4-9-2015

Mutagenesis and Genetic Control of Translesion Synthesis Across an Abasic Site, 8,5'-Cyclopurines and O^2-Alkylthymidine DNA Lesions

Savithri Anurada Kumari Weerasooriya
University of Connecticut - Storrs, savithri.weerasooriya@uconn.edu

Mutagenesis and Genetic Control of Translesion Synthesis Across an Abasic Site, 8,5'-Cyclopurines and O^2-Alkylthymidine DNA Lesions

Savithri Anurada Kumari Weerasooriya, Ph. D.

University of Connecticut, 2015

DNA contains the genetic information of all living organisms. Therefore, its integrity and stability are essential for the long-term viability of the organism. However, DNA is a chemical entity, which gets constantly damaged by both endogenous and exogenous DNA damaging agents. These DNA damages may lead to mutations and eventually cause diseases like cancer. Multiple DNA repair mechanisms have evolved in living organisms to repair DNA damages. Even so, not all DNA damages can be repaired before the replication apparatus encounters the DNA damage. As such, cells developed a damage tolerance pathway known as translesion synthesis (TLS) that allows cells to overcome replication blockage and facilitate bypass of the DNA lesions. This process is carried out by TLS polymerases, of which most belong to the Y family of DNA polymerases. These specialized enzymes are capable of bypassing the damaged DNA, but they also are low fidelity enzymes, frequently associated with mutagenesis and carcinogenesis.

In this dissertation, I have investigated the replication bypass of different DNA lesions, including abasic site, 8,5'-cyclopurines and tobacco-specific nitrosamine-derived O^2-alkylthimidines in *Escherichia coli* and in human embryonic kidney (HEK293T) cells. The comparative replicative assays in bacterial and human cells revealed that bypass of these DNA lesions are significantly different in the two systems and that the mammalian polymerases are more efficient in the TLS of strong replication blocking lesions. Based on

the cytotoxic and mutagenic properties of these lesions, it is evident that replicative bypass efficiency varies with the complexity of the lesion. Moreover, multiple TLS polymerases are involved in the mutagenesis of different DNA lesions. These studies provide important mechanistic details as to how these lesions are bypassed and reveal their mutagenic properties.

Mutagenesis and Genetic Control of Translesion Synthesis Across an Abasic Site, 8,5'-Cyclopurines and O^2-Alkylthymidine DNA Lesions

Savithri Anurada Kumari Weerasooriya

B.Sc., University of Peradeniya, Sri-Lanka, 2007

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2015

Approval Page

Doctor of Philosophy Dissertation

Mutagenesis and Genetic Control of Translesion Synthesis Across an Abasic Site, 8,5'-Cyclopurines and O^2-Alkylthymidine DNA Lesions

Presented by

Savithri A. K. Weerasooriya

Major Advisor _____

Dr. Ashis K. Basu

Associate Advisor _____

Dr. Steven Suib

Associate Advisor _____

Dr. Mark Peczuh

University of Connecticut

2015

DEDICATION

I dedicate this dissertation to my loving father, Mr. Ranjith Weerasooriya, my loving mother Mrs. Hema Kumari, my loving brothers, and to my amazing husband Saminda whose affection and encouragement make me achieve such success.

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my major advisor, Dr. Ashis Basu, for his tremendous support and constant encouragement throughout my time as his student. I have been extremely fortunate to have a supervisor who cared so much not only for his research, but also for his students. As a young researcher, his patience and compassion towards me has been remarkable. His motivation, enthusiasm and immense knowledge are truly inspiring and have helped me to grow as a researcher. I would further extend my gratitude to my advisory committee, Dr. Steven Suib, Dr. Mark Peczuh, Dr. Fatma Selampinar and Dr. Alfredo Angeles-Boza for their generous support, guidance, and motivation.

Lab would not have been a fun place to work without all of my lab mates. I am grateful to all seniors Drs. Paromita, Vijay, Paritosh, and Rajat, who were great mentors. I would like to thank Vijay for all invaluable suggestions and support given to all my projects. Special thanks to Dr. Varsha, for being a wonderful friend, a source of great intellectual and emotional support, and for all the great things we did together. Also, I would like to thank current group members Arindam, Kimberly, Arindom, and Brent for their friendship and support throughout these years. I also thank Nadeeshani, Vidya, and all the undergraduates, Savas, Ashley, Anonya, Vruksha and Hamsa for their help in my projects.

I express my heartfelt gratitude to my husband Saminda for his unconditional love, unwavering support, and patience throughout these years. This journey would not have been possible without you in my life. I am highly indebted to my parents for believing in me and for constant encouragement throughout my life. I would also like to thank my two brothers for their continuous support.

A final thank you to all of my family and friends, both here and back home, for all of your love and support.

Table of Content

List of Figures

List of Tables

1 Literature Survey

1.1 Cancer

Cancer is one of the leading causes of death worldwide. Statistics from World Health Organization (WHO) revealed that 8.2 million deaths related to cancer occurred in the year 2012. WHO also predicted that the death toll will increase to 9 million in 2015.[1]

According to the American Cancer Society, cancer is the second leading cause of death in the United States. Approximately 1,658,370 new cancer cases are expected to be diagnosed in 2015 and estimated deaths from cancer will be 589,430, that is, nearly 1,620 deaths per day.[2] Although this statistics show a decrease in mortality rate from its peak in 1991, there is a continuous effort to create awareness and ultimately combat cancer. Hence, researchers worldwide are motivated to explore avenues such as cancer prevention, early detection, and effective therapeutic intervention to increase the quality of life and cancer survival.

Cancer is the uncontrolled growth of abnormal cells in the body. If not controlled, it can ultimately lead to death. Development of cancer is a complex multistep process, which involves an alteration in the DNA of a normal cell that leads to proliferation, survival, invasion, and metastasis of an abnormal cell mass. This process is known as the carcinogenesis. Nearly all cancers/tumors occur due to a failure of the genes involved in the cell growth and division. Some tumors can be benign while others may exhibit malignancy, and almost all the malignant tumors are cancerous. Distinctive types of cancers exhibit unique characteristics; therefore, prognosis and treatment in each case may be different. Nonetheless, occasionally there are marked resemblances amongst the various cancer types.[3]

It is believed that all cancers share a common pathogenesis. Development of a cancer is considered analogous to outcome of Darwinian evolution occurring amongst cell

populations within the microenvironments of the tissues of an organism. In this regard, carcinogenesis is centered on two constituent processes, the continuous acquirement of heritable genetic variation in individual cells by random mutations and the natural selection of resultant phenotypic diversity. During the selection process, cells that have acquired deleterious mutations and are incapable of proliferating may be eliminated, while a single cell acquired a set of adequately advantageous mutations that allow autonomous proliferation can survive and become malignant.[4]

DNA in normal cells is subjected to damage by various agents, both internal and external of origin, which may lead to mutations. Somatic mutations that arise in the cancer genome may comprise of several types of DNA sequence changes, such as, base substitutions, insertions or deletions of small DNA segments, DNA rearrangements, and gene copy number increase or decrease. Cancerous DNA may also have acquired completely new DNA sequences form exogenous sources like viruses.[5] Moreover, mutation rates increase substantially due to exposure to exogenous mutagenic agents, such as, tobacco carcinogens, chemicals and radiation. Rates of somatic mutations are also increased in several inherited diseases like Fanconi anaemia and Xeroderma Pigmentosum (XP), which have increased potential for development of cancer.[4]

It is generally accepted that accumulation of mutations in oncogenes, tumor suppressor genes, and stability genes are responsible for tumorigenesis. All these genes are critically important in maintaining regular cell cycle. Proto-oncogenes are genes that code for proteins to help regulate cell growth and differentiation. Any mutations that result in activation of the proto-oncogene to become an oncogene, which produce onco-proteins, can eventually lead to uncontrolled cell growth.[6]

When cells encounter stress stimuli such as oncogene activation or DNA damage, tumor suppressor genes are activated to prevent cell cycle progression, to promote DNA repair or apoptosis. *p53* is one of the most important tumor suppressor genes and its activity is tightly regulated in normal cells. Therefore, mutations in such tumor suppressor genes could cause inhibition of cell-cycle arrest or apoptosis, eventually leading to uncontrolled cell growth, resulting in tumorigenesis.[7] Stability genes are genes that are involved in DNA repair processes, mitotic division, and chromosomal segregation. These genes are involved in keeping genetic alterations to a minimum, and thus when mutated, it accelerate the progression of accumulated mutations. Mutations that occur in the germline of these genes will result in hereditary cancer predispositions, while mutations in single somatic cells will lead to sporadic tumors.[8]

The prime objective of every life form is to transfer its genetic material intact and unchanged to the next generation. Therefore, maintaining genome integrity is essential for normal cellular functions. However, endogenous and exogenous damaging agents constantly insult genome stability. To counteract this threat, cells have evolved several systems to detect DNA damage, signal its presence and trigger its repair. Despite the presence of finely tuned DNA repair mechanisms, some of the DNA damages escape repair and interfere replication fork progression. Specialized polymerases known as translesion synthesis polymerases are able to synthesize DNA past such damaged bases with low fidelity leading to mutations. It is now widely accepted that this process of translesion synthesis (TLS) is responsible for increased point mutations in the genome leading to carcinogenesis. Hence it is important to improve our knowledge in the area of translesion synthesis for better understanding of DNA damage bypass, mutagenesis and also new avenues for cancer therapy.

1.2 DNA Damage

DNA contains the genetic information of all living organisms. Therefore, its integrity and stability are essential for survival of the species. In both normal and cancer cells, DNA constantly undergoes a myriad of damages either spontaneously, or through exposure to endogenous and exogenous agents, leading to a plethora of damages in DNA.[9]

Spontaneous DNA lesions are introduced every day during normal cellular processes like DNA replication, DNA repair, or DNA rearrangement, and also as a result of chemical modifications of DNA molecule by hydrolysis, methylation, oxidation, and deamination of nitrogen bases. Abasic site, one of the most common lesions arises as a result of hydrolysis of nucleotides under normal physiological conditions. Deamination, either spontaneous or induced by a reactive species, of cytosine, adenine, guanine, or 5-methylcytosine, converts these bases to miscoding uracil, hypoxanthine, xanthine and thymine, respectively.[9, 10]

All organisms are exposed to ionizing radiation and ultraviolet (UV) radiation from sun, whereas humans are frequently exposed to radiation during medical procedures. These exposures result in DNA damage. In addition to radiation, environmental factors such as exposure to chemical toxins and infections by pathogenic bacteria and viruses have also been shown to induce DNA damage. For humans, chemicals produced by tobacco products are probably the most prevalent environmental carcinogens found in the nineteenth and twentieth centuries, which cause a variety of DNA lesions and ultimately result in tobacco-induced carcinogenesis.[11-13]

Certain cellular metabolites and byproducts generated during metabolism also cause DNA damages. Reactive oxygen species (ROS), produced endogenously during respiration and lipid peroxidation, are the major culprits of tens of thousands of DNA lesions per day.

However, living cells have developed a natural defense system during evolution to fight against ROS-induced DNA damages.[14]

If these DNA lesions are not repaired rapidly, they can either block or stall genome replication and transcription. Furthermore, if incorrectly repaired or bypassed during replication, they might result in mutations or wider-scale genome aberrations that threaten integrity of the genome.

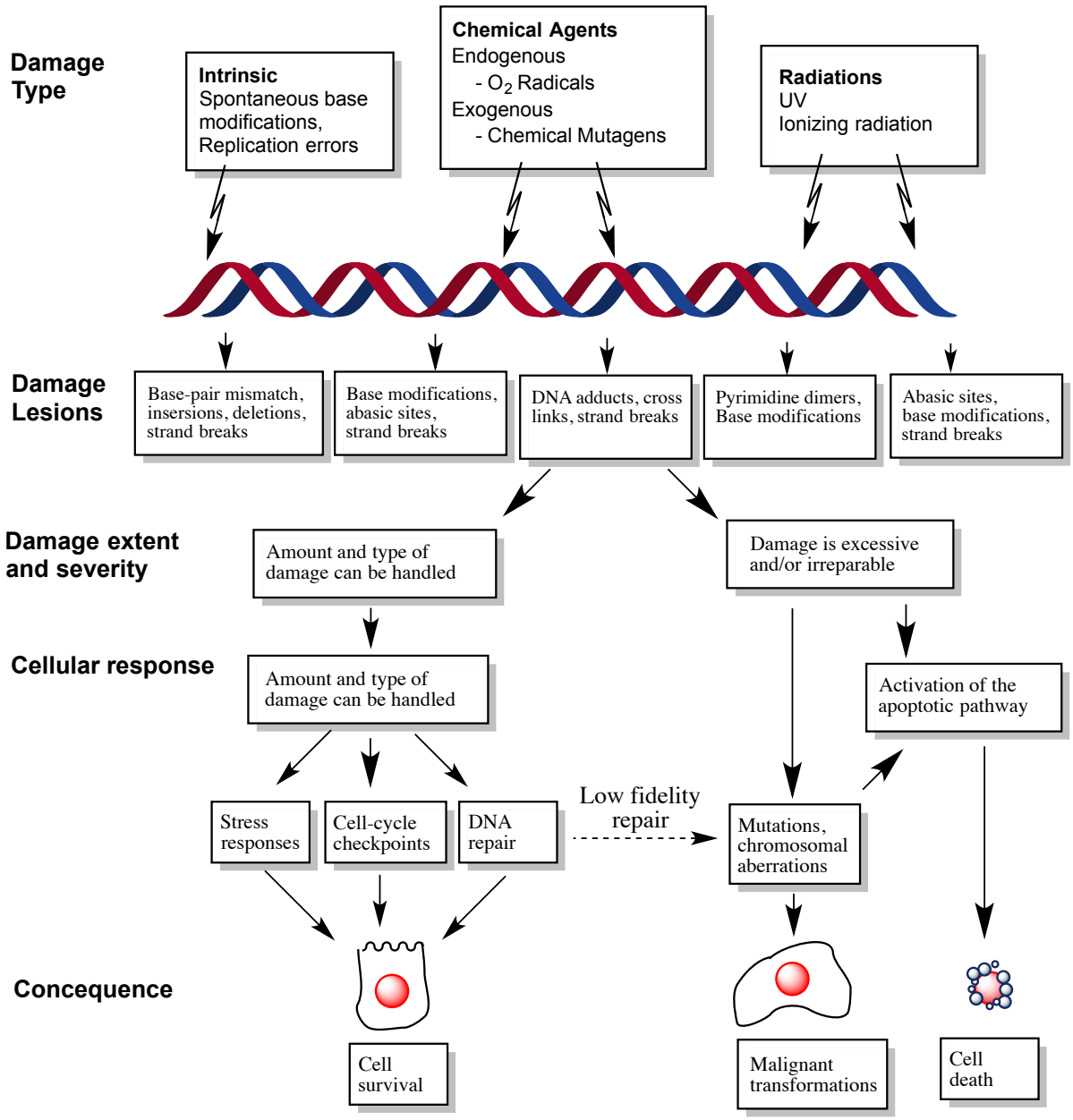


Figure 1: Different types of DNA damage cause different types of lesions, and these, in turn, are handled by the cell in different ways, Adapted from ref[15]

1.2.1 Reactive Oxygen Species (ROS)

Free radials are initially known in chemistry as intermediates in organic and inorganic reactions. In 1954, Gilbert and Rebecca Gersham were the pioneers to publish suggesting that these radicals are important players in biological environments and responsible for deleterious processes in the cell.[16, 17] Soon afterwards in 1956, Herman Denham suggested that these species might play a role in physiological events and particularly in the aging process. Thereafter, numerous studies were conducted in the field of biochemistry to understand the role of free radicals.[16, 18]

Normal cellular processes generate reactive species capable of reacting aggressively with biological macromolecules, including DNA, proteins and lipids. Reactive oxygen species (ROS) are one such important group of reactive species found in cells during normal physiological conditions as well as in abnormal conditions. These are species of oxygen, in which oxygen is reduced to varying degree; therefore, they are in a more reactive state than molecular oxygen. These ROS include superoxide ion radical ($O_2^{-\bullet}$), hydroxyl radical ($OH^{\bullet}$), peroxyl radical ($ROO^{\bullet}$) alkoxyl radicals ($RO^{\bullet}$), and one form of singlet oxygen ($^1O_2$).[16]

ROS originate from both endogenous and exogenous sources. Cellular metabolism and byproducts of metabolism are examples for endogenous sources. In healthy cells, ROS are produced specifically to serve essential biological functions such as cell signaling and cellular defense. However, they also are formed as byproducts, particularly during aerobic respiration, by oxidoreductase enzymes and metal catalyzed oxidations. Exogenous sources for ROS production include xenobiotics and ionizing and UV radiation.[19]

1.2.2 Sources of ROS

During oxidative phosphorylation, about 1-2% of molecular oxygen is converted to ROS, largely to superoxide anion, through a series of consecutive reactions. This occurs due

to “leakage” of electrons form the electron transport chain in mitochondrial membrane. These leaking electrons react with surrounding oxygen to form superoxide radical, which in turn are converted to hydrogen peroxide (H_2O_2) by superoxide dismutase. Unlike superoxide ion, H_2O_2 is able to cross the cellular membrane and generates highly reactive hydroxyl radicals by reacting with metal cations, e.g., iron or copper via Fenton reaction.[20]

$O_2 + e^- \longrightarrow O_2^{-\cdot}$ (Superoxide anion)

$O_2^{-\cdot} + H_2O \longrightarrow HO_2^{\cdot}$ (Hydroperoxyl radical)

$HO_2^{\cdot} + e^- + H \longrightarrow H_2O_2$ (Hydrogen peroxide)

$H_2O_2 + e^- \longrightarrow OH^- + {}^{\cdot}OH$ (Hydroxyl radical)

Fenton reaction;

$H_2O_2 + Fe^{2+} \longrightarrow OH^- + {}^{\cdot}OH + Fe^{3+}$

Cytochrome P450 (CPY), a family of hemo-thiolated enzymes, is the terminal oxidase of the monoxygensae system found in mammalian endoplasmic reticulum (ER). These enzymes catalyze the oxidative metabolism of a variety of endogenous and exogenous lipophilic compounds. Poor coupling of CPY catalytic cycle significantly increases endogenous ROS production. In particular, induction of P450 2E1 and 2B during metabolism of ethanol and phenobarbital, respectively, have been proposed to contribute considerably to ROS production.[20]

Other important endogenous sources of ROS include host inflammatory cells like neutrophils, eosinophils, and macrophages. Most significantly, activated macrophages are

known to stimulate a rapid but transient increase in oxygen uptake called "respiratory burst", which gives rise to a wide variety of ROS, including superoxide anion, hydrogen peroxide, and nitric oxide. For instance, activation of Kupffer cells, specialized macrophages, has been found to be associated with tumor induction due to ROS release.[21] Microsomes are responsible for ROS formation in hyperoxia sites. In other organs, liver in particular, peroxisomes also contribute to H_2O_2 generation.[22]

Furthermore, ROS production can be initiated by exogenous sources including radiations and xenobiotics, including chlorinated compounds. Ionizing radiation such as γ-radiation and non-ionizing UV radiation can produce an array of ROS from ionizing of intercellular water and hydrolytic cleavage of H_2O_2, respectively.[16] The relationship between ionizing radiation and DNA damage via generation of hydroxyl radicals is now well established. Hydroxyl radicals subsequently react with water in two possible pathways:

In the first pathway, ionization of the water molecule results in the formation of a hydrated electron (e^-_{aq}) and $H_2O^{\bullet+}$ that rapidly loses a proton to generate the hydroxyl radical. This hydrated electron slowly decays into $H^\bullet$ and hydroxide anion. It is now known that generation of reactive $^\bullet OH$ through such processes results in persistent attack on the nucleophilic intracellular DNA.[23]

The second process involves excitation of the water molecule followed by homolysis into $H^\bullet$ and the hydroxyl radical ($^\bullet OH$). Amongst many intracellular radicals produced under physiological and pathophysiological conditions, the $^\bullet OH$ radical is one of the most reactive species and one that plays a significant role in a wide variety of DNA damages.[23]

1.2.3 Oxidative stress

Under normal physiological conditions in healthy cells, the balance between the production and elimination of ROS are tightly regulated. This could be either non-

enzymatically by antioxidants like vitamin C, vitamin E, vitamin A, carotene, glutathione or enzymatically by superoxide dismutase, catalase and glutathione peroxidase. However, under certain conditions (such as aging, diseases, or due to exposure to radiation), an imbalance between these two process could occur leading to a condition called "oxidative stress," a state characterized by excessive production of ROS. Oxidative stress can significantly alter critical cellular macromolecules such as DNA, protein and lipids. Although modifications to lipids and proteins can be removed via regular turnover of the molecules, modifications to DNA needs to be repaired.[24] Therefore, the primary deleterious consequence of oxidative stress stems from damage to DNA, as it has been implicated in mutagenesis, carcinogenesis, aging and many neurological diseases.

As elaborated above, hydroxyl radical is considered to be the major culprit for oxidative DNA damage. Damage to DNA by $^{\bullet}OH$ includes, base modification, sugar damage, strand breaks, abasic sites and protein-DNA cross-links. Hydrogen abstractions from deoxyribose sugar moiety, a principal reaction between $^{\bullet}OH$ and DNA, can leads to base loss and single strand breaks in DNA. Moreover, addition of $^{\bullet}OH$ to the π bonds of DNA nucleobases may result in distinct DNA lesions (adducts). In spite of the existence of DNA repair enzymes to repair these lesions efficiently, low levels of adducts may persist in DNA, resulting in mutagenesis and carcinogenesis. Therefore, assessment of mutational properties of such lesions provide better understanding of their biological consequences in a more comprehensive way.[25] Abasic site and 8,5'-cyclo-2'-deoxynucleoside lesions are important lesions formed due to $^{\bullet}OH$ attack to DNA and are discussed in more details in the next section.

1.3 Abasic site

Abasic sites, also called apurinic or apyrimidinic sites (AP-sites), are one of the most common lesions in DNA. [26] AP-sites can be generated via spontaneous hydrolysis of *N*-glycosidic bond. DNA damaging agents such as free radicals and alkylating agents can also form AP sites. Alkylating DNA damaging agents destabilize the *N*-glycosidic bond due to modification of bases generating a better leaving group.[10, 27-32] It has been estimated that a mammalian cell looses at least 10^4 purines per day under normal physiological conditions.[33] In addition, AP sites are also formed as intermediates during base excision repair (BER) mechanism.[34] Interestingly it has been determined that depyrimidination occurs at least 100 times lower than depurination, for reasons that are not well characterized.[35]

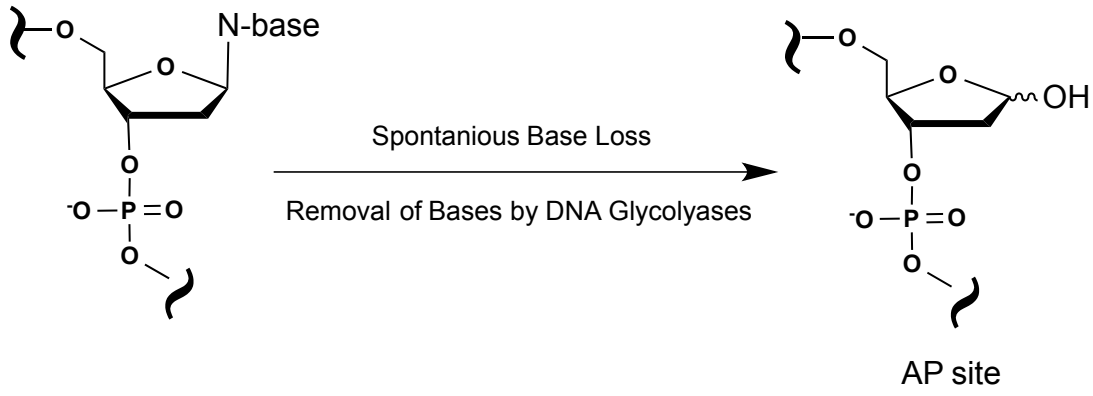


Figure 2: Formation of abasic site

1.3.1 Structural aspects

Several groups have studied the structural aspects of abasic site using a stable tetrahydrofuran or open analogue form in different types of duplex DNA. Based on NMR and restrained molecular dynamics data, it has been observed that AP-site containing DNA assumes more or less regular B-form.[36, 37] However, depending on the adjacent pairs and opposite bases, sequence-specific conformational differences have been observed in DNA duplexes.[36] Most studies on purines opposite abasic sites showed that they remained stacked within the DNA helix. In contrast, pyrimidine opposite abasic site may either remain stacked within the helix or exists as extra-helical depending on the nature of the abasic site and the flanking bases.[38-40] Although there is only a modest affect on helix distortion, thermodynamic and calorimetric data suggest that the abasic sites cause a significant instability of DNA duplex.[41]

1.3.2 Repair

Most of the abasic sites formed are primarily repaired by the base excision repair (BER) mechanism. During BER, AP site is cleaved by AP endonuclease forming a 3'-OH group and a 5'-deoxyribose phosphate residue (5'dRP).[42, 43] Further repair is carried out by at least two distinct BER sub pathways, the short patch repair or long patch repair. AP sites can be incised by AP layase/DNA glycosylases and initiate short patch BER of AP site. AP layase cleaves DNA at 3' side of AP site generating a 5' phosphate and a 3' α-β unsaturated terminus, which are further processed by DNA repair mechanisms.[43, 44]

Though AP sites can be repaired by BER, AP sites formed within the single-stranded DNA (ssDNA) likely would not be repaired as it lacks information on the complementary strand. Further, if BER is attempted in ssDNA of the replication fork, it is more likely to generate single strand breaks (SSB), which are much more toxic than AP sites.[45] Therefore,

AP sites that escape from DNA repair is highly detrimental to cells and a strong block to DNA replication. However, translesion synthesis polymerases are able to bypass AP sites frequently generating mutations.

1.4 8,5'-Cyclopurines

Ionizing radiation and other processes that give rise to reactive oxygen species generate many lesions in DNA, including tandem DNA lesions. This type of DNA damage is found to play an important role in mutagenesis, carcinogenesis and ageing. 8,5'-Cylopurine 2'-deoxynucleosides (cPDNS), which are formed due to ionization radiation, have been detected in DNA derived from many different cells and organisms.[46] They are also considered one of the major oxidative stress response markers. These cyclopurines exists as 5'*R* and 5'*S* diastereomers and are considered unique tandem lesions due to the presence of an additional C8-C5' covalent bond between the purine base and the sugar moiety of the same nucleoside. This additional *N*-glycosidic bond gives extra stability to the damaged nucleoside.[47] NMR studies have shown that the 2-deoxyribose sugar moiety of 5'S cdG has an O4' *exo* conformation and able to maintain Watson-Crick base pairing interaction with complementary cytosine base. However, this O4'- *exo* pseudorotation causes a perturbation in the helical twist and base-pair stacking of DNA, causing thermodynamic destabilization of the DNA double helix.[48] Furthermore, due to the additional covalent bonding and the extra stability, they are resistant to acid-induced hydrolysis, therefore preventing their repair by base excision repair (BER) and can only be repaired by nucleotide excision repair (NER) mechanism.[49] Hence it is proposed that these lesions may be responsible, at least in part, for neurodegenerative diseases like Xeroderma pigmentosum (XP) and Cockayne syndrome (CS) patients with defects in NER.[50]

1.4.1 Chemistry of Cyclopurines Formation

8,5'-Cyclo-2'-deoxyadenosine (cdA) and 8,5'-cyclo-2'-deoxyguanosnie (cdG) are a class of major DNA lesions formed by ionizing radiation and the Fenton reaction induced hydroxyl radical attack on 2'-deoxyadenosine and 2'-deoxyguanosine, respectively. It is evident that these lesions are not formed by singlet oxygen or other types of ROS. Cyclopurines represent concomitant damage to both the base and the sugar moiety of the same nucleotide and are considered tandem DNA lesions.[51]

The formation of cyclopurines involves two steps illustrated in Figure 3. In the first step, hydroxyl radical attack at the C5' of the sugar moiety to abstract an H atom generates the C5' radical. This radical formed at C5' then attack double bond between N7 and C8 of the purine ring of the same nucleoside in the absence of oxygen to form an 8,5' bond. Subsequent oxidation of the N7-centered radical leads to the formation of the 8,5'-cyclopurined. Evidence for this reaction was first presented when 8,5'-cyclo-AMP was formed after adenosine-5'-monophospahte (AMP) was exposed to ionizing radiation in the absence of oxygen. The experimental rate constant for H abstraction was determined as 2.8×10^{9} M^{-1}s-1 and the rate constant for cyclization reaction is 1.6×10^{5} s^{-1}. Since the second reaction is relatively slow, it can be interfered by cellular molecular oxygen to produce a peroxyl radical. However, the inhibitory reaction may not be prominent in the eukaryotic nucleus, where there is substantially low O_2 concentration. This may be the reason that there were no evidences for presence of cyclopurines in the mitochondria, where there is a higher O_2 concentration.[52, 53]

Furthermore, research on cyclopurines have identified two diastereomers (*S*, *R*), in DNA exposed to ionizing radiation in aqueous solutions, and the ratio varies depending on

the nucleoside. Biochemical and structural studies showed that both *R* isomers of cdA and cdG are mostly found in ssDNA, while *S* diastereomers are favored in dsDNA.[54]

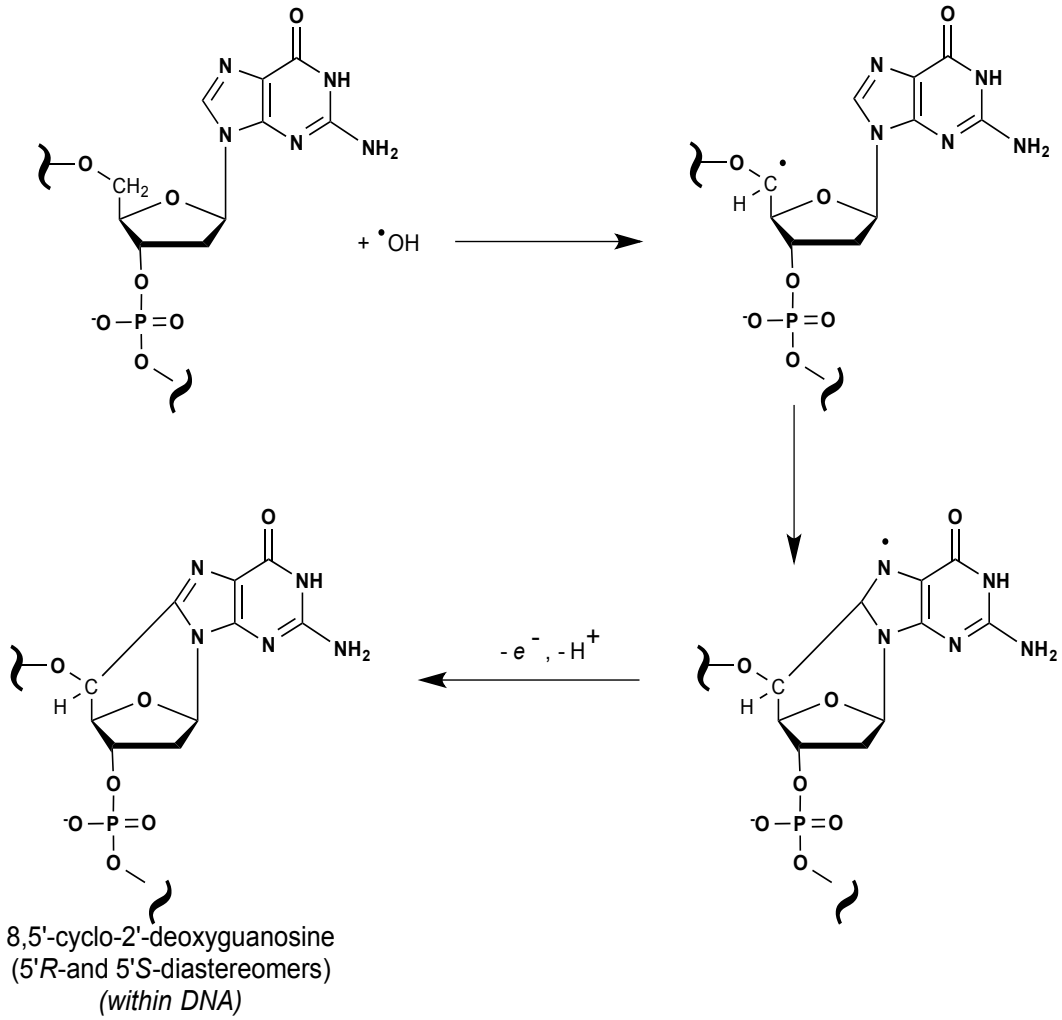


Figure 3: Formation of 8,5'-Cyclopurines

1.4.2 Biological Significance of Cyclopurines

The genomes all organisms are constantly under attack by various DNA damaging agents. Every cell in the body gives rise to thousands of lesions per day. However, DNA is protected by a variety of repair systems, including nucleotide excision repair (NER) and base excision repair (BER). These two systems remove the majority of DNA lesions.

Earlier studies have shown that cyclopurines can only be repaired by NER and may not be repaired by BER due to the presence of an extra covalent bond between C8 and C5' of the same nucleoside.[49, 54-56] Most of the initial investigations on repair were performed with *S* or *R* isomer of cdA and less attention was given to cdG. Kuraoka et al. have searched for glycosylase and nuclease activity of HeLa extract on a plasmid containing either 5'*R* or 5'*S* isomers of cdA. This study showed that glycosylase activity was unsuccessful in the cleavage of 8,5'bond and postulated that BER may be inefficient in repairing these lesions. However, both isomers were removed in the cell-free extract, suggesting that these are repaired by NER. It was also noted that 5'*R* isomer was repaired more efficiently than the 5'*S* isomer of cdA. Moreover, both cyclopyrimidine dimer (CPD) and 5'S-cdA had a comparable NER efficiency.[49].

Gene expression was impaired by a single (5′*S*)-cdA on the transcribed strand of a gene in Chinese hamster ovary and human cells, with efficiency comparable to CPD. This kind of effect on gene expression blockage can be related to those patients with XP whose NER is defective. Therefore it was hypothesized that the accumulation of cyclopurines in an active gene would prevent transcription, which in turn may lead to neurodegeneration related to XP disease. This hypothesis was further tested in human cells transfected with the vector containing the 8,5'-cdA lesion to an active gene and measuring its activity through a

modified luciferase response as well as through direct monitoring of the transcriptional activity of cells from a XP patient with neurodegeneration.[54]

In vitro kinetic study using T7 DNA polymerase and mammalian DNA polymerase δ showed that presence of any diastereomers of cdA hindered primer extension. In the case of T7 DNA polymerase, it was able to incorporate a base opposite the 5' *R* isomer, while pol δ was blocked before the lesion and no extension was observed. This indicates that bypass polymerases can bypass bulky helix distorting lesions.[49] Indeed, purified pol η was able to efficiently bypass 5'*R*-cdA, while 5'*S*-cdG was strongly blocked after incorporation of one base opposite the lesion. This suggested that difference in stereoisomers can influence not only the nuclease activity, but also the translesion synthesis across a specific adduct.[57]

Experimental evidences show that single 5'S-cdA in TATA box sequence was able to block binding of transcription binding factors (TATA box binding proteins) and significantly block gene expression.[58] This observation also supports the notion that accumulation of cyclopurines may be responsible for neuronal death in XP. Apart from transcriptional blockage, 5'*S*-cdA has shown to trigger transcriptional mutagenesis, a situation where RNA polymerase bypasses non-bulky DNA lesions. Characterization of mutant transcripts resulting from bypass of *S*-cdA by RNA polymerase II (Pol II) showed misincorporation of an adenosine opposite to the next nucleotide 5'- to *S*-cdA and to cause multiple nucleotide deletions. This mutational event is also suspected to contribute to neuropathology in XP patients as a result of lesion bypass by RNA polymerase II.[59] Another study found out that levels of 5'S-cdA and 5'R-cdA were higher in *neil1-/-* mice suggesting that NEIL1 plays an important role in their repair.[60]

Recent studies published from our group evaluated the solution structure of site-specifically incorporated *S*-cdG in a duplex opposite dC, dT or dA in the complementary

strand. These studies established that O4'-*exo* (west) pseudorotation was maintained in each pair. *S*-cdG.dC pair was able to maintain Watson-Crick base pair, but *S*-cdG.dT pair adopted a wobble-pair. In contrast, there was no hydrogen bonding observed between *S*-cdG.dA pair.[48]

1.4.3 Cyclopurines and Human Diseases

Xeroderma pigmentosum (XP) is a neurological disease characterized by loss of large neurons in brain, spinal cord, and peripheral nervous system. It is also a cancer-prone disease, and XP patients exhibit increased skin cancer incidences due to UV sensitivity. Based on several scientific reports, it is now established that NER is defective in XP patients. Therefore, one can conclude that oxidatively induced DNA lesions including cyclopurines, which are specifically substrates for NER, might play a role in XP neurologic diseases.[50]

Scientific rationales for cyclopurins being responsible for XP include, (i) cyclopurines are substrates for NER, which is defective in XP patient, (ii) cyclopurines are not repaired by BER, (iii) gene expression in XP cells are strongly blocked by cyclopurines, an effect likely to cause neurodegeneration, (iv) cyclopurines are inherently stable and could accumulate in a cell in the absence of repair.[50, 61]

Cockayne syndrome (CS) is a genetic neurological condition with severe growth failure, premature aging, hypersensitivity to sunlight, microcephaly, and other neurologic abnormalities. CS complementation group A (CS-A), is associated with mutations in CSA gene. Evidence shows that CSA is required for repair of 5'*S*-cdA by NER. Therefore, it was hypothesized that accumulation of 5'*S*-cdA may be implicated in clinical features of CS-A patients.[62, 63] Furthermore, elevated levels of 5'*S*-cdA was observed in $csb^{-/-}$ mice indicating that it may accumulate in CS patients.[64]

In women with BRCA1 mutation, accumulation of cdA and cdG was observed, supporting the notion that BRCA1 may be involved in cyclopurins repair during NER, and if unrepaired may contribute to breast cancer.[65-67] Furthermore, it has been observed that presence of these lesions in non-transcribed regions may contribute toward aging.[58, 59, 68]

1.5 Tobacco Induced Carcinogenesis

Cancer is one of the leading causes of deaths in the United States. According to the United States Public Health Services, "cigarette smoking is the major single cause of cancer in the United States." Although the tobacco usage is declining in most industrial countries, overall consumption is increasing with approximately 5.5 trillion cigarettes smoked each year.[69] World Health Organization (WHO) predicts that, due to current smoking trends and unhealthy lifestyles, the estimated annual death toll will exceed to 12 million, and about 15 million new cancer cases will be detected annually by 2020.[70] Based on International Agency for Research on Cancer (IARC) statistics, numerous cancers are associated with tobacco including lung, oral cavity, nasal cavity, esophagus, stomach, liver, pancreas, bladder, ureter, and kidney.[71] Approximately, 90% of male and 75-80% of female lung cancer deaths reported in US are due to cigarette smoking.[72] Furthermore, tobacco smoke accounts for 30% of all cancer motility in developed countries. Human risks associated with tobacco smoke are not just limited to smokers, but also to non-smokers, including children and infants, as a result of indirect exposure through second hand smoke (SHS).[71]

The link between cigarette smoking and cancer was initially documented in the 1964 in the US Surgeon General's report, which had enormous positive effect on public health in the United States. Since then, smoking prevalence of men has decreased from 51.1% to the current 21.6%, while for women it has decreased from 33.3% to 16.5%. These facts indicate the critical importance of awareness and tobacco control in disease prevention.[73, 74]

Tobacco products contain a variety of chemicals such as nicotine and carcinogens, which are responsible for many preventable and premature deaths worldwide. The impact on human health due to tobacco depends on the type of the tobacco product and also the duration of exposure in one's lifetime. Smokers are exposed to tobacco by direct smoking of

manufactured and/or hand rolled cigarettes, whilst second hand smokers are expose to tobacco smoke from the environment, where cigarette smoking occurs. Smoke-less tobacco or chewing tobacco is another way people get expose to tobacco toxicity.[75]

Nicotine, the major component of all types of tobacco products, is an addictive chemical.[76] It is responsible for activating signaling pathways related to tumor initiation. In addition to nicotine, tobacco products consist of a variety of chemical carcinogens, including polyaromatic hydrocarbons (PAH), tobacco specific nitrosamines, volatile nitrosamines, aromatic amines, aldehydes, volatile hydrocarbons, metals, and many more organic and inorganic compounds. Most of these carcinogens are linked to multiple cancers in humans occurring in various systems like reparatory, digestive and urinary systems of tobacco users. Furthermore, binding of these tobacco smoke constitutes directly (without metabolic activation) to cellular receptors are responsible for activation metabolic pathways such as Akt and PKA, which in turn contribute to tobacco carcinogenicity.[77]

It is inevitable that prevention tobacco usage and smoking cessation are the fundamental approaches against smoking-related cancers. However, due to the powerful addictive property of nicotine, these strategies are only partially successful.[73] Moreover, tobacco companies, in response to the more widely accepted smoke-free laws, have marketed alternative nicotine sources like smokeless tobacco products. However, the levels of carcinogens in some these products are frequently higher than normal tobacco products, thus have not been accepted as safe substitutes. The promotion of smoking cessation and smoke-free environments should still remain as major efforts to reduce tobacco-related cancers.[78] Also it is important to further understand the carcinogens in tobacco products and their mechanisms of cancer induction. This knowledge could lead to alternate methods for cancer prevention.

1.5.1 Tobacco smoke

Tobacco smoke contains at lest 7000 chemicals, of which about 70 chemicals have been identified to cause cancers of various types.[79] Nicotine is the major addictive component in cigarette smoke, and it plays a role in tobacco carcinogenesis. It involves promoting cancer activating signaling pathways to assist cancer growth, angiogenesis, migration and invasion.[80] Furthermore, nicotine can undergo chemical conversion to carcinogenic, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) and, *N*'- nitrosonicotine (NNN) during the process of curing or smoking.[80, 81] Majority of the nicotine is metabolized to cotine by cytochrome P450 and aldehyde oxidase, and rest to other metabolites such as nicotine-*N*-oxides.[75]

Besides nicotine, tobacco smoke contains an array of carcinogens, including polycyclic aromatic hydrocarbons (PAH), aza-arenes, *N*-nitrosamines, aldehydes, volatile hydrocarbons, nitro compounds, and miscellaneous organic and inorganic compounds. Other carcinogens that are not yet evaluated by IARC are also present.[71, 82, 83] These metabolites are detoxified to more water-soluble forms by drug metabolizing enzymes, including cytochrome P450, glutathione S-transferase, and UDP-glucuronosyl, which are then excreted in urine. However, reactive electrophilic intermediates formed during this process attack to nucleophilic centers of DNA generating DNA adducts.[84] This process known as metabolic activation is required to exert carcinogenic functions of reactive intermediates. If these DNA adducts remain unrepaired, they may lead to mutations during translesion synthesis process, the adduct bypass mechanism. Specifically, if these mutations occur in tumor suppressor genes like *TP53* or in oncogenes such as *KRAS,* it could lead to development of cancer.[85, 86]

1.5.2 NNK Carcinogenicity

Nitrosation of nicotine and related alkaloids produce nitrosamines in tobacco. Among nitrosamines in tobacco smoke, NNK ("nicotine-derived nitrosaminoketone") is the strongest systemic carcinogen found in rodents, which induces lung carcinogenesis regardless of the route of administration. NNK is a procarcinogen and require metabolic activation by cytochrome P450 (CPY) enzymes to be carcinogenic.[87] As a carcinogen, it is very potent and was shown to induce lung cancers in rats with a total dose as low as 6 mg/kg (1.8 mg/kg when considered as a part of dose-response trend). This is similar to estimated 1.1 mg/kg dose of NNK from 40 years of smoking.[88]

NNK along with several other nitrosamines were also found to be effective hepatocarcinogens in rats. NNAL, the primary metabolite of NNK, and NNK are the only pancreatic carcinogens found in tobacco products.[89] Scientific evidence shows that metabolically activated NNK or PAH in cervix could lead to cervical carcinogenesis.[90]

During metabolic detoxification process, NNK undergoes α-hydroxylation to form NNAL. Both NNK and NNAL are carcinogenic. Further metabolic processing and detoxification converts NNAL to non-carcinogenic glucuronidated form, [4-(methylnitrosamino)-1-(3-pyridyl)but-1-yl]-*β*-*O*-D-glucosiduronic acid (NNAL-gluc), the ultimate detoxification product, which is rapidly excreted in the urine.[91]

NNK is an unsymmetrical nitrosamine. Metabolic activation of NNK by multiple CPYs generates reactive intermediates that can induce methylation, pyridyloxobutylation and pyridylhydroxybutalation of nucleobases in DNA generating multiple DNA adducts. There are three primary pathways for NNK activation i) carbonyl reduction, ii) pyridine *N*-Oxidation and iii) α-hydroxylation.[92] CPY-catalyzed α-hydroxylation of NNK occurs in two

pathways, (either α-methelhydroxylation or α-methylenehydroxylation) both of which form carcinogenic intermediates.[92, 93]

1.5.3 Methylation Pathway:

The α-methylenehydroxylation of NNK yields methane diazohydroxide and/or methyldiazonium ion, which leads to subsequent DNA methylation forming methyl (Me) DNA adducts.[94]

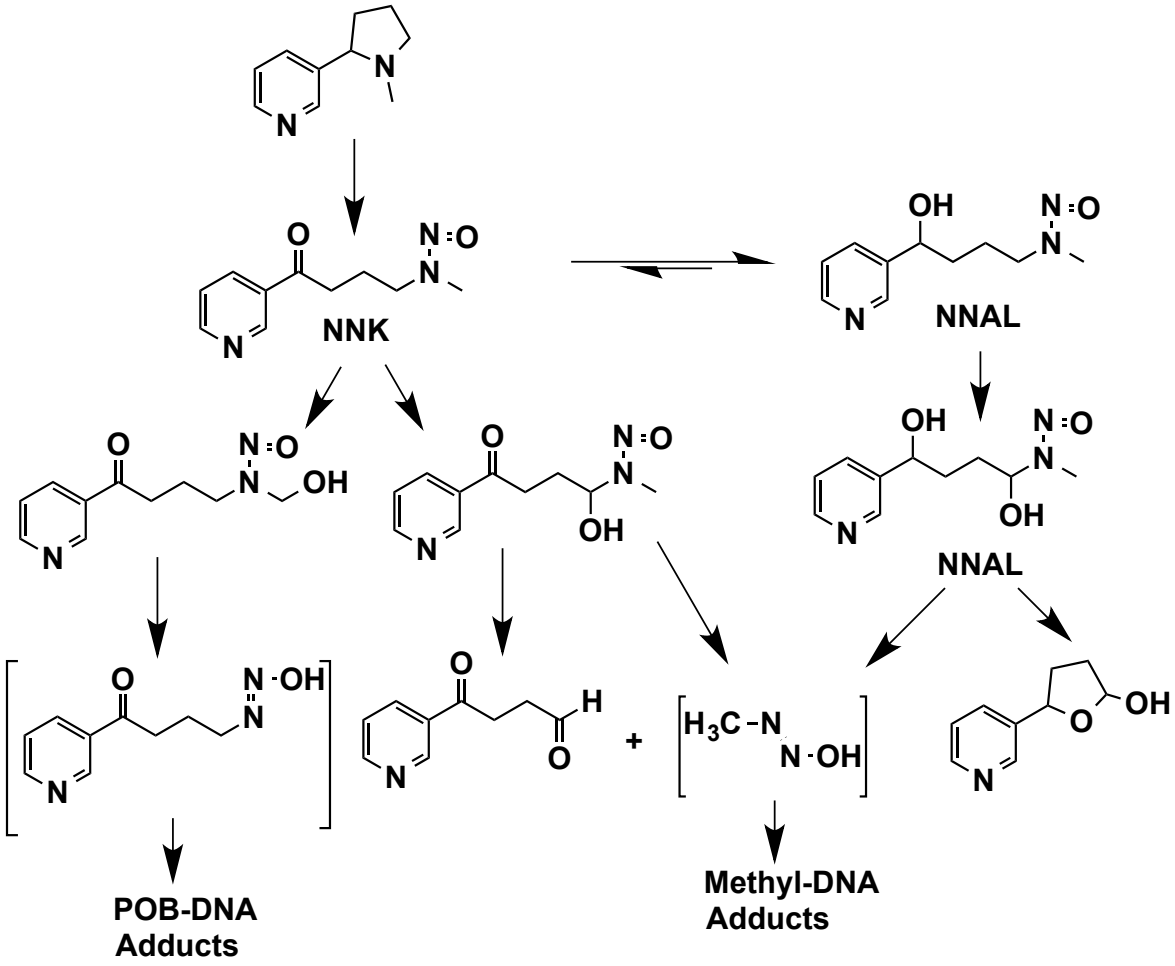


Figure 4: Metabolic activation of NNK . Adapted from ref[95]

Some of the methyl (Me) DNA adducts have been studied and are well characterized. Several methyl DNA adducts have been identified, 7-methylguanine (7-Me-dG), O^6-methylguanine (O^6-Me-dG), O^2-methyl thymidine (O^2-Me-dT), O^4-methylthymine (O^4-Me-dT) and O^2-methylcytidine (O^2-Me-dC).[96] The most abundant of these adducts *in vivo* is 7-Me-dG followed by O^6-Me-dG. Adduct formation from NNK shows great variation in rodent tissues depending on the activating enzymes.[96]

The persistence of a DNA lesion is a critical determinant of lung tumor formation in the A/J mouse. In fact, there is a strong correlation between the persistence of O^6-Me-dG and tumor formation in the lungs of A/J mice following exposure to NNK.[97]

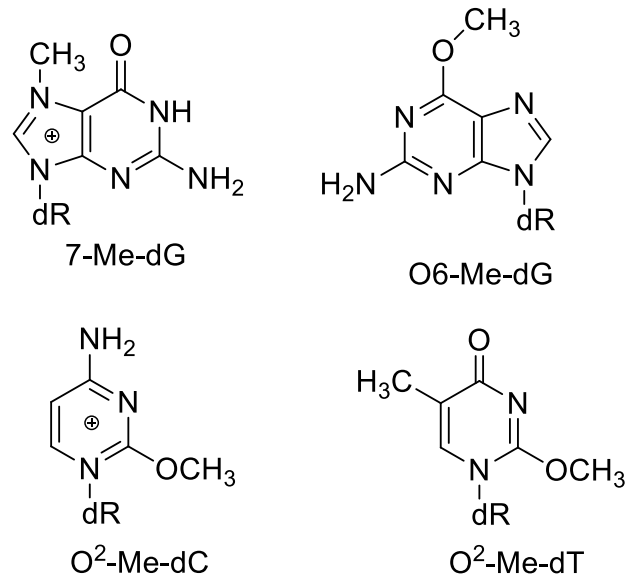


Figure 5: DNA adducts formed from the NNK methylating pathway

The adduct O^6-Me-dG is strongly mutagenic and has shown to induce primarily GC to AT transition mutations in codon 12 of the *K-ras* gene isolated from lung tumors of A/J mouse.[98] This mutation is a known consequence of mispairing of O^6-Me-dG with T. O^6-Me-dG has been shown to play a critical role in NNK-induced lung tumorigenesis in experimental animals as well as in human tissues. However, O^4-Me-dT concentration was found to be low due to its rapid removal in the lung and may have a limited contribution for lung carcinogenesis. Furthermore, 7-Me-dG was found be rapidly removed by base excision repair and seems to have low mutagenic potency as there was no evidence for persistence of 7-Me-dG adduct levels from NNK and incidence of liver tumors in rodents.[99] [100]

1.5.4 POB Pathway:

The α-hydroxylation at the methyl carbon produces α-hydroxymethyl NNK, which in turn undergoes glucuronidation. The latter spontaneously loses formaldehyde, producing pyridyloxobutyldiazohydroxide, which reacts with DNA forming bulky pyridyloxobutylation (POB) adducts in DNA and proteins.[101] Mild hydrolysis of these adducts would yield 4-hydroxy-1-(3-pyridyl)-1-butanone (HPB).[102] The carcinogenic properties of the pyridyloxobutylation pathway and the impact of the adducts formed from this pathway are not well understood. Four stable DNA adducts generated in this pathway have been identified. They are: 7-[4-3-(pyridyl)-4-oxobut-1-yl]-2'-deoxyguanosine (7-POB-dG), *O*2-[4-3-(pyridyl)-4-oxobut-1-yl]-2'deoxycytosine (O^2-POB-dC), *O*2-[4-3-(pyridyl)-4-oxobut-1yl]thymidine (O^2-POB-dT), and *O*6-[4-3-(pyridyl)-4-oxobut-1-yl]-2'-deoxyguanosine (O^6-POB-dG).[103]

These POB adducts have recently been detected in DNA from tissues of NNK-treated rodents. Of these adducts, O^6-POB-dG and O^2-POB-dT are the most stable, while 7-POB-dG and O^2-POB-dC either lose the sugar moiety to generate nucleobase adducts or dealkylate to

release HPB under neutral thermal hydrolysis conditions.[104] Quantitation of the four POB adducts in NNK-treated calf thymus DNA indicates that 7-pobG is formed in the greatest relative amount *in vitro*, followed by O^2-POB-dC, O^2-POB-dT, and O^6-POB-dG.[105]

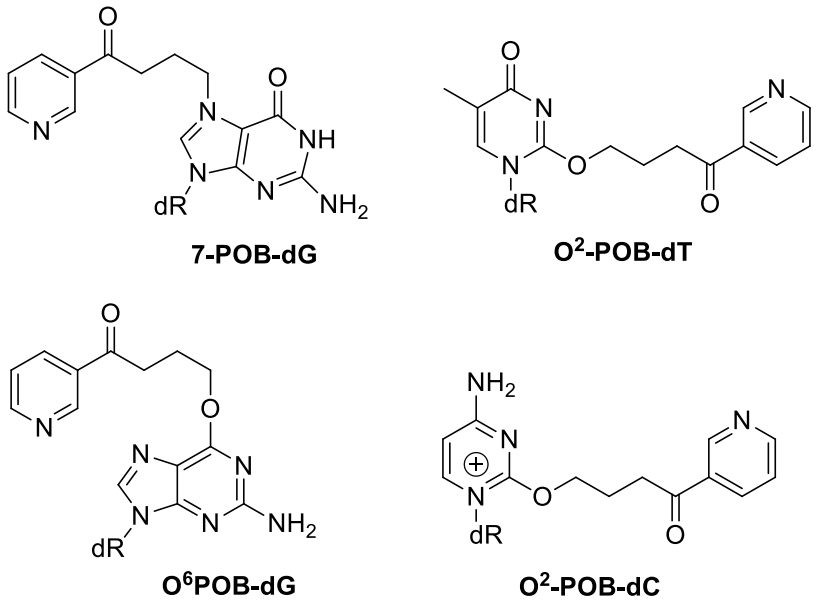


Figure 6: DNA adducts formed from the NNK pyridyloxobutylating pathway

POB DNA adducts have been detected in lung tissue and it persist for a significant period. Pyridyloxobutylating agents are also mutagenic in the Ames Assay. Mutagenicity of POB adducts showed considerable difference in mutations. Spontaneous depurination of 7-POB-dG gives rise to apurinic sites.[104] However, it was suspected that O^2-POB-dT play a role in the etiology of tumors in the respiratory mucosa and esophagus in rats treated with higher doses of NNN.[106]

Furthermore, the cytotoxic and mutagenic properties of the pyridyloxobutylating agent, 4-(acetoxymethylnitrosamino)-1-(3-pyridyl)-1-butanone (NNKOAc), were investigated in repair-proficient and -deficient Chinese hamster ovary cell lines (CHO). Analysis of the mutational spectra in the *hprt* gene showed mainly point mutations at AT base pairs, while GC to AT transitions were found in less frequent.[106]

Despite these studies, there are only limited data as to which POB adducts cause these mutations and whether or not these mutations are critical to tumor initiation. It is known that the pyridyloxobutyl DNA adduct O^6-POB-dG leads to mutations at GC base pairs. Site-specific mutagenesis studies demonstrate that O^6-POB-dG produces GC to AT transitional mutations in bacteria, while this adduct produces not only GC to AT transitional mutations but also GC to TA transversions and deletions in human kidney cells.[107] However, O^6-POB-dG has shown to be efficiently repaired by O^6-alkylguanineDNA-alkyltransferase (AGT).[108] Moreover, *in vitro* bypass of O^2-POB-dT suggested that this adduct miscodes during DNA replication.[109] While some information exists about mutations caused by POB DNA adducts, precise contribution of these adducts in carcinogenesis is yet to be investigated.

1.6 DNA Repair

Thousands of DNA damages arise every day challenging the integrity of the genome. DNA damage interferes with replication and transcription, inducing mutations and chromosomal aberrations. The toxic and mutagenic consequence of the DNA damage can be eliminated or minimized, however, by damage signaling and repair mechanisms, in order to preserve the integrity of the genome. These mechanisms are collectively known as DNA-damage response (DDR) and are coordinated to detect lesions, signal their existence, and trigger repair. DDR responds to DNA damage by slowing or arresting cell-cycle progression by activation cell-cycle “checkpoints”, altering chromatin structure at the damage site, and

also by transcriptional activation and posttranscriptional modification of DNA repair proteins.[110] Defects in these pathways lead to extreme sensitivity towards DNA damaging agents and may lead to diseases. Although the choice of the response system differs depending on the type of the lesion and the cell-cycle phase, they usually occur by a common general mechanism. [110-112]

1.6.1 DNA Damage Signaling and Cell-cycle Checkpoints

Cellular checkpoints coordinate DNA repair with chromosome metabolism and cell-cycle transitions. Theses checkpoints are cellular proteins that get recruited to DNA lesions by repair complexes to generate DNA intermediate structures. These structures subsequently activate checkpoint responses. Following activation, checkpoint signals are amplified and transmitted by the checkpoint transducers to downstream targets such as DNA repair apparatus and cell cycle machinery. [111]

Key DDR-signaling components of mammalian checkpoint machinery are the phosphoinositide 3-kinase like protein kinases, ataxia–telangiectasia mutated (ATM), ataxia–telangiectasia Rad3-related (ATR) and DNA dependent protein kinase catalytic subunit (DNA-PKcs). ATM and DNA-PKcs are predominantly recruited to double-stand breaks (DSBs), while ATR responds to a wide variety of lesions and is activated by replication protein A (RPA)-coated single-stranded DNA (ssDNA). ATM and ATR target transducer kinases, CHK2 and CHK1, act together with ATM and ATR. They reduce cyclin-dependent kinase (CDK) activity by different mechanisms, some of which are facilitated by *p53* (a tumor suppressor gene) transcription factors.[111-113] Once activated, ATM/ATR signaling triggers DNA repair by inducing certain DNA repair pathways. Upon effective removal of the DNA damage, DDR inactivation ensues, allowing continuation of cellular functions.

However, if damage cannot be removed, DDR triggers for apoptosis or cellular senescence. [111]

1.6.2 DNA repair pathways

Repair of different types of DNA lesions rely on different DNA repair pathways. There are several DNA repair pathways, including DNA damage reversal to remove base alterations and also multiple distinct mechanisms for excising the damaged bases, termed base excision repair (BER), nucleotide excision repair (NER), and mismatch repair (MMR). Homologous recombination (HR) and non-homologous end joining (NHEJ) systems repair strand breaks in DNA.[114]

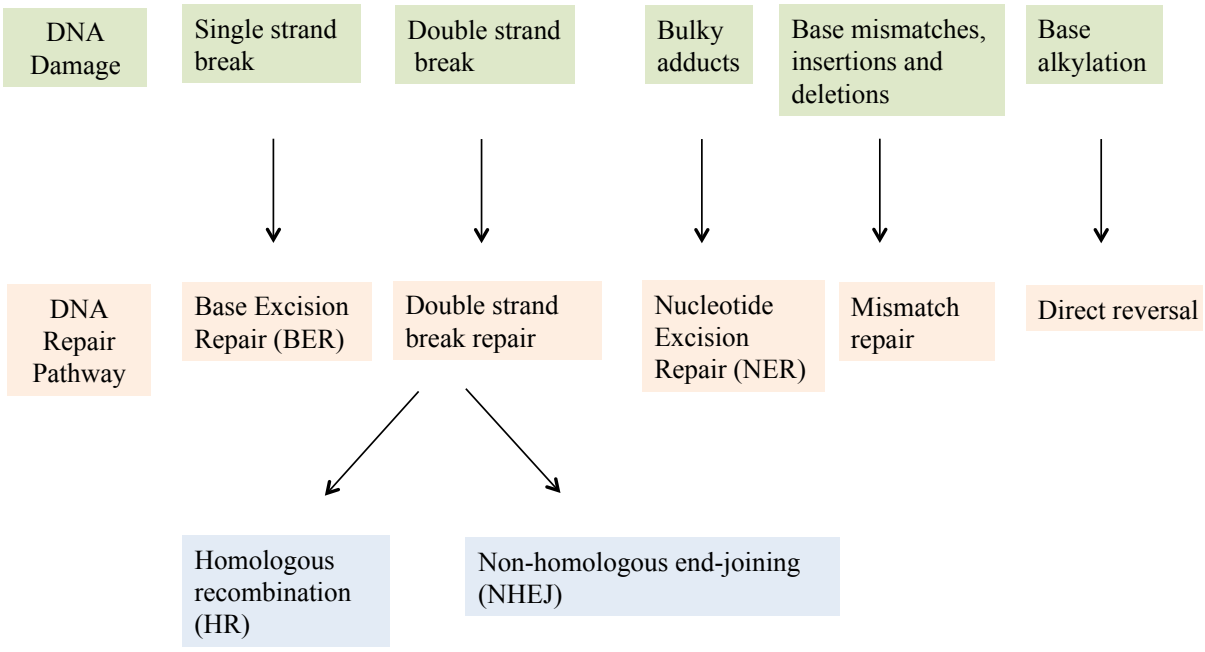


Figure 7: DNA repair pathways. Adapted from ref[114]

1.6.3 Reversion repair

Mammalian reversion repair includes single-step removal of O^6-alkylation lesions by O^6-methylguanine-DNA methyltransferase (MGMT) and DNA-damage reversal by AlkB and its homologues proteins. Apart from protection from alkylating agents, MGMT also is involved in protection from sister chromatid exchange (SCEs) and chromosomal aberrations.[115, 116] AlkB is specifically involved in oxidative demethylation of DNA alkylation damage, such as, 1-methyladenine and 3-methylcytosin in both single-stranded and double-stranded DNA.[117]

1.6.4 Base Excision Repair

Base excision repair is responsible for removing an array of DNA lesions formed by ROS, hydrolytic reactions, alkylating agents and single-strands breaks induced by ionizing radiation. Oxidized DNA lesions such as 8-oxo-7,8-hydroxyguanine (8oxoG) and thymine glycol, alkylated bases including 7-methylguanine, O^6-methylguanine, 3-methylguanine, 3-methyladenine, and lesions formed by hydrolysis like deoxyuracils and abasic sites (AP) are examples of lesions removed by BER. The first few steps in BER are to recognize the damaged base by DNA glycosylases, flip the suspected base out of the DNA helix, and remove it by hydrolyzing the *N*-glycosidic bond. This results in formation of an abasic site, which also can be formed spontaneously. [115, 118-120] Subsequent step involves strand incision at the AP site by AP endonuclease (APE1). After insertion of the appropriate base, repair could occur by one of the two different mechanisms, short-patch or long-patch repair. Short-patch repair is usually the preferred repair pathway in mammals, in which DNA pol β performs a one–nucleotide gap filling reaction and subsequent removal of 5'deoxyrobose -5-phosphate (baseless sugar moiety) by its lyase activity. Finally the nicks are sealed by XRCC1-ligase 3 complex. In addition to pol β, pol δ/ε and proliferating cell nuclear antigen

(PCNA) are used for DNA synthesis in long-patch repair, resulting in a longer repair patch of 2-10 nucleotides. Following DNA synthesis, flap endonuclease (FEN1) is used to displace DNA flap formed during polymerization and gaps are ligated by DNA ligase 1. [115, 120] Although in some human tumors sporadic changes have been reported in *POLB,* there are no disorders in inherited deficiencies in BER. [120]

1.6.5 Nucleotide Excision Repair

In terms of adaptability in lesion recognition, nucleotide excision repair (NER) is the most versatile DNA repair pathway, which is also a biochemically-complicated process. Unlike BER, NER process involves as many as 30 distinct proteins in a large complex known as nucleotide excision repairosome. NER plays an important role specifically in G1 phase to remove bulky, helix-distorting DNA lesions, such as the DNA adducts formed by the UV light as well as various DNA damages caused by carcinogenic and chemotherapeutic agents that are not repaired by BER. However, its activity is not just restricted to G1 phase of cell cycle.[29, 112]

NER operates via two sub-pathways that depend on the initial lesion recognition events, transcription-coupled NER (TCR), which specifically target transcription-blocking lesions and global genomic repair (GGR).[119] These repair sub-pathways essentially follow the same sequences except for the initial DNA damage recognition step. The damage recognition event in TCR is triggered by the lesion-induced stalling of RNA polymerase II (pol II) at the transcriptionally active strand. The damage recognition in GGR involves different set of proteins and is totally independent of transcriptional status of the DNA strand. After the initial damage recognition event, the two sub-pathways converge and use the same set of proteins and sequences for dual incision, repair, and the final ligation steps. A major difference from BER and one of the key aspects of NER is that the damage is excised as a

22-30 base oligonucleotide producing a ssDNA fragment, and the gapped region is subsequently repaired by DNA polymerase and associated factors.[121]

NER is an important and vital DNA repair pathway and defects in NER are associated with several human (NER-deficiency) syndromes including Xeroderma pigmentosum (XP), Cockayne syndrome (CS), trichothiodystrophy (TTD), and UV-sensitive syndrome (UVSS), all of which show characteristic UV-sensitivity.[122]

1.6.6 Mismatch Repair (MMR)

The mismatch repair system removes nucleotide mismatches caused by spontaneous and induced base deamination, oxidation, methylation and replication errors. These replication errors include insertion/deletion loops (ranging from one to ten or more bases) resulting from slippage during replication of repetitive sequences or during recombination. There are four principal steps of MMR comprising: (1) initial mismatch recognition by MutSα (MSH2-MSH6), (2) recruitment of additional MMR factors, (3) search and recognize mismatched newly synthesized strand followed by degradation past the mismatch, and (4) final re-synthesis of the excised strand.[123]

MMR is an important strategy for maintaining genome integrity and stability and is clearly illustrated by MMR deficiency syndrome like hereditary non-polyposis colorectal cancer (HNPCC). Defects in MMR also increase the mutation rate leading to oncogenesis.[124]

1.6.7 DNA Double-strand Break Repair

DNA double strand breaks (DSB) are perhaps the most toxic damage found in genome and are repaired by different pathways including homologous recombination (HR) and non-homologous end-joining (NHEJ), which are error free and error prone, respectively.

DSB can be formed from ionizing radiation or X-rays, free radicals, carcinogens and during replication of SSB.[125]

Upon DSB detection, complex cascade of reactions trigger cell-cycle checkpoints and recruit repair factors to the site of the DNA damage.[126] The choice of the repair pathway for DSB depends on the cell-cycle phase. The predominant DSB repair pathway in G0/G1 phase is NHEJ, due to lack of sister chromatids and highly condensed of structures of chromatin, whereas HR occurs during the late S and G2 phase of the cell-cycle.[127] During NHEJ, two ends of DSB are recognized by Ku proteins, which in turn activate protein kinase DNA-PKcs that recruit end processing enzymes, polymerases and ligase IV. Furthermore, the presence of nucleosomes or other structures may assist NHEJ process as the broken ends are held together at a close proximity within the complex. This process occurs without any sequence homology between the two broken DNA ends. Therefore, it is error prone as few nucleotides could be gained or lost during the process.[9, 128]

In contrast, HR is restricted to S and G2 of the cell-cycle and is a high fidelity process, as it uses sequence homology of the sister chromatids as a template for damage repair. In HR, first DNA ends are resected in the 5'-3' direction by nuclease activity. Then the damaged chromosome come in to contact with the region of sequence homology of the undamaged DNA molecule, which in turn is used as a template for re-synthesis of the damaged strands. After DNA synthesis, ligation, and branch migration, resulting Holiday junctions (DNA crossovers) are resolved to obtain two intact DNA molecules.[126]

NHEJ and HR are the main repair modes for maintaining genome integrity. Defects in DSB repair are associated with many disorders including Ataxia telangiectasia (AT), AT-like disorder, Nijmegen breakage syndrome. Cancer predisposition, immunodeficiency, hypersensitivity to X-rays and chromosomal instability are also some of the conditions

associated with these disorders.[126] Inherited mutations in at least one of these repair proteins (BRCA1/BRCA2) make one highly susceptible to breast cancer incidences.[129]

1.7 Translesion Synthesis (TLS)

In all living organisms DNA is replicated with high efficiency and extraordinary fidelity. These replicative polymerases are comprised of a tight active site, enabling it to accurately match the incoming nucleotide to the template base by the appropriate Watson-Crick base pairing, conferring them the property of high fidelity. In addition, the 3'–5' exonucleases activity of replicative polymerases removes any base that might, on rare occasions, be misinserted.[130] However, DNA is under constant attack by various endogenous and exogenous damaging agents. Despite the existence of highly refined DNA repair pathways, it is likely that some DNA damages escape repair and intervene replication machinery to stall replication fork progression. Stalled replication forks are highly unstable and can cause genome instability by chromosomal aberrations.[131] To avoid such deleterious consequences, cells have developed an important specialized mechanism known as translesion synthesis (TLS), a mechanism that is able to tolerate such DNA damages.[132-134] During TLS process, specialized DNA polymerases insert a base opposite the lesion and are able to continue DNA synthesis past damaged site. These specialized polymerases (TLS pols), most of which belongs to Y family, have low fidelity and operate at a lower speed with low processivity compared to the replicative DNA polymerases. Hence TLS is an error-prone process and is responsible for many of the point mutations in the cell and, in particular, increased point mutations found in the cancer genome.[11, 134] The basic necessity for the presence of TLS pols reflects a trade-off between maintaining genomic instability and accumulation of mutations due to lesion bypass by TLS.[135, 136]

It has been suggested that TLS is a multistep process, involving several "polymerase switches" where multiple TLS polymerase act in a sequential manner for lesion bypass. During this process, when replicative polymerase is stalled at a damaged site, it is displaced and a

specific TLS polymerase is recruited to the primer terminus of the damaged DNA strand. This TLS polymerase, the so called "inserter", inserts a correct or an incorrect base opposite the lesion. The inserted base is subsequently extended past the lesion by an "extender", the same TLS polymerase or a second TLS polymerase to complete TLS.[137] After synthesizing a short tract of DNA, all TLS polymerases are dissociated from the primer terminus and replaced with replicative polymerase to resume DNA synthesis.[138]

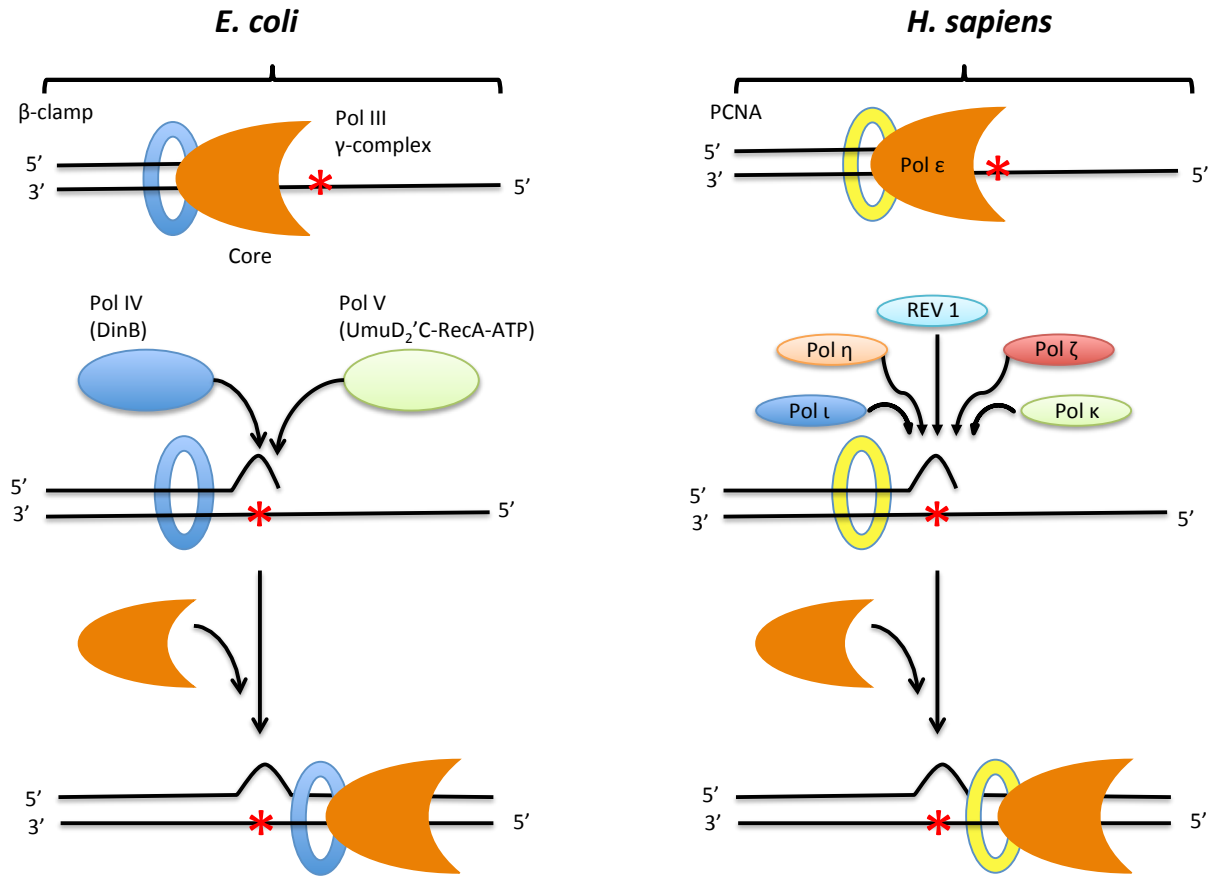


Figure 8: Evolutionary conserved role of DNA polymerase in translesion synthesis (TLS). Adapted from ref[135]

It is unavoidable that all living cells are exposed to endogenous and exogenous damaging agents, which could potentially cause lethal DNA damage. However, TLS pols allow cells to

tolerate such damages accompanied by undesirable mutagenesis. Therefore, it is crucial to strictly regulate activity of TLS pols to avoid any major deleterious consequences. It has been found that all Y family DNA pols are concentrated in the nucleus. During the S phase of the cell-cycle, pols ι, η and REV1 (mammalian Y-family pols) are transferred to replisome, where they co-localize with the polymerase sliding clamp, proliferating cell nuclear antigen (PCNA) and other accessory proteins. Only when necessary, these TLS pols are recruited to the replication fork replacing replicative polymerase.[134, 138, 139]

TLS pols (mostly of the Y-family) are found in all domains of life, and their existence in higher organisms suggests an essential evolutionary conserved role.[135] Y-family TLS polymerases include, archaeal Dpo4 from *Sulfolobus solfataricus*, bacterial, such as the *Escherichia coli (E. coli)* polymerase IV, V and mammalian pols η, ι, κ, REV1. In addition to Y-family polymerases, mammalian polymerase ζ (of the B-family) is also considered a TLS pol. Further *E. coli* polymerase II (high fidelity with exonuclease proof-reading activity) is also implicated in TLS.[132] A brief summery of individual TLS polymerase will be described in the following section.

1.7.1 Y-family DNA polymerase

On the basis of phylogenetic relationships, DNA polymerases have been grouped into six families, family A, B, C, D, X, or Y. Y-family polymerases (Y-family pols) are specialized DNA polymerases capable of incorporation of nucleotides opposite templates that are non-instructional to replicative enzymes. There are five sub families of Y-family pols including, UmuC, DinB, pol ι, pol η, and Rev1.[140] Key features of Y-family pols that differ from other regular pols are: the ability to bypass damaged bases and their reduced fidelity on replicating undamaged templates. This is mainly due to more capacious less constrained catalytic site of these pols.[135, 141, 142]

The general tertiary structure of replicative pols features a right hand topology with palm, finger and thumb domains and is conserved among different polymerase families. In addition to these sub domains, Y-family pols catalytic region is composed of an additional little finger domain. This little finger domain, also known as polymerase-associated domain (PAD), is unique to the Y-family pols, has more sequence variability than the catalytic core, implicated in polymerase specificity for certain lesions, involves in stabilization of template DNA with bulky lesions, and influence overall activity.[143] Furthermore, the decreased processivity and the poor fidelity is associated with the "stubbier" finger and thumb domains of Y-pols, as they make fewer contacts with the incoming nucleotide and the template DNA. Collectively, all these features allow Y-family pols to (mis)incorporate a nucleotide opposite a damaged DNA and extend past the lesion by a TLS polymerase. Additionally, lack of 3'-5' proofreading exonuclease activity of these Y-family pols results in increased mutagenicity during TLS.[135]

The error rate of replicative polymerases of families A, B, and C are extremely low and lies between 10^{-6} to 10^{-8} and further decreased to 10^{-8} to 10^{-10} by auxiliary proteins such as PCNA, replication protein A (RP-A) and postreplicative MMR. On the other hand, the error rate of TLS pols ranges from 10^{-1} to 10^{-3} for replication of undamaged DNA.[144]

1.7.2 Mammalian Y-family DNA polymerase

1.7.2.1 Polymerase eta (pol η)

Pol η encoded by *POLH* is considered to be one of main TLS pols found and plays a key role in the accurate replication of the UV-induced DNA damage, the cyclobutane pyrimidine dimers (CPDs). It is noteworthy that pol η is capable of inserting A-A opposite a *cis-syn* T-T CPD with similar accuracy to unmodified T-T. Nevertheless, on undamaged template pol η is significantly error prone.[145] Defective (mutations) pol η is specifically responsible for clinical

syndrome *Xeroderma pigmentosum variant* (XPV), characterized by hypersensitivity to UV-induced skin pigmentation change and increased sensitivity to skin cancer. It has been suggested that in the absence of pol η, pol κ or ι in cooperation with pol ζ, serve as an error prone TLS polymerase to bypass UV-induced lesions.[146] Recent structural studies on pol η reveal that its active site is large enough to accommodate both of the linked T-T dimer and also stabilize incoming dA opposite the lesion. Furthermore, specialized β clamp (positively charged) in the little finger domain that acts as a splint, ensures proper reading frame without any template slippage. To avoid any erroneous replication of the undamaged template, pol η is displaced from the DNA when three bases beyond the lesion have been inserted.[147]

1.7.2.2 Polymerase iota (pol ι)

Human pol ι encoded by *POLI,* exhibits 10^{-5} fidelity depending on the lesion and is considered unique among Y family pols. Pol ι incorporates nucleotides opposite the four template bases with varying efficiency and fidelity. It incorporates nucleotides opposite template purines with a much higher efficiency and fidelity than that for pyrimidines. When copying dA it has misincorporation fidelity of $1\text{-}2\times10^{-4}$. But while replicating dT, the enzyme misinserts dG 3-10 times more accurately than correct dA. This is mainly due to structural features of key residues in the finger domain that uses Hoogesteen base-paring for DNA synthesis.[148] Additionally, it has been postulated that these specialized configurations of pol ι may have implications on BER.[149] It has been reported that opposite an abasic site and the 3′T of the (6-4) TT photoproducts, pol ι incorporates a G, an A, or a T with an equal efficiencies.[150] Furthermore, pol ι can proficiently incorporate nucleotides opposite an N^2-adducted guanine.[151]

1.7.2.3 Polymerase kappa (pol κ)

Human pol κ encoded by *POLK* belongs to DinB subfamily of Y-family pols.[152] It is one of the highly conserved polymerase in the Y-family with homologs in bacteria, archaea, and eukaryotes. Pol κ is considered to be the most accurate on undamaged DNA amongst Y-family pols and has a lower tendency to make one base deletions as compared to bacterial homolog dinB.[134] Its catalytic site can accommodate only a single Watson-Crick base pair; therefore, pol κ is blocked by photodimers. Nevertheless, it is capable of accurately bypassing many lesions including some N^2-dG adducts, albeit with low efficiency.[136, 153] Furthermore, it was suggested that pol κ cooperates with pol δ during NER being recruited by ubiquitinated PCNA and scaffold proteins.[154]

1.7.2.4 REV1

REV1, which is encoded by *REV1*, is usually restricted to incorporation of dC opposite template dG or a narrow range of lesions including abasic sites and N^2-dG adducts.[155, 156] However, this is an unusual process, where there is no direct pairing of incoming dC with the dG in the template strand. Instead, template dG is swung out of the helix and temporally coordinates it with a specialized loop within the little finger domain. Then Arg324 of REV1 fill the space previously occupied by the dG, or provided by abasic site, forming hydrogen bonding with incoming dCTP. Therefore, this mechanism allows bypass of bulky dG, maintaining the specificity for the correct base and also template-independent dC transferase activity.[135, 136, 157] In addition to the catalytic role of REV1, it has been proposed that key non-catalytic function of REV1 in vertebrates is to help coordinate polymerase switch between the replicative polymerase and TLS polymerase during TLS.[158]

1.7.3 Polymerase zeta (pol ζ)

Pol ζ belongs to family B, which includes high fidelity replicative polymerases pol delta (δ), epsilon (ε) and alpha (α). Unlike replicative pols, pol ζ is error-prone and lacks 3'-5' exonuclease activity, albeit belonging to the same family of high fidelity enzymes. Both human and yeast pol ζ are heterodimeric proteins containing a catalytic subunit (REV3) and the accessory subunit (REV7). Accessory subunit stimulates the activity of REV3 and links pol ζ to REV1 during TLS.[136] *In vitro* kinetic studies have shown that pol ζ is generally inefficient in incorporation of nucleotides opposite the lesion but have higher extension efficiency of the mismatched primer termini.[159]

Both subunits of Pol ζ, REV3 and REV7, are indispensible for UV-induced mutagenesis, abasic site mutagenesis, and chemically-induced mutagenesis. Compared to the other TLS pols, pol ζ's action, however, is not limited to the error-prone bypass of DNA lesions, it also facilitates error-free bypass of a variety of DNA lesions.[136, 140]

1.7.4 Bacterial (*Escherichia coli*) Y-family Polymerase

1.7.5 Polymerase IV (pol IV)

Escherichia coli (*E. coli*) pol IV is encoded by SOS-controlled *DinB* gene and was first identified by Kenyon and Walker in 1980.[160] However, its function remained unclear till late 1990.[161] Later, it was found that it is involved in damaged tolerance pathway in *E. coli*. *DinB* orthologs are found in all domains of life and are the most ubiquitous of Y-family DNA pols. Similar to all Y-family pols, pol IV also lacks 3'-5' exonuclease proofreading activity, and therefore, has low fidelity. Highly elevated pol IV levels, almost 10 fold, are found (from 250 to

2500 molecules/cell) during bacterial SOS induction (a damage tolerance mechanism).[162] Experimental evidences have shown that –1 frameshift mutation occurred within short stretches of G residues with overexpressed *DinB* levels.[163] Furthermore, pol IV is able to bypass various lesions, particularly found in minor grove of DNA, including 8-oxo-dG, O^6-medG, abasic site (AP), *N*-2-acetyaminofluorene (AAF), cyclobutane pyrimidine dimer (CPD), 6-4 photoproducts, and many others.[164] Experimental observations under normal physiological conditions have shown that, pol IV does not contribute to chromosomal mutations.[165] However, pol IV is able to specifically extend replication errors made by pol III (replicative polymerase).[166] In addition to its role in TLS, pol IV appears to be involved in non TLS processes including adaptive mutagenesis, competitive fitness during stationary phase, recovery of arrested transcription events, prevention of hydroxyurea-induced cell death, and also in AP lyase activity.[164]

1.7.6 Polymerase V (pol V)

Pol V encoded by *umuD'$_2$C* is the major TLS polymerase found in *E. coli*. and by far one of the most studied bacterial TLS polymerases. Pol V is able to bypass a variety of lesions with low fidelity leading to increased cellular mutagenesis after DNA damage.[167] It is a heterotrimeric complex that has shown to require RecA nuclear protein filament for optimal activity. Furthermore, it is evident that expression of pol V is limited to SOS induction and occurs about 40-50 min after UV irradiation. Expression of pol V during SOS response is highly regulated by *umuDC* operon, which contains UmuC and UmuD' homodimer, an accessory subunit derived from UmuD. This kind of tight control is achieved at several levels including regulation of transcription level by LexA repressor, requirement of posttranslational modification from UmuD to UmuD', and also by limiting the half- life of the protein.[168]

1.7.7 *Escherichia coli* Polymerase II (pol II)

E. coli pol II is encoded by *polB*, is a high fidelity DNA polymerase that is involved in bacterial TLS. Unlike most TLS pols, pol II possesses 3'-5' exonuclease proofreading activity, although its expression is increased with SOS induction. Compared to other replicative polymerases in *E. coli* (pol I and pol III), pol II can efficiently bypass abasic sites. Even though poll II is a high fidelity polymerase, it can generate -2 frameshift mutations to bypass C8-guanine adduct of *N*-2-acetylaminofluorene (G-AAF). Additionally, pol II is involved in TLS of 3,N^4-ethenocytosine, intra-strand cross-links induced by butadiene, and inter-strand cross-links.[169]

1.8 Bacterial SOS Response

In nature different bacterial species have adapted to live in environments of harsh conditions like high temperature, pH shifts, and extreme chemical stress. As part of adaptation, bacteria respond to such hostile environments by switching global pattern of gene expression and have developed distinct response systems, including stress response, SOS response, and heat-shock response. Similarly, bacterial genome is constantly exposed to endogenous and exogenous damaging agents and DNA replication is impaired due to DNA damage. Bacteria, such as *E. coli,* respond to such DNA damage through SOS response. SOS gene products are involved in multiple DNA metabolism processes, such as, recombination, repair and replication.

During SOS induction, more than 40 genes are expressed, which are regulated by the LexA protein. Under physiological conditions, LexA dimer, in its native state, is bound to 20bp consensus sequence, conventionally called the SOS box in the promoter region of the SOS genes, limiting their expression levels.[170] Upon DNA damage, single-stranded regions of DNA are created downstream from the DNA lesion. At this point, RecA is activated to become a co-

protease and forms a filament around the single-stranded DNA regions in an ATP dependent manner and becomes conformationally active. The activated RecA facilitates self-cleavage of LexA. Cleavage of LexA decreases repressor levels and increases the expression of SOS genes, including three low fidelity polymerases, which are involved in translesion synthesis, pol II, pol IV and pol V. This system, though quite complex and intricate, is also error-prone.[169, 171]

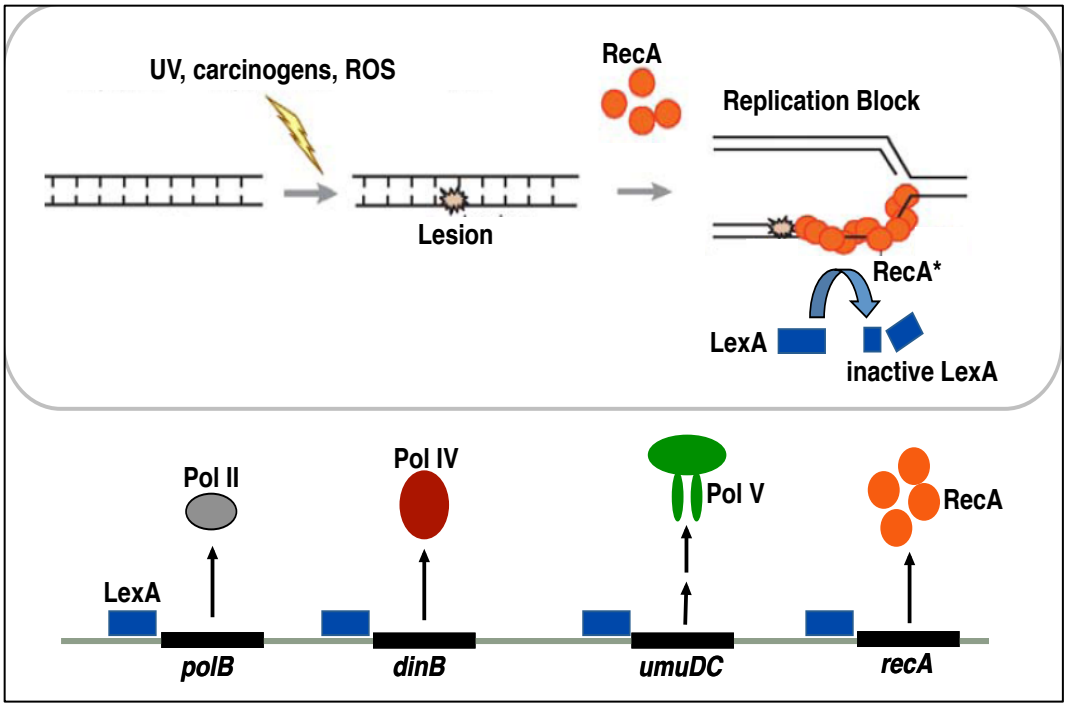


Figure 9: SOS response mechanism in bacteria. Adapted form ref [169]

2 Scope of the Dissertation

Chapter one of this dissertation is focused on the replicative bypass of an abasic site [apurinic/apyrimidinic (AP)] in *Escherichia coli* and in human embryonic kidney (HEK293T) cells. Abasic sites are the most common DNA damages, opposite which dAMP is frequently inserted ('A-rule') in *Escherichia coli*. Nucleotide insertion opposite the AP-site in eukaryotic cells is less certain. Depending on the assay system and the type of cells, a 'C-rule', 'A-rule', or the lack of specificity have been reported. Several studies also show that the sequence context modulates nucleotide insertion opposite AP-site. In this work, using an extra-chromosomal probe I have compared replication of tetrahydrofuran (Z), a stable analog of AP-site, in *E. coli* and human embryonic kidney 293T cells in two sequence contexts. I determined that efficiency of translesion synthesis or viability of the AP-site construct in *E. coli* was less than 1%; yet their magnitude was significantly different for the two sequence. The difference in viability increased even more in pol V-deficient strains. In uninduced *E. coli*, for both sequences targeted one-base deletions occurred as the predominant mutation, which was decreased significantly upon induction of the SOS functions. The full-length products with SOS primarily involved dAMP insertion opposite the AP-site, which was largely carried out by pol V. In contrast to these results in *E. coli*, I found that the viability is 2 to 3 orders of magnitude higher in the human cells, and the 'A-rule' is more rigidly followed as Z→T substitutions occurred as the major mutation for both sequences. Furthermore, in HEK293T cells, I found that targeted one-base deletions are negligible. siRNA knockdown of Rev1 or pol ζ established that both these polymerases are vital for AP-site bypass. However, neither polymerase is indispensable, suggesting roles of additional DNA polymerases in AP-site bypass in human cells.

The focus of chapter 2 is to investigate the replicative bypass of γ radiation or oxidatively induced 8,5'-cyclopuries in *E. coli* and HEK293T cells. Ionizing radiation and other processes that give rise to reactive oxygen species generate many lesions in DNA, including 8,5'-cyclo-2'-deoxyadenosine (*S*-cdA) and 8,5'-cyclo-2'-deoxyguanosine (*S*-cdG) tandem lesions. In this study, I have replicated a plasmid containing either *S*-cdA or *S*-cdG in *Escherichia coli* and HEK293T cells. I found that *S*-cdA is mutagenic and highly genotoxic in *E. coli*. Furthermore, viability and mutagenicity of *S*-cdA are dependent on functional pol V, but mutational frequencies (MF) and types varied in pol II- and pol IV-deficient strains. Furthermore, the analysis of progeny from HEK293T cells showed that *S*-cdG is much more mutagenic compared to *S*-cdA and the major types of mutations are *S*-cdG→T followed by *S*-cdG→A. In the case of *S*-cdA, A→ T substitutions are observed as the predominant type of mutations followed by A → G transitions. Interestingly, I also found that for *S*-cdA, a significant number of semi-targeted mutations occurred 5' to the lesion. Using si-RNA induced knockdown of TLS polymerases in HEK293T cells, I found that pol ι or pol η are involved in the nucleotide incorporation opposite the cyclopurines, whereas pol κ or pol ζ are involved in the extension past the lesions.

Chapter 3 discusses the replicative bypass of O^2-alkylthymidine lesions generated by the tobacco-specific nitrosamines 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK). Metabolic activation of NNK generates a variety of DNA adducts including O^2-methylthymidine (O^2-Me-dT) and O^2-{4-(3-pyridyl-4-oxobut-1-yl]thymidine (O^2-POB-dT) lesions. In this study, I have investigated the mutagenic and genotoxic properties of these O^2-alkylthymidine lesions by replicating single-stranded plasmids containing a site-specifically incorporated O^2-Me-dT or O^2-POB-dT in HEK293T cells. I found that the bulkier O^2-POB-dT is highly genotoxic in human

cells, while O^2-Me-dT showed much lower genotoxicity. Furthermore, I found that in HEK293T cells, O^2-Me-dT and O^2-POB-dT are highly mutagenic and the major type of mutations observed for both lesions are targeted T→A substitutions. Using siRNA induced knockdown of TLS pols, I determined that pol η, pol ζ, and REV1 are involved in the lesion bypass of both O^2-Me-dT and O^2-POB-dT. In conclusion, these studies provide the identities of the DNA polymerases that play important roles in both error-free and error-prone bypass of these lesions. These investigations also provide additional information on the interplay of the various TLS polymerases in lesion bypass.

3 Materials and Methods

3.1 Materials

3.1.1 Chemicals, enzymes, plasmid DNA, cells, siRNA and instruments

All chemicals and reagents, unless otherwise specified, used for gel electrophoresis (e.g., agarose, Tris base, boric acid, Na_2EDTA, urea, acrylamide, *N*,*N*'-methylene bisacrylamide, TEMED, APS), and preparation of biological media (e.g., agar, yeast extract, bactotryptone, sodium chloride) were purchased from Fisher Scientific (Agawam, MA). Phenol (Ultrapure™ Buffer-Saturated Phenol) for phenol-chloroform extraction was obtained from Invitrogen. Hybridization buffer (PerfectHyb™ Hybridization Buffer 1X) chloroform and isoamyl alcohol were purchased from Sigma-Aldrich (St. Louis, MO). [γ-^{32}P] ATP for gel assays and hybridization were obtained from PerkinElmer Life and Analytical Sciences (Boston, MA).

The media required for mammalian cell culture such as fetal bovine serum (FBS), Dulbecco's modified Eagle's medium (DMEM), non-essential amino acids (NEAA), penicillin/streptomycin (PS), Dulbecco's phosphate buffer saline (D-PBS) solution, 0.25% trypsin-EDTA, Opti-MEM I Reduced Serum Medium and transfection reagent (Lipofectamine™ 2000) were purchased from Invitrogen Corp. (Carlsbad, CA).

All enzymes needed for construct preparation and radiolabeling including *Eco*RV, *Eco*RI T4 DNA ligase, T4 polynucleotide kinase (PNK), uracil DNA glycosylase (UDG), and exonuclease III (exo III) were obtained from New England Biolabs (Beverly, MA).

pMS2 plasmid was a gifts from M. Moriya (SUNY, Stony Brook, NY)

Escherichia coli strain DH10B and DH12S were purchased from Invitrogen (Carlsbad, CA). HEK 293T cells were obtained from ATCC (Manassas, VA).

The *E. coli* strains used were AB1157 [F^- *thr-1 araC14 leuB6*(Am) Δ(*gpt-proA*)*62lacY1 tsx-33 supE44*(AS) *galK2*(Oc) *hisG4*(Oc) *rfbD1 mgl51 rpoS396*(Am) *rpsL31*(Str^r) *kdgK51 xylA5 mtl-1 argE3*(Oc) *thi-1*], pol II^- (AB1157 but *polB*Δ*1*::Ω Sm-Sp), pol IV^- (AB1157 but Δ*dinBW2*::cat), GW8017 (AB1157 but *umuDC595*::cat), and pol II^-/pol IV^-/pol V^- (AB1157 but *polB*Δ*1*::Ω Sm-Sp *dinB umuDC595*::cat). All *E. coli* strains were provided by G. Walker (MIT, Cambridge, MA).

The unmodified oligodeoxynucleotides were purchased from Integrated DNA Technology (Coralville, IA) and Midland Certified Reagent Company (Midland, TX). The modified oligonucleotides were prepared as reported. Oligonucleotides were desalted using a Sep-Pak C18 cartridge from Waters Corp. (Milford, MA).

Synthetic siRNA duplexes against *POLH* (SI02663619), *POLK* (SI04930884), *POLI* (SI03033310), *REV1* (SI00115311), and AllStars Negative control siRNA (1027280) were purchased from Qiagen (Valencia, CA), whereas the same for REV3 was purchased from Integrated DNA Technologies (Coralville, IA). Sequences of all the siRNAs are listed in Table 1

Table 1: Sequences of siRNA used for TLS knockdown

Gene	siRNA sequence (sense)
POLH	5’-GGUUGUGAGCAUUCGUGUATT-3’
POLK	5’-CUGUGAGUAAAGAGGUUAATT-3’
POLI	5’-GCCUCAUACAGUGAGAUUATT-3’
REV3	5’-GCAAUUUUGAACCUUAUGGTT-3’
REV1	5’-CGGUGGAAUCGGUUUGGAATT-3’

Total RNA isolation was performed using Qiagen's RNeasy plus kit. Primers required for amplifying cDNA of pol η, κ, ι, REV1 were purchased from Qiagen. Quantitect SYBR RT-PCR kit from Qiagen was used for quantitative RT-PCR.

BioRad Gene Pulser II electroporator was used for transformation of *E. coli* cells. PCR amplification of ds DNA from DH10B transformants was performed on Applied Biosystems Gene Amp PCR System 2700. Extraction of ds DNA from DH10B transformants and DNA sequencing were performed in QIAGEN BioRobotic system and Beckman Coulter CEQ 2000XL capillary DNA sequencers, respectively, located at BioTech Center, University of Connecticut, Storrs, CT. For Real-time quantitative RT-PCR analysis, CFX-96 Real-Time PCR Detection System in Dr. Rahul Kanadia's lab (Physiology and Neurobiology Department, UConn, Storrs) was used. Quantification of ds DNA samples in 96-well UV plate was performed in Molecular Device SPECTRAmax 384-plus UV/VIS micro plate reader also located at UConn Biotech Center.

3.1.2 Synthesis and characterization of oligonucleotides

3.1.3 Abasic site (Z) lesion containing oligonucleotides

Z containing and control dodecamers of the sequences 5'-TGCAGZGTCAGC-3', 5'-TGCAGTGZCAGC-3', and 5'-TGCAGTGTCAGC-3' were synthesized by the Midland Certified Reagent Company, Inc (Midland, TX). The oligonucleotides were purified by C18 reverse phase HPLC followed by denaturing polyacrylamide gel electrophoresis. Mass spectrometric analysis by MALDI-TOF and/or ESI-MS verified the molecular weight of the oligonucleotides.

Additionally, MS and polyacrylamide gel electrophoresis analyses indicated that the Z containing oligonucleotides were more than 99% pure.

3.1.4 *S*-cdA and *S*-cdG lesion containing oligonucleotides

The *S*-cdA -modified oligonucleotide 5'-GTGCA*TGTTTGT-3' and *S*-cdG modified oligonucleotide 5'-GTGCG*TGTTTGT-3' containing the DNA sequence of *p53* codons 272-275 in which the lesions were located in codon 273 were synthesized and characterized as reported.[48,172] The unmodified oligonucleotides were analyzed by MALDI-TOF MS analysis, which gave a molecular ion with a mass within 0.005% of theoretical, whereas adducted oligonucleotides were analyzed by ESI-MS in addition to digestion followed by HPLC analysis.

3.1.5 O^2-Me-dT and O^2-POB-dT containing oligonucleotides

The O^2-POB-dT-modified oligonucleotides 5'-GTGCGT*GTTTGT-3', (where T* represents O^2-POB-dT) containing the DNA sequence of p53 codons 272-275 in which the lesion was located in codon 273, was synthesized and characterized as reported.[173] The synthesis of the phosphoramidite of O^2-Me-dT was performed as reported.[103] Incorporation into the oligodeoxynucleotide 5'-GTGCGT*GTTTGT-3', where T* represents O^2-Me-dT, was accomplished as described for O^2-POB-dT. Oligonucleotides were analyzed by MALDI-TOF MS analysis, which gave a molecular ion with a mass within 0.005% of theoretical. The M+1 for the O^2-Me-dT oligodeoxynucleotide is 3713 and we found an *m/z* of 3712. The M+1 for the O^2-POB-dT oligodeoxynucleotide is 3846 and we found an *m/z* of 3847. Unmodified oligonucleotides were analyzed by MALDI-TOF MS analysis, which gave a molecular ion with a mass within 0.005% of theoretical.

3.2 Methods

3.2.1 Purification and analysis of oligomers with polyacrylamide gel electrophoresis

Purity of the site specifically incorporated lesions containing 12mer oligonucleotides were determined by denaturing (containing 8M urea) polyacrylamide gel electrophoresis (20%). A "cold gel" (with unlabeled oligomers) was performed to remove any contaminant oligomers from the desired 12mer. The unlabeled modified oligomers as well as a standard unmodified 12mer containing loading dye (xylene cyanol and bromophenol blue in formamide) were loaded onto 20% PAGE (50 cm length) containing 8M urea. The gel was run at 2250 V ($45V/cm^2$) for 5~6 hours at room temperature until the leading dye bromophenol blue (which runs approximately with a 6mer) reached the bottom of the gel. The desired oligomer-containing band was visualized using a UV (254 nm) light and excised from the gel. The oligonucleotides were extracted from the excised gel by Spin-X (Costar Corning) columns and eluted with water, and subsequently desalted by passing through Sep-Pak C18 cartridges (Waters) with acetonitrile:water (60:40) . The elute from each sample was dried in a speed-vac and the concentration was determined by UV absorbance measurements. A " hot gel" (with radiolabeled oligomers) was performed to analyze the purity of the purified oligomers. The purified oligomers were 5'-labeled with [γ-^{32} P] ATP, mixed with 2 μL of loading dye and run on 20 % denaturing PAGE using above-mentioned conditions. Finally the ^{32}P-labeled oligomers were analyzed by exposing to autoradiography.

3.2.2 Agarose gel electrophoresis for single-stranded vector analysis

Agarose gel electrophoresis was performed to analyze purity of the plasmid preparations and to determine the construct concentration. The ss-pMS2 vector analysis was performed in 1.1% agarose gels and the ss-M13 vector analysis was performed in 1% agarose gels. The agarose gel (1.1% w/v, 0.66 mg agarose in 60 mL of 1x TBE) was casted on a horizontal gel apparatus and gel was allowed to solidify for 30 min. The sample was prepared by mixing 0.2 µg (2 µL) of the plasmid DNA with 2 µL of 6x loading dye (10 mM Tris-HCl (pH 7.6), 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% glycerol, 60 mM EDTA) diluted to 10 µL with water. For ss-pMS2 analysis, gel was run at 110 V for 3 hours at room temperature and for ss-M13 analysis it was run at 100 V for 1h at room temperature followed by 4h at 80 V in cold room (4°C), in 1x TBE buffer. Following electrophoresis, the gel was stained with ethidium bromide (25 µg/µL in 500 mL water) for five minutes, de-stained for 30 min in 1x TBE, and the DNA bands were visualized by UV-light. Based on the net intensity of the standard sample, the concentration of band of interest was determined as follows.

$$\frac{\text{Known amount of standard DNA (µg)}}{\text{Intensity of the standard DNA band}} = \frac{x \text{ µg of DNA}}{\text{Band intensity of the DNA of interest}}$$

3.2.3 Bacterial electro-competent cell preparation and SOS induction

Four milliliters of overnight *E. coli* culture, prepared from a single colony, was transferred to 200 mL of 1x LB media. *E. coli* cells were grown to 1×10^8 cells/mL (~ 2 h) and then harvested by centrifugation at 5000g for 15 min at 0 ^{o}C. This procedure was repeated twice except the cells were resuspended in 40 mL ice-cold deionized water. The bacterial pellet was resuspended in 1 mL of glycerol/water (10% v/v) and kept on ice until further use. For SOS

response, the following additional steps were introduced after the first centrifugation. The cells were resuspended in 50 mL 10 mM $MgSO_4$ and treated with 50 J/m^2 or 20 J/m^2 of UV light (254 nm) in 25 mL aliquots in 150 × 50 mm plastic petri dishes. The cultures were incubated in Luria broth at 37 ^oC for 40 min in order to express SOS functions maximally. Following SOS induction, these cells were centrifuged, deionized, and resuspended in glycerol/water in a similar manner as described earlier except all manipulations were carried out in subdued light.

3.2.4 M13mp7L2 vector containing site specifically incorporate abasic site (Z)

3.2.5 Large scale M13mp7L2 vector preparation

M13mp7L2 ss-DNA was prepared by infection of *E. coli* cells by phage DNA. A single plaque of M13mp7L2 was picked from 1x LB plate containing IPTG and X-gal and re-suspended in 1 mL of sterile distilled water. Hundred milliliters of 2x YT media was inoculated with 100 μL of GW5100 overnight culture, prepared from a single colony, and infected with 300 μL of phage solution. Following incubation of the infected culture at 37 ^oC, 230 rpm, for 6 h, the phage supernatant was recovered by centrifugation at 10 000 rpm for 30 min. The supernatant was transferred to a 2-L flask. Four grams of PEG8000 (final concentration, Cf, of 4%) and 2.9 g of NaCl (Cf = 0.5M) were added to the collected supernatant. The mixture was kept in ice with stirring for 1.5 h, transferred into centrifuge bottles and centrifuged at 10000 rpm, 0 °C for 30 min. The supernatant was discarded and the pellet was re-suspended in 4 ml of 1x TE (pH=8). From the resulting solution, the ss-M13 DNA was purified by extracting twice with phenol-chloroform, once with phenol: chloroform: iso-amyl alcohol (25:24:1) and precipitated with ethanol. Final DNA pallet was dissolved in 200 μL of 1xTE (pH=8), the DNA concentration was determined by UV-absorbance and analyzed by agarose gel electrophoresis.

3.2.6 Construction of ss-M13mp7L2 genome containing a single abasic site

The ssM13mp7L2 (75 μg) DNA was linearized by digestion with large excess of *Eco*RI (750 units) restriction enzyme in 100 mM Tris-HCI (pH7.5), 5 mM MgCI2 and 50 mM NaCI at room temperature overnight. The sample was analyzed on 1% agarose gel to ensure complete linearization of the vector. After digestion, two fold molar excess of a scaffold 50-mer and 1 M NaCI (Cf= 0.025M) were added to the linear ssDNA at a concentration of 100 ng/μL to form a gapped-duplex by annealing of the scaffold. The annealing was carried out by incubating the mixture for 5 min at 70 °C followed by slow cooling to room temperature over a period of 3-4 h and to 4 °C overnight (16h). The proportion of the circularized vector as determined by the agarose gel electrophoresis was 35-45%. The scaffolded sample was divided in to three equal parts (25 μg each) to perform annealing of the 12mer oligonucleotide, either modified or unmodified. Then a 20 molar excess of 12mer insert, either modified or unmodified, was phosphorylated at its 5' end in the presence of 75.6 U of T4 polynucleotide kinase and ligase buffer (Cf= 1x) at 37 °C for 30 min. The kinase enzyme was then inactivated at 65°C for 20 min. The phosphorylated insert, 3200 U of T4 DNA ligase and 10x ligase buffer (Cf = 1x) were combined with the mixture containing the gapped-duplex and incubated at 16°C for 16 h to allow annealing and ligation of the insert into the appropriate gap on the gapped-duplex. Unligated oligonucleotides were removed by being passed through a Centricon -100 apparatus. Any residual uncut-circular or re-ligated DNA was removed by additional round of *Eco*RI digestion. The scaffold that was annealed to the M13 DNA was removed by adding 50x molar excess of anti-scaffold, then heating the mixture at 90 °C for 45 sec, followed by rapid cooling in ice for 30 min. A “mock ligation” was performed with 5 μg of ssM13mp7L2, in which no oligonucleotide was included. The efficiency of ligation was ~40% for both the control and

modified 12-mer. The crude construct was purified by phenol-chloroform extraction followed by ethanol precipitation. Finally the concentration of the purified construct was assessed on 1% agarose gel by comparing the net intensity of the circular band to the net-intensity of the known control plasmid using the KODAK Digital Science™-1D program.

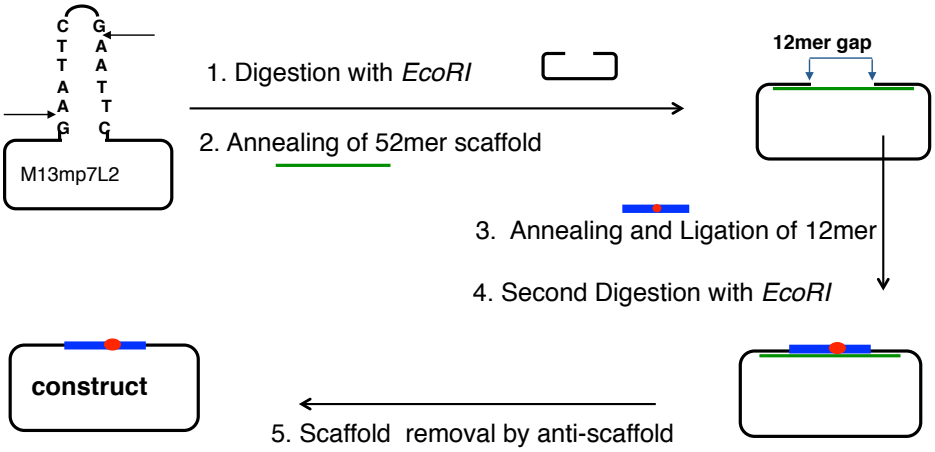


Figure 10: The schematic of single-stranded construct preparation

3.2.7 Transformation of ss-M13mp7L2 vector containing a site specifically incorporated abasic site in *E. coli*

Competent *E. coli* cells were prepared as described above. To induce SOS, the cells were re-suspended in 10 mM $MgSO_4$ solution and aliquots were treated with UV light (254 nm) (50 J/m2) in plastic Petri dishes.

For each transformation, 40 μL of the cell suspension was mixed with 1 μL (100 ng) DNA solution and transferred to the bottom of a cold Bio-Rad Gene-Pulser cuvette. Electroporation of cells was carried out in a Bio-Rad Gene-Pulser apparatus at 25 μF and 2.5 kV with the pulse controller set at 200 Ω. Immediately after electroporation, 1 ml SOC medium was added and the mixture was transferred to 14 mL culture tubes. The cells were incubated for 1 h at 37 ^{o}C to allow for phage replication and subsequently transferred to 1.5 mL centrifuge tubes and centrifuged at 15000 × g (5 min) to isolate the phage-containing supernatant. The supernatant was kept at 4 ^{o}C was used to prepare plates for cytotoxicity and mutation scoring analysis. In order to prepare LB plates containing phage particles, a predetermined volume of the phage solution (diluted to some degree when necessary), was mixed together with 3 mL of molten LB top agar (kept at 45°C), 1 mL of *E. coli* GW5100 culture, and 50 μL of ChromoMax™ IPTG/X-Gal solution and the mixture was poured onto pre-warmed (37 ^{o}C) LB agar plates. The plates were allowed to solidify at room temperature and incubated at 37 ^{o}C overnight (16-18 h) until individual plaques were visible. Progeny plaques were then analyzed by plaque lift hybridization and DNA sequencing. Lesion bypass efficiency was calculated by comparing the transformation efficiency of the AP-site-containing construct with that of the control construct.

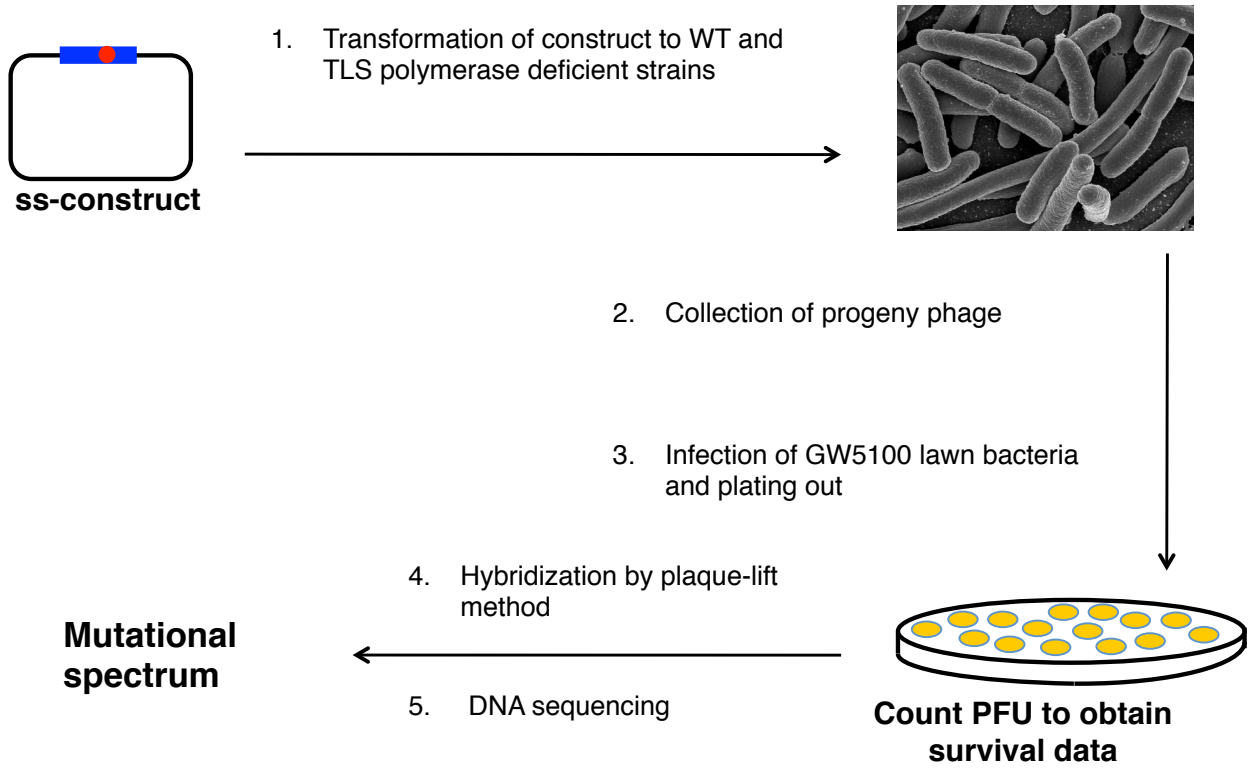


Figure 11: Schematic representation of TLS assay in *E. coli* using single-stranded M13 constructs

3.2.8 Plaque lift hybridization for mutational screening

Plaque lift hybridization method was performed to distinguish possible mutants at the lesion site and nearby location from the plaques containing unmodified vector. Prior to the transfer of plaques to nylon membrane, the plaques containing plates were chilled at 4°C at least for 1h. Then a labeled sterile Magnagraph nylon transfer membrane (Osmonics, Inc.), marked with two asymmetry cuts, was carefully laid on the surface of the plaque-containing plate, with out introducing any air bubbles in-between the surface and the membrane. The two cuts in the filter paper were marked on the outer surface of the LB plate, which were later used to align the plate with the autoradiograph. After 30 seconds, the membrane was carefully lifted from the plate and placed plaque-side up on a paper towel. Three replica filters were made for each plate. The filters were then placed (plaque-side up) sequentially on Whatman filter papers that were saturated with the following solutions: (1) denaturing buffer (0.5 M NaOH,1.5 M NaCI), 30 sec; (2) neutralizing solution (0.5 M Tris-HCI, pH=8, 1.5 M NaCI), 5 min; and (3) wash solution (2x SSC), 12 min. The membranes were allowed to air-dry for 30 minutes, and the liberated M13 DNAs were cross-linked to the membrane with UV light (12000 x 100 $\mu J/cm^2$) using a UV-crosslinker (FB-UVXL- 1000).

The processed membranes were placed (without overlapping) inside 12-inch hybridization roller bottles (Robbins Scientific, Sunnyvale, CA), 6 filters each. To each bottle, 20 mL of pre-hybridization solution (6x SSC, 5x Denhardt's) was added and incubated for 3 h at the hybridization temperature (T_{Hyb}, 51°C) in a rotary hybridization oven (Robbinson scientific). After 3 hours, prehybridization solution was replaced with 20 ml PerfectHybe™ hybridization buffer (Sigma) containing corresponding ^{32}P radiolabeled probe (50 pmol). Hybridization was carried out at T_{Hyb} for 12-16 hours in the rotary oven. Unhybridized probes were removed by

washing filters three times with 20 ml of the wash solution (6x SSC, 0.05% Na pyrophosphate, 0.1% SDS) for 30 min, once at T_{Hyb}-10°C and twice at T_{Hyb}. The washed membranes were air-dried and analyzed by autoradiography. The oligonucleotide probes containing the complementary sequence were used for analysis. Two left and right probes were used to select phages containing the correct insert, and transformants that did not hybridize with both the left and right probes were omitted. Any transformants that hybridized with the 15-mer wild-type probe were subjected to DNA sequence analysis. Lesion bypass efficiency was calculated by comparing the transformation efficiency of the lesion-containing construct to that of the control, whereas mutation frequency (MF) was calculated on the basis of hybridization and sequence analysis.

Figure 12: Probes used for oligonucleotide hybridization

3.2.9 Single-stranded pMS2 vector containing site specifically incorporated abasic site/*S*-cdA/*S*-cdG/O^2-Me-dT or O^2-POB-dT lesion

3.2.10 Large scale preparation of single-stranded pMS2 (sspMS2) vector

Single-stranded phagemid pMS2 DNA was prepared from *E. coli* JM109 with the aid of the helper phage M13K07 (NEB, Beverly, MA) as reported by Moriya.[174] The pMS2 vector contains origins of replication for f1, ColE1, and SV40, which can be used for single stranded replication in *E. coli*, double stranded replication in *E. coli* and mammalian cells. It also contains neomycin-resistant (neo) and ampicillin-resistant (amp) genes, which can be used for the selection of the cells containing the vector in mammalian and bacterial cells, respectively. In addition, it has the SV40 early promoter, SV40 small tumor (T) antigen splice sites, SV40 early polyadenylylation sites, and a multiple cloning site with a hairpin loop, that allow digestion of the single-stranded vector with *Eco*RV

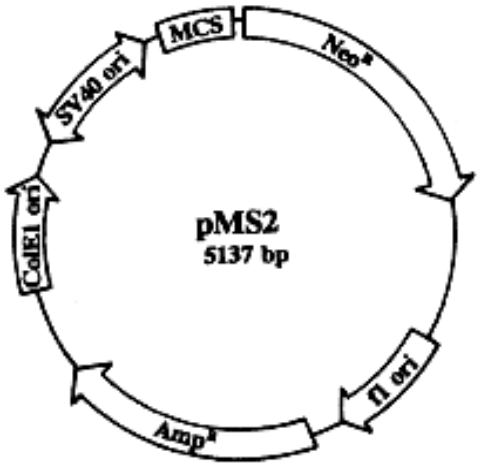


Figure 13: Structure of pMS2. MCS represent a multiple cloning site where hairpin structure is included

A brief description of large-scale preparation of pMS2 vector is described below. Electrocompetent *E. coli* DH12S (Invitrogen) cells were thawed on ice and 50 µL of cells were mixed with 50 ng of double stranded replicative form (RF) of pMS2. Then this mixture was transferred to a pre-chilled 1-mm gapped electroporation cuvette and the electroporation was carried out under following conditions: 1.8 kV, 25 µF and 200 Ω. Immediately after electroporation 1 mL of pre-warmed (37 °C) SOC medium (Invitrogen) was added, transferred to 14 mL culture tube and was grown at 37 °C in an orbital shaker at 225-250 rpm for 1 hour. From the resulting culture, 5.0 µL aliquot was plated on lx YT + ampicillin (100µg/ml) and incubated overnight at 37 °C. The following day, about 100 colonies were scraped from the 1x YT plate using a sterile loop, inoculated to 2 L of ampicillin containing 2xYT media along with 1 mL of M13K07 helper phage (Invitrogen) and incubated at 37 °C. After two hours of growth, 100µg/mL kanamycin was added and the culture was incubated overnight for another 16 hours.

The phage-infected culture was transferred to 250 mL centrifuge bottles, spun down at 10,000 rcf (20 min, 15°C), the resulting supernatant was transferred to another 250 mL centrifuge bottle and centrifuged at the same condition. Supernatant from the second spin was transferred 2L flask containing 4% (w/v) polyethylene glycol (PEG) and 3.5 M NaCl. The mixture was stirred in ice for at least one hour to precipitate the phage particles. The resulting cloudy solution was centrifuged at 10,000 rcf (20 min, 4 °C) and the pellet obtained was re-suspended in the minimum volume (6 mL) of 1x Tris-EDTA (pH 7.6) buffer. Then the phage particles were re-precipitated with ice-cold 4% (w/v) PEG and 4 M NaCl for an hour and centrifuged at 10,000 rcf (20 min, 4°C). The pallet obtained was re-suspended in 2.4mL of freshly prepared 1x Tris-EDTA (pH 7.6) buffer and centrifuged for 10 min at 4°C and 10000 rpm. The resulting supernatant containing the phage particles was incubated with 0.625 µg/mL

proteinase K (Invitrogen) for 16 hours and single-stranded pMS2 was purified from the digested phage solution by a series of phenol: chloroform extractions, followed by ethanol precipitation. Finally sspMS2 was re-suspended in 200 µL sterile deionized water and quantified on the UV spectrometer. After purification, the purity of the ss-pMS2 was determined by the absorbance ratio of 260 nm to 280 nm followed by analysis on 1.1 % agarose gel.

3.2.11 ss-pMS2 construct preparation

The construction of modified pMS2 vectors containing site-specifically incorporated adducts were performed following the method developed by Moriya (Figure 14). For construction of the pMS2 with modified or unmodified 12mer insert, l00 µg (58 pmols) of sspMS2 DNA was digested with 800 units of *Eco*RV at 37 °C for 3 h followed by room temperature overnight (16h). An aliquot of the reaction mixture was run on an l% agarose gel to confirm the complete linearization of the circular vector. An equimolar ratio of a 58-mer scaffold containing deoxyuridinies in place of deoxythymidines was annealed to linear ssDNA in 50 mM NaCl by heating at 90 °C for 5 min followed by slow cooling to 9 °C over 16 h. Annealing of the scaffold generated a gapped duplex that contained a 12mer-gap with complementary sequence of the lesion to be incorporated. An aliquot of the gapped duplex was run on 1% agarose gel with a known amount of control plasmid (ss-pMS2). The amount of the gapped duplex was determined by comparing the net-intensity of the gapped duplex band to the net-intensity of the known control plasmid using the KODAK Digital Science™-1D program. After determining the gapped duplex concentration, a 20-fold molar excess of the modified or unmodified 5'-phosphorylated 12mer insert was added and ligated at 16°C overnight with 2800 units of T4 DNA ligase. Unligated oligonucleotides were removed by being passed through Centricon 100 apparatus. Any residual uncut-circular or religated DNA was removed by additional round of *Eco*RV digestion.

The scaffold was removed by enzymatic digestion with uracil DNA glycosylase (UDG, 05 units/µl) and exonuclease-III (0.5 units/µL) at 16°C overnight. The resulting constructs were extracted with phenol: chloroform and precipitated with ethanol. The final constructs were resuspended in deionized distilled-water and were run on 1% agarose gel. The amount of construct was quantified in a manner similar to the way gapped duplex was quantified.

3.2.12 Transformation of site specifically incorporated lesion (*S*-cdA) containing ss-pMS2 vector in *E. coli* and progeny analysis:

Competent *E. coli* cells were prepared as described earlier. To induce SOS, the cells were re-suspended in 10 mM $MgSO_4$ solution and aliquots were treated with UV light (254 nm) (20 J/m2) in plastic Petri dishes. The combined aliquots were incubated in Luria broth at 37 ^{o}C for 40 min for maximum SOS induction. Following SOS induction, these cells were processed in a similar manner, except the procedure was carried out in subdued light. Then transformation of cells was carried out in a Bio-Rad Gene-Pulser apparatus by mixing vector DNA and un-induced cells or SOS-induced cells. Cells were grown for 1 h at 37 ^{o}C after adding SOC medium and then plated on a 1X YT plates containing Ampicillin and the transformants were analyzed by dot blot hybridization.

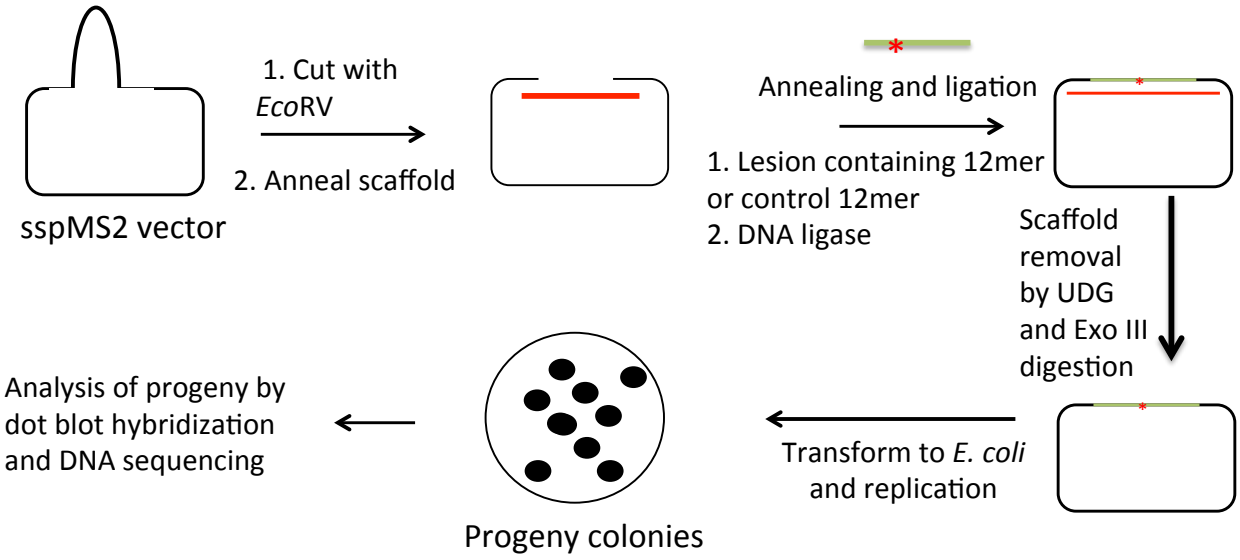
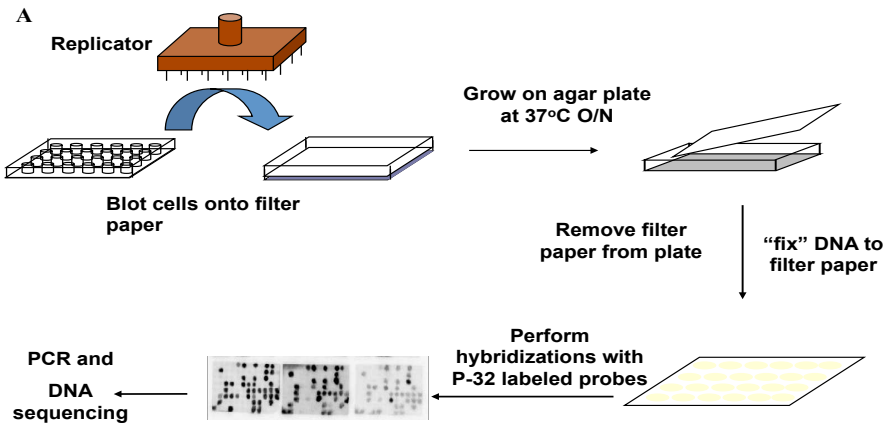


Figure 14: General protocol for making pMS2 construct followed by replication and analysis in *E. coli*

3.2.13 Mutant screening by dot-blot hybridization

Dot- blot hybridization was performed to identify the possible mutations at the lesion-site or at nearby locations in the progeny. In brief, appropriate amounts of liquid culture were plated on lx YT + ampicillin (100µg/mL) plates and incubated at 37 °C overnight (16-18 h). The resulting individual colonies were picked using sterile toothpicks into single wells of a 96-well plate (Fisher Scientific) containing 200 µl lx YT + ampicillin (100 µg/mL). The inoculated 96-well plate was then incubated at 37°C for a minimum of 3 hrs before blotting colonies onto filter paper. Using a 96-well replicator (Boekel Scientific, Feasterville, PA), cultures in the 96-well plate were blotted onto sterile labeled Whatmann chromatography papers placed on lx YT + ampicillin (100 µg/mL) plates, in triplicates. The plates were incubated at 37 °C for 16 h for colonies to be replication on the filter paper. After overnight incubation, following procedure was carried out to fix DNA onto the filter paper. The filters were carefully lifted from the plates and subsequently transferred to (colony side up) 0.5 N NaOH, denaturing solution, for 12 mins, and to 0.5 M Tris-HCl (pH 7.5), neutralizing solution for 7 mins. Then the filters were washed twice with lx SSC for 4 mins with slow agitation (50 rpm) and once with 100% ethanol for 30 seconds before they were baked for 2 hours at 80 °C. In-between each washing steps, except for two SSC washes, filters were drained and damped dry on paper towels (colony side up). Thereafter, each filter was placed separately in a heat sealable bag and sealed with 6 mL of PerfectHybe™ hybridization buffer (Sigma) containing corresponding left, right or wild-type ^{32}P-radiolabeled probe (10 pmol of probe/filter). Liquid in the sealed bags were spread evenly to entirely cover the filters and were allowed to hybridize to corresponding probe at its optimum hybridization temperature for a minimum of 4 hours in the hybridization oven (Problot™ 12). Then the filters were carefully removed from the bags and gently agitated (60 rpm) twice with 2x

SSC+0.1% SDS, preheated to the hybridization temperature (T_{Hyb}), for 10 mins. Finally the filters were removed from the SSC solution, dried and exposed to autoradiography. Two left and right probes were used to select phages containing the correct insert, and transformants that did not hybridize with both the left and right probes were omitted. Any transformants that hybridized with the left and right probes but failed to hybridize with the wild-type probes were subjected to DNA sequence analysis.



B

Template strand

5'-…CTCGGTACCAGCGATGTGCGTGTTTGTATCGCTTGCAGGGG-…3'

Complementary strand

3'-…GAGCCATGGTCGCTACACGCACAAACATAGCGAACGTCCCC-…5'

5'-GTACCAGCGATGTG-3' 5'-TTGTATCGCT TGCAG-3'

Left probe (14-mer) Right probe (15-mer)

5'-ATGTGCGTGTTTGTAT-3'

Wild type probe (16-mer)

C

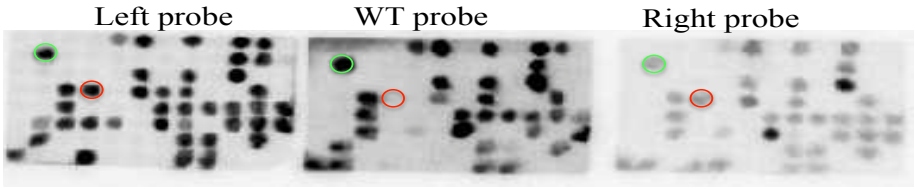


Figure 15: Mutational analysis by dod blot hybridizaton. A. The general scheme for dot blot hybridization; B. Sequences used for mutant analysis; C. A representation of hybridization results used for selection mutants. Green circles shows insert without mutation while red circles shows insert with mutation.

3.2.14 Replication and analysis of site specifically incorporated abasic site/ *S*-cdA/ *S*-cdG/ O^2-Me-dT or O^2-POB-dT containing ss-pMS2 in human embryonic kidney 293T cells and siRNA induced TLS knockdown cells

3.2.15 Replication and analysis of site specifically incorporated lesion containing ss-pMS2 in HEK293T cells

The HEK293T cells were maintained in Dulbecco's modified Eagle's medium supplemented with 4 mM L-glutamine, and adjusted to contain 1.5 g/L sodium bicarbonate, 4.5 g/L glucose, and 10% fetal bovine serum in. When it was about ~80% confluent, cells were washed with PBS (Invitrogen), harvested with 6 mL of Accutase and was seeded in 24-well plate (approximately 1 x 10^5 cells). The cells were grown to ~80% confluency in the 24 well plate and transfected with 100 ng of each construct using Lipofectamine cationic lipid reagent (Invitrogen) according to manufacture's instruction. Subsequent to transfection with control or lesion containing pMS2, the cells were allowed to grow at 37^{o}C, 5% CO_2 for 24 hours and the plasmid DNA was collected and purified by the method of Hirt[175] as described below. It was used to transform *E. coli* DH10B and the colonies were analyzed by dot blot hybridization and DNA sequencing.

3.2.16 Isolation of progeny pMS2 from HEK293T and its amplification

The progeny plasmid-DNA was harvested by a method described by Hirt.[175] Briefly, the HEK293T/17 cells were centrifuged at 14,000 rpm for 30 sec at room temperature. The supernatant was discarded, and pellet was gently re-suspended in 100µL of ice cold buffer P1 (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 100 µg/mL RNase A). The cells were then lysed by adding 100µL of' buffer P2 (200 mM NaOH, 1% SDS), gently inverted a few times and

incubated for 2 min at room temperature. The reaction was neutralized by the addition of ice-cold 120 µl of P3 buffer (3.0 M potassium acetate, pH 5.5) and incubated on ice for 10 min. The cell lysate was centrifuged at 14000 rpm (10 min) at room temperature for 2 min and the resulting supernatant containing the double stranded progeny DNA was transferred to a new tube. The DNA was extracted with phenol/chloroform (3X) and precipitated by ethanol precipitation. Final DNA pallet was air-dried, resuspended in sterile distilled water, and transformed to DH10B cells (Invitrogen) by electroporation. Immediately after transformation, 1 mL of SOC media was added, grown for 1h and plated on 1x YT + ampicillin (100 µg/mL) plates. The resulting colonies were analyzed by dot blot hybridization and DNA sequencing (discussed in previous sections).

3.2.17 siRNA induced knockdown of TLS polymerases in HEK293T/17 cells and mutational analysis of TLS products

Prior to transfection of the control or the lesion containing vectors, synthetic siRNA duplexes were transfected into HEK293T cells using Lipofectamine. HEK293T cells were plated in 6-well plates at 50% confluence. After 24 h incubation, they were transfected with 100 pmoles of siRNA duplex mixed with Lipofectamine, diluted in Opti-MEM (Gibco), per well. One day before transfection of the plasmid, cells (1×10^5) were seeded in 24-well plates at 70% confluence. Cells were then co-transfected with another aliquot of siRNA and either control plasmid or lesion containing plasmid. After 24 h – 36 h incubation period, progeny plasmids were isolated as described earlier.

3.2.18 TLS assay in HEK293T cells and siRNA induces TLS knockdown cells

The lesion containing or the control pMS2 construct was mixed at a 1:2 ratio with a single-stranded pMS2 DNA containing a similar DNA sequence 5'-GTCCGTGTTTGT-3' in the place of the 12mer insert, which was used as an internal control to determine the TLS events. This was instead of 5'-TGCAGTGTCAGC-3' in the region where abasic site was placed and instead of 5'-GTGCGTGTTTGT-3' in the region where O^2-Me-dT or O^2-POB-dT was placed. The mixed DNA was used to transfect either HEK293T cells or TLS knockdown cells and processed as described earlier. Oligonucleotide probes complementary to the wild type and the mutant plasmid were used to analyze the progeny. The mutant DNA (internal control) gave approximately equal number of progeny as the control construct.

3.2.19 Total RNA isolation with RNeasy Plus Mini Kit:

Total RNA was successfully isolated from mammalian cells using the RNeasy Plus Mini Kit from Qiagen, according to manufacture's instructions. A brief description of the protocol is as follows. Seventy-two hours after the first transfection of siRNA duplexes, 350 μL of RLT Plus buffer containing β-mercaptoethanol was added to each well of the 24 well plate to lyse cells, and the lysate was homogenized by vortexing for 1 min. Then the lysate was passed through a gDNA Eliminator spin column by centrifugation, and equal volume of ethanol was added to the flow-through. After that the sample was applied to an RNeasy spin column, centrifuged and flow through was discarded. Subsequently, column was washed with provided washing buffers. Finally, RNeasy spin column was transferred to a sterile RNase free column and RNA was eluted in 50μl of RNase free water.

3.2.20 Quantification of RNA

RNA yield was quantified by spectrophotometric analysis using the convention that 1 absorbance unit at 260nm equals 40μg/ml RNA. Sample purity and concentration were determined by the absorbance measurements at 260 and 280nm. The A260/A280 ratio should be close to 2.0 for pure RNA (ratios between 1.9 and 2.1 are acceptable). UV absorbance measurements were obtained using Nanodrop2000c (Thermo Scientific).

3.2.21 Reverse Transcription Polymerase Chain Reaction (RT- PCR) Analysis.

Total RNA was extracted from the cells 72 h after the first transfection of siRNA duplexes, using the All Prep DNA/RNA/Protein Kit (Qiagen). RT-PCR analysis was performed with 100 ng of total RNA using One Step RT-PCR Kit (Qiagen) according to the manufacturer's instructions. Primer sequences used for RT-PCR are listed in Table 2. Using primers specific to TLS DNA polymerases and control gene GAPDH, the siRNA knockdown efficiency was determined as previously described. Reverse transcription and the PCR initial activation step were performed for 30 min at 50 °C and 15 min at 95 °C, respectively. For PolH, PolK, PolI, and Rev1, amplification was conducted at 94°C for 30s, 55°C for 45s, and 72°C for 60s for 26 cycles and Rev3 was amplified for 32 cycles. Amplification of GAPDH was conducted at 94 °C for 30 s, 55°C for 45s, and 72°C for 45s for 24 cycles. RT-PCR products were analyzed on a 2% agarose gel run at 100 V for 3 h in 1× TBE buffer.

Table 2: Primer sequences used for RT-PCR for quantification of the siRNA induced TLS knockdown

Gene	Primer	
POLH	Forward:	5'-CTACTCGGGAACAGGTACAATG-3'
	Reverse:	5'-GTGGGAGCAGTAAGAGATGAAA-3'
POLK	Forward:	5'-AGCTGTGAGTAAAGAGGTTAAGG-3'
	Reverse:	5'-TGAACCTAGACCCAAGGAGATA-3'
POLI	Forward:	5'-ATTAGGGACAGGAAATTATG-3'
	Reverse:	5'-CAATGTCAGAAGGGAAAAG-3'
REV3	Forward:	5'-GAGAGAAGCACAAGTCCCATAAA-3'
	Reverse:	5'-CTAATGGGCTCTGCTGTAGATC-3'
REV1	Forward:	5'-AATCTACCAGGAGTTGGA-3'
	Reverse:	5'-GTGGAAGGGTTCAGATTAG-3'
GAPDH	Forward:	5'-ACCACAGTCCATGCCATCAC-3'
	Reverse:	5'-TCCACCACCCTGTTGCTGTA-3'

3.2.22 Quantification of knockdown by Real-time Quantitative RT-PCR

Using the RNeasy Mini Kit (Qiagen), total RNA was extracted from the cells 72h after transfection with siRNA. Then the cDNA synthesis and RT-PCR were performed on a Bio-Rad CFX-96 Real-Time PCR Detection System using Quantitect SYBR Green RT-PCR Kit (Qiagen) according to manufacturer's recommended procedure: one cycle for reverse transcription, 30min (50^{o}C); one cycle for PCR initial activation step, 15 min (95^{o}C); 40 cycles for denaturation, annealing and extension, 15 sec (94^{o}C), 30 sec (55^{o}C), 30 sec (72^{o}C) respectively. Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) was used as an internal control for analysis. Relative quantification of gene expression was performed using the comparative cycle threshold (C_T) method ($2^{-\Delta\Delta C_T}$).

4 Chapter 1† : Replicative Bypass of Abasic Site in *Escherichia coli* and Human Cells: Similarities and Differences

†The text and figures in Chapter 1, in part or in full, are a reprint of the material as it appears in Weerasooriya, S.; Jasti, V. P.; Basu, A. K., Replicative Bypass of Abasic Site in *Escherichia coli* and Human Cells: Similarities and Differences. *PLoS ONE* **2014,** 9, (9), e107915.

4.1 Abstract

Abasic [apurinic/apyrimidinic (AP)] sites are the most common DNA damages, opposite which dAMP is frequently inserted ('A-rule') in *Escherichia coli*. Nucleotide insertion opposite the AP-site in eukaryotic cells depends on the assay system and the type of cells. Accordingly, a 'C-rule', 'A-rule', or the lack of specificity has been reported. DNA sequence context also modulates nucleotide insertion opposite AP-site. Herein, we have compared replication of tetrahydrofuran (Z), a stable analog of AP-site, in *E. coli* and human embryonic kidney 293T cells in two different sequences. The efficiency of translesion synthesis or viability of the AP-site construct in *E. coli* was less than 1%, but it was 7- to 8-fold higher in the GZGTC sequence than in the GTGZC sequence. The difference in viability increased even more in pol V-deficient strains. Targeted one-base deletions occurred in 63% frequency in the GZG and 68% frequency in GZC sequence, which dropped to 49% and 21%, respectively, upon induction of SOS. The full-length products with SOS primarily involved dAMP insertion opposite the AP-site, which occurred in 49% and 71% frequency, respectively, in the GZG and GZC sequence. dAMP insertion, largely carried out by pol V, was more efficient when the AP-site was a stronger replication block. In contrast to these results in *E. coli*, viability was 2 to 3 orders of magnitude higher in human cells, and the 'A-rule' was more rigidly followed. The AP-site in the GZG and GZC sequences gave 76% and 89%, respectively, Z→T substitutions. In human cells, targeted one-base deletion was undetectable, and dTMP > dCMP were the next preferred nucleotides inserted opposite Z. siRNA knockdown of Rev1 or pol ζ established that both these polymerases are vital for AP-site bypass, as demonstrated by 36-67% reduction in bypass efficiency. However, neither polymerase was indispensable, suggesting roles of additional DNA polymerases in AP-site bypass in human cells.

4.2 Introduction

AP-sites are non-coding and they strongly block DNA replication by replicative DNA polymerases (pols). A specialized class of pols, known as the translesion synthesis (TLS) enzymes, however, can bypass DNA damages, including AP-sites. [176] But AP-sites lack the information necessary to identify the correct base. So, when a nucleotide is incorporated opposite AP-site, it frequently results in a mutation. [177, 178]

In *Escherichia coli* cells, AP-site bypass is largely SOS-dependent [177] and dAMP is most commonly inserted opposite it, which led to the so-called 'A-rule'. [179] Site-specific studies in *E. coli* showed that the preferential dAMP insertion occurs with SOS. [180] In *Saccharomyces cerevisiae*, dCMP is inserted ('C-rule) when the AP-site is located in the single-stranded gap of a gapped duplex plasmid [181], but in duplex DNA dAMP is inserted preferentially opposite the AP-site in both the leading and lagging strand. [182] In simian kidney cells, some studies suggest a lack of specificity in nucleotide insertion opposite AP-sites. [183-185] In human cells, by contrast, in the single-stranded gap region of a gapped duplex plasmid, the majority of the bypass involved insertion of dAMP opposite AP-sites. [186] The natural AP-site remains in ring-chain equilibrium of the cyclic hemiacetal with an open chain aldehyde form, which is highly labile to heat and alkaline pH. [187] Because of its instability, a tetrahydrofuran (Z), as a stable model of the cyclic form of abasic site, has been used in many studies. A comparison of DNA replication past the natural AP-site and Z, the stable analog, in *E. coli* showed that the biological effects of these two lesions are similar in SOS-induced *E. coli*, but their replication characteristics are distinct in uninduced cells.[188]

The discovery of the Y-family polymerases [189, 190] that can efficiently bypass DNA lesions generated interest in investigating the kinetics of TLS of these enzymes on AP-site

templates. Pol V (UmuD'$_2$C complex) can bypass AP-site efficiently *in vitro* incorporating dAMP preferentially opposite it, whereas pol III and pol IV failed to bypass it on the same time scale.[191] But pol IV incorporates a nucleotide opposite the AP-site and so another polymerase may carry out the extension step.[191] Both yeast and human pol η can efficiently bypass AP-sites and both prefer to insert purine nucleotides.[182, 192] The yeast enzyme, in addition, causes high frequency of -1 and -2 frameshifts.[193] Dpo4 and human pol κ also bypass AP-sites, although with significantly reduced efficiency.[194, 195] For human pol κ, however, accessory proteins such as PCNA, RFC, and RPA increase efficiency by more than an order of magnitude.[195] Human pol ι also bypasses AP-sites inserting either dGTP or dTTP with 10-fold reduced efficiency.[150] Pol δ and REV1, in the presence of pol ζ, can bypass AP-sites.[196, 197] It was suggested that pol δ inserts dAMP preferentially opposite the AP-site, and pol ζ subsequently extends from the inserted nucleotide [196], but arguments against this model have been presented.[197]

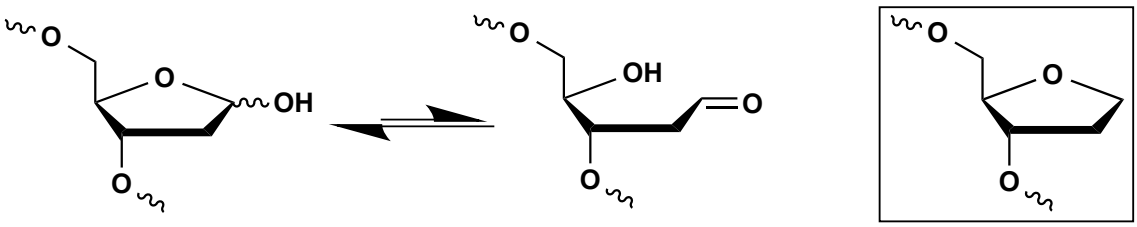
Despite the wealth of data from the *in vitro* experiments described above, the mechanism of AP-site bypass in a cell remains unclear. Additional complications arise from the fact that nucleotide insertion opposite the AP-site is greatly dependent on the DNA sequence context and the type of cells. Consequently, evaluation of data from different laboratories using different sequence contexts and different types of cells could be challenging.

Oligonucleotide Inserts

5’ – TGCAGZGTCAGC **(GZG)**

5’ – TGCAGTGZCAGC **(GZC)**

5’ – TGCAGTGTCAGC **(Ctrl)**



Naturally occurring basic site

Tetrahydrofuran analog of abasic site (Z)

Figure 16: Oligonucleotides and tetrahydrofuran analoge of abasic site used in the study

4.3 Results

4.3.1 PAGE analysis of the purified lesion containing oligonucleotide

The gel purified 12-mer oligodeoxynucleotides containing site-specifically incorporated tetrahydrofuran, adduct in the sequence 5'-TGCAGZGTCAGC-3', and 5'-TGCAGTGZCAGC-3', were analyzed in denaturing 20% polyacrylamide gel electrophoresis (PAGE) containing 8M urea. The unmodified oligodeoxynucleotide containing same sequence, 5'-TGCAGTGTCAGC-3' was used as a control. The 12-mer samples were radiolabeled at 5'-end with [γ-^{32}P] ATP using T4 PNK enzyme and unlabeled free [γ-^{32}P] ATP were removed using Sephadex G-10 columns. A typical autoradiograph result of these adducts is shown below in the Figure 17.

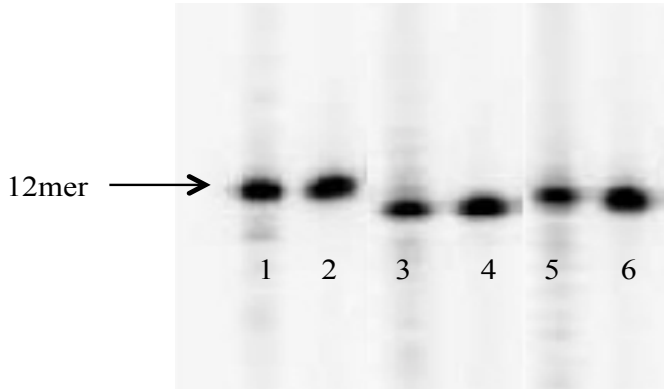


Figure 17: A representative 20% denaturing-PAGE [γ-32P] ATP labeled oligonucleotide samples containing tetrahydrofuran along with the unmodified 12-mer control. Key: 1. Unmodified 12mer-control, 2. Purified 12mer control, 3. Crude oligonucleotide containing GZG sequence, 4. PAGE-purified oligonucleotide containing GZG sequence, 5. Crude oligomer containing GZC sequence, 6. PAGE purified oligomer containing GZC sequence

4.3.2 Construction of the AP-site containing vector and its replication

We used the M13mp7L2 vector [198] to investigate TLS of Z in *E. coli* cells, whereas the same in HEK 293T was carried out using the pMS2 shuttle vector plasmid containing the origins of replication for f1, ColE1, and SV40, and the genes for neo and amp resistance.[199, 200] Both single-stranded vectors contain a hairpin region that, upon digestion with the appropriate restriction enzyme followed by scaffolding with an oligonucleotide, generates the desired gapped DNA. The Z containing 12-mer is ligated to this gap and the scaffold is removed before replication. Each step of the construct preparation process was monitored on 1% agarose gel. Finally the concentration of the each construct was determined by comparing the net intensity of the circular band to the net intensity of a known amount of standard DNA. A representative picture of an agarose gel is shown in Figure 18.

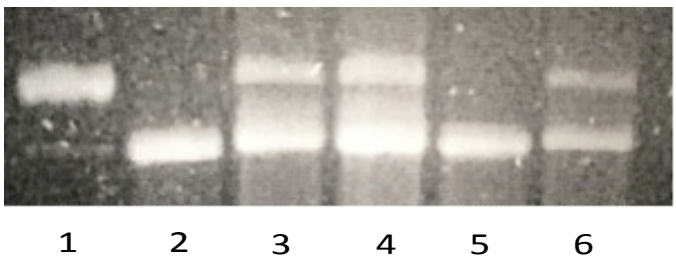


Figure 18: A representative agarose gel analysis of the M13 constructs. Lane 1 shows ss-M13 DNA where as lane 2 show same after digestion with EcoRI. Lanes 3 and 4 represent circular constructs. Lanes 5 and 6 represent the construct after and before scaffold removal, respectively, of a mock ligation mixture

4.3.3 TLS of Z in *E. coli*

Magnitude of TLS or viability was determined by comparing the transformation efficiencies of Z-containing and control constructs. As shown in Figure 19 and Table 3, Z is a highly toxic lesion, and in the absence of SOS-induction, viability was only 0.04% and 0.3%, respectively, in the GZC and GZG sequence. With SOS, viability increased by 50-75%. The increase in viability with SOS is lower than what others have reported with either natural abasic site or Z.[180, 188, 193] Since SOS induction was considered optimal by parallel replication of other lesions [201, 202], we suspect that the sequence context may have played a role in this relatively modest enhancement in viability. What we found noteworthy, however, is that the viability of the GZC construct was 7- to 8-fold lower than that of the GZG sequence, in both uninduced and SOS-induced *E. coli* (Figure 19 and Table 3), which suggests that the GZC sequence poses a stronger replication block than the GZG sequence. Viability followed a similar trend in pol II- or pol IV-deficient cells, although it was increased nearly 3-fold with SOS in the GZG sequence. For the GZC sequence, viability was 3- to 4-fold lower in the pol V-deficient strain, whereas it was 8- to 14-fold lower in the triple knockout (TKO) strain that lacks pol II, pol IV, and pol V (Figure 19) But such pronounced decrease in viability was not detected in the GZG sequence, so that the viability in GZC sequence relative to GZG sequence was reduced by at least 15-fold and 48-fold, respectively, in pol V-deficient and TKO cells lines (Figure 19 and Table 3). These results suggest that pol V plays a particularly critical role in bypassing Z in the stronger replication blocking GZC sequence compared to the GZG sequence and that in the absence of pol V, the other Y-family pols such as pol II and pol IV also have a role in TLS of Z in this sequence context.

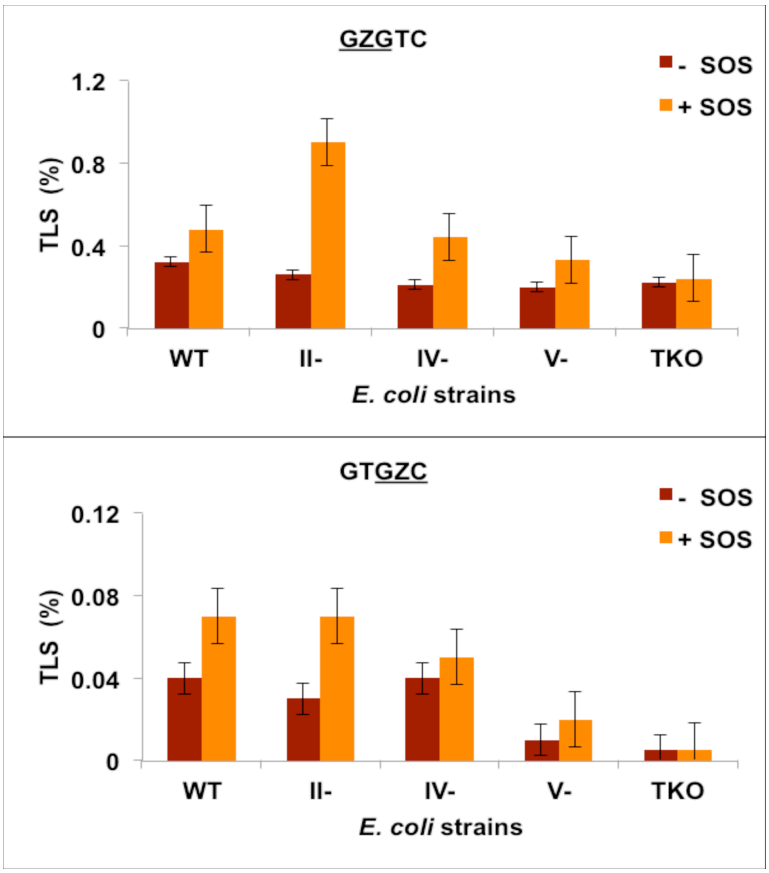


Figure 19:TLS frequencies for GZGTC and GTGZC constructs in wild type and pol II-, pol IV-, pol V-, and triple-knockout *E. coli* strains without and with SOS

Table 3: Viability of abasic site in E. coli cells

Strain	SOS	GZGTC		GTGZC	
WT	-	0.3	± 0.2	0.04	± 0.01
	+	0.48	± 0.06	0.07	± 0.002
pol II-	-	0.26	± 0.03	0.03	± 0.004
	+	0.9	± 0.1	0.07	± 0.04
pol IV-	-	0.21	± 0.06	0.04	± 0.01
	+	0.44	± 0.01	0.05	± 0.01
pol V-	-	0.2	± 0.02	0.01	± 0.01
	+	0.29	± 0.09	0.02	± 0.009
TKO	-	0.22	± 0.01	0.005	± 0.002
	+	0.24	± 0.07	0.005	± 0.006

4.3.4 Mutations resulting from bypass of Z in *E. coli*

Analyses of the progeny plaques were carried out by oligonucleotide hybridization followed by DNA sequencing [201]. In the absence of SOS, the major type of mutation in both sequences was targeted one-base deletions (63% in GZG and 68% in GZC) (Figure 20 and Table 4). Of the rest, most were Z→T (dAMP insertion), which occurred in 35% and 24% frequency, respectively, in GZG and GZC sequences. In the GZC sequence, multiple deletions were detected in 6% progeny, but they were undetectable in GZG sequence. dCMP, dTMP, and dGMP insertions were rare and occurred in <1% frequency. In the GZC sequence, Z→T increased from 24% in uninduced cells to 71% in SOS-induced cells, and one-base deletions dropped to 21%. In the GZG sequence, however, Z→T events increased to 49%, and one-base deletions also occurred at approximately 49% frequency. SOS had relatively minor effect on the other types of mutations. It is worth noting that significant differences in the TLS efficiency and nucleotide incorporation pattern opposite the AP-site depending on the DNA sequence context were reported by others as well [180, 188], and as with these studies, we cannot provide a rationale for the

observed differences in TLS between the GZG and GZC sites. Since 92% and 98%, respectively, of the progeny in the GZC and GZG construct gave Z→T mutations and one-base deletions, for the polymerase-knockout strains, we only focused on these two types of mutations. In Figure 21, we report percent MF normalized by TLS to display both the relative proportion of the two main mutations combined with the bypass efficiency (whereas the % MF only is shown in Figure 22). As shown in Figure 21, Table 5 and Table 6, in pol II- or pol IV-deficient cells, there were only minor changes in the relative proportion of the Z→T mutations or one-base deletions in GZG sequence. However, there were significant changes in the relative proportion of these two events in the GZC sequence. Without SOS, Z→T mutations increased from 24% in the wild type strain to 32% and 39%, respectively, in the pol II- and pol IV-deficient strains. With SOS, Z→T mutations increased from 71% in the wild type strain to 83% and 76%, respectively, in the pol II- and pol IV-deficient strains. In uninduced pol V-deficient strain also (Figure 21 and Table 7), Z→T mutations occurred at a higher frequency (38% in both sequences) than in the wild type cells (24% and 35%, respectively, in GZC and GZG sequence). In contrast, in both sequence contexts, the relative proportion of Z→T mutations dropped with SOS in pol V-deficient and TKO cells, except for GZG in TKO strain, where it remained approximately the same (Figure 21, Table 7 and Table 8). These results indicate the major function of pol V is to insert dAMP opposite AP-sites, particularly with SOS when this pol is present in substantial concentration. But our results also suggest that in addition to the TLS pols, replicative pols are able to bypass AP-sites. Furthermore, substantial dAMP insertion occurred in the absence of pol V as well as other TLS pols, which suggests that replicative pols are able to insert dAMP across AP-sites. While the results of this study show pol V's primary role is to insert dAMP opposite AP-sites, it does not point to a specific pol for the one-base deletions.

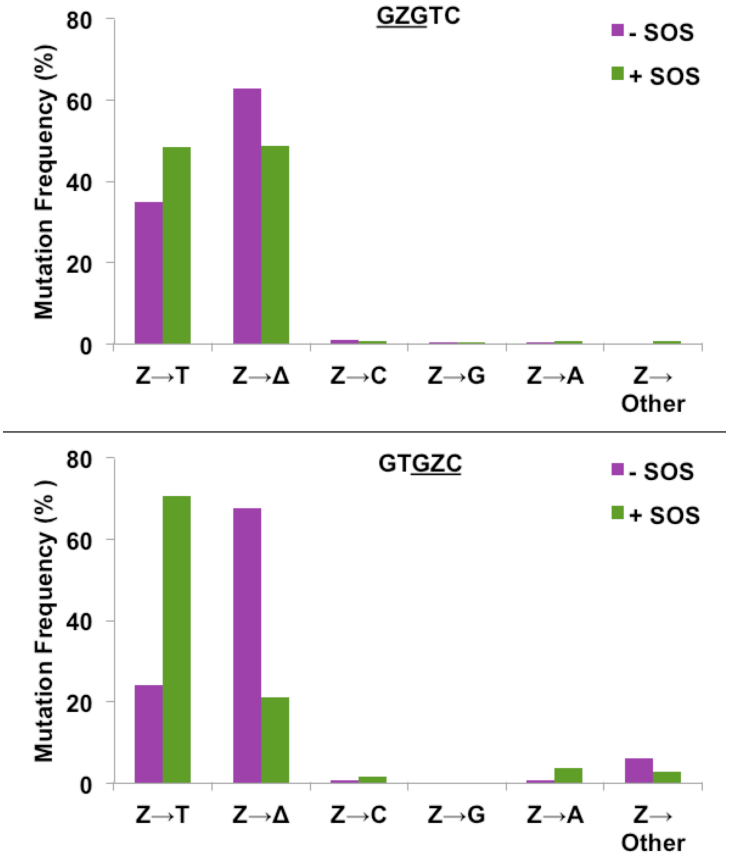


Figure 20: Mutations induced by Z in GZGTC and GTGZC sequence contexts in *E. coli* with or without SOS

Table 4: Mutational frequency in wild type *E. coli* cells[†]

Strain	Lesion	SOS	Trial	Total plaques screened	Z →T (%)	Z →Δ (%)	Z →C (%)	Z →G (%)	Z→A (%)	Z→ Other (%)
WT	GZGTC	-	1	85	31[a] (37)	52 (61)	2 (2)	0 (0.0)	0 (0.0)	0 (0.0)
			2	101	34[b] (34)	65 (64)	0 (0.0)	1 (1)	1[c] (1)	0 (0.0)
			Total	**186**	**65 (35)**	**117 (63)**	**2 (1)**	**1 (0.5)**	**1 (0.5)**	**0 (0.0)**
		+	1	156	74 (47)	77 (49)	1 (1)	1[d] (1)	1 (1)	2[e] (1)
			2	124	62 (50)	60 (48)	1 (1)	0 (0.0)	1 (1)	0 (0.0)
			Total	**280**	**136 (49)**	**133 (49)**	**2 (0.7)**	**1 (0.4)**	**2 (0.7)**	**2 (0.7)**
	GTGZC	-	1	54	11[f] (20)	41[g] (76)	0 (0.0)	0 (0.0)	0 (0.0)	2[h] (4)
			2	57	16 (28)	34[i] (60)	1 (2)	0 (0.0)	1[j] (2)	5[h,k,l] (9)
			Total	**107**	**27 (24)**	**74 (68)**	**1 (1)**	**0 (0.0)**	**1 (1)**	**7 (6)**
		+	1	71	51[m] (72)	13 (19)	2 (3)	0 (0.0)	3[j,n] (4)	2[k,o] (3)
			2	38	26 (68)	10 (26)	0 (0.0)	0 (0.0)	1[j] (3)	1[p] (3)
			Total	**109**	**77 (71)**	**23 (21)**	**2 (2)**	**0 (0.0)**	**4 (4)**	**3 (3)**

[†]The superscript indicates one or more mutants containing mutation elsewhere shown below and the number in parenthesis shows the number of events detected.

(a) TGC AGT TTC CGT (1), (b) CGC AGT GTC CGT (1), (c) TGC AGA ATC CGT (1), (d) TAG AAG GTC ACG (1), (e) TG _____ AGC, _____ TG AGC (1),(f) TGC AG _ GTC AGC (1), TGC A _ _ GTC AGC (1), TGC A _ _ CTC AGC (1),(g) TGC AG _ G _ C AGC (1), T_ G AGT G_C AGC (1), (h)TGC AG _____ (2 in expt 1 and 2 in expt 2), (i) TGC AGT _ _C AGC, TGC AGT G_ C CGC (1),(j) TGC AGT _ AC AGC (2 in expt 1 and 1 in expt 2), (k) _____ ___ _GC (1), (l)TGC _____ AGC, TG____ _ _C AGC (1) , (m) TGC AGA GTC AGC (1), (n) _____ _ AC AGC (1),(o) TGC A _ _ _ _C AGC (1), (p) _____ _ _C AGC (1)

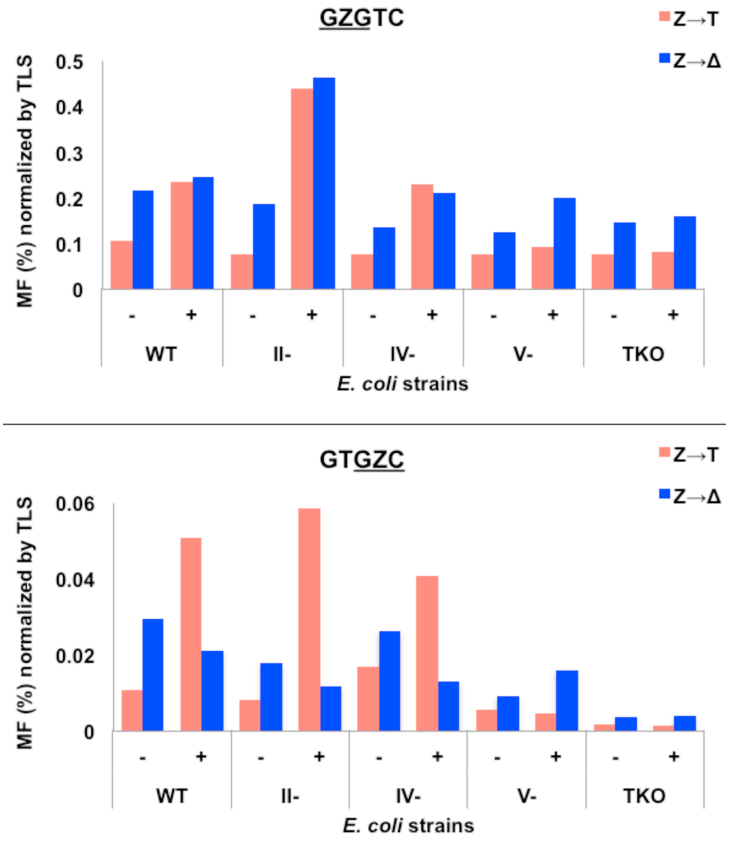


Figure 21: A comparison of the frequency of Z→T versus targeted Z deletion (i.e., Z→Δ) multiplied with % TLS for GZGTC and GTGZC constructs in wild type and pol II-, pol IV-, pol V-, and triple-knockout E. coli strains without and with SOS

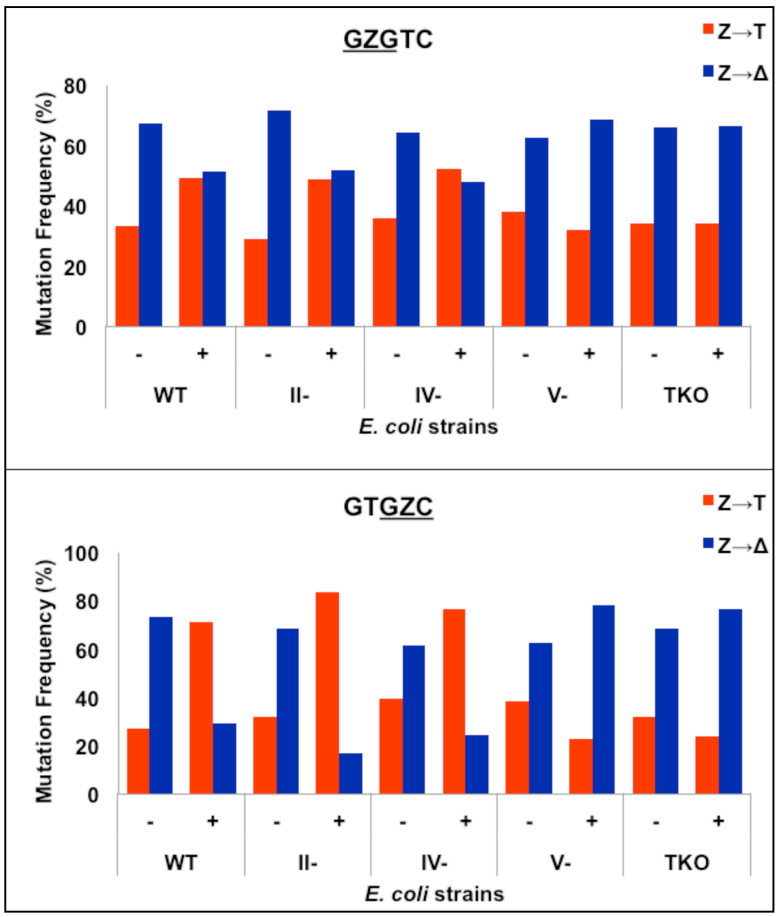


Figure 22: A comparison of the frequency of Z→T versus targeted Z deletion (i.e., Z→Δ) for GZGTC and GTGZC constructs in wild type and pol II-, pol IV-, pol V-, and triple-knockout E. coli strains without and with SOS

Table 5: Mutational frequency in pol II- deficient *E. coli* strain

Strain	Lesion	SOS	Trial	Total plaques signals	Z→T (%)	Z→Δ(%)
pol II-	**GZGTC**	-	1	86	27 (31)	59 (69)
			2	268	75 (28)	193 (72)
			Total	**354**	**102 (29)**	**252 (71)**
		+	1	211	99 (47)	112 (53)
			2	63	29 (46)	34 (54)
			3	117	62 (53)	55 (47)
			Total	**391**	**190 (49)**	**201 (51)**
	GTGZC	-	1	25	9 (36)	16 (64)
			2	13	3 (23)	10 (77)
			Total	**38**	**12 (32)**	**26 (68)**
		+	1	45	33 (73)	12 (27)
			2	85	78 (92)	7 (8)
			3	56	44 (79)	12 (21)
			Total	**186**	**155 (83)**	**31 (17)**

Table 6: Mutational frequency in pol IV- deficient *E. coli* strain

Strain	Lesion	SOS	Trial	Total plaques screened	Z→T (%)	Z→Δ (%)
pol IV-	**GZGTC**	-	1	115	39 (34)	76 (66)
			2	89	34 (38)	55 (62)
			Total	**204**	**73 (36)**	**131 (64)**
		+	1	175	87 (50)	88 (50)
			2	190	103 (54)	87 (46)
			Total	**365**	**190 (52)**	**175 (48)**
	GTGZC	-	1	30	12 (40)	18 (60)
			2	29	11 (38)	18 (62)
			Total	**59**	**23 (39)**	**36 (61)**
		+	1	80	60 (75)	20 (25)
			2	37	29 (78)	8 (22)
			Total	**117**	**89 (76)**	**28 (24)**

Table 7: Mutational frequency in pol V- deficient *E. coli* strain

Strain	Lesion	SOS	Trial	Total plaques screened	Z→T (%)	Z→Δ (%)
pol V-	**GZGTC**	-	1	77	29 (38)	48 (62)
			2	146	53 (36)	93 (64)
			3	225	87 (39)	138 (61)
			Total	**448**	**169 (38)**	**279 (62)**
		+	1	34	10 (29)	24 (71)
			2	236	78 (33)	158 (67)
			3	117	35 (30)	82 (70)
			Total	**387**	**123 (32)**	**264 (68)**
	GTGZC	-	1	18	5 (28)	13 (72)
			2	40	17 (43)	23 (58)
			Total	**58**	**22 (38)**	**36 (62)**
		+	1	20	4 (20)	16 (80)
			2	59	12 (20)	47 (80)
			3	87	21 (24)	66 (76)
			Total	**166**	**37 (22)**	**129 (78)**

Table 8: Mutational frequency in triple knockout *E. coli* strain

Strain	Lesion	SOS	Trial	Total Plaques screened	Z→T (%)	Z→Δ (%)
TKO	**GZGTC**	-	1	137	48 (35)	89 (65)
			2	178	59 (33)	119 (67)
			Total	**315**	**107 (34)**	**208 (66)**
		+	1	73	23 (32)	50 (68)
			2	92	33 (36)	59 (64)
			Total	**165**	**56 (34)**	**109 (66)**
	GTGZC	-	1	11	3 (27)	8 (73)
			2	58	19 (33)	39 (67)
			Total	**69**	**22 (32)**	**47 (68)**
		+	1	17	5 (29)	12 (71)
			2	21	4 (19)	17 (81)
			Total	**38**	**9 (24)**	**29 (76)**

4.3.5 TLS of AP-site in human cells.

To determine the magnitude of TLS in HEK 293T cells, we have mixed 2:1 ratio of the AP-site construct and unmodified plasmid that contained a different sequence at the Z-containing 12-mer region. The unmodified plasmid was used as an internal control. The percentages of the colonies originating from the lesion-containing construct relative to the unmodified plasmid, reflecting the percentage of TLS, were determined by oligonucleotide hybridization. In stark contrast to *E. coli* cells, which gave less than 1% progeny for the Z-containing construct even with SOS, the frequency of TLS in HEK 293T cells was 23% and 33%, respectively, for the GZC and GZG sequence contexts (Figure 23 and Table 9). Even though Z in the GZC sequence posed a stronger replication block relative to GZG sequence in both *E. coli* and human cells, the abundance of TLS pols in the latter resulted in 2 to 3 order of magnitude more efficient TLS.

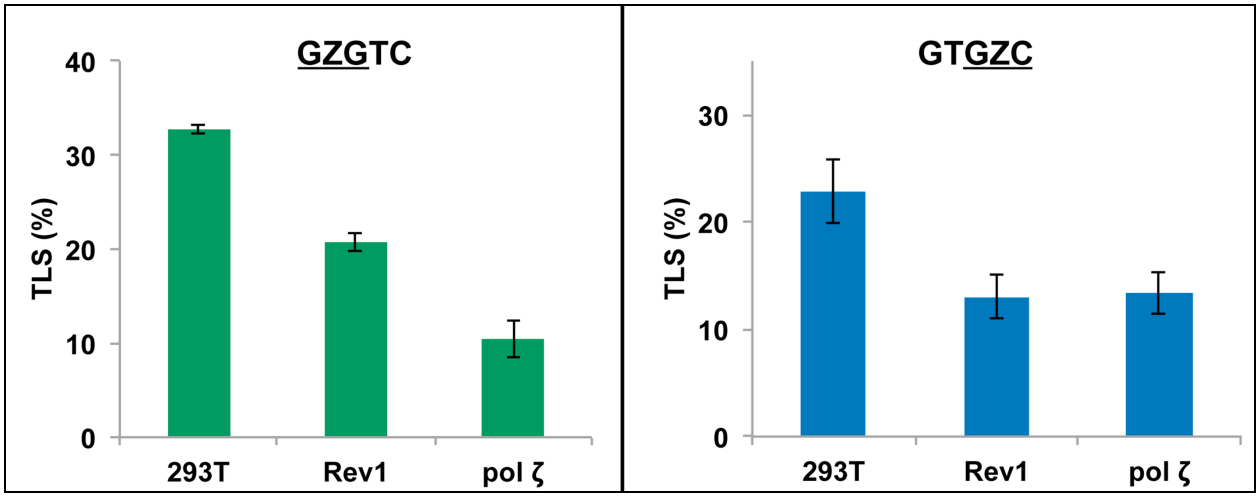


Figure 23: Effects of siRNA knockdowns of pol ζ and Rev1 on the extent of replicative bypass of Z for GZGTC and GTGZC constructs. Percent TLS in the pol knockdowns was measured using an internal control of unmodified plasmid containing a different sequence near the lesion site. When control siRNA was used, the % bypass remained the same as in HEK293T cells.

Table 9: TLS % in different polymerase knockdown HEK293T cells

	Strain	Trial	TLS %
GZGTC	293T	1	33
		2	32
		AVG	**32.5±0.5**
	Rev 1	1	22
		2	20
		AVG	**21±1**
	pol ζ	1	13
		2	8
		3	12
		4	8
		AVG	**10±2**
	Strain	**Trial**	**TLS %**
GTGZC	293T	1	20
		2	26
		AVG	**23±3**
	Rev 1	1	11
		2	15
		AVG	**13±2**
	pol ζ	1	11
		2	15
		3	14
		AVG	**13±2**

4.3.6 AP-site mutagenesis in HEK293T cells

Figure 24 shows the percent MF normalized by TLS to display the relative proportions of various types of mutations combined with the bypass efficiencies in HEK 293T cells. Unlike the progeny from *E. coli* cells, the TLS of Z more rigidly followed the 'A-rule.' In the GZG and GZC sequence contexts, 89% and 76% Z→T, respectively, were detected (Figure 24, Figure 22, Table 10 and Table 11). Moreover, unlike in *E. coli*, targeted one-base deletions were undetectable in both sequence contexts. The lesion in GZG and GZC sequences, respectively, yielded 6% and 13% Z→A, showing that dTMP insertion opposite Z was the second most prevalent event. dCMP insertion opposite the AP-site occurred at 2.5% and 10% frequency, respectively, in the GZG and GZC sequence. Of all the nucleotides inserted opposite the AP-site, dGMP was least favored, which occurred at 2% frequency in GZG sequence, but it was undetectable in the GZC sequence.

4.3.7 Contribution of pol ζ and Rev1 in TLS of AP-site

In yeast, pol ζ was found essential for AP-site bypass.[182] It was also determined that Rev1 can proficiently insert a C opposite an AP-site, but it is unable to extend the replication product from there, whereas pol η is highly inefficient at both the insertion and extension steps of AP-site bypass.[182, 196, 203] To investigate the roles of pol ζ and Rev1 in AP-site bypass in HEK 293T cells, we employed siRNA knockdown approach to constrain the expression of these pols. The extent of siRNA knockdown was determined by RT-PCR and by Western blotting analysis. For each pol, the knockdown was at least 75% efficient. When the cells were treated with a control siRNA, the extent of TLS of AP-site remained the same. By contrast, in the GZC sequence, knockdown of either pol ζ or Rev1 resulted in more than 40% reduction in TLS(Figure 23 and Table 9). In the GZG sequence, though, 36% and 69% reduction in TLS occurred

in the Rev1 and pol ζ knockdown cells, respectively(Figure 23 and Table 9). Therefore, similar to the result in yeast [182], in human cells, pol ζ plays a key role; yet, it does not appear to be indispensable for AP-site bypass, because 30-57% bypass occurred in the absence of pol ζ. Whether the remaining low concentration of the enzyme in pol ζ knockdown cells is enough to carry out this level of TLS is a question we cannot address at this time. Rev1 also is important for TLS in human cells, but it does not insert dCMP opposite the AP-site. Rev1's likely role in AP-site bypass is the scaffolding function of the C-terminal domain that interacts with the Rev1-interacting region of other pols, including the Rev7 subunit of pol ζ.[204] The type of mutations in the Rev1 and pol ζ knockdown cells, however, did not change appreciably from that in HEK 293T cells (Figure 24, Figure 25, Table 10 and Table 11).

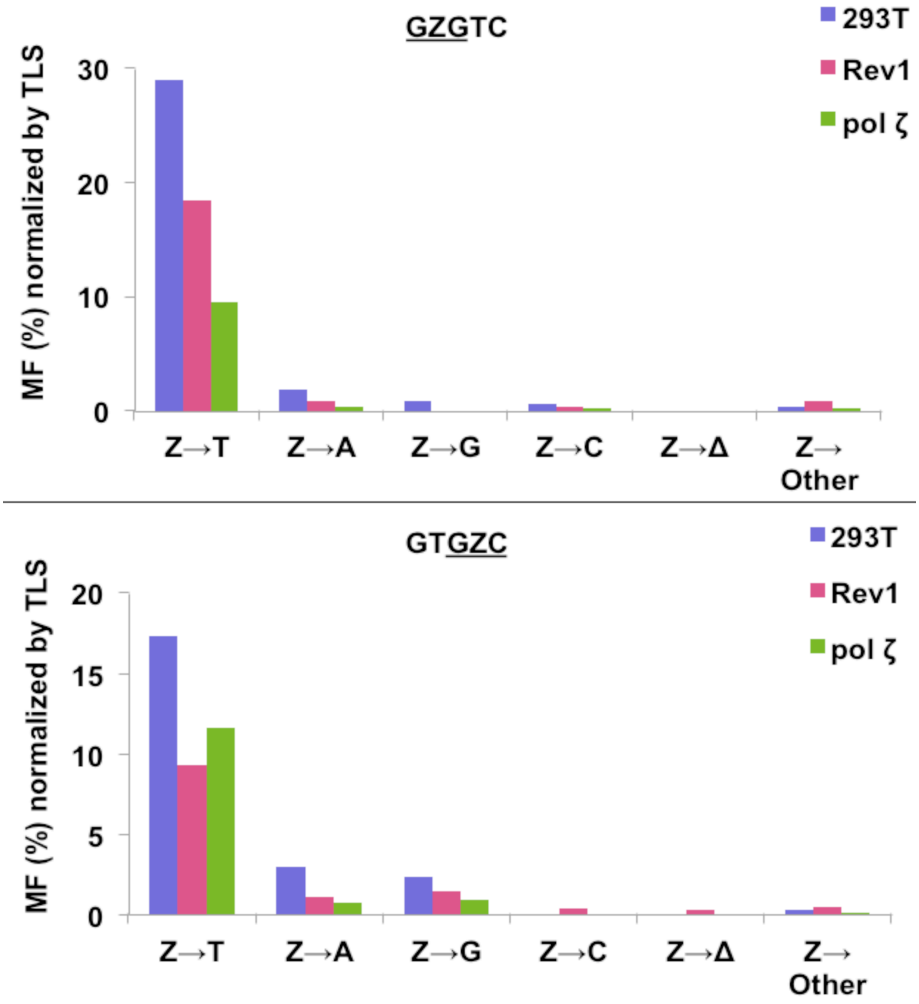


Figure 24: Percent mutations induced by Z in GZGTC and GTGZC sequence contexts multiplied with % TLS for GZGTC and GTGZC constructs in HEK293T cells without or with siRNA knockdowns of pol ζ and Rev1

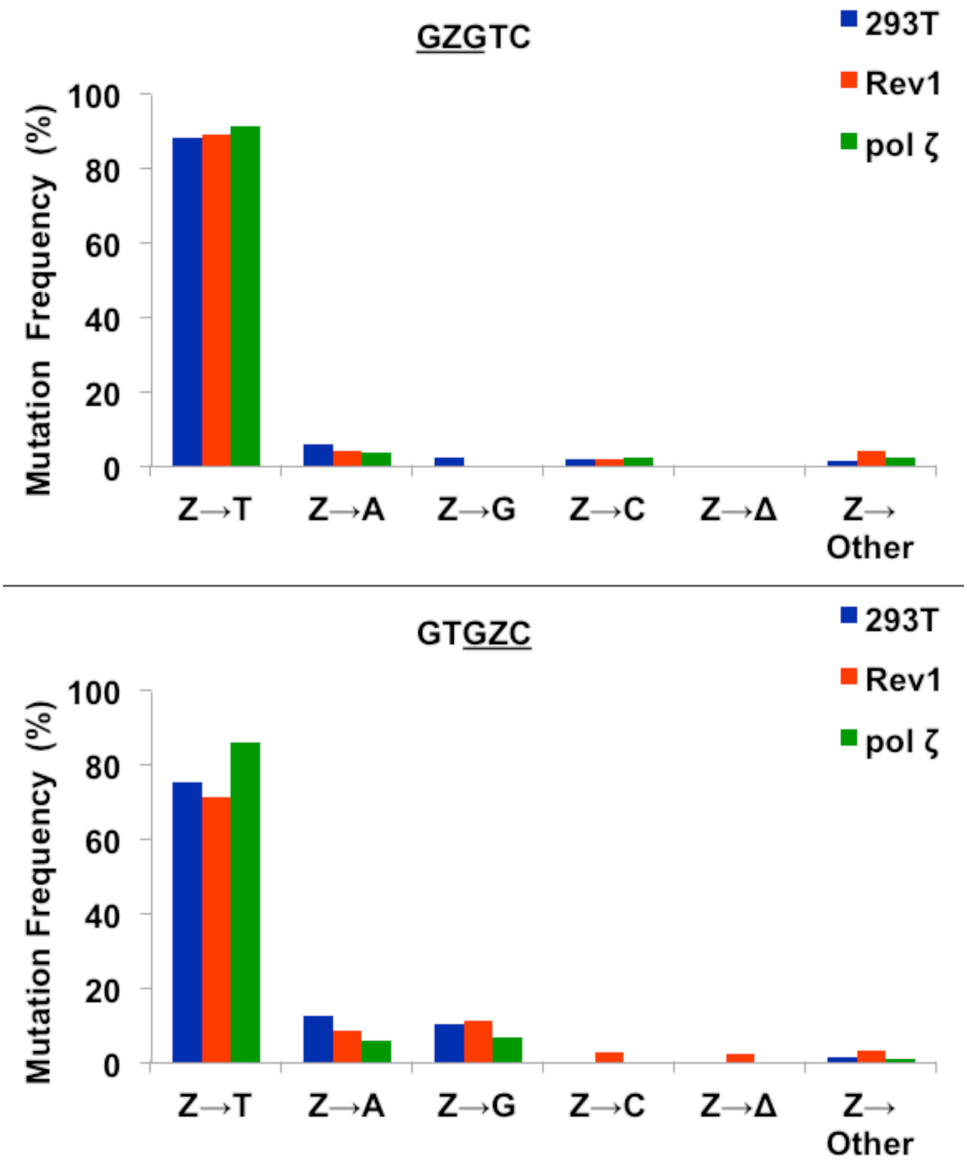


Figure 25: Percent mutations induced by Z in GZGTC and GTGZC sequence contexts in HEK 293T cells without or with siRNA knockdowns of pol ζ and Rev1

Table 10: Mutation frequency of GZGTC in TLS polymerase knockdown HEK293T cells[†]

Lesion		Trial	Total Signals	Z →T	(%)	Z→A	(%)	Z→G	(%)	Z→C	(%)	Z→Del	(%)	Other	(%)
GZGTC	**293T**	1	43	37[a,b]	(86)	3	(7)	3[c]	(7)	0	(0)	0	(0)	0	(0)
		2	61	54[d]	(89)	4	(7)	0	(0)	1	(2)	0	(0)	2[e]	(3)
		3	53	48[f]	(91)	2	(4)	1	(2)	2	(4)	0	(0)	0	(0)
	Total		**157**	**139**	**(89)**	**9**	**(6)**	**4**	**(3)**	**3**	**(2)**	**0**	**(0)**	**2**	**(1)**
	Rev 1	1	41	37[b]	(90)	2	(5)	0	(0)	1	(2)	0	(0)	1[g]	(2)
		2	53	47[a]	(89)	2	(4)	0	(0)	1	(2)	0	(0)	3[g]	(6)
	Total		**94**	**84**	**(89)**	**4**	**(4)**	**0**	**(0)**	**2**	(2)	**0**	**(0)**	**4**	**(4)**
	Pol ζ	1	82	74[b]	(90)	3	(4)	0	(0)	2	(2)	0	(0)	3[g]	(4)
		2	46	43[b,h]	(94)	2	(4)	0	(0)	1	(2)	0	(0)	0	(0)
	Total		**128**	**117**	**(91)**	**5**	**(4)**	**0**	**(0)**	**3**	**(2)**	**0**	**(0)**	**3**	**(2)**

[†]The superscript indicates one or more mutants containing mutation elsewhere shown below and the number in parenthesis shows the number of events detected.

(a) TGC CGT GTC AGC (1), (b) TGC ATT GTC AGC (1) (c) TGC GAG GTC AGC (2), (d) TGC _ _ T GTC AGC (1), (e)_ _ _ _ _ _ GTC AGC (1), (f) TGC GAT GTC AGC (2), (g) _ _ _ _ _ _ _ _ _ C AGC (1 in expt 1, 3 in expt 2 of Rev1; 3 in expt 1 of pol ζ), (h) _ _ _ AGT GTC AGC (1)

Table 11: Mutation frequency of GTGZC in TLS polymerase knockdown HEK293T cells[†]

Lesion		Trial	Total Signals	Z →T	(%)	Z→A	(%)	Z→G	(%)	Z→C	(%)	Z→ Del	(%)	Other	(%)
GTGZC	**293T**	1	27	20	(74)	4	(15)	3	(11)	0	(0)	0	(0)	0	(0)
		2	51	39[i]	(77)	6	(12)	5	(10)	0	(0)	0	(0)	1[j]	(2)
	Total		**78**	**59**	(76)	**10**	**(13)**	**8**	**(10)**	**0**	**(0)**	**0**	**(0)**	**1**	**(1)**
	REV1	1	78	56[i]	(72)	9	(12)	10	(13)	2[k]	(3)	0	(0)	1[l]	(1)
		2	39	25[i]	(64)	4	(10)	6	(15)	1[m]	(3)	3	(8)	0	(0)
		3	59	45[i,n]	(76)	2	(3)	4	(7)	2	(3)	1	(2)	5[l,o]	(9)
	Total		**176**	**126**	**(72)**	**15**	**(9)**	**20**	**(11)**	**5**	**(3)**	**4**	**(2)**	**6**	**(3)**
	pol ζ	1	54	48	(89)	3	(6)	3	(6)	0	(0)	0	(0)	0	(0)
		2	34	28	(82)	2[p]	(6)	3[q]	(9)	0	(0)	0	(0)	1[o]	(3)
	Total		**88**	**76**	**(86)**	**5**	**(6)**	**6**	**(7)**	**0**	**(0)**	**0**	**(0)**	**1**	**(1)**

[†]The superscript indicates one or more mutants containing mutation elsewhere shown below and the number in parenthesis shows the number of events detected.

(i) TGC AAT GTC AGC (1 in 293T, 1 each in expt 1 & 3 and 2 in expt 2 of Rev1), (j) TG - - - - G- C AGC (1) ,(k) TGC AG- CCC AGC (1) ,(l) - - - - - - - - - C AGC (4), (m) TGC AGT ACC AGC (1), (n) _ _ _ _ _ _ TT C ACG(1), TGC TGT GTC AGC(1), _ _ _ AGT GTC AGC (1), (o) T_ _ _ _ _ _ _ _ AGC (1) , (p) _ _ _ GAT GAC AGC (1), (q) TGC AAT GGC AGC (1)

4.4 Discussion and Conclusions

Analysis of bypass efficiency in *E. coli* indicated that TLS of AP-site is a particularly rare event and that pol V is required for 60-70% of the bypass in the GZC sequence, but it plays a less critical role in the GZG sequence. *In vitro* studies have shown that pol V can insert dAMP opposite an AP-site much more proficiently than pol III or pol IV [191], which was also indicated in our *in vivo* experiments. Thermodynamic studies have established that an A opposite the AP-site is enthalpically more favorable than a T [205], but it is likely that a pol can override thermodynamic considerations in the active site of the enzyme. In the absence of SOS in the current study, the major bypass involved one-base deletions. The mechanism of one-base deletion is not clear. Although dCMP insertion opposite the AP-site followed by slippage in GZN sequence to pair with the 5'G can be postulated, it cannot account for the lack of slippage when dAMP is incorporated opposite AP-site in TZG sequence used in another study.[188] Perhaps the GC pair can stabilize the slippage more efficiently than the AT pair. The role of pol II, pol IV, or both in one-base deletion is likely, but more experiments are needed to rationalize the context effects. With SOS, the full-length products were increased to 49% in GZG and 71% in GZC sequence, with predominantly dAMP inserted opposite the AP-site. In the GZG sequence, one-base deletions occurred at the same frequency as dAMP insertions. Livneh and coworkers have analyzed TLS of AP-site in a GZG sequence, but except for the immediate neighbors, the sequence was different from this study.[206] Furthermore, the lesion was located in the single-stranded gap region of a gapped duplex plasmid, and TLS gave 76% dAMP insertions opposite the AP-site and 21% targeted one-base deletions with SOS. While both pol IV and pol V can bypass Z *in vitro* [206, 207], pol V is much more efficient than pol IV [191], and pol V and not pol IV was suggested to play a major role *in vivo.*[206] Besides, suggestions have been made that TLS by

pol V results in base substitutions, whereas deletions are more likely to result from bypass by pol IV or even pol III, although the latter is much less efficient in bypassing AP-sites.[191, 206] In our study a clear-cut role of pol IV in deletion was not identified.

TLS studies on AP-site from different laboratories have shown that less than 1% TLS in *E. coli* occur in the absence of SOS, which increased 2- to 10-fold with SOS, whereas about 6% TLS takes place in yeast cells in double-stranded DNA.[180, 182, 188] In contrast, in the single-stranded gap region of a gapped duplex plasmid, 20-90% TLS of AP-site have been reported to occur in human cells.[208] Our comparative study in *E. coli* and human cells using the same sequence context is consistent with these reports. In contrast to *E. coli*, in which less than 1% TLS occurred, in HEK 293T cells we observed 23% and 33% TLS in the GZC and GZG sequence contexts. As in *E. coli*, GZC site was a stronger replication block than GZG, but much higher TLS in both sites suggests that the pols in human cells are more proficient in bypassing AP-sites. dAMP insertion opposite the AP-site was the dominant incidence in both sites with >75% Z→T substitutions. The second and third most prevalent events were dTMP and dCMP insertions, respectively, whereas dGMP insertion occurred at a very low frequency. One-base deletion, the major outcome noted in *E. coli*, was undetectable. This is noteworthy because *in vitro* studies on AP-site templates show that pol β and pol λ can promote template slippage to generate -1 frameshifts.[209] Our result, therefore, implies that these X-family pols are not involved in TLS of AP-site during replication in HEK 293T cells, although they might be important for re-synthesis step of repair pathways. Evidently, the bypass polymerases in mammalian cells can avoid the deleterious consequence of frameshift mutations, which parallels our earlier observation in simian kidney cells.[210] Using a gapped plasmid vector system in human adenocarcinoma H1299 cells, insertion of dAMP opposite Z was reported [208], as we have

observed in HEK 293T cells. We also determined that pol ζ is a critical bypass pol for TLS of AP-site in human cells, since 44-67% drop in TLS have occurred in pol ζ knockdown cells. Likewise, Rev1 also is important for AP-site bypass, as shown by 36-43% reduction in TLS in Rev1 knockdown cells. Since Rev1 does not serve as a deoxycytidyl transferase in TLS of AP-sites, it probably acts as a structural element for another pol such as pol ζ. But, unlike in yeast [182], neither pol ζ nor Rev1 was deemed indispensable in human cells. However, future experiments with pol ζ- and Rev1-knockout cells will be required to address this question with certainty. It was suggested that AP-site bypass in human cells requires at least one of the replicative pols, α, δ, or ε.[208] Although we did not investigate the roles of the replicative pols in AP-site bypass in human cells, our results are not inconsistent with this notion.

5 Chapter 2[†]: Mutagenicity and Genotoxicity of (5'*S*)-8,5'-Cyclo-2'-Deoxynucleosides in *Escherichia coli* and Human Embryonic Kidney (HEK293T) cells

[†]The text and figures in Chapter 2, in part or in full, are a reprint of the material as it appears in Pednekar, V.; Weerasooriya, S.; Jasti, V. P.; Basu, A. K., Mutagenicity and Genotoxicity of (5 ' S)-8,5 '-Cyclo-2 '-deoxyadenosine in *Escherichia coli* and Replication of (5'S)-8,5 '-Cyclopurine-2 '-deoxynucleosides in Vitro by DNA Polymerase IV, Exo-Free Klenow Fragment, and Dpo4. *Chemical Research in Toxicology* **2014,** 27, (2), 200-210.

5.1 Abstract

Ionizing radiation and other processes that give rise to reactive oxygen species generate many lesions in DNA, including 8,5'-cyclo-2'-deoxyadenosine (cdA) and 8,5'-cyclo-2'-deoxyguanosine (cdG) tandem lesions. Herein the result of replication of a plasmid containing *S*-cdA in *Escherichia coli* is reported. We found *S*-cdA mutagenic and highly genotoxic. Viability and mutagenicity of *S*-cdA, as we have previously noted with *S*-cdG, are dependent on functional pol V, but mutational frequencies (MF) and types varied in pol II- and pol IV-deficient strains. In repair and replication-competent strain, both *S*-cdA→T and *S*-cdA→G substitutions occurred in equal frequency, but the frequency of the latter mutation dropped in pol IV-deficient strain, especially when it was SOS-induced. This suggests that pol IV plays a role in *S*-cdA→G mutations. MF increased significantly in pol II-deficient strain, suggesting its likely role in error-free translesion synthesis. Furthermore, the analysis of progeny from HEK293T cells showed that *S*-cdG (~81%) was much more mutagenic compared to *S*-cdA (21%), and the major types of mutations induced were *S*-cdG→T (43.4%) followed by *S*-cdG→A (28.5%). In the case of *S*-cdA, A→ T (7.9%) substitutions were observed as predominant mutation followed by A → G (4.9%). Interestingly, for *S*-cdA a significant number of semi-targeted mutations occurred 5' to the lesion, whereas, for *S*-cdG, it was observed at a low frequency of about 2%. si-RNA induced knockdown of TLS polymerases indicated that pol ι or pol η may be involved in the nucleotide incorporation opposite the cyclopurines and pol κ or pol ζ may be involved in the extension past the lesions.

5.2 Introduction

The tandem DNA lesions, 8,5'-cyclo-2'-deoxyadenosine (cdA) and 8,5'-cyclo-2'-deoxyguanosine (cdG) diastereomers (Figure 1), have been detected in DNA derived from various cells and organisms.[52] A characteristic of these lesions is that both the 2'-deoxyribose and the purine base are damaged, and the C5'-C8 intramolecular cyclization induces an unusual O4'-*exo* (west) pseudorotation in DNA.[48] The O4'-exo (west) pseudorotation causes a perturbation in the helical twist and base pair stacking of DNA, resulting in thermodynamic destabilization of the DNA duplex. However, Watson-Crick base-pairing can be maintained. Owing to the presence of the C5'-C8 covalent bond between the base and the sugar, base excision repair cannot excise cdA and cdG and, indeed, no glycosylase has been found to repair these tandem lesions.[49, 54] But nucleotide excision repair (NER) system can repair these DNA damages,[49, 54] and the efficiency of repair varies with the base located opposite them in the complementary strand.[56] These lesions are suspected to play a role in neurologic diseases in Xeroderma Pigmentosum patients with defects in NER.[50] *S*-cdA accumulates in genomic DNA of csb-/- mice, signifying that the cyclopurine lesions may accumulate in Cockayne syndrome patients.[211] *S*-cdA is also a strong block of gene expression in human cells.[54]

In a previous investigation, we found that *S*-cdG is highly toxic and mutagenic in *Escherichia coli.*[212] High toxicity of this lesion was unexpected, since the C5'-C8 bond locks the base in the *anti*-orientation but does not adversely affect Watson-Crick base-paring.[48] We also determined that error-prone bypass by DNA polymerase V (pol V) is essential for its viability and mutagenesis. Replication blocking and mutagenic properties *S*-cdG and *S*-cdA were reported also by others.[213] In *E. coli*, pol V is one of the three SOS polymerases, which carry out translesion synthesis (TLS) of replication-blocking lesions.[139] Many TLS polymerases, including

pol IV and pol V in *E. coli*, belong to the Y-family of specialized DNA polymerases,[139, 190] whereas pol II belongs to the B-family.[214] The TLS polymerases have more spacious active sites allowing them to accommodate many DNA lesions. Unlike the replicative polymerases, the lack of proofreading exonucleolytic function of the TLS polymerases permits them to continue DNA synthesis, albeit with a lower fidelity. For *S*-cdG and *S*-cdA, in vitro primer extension studies were performed with *Saccharomyces cerevisiae* and human polymerase η, and both bypass these lesions accurately and efficiently.[215]

In the current work, we have evaluated error-free and error-prone TLS of *S*-cdA in *E. coli* and mutagenic properties of both *S*-cdA and *S*-cdG in human embryonic kidney cells (HEK 293T) that are either untreated or siRNA knockdown in TLS polymerases. We show that, like *S*-cdG, *S*-cdA is a strong block of replication and its bypass is entirely dependent on pol V in *E. coli*. Further in mammalian cells *S*-cdG is more mutagenic as compared to *S*-cdA and multiple TLS polymerases may be involved in lesion bypass.

Oligonucleotides

5’ – GTGCGXGTTTGT – 3’

X = dG (or) *S*-cdG (or) *S*-cdA

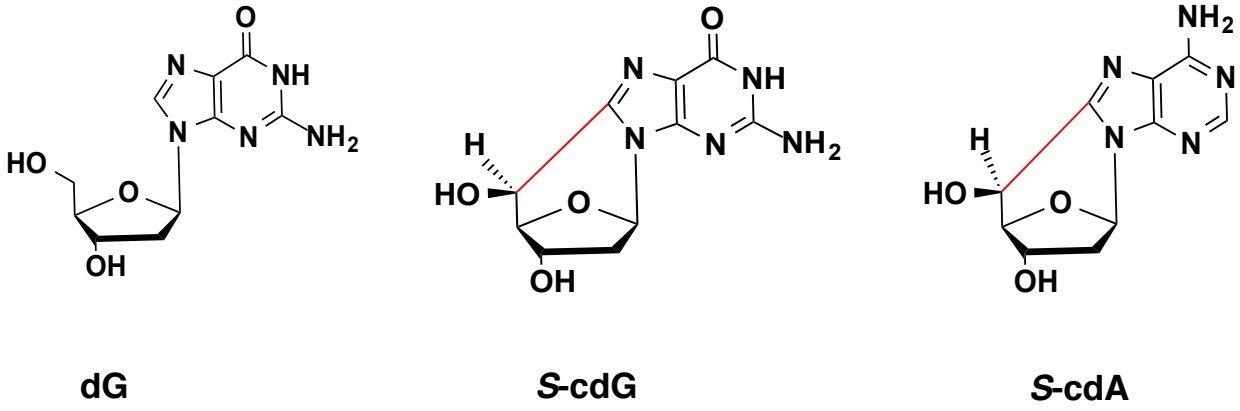


Figure 26: : Oligonucleotides and lesions used in the study

5.3 Results

5.3.1 Viability of *S*-cdA in *E. coli*

Compared to the control, viability of *S*-cdA construct was 0.8% in *E. coli* with normal repair and replication functions, which increased ten-fold with SOS. Viability was even less in pol II- and pol IV- deficient strains. Upon induction of SOS, in pol II-deficient strain viability increased only four-fold, whereas in pol IV-deficient strain viability increased nearly seventeen-fold. Furthermore, no progeny was recovered from pol V-deficient strain or the strain deficient in all three SOS polymerases. These results are similar to what was observed with replication of *S*-cdG construct reported earlier, and we conclude that pol V ($UmuD'_2C$) is required for replication of both *S*-cdA and *S*-cdG. The role of pol II may be more complex in that it plays a secondary role in bypassing these lesions but it is unable to bypass them independently.

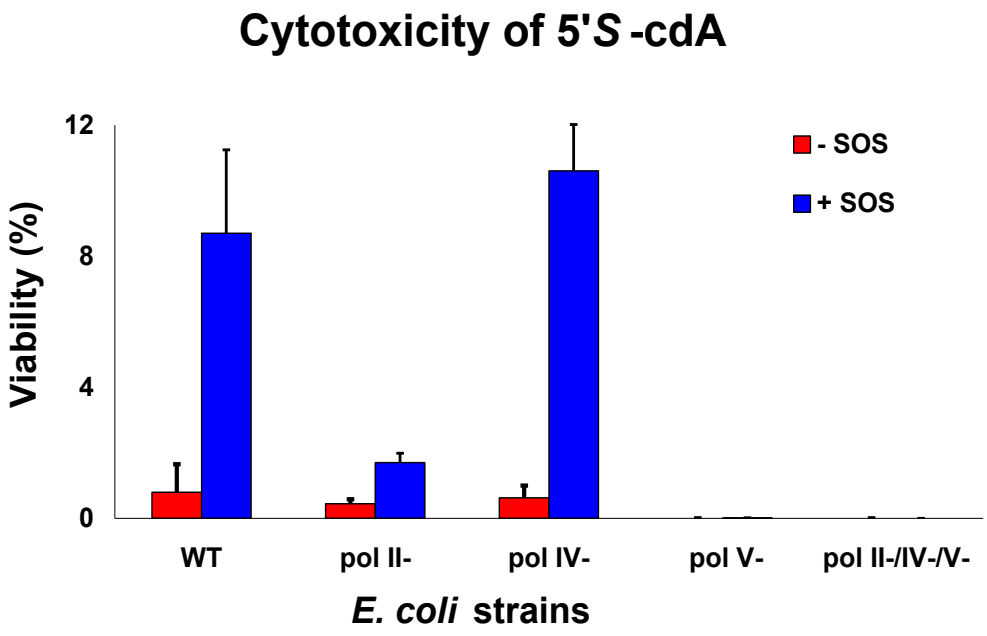


Figure 27: Viability of S-cdA in *E. coli*. Viability was determined by comparing transformation efficiency of the *S*-cdA plasmid with that of the control construct. The data represent average of four independent experiments. SOS was induced with 20 J/m^2 UV irradiation.

Table 12: Viability of S-cdA in *E. coli*[a]

Polymerase knocked out	**-SOS**	**+SOS[b]**
None	0.5 ± 0.2	8.7 ± 2.5
pol II	0.6 ± 0.1	1.7 ± 0.3
pol IV	0.4 ± 0.2	10.6 ± 1.4
pol V	<0.001	<0.005
pol II / pol IV / pol V	<0.001	<0.001

[a]Viability was determined by comparing transformation efficiency of the *S*-cdA plasmid with that of the control construct. The data represent average of four independent experiments.

[b]SOS was induced with 20 J/m^2 UV irradiation.

5.3.2 Mutagenicity of *S*-cdA in *E. coli*

The progeny from each transfection was analyzed by oligonucleotide hybridization followed by DNA sequencing. Mutational frequency (MF) in repair and replication proficient strain was 6.2% and 8.4% without and with SOS, respectively, and equal fraction of mutants carried *S*-cdA→T and *S*-cdA→G in each case. In pol II-deficient strain, the MF increased to more than 15% in the absence of SOS, which did not increase with SOS and, in fact, dropped slightly to 13%. But the increase in MF in the pol II-deficient strain was mainly due to the increase in the *S*-cdA→T events. In contrast, MF was lower in pol IV-deficient strain with 4.2% and 5.6% without and with SOS, respectively, and the major type of mutation with SOS was *S*-cdA→T. Although a clear picture of the type of errors by individual polymerases did not emerge, it seems likely that pol IV plays a role in most *S*-cdA→G mutations, whereas both pol IV and pol V may be involved in *S*-cdA→T mutations. The role of pol II is less certain, but it seems to play a role in error-free bypass. Increased MF in pol II-deficient strain in the absence of SOS was also observed with *S*-cdG.

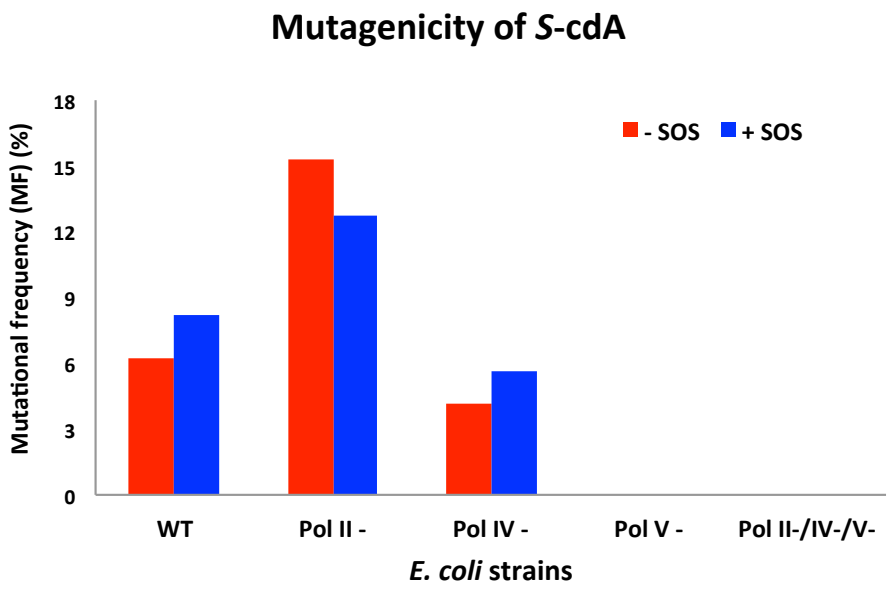


Figure 28: Total mutation frequency of *S*-cdA in different *E. coli* strains with out or with SOS induction. Data from several transformations have been combined (shown in table 13). SOS was induced with 20 J/m^2 UV radiaton.

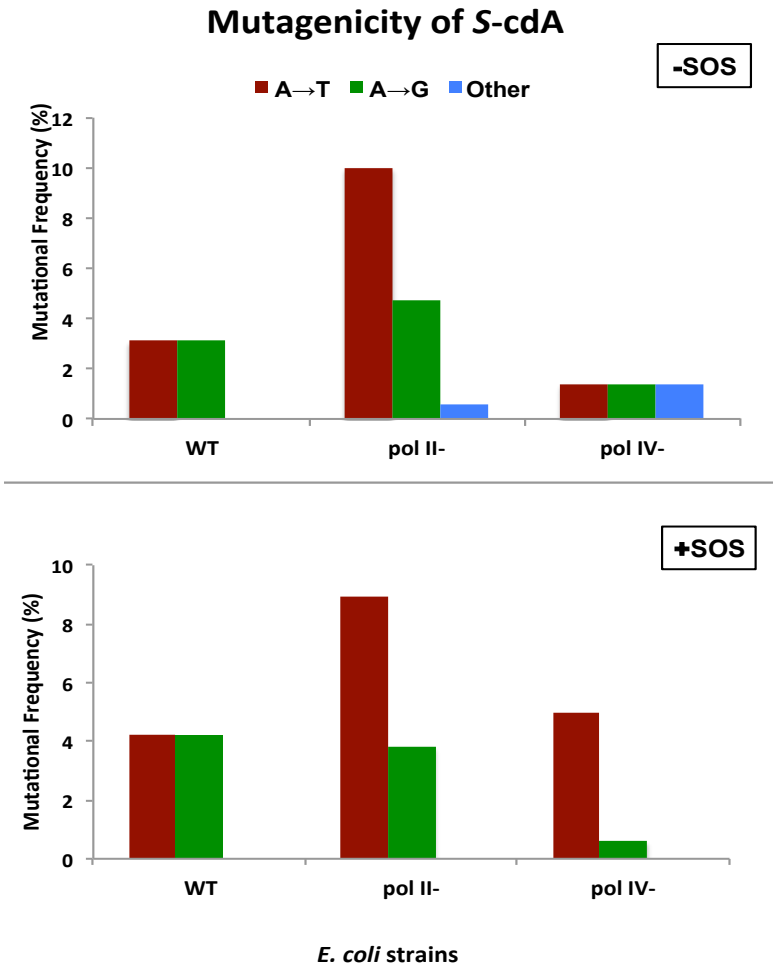


Figure 29: Progeny analysis of the replication of *S*-cdA construct in different *E. coli* strains with out and with SOS. Data from several transformations have been combined (shown in table 13). SOS was induced with 20 J/m^2 UV radiaton

Table 13: Mutations induced by *S*-cdA

Polymerase Knocked out	SOS[a]	Expt No.	No. of. Colonies Screened	No. of Mutations (%)	No. of A*→T (%)	No. of A*→G (%)	Other Mutations (%)
None	-	1	56	5	2	3	0
		2	57	3	3	0	0
		3	49	2	0	2	0
		Total	162	10 (6.2)	5 (3.1)	5 (3.1)	0
	+	1	80	6	3	3	0
		2	86	8	4	4	0
		Total	166	14 (8.4)	7 (4.2)	7 (4.2)	0
pol II	-	1	47	8	6	2	0
		2	53	6	4	2	0
		3	50	9	5	3	1[c]
		Total	150	23 (15.3)	15 (10.0)	7 (4.7)	1[c] (0.6)
	+	1	79	10	6	4	0
		2	78	10	8	2	0
		Total	157	20 (12.7)	14 (8.9)	6 (3.8)	0
pol IV	-	1	72	4	1	1	2[d]
		2	73	2	1	1	0
		Total	145	6 (4.2)	2 (1.4)	2 (1.4)	2[d] (1.4)
	+	1	85	4	3	1	0
		2	76	5	5	0	0
		Total	161	9 (5.6)	8 (5.0)	1 (0.6)	0

[a]SOS was induced with 20 J/m^2 UV irradiation.

[b]For each strain MF of control construct was <1% (data not shown).

[c]A*T → TG

[d]A* → C

5.3.3 Mutagenicity of *S*-cdA and *S*-cdG in HEK293T cells

We investigated the mechanism of translesion synthesis past both *S*-cdG and *S*-cdA in mammalian cells by constructing ss-pMS2 vectors containing a site specifically incorporated *S*-cdA, *S*-cdG and the corresponding unmodified oligomers. The modified and control vectors were replicated in HEK293T cells and the progeny double stranded plasmids were recovered. The recovered plasmids were further amplified in DH10B *E. coli* cells and the resulting colonies were analyzed the by dot-blot hybridization followed by DNA sequencing. Two left and right probes were used to select plasmids containing the correct insert, and transformants that did not hybridize with both the left and right probes were omitted. Any transformant that hybridized with the left and right probes but failed to hybridize with the corresponding wild-type probe was subjected to DNA sequence analysis. The wild type probes used would recognize any progeny that contains a G at the *S*-cdG site or an A at the *S*-cdA site, indicating correct read-through past the lesion by DNA polymerase and are considered non-mutants. The progeny that failed to hybridize to the respective wild type probes were considered mutants and the type of mutations were determined by DNA sequencing.

The mutational frequency (MF) was estimated by comparing the number of mutants obtained to the total number of colonies screened. As shown in Figure 30 and Table 14, we observed that *S*-cdG exhibited significant mutagenicity in HEK293T cells showing a total mutation frequency (MF) of 81%. In contrast, for *S*-cdA, 79% progeny contained an A opposite the *S*-cdA site indicating accurate lesion bypass by DNA polymerases and showed a total MF of only 21% (Figure 30 and Table 15). The predominant types of mutations caused by *S*-cdG were base substitutions, and G → T (43.4%) transversions were the major event followed by G →A (28.5 %) transitions. In addition, G → C (7.1%) transversions and semi-targeted mutations also

occurred (Figure 30 and Table 1). In the case of *S*-cdA, A→ T (7.9%) substitutions were the major mutations followed by A → G (4.9%) transitions (Figure 30 and Table 15). A striking feature of the mutational spectrum of *S*-cdA was that a significant number of mutations occurred 5' to the lesion, which we describe as semi-targeted mutations. The most prevalent semi-targeted mutations involved a C → T transition immediately 5' to *S*-cdA at a frequency (4.9%) similar to that of A → G targeted mutations (4.9%) (Figure 30 and Table 15). However, for *S*-cdG, semi-targeted and other mutations occurred at a low frequency of about 2%, which accounts to less than 3% of the targeted base substitution events (Figure 30). It is also important to note that no mutations were detected within the 12mer sequence of the control constructs of approximately 300 progeny analyzed.

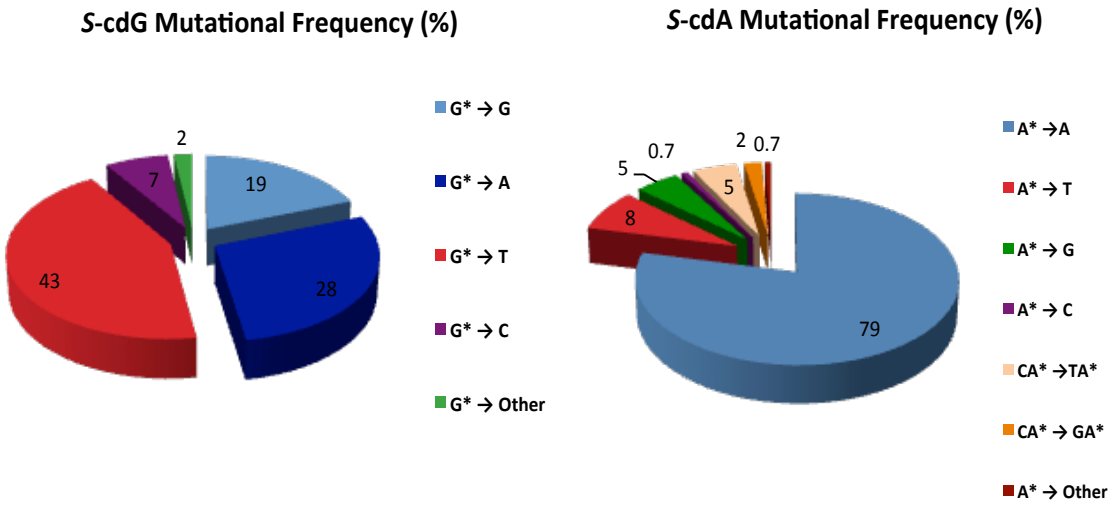


Figure 30: Analysis of progeny derived from replicating either *S*-cdG or *S*-cdA containing ss-pMS2 construct in HEK293T cells treated with NC siRNA. The data from three independent experiments have been combined.

5.3.4 Mutational specificity of *S*-cdG and *S*-cdA in TLS polymerase knockdown HEK293T cells

We employed a siRNA induced TLS polymerase knockdown approach to find out the involvement of specific TLS polymerases in the lesion bypass. The siRNA against a specific TLS polymerase was introduced to HEK293T cells and the knockdown efficiency was determined by quantitative reverse transcription PCR and Western blot analysis. Consistent with a previous study from our group, knockdown efficiency for individual polymerase was determined to be at least 70% efficient.[216]

Relative to untreated or treated with a negative control siRNA (NC-siRNA) HEK293T cells, we observed only a modest effect on the total MF in TLS pols knockdown cells (Figure 31). The total MF of *S*-cdG (of ~81%) increased by ~12% to 90% in pol κ knockdown cells compared to the NC-siRNA treated (or untreated) cells, and this was mainly due to a significant increase in G → A mutations, which increased by 33%. Whereas the total MF of *S*-cdG in pol ι remained approximately the same as in NC-treated cells, there was a ~7-8% increase in the total MF in pol η, REV1 or pol ζ knockdown cells, and the MF increased to ~86%. (Figure 31, Figure 32, and Table 14). However, in pol ζ-knockdown cells, targeted G → T mutations were increased by 18% compared to the same in NC-treated HEK293T, whereas in pol η knockdown cells the increase in MF was due to an increase in semi-targeted mutations (Figure 32 and Table 14). The reduction in total MF in pol ι knockdown cells was due to a 20% reduction in G → A mutations (Figure 32 and Table 14). Above results suggest that pol η, pol κ, pol ι and pol ζ likely contribute to mutagenicity of *S*-cdG in mammalian cells, although no single TLS polymerase is critical for the error-prone bypass of *S*-cdG.

Similarly, compared to NC-treated HEK293T cells, *S*-cdA exhibited no significant change in MF with knockdown of either pol η or REV1 (Figure 31, Figure 32 and Table 14). In stark contrast to the increase in MF in the case of *S*-cdG in pol κ and pol ζ knockdown cells, MF decreased by 39% and 30%, respectively, in the progeny from the *S*-cdA construct. This reduction was mainly attributed to decrease in A→T mutations. In addition to this, decrease of semi-targeted mutations in pol κ knockdown cells and A → G mutation in pol ζ knockdown cells were also observed. Further, A → C events were undetected in pol κ or pol ζ knockdown cells (Figure 32 and Table 15).

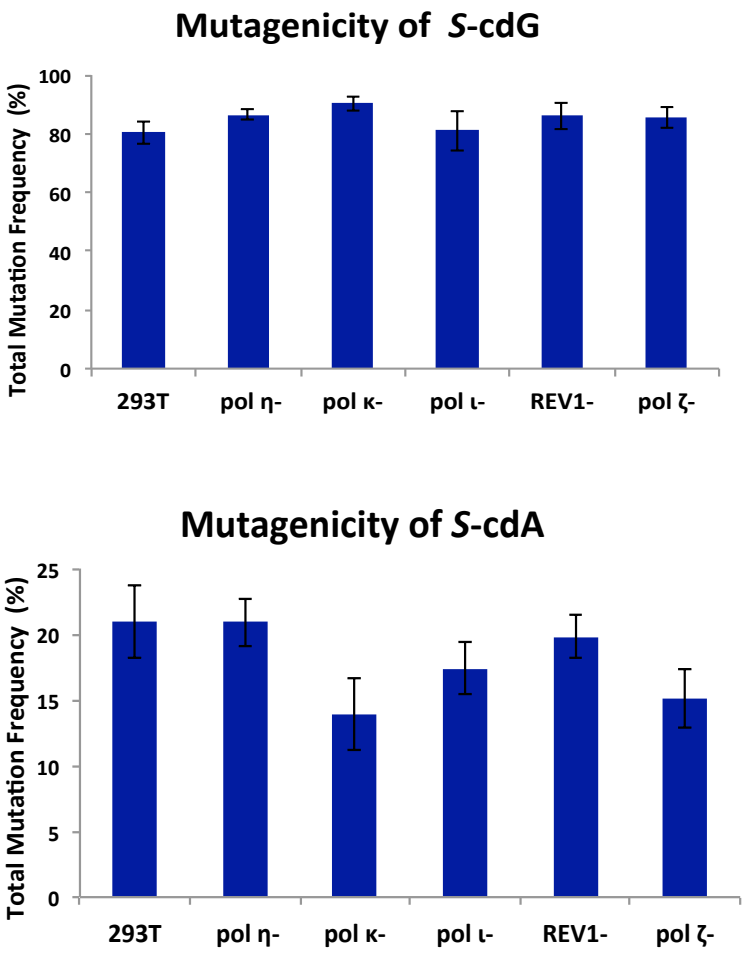


Figure 31: Effect of siRNA knockdowns of TLS pols on mutational frequency of *S*-cdG and *S*-cdA in HEK293T cells. The 293T represents HEK293T cells treated with NC siRNA. Data from several independent transformations have been combined (shown in table 14 and 15).

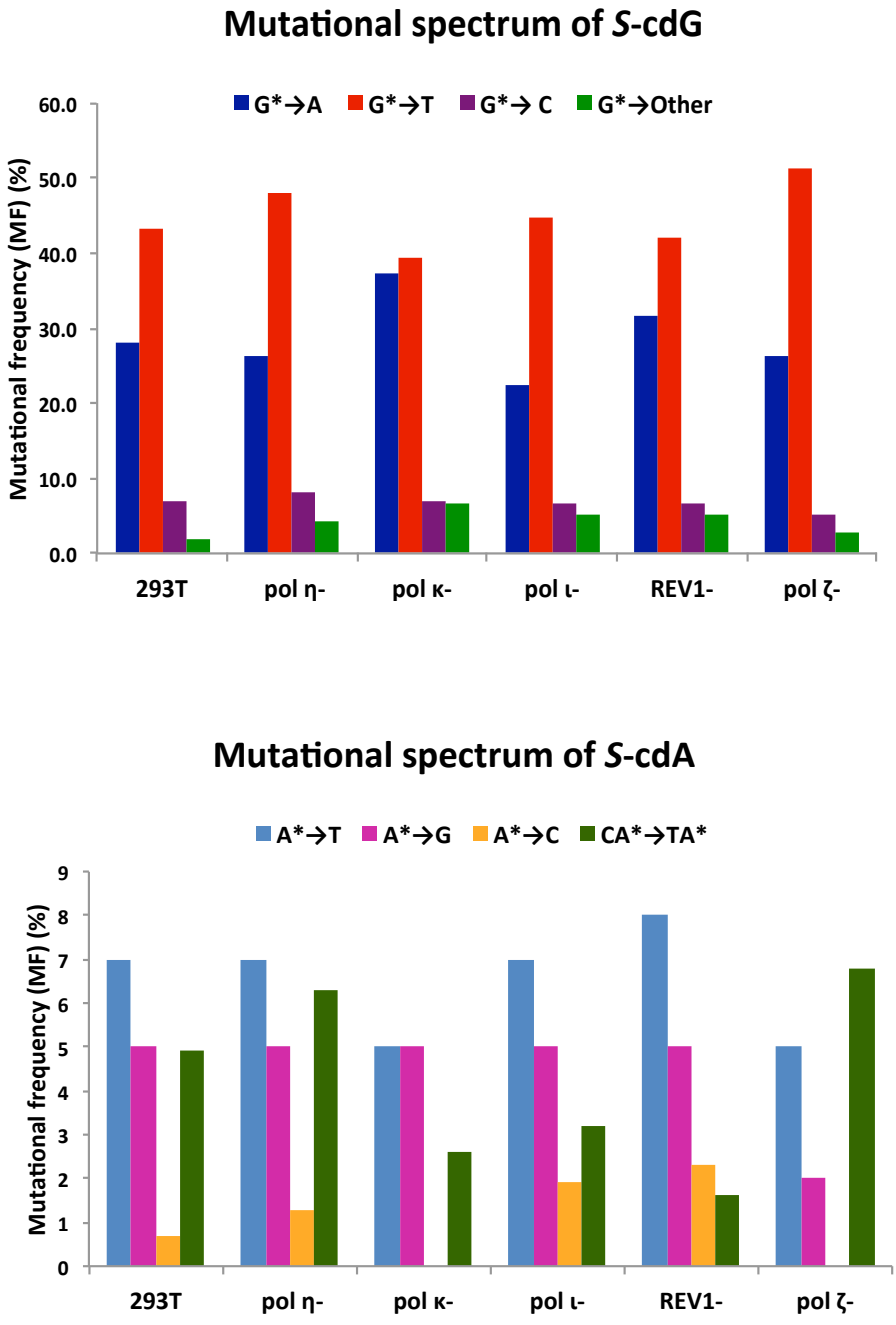


Figure 32: Mutational frequency of *S*-cdG and *S*-cdA in HEK293T cells treated with NC siRNA (293T) and siRNA against TLS polymerase. The data represent combined mutational frequency from several independent experiments (Shown in table 14 and 15).

Table 14: Mutational frequency of *S*-cdG in HEK293T cells treated with siRNA against TLS polymerase[†]

siRNA	Exp	Total colonies screened	Total mutants (%)	G* → G	(%)	A	(%)	T	(%)	C	(%)	Other	(%)
NC-siRNA	1	123	94 (76.4)	29	(23.6)	34[a]	(27.6)	50[b]	(40.7)	8[c]	(6.5)	2[d]	(1.6)
	2	141	116 (82.3)	25	(17.7)	37[e]	(26.2)	69[f]	(48.9)	8[g]	(5.7)	2[h]	(1.4)
	3	86	72 (83.7)	14	(16.3)	27[i]	(31.4)	33[j]	(38.4)	9	(10.5)	3[k]	(3.5)
	Total	**350**	**282 (80.6)**	**68**	**(19.4)**	**98**	**(28.0)**	**152**	**(43.4)**	**25**	**(7.14)**	**7**	**(2.0)**
POLH	1	138	122 (88.4)	16	(11.6)	41[a]	29.7)	67[b]	(48.6)	8	(8.8)	6[c]	(4.3)
	2	110	94 (85.5)	16	(14.5)	26[e]	(23.6)	51[f]	(46.4)	11[g]	(10.0)	6[h]	(5.5)
	3	115	98 (85.2)	17	(14.8)	28[i]	(24.3)	56[j]	(48.7)	10	(8.7)	4[k]	(3.5)
	Total	**363**	**314 (86.5)**	**49**	**(13.5)**	**95**	**(26.2)**	**174**	**(47.9)**	**29**	**(8.0)**	**16**	**(4.4)**
POLK	1	92	81 (88.0)	11	(12.0)	30[a]	(32.6)	41[b]	(44.6)	4	(4.3)	6[c]	(6.5)
	2	95	88 (92.6)	7	(7.4)	41[d]	(43.1)	31	(32.6)	10	(10.5)	6[f]	(6.3)
	3	98	88 (89.8)	10	(10.0)	35[g]	(35.7)	40[h]	(40.8)	6	(6.1)	7[i]	(7.1)
	Total	**285**	**257 (90.2)**	**28**	**(10.0)**	**106**	**(37.2)**	**112**	**(39.3)**	**20**	**(7.0)**	**19**	**(6.7)**
POLI	**1**	115	85 (73.9)	30	(26.1)	25[a]	(21.7)	47[b]	(40.9)	9[c]	(7.8)	4[d]	(3.5)
	2	115	98 (85.2)	19	(16.5)	28[e]	(24.3)	53	(46.1)	8[f]	(7.0)	7[g]	(6.1)
	3	109	94 (86.2)	15	(13.8)	23[h]	(21.1)	52[i]	(47.7)	17	(15.6)	2[j]	(1.8)
	Total	**339**	**275 (81.7)**	**64**	**(18.9)**	**76**	**(22.4)**	**152**	(44.8)	**34**	**(10.0)**	**13**	**(3.8)**
REV1	1	105	87 (82.9)	18	(17.1)	31	(29.5)	46[a]	(43.8)	4	(3.8)	6[b]	(5.7)
	2	101	90 (89.1)	11	(10.9)	34[c]	(33.7)	41[d]	(40.6)	10[e]	(9.9)	5[f]	(5.0)
	Total	**206**	**177 (86.0)**	**29**	**(14.0)**	**65**	**(31.6)**	**87**	(42.2)	**14**	**(6.8)**	**11**	**(5.3)**
REV3L	1	107	89 (83.2)	18	(16.8)	22	(20.6)	55[a]	(51.4)	8	(7.5)	4[b]	(3.7)
	2	102	90 (88.2)	12	(11.8)	33	(32.4)	52[c]	(51.0)	3	(2.9)	2[d]	(2.0)
	Total	**209**	**179 (85.6)**	**30**	**(14.4)**	**55**	**(26.3)**	**107**	(51.2)	**11**	**(5.3)**	**6**	**(2.9)**

[†]The superscript indicates one or more mutants containing mutation elsewhere shown in Table 16 and the number in parenthesis shows the number of events detected.

Table 15: Mutational frequency of *S*-cdA in HEK293T cells treated with siRNA against TLS polymerase knockdown cells[†]

siRNA	Exp	Total colonies screened	Total mutants (%)	A*→				Semi-targeted and other		
				A (%)	T (%)	G (%)	C (%)	CA*→TA*(%)	CA*→GA*(%)	Other (%)
NC-siRNA	1	88	21 (23.9)	67 (76.1)	5 (5.7)	5[a] (5.7)	1 (1.1)	5 (6.0)	3 (3.4)	2[b] (2.3)
	2	93	17 (18.3)	76 (81.7)	5[c] (5.4)	3 (3.2)	0 0.0	8 (8.6)	1 (1.1)	0 0.0
	3	124	26 (21.0)	98 (79.0)	14[d] (11.3)	7[e] (5.6)	1[f] (0.8)	2 (1.6)	2 (1.6)	0 0.0
	Total	**305**	**64 (21.0)**	**241 (79.0)**	**24 (7.9)**	**15 (4.9)**	**2 (0.7)**	**15 (4.9)**	**6 (2.0)**	**2 (0.7)**
POLH	1	80	18 (22.5)	62 (77.5)	5[a] (6.3)	5 (6.3)	1 (1.3)	6 (7.5)	1 (1.3)	0
	2	118	25 (21.2)	93 (78.8)	7[b] (5.9)	7 (5.9)	1 (0.8)	9[c] (7.6)	1 (0.8)	0
	3	121	23 (19.0)	98 (81.0)	11[d] (9.1)	4 (3.3)	2 (1.7)	5 (4.1)	1 (0.8)	0
	Total	**319**	**66 (20.7)**	**253 (79.3)**	**23 (7.2)**	**16 (5.0)**	**4 (1.3)**	**20 (6.3)**	**3 (0.9)**	**0**
POLK	1	101	14 (13.9)	87 (86.1)	4[a] (4.0)	8 (7.9)	0	1 (1.0)	0 0.0	1[b] (1.0)
	2	133	16 (12.0)	114 (85.7)	9 (6.8)	3 (2.3)	0	5 (3.8)	1 (0.8)	1[c] (0.8)
	Total	**234**	**30 (12.8)**	**201 (85.9)**	**13 (5.6)**	**11 (4.7)**	**0**	**6 (2.6)**	**1 (0.4)**	**2 (0.9)**
POLI	1	50	8 (16.0)	42 (84.0)	3[a] (6.0)	5 (10.0)	0	0 0.0	0	0 0.0
	2	106	20 (18.9)	86 (81.1)	8 (7.5)	1 (0.9)	3 (2.8)	5 (4.7)	2 (1.9)	1[b] (0.9)
	Total	**156**	**28 (17.9)**	**128 (82.1)**	**11 (7.1)**	**6 (3.8)**	**3 (1.9)**	**5 (3.2)**	**2 (1.3)**	**1 (0.6)**
REV1	1	43	9 (20.9)	34 (79.1)	3[a] (7.0)	3 (7.0)	1 (2.3)	2 (4.7)	0	0 0.0
	2	86	16 (18.6)	70 (81.4)	8[b] (9.3)	2 (2.3)	2 (2.3)	0	3 (3.5)	1[c] (1.2)
	Total	**129**	**25 (19.4)**	**104 (80.6)**	**11 (8.5)**	**5 (3.9)**	**3 (2.3)**	**2 (1.6)**	**3 (2.3)**	**1 (0.8)**
REV3L	1	125	17 (13.6)	108 (86.4)	4 (3.2)	2[a] (1.6)	0	10 (8.0)	1 (0.8)	0
	2	66	11 (16.7)	55 (83.3)	5 (7.6)	2 (3.0)	0	3 (4.5)	1 (1.5)	0
	Total	**191**	**28 (14.7)**	**163 (85.3)**	**9 (4.7)**	**4 (2.1)**	**0**	**13 (6.8)**	**2 (1.0)**	**0**

[†]The superscript indicates one or more mutants containing mutation elsewhere shown in Table 17 and the number in parenthesis shows the number of events detected.

Table 16: Semi-targeted and other mutations induced by *S*-cdG in TLS polymerase knockdown HEK293T cells

siRNA	Semi targeted and other mutations as indicated by superscript
NC-siRNA	(a) ATT CGT GTT TGT [1] (b) GTT TAT GTT TGT [1], GTG _AT GTT TGT [1], CGA GATGTTTGT [1], 5'GTG C_T GTT TGT [1] (c) ATGCTTGTTTGT [1] (d) GTG TTA ATT TGT [1], GTG CTTT TTTGT [2] (e) GTG CGG GTT TGT [1], GTG CGA GTT TGT [1] (F) GTG GCT GTT TGT [1] (g) GTT GAT GTT TGT [1], GTT TAT GTT TGT [1]
POLH	(a) GTG CAA GTT TGT [1], GTA GAT GTT TGT [1], GTG GAT GTT TGT [1] (b) GTG ATT GTT TGT [1], GTT GTT GTT TGT [2] (c) GTG TGT GTT TGT [1], GTG GGT GTT TGT [5] (d) GTG _GT GTT TGT [1], GTG TGT GTT TGT [1] (e) ATG CAT GTT TGT [1] (f) GTG GTG GTT TGT [1], GTA CTT GTT TGT [2], ATG CTT GTT TGT [1] (g) GTG TCT GTT TGT [1], GTT GCT GTT TGT [1] (h) GTG GGT GTT TGT [1], GTG TGT GTT TGT [1] (i) GTT CAT GTT TGT [1] (j) GTG TTT GTT TGT [1] (k) GTG GGT GTT TGT [2], GTG TGT GTT TGT
POLK	(a) GTT GAT GTT TGT [1], (b) GTG CTT CTT TGT [1], GTG GTT GTT TGT [2], (c) GTG CGC TTG TGT [1], GTG TGT GTT TGT [2], ___ ___ GTT TGT [1], GTG GGT GTT TGT [2] (d) GTA CTT GTT TGT [1] (e) GTG TCT GTT TGT [3], GTG GCT GTT TGT 1], (f) GTG GGT GTT TGT [3], GTG TGT GTT TGT [2], GTT GGT GTT TGT [1] (g) GTG GAT GTT TGT [2], GTA GAT GTT TGT [1] (h) GTT GTT GTT TGT [1] (i) GTG GGT GTT TGT [2], GTT GGT GTT TGT [1], GTG TGT GTT TGT [1], TTG _ GT GTT TGT [1]
POLI	(a) ATG CAT GTT TGT [1], GAG CAT GTT TGT [1] (b) GTT GTT GTT TGT [1], GTT CTT GTT TGT [1] (c) GTG TCT GTT TGT [1] (d) GTG – GA GTT TGT [1], GTG TGT GTT TGT [2], GTG – GT GTT TGT [1] (e) GTG GAT GTT TGT [1], GTA CAT GTT TGT [1] (f) GTT GCT GTT TGT [1], GTG TCT GTT TGT [1] (g) GTG GGT GTT TGT [2], GTT GGT GTT TGT [2], GTG TGT GTT TGT [2], GTG _ GT GTT TGT [1] (h) GTG GAT GTT TGT [2], GTG -AT GTT TGT [1] (i) GTG GTT GTTTGT [1], GTA CTT GTT TGT [1] (j) GTG TGT GTT TGT [1], GTT GGT GTT TGT [1]
REV1	(a) GTT GTT GTT TGT [1], TTG CTT GTT TGT [1], (b) GTT GGT GTT TGT [4], GTG GGT GTT TGT [2] (c) GTT TAT GTT TGT [1] (d) GTG ATT GTT TGT [1] (e) GTG TCT GTT TGT [1] (f) GTG TGT GTT TGT [1], GTG GGT GTT TGT [2], GTG AGT GTTTGT [1] ------ GTT TGT [1]
REV3L	(a) GTG ATT GTT TGT [2] (b) GTT GGT GTT TGT [4], GTG CGT GTG CTT [1] (c) GTT GTT GTT TGT [2] (d) GTG GGT GTT TGT [1], GTG – GT GTT TGT [1]

Table 17: Semi-targeted and other mutations induced by *S*-cdA in TLS polymerase knockdown HEK293T cells

siRNA	Semi-targeted and other mutations as indicated by superscript
NC-siRNA	(a) ATT CGT GTT TGT (b) GTG _AT GTT TGT [1], GTG C_T GTT TGT [1] (c) ATG CTG TTTGT [1] (d) GTG TTA ATT TGT [1], GTG CTT TTT TGT [2] (e) GTG CGG GTT TGT [1], GTG CGA GTT TGT [1] (f) 5'GTG GCT GTT TGT [1] (g) GTT GAT GTT TGT [1], GTT TAT GTT TGT [1]
POLH	(a) GTG ATT GTT TGT (b) GTG TTT GTT TGT (c) GTG TAT ATT TGT
POLK	(a) GTG CTT TTT TGT [1] (b) GTG CAT GAT TGT [1], (c) GTC CAT GTT TGT
POLI	(a) GTG TGT GTT TGT [1] (b) GTT CAT GTT TGT [1]
REV1	(a) GTG GTT GTT TGT [1] (b) GTG TTT GTT TGT [4], GTG GTT GTT TGT [1] (c) GTG CAT ATT TGT [1]
REV3L	(a) GTG TGT GTT TGT [1]

5.4 Discussion and Conclusions

5.4.1 Mutagenicity and cytotoxicity of *S*-cdA in *E. coli*

Both *S*-cdG and *S*-cdA contain an extra covalent bond between the base and the sugar. Therefore, it is not surprising that some of their chemical and biological characteristics are similar. Even so, they exhibit distinct biological properties as well. In terms of similarities, *S*-cdG and *S*-cdA are highly genotoxic in *E. coli*, with viability less than 1% relative to control, which increased several-fold with SOS (Table 1 and ref[212]). Furthermore, they are mutagenic in *E. coli*, but *S*-cdG is considerably more mutagenic than *S*-cdA (Table 2 and ref[212]). TLS of both lesions is entirely dependent on pol V since in pol V-deficient strains viability dropped to nearly 0%. With respect to their differences, in *E. coli*, the major class of mutations induced by *S*-cdG is G → A transitions,[212, 213] whereas *S*-cdA induced both A → T and A → G base substitutions (Table 2). *S*-cdA induced A → T transversions and A → G transitions in equal frequency in cells with normal repair and replication functions, which remained the same with SOS induction, but A → T was the dominant mutation in pol II-deficient and SOS-induced pol IV-deficient strains. The A → G mutations dropped considerably in pol IV- deficient strains, suggesting that pol IV may be able to incorporate dCMP opposite *S*-cdA, but it needs pol V for further extension. In contrast, pol II seems to play a role in error-free replication, but it also is dependent on pol V.

It is worth noting that another study of TLS of *S*-cdG and *S*-cdA in *E. coli* has been published.[213] However, this investigation concentrated only on SOS-induced *E. coli*, whereas our study also determined the bypass and mutagenicity in the absence of SOS. While the pattern of bypass efficiency and mutagenicity in SOS-induced *E. coli* of the earlier report and our results show similar trend, there are significant quantitative differences between the results of these investigations. We believe that the difference stems from the differences between the two

experimental systems. The reported study exhibited a high background on both the bypass efficiency and error rate,[213] whereas the background in our approach has been very low. As a result, we could determine the absolute requirement of pol V in TLS of *S*-cdA (this study) and *S*-cdG (ref [212]), which the other study was unable to ascertain. The bypass efficiency of *S*-cdA in the pol V-deficient strain was 13% in this earlier work,[213] whereas we determined it to be less than 1%. Furthermore, the substantial reduction in MF of A → G mutations in pol IV- deficient strain noted in the current work was not identified, presumably due to the high background MF. However, some of the differences between the studies may also have been due to the location of the lesions in different sequence contexts.

5.4.2 Mutagenicity in HEK293T cells

Our results suggested that both *S*-cdA and *S*-cdG is genotoxic and mutagenic lesions in HEK293T cells. Similar to the results in *E. coli*,[212] *S*-cdG (~81%) was much more mutagenic than *S*-cdA[202] (21%) in human cells (Figure 30). However, both lesions showed increased mutagenicity in HEK293T cells than in *E. coli* cells.[202, 212, 213] The major mutation induced by *S*-cdA was A→T transversions at a frequency of ~8%, followed by 5% A→G transitions, a pattern similar to what was observed in *E. coli*.[202] Unlike in *E. coli*[212], however, in which G→A (26%) was the major base substitution mutations induced by *S*-cdG, in HEK293T cells, G →T tranversions were the predominant mutations, which occurred at a frequency of 43%. Another study reported similar pattern of mutations in human cells.[217] The second and third most prevalent events were G → A (28%) and G →C (7%) mutations as shown in Figure 30. These results indicate that cyclopurine lesions are much more mutagenic in human cells in comparison to bacteria and it seems likely that that the abundance of TLS pols in mammalian cells allows for

efficient replication across these lesions, albeit with low fidelity. We speculate that these endogenous lesions play an important role in the development of human diseases owing to their high mutagenicity.

We employed siRNA knockdown approach to investigate the effect of individual TLS pol in lesion bypass of cyclopurines. We found that the MF of both *S*-cdG and *S*-cdA in pol knockdown cells was comparable to that in HEK293T cells, although minor differences were observed in the mutational patterns of the these lesions. Approximately ~10% increase in the total MF of *S*-cdG occurred in pol κ knockdown cells. This increase in the MF was due to increase in G→A events by 33% as compared to HEK293T cells (Figure 3, and Table 1).We hypothesize that pol κ, which was reported to act as an extender for several other lesions,[196, 218, 219] may be, at least partially, involved in error-free bypass of *S*-cdG *in vivo*. In contrast, it was noted that the MF of *S*-cdA decreased in pol κ and pol ζ knockdown cells compared to that in NC-treated HEK293T cells. Hence it is conceivable that these pols are involved in error-prone bypass of *S*-cdA. TLS is a multistep process involving several polymerase switches, where multiple polymerases act in a sequential manner for lesion bypass.[135, 220] It is generally considered that pol ζ is involved in the extension step, subsequent to incorporation of a base opposite a DNA lesion.[221] Therefore, one can hypothesize that pol ζ is involved in the extension of the A*:A mismatch pair more efficiently than the correct A*:T pair for *S*-cdA. Furthermore, in the absence of pol ζ, pol κ may take part in the extension of the mismatched primer termini of *S*-cdA. In a previous *in vitro* study by You et al,[217] pol κ exhibited its ability to extend past the cyclopurines, which is consistent with what we have observed in the current study. However, in the same study, it was also reported that pol κ showed no significant contribution for cyclopurine

bypass *in vivo*. This difference may stem from the effect of the sequence context[222, 223] as well as difference between the two experimental systems.

Furthermore, MF of both *S*-cdA and *S*-cdG was not affected in pol η knockdown cells. Swanson et.al[215] suggested a role of pol η in the error-free bypass of the cyclopurine lesions *in vitro*, which was not supported by the current study or the in the cellular studies performed by You et al.[217] The difference between the *in vitro* study using purified DNA polymerases and cellular studies may arise due to several reasons. Presences of multiple TLS pols in cells, which may act collectively with accessory proteins, in the presence of all four nucleotides, modulate the fidelity and efficiency of lesion bypass. On the other hand, *in vitro* experiments with purified pols are often carried out with a single polymerase and one type of nucleotide, which could affect the nature of lesion bypass experiments.[217] Nevertheless, it is also uncertain whether the residual pol η might be sufficient for lesion bypass.

In pol ι knockdown cells, the targeted G→A substitution observed for *S*-cdG decreased by 20% compared to that in HEK293T cells. This indicates that pol ι preferentially incorporates dTMP opposite *S*-cdG, leading to erroneous replication, which was also observed by You et al[217]. Furthermore, knockdown of either REV1 or pol ζ displayed no significant effect on the *S*-cdG mutagenesis except a slight increase in the MF. The MF of *S*-cdA was not affected by siRNA-induced knockdown of pol ι or REV1 indicating that these pols may not be involved in TLS of *S*-cdA.

Based on our findings, it was evident that both *S*-cdA and *S*-cdG were highly mutagenic lesions in mammalian cells. Inability to distinguish critical roles of individual TLS pols during bypass of cyclopurines suggests that it is a complex process and more than one TLS pols are

involved in the lesion bypass. Yet, based on the current data, one can hypothesize that pol ι or pol η is involved in the nucleotide incorporation opposite the cyclopurines and pol κ or pol ζ is involved in the extension past the lesions. Therefore, these results support the current hypothesis that multiple pols take part in the lesion bypass.[224-226] Additional experiments will be needed to confirm the involvement of specific TLS pols for TLS of cyclopurines.

6 Chapter 3: Mutagenicity and Cytotoxicity of Tobacco-Specific Nitrosamine, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) - Induced O^2-Alkylthymidine DNA lesions

6.1 Abstract

The tobacco-specific nitrosamines 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) is a potent human carcinogen. Metabolic activation of NNK generates a variety of DNA adducts including O^2-Methylthymidine (O^2-Me-dT) and O^2-{4-(3-pyridyl-4-oxobut-1-yl]thymidine (O^2-POB-dT) lesions. To investigate the mutagenic and cytotoxic properties of these O^2-Alkylthymidine lesions, we have replicated single-stranded plasmids containing a site specifically incorporated O^2-Me-dT or O^2-POB-dT in human embryonic kidney 293T (HEK293T) cells. The bulkier O^2-POB-dT was found highly genotoxic in mammalian cells, showing bypass efficiency of only 26%, while O^2-Me-dT showed much lower genotoxicity as the bypass efficiency was ~55%. In mammalian cells, O^2-Me-dT and O^2-POB-dT induced 64% and 53% mutations, respectively, and the major type of mutations observed for both lesions was targeted T→A followed by T→G. A low level T→C mutation was also observed for O^2-POB-dT, which was undetected in the progeny from the O^2-Me-dT construct. siRNA induced knockdown of translesion synthesis (TLS) polymerases (pols) indicated that pol η, pol ζ, and REV1 are involved in the lesion bypass of both O^2-Me-dT and O^2-POB-dT as the TLS efficiency was decreased with knockdown of each pol. The mutation frequency (MF) of O^2-Me-dT was decreased in pol ζ and REV1 knockdown cells by 24% and 25%, respectively, while for O^2-POB-dT, it was decreased by 44% only in pol ζ knockdown cells, indicating that these TLS pols are involved in the mutagenesis of O^2-alkylthymidine lesions in mammalian cells. Therefore, this study provides important mechanistic details about how these lesions are bypassed in mammalian cells as well as their mutagenic properties.

6.2 Introduction

Tobacco products contain an array of chemicals including nicotine and several chemical carcinogens that trigger a cascade of events leading to tobacco-induced carcinogenesis.{Hecht, 1999 #518} Although nicotine is the major constituent of all tobacco products, nicotine itself is not carcinogenic;[227] however, it is responsible for activating tumor-initiation signaling pathways. Amongst the carcinogens present in tobacco, the likely candidates for causing cancers are tobacco-specific nitrosamines, including (4-(methylnitrosamine)-1-(3-pyridyl)-1-butanone (NNK)) and *N*'-nitrosonornicotine (NNN).[78, 85, 227, 228] Intermediates formed during metabolic activation of these compounds by P450 may covalently bind to DNA forming DNA adducts.[70, 71, 229, 230] NNK thus far is the only carcinogen found in tobacco products that is specific to lung carcinogenesis, irrespective of the route of administration.[93, 231]

NNK can be metabolically activated by either methylation pathway to generate methylating agents or by pyridyloxobutylation pathway to generate pyridyloxobutyalting agents. These agents react with DNA generating methyl (Me) and 4-(3-pyridyl)-4-oxobutyl (POB) adducts, respectively. The most common methylated adducts formed in the methylation pathway include 7-Me-dG and O^6-Me-dG;[97, 103, 232] however, other methylation products such as O^2-Me-dC and O^2-Me-dT have also been detected.[233, 234] While mutagenicity and carcinogenicity of NNK-induced O^6-Me-dG is well established,[231, 235, 236] the biological consequence of other methylated adducts are yet to be investigated.

Pyridyloxobutyalting pathway generates four POB adducts including O^2-POB-dT, O^2-POB-dC, O^6-POB-dG and 7-POB-dG, which were detected in NNK treated rats and mice.[104, 237, 238] A/J mice treated with a pyridyloxobutyalting agent, 4(acetoxymethyl-nitrosamino)-1-(3-pyridyl)-1-butanone (NNKOAc), a presumed electrophilic metabolite, showed that three

pyridyloxobutyalting DNA adducts, O^2-POB-dT, O^6-POB-dG and 7-POB-dG, form and persist in lung DNA at significant levels, suggesting their contribution in lung carcinogenesis.[239] Further, O^2-POB-dT was identified as the most abundant POB adduct in lung tissues of A/J female mice[240] as well as the most persistent adduct in lung and liver of male F344 rats treated with NNK.[241] NNKOAc treated CHO cells produced point mutations mainly at the AT base pairs, suggesting that O^2-POB-dT might be responsible for these mutation.[106]

In vivo mutagenesis study of O^2-Me-dT and O^2-POB-dT using a single-stranded plasmid showed that these lesions were highly mutagenic in *E. coli* cells.[242] Both survival and the mutagenicity were increased with SOS induction, indicating involvement of TLS pols for their bypass and mutagenesis. Replication of O^2-alkylthymidine lesions with varying side chains in *E. coli* showed that bypass efficiency decreased as the chain length of the alkyl group increased[243] and pol V was indispensable for T→A mutations.[242, 243] *In vitro* kinetics using *E. coli* DNA polymerase I (Kf), *Sulfolobus solfataricus* DNA polymerase IV (Dpo4), human polymerase kappa (pol κ) and yeast polymerase eta (pol η) established that O^2-Me-dT is a strong block to DNA synthesis and that dATP is preferentially incorporated opposite O^2-Me-dT.[109, 244] Furthermore, bypass of O^2-POB-dT also is inefficient by Kf- and Dpo4, and dTTP is the most preferred nucleotide opposite O^2-POB-dT, indicating low fidelity bypass.[109]

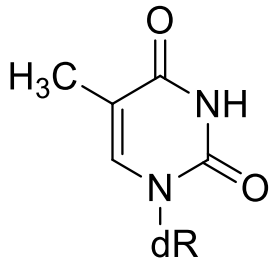
In vivo mutagenesis studies in *E. coli* and *in vitro* replication data showed that O^2-Me-dT and O^2-POB-dT are highly mutagenic and replication blocking lesions. However, their mutagenic and cytotoxic properties in mammalian cells are not well characterized. In order to investigate the replication properties of these two lesions in mammalian cells, we have replicated single-stranded pMS2 constructs containing a single O^2-Me-dT or O^2-POB-dT in human embryonic kidney 293T cells (HEK293T). Moreover, we also investigated the involvement of TLS pols,

including pol η, κ, ι, ζ and REV1 for lesion bypass and mutagenesis in HEK293T cells by siRNA-induced knockdown of these TLS pols.

Oligonucleotides

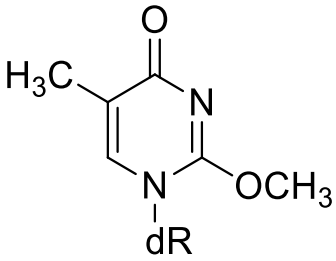
5' – GTGCGXGTTTGT – 3'

X = dT (or) O^2-Me-dT (or) O^2-POB-dT



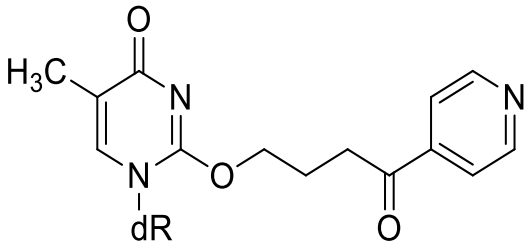
dT

(Thymidine)



O^2-Me-dT

(O^2-Methylthymidine)



O^2-POB-dT

(O^2-Pyridyloxobutylthymidine)

Figure 33: Oligonucleotides and lesions used for the study. Oligonucleotides were provided by Dr. T. Spratt (Penn State College of Medicine)

6.3 Results

6.3.1 Inhibition of replication by O^2-Me-dT and O^2-POB-dT adducts and contribution of pols η, ζ and REV1 for lesion bypass.

To investigate the mutagenic and cytotoxic properties of two O^2-alkyl-dT adducts formed by tobacco-specific nitrosamine during methylation and pyridyloxobutylation pathway, we have constructed single-stranded pMS2 plasmids containing a single O^2-Me-dT or O^2-POB-dT, which were replicated in HEK293T cells. Replication of adduct containing single-stranded vectors allow to study the effect of TLS polymerases in the absence of repair, as ss-plasmids are inefficiently repaired by DNA repair pathways.

To determine the roles of individual TLS polymerases in the bypass of these two lesions, we employed siRNA knockdown approach to suppress their expression. In agreement with previous studies, our real time and RT-PCR and Western blot analysis showed that, for each polymerase, the knockdown was at least 70% efficient.[216] Prior to replication of the lesion containing vectors, HEK293T cell were treated with siRNA for a specific TLS polymerase and allowed 48 h to reduce their expression. A non-specific siRNA was used as the negative control. After 48 h, another aliquot of siRNA was added along with the lesion containing vectors, to ensure efficient knockdown of the desired pol during replication of the ss-vector.

The relative TLS efficiency of the lesion bypass was determined by a previously reported procedure. Briefly, the lesion-containing ss-pMS2 construct was mixed in a 2:1 ratio with another modified ss-pMS2 plasmid, in which the 12mer insert contained a C instead of G, two nucleotides 5' to the lesion site and cotransfected into siRNA treated cells. This unmodified DNA containing vector was used as an internal control to determine the lesion bypass efficiency.

After transfection, the plasmids were allowed to replicate for 24 h, the progeny plasmids were recovered, transformed to *E. coli* DH 10B cells and the resulting colonies were analyzed by oligonucleotide hybridization. TLS efficiency for lesion bypass was estimated as the percentage of the colonies originating from the lesion bearing plasmid relative to the internal control plasmid containing colonies.

As shown in Figure 34 (Table 18 and Table 19) the TLS efficiencies of O^2-Me-dT and O^2-POB-dT in HEK293T cells transfected with negative control siRNA were 55% and 26% respectively, relative to 100% progeny generated from the control construct. This is similar to what we have observed in *E. coli* in that the bulkier O^2-POB-dT is more toxic. In fact, in mammalian cells it is about 2 fold more toxic than O^2-Me-dT. Furthermore, the abundance of TLS pols in mammalian cells resulted in ~12 to ~9 folds higher TLS efficiency as compared to *E. coli*.

Relative to HEK293T cells that were treated with NC-siRNA (or untreated), the TLS efficiency for O^2-Me-dT decreased by 45% - 53% in pol η, ζ and REV1 knockdown cells (Figure 34 and Table 18). Similarly, the relative TLS % of O^2-POB-dT dropped by 35% - 46% in pol η, ζ and REV1 knockdown cells (Figure 34 and Table 19). However knockdown of pol κ or ι did not show significant difference in TLS of both O^2-Me-dT and O^2-POB-dT (Figure 34, Table 18 and Table 19). These results indicate that multiple TLS pol play important roles in bypassing O^2-Me-dT and O^2-POB-dT, albeit none of the pols are indispensible for lesion bypass. However, it is uncertain whether the remaining low concentrations of enzymes in knockdown cells were able to carry out TLS across the lesion.

6.3.2 Mutational specificity of O^2-Me-dT and O^2-POB-dT in HEK293T cells

Figure 2 shows the DNA sequence analysis results of the TLS products of O^2-Me-dT and O^2-POB-dT from the HEK293T cells. In stark contrast to less than 4% mutagenicity observed in *E. coli* cells, these results showed that both O^2-Me-dT and O^2-POB-dT were highly mutagenic in mammalian cells. For O^2-Me-dT, 64% of the progeny analyzed were mutants and only ~36% progeny yielded accurate replication across the lesion in cells transfected with NC siRNA (Table 20). Similarly, O^2-POB-dT generated ~53% mutated progeny (Figure 35 and Table 21). Therefore, we conclude that, though O^2-POB-dT is a strong block to DNA replication, its replication is more accurate relative to O^2-Me-dT in human cells. Most mutations observed were targeted one-base substitution, although a low frequency of semi-targeted mutations were also observed (Figure 36, Table 20 and Table 21). Unlike in *E. coli*, the major type of mutation observed for both O^2-Me-dT and O^2-POB-dT in HEK293T cells was T → A, 56% and 47%, respectively (Figure 36, Table 20 and Table 21). dCMP incorporation opposite the O^2-Me-dT lesion was the second most prevalent mutation, which occurred at 7% frequency, however, it occurred only at ~2% for O^2-POB-dT (Figure 36, Table 20 and Table 21). A low level of T → C (1.5%) and semi-targeted (2.3%) mutations were observed for O^2-POB-dT, but T → C events were undetected for O^2-Me-dT (Figure 36, Table 20 and Table 21). Also it is important to note that no mutations were observed within the 12mer sequence of the control construct in approximately 300 colonies that were analyzed. (Results not shown)

6.3.3 Contribution of pols η, ζ and REV1 in mutagenesis of O^2-Me-dT and O^2-POB-dT

To determine the possible roles of TLS polymerase in O^2-alky-dT mutagenesis, we analyzed the TLS products obtained from the TLS pol knockdown cells. MF for O^2-Me-dT dropped in pol ζ and REV1 knockdown cells by approximately 24% - 25% relative to HEK293T cells treated with NC siRNA (Figure 35 and Table 20). In contrast, a significant 44% reduction in MF was observed in pol ζ knockdown cells for O^2-POB-dT, though no significant change was observed in REV1 knockdown cells (Figure 35 and Table 21). However, for both O^2-Me-dT and O^2-POB-dT, the decrease in MF in different pol knockdown strains was mainly due to reduced T → A mutation (Figure 36, Table 20 and Table 21). Furthermore, a modest 17% decrease in MF was noted with knockdown of pol η for O^2-POB-dT, while knockdown of neither pol κ nor pol ι had a significant effect on MF for the same (Figure 35 and Table 20). For O^2-Me-dT, intermediate MF values of 56%, 55% and 57%, respectively, were observed in pol η, κ and ι knockdown cells (Figure 35 and Table 21).

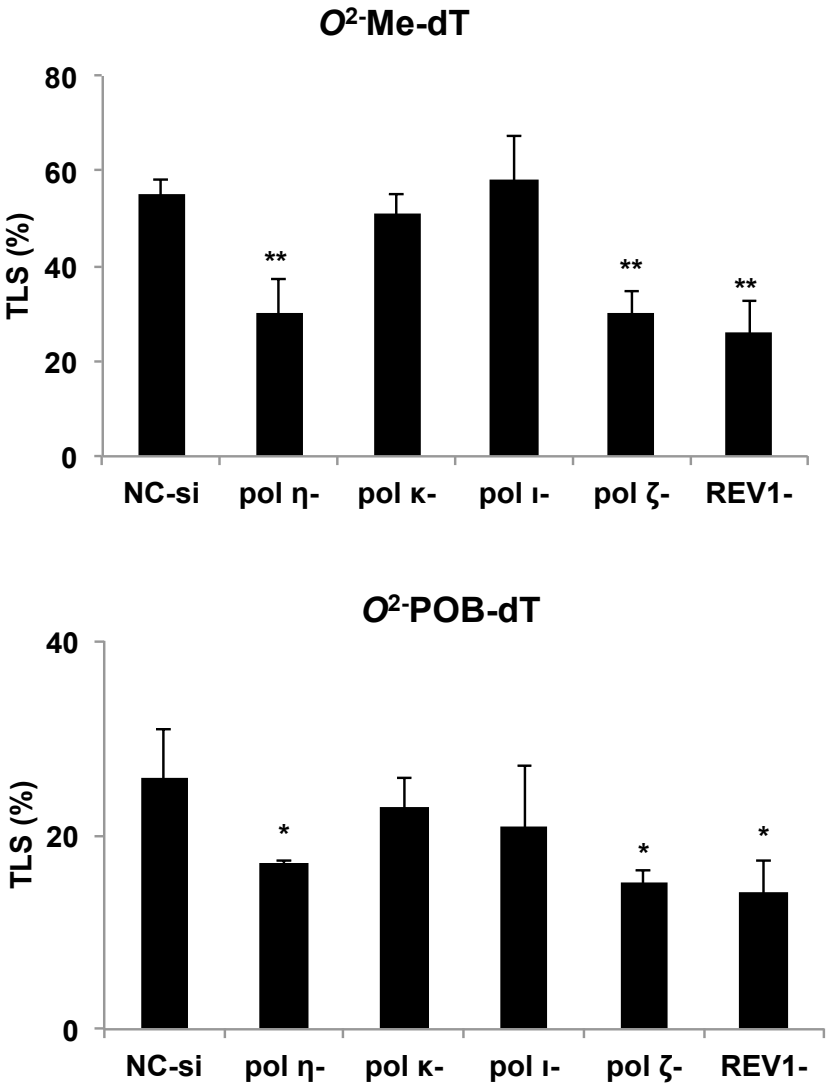


Figure 34: Effect of siRNA-induced knockdowns of TLS pols on the replicative bypass of O^2-Me-dT and O^2-POB-dT in HEK293T cells. The TLS % in various pol knockdown s was estimated using an internal control of an unmodified plasmid containing a mutation two nucleotides 5' to the lesion site. The data represent the mean and the standard deviation of results from at least 3 independent experiments ± S.D. HEK293T cells were treated with negative control (NC-si) siRNA whereas the other single knockdowns are as indicated. The statistical significance between NC-siRNA treated HEK293T and TLS pols knockdowns were calculated using a two-tailed, unpaired Student's *t* test (*$p < 0.05$: **$p < 0.01$)

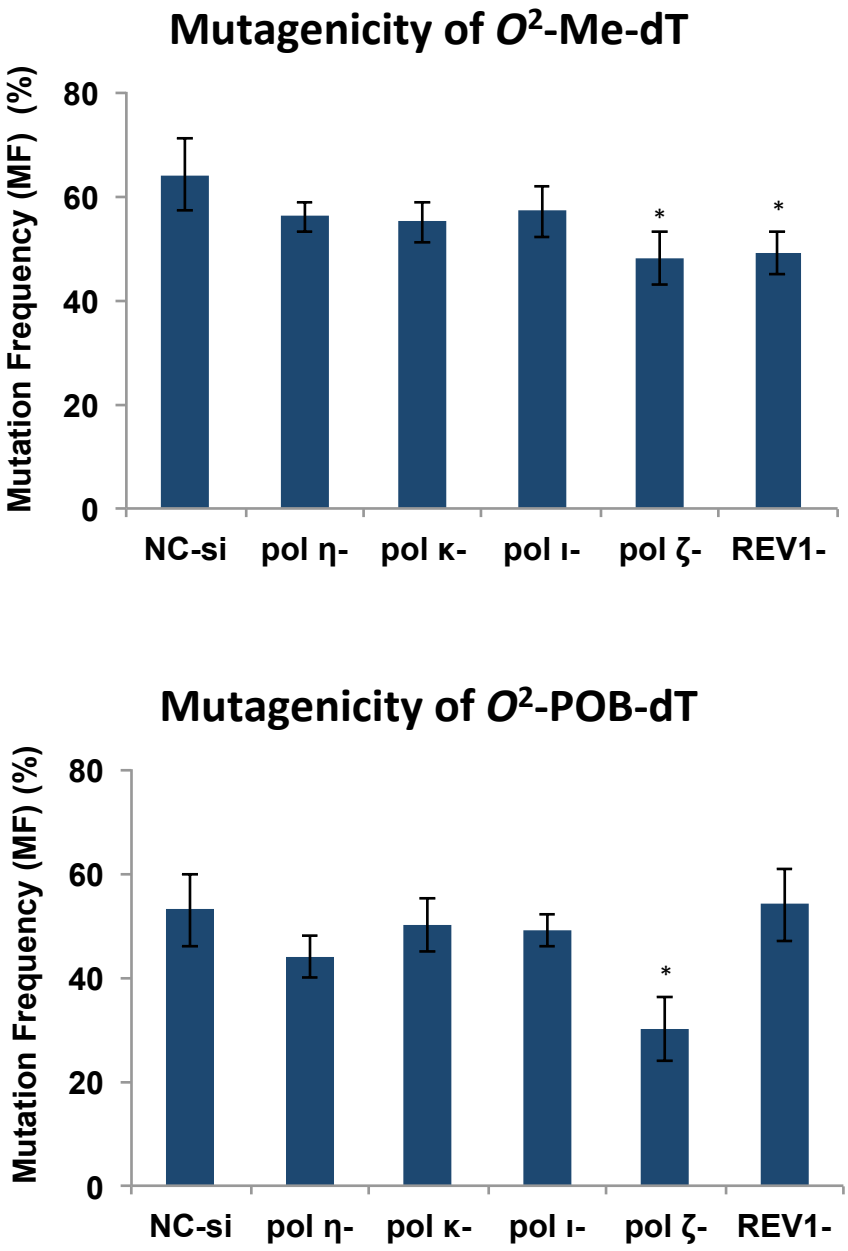


Figure 35: Total mutational frequency (MF) of O^2-Me-dT and O^2-POB-dT in HEK293T cells treated with negative control siRNA (NC-si) or siRNA for TLS pols. The data represent the average of at least three independent experiments (except for pol ι, as shown in Table 20 and 21) ±S.D. The *p* values were calculated by using two-tailed, unpaired Student's *t* test. (*$p < 0.05$).

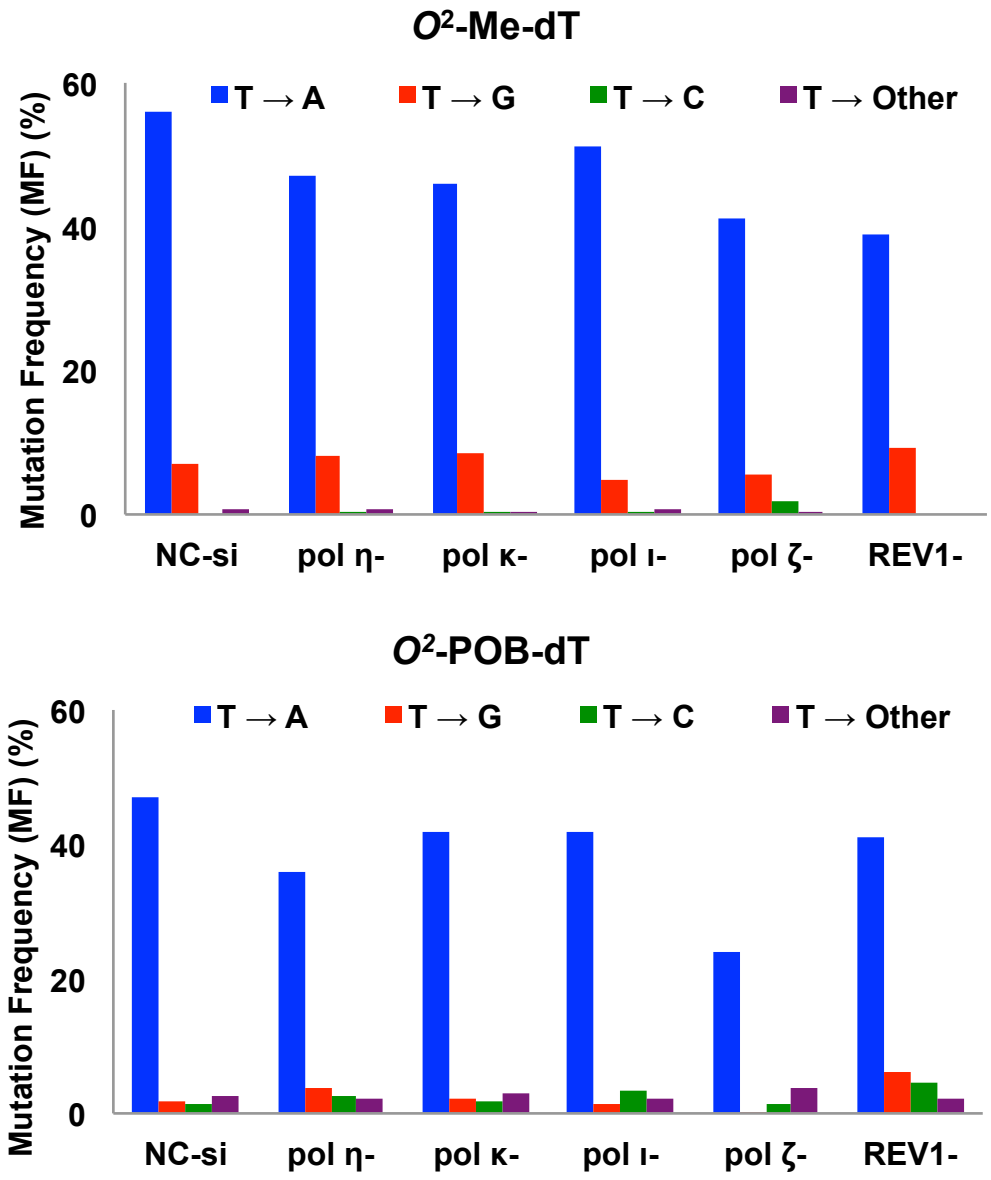


Figure 36: Types of mutations observed in HEK293T cells treated with NC siRNA (NC-si) or siRNA for individual TLS pols. The data represent the average of at least three independent experiments (except for pol ι, as shown in Table 20 and 21).

Table 18: TLS % of O^2-Me-dT in HEK293T cells in which the expression of specific TLS DNA pol was knocked-down with siRNA

Adduct	siRNA	Trial	TLS %
O^2-Me-dT	NC-si	1	56
		2	51
		3	58
		AVG	**55 (±3)**
	POLH	1	34
		2	22
		3	25
		4	38
		AVG	**30 (±7)**
	POLK	1	50
		2	56
		3	48
		AVG	**51 (±4)**
	POLI	1	62
		2	64
		3	47
		AVG	**58 (±9)**
	REV3L	1	34
		2	25
		3	31
		AVG	**30 (±5)**
	REV1	1	30
		2	30
		3	19
		AVG	**26 (±6)**

Table 19: TLS % of O^2-POB-dT in HEK293T cells in which the expression of specific TLS DNA pol was knocked-down with siRNA

Adduct	siRNA	Trial	TLS %
O^2-POB-dT	NC-si	1	25
		2	31
		3	21
		AVG	**26 (±5)**
	POLH	1	17
		2	16.7
		3	17.4
		AVG	**17 (±0.4)**
	POLK	1	26
		2	24
		3	19
		4	25
		5	22
		AVG	**23 (±3)**
	POLI	1	26
		2	22
		3	14
		AVG	**21 (±6)**
	REV3L	1	13
		2	16
		3	15
		AVG	**15 (±1)**
	REV1	1	14
		2	16
		3	10
		AVG	**14 (±3)**

Table 20: DNA sequence analysis of TLS across a O^2-Me-dT in HEK293T cells, in which the expression of defined TLS DNA polymerases was knocked-down using siRNA[†]

Lesion	siRNA	Trial	Total colonies screened	Mutation (%)	T →				
					T (%)	A (%)	G (%)	C (%)	Others (%)
O^2-Me-dT	NC-si	1	119	76 (63.9)	43 (36.1)	69 (58.0)	5 (4.2)	0 0.0	2[a] (1.7)
		2	108	62 (57.4)	46 (42.6)	53 (49.1)	8 (7.4)	0 0.0	1[b] (0.9)
		3	155	111 (71.6)	44 (28.4)	96 (61.9)	15 (9.7)	0 0.0	0 0.0
		Avg		**(64.3) ±7**	**(35.7) ±7**	**(56.3) ±7**	**(7.1) ±3**	**0.0**	**(0.9) ±1**
	POLH	1	111	63 (56.8)	48 (43.2)	53 (47.7)	10 (9.0)	0 0.0	0 0.0
		2	136	72 (52.9)	64 (47.1)	59 (43.4)	12 (8.8)	1 (0.7)	0 0.0
		3	147	85 (57.8)	62 (42.2)	73 (49.7)	9 (6.1)	0 0.0	3[c] (2.0)
		Avg		**(55.8) ±3**	**(44.2) ±3**	**(46.9) ±3**	**(8.0) ±2**	**(0.2) ±0.4**	**(0.7) ±1**
	POLK	1	119	64 (53.8)	55 (46.2)	53 (44.5)	10 (8.4)	1 (0.8)	0 0.0
		2	74	44 (59.5)	30 (40.5)	35 (47.3)	8 (10.8)	0 0.0	1[d] (1.4)
		3	148	78 (52.7)	70 (47.3)	69 (46.6)	9 (6.1)	0 0.0	0 0.0
		Avg		**(55.3) ±4**	**(44.7) ±4**	**(46.1) ±1**	**(8.4) ±2**	**(0.3) ±0.5**	**(0.5) ±0.8**
	POLI	1	66	35 (53.0)	31 (47.0)	31 (47.0)	3 (4.5)	0 0.0	1[e] (1.5)
		2	135	81 (60.0)	54 (40.0)	73 (54.1)	7 (5.2)	1 (0.7)	0 0.0
		Avg		**(56.5) ±5**	**(43.5) ±5**	**(50.6) ±5**	**(4.9) ±2**	**(0.4) ±0.4**	**(0.8) ±0.8**
	REV3L	1	133	58 (43.6)	75 (56.4)	49 (36.8)	8 (6.0)	1 (0.8)	0 0.0
		2	123	66 (53.7)	57 (46.3)	56 (45.5)	8 (6.5)	1 (0.8)	1[f] (0.8)
		3	81	38 (46.9)	43 (53.1)	32 (39.5)	3 (3.7)	3 (3.7)	0 0.0
		Avg		**(48.1) ±5**	**(51.9) ±5**	**(40.6) ±6**	**(5.4) ±1**	**(1.8) ±2**	**(0.3) ±0.5**
	REV1	1	60	32 (53.3)	28 (46.7)	25 (41.7)	7 (11.7)	0 0.0	0 0.0
		2	73	34 (46.6)	39 (53.4)	29 (39.7)	5 (6.8)	0 0.0	0 0.0
		3	95	44 (46.3)	51 (53.7)	35 (36.8)	9 (9.5)	0 0.0	0 0.0
		Avg		**(48.7) ±4**	**(51.3) ±4**	**(39.4) ±2**	**(9.3) 3**	**0.0**	**0.0**

[†]The superscript indicates one or more mutants containing mutation elsewhere shown below and the number in parenthesis shows the number of events detected.

(a) CGT* → TGT [1], GTG→_A_ [1] , (b) T*G → TT [1], (c) T*G →T_ [2]; GTG CGT → _CG CTG [1], (d) CGT* → _ _A [1], (e) GT* → TT [1], (f) GCGT* → ACG

Table 21: DNA sequence analysis of TLS across a O^2-POB-dT in HEK293T cells, in which the expression of defined TLS DNA polymerases was knocked-down using siRNA[†]

Lesion	siRNA	Trial	Total colonies screened	Mutation (%)	T →				
					T (%)	A (%)	G (%)	C (%)	Others (%)
O^2-POB-dT	NC-si	1	110	63 (57.3)	47 (42.7)	56 (50.9)	1 (0.9)	3 (2.7)	3[a] (2.7)
		2	105	47 (44.8)	58 (55.2)	42 (40.0)	3 (2.9)	1 (1.0)	1[b] (1.0)
		3	150	85 (56.7)	65 (43.3)	77 (51.3)	2 (1.3)	1 (0.7)	5[c] (3.3)
		Avg		**(52.9) ±7**	**(47.1) ±7**	**(47.4) ±6**	**(1.7) ±1**	**(1.5) ±1**	**(2.3) ±1**
	POLH	1	97	38 (39.2)	59 (60.8)	32 (33.0)	4 (4.1)	1 (1.0)	1[d] (1.0)
		2	130	60 (46.2)	70 (53.8)	45 (34.6)	6 (4.6)	4 (3.1)	5[e] (3.8)
		3	133	62 (46.6)	71 (53.4)	53 (39.8)	3 (2.3)	4 (3.0)	2[f] (1.5)
		Avg		**(44.0) ±4**	**(56.0) ±4**	**(35.8) ±4**	**(3.7) ±1**	**(2.4) ±1**	**(2.1) ±2**
	POLK	1	102	55 (53.9)	47 (46.1)	42 (41.2)	3 (2.9)	2 (2.0)	8[g] (7.8)
		2	109	48 (44.0)	61 (56.0)	45 (41.3)	2 (1.8)	1 (0.9)	0 0.0
		3	46	25 (54.3)	21 (45.7)	20 (43.5)	0 0.0	3 (6.5)	2[h] (4.3)
		4	63	29 (46.0)	34 (54.0)	27 (42.9)	2 (3.2)	0 0.0	0 0.0
		Avg		**(49.6) ±5**	**(50.5) ±6**	**(42.2) ±1**	**(2.0) ±1**	**(2.4) ±3**	**(3.0) ±4**
	POLI	1	122	62 (50.8)	60 (49.2)	53 (43.4)	1 (0.8)	5 (4.1)	3[i] (2.5)
		2	121	55 (45.5)	66 (54.5)	48 (39.7)	2 (1.7)	3 (2.5)	2[j] (1.7)
		Avg		**(48.2) ±3**	**(51.9) ±3**	**(41.6) ±2**	**(1.3) ±0.6**	**(3.3) ±1**	**(2.1) ±0.7**
	REV3L	1	88	25 (28.4)	63 (71.6)	21 (23.9)	0 0.0	2 (2.3)	2[k] (2.3)
		2	80	29 (36.3)	51 (63.8)	26 (32.5)	1 (1.3)	1 (1.3)	1[l] (1.3)
		3	38	9 (23.7)	29 (76.3)	6 (15.8)	0 0.0	0 0.0	3[m] (7.9)
		Avg		**(29.5) ±6**	**(70.6) ±6**	**(24.1) ±8**	**(0.4) ±0.8**	**(1.2) ±1**	**(3.8) ±4**
	REV1	1	60	37 (61.7)	23 (38.3)	26 (43.3)	3 (5.0)	7 (11.7)	1[n] (1.7)
		2	48	22 (45.8)	26 (54.2)	19 (39.6)	1 (2.1)	2 (4.2)	0 0.0
		3	26	14 (53.8)	12 (46.2)	11 (42.3)	3 (11.5)	0 0.0	0 0.0
		4	56	30 (53.6)	26 (46.4)	22 (39.3)	3 (5.4)	1 (1.8)	4[o] (7.1)
		Avg		**(53.7) ±6**	**(46.3) ±5**	**(41.1) ±2**	**(6.0) ±4**	**(4.4) ±5**	**(2.2) ±3**

[†]The superscript indicates one or more mutants containing mutation elsewhere shown below and the number in parenthesis shows the number of events detected.

(a) GT* → AA [1]; GT* → AT [1]; GTG CGT* G → _ _ _ _ _ _ G [1], (b) T*G → TT [1], (c) G T* →TA [3]; CG T* → GGT [1]; G T* → TG [1], (d) T*G → TT [1], (e) CG T* → TGT [1]; GTG CG T* → _ _ _ _ _ T [1]; GT*→ TT [2]; GTG CGT* GTT → Long deletion [1], (f) GT* → TA [1]; GT* → CA [1], (g) GT* → AT [2]; GTG CGT* G → _ _ _ _ _ _ _ [2]; GTG CG T* → _ _ _ CAA [1]; CG T* → TGT [1];, GT* → TA [1]; G T* → TT [1], (h) CG T* → AGA [2], (i) GTG CG T* → _ _ _ _ AG [1]; CG T* → TGT [1]; G T* → TC [1], (j) G T* → AA [1]; G T* → CA [1], (k) G T* → TT [2], (l) G T* → TT [1], (m) G T* → TA [2]; CG T* → TGA [1], (n) G T* → TA [1], (o) CGT →TGA [2]; CG T* →GAA [1]; G T* → TA [1]

6.4 Discussion and Conclusions

6.4.1 Involvement of pol η, ζ and REV1 for bypass of O^2-Me-dT and O^2-POB-dT

It is well established that NNK, a tobacco specific nitrosamine, is a potent carcinogen in human and animal models.[103, 228, 245] During metabolic activation of NNK, multiple reactive intermediates are generated that react with DNA, forming adducts, which are believed to play a role in chemical carcinogenesis.[173] In this study we investigated the replication properties of O^2-Me-dT and O^2-POB-dT, two O^2-alkyl thymidine lesions formed during metabolic activation of tobacco specific nitrosamine, in human cells. From our data (Figure 34), it is evident that TLS polymerase examined in this work including Y-family pols η, and REV1 and B-family pol ζ, have a role in lesion bypass as indicated by a decrease in TLS efficiency with knockdown of any of these TLS polymerases. The TLS efficiency for O^2-Me-dT and O^2-POB-dT in NC-siRNA treated cells were 55% and 26%, respectively Figure 34. We observed a similar pattern of cytotoxicity in SOS uninduced *E. coli,* where the bypass efficiencies were 4.5% and 2.7%, respectively, for O^2-Me-dT and O^2-POB-dT, using the same constructs.[242] These TLS efficiencies are approximately 10-12 fold lower than the bypass efficiencies in mammalian cells. Similar increase in TLS efficiency in human cells compared to *E. coli* has been observed in bypass of abasic site.[222, 246] This indicates that TLS efficiency is grater in mammalian cells than in *E. coli* cells, presumably due to an abundance of bypass pols in the former. Furthermore, we observed that the bulkier O^2-POB-dT was highly cytotoxic and bypassed less efficiently in human cells than O^2-Me-dT. In a separate study using *E. coli,* it was established that the replication bypass efficiency was decreased with longer chain lengths of O^2-alkydT,[243] which is in agreement with the results we observed for O^2-Me-dT and O^2-POB-dT in HEK293T cells.

We observed that the TLS efficiency for O^2-Me-dT in pol η, ζ and REV1 knockdown cells decreased by approximately 45% - 53% and for O^2-POB-dT by 35% - 46% (Figure 34). This indicates an involvement of multiple TLS pols for lesion bypass, which supports the current hypothesis, that more than one polymerase may be involved in multiple pol switches during TLS process.[226, 247-249] The TLS efficiency dropped by 45% and 35%, respectively, for O^2-Me-dT and O^2-POB-dT in pol η knockdown cells (Figure 34). Pol η is considered as one of the main pol involved in bypass of many lesions due to its capacious catalytic site.[133, 145, 192] Reduced TLS efficiency for both lesions in pol η knockdown cells indicates involvement of this pol for lesion bypass. This is comparable to what we observed in *E. coli*,[242] that is pol V, ortholog[250] of pol η, is essential for bypass of these lesions.

We also established that pol ζ is essential for TLS of both O^2-Me-dT and O^2-POB-dT since the TLS efficiency was reduced in pol ζ knockdown cells by 45% and 42%, respectively (Figure 34). Pol ζ has been reported to perform extension[220] after incorporation of a nucleotide opposite a lesion during TLS process. Therefore, it is conceivable that, it could function efficiently in the extension of the nucleotides that were preferentially incorporated by pol η and by pol κ or ι. Likewise, REV1 is equally important for O^2-Me-dT and O^2-POB-dT bypass, as shown by 46% - 53% reduction in TLS in REV1 knockdown cells. Based on experimental evidences it is accepted that REV1 functionally interact with all other Y family pols during TLS process.[251-253] Therefore, it is likely that REV1 acts as a structural element to form a bridge between the inserter, specifically pol η, and the extender, pol ζ, during TLS of O^2-Me-dT and O^2-POB-dT.

6.4.2 Mutagenicity of O^2-Me-dT and O^2-POB-dT and error prone bypass by pol η, ζ and REV1

Analysis of the progeny showed that both O^2-Me-dT and O^2-POB-dT are highly mutagenic lesions in HEK 293T cells and showed MF of 65% and 53%, respectively (Figure 35). This was highly different to what we have observed in SOS un-induced *E. coli*, where mutagenicity was less than 4%.[242] This may be due to higher TLS efficiency in mammalian cells could have resulted in low fidelity replication of the two lesions. Further, based on the MF data, it was evident that, although bulkier O^2-POB-dT was a highly replication-blocking lesion, it was bypassed accurately as compared to its analogue, O^2-Me-dT. It could be rationalized that, owing to the small size of the alkyl chain of the O^2-Me-dT, it is easily accommodated by the capacious catalytic site[135] of the TLS pols, thereby replicated efficiently, albeit with low fidelity.

In HEK293T cells treated with NC-siRNA, the major type of mutation induced by both O^2-Me-dT and O^2-POB-dT is T→ A transversion, at a frequency of 56.3% and 47.4%, respectively (Figure 36). The second most prevalent event for O^2-Me-dT was T→ G (7.1%) followed by low levels of semi-targeted mutations (0.9%). O^2-POB-dT generated approximately same amounts of T→ G, T→ C and other semi-targeted mutations, which ranged from 1.5% to 2.3% (Figure 36, Table 20 and Table 21). It has been reported that in NNK treated, AT → TA mutations occurred 2 to 3 times more often than AT → CG or AT → GC mutation.[254, 255] Furthermore, *in vitro* kinetics showed that dTTP incorporation opposite O^2-POB-dT was more efficient than dCTP or dGTP by both *E. coli* Kf⁻ and *Sulfolobus solfataricus* DNA pol IV (Dpo4), two model DNA pols.[109] Our *in vivo* results for O^2-POB-dT are consistent with the mutational spectra obtained in NNK-treated rodents as well as *in vitro* kinetic data. Further, *in*

vitro replication study on O^2-Me-dT, in a different sequence context, showed that it was a strong block to DNA replication by *E. coli* DNA polymerase I (Kf^-), yeast pol η, and human pol κ. Also, Kf^- and yeast pol η preferentially incorporated the correct nucleotide, dAMP, while human pol κ preferred dCMP opposite the lesion.[244] The involvement of pol κ in T → G mutation was not distinct in our study. This may be due to difference in the sequence context as well as the presence of multiple pols *in vivo* as opposed to a single pol in *in vitro*.

The targeted T → A events changed in different pol knockdown cells for both O^2-Me-dT and O^2-POB-dT, however, the pattern of the mutations remained comparable in most knockdown experiments (Figure 36). Although, the TLS efficiency was reduced in pol η knockdown cells, the MF was only slightly reduced for both lesions. It has been shown that pol η involves in error prone bypass of different DNA lesions.[192] Therefore, it could be postulated that it involved in the mis-incorporation of dTMP opposite both lesions. However, other than pol η, pol ι or κ could also be involved in mis-incorporation opposite the lesion. Nonetheless, it is challenging at this time to distinguish the role of each TLS pol in nucleotide insertion step, as residual enzyme may also be able to carry out TLS across the lesions. Evidently, further experiments with TLS pols knockout cells will be needed to distinguish the role of individual pol for lesion bypass.

In pol ζ knockdown cells, the MF for O^2-Me-dT and O^2-POB-dT decreased appreciably by 25% and 44%, respectively, in addition to the drop in TLS efficiency (Figure 35). Pol ζ is relatively tolerant of abnormal structure at the primer terminus.[256] Hence it is plausible that, it involved in the extension of T*: T miss-matched base pair more efficiently than T*: A pair, leading to error prone replication. This was more prominent across the TLS of bulkier O^2-POB-dT as compared to O^2-Me-dT. Further, the MF in REV1 knockdown cells was also reduced to a

similar extent as MF in pol ζ knockdown cells for O^2-Me-dT, indicating that REV1 plays an indispensable structural role in pol ζ dependent bypass of O^2-Me-dT (Figure 35).

In conclusion, the O^2-Me-dT and O^2-POB-dT formed by tobacco specific nitrosamine were found to be cytotoxic and highly mutagenic in HEK293T cells. Among these two, bulkier O^2-POB-dT was strong replication blocking lesion compared to O^2-Me-dT. T → A was observed as the major type of mutation, although low frequencies of T → G, T → C, and semi- targeted mutations were also observed for both lesions. siRNA knockdown of TLS pols established that pols η, ζ and REV1 involve in error prone replication bypass of O^2-Me-dT and O^2-POB-dT.

7 Summary and Future Work

In this dissertation, I have investigated the mutagenicity and translesion synthesis of abasic site, 8,5'-Cycopurine and O^2-alkylthymidine DNA lesions in *E. coli* and HEK293T cells. These studies provided important mechanistic details about lesion bypass and mutagenesis of these lesions; it also opened up several attractive avenues for future studies.

In chapter one of this dissertation, I have investigated the replicative bypass of an abasic (apurinic/apyrimididnic (AP)) site, one of the most common lesions in DNA. Since I found a major effect of the sequence context in TLS of abasic site, it will be interesting in the future to investigate *in vitro* replication properties of the abasic site in the two sequences for better understanding of the effect of the local sequences in the replication bypass. Moreover, molecular dynamics and thermodynamic investigations will also provide important information for plausible mechanism responsible for this difference and would strengthen the current observations.

In human cells, I employed the siRNA knockdown approach to constrain the expression of pol ζ and REV1 . I established that both these polymerases are vital for AP-site bypass. However, neither polymerase was indispensable, suggesting roles of additional DNA polymerases in AP-site bypass in human cells. Therefore, it will be interesting in the future to investigate the effect of other TLS polymerase as well as the effect of replicative polymerase in AP-site bypass. The latter could be achieved by replicating the lesion-containing vectors in TLS polymerase deficient mammalian cells and also in replicative polymerase inhibited (by aphidicolin) cells. These studies will provide important mechanistic details about AP-site bypass and mutagenesis. Furthermore, it will strengthen the understanding of different mechanisms for lesion bypass.

In the second chapter I have investigated the translesion synthesis of 8,5'-Cyclopurines in the *E. coli* and in HEK293T cells. Reports have been published on the synthesis of 8,5'-cyclodeoxyadenosine triphosphates and its effects on DNA replication. Due to highly cytotoxic and mutagenic nature of these cyclopurines in *E. coli*, it will be interesting to evaluate the role of 8,5'-Cyclopurine-2'-deoxynucleoside triphosphates as potential antibacterial agents. This could be achieved by using these triphosphates as incoming nucleotides for *in vitro* replication assays. If as anticipated, the DNA synthesis is inhibited by these analogues, they can be used as viable candidates for development of new anti-bacterial drugs.

Similar to *E. coli*, both *S*-cdA and *S*-cdG exhibited significant mutagenicity in HEK293T cells and I found that multiple TLS polymerase are involved in the mutagenesis of cyclopurines in human cells. However, the specific role of individual polymerases in the mutagenesis of the 8,5'-cyclopurines did not immerge from the current study. Furthermore, the involvement of replicative enzymes in the mutagenesis of these lesions is yet to be investigated. Therefore, in the future it will be interesting to study the effect of replicative enzymes in the mutagenicity of these lesions. Moreover, *in vitro* replication assays with individual polymerases in the presence of all four nucleotides will give important mechanistic information on TLS of these lesions. This work would pave the way towards better understanding of the roles cyclopurines in the etiology of human diseases.

In the third chapter I investigated the translesion synthesis across tobacco specific nitrosamine derived O^2-alkylthymidine lesions. Tobacco specific nitrosamine, 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) is one of the most important human carcinogens present in tobacco smoke and tobacco products. I established that O^2-POB-dT is more toxic than O^{2-}Me-dT in human cells and that both lesions are highly mutagenic in

mammalian cells inducing mainly T→ A substitutions. siRNA knockdown of TLS polymerases showed that pol eta, pol zeta, and REV1 have a role in lesion bypass as well as in the mutagenesis of these lesions. Therefore it will be interesting to perform *in vitro* kinetic assays to investigate the fidelity of TLS polymerase in the incorporation and the extension past these lesions. These studies will provide important mechanistic details of NNK-induced carcinogenicity.

8 References

1. WHO Cancer Fact Sheet N^{0}297. http://www.who.int/mediacentre/factsheets/fs297/en/

2. *Cancer Prevention & Early Detection Facts & Figures 2013* American Cancer Society: Atlanta, 2013.

3. Umar, A.; Dunn, B. K.; Greenwald, P., Future directions in cancer prevention. *Nature Reviews Cancer* **2012,** 12, (12), 835-848.

4. Stratton, M. R.; Campbell, P. J.; Futreal, P. A., The cancer genome. *Nature* **2009,** 458, (7239), 719-724.

5. Talbot, S. J.; Crawford, D. H., Viruses and tumours - an update. *European Journal of Cancer* **2004,** 40, (13), 1998-2005.

6. Balmain, A.; Brown, K., Oncogene Activation in Chemical Carcinogenesis. In *Advances in Cancer Research*, George, K.; Sidney, W., Eds. Academic Press: 1988; Vol. Volume 51, pp 147-182.

7. Bode, A. M.; Dong, Z., Post-translational modification of p53 in tumorigenesis. *Nat Rev Cancer* **2004,** 4, (10), 793-805.

8. Vogelstein, B.; Kinzler, K. W., Cancer genes and the pathways they control. *Nature medicine* **2004,** 10, (8), 789-799.

9. Hoeijmakers, J. H. J., Genome maintenance mechanisms for preventing cancer. *Nature* **2001,** 411, (6835), 366-374.

10. Lindahl, T., Instability and decay of the primary structure of DNA. *Nature* **1993,** 362, (6422), 709-715.

11. Lange, S. S.; Takata, K.-i.; Wood, R. D., DNA polymerases and cancer. *Nature Reviews Cancer* **2011,** 11, (2), 96-110.

12. Urcan, E.; Scherthan, H.; Styllou, M.; Haertel, U.; Hickel, R.; Reichl, F.-X., Induction of DNA double-strand breaks in primary gingival fibroblasts by exposure to dental resin composites. *Biomaterials* **2010,** 31, (8), 2010-2014.

13. Toller, I. M.; Neelsen, K. J.; Steger, M.; Hartung, M. L.; Hottiger, M. O.; Stucki, M.; Kalali, B.; Gerhard, M.; Sartori, A. A.; Lopes, M.; Mller, A., Carcinogenic bacterial pathogen Helicobacter pylori triggers DNA double-strand breaks and a DNA damage response in its host cells. *Proceedings of the National Academy of Sciences* **2011,** 108, (36), 14944-14949.

14. Finkel, T.; Holbrook, N. J., Oxidants, oxidative stress and the biology of ageing. *Nature* **2000,** 408, (6809), 239-247.

15. Shiloh, Y., ATM and related protein kinases: safeguarding genome integrity. *Nat Rev Cancer* **2003,** 3, (3), 155-168.

16. Kohen, R.; Nyska, A., Invited Review: Oxidation of Biological Systems: Oxidative Stress Phenomena, Antioxidants, Redox Reactions, and Methods for Their Quantification. *Toxicologic Pathology* **2002,** 30, (6), 620-650.

17. Gilbert, D. L., Perspective on the history of oxygen and life. In *Oxygen and living processes*, Springer: 1981; pp 1-43.

18. Harman, D., Aging: a theory based on free radical and radiation chemistry. **1955**.

19. Hussain, S. P.; Hofseth, L. J.; Harris, C. C., Radical causes of cancer. *Nature Reviews Cancer* **2003,** 3, (4), 276-285.

20. Valko, M.; Izakovic, M.; Mazur, M.; Rhodes, C. J.; Telser, J., Role of oxygen radicals in DNA damage and cancer incidence. *Molecular and cellular biochemistry* **2004,** 266, (1-2), 37-56.

21. Franco, R.; Schoneveld, O.; Georgakilas, A. G.; Panayiotidis, M. I., Oxidative stress, DNA methylation and carcinogenesis. *Cancer letters* **2008,** 266, (1), 6-11.

22. Puntarulo, S.; Cederbaum, A. I., Production of reactive oxygen species by microsomes enriched in specific human cytochrome P450 enzymes. *Free Radical Biology and Medicine* **1998,** 24, (7), 1324-1330.

23. Le Caër, S., Water radiolysis: influence of oxide surfaces on H2 production under ionizing radiation. *Water* **2011,** 3, (1), 235-253.

24. Evans, M. D.; Dizdaroglu, M.; Cooke, M. S., Oxidative DNA damage and disease: induction, repair and significance. *Mutation Research/Reviews in Mutation Research* **2004,** 567, (1), 1-61.

25. Wang, D.; Kreutzer, D. A.; Essigmann, J. M., Mutagenicity and repair of oxidative DNA damage: insights from studies using defined lesions. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **1998,** 400, (1–2), 99-115.

26. Lindahl, T.; Nyberg, B., Rate of depurination of native deoxyribonucleic acid. *Biochemistry* **1972,** 11, (19), 3610-8.

27. Loeb, L. A.; Preston, B. D., Mutagenesis by apurinic/apyrimidinic sites. *Annual review of genetics* **1986,** 20, (1), 201-230.

28. Schaaper, R. M.; Danforth, B. N.; Glickman, B. W., Mechanisms of spontaneous mutagenesis: an analysis of the spectrum of spontaneous mutation in the Escherichia coli lacI gene. *J Mol Biol* **1986,** 189, (2), 273-84.

29. Friedberg, E. C.; Aguilera, A.; Gellert, M.; Hanawalt, P. C.; Hays, J. B.; Lehmann, A. R.; Lindahl, T.; Lowndes, N.; Sarasin, A.; Wood, R. D., DNA repair: from molecular mechanism to human disease. *DNA Repair (Amst)* **2006,** 5, (8), 986-96.

30. Schaaper, R. M.; Glickman, B. W.; Loeb, L. A., Role of depurination in mutagenesis by chemical carcinogens. *Cancer Res* **1982,** 42, (9), 3480-5.

31. Loeb, L. A.; Preston, B. D.; Snow, E. T.; Schaaper, R. M., Apurinic sites as common intermediates in mutagenesis. *Basic Life Sci* **1986,** 38, 341-7.

32. Schaaper, R. M.; Koplitz, R. M.; Tkeshelashvili, L. K.; Loeb, L. A., Metal-induced lethality and mutagenesis: possible role of apurinic intermediates. *Mutat Res* **1987,** 177, (2), 179-88.

33. Lindahl, T.; Nyberg, B., Rate of depurination of native deoxyribonucleic acid. *Biochemistry* **1972,** 11, (19), 3610-3618.

34. Krokan, H.; Standal, R.; Slupphaug, G., DNA glycosylases in the base excision repair of DNA. *Biochem. J* **1997,** 325, 1-16.

35. Lindahl, T.; Karlstrom, O., Heat-induced depyrimidination of deoxyribonucleic acid in neutral solution. *Biochemistry* **1973,** 12, (25), 5151-5154.

36. Cuniasse, P.; Sowers, L.; Eritja, R.; Kaplan, B.; Goodman, M.; Cognet, J.; LeBret, M.; Guschlbauer, W.; Fazakerley, G., An abasic site in DNA. Solution conformation determined by proton NMR and molecular mechanics calculations. *Nucleic acids research* **1987,** 15, (19), 8003-8022.

37. Kalnik, M. W.; Chang, C. N.; Johnson, F.; Grollman, A. P.; Patel, D. J., NMR studies of abasic sites in DNA duplexes: deoxyadenosine stacks into the helix opposite acyclic lesions. *Biochemistry* **1989,** 28, (8), 3373-3383.

38. Goljer, I.; Kumar, S.; Bolton, P. H., Refined solution structure of a DNA heteroduplex containing an aldehydic abasic site. *Journal of Biological Chemistry* **1995,** 270, (39), 22980-22987.

39. Cuniasse, P.; Fazakerley, G. V.; Guschlbauer, W.; Kaplan, B. E.; Sowers, L. C., The abasic site as a challenge to DNA polymerase: A nuclear magnetic resonance study of G, C and T opposite a model abasic site. *Journal of Molecular Biology* **1990,** 213, (2), 303-314.

40. Morden, K. M.; Chu, Y. G.; Martin, F. H.; Tinoco Jr, I., Unpaired cytosine in the deoxyoligonucleotide duplex dCA3CA3G. cntdot. dCT6G is outside of the helix. *Biochemistry* **1983,** 22, (24), 5557-5563.

41. Vesnaver, G.; Chang, C.-N.; Eisenberg, M.; Grollman, A. P.; Breslauer, K. J., Influence of abasic and anucleosidic sites on the stability, conformation, and melting behavior of a DNA duplex: correlations of thermodynamic and structural data. *Proceedings of the National Academy of Sciences* **1989,** 86, (10), 3614-3618.

42. Chen, D. S.; Herman, T.; Demple, B., Two distinct human DNA diesterases that hydrolyze 3' blocking deoxyribose fragments from oxidized DNA. *Nucleic acids research* **1991,** 19, (21), 5907-5914.

43. Aspinwall, R.; Rothwell, D. G.; Roldan-Arjona, T.; Anselmino, C.; Ward, C. J.; Cheadle, J. P.; Sampson, J. R.; Lindahl, T.; Harris, P. C.; Hickson, I. D., Cloning and characterization of a functional human homolog of Escherichia coli endonuclease III. *Proceedings of the National Academy of Sciences* **1997,** 94, (1), 109-114.

44. Nilsen, L.; Forstrom, R. J.; Bjoras, M.; Alseth, I., AP endonuclease independent repair of abasic sites in Schizosaccharomyces pombe. *Nucleic acids research* **2012,** 40, (5), 2000-2009.

45. Sung, J.-S.; Demple, B., Roles of base excision repair subpathways in correcting oxidized abasic sites in DNA. *FEBS Journal* **2006,** 273, (8), 1620-1629.

46. Jaruga, P.; Birincioglu, M.; Rodriguez, H.; Dizdaroglu, M., Mass Spectrometric Assays for the Tandem Lesion 8,5'-Cyclo-2'-deoxyguanosine in Mammalian DNA. *Biochemistry* **2002,** 41, (11), 3703-3711.

47. Das, R. S.; Samaraweera, M.; Morton, M.; Gascón, J. A.; Basu, A. K., Stability of N-Glycosidic Bond of (5′ S)-8,5′ -Cyclo-2′ -deoxyguanosine. *Chemical Research in Toxicology* **2012,** 25, (11), 2451-2461.

48. Huang, H.; Das, R. S.; Basu, A. K.; Stone, M. P., Structure of (5'S)-8,5'-cyclo-2'-deoxyguanosine in DNA. *J Am Chem Soc* **2011,** 133, (50), 20357-68.

49. Kuraoka, I.; Bender, C.; Romieu, A.; Cadet, J.; Wood, R. D.; Lindahl, T., Removal of oxygen free-radical-induced 5',8-purine cyclodeoxynucleosides from DNA by the nucleotide excision-repair pathway in human cells. *Proc Natl Acad Sci U S A* **2000,** 97, (8), 3832-7.

50. Brooks, P. J., The 8,5'-cyclopurine-2'-deoxynucleosides: Candidate neurodegenerative DNA lesions in xeroderma pigmentosum, and unique probes of transcription and nucleotide excision repair. *DNA Repair* **2008,** 7, (7), 1168-1179.

51. Cooke, M. S.; Evans, M. D.; Dizdaroglu, M.; Lunec, J., Oxidative DNA damage: mechanisms, mutation, and disease. *The FASEB Journal* **2003,** 17, (10), 1195-1214.

52. Jaruga, P.; Dizdaroglu, M., 8,5'-Cyclopurine-2'-deoxynucleosides in DNA: mechanisms of formation, measurement, repair and biological effects. *DNA Repair (Amst)* **2008,** 7, (9), 1413-25.

53. Dizdaroglu, M.; Jaruga, P.; Birincioglu, M.; Rodriguez, H., Free radical-induced damage to DNA: mechanisms and measurement1,2. *Free Radical Biology and Medicine* **2002,** 32, (11), 1102-1115.

54. Brooks, P. J.; Wise, D. S.; Berry, D. A.; Kosmoski, J. V.; Smerdon, M. J.; Somers, R. L.; Mackie, H.; Spoonde, A. Y.; Ackerman, E. J.; Coleman, K.; Tarone, R. E.; Robbins, J. H., The oxidative DNA lesion 8,5'-(S)-cyclo-2'-deoxyadenosine is repaired by the nucleotide excision repair pathway and blocks gene expression in mammalian cells. *J Biol Chem* **2000,** 275, (29), 22355-62.

55. Theruvathu, J. A.; Jaruga, P.; Dizdaroglu, M.; Brooks, P., The oxidatively induced DNA lesions 8, 5'-cyclo-2'-deoxyadenosine and 8-hydroxy-2'-deoxyadenosine are strongly resistant to acid-induced hydrolysis of the glycosidic bond. *Mechanisms of ageing and development* **2007,** 128, (9), 494-502.

56. Pande, P.; Das, R. S.; Sheppard, C.; Kow, Y. W.; Basu, A. K., Repair efficiency of (5'S)-8,5'-cyclo-2'-deoxyguanosine and (5'S)-8,5'-cyclo-2'-deoxyadenosine depends on the complementary base. *DNA Repair (Amst)* **2012,** 11, (11), 926-31.

57. Kuraoka, I.; Robins, P.; Masutani, C.; Hanaoka, F.; Gasparutto, D.; Cadet, J.; Wood, R. D.; Lindahl, T., Oxygen free radical damage to DNA. Translesion synthesis by human DNA polymerase eta and resistance to exonuclease action at cyclopurine deoxynucleoside residues. *Journal of Biological Chemistry* **2001,** 276, (52), 49283-49288.

58. Marietta, C.; Gulam, H.; Brooks, P., A single 8, 5'-cyclo-2'-deoxyadenosine lesion in a TATA box prevents binding of the TATA binding protein and strongly reduces transcription in vivo. *DNA Repair* **2002,** 1, (11), 967-975.

59. Marietta, C.; Brooks, P. J., Transcriptional bypass of bulky DNA lesions causes new mutant RNA transcripts in human cells. *EMBO reports* **2007,** 8, (4), 388-393.

60. Jaruga, P.; Xiao, Y.; Vartanian, V.; Lloyd, R. S.; Dizdaroglu, M., Evidence for the Involvement of DNA Repair Enzyme NEIL1 in Nucleotide Excision Repair of (5'*R*)- and (5'*S*)-8,5'-Cyclo-2'-deoxyadenosines. *Biochemistry* **2010,** 49, (6), 1053-1055.

61. Jaruga, P.; Dizdaroglu, M., 8, 5′ -Cyclopurine-2′ -deoxynucleosides in DNA: mechanisms of formation, measurement, repair and biological effects. *DNA Repair* **2008,** 7, (9), 1413-1425.

62. D'Errico, M.; Parlanti, E.; Teson, M.; Degan, P.; Lemma, T.; Calcagnile, A.; Iavarone, I.; Jaruga, P.; Ropolo, M.; Pedrini, A. M.; Orioli, D.; Frosina, G.; Zambruno, G.; Dizdaroglu, M.; Stefanini, M.; Dogliotti, E., The role of CSA in the response to oxidative DNA damage in human cells. *Oncogene* **2007,** 26, (30), 4336-4343.

63. de Waard, H.; de Wit, J.; Gorgels, T. G. M. F.; van den Aardweg, G.; Andressoo, J.-O.; Vermeij, M.; van Steeg, H.; Hoeijmakers, J. H. J.; van der Horst, G. T. J., Cell type-specific hypersensitivity to oxidative damage in CSB and XPA mice. *DNA Repair* **2003,** 2, (1), 13-25.

64. Kirkali, G. l.; de Souza-Pinto, N. C.; Jaruga, P.; Bohr, V. A.; Dizdaroglu, M., Accumulation of (5'-S)-8,5'-cyclo-2'-deoxyadenosine in organs of Cockayne syndrome complementation group B gene knockout mice. *DNA Repair* **2009,** 8, (2), 274-278.

65. Rodriguez, H.; Jaruga, P.; Leber, D.; Nyaga, S. G.; Evans, M. K.; Dizdaroglu, M., Lymphoblasts of Women with BRCA1 Mutations Are Deficient in Cellular Repair of 8,5,Äò-Cyclopurine-2'-deoxynucleosides and 8-Hydroxy-2,Äò-deoxyguanosine,Ä†. *Biochemistry* **2007,** 46, (9), 2488-2496.

66. Nyaga, S. G.; Jaruga, P.; Lohani, A.; Dizdaroglu, M.; Evans, M. K., Accumulation of oxidatively induced DNA damage in human breast cancer cell lines following treatment with hydrogen peroxide. *Cell Cycle* **2007,** 6, (12), 1472-1478.

67. Anderson, K. M.; Jaruga, P.; Ramsey, C. R.; Gilman, N. K.; Green, V. M.; Rostad, S. W.; Emerman, J. T.; Dizdaroglu, M.; Malins, D. C., Structural Alterations in Breast Stromal and Epithelial DNA: The Influence of 8,5-cyclo-2-Deoxyadenosine. *Cell Cycle* **2006,** 5, (11), 1240-1244.

68. Wang, J.; Clauson, C. L.; Robbins, P. D.; Niedernhofer, L. J.; Wang, Y., The oxidative DNA lesions 8,5'-cyclopurines accumulate with aging in a tissue-specific manner. *Aging cell* **2012,** 11, (4), 714-716.

69. Organization, W. H., The World health report: 2002: Reducing the risks, promoting healthy life. **2002**.

70. Xue, J.; Yang, S.; Seng, S., Mechanisms of Cancer Induction by Tobacco-Specific NNK and NNN. *Cancers* **2014,** 6, (2), 1138-1156.

71. Hang, B., Formation and repair of tobacco carcinogen-derived bulky DNA adducts. *Journal of nucleic acids* **2010,** 2010.

72. WHO report on the global tobacco epidemic, 2008. In 2008.

73. Hecht, S. S.; Szabo, E., Fifty Years of Tobacco Carcinogenesis Research: From Mechanisms to Early Detection and Prevention of Lung Cancer. *Cancer prevention research (Philadelphia, Pa.)* **2014,** 7, (1), 1-8.

74. Health, U. D. o.; Services, H., Preventing tobacco use among youth and young adults: A report of the Surgeon General. *Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health* **2012,** 2.

75. Hukkanen, J.; Jacob, P.; Benowitz, N. L., Metabolism and disposition kinetics of nicotine. *Pharmacological reviews* **2005,** 57, (1), 79-115.

76. Hecht, S. S., Lung carcinogenesis by tobacco smoke. *International Journal of Cancer* **2012,** 131, (12), 2724-2732.

77. Chen, R.-J.; Chang, L. W.; Lin, P.; Wang, Y.-J., Epigenetic Effects and Molecular Mechanisms of Tumorigenesis Induced by Cigarette Smoke: An Overview. *Journal of Oncology* **2011,** 2011, 14.

78. Health, U. D. o.; Services, H., The health consequences of smoking-50 years of progress: A report of the surgeon general. *Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health* **2014,** 17.

79. Hecht, S. S., Research Opportunities Related to Establishing Standards for Tobacco Products Under the Family Smoking Prevention and Tobacco Control Act. *Nicotine & Tobacco Research* **2012,** 14, (1), 18-28.

80. Warren, G. W.; Singh, A. K., Nicotine and lung cancer. *Journal of carcinogenesis* **2013,** 12, (1), 1.

81. Tutka, P.; Mosiewicz, J.; Wielosz, M., Pharmacokinetics and metabolism of nicotine. *Pharmacol Rep* **2005,** 57, (2), 143-153.

82. Pfeifer, G. P.; Denissenko, M. F.; Olivier, M.; Tretyakova, N.; Hecht, S. S.; Hainaut, P., Tobacco smoke carcinogens, DNA damage and p 53 mutations in smoking-associated cancers. *Oncogene* **2002,** 21, (48), 7435-7451.

83. Hoffmann, D.; Hoffmann, I.; El-Bayoumy, K., The Less Harmful Cigarette: A Controversial Issue. A Tribute to Ernst L. Wynder. *Chemical Research in Toxicology* **2001,** 14, (7), 767-790.

84. Guengerich, F. P., Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chemical Research in Toxicology* **2001,** 14, (6), 611-650.

85. Hecht, S. S., Cigarette smoking: Cancer risks, carcinogens, and mechanisms. *Langenbeck's Archives of Surgery* **2006,** 391, (6), 603-613.

86. Tang, D.; Phillips, D. H.; Stampfer, M.; Mooney, L. A.; Hsu, Y.; Cho, S.; Tsai, W. Y.; Ma, J.; Cole, K. J.; Ni Sh, M.; Perera, F. P., Association between carcinogen-DNA adducts in white blood cells and lung cancer risk in the physicians health study. *Cancer Research* **2001,** 61, (18), 6708-6712.

87. Hecht, S. S.; Rivenson, A.; Braley, J.; DiBello, J.; Adams, J. D.; Hoffmann, D., Induction of oral cavity tumors in F344 rats by tobacco-specific nitrosamines and snuff. *Cancer Res* **1986,** 46, (8), 4162-6.

88. Belinsky, S. A.; Foley, J. F.; White, C. M.; Anderson, M. W.; Maronpot, R. R., Dose-response relationship between O6-methylguanine formation in Clara cells and induction of pulmonary neoplasia in the rat by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Cancer Res* **1990,** 50, (12), 3772-80.

89. Rivenson, A.; Hoffmann, D.; Prokopczyk, B.; Amin, S.; Hecht, S. S., Induction of lung and exocrine pancreas tumors in F344 rats by tobacco-specific and Areca-derived N-nitrosamines. *Cancer Res* **1988,** 48, (23), 6912-7.

90. Prokopczyk, B.; Cox, J. E.; Hoffmann, D.; Steven E., S. E., Identification of Tobacco-Specific Carcinogen in the Cervical Mucus of Smokers and Nonsmokers. *Journal of the National Cancer Institute* **1997,** 89, (12), 868-873.

91. Hecht, S. S., Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat Rev Cancer* **2003,** 3, (10), 733-44.

92. Wiener, D.; Doerge, D. R.; Fang, J.-L.; Upadhyaya, P.; Lazarus, P., Characterization of N-glucuronidation of the lung carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) in human liver: importance of UDP-glucuronosyltransferase 1A4. *Drug metabolism and disposition* **2004,** 32, (1), 72-79.

93. Hecht, S. S., Progress and challenges in selected areas of tobacco carcinogenesis. *Chem Res Toxicol* **2008,** 21, (1), 160-71.

94. Murphy, S. E.; Spina, D. A.; Nunes, M. G.; Pullo, D. A., Glucuronidation of 4-[(Hydroxymethyl) nitrosamino]-1-(3-pyridyl)-1-butanone, a Metabolically Activated Form of 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone, by Phenobarbital-Treated Rats. *Chemical Research in Toxicology* **1995,** 8, (5), 772-779.

95. Balbo, S.; Johnson, C. S.; Kovi, R. C.; James-Yi, S. A.; O'Sullivan, M. G.; Wang, M.; Le, C. T.; Khariwala, S. S.; Upadhyaya, P.; Hecht, S. S., Carcinogenicity and DNA adduct formation of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and enantiomers of its metabolite 4-

(methylnitrosamino)-1-(3-pyridyl)-1-butanol in F-344 rats. *Carcinogenesis* **2014,** 35, (12), 2798-806.

96. Belinsky, S. A.; White, C. M.; Boucheron, J. A.; Richardson, F. C.; Swenberg, J. A.; Anderson, M., Accumulation and persistence of DNA adducts in respiratory tissue of rats following multiple administrations of the tobacco specific carcinogen 4-(N-methyl-N-nitrosamino)-1-(3-pyridyl)-1-butanone. *Cancer Res* **1986,** 46, (3), 1280-4.

97. Peterson, L. A.; Hecht, S. S., O6-methylguanine is a critical determinant of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone tumorigenesis in A/J mouse lung. *Cancer Res* **1991,** 51, (20), 5557-64.

98. Ronai, Z. A.; Gradia, S.; Peterson, L. A.; Hecht, S. S., G to A transitions and G to T transversions in codon 12 of the Ki-ras oncogene isolated from mouse lung tumors induced by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and related DNA methylating and pyridyloxobutylating agents. *Carcinogenesis* **1993,** 14, (11), 2419-22.

99. Liu, L.; Castonguay, A.; Gerson, S. L., Lack of correlation between DNA methylation and hepatocarcinogenesis in rats and hamsters treated with 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Carcinogenesis* **1992,** 13, (11), 2137-2140.

100. Plosky, B.; Samson, L.; Engelward, B. P.; Gold, B.; Schlaen, B.; Millas, T.; Magnotti, M.; Schor, J.; Scicchitano, D. A., Base excision repair and nucleotide excision repair contribute to the removal of N-methylpurines from active genes. *DNA Repair* **2002,** 1, (8), 683-696.

101. Carmella, S. G.; Hecht, S. S., Formation of hemoglobin adducts upon treatment of F344 rats with the tobacco-specific nitrosamines 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and N'-nitrosonornicotine. *Cancer Research* **1987,** 47, (10), 2626-2630.

102. Hecht, S. S.; Spratt, T. E.; Trushin, N., Evidence for 4-(3-pyridyl)-4-oxobutylation of DNA in F344 rats treated with the tobacco-specific nitrosamines 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and N'-nitrosonornicotine. *Carcinogenesis* **1988,** 9, (1), 161-5.

103. Peterson, L. A., Formation, repair, and genotoxic properties of bulky DNA adducts formed from tobacco-specific nitrosamines. *Journal of nucleic acids* **2010,** 2010.

104. Wang, M.; Cheng, G.; Sturla, S. J.; Shi, Y.; McIntee, E. J.; Villalta, P. W.; Upadhyaya, P.; Hecht, S. S., Identification of adducts formed by pyridyloxobutylation of deoxyguanosine and DNA by 4-(acetoxymethylnitrosamino)-1-(3-pyridyl)-1-butanone, a chemically activated form of tobacco specific carcinogens. *Chem Res Toxicol* **2003,** 16, (5), 616-26.

105. Sturla, S. J.; Scott, J.; Lao, Y.; Hecht, S. S.; Villalta, P. W., Mass spectrometric analysis of relative levels of pyridyloxobutylation adducts formed in the reaction of DNA with a chemically activated form of the tobacco-specific carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Chem Res Toxicol* **2005,** 18, (6), 1048-55.

106. Li, L.; Perdigao, J.; Pegg, A. E.; Lao, Y.; Hecht, S. S.; Lindgren, B. R.; Reardon, J. T.; Sancar, A.; Wattenberg, E. V.; Peterson, L. A., The influence of repair pathways on the cytotoxicity and mutagenicity induced by the pyridyloxobutylation pathway of tobacco-specific nitrosamines. *Chemical Research in Toxicology* **2009,** 22, (8), 1464-1472.

107. Pauly, G. T.; Peterson, L. A.; Moschel, R. C., Mutagenesis by O(6)-[4-oxo-4-(3-pyridyl)butyl]guanine in Escherichia coli and human cells. *Chem Res Toxicol* **2002,** 15, (2), 165-9.

108. Wang, L.; Spratt, T. E.; Liu, X.-K.; Hecht, S. S.; Pegg, A. E.; Peterson, L. A., Pyridyloxobutyl adduct O 6-[4-oxo-4-(3-pyridyl) butyl] guanine is present in 4-(acetoxymethylnitrosamino)-1-(3-pyridyl)-1-butanone-treated DNA and is a substrate for O 6-alkylguanine-DNA alkyltransferase. *Chemical Research in Toxicology* **1997,** 10, (5), 562-567.

109. Gowda, A. S.; Krishnegowda, G.; Suo, Z.; Amin, S.; Spratt, T. E., Low fidelity bypass of O(2)-(3-pyridyl)-4-oxobutylthymine, the most persistent bulky adduct produced by the tobacco specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone by model DNA polymerases. *Chem Res Toxicol* **2012,** 25, (6), 1195-202.

110. Rouse, J.; Jackson, S. P., Interfaces between the detection, signaling, and repair of DNA damage. *Science* **2002,** 297, (5581), 547-551.

111. Jackson, S. P.; Bartek, J., The DNA-damage response in human biology and disease. *Nature* **2009,** 461, (7267), 1071-1078.

112. Branzei, D.; Foiani, M., Regulation of DNA repair throughout the cell cycle. *Nature Reviews Molecular Cell Biology* **2008,** 9, (4), 297-308.

113. Meek, D. W., Tumour suppression by p53: a role for the DNA damage response? *Nature Reviews Cancer* **2009,** 9, (10), 714-723.

114. Lord, C. J.; Ashworth, A., The DNA damage response and cancer therapy. *Nature* **2012,** 481, (7381), 287-294.

115. Christmann, M.; Tomicic, M. T.; Roos, W. P.; Kaina, B., Mechanisms of human DNA repair: an update. *Toxicology* **2003,** 193, (1,Äi2), 3-34.

116. Kaina, B., Mechanisms and consequences of methylating agent-induced SCEs and chromosomal aberrations: a long road traveled and still a far way to go. *Cytogenetic and Genome Research* **2004,** 104, (1-4), 77-86.

117. Dinglay, S.; Trewick, S. C.; Lindahl, T.; Sedgwick, B., Defective processing of methylated single-stranded DNA by E. coli AlkB mutants. *Genes & development* **2000,** 14, (16), 2097-2105.

118. Brenerman, B. M.; Illuzzi, J. L.; Wilson, D. M., Base excision repair capacity in informing healthspan. *Carcinogenesis* **2014**, bgu225.

119. Sancar, A.; Lindsey-Boltz, L. A.; Unsal-Kacmaz, K.; Linn, S., Molecular Mechanisms Of Mammalian DNA Repair and the DNA Damage Checkpoints. *Annual review of biochemistry* **2004,** 73, (1), 39-85.

120. Hoeijmakers, J. H., Genome maintenance mechanisms for preventing cancer. *Nature* **2001,** 411, (6835), 366-374.

121. Friedberg, E. C., DNA damage and repair. *Nature* **2003,** 421, (6921), 436-440.

122. Vermeulen, W.; de Boer, J.; Citterio, E.; Van Gool, A.; van der Horst, G.; Jaspers, N.; De Laat, W.; Sijbers, A.; Van der Spek, P.; Sugasawa, K., Mammalian nucleotide excision repair and syndromes. *Biochemical Society Transactions* **1997,** 25, (1), 309.

123. Marinus, M. G., DNA Mismatch Repair. *EcoSal Plus* **2012,** 2012, 10.1128/ecosalplus.7.2.5.

124. Hsieh, P.; Yamane, K., DNA mismatch repair: Molecular mechanism, cancer, and ageing. *Mechanisms of ageing and development* **2008,** 129, (7-8), 391-407.

125. Brandsma, I.; Gent, D. C., Pathway choice in DNA double strand break repair: observations of a balancing act. *Genome Integrity* **2012,** 3, 9-9.

126. Khanna, K. K.; Jackson, S. P., DNA double-strand breaks: signaling, repair and the cancer connection. *Nature genetics* **2001,** 27, (3), 247-254.

127. Takata, M.; Sasaki, M. S.; Sonoda, E.; Morrison, C.; Hashimoto, M.; Utsumi, H.; Yamaguchi,ÄêIwai, Y.; Shinohara, A.; Takeda, S., Homologous recombination and non,Äêhomologous end,Äêjoining pathways of DNA double,Äêstrand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *The EMBO Journal* **1998,** 17, (18), 5497-5508.

128. Lieber, M. R., The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* **2010,** 79, 181-211.

129. Antoniou, A.; Pharoah, P.; Narod, S.; Risch, H. A.; Eyfjord, J. E.; Hopper, J.; Loman, N.; Olsson, H. k.; Johannsson, O.; Borg, Ö., Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies. *The American Journal of Human Genetics* **2003,** 72, (5), 1117-1130.

130. Johansson, E.; Dixon, N., Replicative DNA polymerases. *Cold Spring Harb Perspect Biol* **2013,** 5, (6).

131. Ghosal, G.; Chen, J., DNA damage tolerance: a double-edged sword guarding the genome. *Transl Cancer Res* **2013,** 2, (3), 107-129.

132. Friedberg, E. C., Suffering in silence: the tolerance of DNA damage. *Nature Reviews Molecular Cell Biology* **2005,** 6, (12), 943-953.

133. Waters , L. S.; Minesinger, B. K.; Wiltrout, M. E.; D'Souza, S.; Woodruff, R. V.; Walker, G. C., Eukaryotic Translesion Polymerases and Their Roles and Regulation in DNA Damage Tolerance. *Microbiology and Molecular Biology Reviews* **2009,** 73, (1), 134-154.

134. Prakash, S.; Johnson, R. E.; Prakash, L., Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annu Rev Biochem* **2005,** 74, 317-53.

135. Sale, J. E.; Lehmann, A. R.; Woodgate, R., Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nature Reviews Molecular Cell Biology* **2012,** 13, (3), 141-152.

136. Sale, J. E., Translesion DNA synthesis and mutagenesis in eukaryotes. *Cold Spring Harbor Perspectives in Biology* **2013,** 5, (3), a012708.

137. Guo, C.; Kosarek-Stancel, J. N.; Tang, T.-S.; Friedberg, E. C., Y-family DNA polymerases in mammalian cells. *Cellular and Molecular Life Sciences* **2009,** 66, (14), 2363-2381.

138. Andersen, P. L.; Xu, F.; Xiao, W., Eukaryotic DNA damage tolerance and translesion synthesis through covalent modifications of PCNA. *Cell Res* **2008,** 18, (1), 162-73.

139. Ohmori, H.; Friedberg, E. C.; Fuchs, R. P.; Goodman, M. F.; Hanaoka, F.; Hinkle, D.; Kunkel, T. A.; Lawrence, C. W.; Livneh, Z.; Nohmi, T.; Prakash, L.; Prakash, S.; Todo, T.; Walker, G. C.; Wang, Z.; Woodgate, R., The Y-family of DNA polymerases. *Mol Cell* **2001,** 8, (1), 7-8.

140. Prakash, S.; Johnson, R. E.; Prakash, L., Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annu. Rev. Biochem.* **2005,** 74, 317-353.

141. Friedberg, E. C.; Fischhaber, P. L.; Kisker, C., Error-Prone DNA Polymerases: Novel Structures and the Benefits of Infidelity. *Cell* **2001,** 107, (1), 9-12.

142. Goodman, M. F.; Woodgate, R., Translesion DNA Polymerases. *Cold Spring Harbor Perspectives in Biology* **2013,** 5, (10).

143. Boudsocq, F. o.; Kokoska, R. J.; Plosky, B. S.; Vaisman, A.; Ling, H.; Kunkel, T. A.; Yang, W.; Woodgate, R., Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. *Journal of Biological Chemistry* **2004,** 279, (31), 32932-32940.

144. Knobel, P. A.; Marti, T. M., Translesion DNA synthesis in the context of cancer research. *Cancer cell international* **2011,** 11, (39).

145. Matsuda, T.; Bebenek, K.; Masutani, C.; Hanaoka, F.; Kunkel, T. A., Low fidelity DNA synthesis by human DNA polymerase-[eta]. *Nature* **2000,** 404, (6781), 1011-1013.

146. Wang, Y.; Woodgate, R.; McManus, T. P.; Mead, S.; McCormick, J. J.; Maher, V. M., Evidence that in xeroderma pigmentosum variant cells, which lack DNA polymerase eta, DNA polymerase iota causes the very high frequency and unique spectrum of UV-induced mutations. *Cancer Res* **2007,** 67, (7), 3018-26.

147. Biertümpfel, C.; Zhao, Y.; Kondo, Y.; Ran-Maiques, S.; Gregory, M.; Lee, J. Y.; Masutani, C.; Lehmann, A. R.; Hanaoka, F.; Yang, W., Structure and Mechanism of Human DNA Polymerase eta. *Nature* **2010,** 465, (7301), 1044-1048.

148. Nair, D. T.; Johnson, R. E.; Prakash, S.; Prakash, L.; Aggarwal, A. K., Replication by human DNA polymerase-iota occurs by Hoogsteen base-pairing. *Nature* **2004,** 430, (6997), 377-80.

149. Petta, T. B.; Nakajima, S.; Zlatanou, A.; Despras, E.; Couve-Privat, S.; Ishchenko, A.; Sarasin, A.; Yasui, A.; Kannouche, P., Human DNA polymerase iota protects cells against oxidative stress. *EMBO J* **2008,** 27, (21), 2883-95.

150. Johnson, R. E.; Washington, M. T.; Haracska, L.; Prakash, S.; Prakash, L., Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature* **2000,** 406, (6799), 1015-9.

151. Pence, M. G.; Blans, P.; Zink, C. N.; Hollis, T.; Fishbein, J. C.; Perrino, F. W., Lesion bypass of N2-ethylguanine by human DNA polymerase iota. *J Biol Chem* **2009,** 284, (3), 1732-40.

152. Gerlach, V. L.; Aravind, L.; Gotway, G.; Schultz, R. A.; Koonin, E. V.; Friedberg, E. C., Human and mouse homologs of Escherichia coli DinB (DNA polymerase IV), members of the UmuC/DinB superfamily. *Proc Natl Acad Sci U S A* **1999,** 96, (21), 11922-7.

153. Jia, L.; Geacintov, N. E.; Broyde, S., The N-clasp of human DNA polymerase kappa promotes blockage or error-free bypass of adenine- or guanine-benzo[a]pyrenyl lesions. *Nucleic acids research* **2008,** 36, (20), 6571-6584.

154. Ogi, T.; Lehmann, A. R., The Y-family DNA polymerase Œ ∫ (pol Œ ∫) functions in mammalian nucleotide-excision repair. *Nature cell biology* **2006,** 8, (6), 640-642.

155. Zhang, Y.; Wu, X.; Rechkoblit, O.; Geacintov, N. E.; Taylor, J.-S.; Wang, Z., Response of human REV1 to different DNA damage: preferential dCMP insertion opposite the lesion. *Nucleic acids research* **2002,** 30, (7), 1630-1638.

156. Haracska, L.; Prakash, S.; Prakash, L., Yeast Rev1 protein is a G template-specific DNA polymerase. *Journal of Biological Chemistry* **2002,** 277, (18), 15546-15551.

157. Nair, D. T.; Johnson, R. E.; Prakash, L.; Prakash, S.; Aggarwal, A. K., Protein-template-directed synthesis across an acrolein-derived DNA adduct by yeast Rev1 DNA polymerase. *Structure* **2008,** 16, (2), 239-245.

158. Edmunds, C. E.; Simpson, L. J.; Sale, J. E., PCNA ubiquitination and REV1 define temporally distinct mechanisms for controlling translesion synthesis in the avian cell line DT40. *Mol Cell* **2008,** 30, (4), 519-529.

159. Johnson, R. E.; Washington, M. T.; Haracska, L.; Prakash, S.; Prakash, L., Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature* **2000,** 406, (6799), 1015-1019.

160. Kenyon, C. J.; Walker, G. C., DNA-damaging agents stimulate gene expression at specific loci in Escherichia coli. *Proc Natl Acad Sci U S A* **1980,** 77, (5), 2819-23.

161. Brotcorne-Lannoye, A.; Maenhaut-Michel, G., Role of RecA protein in untargeted UV mutagenesis of bacteriophage lambda: evidence for the requirement for the dinB gene. *Proc Natl Acad Sci U S A* **1986,** 83, (11), 3904-8.

162. Kim, S.-R.; Matsui, K.; Yamada, M.; Gruz, P.; Nohmi, T., Roles of chromosomal and episomal dinB genes encoding DNA pol IV in targeted and untargeted mutagenesis in Escherichia coli. *Molecular Genetics and Genomics* **2001,** 266, (2), 207-215.

163. Kim, S. R.; Maenhaut-Michel, G.; Yamada, M.; Yamamoto, Y.; Matsui, K.; Sofuni, T.; Nohmi, T.; Ohmori, H., Multiple pathways for SOS-induced mutagenesis in Escherichia coli: an

overexpression of dinB/dinP results in strongly enhancing mutagenesis in the absence of any exogenous treatment to damage DNA. *Proc Natl Acad Sci U S A* **1997,** 94, (25), 13792-7.

164. Fuchs, R. P.; Fujii, S., Translesion DNA Synthesis and Mutagenesis in Prokaryotes. *Cold Spring Harbor Perspectives in Biology* **2013,** 5, (12), a012682.

165. Kuban, W.; Jonczyk, P.; Gawel, D.; Malanowska, K.; Schaaper, R. M.; Fijalkowska, I. J., Role of Escherichia coli DNA Polymerase IV in In Vivo Replication Fidelity. *Journal of Bacteriology* **2004,** 186, (14), 4802-4807.

166. Kuban, W.; Banach-Orlowska, M.; Bialoskorska, M.; Lipowska, A.; Schaaper, R. M.; Jonczyk, P.; Fijalkowska, I. J., Mutator Phenotype Resulting from DNA Polymerase IV Overproduction in Escherichia coli: Preferential Mutagenesis on the Lagging Strand. *Journal of Bacteriology* **2005,** 187, (19), 6862-6866.

167. McDonald, J. P.; Vaisman, A.; Kuban, W.; Goodman, M. F.; Woodgate, R., Mechanisms employed by Escherichia coli to prevent ribonucleotide incorporation into genomic DNA by Pol V. *PLoS genetics* **2012,** 8, (11), e1003030.

168. Fuchs, R. P.; Fujii, S., Translesion DNA Synthesis and Mutagenesis in Prokaryotes. *Cold Spring Harbor Perspectives in Biology* **2013,** 5, (12).

169. Nohmi, T., Environmental stress and lesion-bypass DNA polymerases. *Annu. Rev. Microbiol.* **2006,** 60, 231-253.

170. Janion, C., Inducible SOS Response System of DNA Repair and Mutagenesis in Escherichia coli. *International Journal of Biological Sciences* **2008,** 4, (6), 338-344.

171. Kreuzer, K. N., DNA Damage Responses in Prokaryotes: Regulating Gene Expression, Modulating Growth Patterns, and Manipulating Replication Forks. *Cold Spring Harbor Perspectives in Biology* **2013,** 5, (11).

172. Hollstein, M.; Shomer, B.; Greenblatt, M.; Soussi, T.; Hovig, E.; Montesano, R.; Harris, C. C., Somatic Point Mutations in the p53 Gene of Human Tumors and Cell Lines: Updated Compilation. *Nucleic acids research* **1996,** 24, (1), 141-146.

173. Hecht, S. S., Biochemistry, Biology, and Carcinogenicity of Tobacco-Specific N-Nitrosamines. *Chemical Research in Toxicology* **1998,** 11, (6), 559-603.

174. Moriya, M., Single-stranded shuttle phagemid for mutagenesis studies in mammalian cells: 8-oxoguanine in DNA induces targeted G.C-->T.A transversions in simian kidney cells. *Proc Natl Acad Sci U S A* **1993,** 90, (3), 1122-6.

175. Hirt, B., Selective extraction of polyoma DNA from infected mouse cell cultures. *J Mol Biol* **1967,** 26, (2), 365-9.

176. Choi, J. Y.; Lim, S.; Kim, E. J.; Jo, A.; Guengerich, F. P., Translesion Synthesis across Abasic Lesions by Human B-Family and Y-Family DNA Polymerases alpha, delta, eta, I, K, and REV1. *Journal of Molecular Biology* **2010,** 404, (1), 34-44.

177. Schaaper, R. M.; Glickman, B. W.; Loeb, L. A., Mutagenesis resulting from depurination is an SOS process. *Mutat Res* **1982,** 106, (1), 1-9.

178. Loeb, L. A.; Preston, B. D., Mutagenesis by apurinic/apyrimidinic sites. *Annu Rev Genet* **1986,** 20, 201-30.

179. Strauss, B. S., The "A" rule revisited: polymerases as determinants of mutational specificity. *DNA Repair (Amst)* **2002,** 1, (2), 125-35.

180. Lawrence, C. W.; Borden, A.; Banerjee, S. K.; LeClerc, J. E., Mutation frequency and spectrum resulting from a single abasic site in a single-stranded vector. *Nucleic Acids Res* **1990,** 18, (8), 2153-7.

181. Gibbs, P. E.; Lawrence, C. W., Novel mutagenic properties of abasic sites in Saccharomyces cerevisiae. *J Mol Biol* **1995,** 251, (2), 229-36.

182. Pages, V.; Johnson, R. E.; Prakash, L.; Prakash, S., Mutational specificity and genetic control of replicative bypass of an abasic site in yeast. *Proc Natl Acad Sci U S A* **2008,** 105, (4), 1170-5.

183. Gentil, A.; Renault, G.; Madzak, C.; Margot, A.; Cabral-Neto, J. B.; Vasseur, J. J.; Rayner, B.; Imbach, J. L.; Sarasin, A., Mutagenic properties of a unique abasic site in mammalian cells. *Biochem Biophys Res Commun* **1990,** 173, (2), 704-10.

184. Gentil, A.; Cabral-Neto, J. B.; Mariage-Samson, R.; Margot, A.; Imbach, J. L.; Rayner, B.; Sarasin, A., Mutagenicity of a unique apurinic/apyrimidinic site in mammalian cells. *J Mol Biol* **1992,** 227, (4), 981-4.

185. Cabral Neto, J. B.; Cabral, R. E.; Margot, A.; Le Page, F.; Sarasin, A.; Gentil, A., Coding properties of a unique apurinic/apyrimidinic site replicated in mammalian cells. *J Mol Biol* **1994,** 240, (5), 416-20.

186. Avkin, S.; Adar, S.; Blander, G.; Livneh, Z., Quantitative measurement of translesion replication in human cells: evidence for bypass of abasic sites by a replicative DNA polymerase. *Proc Natl Acad Sci U S A* **2002,** 99, (6), 3764-9.

187. Brammer, K. W.; Jones, A. S.; Mian, A. M.; Walker, R. T., Study of the use of alkaline degradation of DNA derivative as a procedure for the determination of nucleotide distribution. *Biochim Biophys Acta* **1968,** 166, (3), 732-4.

188. Kroeger, K. M.; Goodman, M. F.; Greenberg, M. M., A comprehensive comparison of DNA replication past 2-deoxyribose and its tetrahydrofuran analog in Escherichia coli. *Nucleic Acids Res* **2004,** 32, (18), 5480-5.

189. Friedberg, E. C.; Wagner, R.; Radman, M., Specialized DNA polymerases, cellular survival, and the genesis of mutations. *Science* **2002,** 296, (5573), 1627-30.

190. Yang, W.; Woodgate, R., What a difference a decade makes: insights into translesion DNA synthesis. *Proc Natl Acad Sci U S A* **2007,** 104, (40), 15591-8.

191. Tang, M.; Pham, P.; Shen, X.; Taylor, J. S.; O'Donnell, M.; Woodgate, R.; Goodman, M. F., Roles of E. coli DNA polymerases IV and V in lesion-targeted and untargeted SOS mutagenesis. *Nature* **2000,** 404, (6781), 1014-8.

192. Zhang, Y.; Yuan, F.; Wu, X.; Rechkoblit, O.; Taylor, J. S.; Geacintov, N. E.; Wang, Z., Error-prone lesion bypass by human DNA polymerase eta. *Nucleic Acids Res* **2000,** 28, (23), 4717-24.

193. Fang, H.; Taylor, J. S., Serial analysis of mutation spectra (SAMS): a new approach for the determination of mutation spectra of site-specific DNA damage and their sequence dependence. *Nucleic Acids Res* **2008,** 36, (18), 6004-12.

194. Boudsocq, F.; Iwai, S.; Hanaoka, F.; Woodgate, R., Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4): an archaeal DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic poleta. *Nucleic Acids Res* **2001,** 29, (22), 4607-16.

195. Haracska, L.; Unk, I.; Johnson, R. E.; Phillips, B. B.; Hurwitz, J.; Prakash, L.; Prakash, S., Stimulation of DNA synthesis activity of human DNA polymerase kappa by PCNA. *Mol Cell Biol* **2002,** 22, (3), 784-91.

196. Haracska, L.; Unk, I.; Johnson, R. E.; Johansson, E.; Burgers, P. M.; Prakash, S.; Prakash, L., Roles of yeast DNA polymerases delta and zeta and of Rev1 in the bypass of abasic sites. *Genes Dev* **2001,** 15, (8), 945-54.

197. Gibbs, P. E.; McDonald, J.; Woodgate, R.; Lawrence, C. W., The relative roles in vivo of Saccharomyces cerevisiae Pol eta, Pol zeta, Rev1 protein and Pol32 in the bypass and mutation induction of an abasic site, T-T (6-4) photoadduct and T-T cis-syn cyclobutane dimer. *Genetics* **2005,** 169, (2), 575-82.

198. Banerjee, S. K.; Christensen, R. B.; Lawrence, C. W.; LeClerc, J. E., Frequency and spectrum of mutations produced by a single cis-syn thymine-thymine cyclobutane dimer in a single-stranded vector. *Proc Natl Acad Sci U S A* **1988,** 85, (21), 8141-5.

199. Moriya, M.; Ou, C.; Bodepudi, V.; Johnson, F.; Takeshita, M.; Grollman, A. P., Site-specific mutagenesis using a gapped duplex vector: a study of translesion synthesis past 8-oxodeoxyguanosine in E. coli. *Mutat Res* **1991,** 254, (3), 281-8.

200. Pandya, G. A.; Moriya, M., 1,N6-ethenodeoxyadenosine, a DNA adduct highly mutagenic in mammalian cells. *Biochemistry* **1996,** 35, (35), 11487-92.

201. Raychaudhury, P.; Basu, A. K., Genetic requirement for mutagenesis of the G[8,5-Me]T cross-link in Escherichia coli: DNA polymerases IV and V compete for error-prone bypass. *Biochemistry* **2011,** 50, (12), 2330-8.

202. Pednekar, V.; Weerasooriya, S.; Jasti, V. P.; Basu, A. K., Mutagenicity and Genotoxicity of (5 ' S)-8,5 '-Cyclo-2 '-deoxyadenosine in Escherichia coli and Replication of (5 ' S)-8,5 '-Cyclopurine-2 '-deoxynucleosides in Vitro by DNA Polymerase IV, Exo-Free Klenow Fragment, and Dpo4. *Chemical Research in Toxicology* **2014,** 27, (2), 200-210.

203. Haracska, L.; Washington, M. T.; Prakash, S.; Prakash, L., Inefficient bypass of an abasic site by DNA polymerase eta. *Journal of Biological Chemistry* **2001,** 276, (9), 6861-6866.

204. Wojtaszek, J.; Lee, C. J.; D'Souza, S.; Minesinger, B.; Kim, H.; D'Andrea, A. D.; Walker, G. C.; Zhou, P., Structural Basis of Rev1-mediated Assembly of a Quaternary Vertebrate Translesion Polymerase Complex Consisting of Rev1, Heterodimeric Polymerase (Pol) zeta, and Pol kappa. *Journal of Biological Chemistry* **2012,** 287, (40), 33836-33846.

205. Malina, J.; Brabec, V., Thermodynamic Impact of Abasic Sites on Simulated Translesion DNA Synthesis. *Chemistry A European Journal* **2014,** 20, 7566-7570.

206. Maor-Shoshani, A.; Hayashi, K.; Ohmori, H.; Livneh, Z., Analysis of translesion replication across an abasic site by DNA polymerase IV of Escherichia coli. *DNA Repair* **2003,** 2, (11), 1227-1238.

207. Reuven, N. B.; Arad, G.; Maor-Shoshani, A.; Livneh, Z., The mutagenesis protein UmuC is a DNA polymerase activated by UmuD ', RecA, and SSB and is specialized for translesion replication. *Journal of Biological Chemistry* **1999,** 274, (45), 31763-31766.

208. Avkin, S.; Adar, S.; Blander, G.; Livneh, Z., Quantitative measurement of translesion replication in human cells: Evidence for bypass of abasic sites by a replicative DNA polymerase. *Proceedings of the National Academy of Sciences* **2002,** 99, (6), 3764-3769.

209. Blanca, G.; Villani, G.; Shevelev, I.; Ramadan, K.; Spadari, S.; Hubscher, U.; Maga, G., Human DNA polymerases lambda and beta show different efficiencies of translesion DNA synthesis past abasic sites and alternative mechanisms for frameshift generation. *Biochemistry* **2004,** 43, (36), 11605-11615.

210. Kalam, M. A.; Basu, A. K., Mutagenesis of 8-oxoguanine adjacent to an abasic site in simian kidney cells: tandem mutations and enhancement of G-->T transversions. *Chem Res Toxicol* **2005,** 18, (8), 1187-92.

211. Kirkali, G.; de Souza-Pinto, N. C.; Jaruga, P.; Bohr, V. A.; Dizdaroglu, M., Accumulation of (5'S)-8,5'-cyclo-2'-deoxyadenosine in organs of Cockayne syndrome complementation group B gene knockout mice. *DNA Repair (Amst)* **2009,** 8, (2), 274-8.

212. Jasti, V. P.; Das, R. S.; Hilton, B. A.; Weerasooriya, S.; Zou, Y.; Basu, A. K., (5S)-8,5-Cyclo-2-deoxyguanosine Is a Strong Block to Replication, a Potent pol V-Dependent Mutagenic Lesion, and Is Inefficiently Repaired in Escherichia coli. *Biochemistry* **2011,** 50, (19), 3862-3865.

213. Yuan, B.; Wang, J.; Cao, H.; Sun, R.; Wang, Y., High-throughput analysis of the mutagenic and cytotoxic properties of DNA lesions by next-generation sequencing. *Nucleic acids research* **2011,** 39, (14), 5945-5954.

214. Rangarajan, S.; Woodgate, R.; Goodman, M. F., A phenotype for enigmatic DNA polymerase II: a pivotal role for pol II in replication restart in UV-irradiated Escherichia coli. *Proc Natl Acad Sci U S A* **1999,** 96, (16), 9224-9.

215. Swanson, A. L.; Wang, J.; Wang, Y., Accurate and Efficient Bypass of 8,5′ -Cyclopurine-2′ -Deoxynucleosides by Human and Yeast DNA Polymerase η. *Chemical Research in Toxicology* **2012,** 25, (8), 1682-1691.

216. Pande, P.; Malik, C. K.; Bose, A.; Jasti, V. P.; Basu, A. K., Mutational Analysis of the C8-Guanine Adduct of the Environmental Carcinogen 3-Nitrobenzanthrone in Human Cells: Critical Roles of DNA Polymerases Œ Σ and Œ ∫ and Rev1 in Error-Prone Translesion Synthesis. *Biochemistry* **2014,** 53, (32), 5323-5331.

217. You, C.; Swanson, A. L.; Dai, X.; Yuan, B.; Wang, J.; Wang, Y., Translesion Synthesis of 8,5′ -Cyclopurine-2′ -deoxynucleosides by DNA Polymerases η, ι, and ζ. *Journal of Biological Chemistry* **2013,** 288, (40), 28548-28556.

218. Washington, M. T.; Carlson, K. D.; Freudenthal, B. D.; Pryor, J. M., Variations on a theme: eukaryotic Y-family DNA polymerases. *Biochimica et biophysica acta* **2010,** 1804, (5), 1113-1123.

219. Washington, M. T.; Minko, I. G.; Johnson, R. E.; Wolfle, W. T.; Harris, T. M.; Lloyd, R. S.; Prakash, S.; Prakash, L., Efficient and error-free replication past a minor-groove DNA adduct by the sequential action of human DNA polymerases iota and kappa. *Mol Cell Biol* **2004,** 24, (13), 5687-93.

220. Shachar, S.; Ziv, O.; Avkin, S.; Adar, S.; Wittschieben, J.; Reißner, T.; Chaney, S.; Friedberg, E. C.; Wang, Z.; Carell, T.; Geacintov, N.; Livneh, Z., Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals. *The EMBO Journal* **2009,** 28, (4), 383-393.

221. Boiteux, S.; Guillet, M., Abasic sites in DNA: repair and biological consequences in Saccharomyces cerevisiae. *DNA Repair* **2004,** 3, (1), 1-12.

222. Weerasooriya, S.; Jasti, V. P.; Basu, A. K., Replicative Bypass of Abasic Site in *Escherichia coli* and Human Cells: Similarities and Differences. *PLoS ONE* **2014,** 9, (9), e107915.

223. Jain, V.; Vaidyanathan, V. G.; Patnaik, S.; Gopal, S.; Cho, B. P., Conformational Insights into the lesion and sequence effects for arylamine-Induced translesion DNA synthesis: 19F NMR, surface plasmon resonance, and primer kinetic studies. *Biochemistry* **2014,** 53, (24), 4059-4071.

224. Bresson, A.; Fuchs, R. P. P., Lesion bypass in yeast cells: Pol η participates in a multi-DNA polymerase process. *The EMBO Journal* **2002,** 21, (14), 3881-3887.

225. Prakash, S.; Prakash, L., Translesion DNA synthesis in eukaryotes: a one- or two-polymerase affair. *Genes Dev* **2002,** 16, (15), 1872-83.

226. Livneh, Z.; Ziv, O.; Shachar, S., Multiple two-polymerase mechanisms in mammalian translesion DNA synthesis. *Cell Cycle* **2010,** 9, (4), 729-35.

227. Hoffmann, D.; Hoffmann, I., The changing cigarette, 1950-1995. *J Toxicol Environ Health* **1997,** 50, (4), 307-64.

228. Hecht, S. S., Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nature Reviews Cancer* **2003,** 3, (10), 733-744.

229. Hecht, S. S., Biochemistry, Biology, and Carcinogenicity of Tobacco-Specific N-Nitrosamines†. *Chemical Research in Toxicology* **1998,** 11, (6), 559-603.

230. Ding, X.; Kaminsky, L. S., Human Extrahepatic Cytochromes P450: Function in Xenobiotic Metabolism and Tissue-Selective Chemical Toxicity in the Respiratory and Gastrointestinal Tracts*. *Annual review of pharmacology and toxicology* **2003,** 43, (1), 149-173.

231. Hecht, S. S., Biochemistry, biology, and carcinogenicity of tobacco-specific N-nitrosamines. *Chem Res Toxicol* **1998,** 11, (6), 559-603.

232. Hecht, S. S.; Trushin, N.; Castonguay, A.; Rivenson, A., Comparative tumorigenicity and DNA methylation in F344 rats by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and N-nitrosodimethylamine. *Cancer Research* **1986,** 46, (2), 498-502.

233. Shrivastav, N.; Li, D.; Essigmann, J. M., Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation. *Carcinogenesis* **2010,** 31, (1), 59-70.

234. Den Engelse, L.; Menkveld, G. J.; De Brij, R. J.; Tates, A. D., Formation and stability of alkylated pyrimidines and purines (including imidazole ring-opened 7-alkylguanine) and alkylphosphotriesters in liver DNA of adult rats treated with ethylnitrosourea or dimethylnitrosamine. *Carcinogenesis* **1986,** 7, (3), 393-403.

235. Morse, M. A.; Amin, S. G.; Hecht, S. S.; Chung, F.-L., Effects of aromatic isothiocyanates on tumorigenicity, O6-methylguanine formation, and metabolism of the tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in A/J mouse lung. *Cancer Research* **1989,** 49, (11), 2894-2897.

236. Peterson, L. A.; Hecht, S. S., O6-methylguanine is a critical determinant of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone tumorigenesis in A/J mouse lung. *Cancer Research* **1991,** 51, (20), 5557-5564.

237. Wang, L.; Spratt, T. E.; Liu, X. K.; Hecht, S. S.; Pegg, A. E.; Peterson, L. A., Pyridyloxobutyl adduct O6-[4-oxo-4-(3-pyridyl)butyl]guanine is present in 4-(acetoxymethylnitrosamino)-1-(3-pyridyl)-1-butanone-treated DNA and is a substrate for O6-alkylguanine-DNA alkyltransferase. *Chem Res Toxicol* **1997,** 10, (5), 562-7.

238. Hecht, S. S.; Villalta, P. W.; Sturla, S. J.; Cheng, G.; Yu, N.; Upadhyaya, P.; Wang, M., Identification of O2-substituted pyrimidine adducts formed in reactions of 4-(acetoxymethylnitrosamino)- 1-(3-pyridyl)-1-butanone and 4-(acetoxymethylnitros- amino)-1-(3-pyridyl)-1-butanol with DNA. *Chem Res Toxicol* **2004,** 17, (5), 588-97.

239. Urban, A. M.; Upadhyaya, P.; Cao, Q.; Peterson, L. A., Formation and Repair of Pyridyloxobutyl DNA Adducts and Their Relationship to Tumor Yield in A/J Mice. *Chemical Research in Toxicology* **2012,** 25, (10), 2167-2178.

240. Balbo, S.; Upadhyaya, P.; Villalta, P. W.; Qian, X.; Kassie, F., DNA adducts in aldehyde dehydrogenase-positive lung stem cells of A/J mice treated with the tobacco specific lung carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK). *Chem Res Toxicol* **2013,** 26, (4), 511-3.

241. Lao, Y.; Yu, N.; Kassie, F.; Villalta, P. W.; Hecht, S. S., Formation and accumulation of pyridyloxobutyl DNA adducts in F344 rats chronically treated with 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and enantiomers of its metabolite, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol. *Chem Res Toxicol* **2007,** 20, (2), 235-45.

242. Jasti, V. P.; Spratt, T. E.; Basu, A. K., Tobacco-Specific Nitrosamine-Derived O2-Alkylthymidines Are Potent Mutagenic Lesions in SOS-Induced Escherichia coli. *Chemical Research in Toxicology* **2011,** 24, (11), 1833-1835.

243. Zhai, Q.; Wang, P.; Cai, Q.; Wang, Y., Syntheses and characterizations of the in vivo replicative bypass and mutagenic properties of the minor-groove O2-alkylthymidine lesions. *Nucleic acids research* **2014**.

244. Andersen, N.; Wang, J.; Wang, P.; Jiang, Y.; Wang, Y., In-Vitro Replication Studies on O2-Methylthymidine and O4-Methylthymidine. *Chemical Research in Toxicology* **2012,** 25, (11), 2523-2531.

245. Hecht, S. S., Tobacco smoke carcinogens and lung cancer. *Journal of the National Cancer Institute* **1999,** 91, (14), 1194-1210.

246. Avkin, S.; Goldsmith, M.; Velasco-Miguel, S.; Geacintov, N.; Friedberg, E. C.; Livneh, Z., Quantitative Analysis of Translesion DNA Synthesis across a Benzo[a]pyrene-Guanine Adduct in Mammalian Cells: The Role of DNA Polymerase Kappa. *Journal of Biological Chemistry* **2004,** 279, (51), 53298-53305.

247. Lehmann, A. R.; Niimi, A.; Ogi, T.; Brown, S.; Sabbioneda, S.; Wing, J. F.; Kannouche, P. L.; Green, C. M., Translesion synthesis: Y-family polymerases and the polymerase switch. *DNA Repair* **2007,** 6, (7), 891-899.

248. Friedberg, E. C.; Lehmann, A. R.; Fuchs, R. P. P., Trading Places: How Do DNA Polymerases Switch during Translesion DNA Synthesis? *Mol Cell* **2005,** 18, (5), 499-505.

249. Plosky, B. S.; Woodgate, R., Switching from high-fidelity replicases to low-fidelity lesion-bypass polymerases. *Current Opinion in Genetics and Development* **2004,** 14, (2), 113-119.

250. Lee, C. H.; Chandani, S.; Loechler, E. L., Homology modeling of four Y-family, lesion-bypass DNA polymerases: the case that E. coli Pol IV and human Pol kappa are orthologs, and E. coli Pol V and human Pol eta are orthologs. *J Mol Graph Model* **2006,** 25, (1), 87-102.

251. Pustovalova, Y.; Bezsonova, I.; Korzhnev, D. M., The C-terminal domain of human Rev1 contains independent binding sites for DNA polymerase eta and Rev7 subunit of polymerase zeta. *FEBS Lett* **2012,** 586, (19), 3051-6.

252. Guo, C.; Fischhaber, P. L.; Luk‚ÄêPaszyc, M. J.; Masuda, Y.; Zhou, J.; Kamiya, K.; Kisker, C.; Friedberg, E. C., Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. *The EMBO Journal* **2003,** 22, (24), 6621-6630.

253. Ohashi, E.; Murakumo, Y.; Kanjo, N.; Akagi, J.-i.; Masutani, C.; Hanaoka, F.; Ohmori, H., Interaction of hREV1 with three human Y-family DNA polymerases. *Genes to Cells* **2004,** 9, (6), 523-531.

254. Hashimoto, K.; Ohsawa, K.; Kimura, M., Mutations induced by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) in the lacZ and cII genes of Muta Mouse. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **2004,** 560, (2), 119-131.

255. Sandercock, L. E.; Hahn, J. N.; Li, L.; Luchman, H. A.; Giesbrecht, J. L.; Peterson, L. A.; Jirik, F. R., Mgmt deficiency alters the in vivo mutational spectrum of tissues exposed to the tobacco carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK). *Carcinogenesis* **2008,** 29, (4), 866-874.

256. Lawrence, C. W.; Maher, V. M., *Mutagenesis in eukaryotes dependent on DNA polymerase zeta and Rev1p*. 2001; Vol. 356, p 41-46.