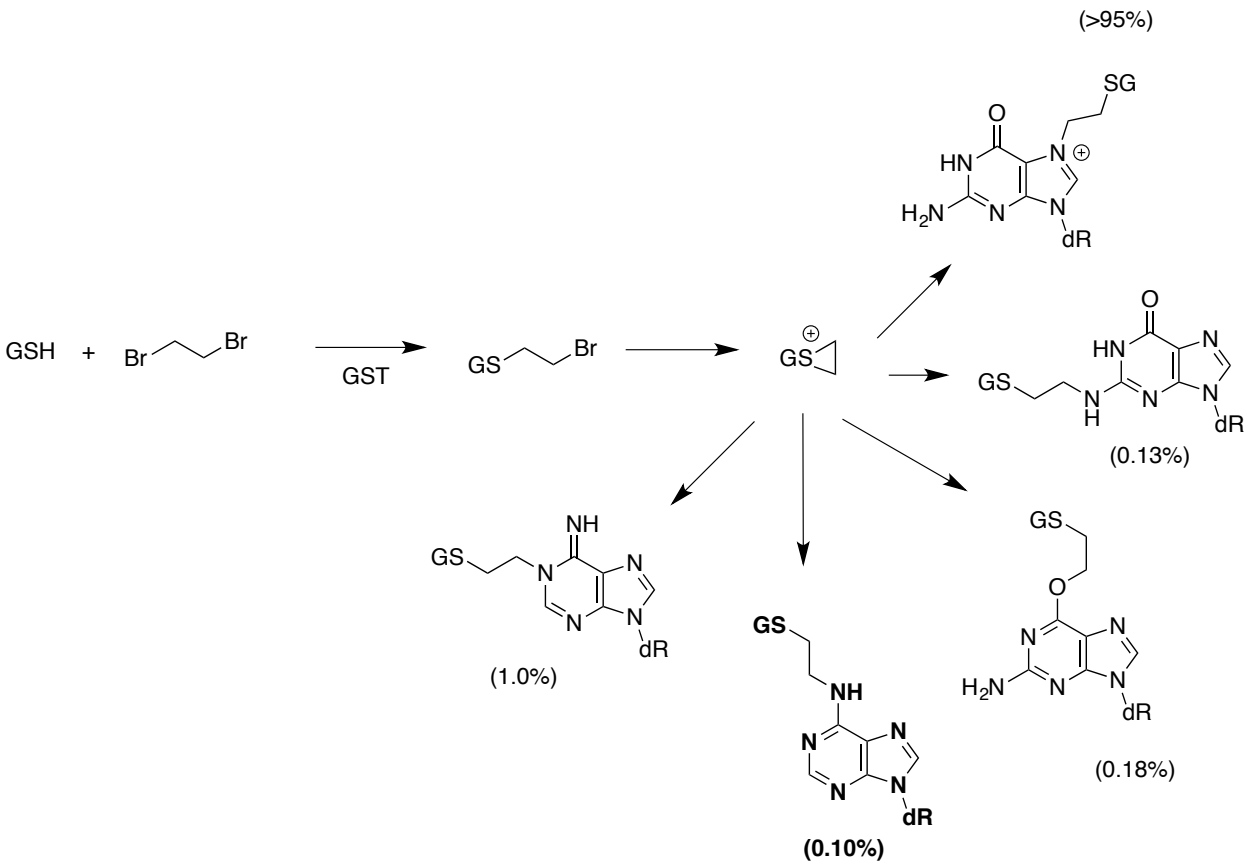


Figure 3.7 Structures and Relative Quantities of Glutathione-DNA Crosslinks Formed from Dibromoethane (99). *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH is shown in bold. dR: 2´-deoxyribose.

blocked by *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH, showing no replication past the adduct. These results were similar to previous studies showing hPol *η* to be the most efficient for bypass of bulky adducts in the major groove of DNA (46,142,143). Dpo4, although sometimes considered to be a homolog of hPol *κ*, had similar activity as hPol *η* (144).

In single-base insertion assays, dTTP was the nucleoside triphosphate most readily incorporated opposite the adduct (Figure 3.5). Steady-state kinetic analysis showed that misincorporation was somewhat higher than in controls for hPol *η* and hPol *ι*, with less than a five-fold increase in misinsertion frequency for the dNTPs (Table 3.1). This increase was similar to our previous results with hPol *η* and hPol *ι* with the *S*-[4-(*N*^{6}-deoxyadenosinyl)-2,3-dihydroxylbutyl]GSH adduct (141). hPol *κ* was severely blocked by the adduct, and no misincorporation was detected opposite the adduct in the work with hPol *κ*.

In order to better understand the insertion events opposite and past the adduct, LC-MS/MS analysis was performed to sequence the synthesized extension product. This protocol (114) utilizes an uracil residue in the primer strand, which allows the extension product to be cut using UDG/piperdine. Following cleavage, the products can be analyzed by HPLC-ESI-MS/MS, identified, and quantified. This experiment was performed only with hPol *η*, due to its high level of misinsertion of each of the dNTPs (Table 3.2). The major products formed included the correct base T inserted opposite the *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH with or without a blunt-end addition. A minor product corresponding to 6% of the overall product represented a misinsertion of C opposite the adduct, approximately the same percentage as the misinsertion predicted by steady-state kinetic analysis. An additional product was a -1 frameshift deletion, accounting for 3% of the product. The -1 frameshift product corresponds to the addition of a G opposite the C 5´ to the adduct and is consistent with the preferential insertion of G in the steady-state kinetic analyses (Table 1), in which a frameshift would not have been detected. Apparently if A is inserted (Table 1), it is not extended to generate full-length product.

Other N^6-adenyl adducts are known in the literature. The role of N^6-methyldeoxyadenosine is well-established in epigenetic mechanisms but to my knowledge no information is available regarding miscoding. Treatment of DNA with classical small alkylating agents does not yield much in the way of N^6-alkyl dA products (145). An adduct was found to be formed from endogenous *N*-nitroso compounds to generate the lesion N^6-carboxymethyl-2´-deoxyadenosine (146). However, this lesion did not block DNA replication or induce mutations *in vitro* or in *E. coli* (147,148). However, some larger electrophiles do generate N^6-dA DNA adducts (149), including some polycyclic hydrocarbons and aryl amines, which have been studied in some detail (150,151). In work from our group discussed earlier (141), the N^6-adenyl adduct formed by conjugation of butadiene diepoxide (and hydrolysis), the DNA polymerases studied here (*η*, *ι*, *κ*, Dpo4) were not very affected, but this entity linked to GSH caused hPol *κ* to insert C. Studies with two other butadiene-related adducts showed that N^6-(2-hydroxy-3-buten-1-yl)2'-deoxyadenosine retarded polymerization but did not affect fidelity, but one isomer of N^6,N^6-(2,3-hydroxybutan-1,4-diyl)-2´-deoxyadenosine caused lack of selectivity in nucleotide incorporation with hPols *η* and *κ* (152). The adduct N^6-oxopropenyl 2´-deoxyadenosine retarded the human translesion synthesis DNA polymerases but did not lead to error-prone polymerization (153).

One open question is how misincorporation at N^6-[2-[deoxyadenosinyl]ethyl]GSH compares with the other DBE-GSH adducts we have characterized (Figure 3.7). Three of these (the corresponding N^7-, N^2-, and O^6-dG adducts) had been synthesized earlier in this laboratory (136) but had only been examined with two *E. coli* DNA polymerases, pol I (Kenow fragment) and pol II exo$^-$ (154). (The human translesion synthesis DNA polymerases were yet unknown, and we do not currently have these three modified oligonucleotides in hand.) However, all three of the adducts showed some blockage and miscoding with *E. coli* Pol I and Pol II (154). We are still interested in establishing the relative abilities of the different adducts to retard polymerization and miscode with the human DNA polymerases, as well as their functions in human cells.

In conclusion, we identified the *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH adduct formed in calf thymus DNA. hPols *η* and *ι*, *S. solfataricus* Dpo4, and *E. coli* Pol I Klenow fragment were capable of bypassing the adduct without significant obstruction. The incorporation proceeds with relatively low but definite misincorporation by hPols *η* and *ι* and *S. solfataricus* Dpo4. These results help elucidate the relevance of individual GSH-coupled adducts in relation to the overall mutagenic effects of DBE.

CHAPTER 4

Degradation and Bypass of DNA-Protein Crosslinks formed by AGT at the N^6-Position of Adenosine

4.1 Introduction

The genomes of cellular organisms are under constant stress from endogenous and exogenous stress (74). This stress can result in the numerous types of DNA damage, including oxidative damage, alkylation, base mismatches, bulky lesions, intrastrand and interstand DNA crosslinks, and DNA-protein crosslinks. One under-examined class of DNA lesion is DNA-protein crosslinks (DPCs). This class of DNA lesions is formed by the covalent attachment of a protein to the nucleobase, sugar, or phosphodiester moiety of a DNA molecule. This type of DNA lesion can be detrimental due to its large size, which prohibits DNA replication and RNA transcription from ensuing. An example is DPCs formed from O^6-alkylguanine DNA alkyltransferase (AGT) (93). There are a few notable endogenous causes of DPCs. Topoisomerases cause DPCs during their enzymatic mechanism to unwind DNA (96-98). DNA polymerase β also can form DPCs with abasic sites while excising it from a DNA strand (155). DPCs can be formed via exogenous crosslinking agents as well, such as UV light, ionizing radiation, certain transition state metals, platinum-based chemotherapeutics (cisplatin), and various carcinogens such as formaldehyde.

Little is known about the mechanisms for repair of DPC lesions, with the exception of DPCs produced with topoisomerases. After proteolysis of the topoisomerase to small peptide fragments, the repair protein tyrosyl-DNA phosphodiesterase-1 or -2 (TDP1 and TDP2) hydrolyzes the phosphotyrosyl bond between the protein and DNA (96,98). For all other types of DPCs, it is thought that they are repaired either through DNA degradation and recombination or nucleotide-excision repair (NER). Nuclease dependent repair of DPCs utilizes the MRN complex to cleave the DNA on both sides of the protein crosslink, forming a double-stranded break in the DNA (156,157). The DNA can then be further processed by the canonical mechanisms of DNA

double-stranded break repair, homologous recombination, or non-homologous end joining. The other pathway for repair of DPCs involves NER, which cleaves the DNA on both sides of the adduct to remove the damaged strand DNA. DNA polymerases utilize the complimentary strand of DNA to resynthesize the excised portion to complete the repair of DNA. As well, it has been widely postulated that translesion DNA polymerases could bypass DPCs when encountered during DNA replication. While this mechanism tends to be more error-prone, it is less complex compared to either of the aforementioned repair pathways and would prevent stalled replication forks. However, if either NER or TLS occur, larger proteins (>10 kDa) would block these pathways from proceeding (100,101). These DPCs would require a nuclear protease to degrade the protein to a smaller size that would allow either pathway to ensue.

Recent research has discovered a yeast protease, Wss1, that can perform this degradation of DPCs (102). A human homologue, SPRTN, was found to have the same activity (110,112). The name, SPRTN, is derived from the putative zinc metalloprotease domain SprT located in the N-terminal part of the protein but the enzyme was first portrayed as a facilitator for p97 binding in regulating DNA damage response (103,104). The protease activity of SPRTN was later characterized as a zinc-dependent metalloprotease that is only active in the presence of DNA. It has the ability to degrade itself and other DNA-binding proteins such as topoisomerases, histones, and Rad5 but had no activity on non-DNA binding proteins (110,112,113).

The available literature suggests that SPRTN may be the protease required for TLS to occur past DPCs found during DNA replication. The purpose of this project was to determine if SPRTN can cleave DPCs formed from AGT. We synthesized an oligonucleotide that contains an AGT crosslink at the N^6-position of adenosine in a DNA primer-template complex. The DPC was reacted with SPRTN to allow degradation of the protein before testing bypass of the DPC by the human DNA polymerase η.

4.2 Results

4.2.1 Synthesis of S-[2-(N^6-Deoxyadenosinyl)ethyl]AGT Peptide-Containing Oligonucleotides

DNA oligonucleotides were synthesized with synthetic peptides from the active site of AGT to model potential bypass substrates for Y-family DNA polymerases. Three different peptide adducts were synthesized: a tripeptide (PCH, 3mer), heptapeptide (LIPCHRV, 7mer), and a decapentepeptide (PVPILIPCHRVVSSS, 15mer). The 15-mer peptide was altered at its 13th position, where in the wild-type human protein there is another cysteine. To prevent multiple reactable thiol groups, this residue was altered into a serine residue. These peptides were purchased with the N-terminus protected with an acetyl group and the C-terminus protected with an amide group to prevent reactions from occurring at the termini. To crosslink the peptides to the DNA, the peptides were first treated with MSH to selectively convert the cysteines into dehydroalanines.

The 19-mer DNA oligonucleotide containing a 6-chloropurine was reacted with cystamine to form an adenosine with the cystamine linked to the N^6-position. The disulfide on the cystamine was then reduced to form the aminoethylthiol. The modified DNA oligonucleotide was then incubated with the peptides to react through a Michael addition to form the DNA-peptide crosslink with an ethyl linkage. The oligonucleotides were purified by HPLC and their identities were confirmed by LC-MS.

4.2.2 Bypass of AGT Peptide Crosslinks

The ability of Y-family DNA polymerases to bypass DNA-peptide crosslinks was assessed by performing *in vitro* bypass assays. The replication assays were performed with duplex DNA with a FAM-labeled primer using hPol κ, η, and ι, as well as with *S. solfataricus* Dpo4 as the translesion DNA polymerases.

Full-length primer extension assays were first performed to study the ability for the enzymes to bypass the DNA-peptide crosslinks (Figure 4.1). These assays were performed in the

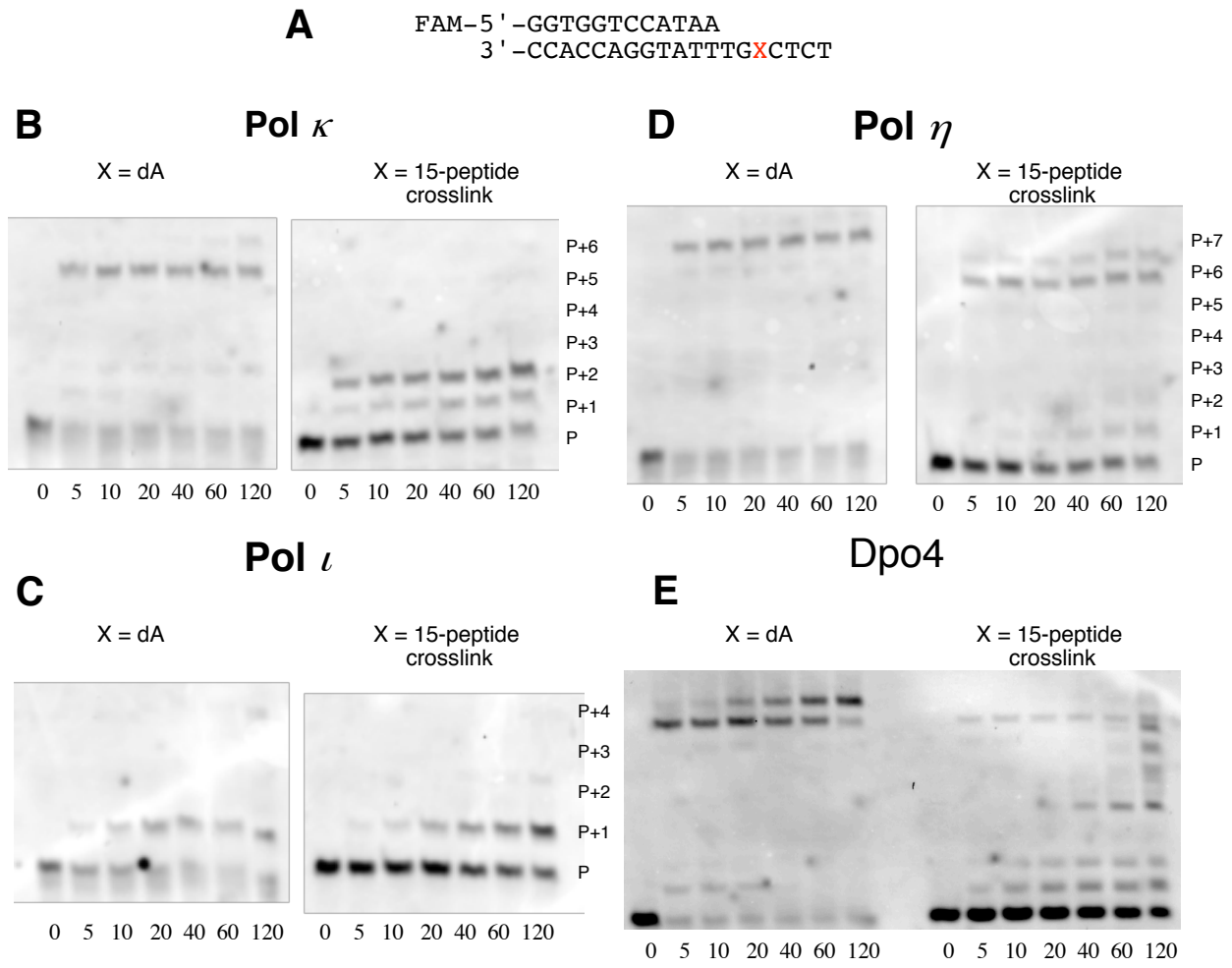


Figure 4.1 Primer Extension Assays in the Presence of All Four dNTPs by Y-Family DNA Polymerases. (A) Primer and template sequences where X is dA or N^6-dA-ethylene-15mer peptide from the active site of AGT. Reactions were done with FAM-labeled primers in the presence of hPol κ (B), hPol ι (C), hPol η (D), and Dpo4 (E). Reactions were conducted over a time course from 0-40 minutes in the presence of 20 nM enzyme with the exception of hPol ι (40 nM).

presence of all four dNTPs and tested the ability of the polymerases to incorporate the two bases prior to the crosslink, opposite the adduct, and to complete the replication to the end of the DNA template. In order to test the ability of the polymerases to bypass bulky peptide crosslinks, these assays were performed using the 15-mer peptide crosslink (Figure 4.1). hPol *η* and Dpo4 were the most efficient of the enzymes to bypass the peptide crosslink. However, hPol *κ* was strongly blocked by the adduct. These polymerases were able to insert the two bases prior to the adduct site, but no detectable levels of bypass were observed after a 40-minute incubation at 37 °C. hPol *ι* also showed no bypass of the adduct (P+2) even using higher enzyme concentrations (40 nM), but only slight extension of the control template was detected after 40 min.

To understand the misincorporation of the DNA polymerases opposite the peptide crosslinks, primer extension assays in the presence of individual dNTPs were performed (Figure 4.2). These assays were performed with a 14-mer DNA primer to allow the first nucleotide incorporation to be inserted opposite the DNA-peptide crosslink. Similar to the earlier assay, hPol *κ* was blocked by each of the peptide crosslinks on the DNA compared to the control template. It was able to incorporate small amounts of dTTP for each of the 3-, 7-, and 15-mer peptide crosslinks, and interestingly, hPol *κ* appeared to have increased ability to incorporate dTTP opposite the 15-mer peptide crosslink compared to the shorter crosslinks. hPol *ι* had similar levels of incorporation opposite the 3-mer and 7-mer peptide crosslinks as its control template, showing almost full incorporation of dTTP with slight misincorporation of dATP followed by dCTP and dGTP. The 15-mer peptide crosslink caused a decrease in each of the incorporations by hPol *ι* in while causing the polymerase to have a similar level of misincorporation as the smaller peptide crosslinks. hPol *η* was able to perform multiple incorporations by each of the dNTPs opposite the control template. It was able to insert multiple dNTPs opposite each of the peptide crosslinks as well, but the levels of the incorporation decreased as the peptide crosslink length increased. Dpo4 also inserted multiple nucleotides opposite the control template for each of dATP, dCTP, and

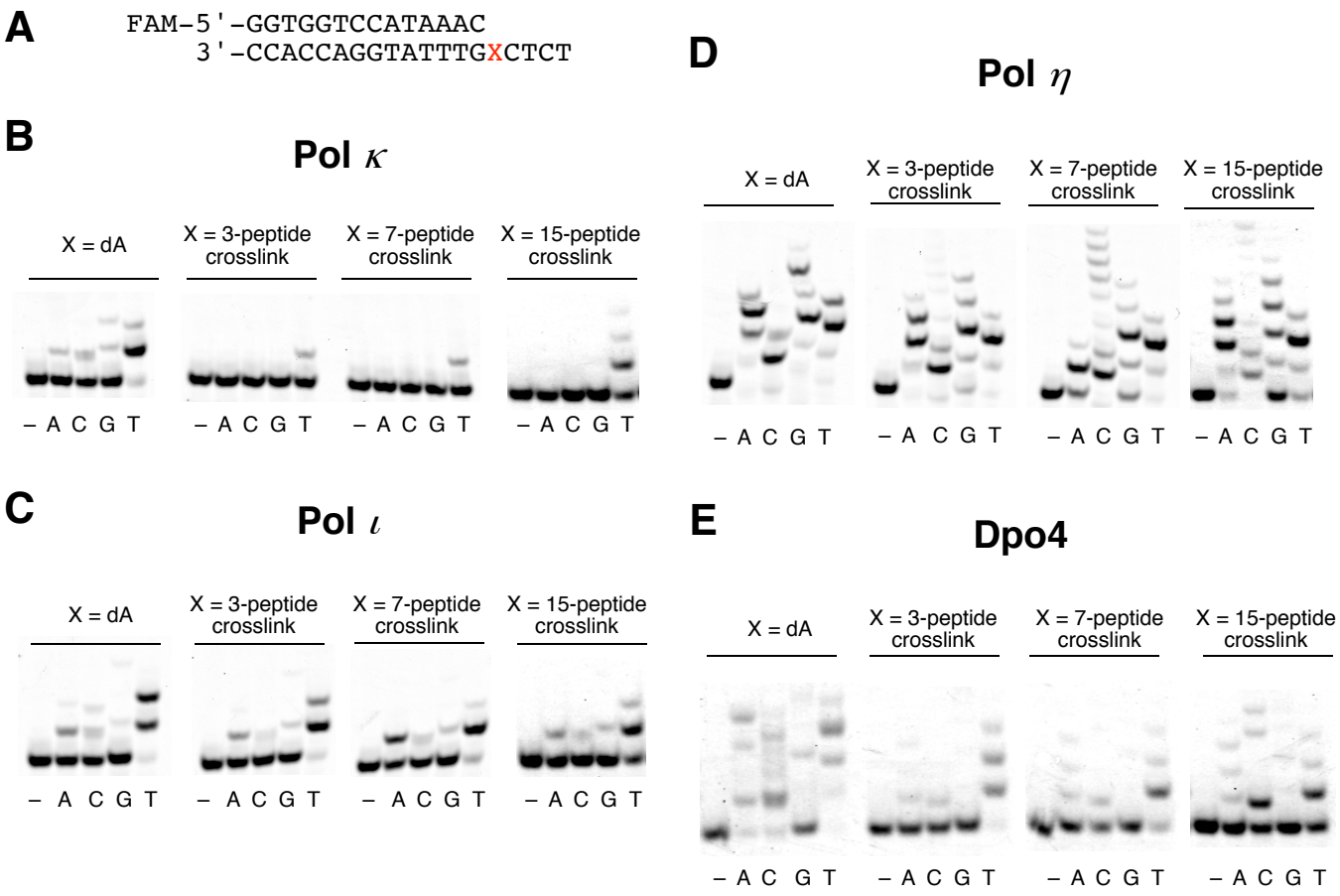


Figure 4.2 Primer Extension in the Presence of Individual dNTPs by Y-Family DNA Polymerases. (A) Primer and template sequences where X is dA or N^6-dA-ethylene-peptides from the active site of AGT. Peptides used were 3, 7, and 15 amino acids in length. Reactions were done with FAM-labeled primers in the presence of hPol κ (B), hPol ι (C), hPol η (D), and Dpo4 (E). Reactions were conducted for 5 minutes in the presence of 20 nM enzyme with the exception of hPol ι (40 nM). *Work done by Dr. K. Johnson.*

Table 4.1 Steady-State Kinetic Analysis of DNA Primer Single-Base Insertion Reaction by hPol *η*.

Template	*dNTP*	k_{cat} *(min^{-1})*	K_m *(μM)*	k_{cat}/K_m *(min^{-1} μM^{-1})*	*f*[a]
dA	T	1.6 ± 0.1	0.87 ± 0.15	1.8	1
	A	0.79 ± 0.03	20 ± 3	0.040	0.022
	C	0.50 ± 0.03	83 ± 15	0.0060	0.0033
	G	0.26 ± 0.01	9.8 ± 1.7	0.027	0.015
N^6-Adenyl-peptide	T	0.33 ± 0.03	8.1 ± 2.2	0.041	1
	A	0.13 ± 0.01	240 ± 50	0.00054	0.013
	C	0.093 ± 0.007	92 ± 18	0.0010	0.024
	G	0.037 ± 0.003	28 ± 8	0.0013	0.032

[a] Misinsertion frequency $f = [k_{cat}/K_m]_{incorrect} / [k_{cat}/K_m]_{correct}$

dTTP, with minor incorporation of dGTP. Opposite the 3-mer and 7-mer peptide, Dpo4 completely incorporated dTTP with minor levels of misincorporation of dATP and dCTP. However, it was partially blocked when incorporating dTTP opposite the 15-mer peptide crosslink and had increased levels of misincorporation of dCTP compared to the smaller peptide crosslinks.

Based on this evidence, as well as previously published work (141,158), hPol *η* was utilized to test the catalytic efficiency for the incorporation opposite the 15-mer peptide-DNA crosslink due to its ability to bypass peptide crosslinks at the N^6 position of adenosine. Steady-state kinetic analysis was performed using the 15-mer peptide crosslink to have a better understanding of the misincorporation levels by the polymerase (Table 4.1). hPol *η* had a significant decline in the catalytic efficiency for the correct incorporation opposite the 15-mer peptide crosslink compared to the control template, decreasing 44-fold. The misincorporation levels for the bypass of the crosslink were similar to the levels for the control template, showing slight increases for dCTP (0.33% to 2.4%) and dGTP (1.5% to 3.2%).

4.2.3 Degradation of AGT by SPRTN

Recombinant wt-SPRTN was expressed and purified in an *E. coli* protein expression system. The protein construct was designed based on a previously reported construct (110), containing an N-terminal Hisx6- plus Z-basic tag and a C-terminal Hisx6 tag. The gene construct for the recombinant protein was transformed into BL21 ArcticExpress (DE3) cells before expressing the protein at 20 °C for 18 h. The protein was purified by FPLC using a Ni-NTA column before removal of the N-terminal tag and further purification by size-exclusion chromatography. The identity of the protein was confirmed western blot and purity of the protein was checked by SDS-PAGE using a total protein stain. SPRTN was also shown to be active by self-digestion assays (Figure A.6).

To address the ability of SPRTN to degrade AGT-DNA crosslinks, protease assays were first performed with free AGT and SPRTN to determine if SPRTN has protease activity with AGT.

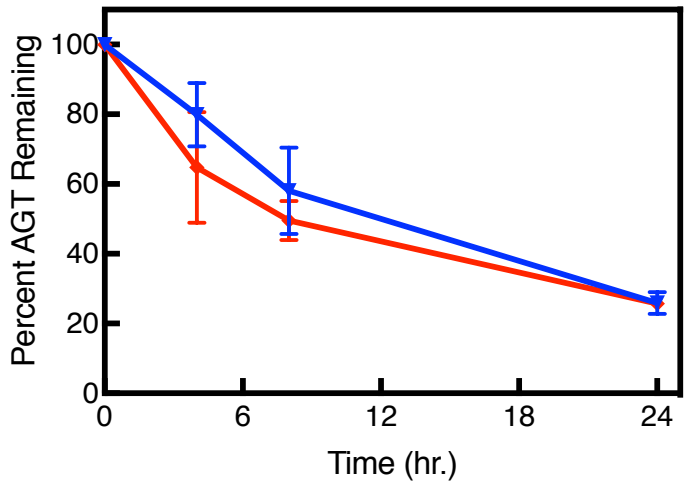


Figure 4.3 Proteolytic Degradation of AGT by SPRTN. AGT was incubated in the presence of either single-stranded DNA (red) or double-stranded DNA (blue) over a 24 hour time course with SPRTN. Incubations were performed in duplicate. Error bars denote standard deviation

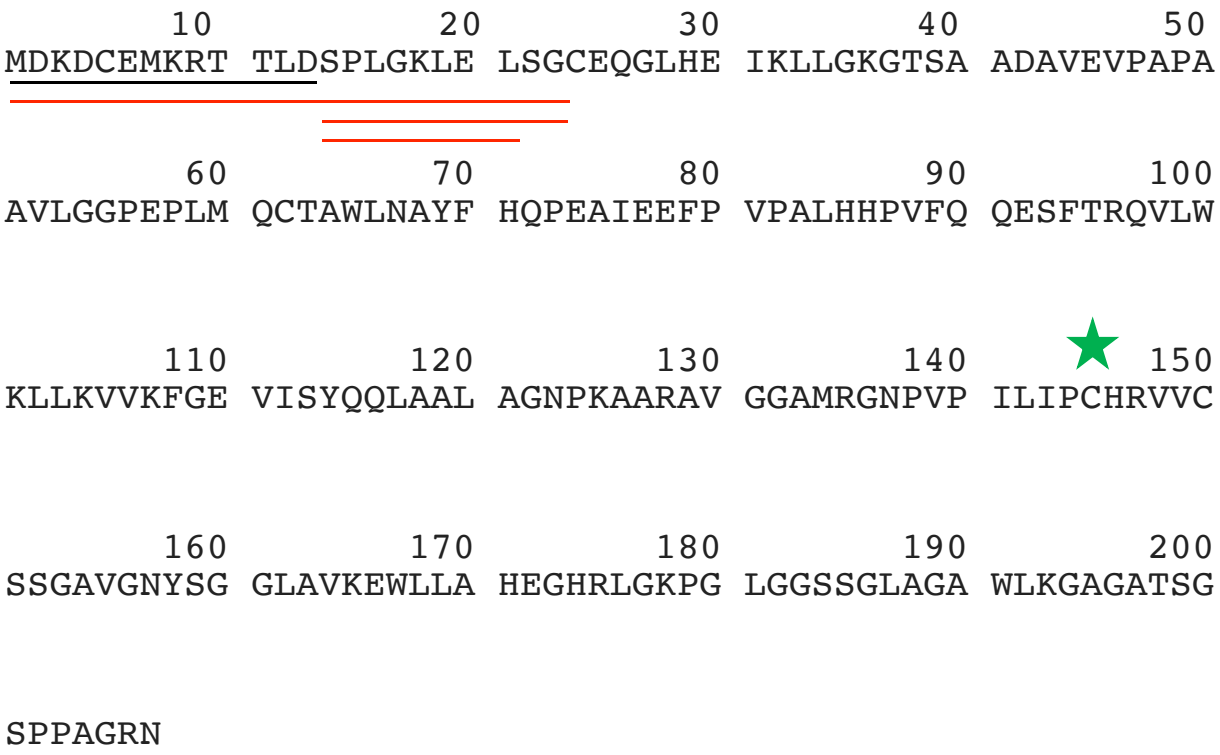


Figure 4.4 Proteomic Analysis of AGT Degradation by SPRTN. The black underline denotes the peptide that was present in all three timepoints. The additional peptides underlined in red were present only after a 16 h incubation with SPRTN. The position of the crosslinked DNA is marked by the green star.

SPRTN and AGT were incubated in the presence of both dsDNA or ssDNA (Figure 4.3). The reactions were done over a time course to 24 h to monitor the digestion of the AGT over time. SPRTN degraded the AGT in both dsDNA and ssDNA to similar levels and was found to degrade most of the protein by 24 h.

An AGT-DNA crosslink was then synthesized to observe the ability of SPRTN to degrade the crosslink. AGT was incubated with dibromoethane (DBE and the 50-mer duplex DNA to form the an AGT crosslink on DNA. Gel analysis identified that the crosslink was formed on DNA. SPRTN was then incubated with the DPC over a period of 16 h to complete digestion of the AGT protein. Proteomic analysis of the degraded AGT yielded peptides originating from the N-terminus of the protein. SPRTN cleaved AGT at three different sites, between an aspartic acid and serine, a leucine and serine, and a glycine and cysteine. Only one peptide, MDKDCEMKRTTLD, was able to be identified at the 2 and 4 h timepoints, but at the 16 h timepoint there were four different peptides identified including the peptide found earlier. The new peptides were formed from cleavages further into the AGT protein compared to the first peptide.

4.2.4 Synthesis, Degradation, and Bypass of a Site Specific AGT-DNA Crosslink

To determine if hPol *η* can bypass the digested AGT peptide-DNA crosslink, a site specific AGT crosslink was formed on a primer-template DNA duplex. The chemical synthesis scheme used for the peptide crosslinks from the active site of AGT was attempted, but the yield of the reaction was extremely low when attempted with the full AGT protein. Therefore, a second chemical synthesis scheme was developed to synthesize the protein crosslink using click chemistry. Using the same 6-chloropurine-containing oligonucleotide, the DNA was treated with propargyl amine to create an alkyne linker on the N^6 position of adenosine. This was then annealed to a 12-mer DNA primer before undergoing the copper-catalyzed azide-alkyne cycloaddition using an azide with an iodoacetamide linker. With the linker attached, the DNA duplex was incubated with AGT to form the crosslink at the active site cysteine (C145) of the

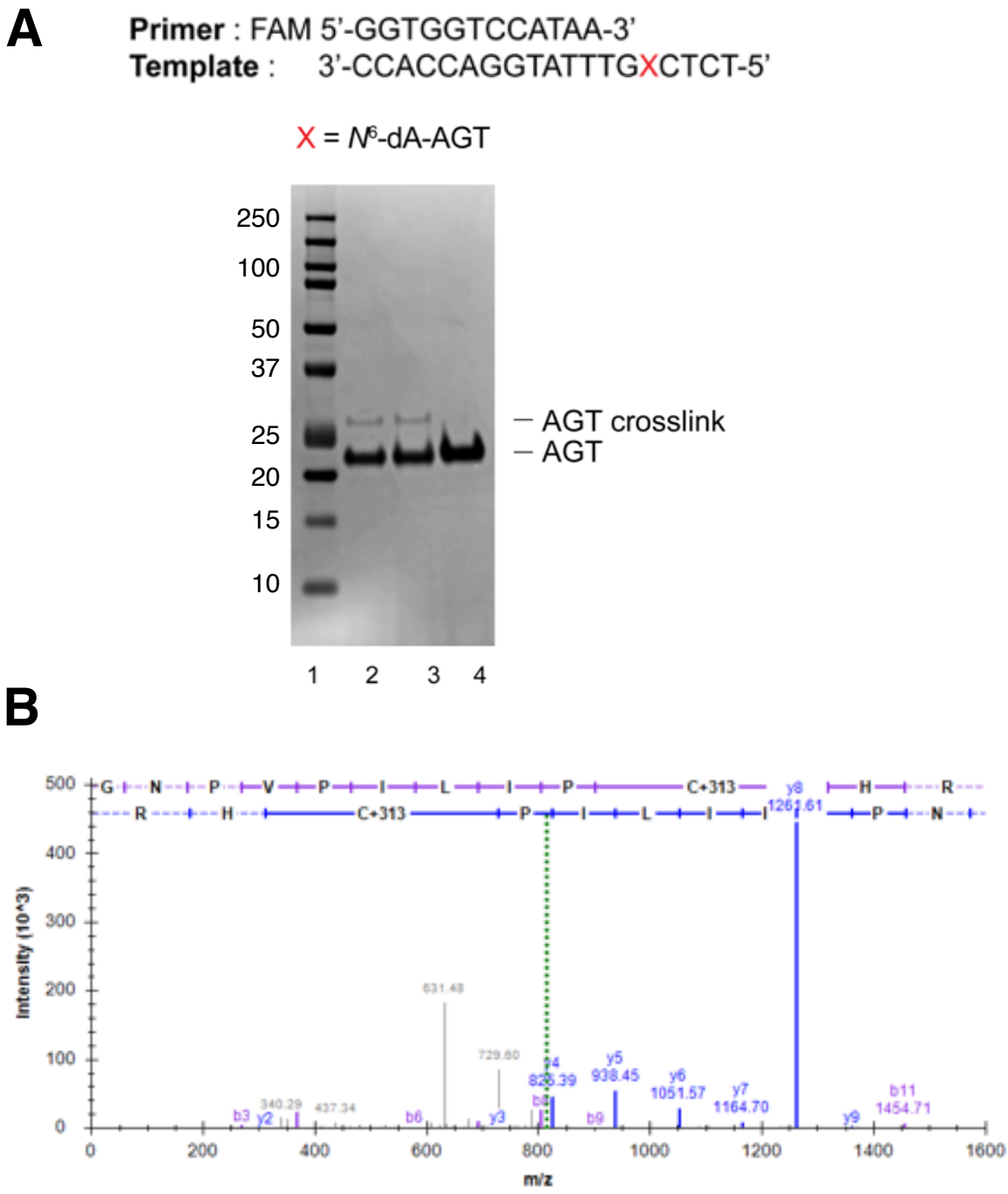


Figure 4.5 Formation of AGT-DNA Protein Crosslink. (A) SDS-PAGE analysis showing the formation of the N^6-dA-AGT crosslink, showing formation of the crosslink after 4 hours (lane 2) and 24 h (lane 3) compared to an AGT standard (lane 4). (B) LC-MS/MS analysis confirming the identity of the AGT crosslink at the active site cysteine residue of AGT. *Work done by Dr. P. Ghodke.*

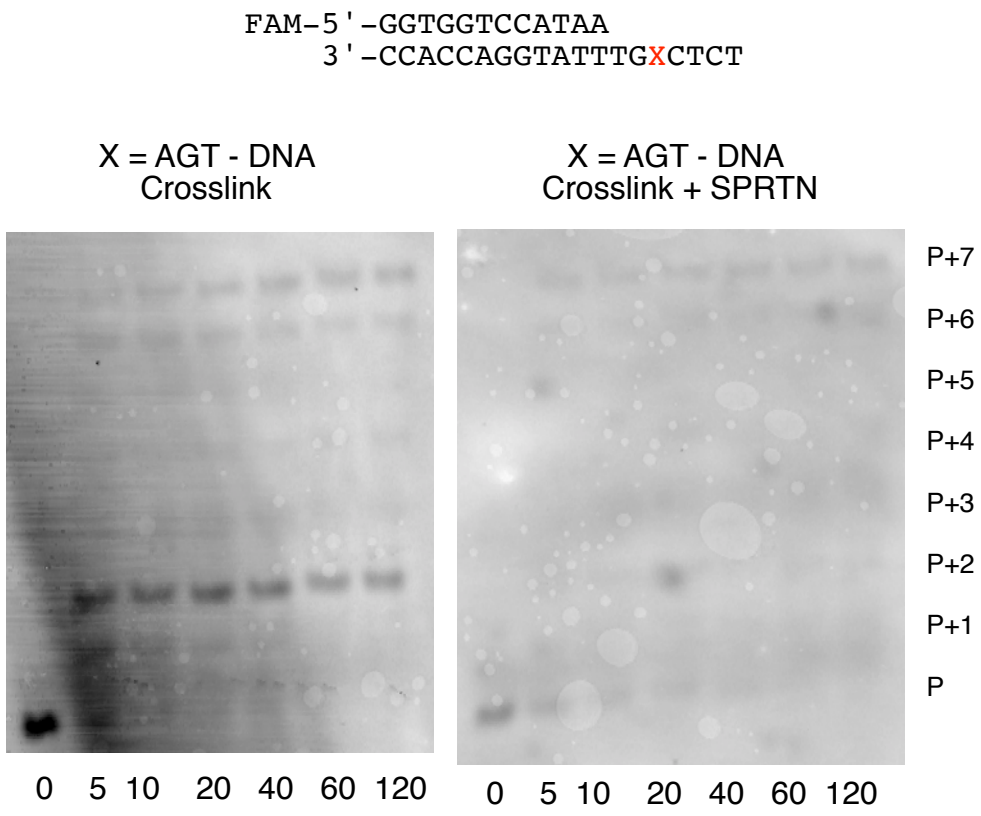


Figure 4.6 Bypass of AGT Protein Crosslinked to DNA. AGT-DNA crosslink was incubated with (right) or without (left) with SPRTN for 48 h to degrade the protein. Full primer extension was then performed by the addition of hPol *η*. Reactions were conducted for varying amounts of time (indicated in minutes) in the presence of 20 nM enzyme.

protein. Gel and proteomic analysis both confirmed the formation of the AGT crosslink on the DNA duplex (Figure 4.4). The crosslink was then purified by FPLC to remove all free DNA from the sample using Ni^{2+}-IMAC to elute the bound protein (Figure A.7). Due to tight binding of free DNA to AGT, 1.5 M NaCl was used to weaken the interaction between the two as well as a long wash step prior to elution of the crosslink.

The AGT-DNA crosslink was then incubated with SPRTN for 48 h to degrade the crosslink. The SPRTN reaction product, as well as the AGT-DNA crosslink, were incubated with hPol *η* in the presence of all four dNTPs to allow full primer extension of the adduct (Figure 4.5). The AGT-DNA crosslink caused the polymerase to stall after two nucleotide incorporations (P+2), signifying that the protein blocked the replication of the template at the site of the crosslink, as well as a minor band near the top of the gel showing full extension of the DNA primer. This minor product is assumed to be the result of contaminating free DNA as shown by evidence by ^{32}P labeling (Figure A.8). After treatment with SPRTN the P+2 band disappeared, and full extension of the primer strand was observed.

4.3 Discussion

DNA-protein crosslinks have been a recent focus in the field of DNA damage and repair for the interest in the deleterious effects that they may cause as well as the recent discovery of the nuclear protease, SPRTN, that can degrade them. Although protein crosslinks formed by proteins such as GADPH and histones did not cause any negative effects to cells (94,95), there has been overwhelming data showing AGT crosslinks to DNA are both highly cytotoxic and mutagenic to both bacterial and mammalian cells (93,131,159). This has led to significant interest in determining the mechanism of these effects of DPCs formed by AGT. While previous research has observed the ability of SPRTN to degrade nuclear proteins, our studies are the first to actually involve the degradation of a DPC, specifically the crosslinks formed by AGT. We observed the

results of a DPC forming at the N^6 position of adenosine and elucidated the effects that the crosslink in respect to SPRTN digestion and replication post-proteolysis.

Site-specific DNA crosslinks were first synthesized with synthetic peptides derived from the active site of AGT. Three peptides of increasing size (3 to 15 aa) were used for the synthesis to simulate the peptides to remain after proteolytic decay in cells. The peptides were centered around the active site cysteine which was determined to be the site of adduction when forming the crosslinks. These peptides were crosslinked using a novel synthetic chemistry scheme resulting in their conjugation onto the N^6 position of adenosine with an ethyl linkage, mimicking the same linkage that was found to be deleterious in cells.

Replication assays were performed with the peptide crosslinks on a 19-mer DNA oligonucleotide used previously in the laboratory (141,158). Qualitative analysis of the bypass ability of various translesion DNA polymerases showed that hPol η was the most efficient to fully bypass the 15-mer peptide crosslink (Figure 4.1). Dpo4 also was able to efficiently bypass the lesion, but was found to stall after the fourth nucleotide incorporation. Neither hPol ι nor hPol κ were found to be able to bypass the crosslink, showing stalling at both the P+1 and P+2 positions. These results were consistent with previous studies that hPol η was the major polymerase capable of bypassing bulky major groove adducts (46,142,143). Showing that hPol η is capable of bypassing the 15-mer peptide crosslink, we deemed it unnecessary to perform full-bypass assays using the smaller peptide crosslinks.

Single-base insertions were performed with each of the peptide crosslinks to get an understanding of the misincorporation for each of the translesion DNA polymerases (Figure 4.2). Consistent with the primer assay using all four dNTPs, hPol κ was significantly blocked by each of the peptide crosslinks and was only able to perform minor incorporation of the correct dTTP. hPol ι was able to efficiently incorporate dTTP opposite the 3- and 7-mer peptide crosslinks with minor misincorporation of dATP, but its total incorporation decreased in the presence of the 15-

mer peptide crosslink. High levels of incorporation were detected using hPol *η* in the presence with each of the dNTPs, most likely due to the high dNTP concentration and reaction duration used for this experiment. To gain a better understanding of the role of hPol *η* to perform misinsertions, steady-state kinetic analysis was completed using the 15-mer peptide crosslink (Table 4.1). This analysis determined the catalytic efficiency for hPol *η* to incorporate dTTP opposite the crosslink decreased 45-fold compared to the undamaged control template. The misincorporation for the enzyme was only slightly higher for the crosslink compared to the control, increasing from approximately 4% to 7% for all of the dNTPs.

Following the discovery of SPRTN in late 2016, we moved forward in expressing and purified the protease to test its ability to degrade AGT. SPRTN was first tested with unbound AGT to get a grasp on the rate of degradation of the protein. We found that SPRTN degraded AGT to ~20% of its total level within a 24 h incubation with either single- or double-stranded DNA. From here we synthesized AGT-DNA crosslinks using dsDNA and DBE and monitored the proteolysis of the AGT crosslink. Proteomic analysis was performed to determine the locations of cleavage on AGT by SPRTN. Preliminary results showed cleavage of AGT on the N-terminus of the protein, with the first peptide appearing within 2 h of the SPRTN incubation and further peptides being detected at 16 h. The proteolytic activity of SPRTN at the N-terminus of AGT was identical to previously reported degradation of histones at the N-terminus by SPRTN (110).

To fully understand the ability of SPRTN to degrade AGT-DNA crosslinks, we synthesized a site-specific AGT crosslink on a DNA primer-template duplex. Initial attempts to synthesize the crosslink using peptide synthesis scheme proved to be unsuccessful at resulting in detectable levels of the protein crosslink. To obtain useable amounts of the protein crosslink, a new synthesis scheme was designed utilizing click chemistry to attach an iodoacetamide linker to the DNA (Figure 4.5). This group selectively reacted with the cysteine 145 of the protein, resulting in the AGT crosslink, albeit with a longer linker than formed through environmentally. Purifying

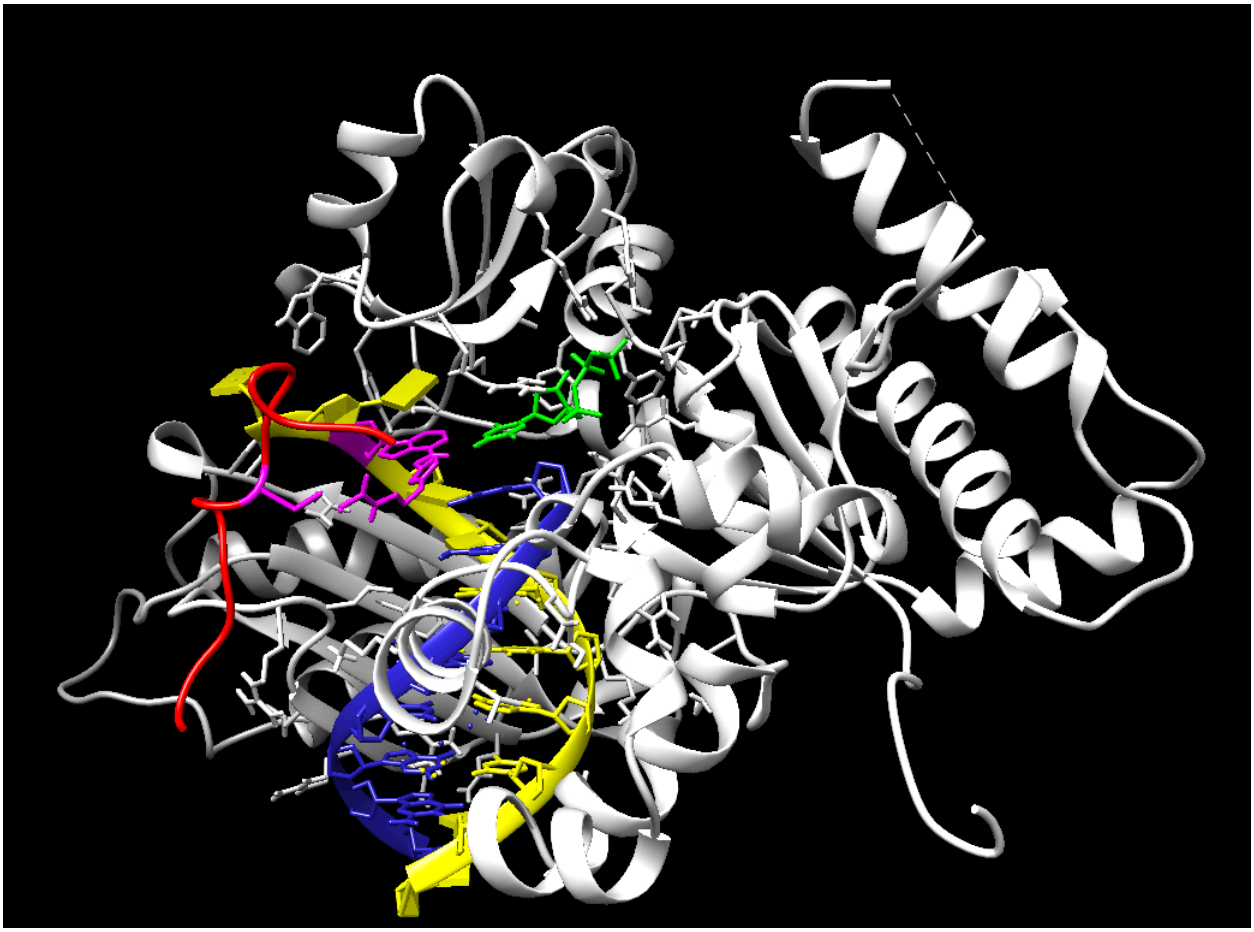


Figure 4.7 Computational Modeling of DPC Fitting into the Active Site of Human Polymerase *η*. Structure of the polymerase is shown in white, based on PDB file 5DG9. The DNA duplex is shown in the active site of the enzyme, showing the template strand in yellow and the primer strand in blue. The incoming dTTP (green) is shown lining up opposite the adenosine with the click-reaction linker (pink) adducted at the N^6-position of the adenosine. The 15-mer peptide (red) can be situated outside of the enzyme active site. Modeling was done in collaboration with Dr. Pradeep Pallan.

the protein crosslink proved to be equally, if not more, challenging to remove the unmodified DNA from the sample. Multiple attempts using affinity columns to separate the crosslink from unmodified DNA proved unsuccessful. A primer extension assay in the presence of all four dNTPs using hPol *η* did show that the protein blocked the majority of the extension prior to the crosslink (P+3), but a minor band was detected showing full primer extension (Figure 4.6). ^{32}P assays were utilized to confirm the presence of unmodified template DNA in the sample used for the replication (Figure A.8), leading us to believe that the full extension of the primer was due to the unmodified DNA as the large size of AGT would prevent the activity of the polymerase. However, after treatment of the AGT-DNA crosslink with SPRTN the stalled bands disappeared on the gel, showing the ability of the degradation of the AGT crosslink to allow replication bypass to occur. Molecular modeling experiments were also performed to test if hPol *η* would be able to accommodate AGT in its active site. Using the crystal structure of hPol *η* performing replication, we modeled in the linker used for the protein crosslink. The entire protein was not able to fit in with the DNA without having steric clashing with hPol *η*. Modeling in the 15-mer peptide with the click chemistry linker was able to fit however, showing the peptide sitting in a cleft outside of the protein away from the DNA (Figure 4.7).

Our results show that SPRTN is capable of degrading DPCs formed with AGT. The degradation of this protein appears to allow translesion DNA synthesis to occur by hPol *η*. The ability of hPol *η* to perform this replication is not surprising due previous literature showing its ability to bypass DNA-peptide crosslinks located in the major groove of DNA. Other research has found that these polymerases are capable of bypassing peptides of similar lengths as the 15-mer peptide described here, although the model linkers normally used are longer which may help accommodate the peptide outside the active site of the polymerase (143,160,161). Interestingly, misincorporation by hPol *η* was still low when bypassing the 15-mer peptide crosslink during replication. This would lead us to believe that the elevated mutations formed by AGT-DNA crosslinks must originate via another mechanism, such as by complications in the DNA repair

processes involved in its removal, but further experimentation would be required to determine this mechanism.

In conclusion, we performed DNA replication assays opposite DNA-peptide crosslinks originating from the active site of AGT located at the N^6 position of adenosine. Dpo4 and hPol η were able to bypass the largest peptide crosslink while hPols κ and ι were hindered, and the bypass was shown to have low levels of misincorporation. We were able to determine that SPRTN is capable of degrading AGT and AGT-DNA crosslinks. This degradation was then shown to alleviate stalled replication opposite AGT-DNA crosslinks by allowing TLS DNA replication to occur.

CHAPTER 5

Conclusions and Future Directions

5.1 Conclusions

The study of DNA-protein crosslinks is complicated by the diversity of the forms in which they exist. DPCs can occur with all four dNTPs, some with multiple locations. The agent involved in generating the crosslink can be short or long as well as flexible or rigid. Proteomic analysis has also identified hundreds of proteins that can be crosslinked to DNA. Therefore, it is very difficult to characterize exactly what role DPCs have in the field of DNA damage because each crosslink can be different.

Little effort has been put into trying to characterize the effects of DPCs in cells. Our lab has investigated the effects of four proteins using *bis*-electrophiles as the crosslinking agent. Perplexingly, we found that two, AGT and GSH, had cytotoxic and mutagenic effects while the other two, GADPH and histone H2b, had no effects (94,95). These results lead us to believe that either there are multiple pathways by which DPCs are handled or that some proteins are inherently detrimental when crosslinked. My dissertation has focused on the elucidation of the mechanism for how DPCs formed using either GSH or AGT are mutagenic. To investigate this, the effects of crosslinks at the N^6 position of adenosine were synthesized and tested for misincorporation during replication by various translesion DNA polymerases. Degradation assays were also performed on the AGT protein crosslink to investigate the necessity of a nuclear protease to allow for bypass of the protein crosslink.

Chapter III describes the role of the *S*-[2-(N^6-deoxyadenosinyl)ethyl]glutathione in DNA. This lesion was previously not known to exist in DNA, presumably because previous detection limits were not low enough to detect the adduct. *S*-[2-(N^6-Deoxyadenosinyl)ethyl]glutathione was proposed to exist after the N^6 adenosine adduct was detected using AGT as the crosslinked protein. Detection of this adduct was identified by reacting GSH and calf thymus DNA. The

reaction product was degraded using a mixture of nucleases to degrade the DNA prior to quantitation by LC-MS and comparison to a synthetic standard. Its level of formation was found to be similar to those of the other non-labile adducts. A site-specific GSH crosslink was synthesized with N^6-dA to perform replication assays by treating GSH with *N*-(BOC)-protected 2-bromoethylamine. After deprotection of the modified GSH, it was incubated with a 6-chloropurine-containing oligonucleotide to synthesize the lesion. Replication assays showed that hPol *η* was the major polymerase capable of bypassing the lesion. The misincorporation levels opposite hPol *η* were modestly higher, from 1-3% to 6-15% for each of the dNTPs. Sequencing the extension products by mass spectrometry revealed similar levels of misincorporation, but an interesting discovery was the presence of frameshift mutations in the extension products. This type of mutation is vastly more detrimental to the cell compared to substitution reactions due to its potential to alter the codon frames during transcription and would help clarify the severity the genotoxic effects of the GSH crosslinks (130).

My results have helped shed light on the total number of adduction sites for GSH-DNA crosslinks. While it was suspected for years that the N^6-dA crosslink exists, these are the first results confirming that hypothesis. The levels of the N^6-dA crosslink were similar as compared to the other sites of adduction when normalizing for the amount of GSH and dibromoethane used to synthesize the crosslink. Studying the ability of TLS polymerases to bypass this crosslink was also quite important in order to determine the mechanism for how DPCs are mutagenic. GSH-DNA crosslinks were known to be mutagenic from mouse experiments (130), but there was no direct analysis to determine what mutations were formed or the roles for each of the DNA adducts formed by dibromoethane and GSH crosslinks to cause mutations. The results presented in Chapter III demonstrate that these mutations can directly come from translesion DNA synthesis as compared to DNA repair of the lesion or other potential mechanisms. A previous paper studying N^6-dA GSH crosslinks using butadiene diexpoxide found that the misincorporation by hPol *η* was lower as compared to the results shown in Chapter III but did find high misincorporation levels by

hPol κ (141). The logical explanation for the differences in the hPol κ rates would be that the presence of the longer linker allowed its activity as compared to the shorter linker, but the earlier work also used higher DNA concentrations as compared to the work reported here (5 μM compared to 120 nM).

Chapter IV describes the effects of AGT-DNA crosslinks at the N^6 position of adenosine. The goal of this project was to determine if a nuclear protease is capable of degrading AGT-DNA crosslinks to allow DNA replication to progress, and subsequently, if this replication was error-prone. The first solution to this question was through the chemical synthesis of synthetic peptides from the active site of AGT attached to adenosine at the N^6 position. This strategy would elucidate the ability of the translesion polymerases, notably hPol η, to bypass potential degradation products. As predicted, hPol η was capable of bypassing the synthetic peptide crosslinks, although its catalytic efficiency decreased 45-fold compared to the control template. Surprisingly however, the misincorporation levels by hPol η opposite the crosslink were quite similar as compared to the control template. With the discovery of SPRTN as a nuclear protease, an effort to determine its ability to degrade AGT and AGT-DNA crosslinks was pursued. Expression of SPRTN was obtained by developing a gene construct based on a previously reported manuscript (110). The protein was designed to contain an N-terminal Z-basic plus a Hisx6 tag, as well as a C-terminal Hisx6 tag. The gene was inserted into a pNIC-ZB expression vector and transformed into *E. coli* designed for low temperature expression. The protein was expressed at 20 °C and purified using metal affinity chromatography prior to size exclusion chromatography. SPRTN was found to be active both in the presence of double-stranded DNA and single-stranded DNA, both for the proteolytic degradation of itself and other proteins. It was found that SPRTN can cleave AGT as well as AGT-DNA crosslinks. In proteomic analysis, the only peptides that were located were at the N-terminus of the protein and far from the active site (cysteine-145). In order to establish if SPRTN could degrade AGT-crosslinks to allow for TLS to occur, a site-specific AGT-DNA crosslink was synthesized and treated with SPRTN prior to reaction with hPol η to analyze

replication bypass. Initial steps to synthesize the crosslink using the same synthesis scheme as the peptide crosslinks proved unsuccessful. A more robust system—using click chemistry and iodoacetamide to link AGT to DNA—was found to be successful, albeit with low yields (~10%). Techniques for purifying this crosslink by binding the protein by its His tag proved rather difficult, in that AGT was capable of binding DNA throughout the elution phase on column chromatography. The AGT crosslink was treated with SPRTN, and the AGT crosslink (before and after SPRTN treatment) was incubated with hPol η in full extension assays. The AGT-DNA crosslink predominantly blocked replication by hPol η, but a minor band was detected representing full extension of the primer. This extension, however, is believed to be the result of contaminating free DNA in the sample, as shown by ^{32}P labeling of the DNA template. Treatment with SPRTN prior to replication resulted in the loss of the dominant band representing blocked replication, indicating that proteolytic degradation of the AGT-DNA crosslink allows TLS to occur.

The mechanism of how AGT crosslinks are cytotoxic and mutagenic is still not completely known. Increased cell death can be explained by the lack of repair and ability of replication machinery to bypass the lesion, stalling DNA replication/transcription and thus destabilizing the cell to trigger cell death pathways. Increased DNA mutation is less clear, in that multiple mechanisms could explain the phenomenon. While the results of the kinetic analysis of hPol η did not show a substantial rise in the mutation frequency opposite the 15mer peptide crosslink, this analysis was only done on misincorporation of one of five adducts formed by AGT (92). The N^7-dG adduct is the most abundant of all of AGT crosslinks, but its role is less certain due to the lability of the adduct to hydrolyze to produce an abasic site. However, based on the misincorporation frequency with abasic sites by TLS pols, one would expect to incorporation of dA opposite the abasic site. The other lesions on guanine (the N^1, N^2, and O^6) positions have yet to be studied regarding misincorporation frequency. Based on the location of the N^2-dG lesion in the minor groove of DNA, it should be susceptible to bypass by hPol κ instead of hPol η, and there may be a different result compared to what was found with the N^6-dA adduct.

An alternative hypothesis would be that the mutations arise from complications during DNA repair, notably occurring during DNA replication. Complex DNA repair pathways involving homologous recombination or nonhomologous end-joining could be the cause for these mutations due to the complexity of double strand DNA breaks. However, these complex pathways would not be likely to form the base pair mutations observed with ethylene dibromide (93,131-133). Other repair by pathways such as nucleotide excision repair would require proteolytic degradation by a nuclear protease such as SPRTN prior to being able to access the DNA for repair. During DNA replication, this process could be cumbersome and require an extended period of time to sufficiently degrade the protein. AGT may be unique in its complexion or size after SPRTN degradation and still render the cell unable to be repaired by NER, causing a longer delay before the cell switches to an alternative pathway for repair. Either of these events could result in a lengthy stalled replication forks, causing additional damage to occur on the surrounding DNA and resulting in DNA mutations. Another possibility is that AGT is unique in that SPRTN is unable to digest AGT crosslinks on DNA, but the results from Chapter IV show that SPRTN is capable of degrading AGT, at least *in vitro.*

During the work on this dissertation, multiple strategies for the chemical synthesis and purification of DNA crosslinks to either peptides or proteins have been attempted. The initial strategy, performed in Chapter III, utilized a simple nucleophilic substitution reaction to react a thiol with an alkyl halide containing a terminal amine. This substituted peptide would then react with the oligonucleotide containing the 6-chloropurine nucleobase to form the crosslink, performing another nucleophilic substitution reaction of the free amine with the chlorine. This strategy worked well using the peptide GSH, but attempts using the longer peptides from the active site of AGT proved difficult in that the alkyl halide also reacted with histidine residues. The second strategy for the AGT peptides (Chapter IV) used MSH to convert the thiol of the peptide to a dehydroalanine. The DNA oligonucleotide containing the 6-chloropurine was reacted with cystamine to perform a nucleophilic substitution reaction, as described above prior to reduction

and reaction with the dehydroalanine-containing peptide. The yield of this reaction was lower compared to the first reaction but proved to be specific for the adduction at the cysteine residue. For the synthesis of the AGT-DNA crosslink, neither technique proved successful. Both alkylation and conversion to a dehydroalanine of the active site cysteine caused the protein to precipitate, preventing viable crosslink formation. The third technique, using the click chemistry linker (first treating the DNA with propargyl amine and subsequent addition of the iodoacetamide to the linker), was successful in the reaction with the protein, although with low (~10%) yields.

Several different approaches were also used for purification of the crosslinks. The GSH crosslink was purified by gel electrophoresis because attempts with HPLC proved unsuccessful from separating the crosslinked DNA from the starting material. The AGT peptide crosslinks were purified using an HPLC protocol with a gradient of aqueous 100 mM triethylammonium acetate (pH 7.6) to 100 mM triethylammonium acetate (pH 7.6) in acetonitrile to separate the oligonucleotide-peptide crosslinks, resulting in better yields compared to gel recovery.

5.2 Future Directions

Difficulties in the synthesis and purification of the site-specific AGT-DNA crosslink have been challenging. All purification techniques tested to separate the crosslink from free DNA have been shown to be insufficient to completely remove all unreacted DNA and have also been plagued by the constant loss of product during purification steps. This lack of removal of free oligonucleotide has caused problems when trying to determine the effect of the proteolytic degradation of AGT by SPRTN to allow translesion synthesis past the protein crosslink. Proteomic analysis of the AGT peptides bound to DNA after this degradation has also proved to be difficult. Part of this problem has been in the recovery of peptides. Presently, it is unclear if SPRTN cleaves proteins by cleaving nonspecifically at the ends of the protein as opposed to the middle. If the cleavage products are short peptides or single amino acids, proteomic analysis could prove

troublesome for detecting peptides. As well, nonspecific cleavage products could cause inconsistent results for a specific size or identity for the remaining peptide.

Studying the effects of protein and peptide crosslinks at the N^6 position of adenosine have been useful in determining how mutations arise from these crosslinks. With the availability of the 2-fluoroinosine oligonucleotide, studying the role of crosslinked peptides at this position would also be useful in studying the misincorporation by both GSH and AGT peptide crosslinks. Due to the position of N^2 in the minor groove of DNA compared the N^6 in the major groove, it would be of interest to determine if there is a different effect. It is assumed that hPol κ would play a more dominant role in the bypass of these crosslinks, and a comparison of the misincorporation rates of hPol κ and hPol η would be valuable. Studying the effects of these crosslinks at the N^1 and O^6 positions of guanine would also be of interest, but plans for their synthesis have not been developed yet.

Studying the N^7-dG crosslinks would also be of interest in that these are the major crosslinks formed with both AGT and GSH. Synthesizing site-specific crosslinks at this site may be feasible for study by utilizing 2′-deoxy-2′-fluoroarabinonucleic acid to stabilize the cation. Studying the ability of translesion DNA polymerases to bypass lesions is of great interest, not only of peptide crosslinks but using even the methyl adduct. Determining the bypass here would be of great interest, but further analysis of the stability of the N^7-dG crosslinks in cells would be of greater interest in determining the rate of hydrolysis to abasic sites *in vivo*. This would be necessary in determining the requirement of the N^7-dG crosslinks to be bypassed or repaired versus the cell needing to repair the increased number of abasic sites instead.

Another possible direction of future interest would be the study of the bypass and misincorporation of other proteins that are found to become crosslinked to DNA, such as GAPDH or histone H2b. SPRTN degradation of these proteins could be quantitated and synthetic peptides could be designed to mimic the bypass of peptide crosslinks at various sites on DNA. If the bypass of these peptides proved significantly different compared to AGT and GSH, it could lead to the

conclusion that the detrimental effects of the AGT peptide crosslinks are sequence specific as compared to size or vice versa.

A different potential direction for this project involves the synergistic effects of the DNA replication machinery in bypassing DNA-peptide crosslinks. One approach would be to perform the *in vitro* replication assays for the crosslink bypass assays using multiple polymerases. This may be useful in determining a role for hPol ι in replicating opposite the crosslink prior to switching to another polymerase such as hPol κ, as shown for other DNA lesions (162). While this type of experiment is not very useful in quantitative analysis (e.g., steady-state kinetics), it could elucidate a secondary bypass mechanism for hPol ι, which has not been shown to be capable of performing full primer extension itself. The addition of hPol ζ or REV1 can also be done with these polymerases to elucidate their roles in extension past the crosslink. Analysis of the synergistic effects of the polymerases in bypass of the crosslink could, in principle, be performed in cells. Transfecting a plasmid containing a site-specific crosslink into mammalian cells could provide information on how the components of the replication machinery work together. Determining individual roles of the polymerases could be done by knockdown or knockout assays using siRNA or CRISPR-Cas9 protocols. After allowing the cells to finish replication of the transfected plasmid, the plasmid could be extracted and amplified in a bacterial system prior to DNA sequencing. This technique has been used or the analysis of mutation frequencies for various DNA lesions (148,163-165) and could be utilized for studying various protein crosslinks as well.

The discovery of SPRTN as the nuclear protease was seminal in determining the role of proteolytic degradation of DPCs. Because the protease is known to degrade DPCs, it can be determined how the proteolytic degradation of the crosslinked proteins affects repair and bypass of DPCs. Further experiments can be performed to determine the role of SPRTN to affect the cytotoxic and mutagenic effects of AGT-DNA crosslinks as seen earlier (93,131). It would be of interest to determine the effect of overexpression versus knockdown of SPRTN on the detrimental effects in cells expressing AGT and treated with dibromoethane. Further, proteomic analysis has

been performed to determine which proteins can form DPCs in cells (85,87). It may be of interest to do the analysis with and without the expression of SPRTN to determine if there are any changes in the relative quantitation of certain proteins detected. If one protein were to be detected at much higher levels in the SPRTN-expressing cells, this protein could be resistant to SPRTN degradation and would have the potential to have more severe effects when forming DPCs.

Finally, crosslinking agents such as cisplatin are commonly used in chemotherapeutic treatments. While this treatment is normally associated with forming DNA-DNA interstrand crosslinks, they will form DPCs in these patients. Known mutation of SPRTN is involved in Ruijs-Aalfs Syndrome, and it would be interesting to know if people with this condition or other variants of SPRTN have different sensitivity to cisplatin and other similar chemotherapeutic agents.

REFERENCES

1. Meselson, M., and Stahl, F. W. (1958) The Replication of DNA in *Escherichia coli. Proc. Natl. Acad. Sci. U. S. A.* 44, 671-682.
2. Kunkel, T. A., and Bebenek, K. (2000) DNA replication fidelity. *Annu. Rev. Biochem.* 69, 497-529.
3. Joyce, C. M., and Steitz, T. A. (1994) Function and structure relationships in DNA polymerases. *Annu. Rev. Biochem.* 63, 777-822.
4. Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Jr., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., and et al. (1993) Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. *Proc. Natl. Acad. Sci. U. S. A.* 90, 6320-6324.
5. Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992) Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* 256, 1783-1790.
6. Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., and Steitz, T. A. (1985) Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* 313, 762-766.
7. Guengerich, F. P. (2006) Interactions of carcinogen-bound DNA with individual DNA polymerases. *Chem. Rev.* 106, 420-452.
8. Steitz, T. A. (1999) DNA polymerases: structural diversity and common mechanisms. *J. Biol. Chem.* 274, 17395-17398.
9. van Loon, B., Woodgate, R., and Hubscher, U. (2015) DNA polymerases: Biology, diseases and biomedical applications. *DNA Repair (Amst)* 29, 1-3.
10. Choi, J. Y., and Guengerich, F. P. (2004) Analysis of the effect of bulk at N^2-alkylguanine DNA adducts on catalytic efficiency and fidelity of the processive DNA polymerases bacteriophage T7 exonuclease- and HIV-1 reverse transcriptase. *J. Biol. Chem.* 279, 19217-19229.
11. Loeb, L. A. (1991) Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res.* 51, 3075-3079.
12. Albertson, T. M., Ogawa, M., Bugni, J. M., Hays, L. E., Chen, Y., Wang, Y., Treuting, P. M., Heddle, J. A., Goldsby, R. E., and Preston, B. D. (2009) DNA polymerase epsilon and delta proofreading suppress discrete mutator and cancer phenotypes in mice. *Proc. Natl. Acad. Sci. U. S. A.* 106, 17101-17104.
13. Waga, S., Bauer, G., and Stillman, B. (1994) Reconstitution of complete SV40 DNA replication with purified replication factors. *J. Biol. Chem.* 269, 10923-10934.
14. Singh, B., Li, X., Owens, K. M., Vanniarajan, A., Liang, P., and Singh, K. K. (2015) Human REV3 DNA Polymerase ζ Localizes to Mitochondria and Protects the Mitochondrial Genome. *PLoS One* 10, e0140409.
15. Makarova, A. V., and Burgers, P. M. (2015) Eukaryotic DNA polymerase zeta. *DNA Repair (Amst)* 29, 47-55.

16. Baltimore, D. (1970) RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 226, 1209-1211.

17. Temin, H. M., and Mizutani, S. (1970) RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* 226, 1211-1213.

18. Wood, R. D., and Doublie, S. (2016) DNA polymerase theta (POLQ), double-strand break repair, and cancer. *DNA Repair (Amst)* 44, 22-32.

19. Mateos-Gomez, P. A., Gong, F., Nair, N., Miller, K. M., Lazzerini-Denchi, E., and Sfeir, A. (2015) Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination. *Nature* 518, 254-257.

20. Martomo, S. A., Saribasak, H., Yokoi, M., Hanaoka, F., and Gearhart, P. J. (2008) Reevaluation of the role of DNA polymerase theta in somatic hypermutation of immunoglobulin genes. *DNA Repair (Amst)* 7, 1603-1608.

21. Takata, K., Shimizu, T., Iwai, S., and Wood, R. D. (2006) Human DNA polymerase N (POLN) is a low fidelity enzyme capable of error-free bypass of 5S-thymine glycol. *J. Biol. Chem.* 281, 23445-23455.

22. Yamanaka, K., Minko, I. G., Takata, K., Kolbanovskiy, A., Kozekov, I. D., Wood, R. D., Rizzo, C. J., and Lloyd, R. S. (2010) Novel enzymatic function of DNA polymerase nu in translesion DNA synthesis past major groove DNA-peptide and DNA-DNA cross-links. *Chem. Res. Toxicol.* 23, 689-695.

23. Belousova, E. A., and Lavrik, O. I. (2015) DNA polymerases beta and lambda and their roles in cell. *DNA Repair (Amst)* 29, 112-126.

24. Matsumoto, Y., and Kim, K. (1995) Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair. *Science* 269, 699-702.

25. Gu, H., Marth, J. D., Orban, P. C., Mossmann, H., and Rajewsky, K. (1994) Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science* 265, 103-106.

26. Blanca, G., Shevelev, I., Ramadan, K., Villani, G., Spadari, S., Hubscher, U., and Maga, G. (2003) Human DNA polymerase lambda diverged in evolution from DNA polymerase beta toward specific Mn(++) dependence: a kinetic and thermodynamic study. *Biochemistry* 42, 7467-7476.

27. Lee, J. W., Blanco, L., Zhou, T., Garcia-Diaz, M., Bebenek, K., Kunkel, T. A., Wang, Z., and Povirk, L. F. (2004) Implication of DNA polymerase lambda in alignment-based gap filling for nonhomologous DNA end joining in human nuclear extracts. *J. Biol. Chem.* 279, 805-811.

28. Fan, W., and Wu, X. (2004) DNA polymerase lambda can elongate on DNA substrates mimicking non-homologous end joining and interact with XRCC4-ligase IV complex. *Biochem. Biophys. Res. Commun.* 323, 1328-1333.

29. Covo, S., Blanco, L., and Livneh, Z. (2004) Lesion bypass by human DNA polymerase mu reveals a template-dependent, sequence-independent nucleotidyl transferase activity. *J. Biol. Chem.* 279, 859-865.

30. Gilfillan, S., Dierich, A., Lemeur, M., Benoist, C., and Mathis, D. (1993) Mice lacking TdT: mature animals with an immature lymphocyte repertoire. *Science* 261, 1175-1178.

31. Trincao, J., Johnson, R. E., Escalante, C. R., Prakash, S., Prakash, L., and Aggarwal, A. K. (2001) Structure of the catalytic core of *S. cerevisiae* DNA polymerase eta: implications for translesion DNA synthesis. *Mol. Cell* 8, 417-426.

32. Wilson, R. C., Jackson, M. A., and Pata, J. D. (2013) Y-family polymerase conformation is a major determinant of fidelity and translesion specificity. *Structure* 21, 20-31.

33. Boudsocq, F., Kokoska, R. J., Plosky, B. S., Vaisman, A., Ling, H., Kunkel, T. A., Yang, W., and Woodgate, R. (2004) Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. *J. Biol. Chem.* 279, 32932-32940.

34. Naryzhny, S. N. (2008) Proliferating cell nuclear antigen: a proteomics view. *Cell. Mol. Life Sci.* 65, 3789-3808.

35. Szuts, D., Marcus, A. P., Himoto, M., Iwai, S., and Sale, J. E. (2008) REV1 restrains DNA polymerase zeta to ensure frame fidelity during translesion synthesis of UV photoproducts *in vivo. Nucleic Acids Res.* 36, 6767-6780.

36. Edmunds, C. E., Simpson, L. J., and Sale, J. E. (2008) PCNA ubiquitination and REV1 define temporally distinct mechanisms for controlling translesion synthesis in the avian cell line DT40. *Mol. Cell* 30, 519-529.

37. Johnson, R. E., Prakash, S., and Prakash, L. (1999) Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Pol eta. *Science* 283, 1001-1004.

38. Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., Takio, K., and Hanaoka, F. (1999) The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta. *Nature* 399, 700-704.

39. Masutani, C., Kusumoto, R., Iwai, S., and Hanaoka, F. (2000) Mechanisms of accurate translesion synthesis by human DNA polymerase eta. *EMBO J.* 19, 3100-3109.

40. Silverstein, T. D., Johnson, R. E., Jain, R., Prakash, L., Prakash, S., and Aggarwal, A. K. (2010) Structural basis for the suppression of skin cancers by DNA polymerase eta. *Nature* 465, 1039-1043.

41. Zhao, Y., Biertumpfel, C., Gregory, M. T., Hua, Y. J., Hanaoka, F., and Yang, W. (2012) Structural basis of human DNA polymerase eta-mediated chemoresistance to cisplatin. *Proc. Natl. Acad. Sci. U S A* 109, 7269-7274.

42. Irimia, A., Eoff, R. L., Guengerich, F. P., and Egli, M. (2009) Structural and functional elucidation of the mechanism promoting error-prone synthesis by human DNA polymerase kappa opposite the 7,8-dihydro-8-oxo-2'-deoxyguanosine adduct. *J. Biol. Chem.* 284, 22467-22480.

43. Yoon, J. H., Bhatia, G., Prakash, S., and Prakash, L. (2010) Error-free replicative bypass of thymine glycol by the combined action of DNA polymerases kappa and zeta in human cells. *Proc. Natl. Acad. Sci. U S A* 107, 14116-14121.

44. Bose, A., Millsap, A. D., DeLeon, A., Rizzo, C. J., and Basu, A. K. (2016) Translesion Synthesis of the *N*(2)-2'-Deoxyguanosine Adduct of the Dietary Mutagen IQ in Human Cells: Error-Free Replication by DNA Polymerase kappa and Mutagenic Bypass by DNA Polymerases η, ζ, and Rev1. *Chem. Res. Toxicol.* 29, 1549-1559.

45. Suzuki, N., Ohashi, E., Kolbanovskiy, A., Geacintov, N. E., Grollman, A. P., Ohmori, H., and Shibutani, S. (2002) Translesion synthesis by human DNA polymerase kappa on a DNA template containing a single stereoisomer of dG-(+)- or dG-(-)-anti-N(2)-BPDE (7,8-

dihydroxy-anti-9,10-epoxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene). *Biochemistry* 41, 6100-6106.

46. Rechkoblit, O., Zhang, Y., Guo, D., Wang, Z., Amin, S., Krzeminsky, J., Louneva, N., and Geacintov, N. E. (2002) Trans-lesion synthesis past bulky benzo[*a*]pyrene diol epoxide N2-dG and N6-dA lesions catalyzed by DNA bypass polymerases. *J. Biol. Chem.* 277, 30488-30494.

47. Ogi, T., and Lehmann, A. R. (2006) The Y-family DNA polymerase kappa (pol kappa) functions in mammalian nucleotide-excision repair. *Nat. Cell Biol.* 8, 640-642.

48. Ogi, T., Limsirichaikul, S., Overmeer, R. M., Volker, M., Takenaka, K., Cloney, R., Nakazawa, Y., Niimi, A., Miki, Y., Jaspers, N. G., Mullenders, L. H., Yamashita, S., Fousteri, M. I., and Lehmann, A. R. (2010) Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. *Mol. Cell* 37, 714-727.

49. Williams, H. L., Gottesman, M. E., and Gautier, J. (2012) Replication-independent repair of DNA interstrand crosslinks. *Mol. Cell* 47, 140-147.

50. Nair, D. T., Johnson, R. E., Prakash, S., Prakash, L., and Aggarwal, A. K. (2004) Replication by human DNA polymerase-iota occurs by Hoogsteen base-pairing. *Nature* 430, 377-380.

51. Frank, E. G., and Woodgate, R. (2007) Increased catalytic activity and altered fidelity of human DNA polymerase iota in the presence of manganese. *J. Biol. Chem.* 282, 24689-24696.

52. Pence, M. G., Blans, P., Zink, C. N., Hollis, T., Fishbein, J. C., and Perrino, F. W. (2009) Lesion bypass of N2-ethylguanine by human DNA polymerase iota. *J. Biol. Chem.* 284, 1732-1740.

53. Tissier, A., McDonald, J. P., Frank, E. G., and Woodgate, R. (2000) Pol iota, a remarkably error-prone human DNA polymerase. *Genes Dev.* 14, 1642-1650.

54. Zhang, Y., Yuan, F., Wu, X., and Wang, Z. (2000) Preferential incorporation of G opposite template T by the low-fidelity human DNA polymerase iota. *Mol. Cell. Biol.* 20, 7099-7108.

55. Haracska, L., Prakash, S., and Prakash, L. (2002) Yeast Rev1 protein is a G template-specific DNA polymerase. *J. Biol. Chem.* 277, 15546-15551.

56. Zhang, Y., Wu, X., Rechkoblit, O., Geacintov, N. E., Taylor, J. S., and Wang, Z. (2002) Response of human REV1 to different DNA damage: preferential dCMP insertion opposite the lesion. *Nucleic Acids Res.* 30, 1630-1638.

57. Washington, M. T., Minko, I. G., Johnson, R. E., Haracska, L., Harris, T. M., Lloyd, R. S., Prakash, S., and Prakash, L. (2004) Efficient and error-free replication past a minor-groove N^2-guanine adduct by the sequential action of yeast Rev1 and DNA polymerase zeta. *Mol. Cell. Biol.* 24, 6900-6906.

58. Nair, D. T., Johnson, R. E., Prakash, L., Prakash, S., and Aggarwal, A. K. (2008) Protein-template-directed synthesis across an acrolein-derived DNA adduct by yeast Rev1 DNA polymerase. *Structure* 16, 239-245.

59. Guo, C., Fischhaber, P. L., Luk-Paszyc, M. J., Masuda, Y., Zhou, J., Kamiya, K., Kisker, C., and Friedberg, E. C. (2003) Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. *EMBO J.* 22, 6621-6630.

60. Ohashi, E., Murakumo, Y., Kanjo, N., Akagi, J., Masutani, C., Hanaoka, F., and Ohmori, H. (2004) Interaction of hREV1 with three human Y-family DNA polymerases. *Genes Cells* 9, 523-531.

61. Tissier, A., Kannouche, P., Reck, M. P., Lehmann, A. R., Fuchs, R. P., and Cordonnier, A. (2004) Co-localization in replication foci and interaction of human Y-family members, DNA polymerase pol eta and REVl protein. *DNA Repair (Amst)* 3, 1503-1514.

62. 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.

63. Sale, J. E., Lehmann, A. R., and Woodgate, R. (2012) Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nat. Rev. Mol. Cell Biol.* 13, 141-152.

64. Grunberger, D., and Weinstein, I. B. (1979) Biochemical effects of the modification of nucleic acids by certain polycyclic aromatic carcinogens. *Prog. Nucleic Acid Res. Mol. Biol.* 23, 105-149.

65. Perera, F. P., Santella, R. M., Brenner, D., Young, T. L., and Weinstein, I. B. (1988) Application of biological markers to the study of lung cancer causation and prevention. *IARC Sci Publ*, 451-459.

66. Lerman, L. S. (1961) Structural considerations in the interaction of DNA and acridines. *J. Mol. Biol.* 3, 18-30.

67. Hayes, J. D., and McLellan, L. I. (1999) Glutathione and glutathione-dependent enzymes represent a co-ordinately regulated defence against oxidative stress. *Free Radic Res* 31, 273-300.

68. Tano, K., Shiota, S., Collier, J., Foote, R. S., and Mitra, S. (1990) Isolation and structural characterization of a cDNA clone encoding the human DNA repair protein for O^6-alkylguanine. *Proc. Natl. Acad. Sci. U S A* 87, 686-690.

69. Scharer, O. D. (2003) Chemistry and biology of DNA repair. *Angewandte Chemie* 42, 2946-2974.

70. Ross, W. C. J. (1962) *Relative Reactivity of Different Nucleophilic Centres Towards Alkylating Agents*, Butterworths, London.

71. Weisburger, E. K. (1977) Carcinogenicity studies on halogenated hydrocarbons. *Environ. Health Perspect.* 21, 7-16.

72. Kitchin, K. T., Brown, J. L., and Kulkarni, A. P. (1993) Predicting rodent carcinogenicity of halogenated hydrocarbons by in vivo biochemical parameters. *Teratog. Carcinog. Mutagen.* 13, 167-184.

73. Oliveira, P. A., Colaco, A., Chaves, R., Guedes-Pinto, H., De-La-Cruz, P. L., and Lopes, C. (2007) Chemical carcinogenesis. *An. Acad. Bras. Cienc.* 79, 593-616.

74. Morita, R., Nakane, S., Shimada, A., Inoue, M., Iino, H., Wakamatsu, T., Fukui, K., Nakagawa, N., Masui, R., and Kuramitsu, S. (2010) Molecular mechanisms of the whole DNA repair system: a comparison of bacterial and eukaryotic systems. *J. Nucleic Acids* 2010, 179594.

75. Gilman, A. (1963) The initial clinical trial of nitrogen mustard. *Am. J. Surg.* 105, 574-578.

76. Gold, L. S., Backman, G. M., Hooper, N. K., and Peto, R. (1987) Ranking the potential carcinogenic hazards to workers from exposures to chemicals that are tumorigenic in rodents. *Environ. Health Perspect.* 76, 211-219.

77. Huff, J. E. (1983) 1,2-Dibromoethane (ethylene dibromide). *Environ. Health Perspect.* 47, 359-363.

78. Van Duuren, B. L., Goldschmidt, B. M., Loewengart, G., Smith, A. C., Melchionne, S., Seldman, I., and Roth, D. (1979) Carcinogenicity of halogenated olefinic and aliphatic hydrocarbons in mice. *J. Natl. Cancer Inst.* 63, 1433-1439.

79. National Toxicology, P. (1982) Carcinogenesis Bioassay of 1,2-Dibromoethane (CAS No. 106-93-4) in F344 Rats and B6C3F1 Mice (Inhalation Study). *Natl. Toxicol. Program Tech. Rep. Ser.* 210, 1-163.

80. Wong, L. C., Winston, J. M., Hong, C. B., and Plotnick, H. (1982) Carcinogenicity and toxicity of 1,2-dibromoethane in the rat. *Toxicol. Appl. Pharmacol.* 63, 155-165.

81. 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.

82. 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.

83. Fu, Y. V., Yardimci, H., Long, D. T., Ho, T. V., Guainazzi, A., Bermudez, V. P., Hurwitz, J., van Oijen, A., Scharer, O. D., and Walter, J. C. (2011) Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell* 146, 931-941.

84. Kuo, H. K., Griffith, J. D., and Kreuzer, K. N. (2007) 5-Azacytidine induced methyltransferase-DNA adducts block DNA replication *in vivo*. *Cancer Res.* 67, 8248-8254.

85. 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.

86. 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.

87. Ming, X., Groehler, A. t., 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.

88. van Bladeren, P. J., Breimer, D. D., Rotteveel-Smijs, G. M., de Jong, R. A., Buijs, W., van der Gen, A., and Mohn, G. R. (1980) The role of glutathione conjugation in the mutagenicity of 1,2-dibromoethane. *Biochem. Pharmacol.* 29, 2975-2982.

89. Ozawa, N., and Guengerich, F. P. (1983) Evidence for formation of an *S*-[2-(N^7-guanyl)ethyl]glutathione adduct in glutathione-mediated binding of the carcinogen 1,2-dibromoethane to DNA. *Proc. Natl. Acad. Sci. U S A* 80, 5266-5270.

90. Peterson, L. A., Harris, T. M., and Guengerich, F. P. (1988) Evidence for an episulfonium ion intermediate in the formation of *S*-[2-(N^7-guanyl)ethyl]glutathione in DNA. *J. Am. Chem. Soc.* 110, 3284-3291.

91. Abril, N., Luque-Romero, F. L., Christians, F. C., Encell, L. P., Loeb, L. A., and Pueyo, C. (1999) Human O^6-alkylguanine-DNA alkyltransferase: protection against alkylating agents and sensitization to dibromoalkanes. *Carcinogenesis* 20, 2089-2094.

92. Chowdhury, G., Cho, S. H., Pegg, A. E., and Guengerich, F. P. (2013) Detection and characterization of 1,2-dibromoethane-derived DNA crosslinks formed with O^6-alkylguanine-DNA alkyltransferase. *Angewandte Chemie* 52, 12879-12882.

93. Liu, L., Pegg, A. E., Williams, K. M., and Guengerich, F. P. (2002) Paradoxical enhancement of the toxicity of 1,2-dibromoethane by O^6-alkylguanine-DNA alkyltransferase. *J. Biol. Chem.* 277, 37920-37928.

94. Loecken, E. M., and Guengerich, F. P. (2008) Reactions of glyceraldehyde 3-phosphate dehydrogenase sulfhydryl groups with bis-electrophiles produce DNA-protein cross-links but not mutations. *Chem. Res. Toxicol.* 21, 453-458.

95. Loecken, E. M., Dasari, S., Hill, S., Tabb, D. L., and Guengerich, F. P. (2009) The bis-electrophile diepoxybutane cross-links DNA to human histones but does not result in enhanced mutagenesis in recombinant systems. *Chem. Res. Toxicol.* 22, 1069-1076.

96. Debethune, 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.

97. Das, B. B., Huang, S. Y., Murai, J., Rehman, I., Ame, J. C., Sengupta, S., Das, S. K., Majumdar, P., Zhang, H., Biard, D., Majumder, H. K., Schreiber, V., and Pommier, Y. (2014) PARP1-TDP1 coupling for the repair of topoisomerase I-induced DNA damage. *Nucleic Acids Res.* 42, 4435-4449.

98. Zeng, Z., Cortes-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.

99. Cmarik, J. L., Humphreys, W. G., Bruner, K. L., Lloyd, R. S., Tibbetts, C., and Guengerich, F. P. (1992) Mutation spectrum and sequence alkylation selectivity resulting from modification of bacteriophage M13mp18 DNA with *S*-(2-chloroethyl)glutathione. Evidence for a role of *S*-(2-N^7-guanyl)ethyl)glutathione as a mutagenic lesion formed from ethylene dibromide. *J. Biol. Chem.* 267, 6672-6679.

100. Reardon, J. T., and Sancar, A. (2006) Repair of DNA-polypeptide crosslinks by human excision nuclease. *Proc. Natl. Acad. Sci. U S A* 103, 4056-4061.

101. Nakano, T., Katafuchi, A., Matsubara, M., Terato, H., Tsuboi, T., Masuda, T., Tatsumoto, T., Pack, S. P., Makino, K., Croteau, D. L., Van Houten, B., Iijima, K., Tauchi, H., and Ide, H. (2009) Homologous recombination but not nucleotide excision repair plays a pivotal role in tolerance of DNA-protein cross-links in mammalian cells. *J. Biol. Chem.* 284, 27065-27076.

102. 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.

103. 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.

104. 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.

105. Kim, M. S., Machida, Y., Vashisht, A. A., Wohlschlegel, J. A., Pang, Y. P., and Machida, Y. J. (2013) Regulation of error-prone translesion synthesis by Spartan/C1orf124. *Nucleic Acids Res.* 41, 1661-1668.

106. Toth, A., Hegedus, L., Juhasz, S., Haracska, L., and Burkovics, P. (2017) The DNA-binding box of human SPARTAN contributes to the targeting of Poleta to DNA damage sites. *DNA Repair (Amst)* 49, 33-42.

107. 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., Hogel, J., Degoricija, M., Fertig, R., Burkhalter, M. D., Hofmann, K., Thiele, H., Altmuller, J., Nurnberg, G., Nurnberg, P., Bahlo, M., Martin, G. M., Aalfs, C. M., Oshima, J., Terzic, J., Amor, D. J., Dikic, I., Ramadan, K., and Kubisch, C. (2014) Mutations in SPRTN cause early onset hepatocellular carcinoma, genomic instability and progeroid features. *Nat. Genet.* 46, 1239-1244.

108. 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.

109. Machida, Y., Kim, M. S., and Machida, Y. J. (2012) Spartan/C1orf124 is important to prevent UV-induced mutagenesis. *Cell Cycle* 11, 3395-3402.

110. 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.

111. 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.

112. 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.

113. Morocz, M., Zsigmond, E., Toth, R., Enyedi, M. Z., Pinter, 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.

114. Zang, H., Goodenough, A. K., Choi, J. Y., Irimia, A., Loukachevitch, L. V., Kozekov, I. D., Angel, K. C., Rizzo, C. J., Egli, M., and Guengerich, F. P. (2005) DNA adduct bypass polymerization by *Sulfolobus solfataricus* DNA polymerase Dpo4: analysis and crystal structures of multiple base pair substitution and frameshift products with the adduct 1,N^2-ethenoguanine. *J. Biol. Chem.* 280, 29750-29764.

115. Pence, M. G., Choi, J. Y., Egli, M., and Guengerich, F. P. (2010) Structural basis for proficient incorporation of dTTP opposite O^6-methylguanine by human DNA polymerase iota. *J. Biol. Chem.* 285, 40666-40672.

116. Patra, A., Nagy, L. D., Zhang, Q., Su, Y., Muller, L., Guengerich, F. P., and Egli, M. (2014) Kinetics, structure, and mechanism of 8-Oxo-7,8-dihydro-2'-deoxyguanosine bypass by human DNA polymerase eta. *J. Biol. Chem.* 289, 16867-16882.

117. Furge, L. L., and Guengerich, F. P. (1997) Analysis of nucleotide insertion and extension at 8-oxo-7,8-dihydroguanine by replicative T7 polymerase exo- and human immunodeficiency virus-1 reverse transcriptase using steady-state and pre-steady-state kinetics. *Biochemistry* 36, 6475-6487.

118. Lowe, L. G., and Guengerich, F. P. (1996) Steady-state and pre-steady-state kinetic analysis of dNTP insertion opposite 8-oxo-7,8-dihydroguanine by *Escherichia coli* polymerases I exo- and II exo. *Biochemistry* 35, 9840-9849.

119. Harris, C. M., Zhou, L., Strand, E. A., and Harris, T. M. (1991) New strategy for the synthesis of oligodeoxynucleotides bearing adducts at exocyclic amino sites of purine nucleosides. *J. Am. Chem. Soc.* 113, 4328-4329.

120. Humphreys, W. G., Kim, D. H., Cmarik, J. L., Shimada, T., and Guengerich, F. P. (1990) Comparison of the DNA-alkylating properties and mutagenic responses of a series of *S*-(2-haloethyl)-substituted cysteine and glutathione derivatives. *Biochemistry* 29, 10342-10350.

121. Edara, S., Kanugula, S., Goodtzova, K., and Pegg, A. E. (1996) Resistance of the human O^6-alkylguanine-DNA alkyltransferase containing arginine at codon 160 to inactivation by O^6-benzylguanine. *Cancer Res.* 56, 5571-5575.

122. Olson, W. A., Habermann, R. T., Weisburger, E. K., Ward, J. M., and Weisburger, J. H. (1973) Induction of stomach cancer in rats and mice by halogenated aliphatic fumigants. *J. Natl. Cancer Inst.* 51, 1993-1995.

123. Ramsey, J. C., Park, C. N., Ott, M. G., and Gehring, P. J. (1979) Carcinogenic risk assessment: ethylene dibromide. *Toxicol. Appl. Pharmacol.* 47, 411-414.

124. Hill, D. L., Shih, T. W., Johnston, T. P., and Struck, R. F. (1978) Macromolecular binding and metabolism of the carcinogen 1,2-dibromoethane. *Cancer Res.* 38, 2438-2442.

125. Rannug, U., and Beije, B. (1979) The mutagenic effect of 1,2-dichloroethane on *Salmonella typhimurium. II*. Activation by the isolated perfused rat liver. *Chem. Biol. Interact.* 24, 265-285.

126. Rannug, U., Sundvall, A., and Ramel, C. (1978) The mutagenic effect of 1,2-dichloroethane on *Salmonella typhimurium I*. Activation through conjugation with glutathion *in vitro*. *Chem. Biol. Interact.* 20, 1-16.

127. Rannug, U. (1980) Genotoxic effects of 1,2-dibromoethane and 1,2-dichloroethane. *Mutat. Res.* 76, 269-295.

128. Wheeler, J. B., Stourman, N. V., Armstrong, R. N., and Guengerich, F. P. (2001) Conjugation of haloalkanes by bacterial and mammalian glutathione transferases: mono- and vicinal dihaloethanes *Chem. Res. Toxicol.* 14, 1107-1117.

129. Kim, D. H., Humphreys, W. G., and Guengerich, F. P. (1990) Characterization of S-[2-(N^1-adenyl)ethyl]glutathione as an adduct formed in RNA and DNA from 1,2-dibromoethane. *Chem. Res. Toxicol.* 3, 587-594.

130. Cho, S. H., and Guengerich, F. P. (2013) *In vivo* roles of conjugation with glutathione and O^6-alkylguanine DNA-alkyltransferase in the mutagenicity of the bis-electrophiles 1,2-dibromoethane and 1,2,3,4-diepoxybutane in mice. *Chem. Res. Toxicol.* 26, 1765-1774.

131. Abril, N., and Margison, G. P. (1999) Mammalian cells expressing *Escherichia coli* O^6-alkylguanine-DNA alkyltransferases are hypersensitive to dibromoalkanes. *Chem. Res. Toxicol.* 12, 544-551.

132. Abril, N., Luque-Romero, F. L., Prieto-Alamo, M.-J., Rafferty, J. A., Margison, G. P., and Pueyo, C. (1997) Bacterial and mammalian DNA alkyltransferases sensitize *Escherichia coli* to the lethal and mutagenic effects of dibromoalkanes. *Carcinogenesis* 18, 1883-1888.

133. Abril, N., Luqueromero, F. L., Prieto-Alamo, M. J., Margison, G. P., and Pueyo, C. (1995) *ogt* alkyltransferase enhances dibromoalkane mutagenicity in excision repair-deficient *Escherichia coli* K-12. *Mol. Carcinog.* 12, 110-117.

134. Liu, L., Hachey, D. L., Valadez, G., Williams, K. M., Guengerich, F. P., Loktionova, N. A., Kanugula, S., and Pegg, A. E. (2004) Characterization of a mutagenic DNA adduct formed from 1,2-dibromoethane by O^6-alkylguanine-DNA alkyltransferase. *J. Biol. Chem.* 279, 4250-4259.

135. Dieckhaus, C. M., Fernandez-Metzler, C. L., King, R., Krolikowski, P. H., and Baillie, T. A. (2005) Negative ion tandem mass spectrometry for the detection of glutathione conjugates. *Chem. Res. Toxicol.* 18, 630-638.

136. Kim, M. S., and Guengerich, F. P. (1997) Synthesis of oligonucleotides containing the ethylene dibromide-derived DNA adducts *S*-[2-(N^7-guanyl)ethyl]glutathione, *S*-[2-(N^2-guanyl)ethyl]glutathione, and *S*-[2-(O^6-guanyl)ethyl]glutathione at a single site. *Chem. Res. Toxicol.* 10, 1133-1143.

137. Christov, P. P., Angel, K. C., Guengerich, F. P., and Rizzo, C. J. (2009) Replication past the N^5-methyl-formamidopyrimidine lesion of deoxyguanosine by DNA polymerases and an improved procedure for sequence analysis of in vitro bypass products by mass spectrometry. *Chem. Res. Toxicol.* 22, 1086-1095.

138. Guengerich, F. P. (2005) Activation of alkyl halides by glutathione transferases. *Methods Enzymol.* 401, 342-353.

139. Koga, N., Inskeep, P. B., Harris, T. M., and Guengerich, F. P. (1986) *S*-[2-(N^7-Guanyl)ethyl]glutathione, the major DNA adduct formed from 1,2-dibromoethane. *Biochemistry* 25, 2192-2198.

140. Inskeep, P. B., Koga, N., Cmarik, J. L., and Guengerich, F. P. (1986) Covalent binding of 1,2-dihaloalkanes to DNA and stability of the major DNA adduct, *S*-[2-(N^7-guanyl)ethyl]glutathione. *Cancer Res.* 46, 2839-2844.

141. Cho, S. H., and Guengerich, F. P. (2013) Replication past the butadiene diepoxide-derived DNA adduct *S*-[4-(N^6-deoxyadenosinyl)-2,3-dihydroxybutyl]glutathione by DNA polymerases. *Chem. Res. Toxicol.* 26, 1005-1013.

142. Choi, J. Y., Chowdhury, G., Zang, H., Angel, K. C., Vu, C. C., Peterson, L. A., and Guengerich, F. P. (2006) Translesion synthesis across O^6-alkylguanine DNA adducts by recombinant human DNA polymerases. *J. Biol. Chem.* 281, 38244-38256.

143. Wickramaratne, S., Ji, S., Mukherjee, S., Su, Y., Pence, M. G., Lior-Hoffmann, L., Fu, I., Broyde, S., Guengerich, F. P., Distefano, M., Scharer, O. D., Sham, Y. Y., and Tretyakova, N. (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.

144. Walsh, J. M., Ippoliti, P. J., Ronayne, E. A., Rozners, E., and Beuning, P. J. (2013) Discrimination against major groove adducts by Y-family polymerases of the DinB subfamily. *DNA Repair (Amst)* 12, 713-722.

145. Lawley, P. D. (1984) Carcinogenesis by alkylating agents. in *Chemical Carcinogens* (Searle, C. E. ed.), 2 Ed., Amer. Chem. Soc., Washington, D.C. pp 324-484.

146. Wang, J., and Wang, Y. (2010) Synthesis and characterization of oligodeoxyribonucleotides containing a site-specifically incorporated N^6-carboxymethyl-2'-deoxyadenosine or N4-carboxymethyl-2'-deoxycytidine. *Nucleic Acids Res.* 38, 6774-6784.

147. Swanson, A. L., Wang, J., and Wang, Y. (2011) *In vitro* replication studies of carboxymethylated DNA lesions with *Saccharomyces cerevisiae* polymerase eta. *Biochemistry* 50, 7666-7673.

148. Yuan, B., Wang, J., Cao, H., Sun, R., and Wang, Y. (2011) High-throughput analysis of the mutagenic and cytotoxic properties of DNA lesions by next-generation sequencing. *Nucleic Acids Res.* 39, 5945-5954.

149. Osborne, M. R. (1984) DNA interactions of reactive intermediates derived from carcinogens. in *Chemical Carcinogens* (Searle, C. E. ed.), 2 Ed., American Chemical Society, Washington, D.C/. pp 485-524.

150. Khalili, H., Zhang, F. J., Harvey, R. G., and Dipple, A. (2000) Mutagenicity of benzo[*a*]pyrene-deoxyadenosine adducts in a sequence context derived from the p53 gene. *Mutat. Res.* 465, 39-44.

151. Wang, L., Wu, M., Yan, S. F., Patel, D. J., Geacintov, N. E., and Broyde, S. (2005) Accommodation of a 1S-(-)-benzo[c]phenanthrenyl-N^6-dA adduct in the Y-family Dpo4 DNA polymerase active site: structural insights through molecular dynamics simulations. *Chem. Res. Toxicol.* 18, 441-456.

152. Kotapati, S., Wickramaratne, S., Esades, A., Boldry, E. J., Quirk Dorr, D., Pence, M. G., Guengerich, F. P., and Tretyakova, N. Y. (2015) Polymerase Bypass of N^6-Deoxyadenosine Adducts Derived from Epoxide Metabolites of 1,3-Butadiene. *Chem. Res. Toxicol.* 28, 1496-1507.

153. Maddukuri, L., Shuck, S. C., Eoff, R. L., Zhao, L., Rizzo, C. J., Guengerich, F. P., and Marnett, L. J. (2013) Replication, repair, and translesion polymerase bypass of N^6-oxopropenyl-2'-deoxyadenosine. *Biochemistry* 52, 8766-8776.

154. Kim, M. S., and Guengerich, F. P. (1998) Polymerase blockage and misincorporation of dNTPs opposite the ethylene dibromide-derived DNA adducts *S*-[2-(N^7-guanyl)ethyl]glutathione, S-[2-(N2-guanyl)ethyl]glutathione, and *S*-[2-(O^6-guanyl)ethyl]glutathione. *Chem. Res. Toxicol.* 11, 311-316.

155. Sung, J. S., DeMott, M. S., and Demple, B. (2005) Long-patch base excision DNA repair of 2-deoxyribonolactone prevents the formation of DNA-protein cross-links with DNA polymerase beta. *J. Biol. Chem.* 280, 39095-39103.

156. 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.

157. 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.

158. Sedgeman, C. A., Su, Y., and Guengerich, F. P. (2017) Formation of *S*-[2-(N^6-Deoxyadenosinyl)ethyl]glutathione in DNA and Replication Past the Adduct by Translesion DNA Polymerases. *Chem. Res. Toxicol.* 30, 1188-1196.

159. Kalapila, A. G., and Pegg, A. E. (2010) Alkyltransferase-mediated toxicity of *bis*-electrophiles in mammalian cells. *Mutat. Res.* 684, 35-42.

160. 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 C^5 of thymidine. *J. Biol. Chem.* 290, 775-787.

161. Pande, P., Ji, S., Mukherjee, S., Scharer, O. D., Tretyakova, N. Y., and Basu, A. K. (2017) Mutagenicity of a Model DNA-Peptide Cross-Link in Human Cells: Roles of Translesion Synthesis DNA Polymerases. *Chem. Res. Toxicol.* 30, 669-677.

162. Wolfle, W. T., Johnson, R. E., Minko, I. G., Lloyd, R. S., Prakash, S., and Prakash, L. (2005) Human DNA polymerase iota promotes replication through a ring-closed minor-groove adduct that adopts a syn conformation in DNA. *Mol. Cell. Biol.* 25, 8748-8754.

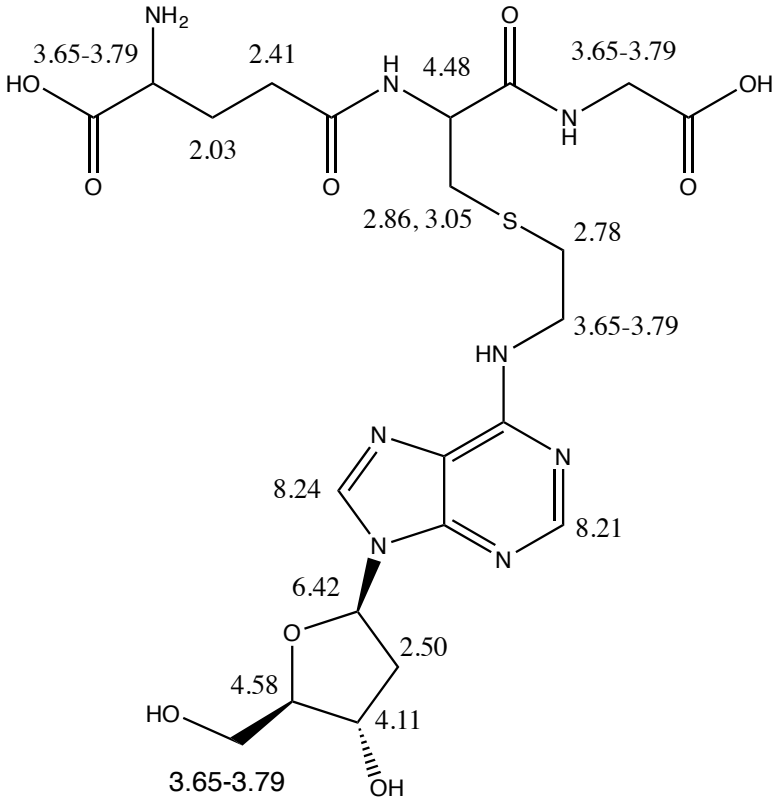
163. Chang, S. C., Fedeles, B. I., Wu, J., Delaney, J. C., Li, D., Zhao, L., Christov, P. P., Yau, E., Singh, V., Jost, M., Drennan, C. L., Marnett, L. J., Rizzo, C. J., Levine, S. S., Guengerich, F. P., and Essigmann, J. M. (2015) Next-generation sequencing reveals the biological significance of the N^2,3-ethenoguanine lesion *in vivo*. *Nucleic Acids Res.* 43, 5489-5500.

164. Lee, D. F., Lu, J., Chang, S., Loparo, J. J., and Xie, X. S. (2016) Mapping DNA polymerase errors by single-molecule sequencing. *Nucleic Acids Res.* 44, e118.

165. de Paz, A. M., Cybulski, T. R., Marblestone, A. H., Zamft, B. M., Church, G. M., Boyden, E. S., Kording, K. P., and Tyo, K. E. J. (2018) High-resolution mapping of DNA polymerase fidelity using nucleotide imbalances and next-generation sequencing. *Nucleic Acids Res.*

APPENDIX

LIST OF FIGURES



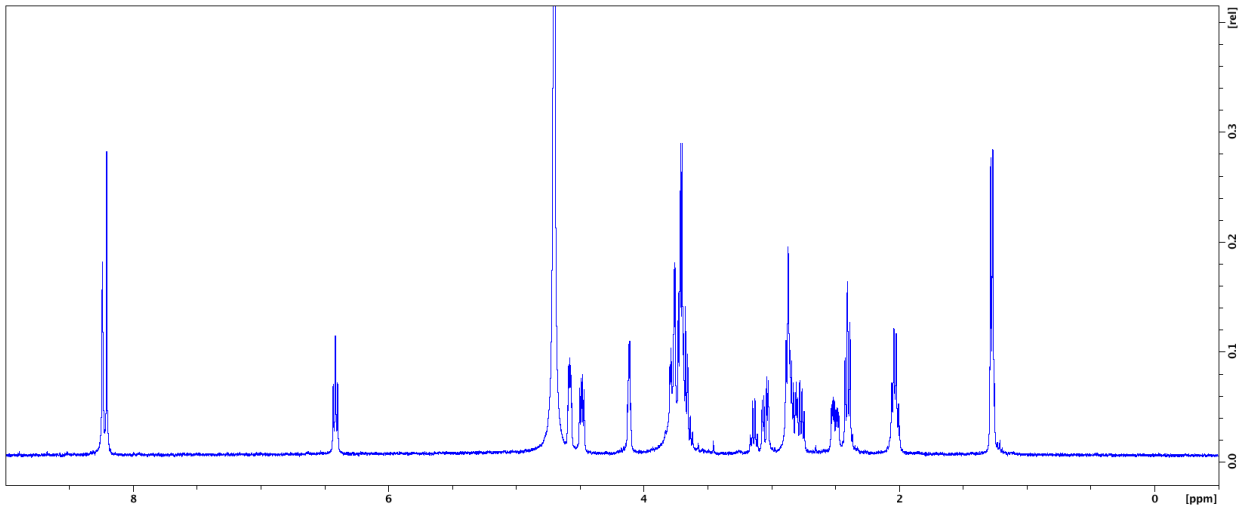
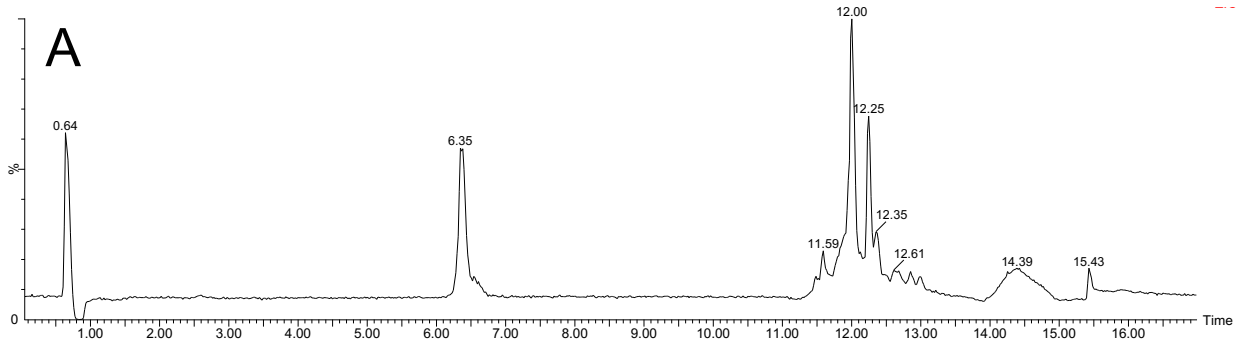
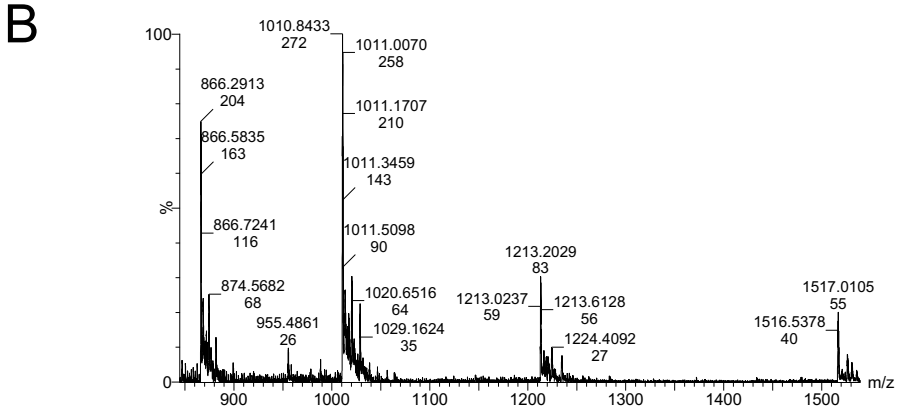


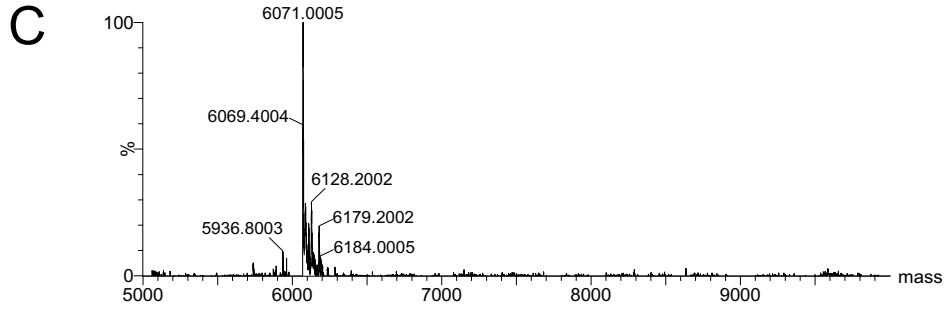
Figure A.1 ^{1}H-NMR spectrum of *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH standard in $^{2}H_2O$ (400 MHz).



t_R, min



m/z



m/z

Figure A.2 LC-MS Characterization of the *S*-[2-(*N*^{6}-deoxyadenosinyl)ethyl]GSH-containing Oligonucleotide. UPLC chromatogram (A), mass spectrum of product at 6.35 min (B), and deconvoluted mass spectrum of product (C).

```
5′- ATGGAACGCGCACGCCGTCCTGCATAATACTTTGAGAGAATGGCAATCCA -3′
3' -TACCTTGCGCGTGCGGCAGGACGTATTATGAAACTCTCTTACCGTTAGGT -5′
```

Figure A.3 DNA Sequence Identity for Randomized 50mer DNA Duplex. Sequence was generated randomly to contain a G:C content of 50%.

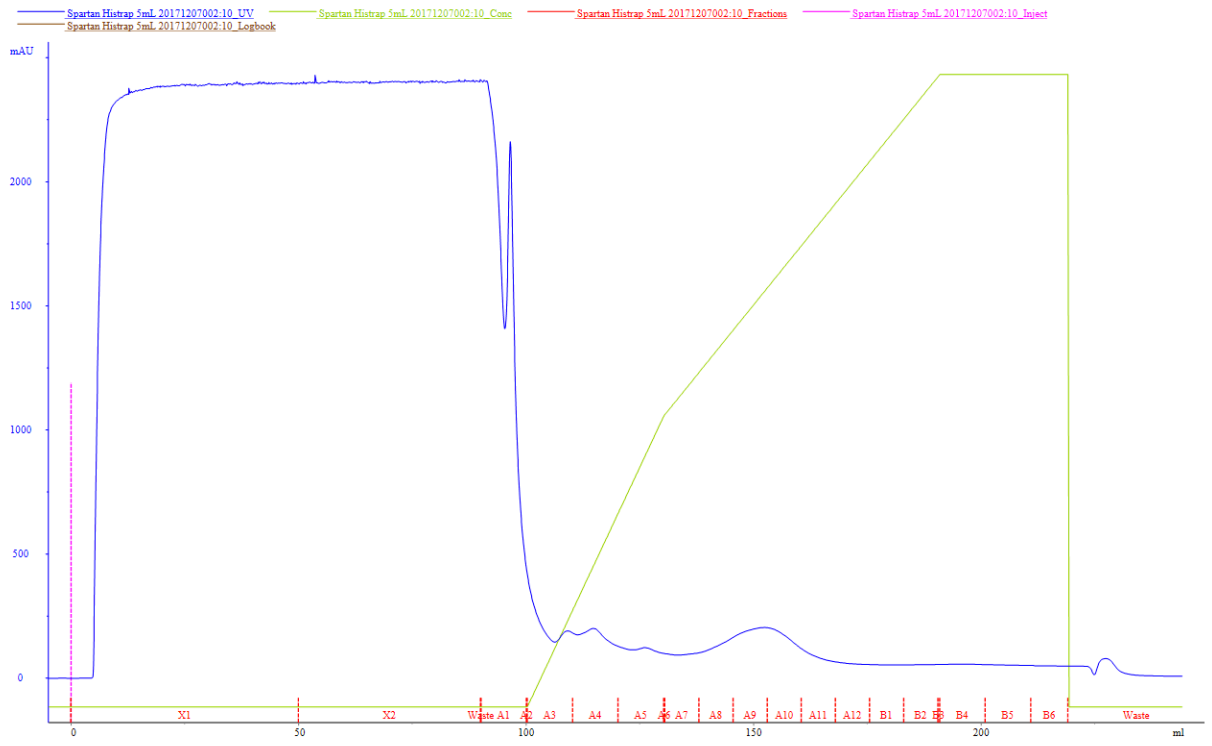


Figure A.4 First Stage Purification of SPRTN. SPRTN was initially purified using a Ni^{2+}-NTA FPLC column using a gradient from 10 mM to 300 mM imidazole. Elution of the protein occurred at 150 mL. UV absorbance was monitored at 280 nm.

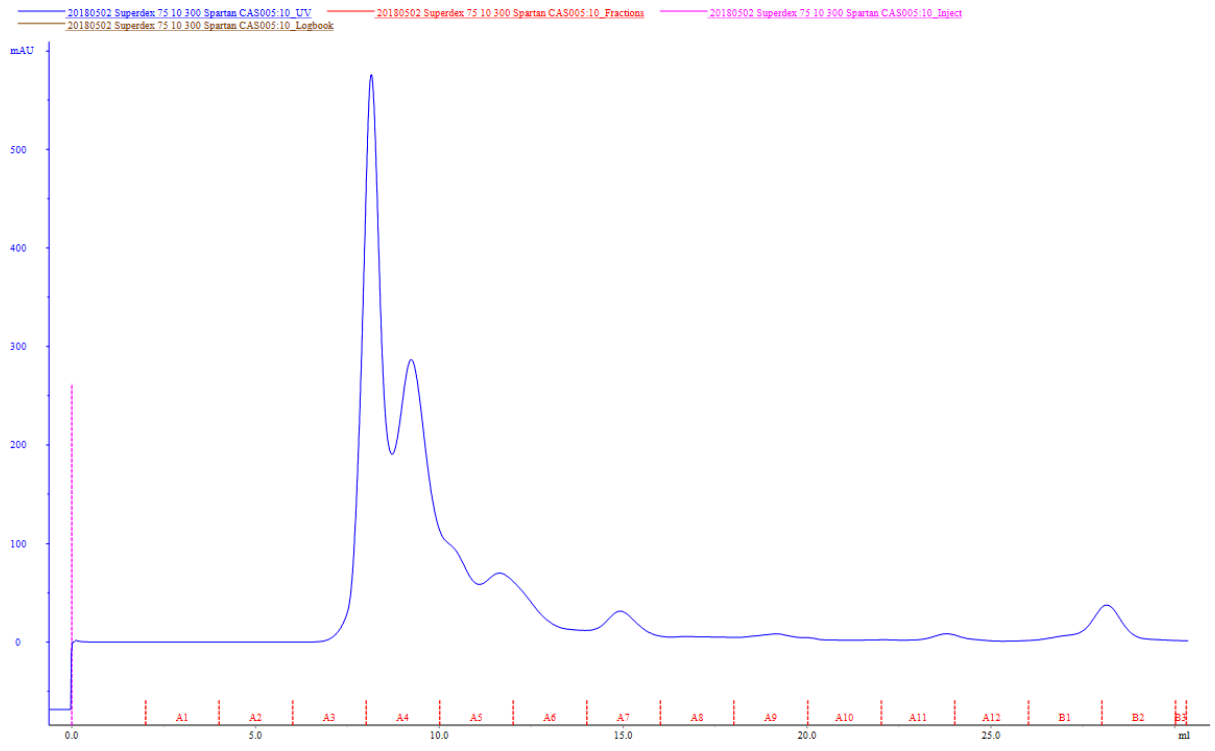


Figure A.5 Second Stage Purification of SPRTN. After removal of the Z-basic tag from the protein, SPRTN was purified by size-exclusion chromatography on the FPLC using a Superdex™ 75 10/300 column. Elution of the protein was at 9 mL. UV absorbance was monitored at 280 nm.

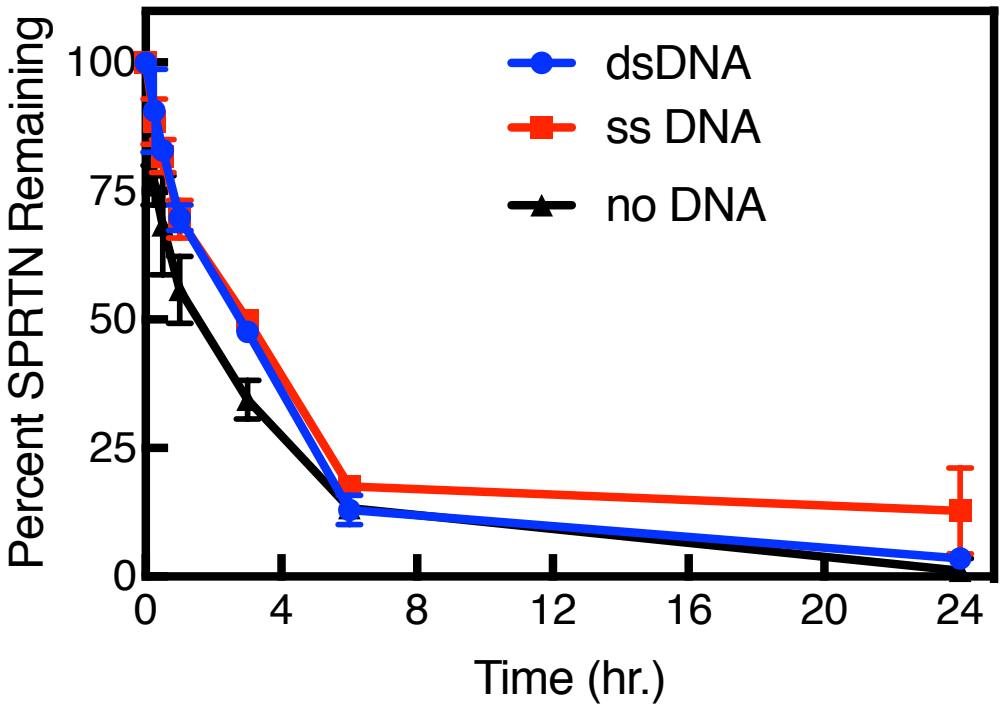


Figure A.6 Time Course for SPRTN Protease Activity. SPRTN activity was monitored for self-degradation in the presence of dsDNA (blue), ssDNA (red), or no DNA (black).

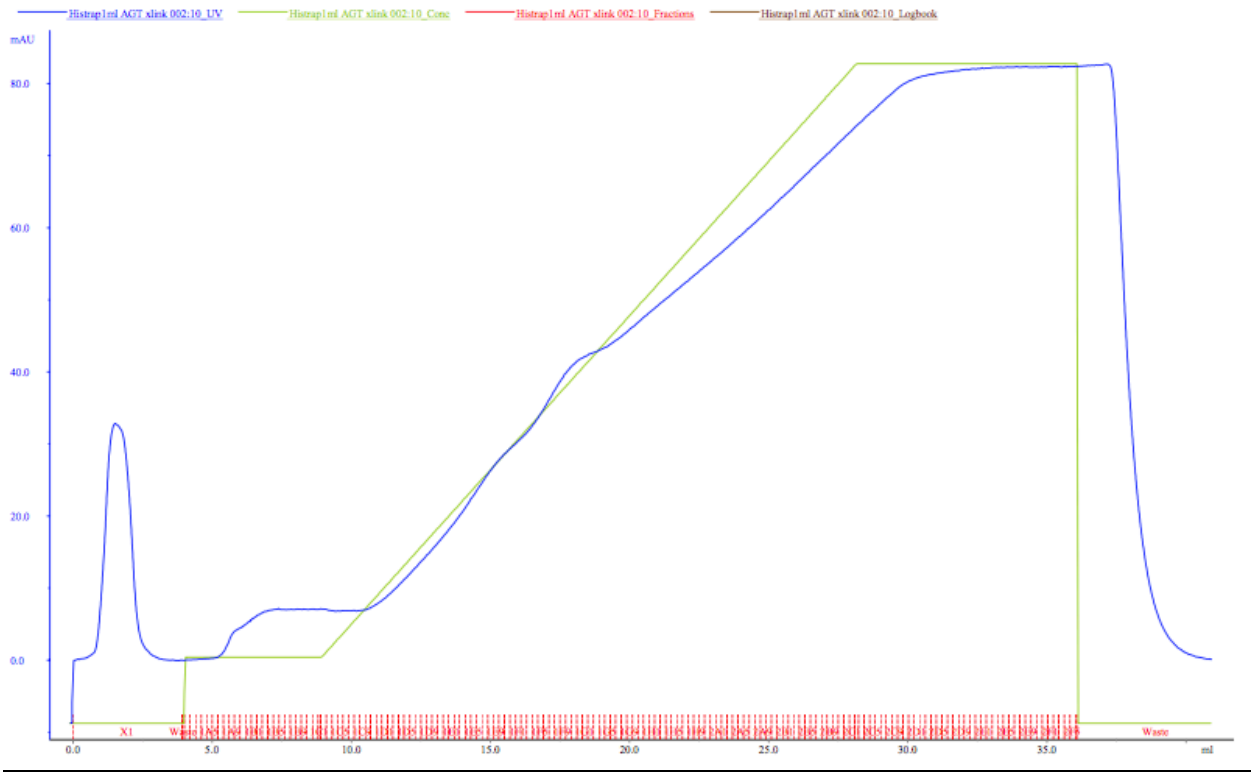


Figure A.7 Purification of AGT-DNA Crosslink by FPLC. The reaction product was loaded onto Ni^{2+}-NTA FPLC column and eluted using a gradient from 5 mM to 100 mM imidazole. Elution of the AGT crosslink was at 18 mL. UV absorbance was monitored at 280 nm.

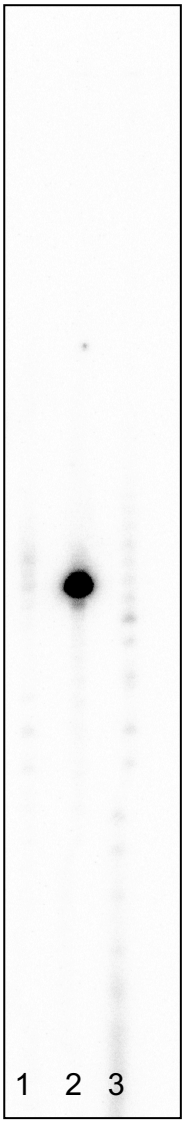


Figure A.8 Detection of Unmodified DNA template in AGT-DNA Crosslink by ^{32}P-labeling. Band corresponding to contaminating DNA in purified AGT-DNA crosslink (Lane 1) found at similar level as unmodified template DNA (Lane 2) and primer DNA (lane 3).