GLYCOSYLASE RECOGNITION OF DAMAGED DNA: BALANCING REPAIR AND MUTAGENESIS

by

Derek Lyons

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in The University of Michigan
2012

Doctoral Committee:

Associate Professor Patrick O'Brien, Chair
Professor Nils G. Walter
Associate Professor Bruce A. Palfey
Assistant Professor Oleg V. Tsodikov

DEDICATION

To love, passion, and life-long learning - to my wife who inspires these qualities everyday

ACKNOWLEDEMENTS

Graduate school has been an enlightening adventure. Never have I been influenced and inspired to such an extent in my life. First, I thank Pat for providing unwavering commitment to mentoring a young scientist. His dedication to teaching was instrumental in my development as a scientist. Pat gave me the freedom to make mistakes, but was always available for guidance. His support ensured I flourished from every flounder. It has been extremely rewarding to see the O'Brien laboratory grow over these years. I've always felt like part of a team – a team that I could contribute to, and likewise benefit from. Michael B. was a tremendous asset, always a steadfast sounding board for ideas and a constant companion in the laboratory. I thank my committee members for support and guidance. It is especially invigorating to be surrounded with great minds, scientists who see the big implications as well as understand the details.

The Department of Biological Chemistry has proven itself as an incredible environment for a young scientist. There was always an electricity in the air that made it exciting to be there. Likewise, I loved being a part of the Chemical Biology doctoral program, such an innovative and exciting program. Innovation is a common theme of the University of Michigan and Ann Arbor in general. I found the University of Michigan was a perfect place to find myself. Surrounded by so many like-minded people, I have never had the chance to grow and develop as a person in such a short time. The people I met here in Michigan, my friends, are some of the most genuine people I have ever

known and I will never forget the great times we had…the Big House. My family was a welcomed source of endless encouragement. My greatest asset of my time in graduate school was my wife Hannah. She provided more than support and motivation. Her toughness and determination taught me to focus, her infinite understanding and limitless patience allowed me to get the most out of my time in graduate school. Cheers to our first of many great adventures.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

Genomic Stability Requires DNA Repair

The human genome contains 12.6 billion deoxynucleotides. Separated into 23 chromosome pairs, the largest chromosome in humans is a polymer nearly half a billion deoxynucleotides long. Maintaining the integrity of a such a large biomolecule poses a significant challenge to the cell, as the genome is constantly bombarded with reactive chemicals (1). Chemical reactions with endogenenous and exogenous sources lead to the modification of up to 100,000 nucleotides a day (2). These covalent modifications can interfere with fundamental processes of cell function such as transcription and replication (3).

Many types of DNA damage are mutagenic, promoting permanent alterations to the genome (4). Mutations lead to diversity among our species, and genetic diversity can provide circumstantial advantages (5). Conversely, mutations are often detrimental to cell function and are intimately associated with disease (6). Remarkably, the human cell restricts the rate of mutations to a mere $1.4x10^{-10}$ mutations/basepair/generation (7).

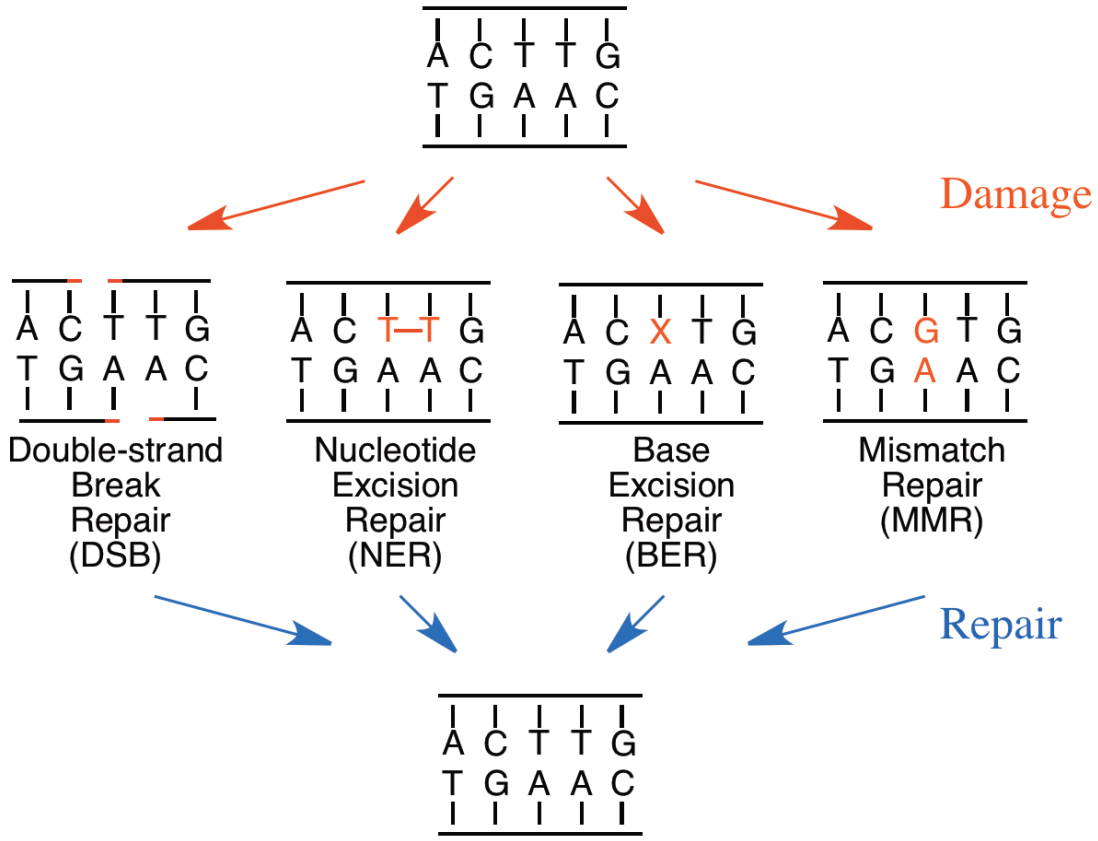

Figure 1-1. DNA damage can manifest itself in many forms and requires several pathways for accurate repair. Double strand break (DSB) repair is carried out by either homologous recombination (HR) or non-homologous end joining (NHEJ). Nucleotide excision repair (NER) removes DNA helix distorting lesions such as bulky adducts and crosslinked nucleotides. Base lesions resulting from relatively small modifications that do not impose significant distortions in the DNA are repaired the base excision repair (BER) pathway. Mismatched nucleotides are one type of error restored by the mismatch repair (MMR) pathway. Additionally, MMR repairs unpaired nucleotide bulges and other secondary structured loops that can occur during replication or recombination events.

At least 125 genes, making up six major pathways of DNA repair, contribute to maintaining an incredibly low rate of mutations (8) (Figure 1-1)[1]. Most pathways of DNA repair function by removing the damaged DNA. Repair synthesis replaces the damaged DNA, taking advantage of the duplex structure of DNA to restore the correct nucleotide sequence. (9). Specifically, the base excision repair (BER) pathway is

[1]Not illustrated in Figure 1-1 is repair by direct reversal. This unique process reverses covalent modifications to DNA without requiring the commonly employed cycle of removal and replacement of damaged or improper nucleotides (10).

responsible for removal and replacement of individual damaged nucleotides, where dimerized, cross-linked, and bulky lesions are repaired by nucleotide excision repair (NER). Mismatch repair (MMR) can be coupled with DNA replication to provide repair of mismatched and unpaired nucleotides that result from polymerase errors during replication (11). Homologous recombination (HR) and non-homologous end joining (NHEJ) represent repair of double strand breaks. In short, each repair pathway specializes in repairing distinct classes of DNA damage.

Insufficient DNA repair activity in individuals manifests as premature aging, increased occurrence of cancer, and shortened lifespan (12). These hallmarks of insufficient DNA repair activity can originate from inactivation of DNA repair genes by both hereditary and somatic mutations. Additionally, there is abundant evidence for epigenetic misregulation contributing to DNA repair dysfunction by altering expression of DNA repair genes (13-15). Dysfunctional DNA repair can lead to an accumulation of DNA damage, resulting in an increased rate of mutations. Acceleration of the mutation rate, termed genomic instability, can trigger global cellular responses to enter senescence or apoptosis (16). Importantly, it is the mutations that inhibit cellular responses to genome instability that are known to drive development of carcinogenesis (17). The function of DNA repair pathways and the impact of DNA repair dysfunction will be discussed further.

Repair of Double Strand DNA Breaks

Double strand breaks are particularly cytotoxic due to fragmentation of chromosomal DNA leading to rearrangement or loss of large amounts of genomic material. Although rare, double strand breaks pose a monumental threat to genome

stability, and appropriately, attract significant attention from the cell. The eukaryotic DNA damage response can halt cell cycle progression and provoke significant changes in gene expression to give extra time for repair of breaks. Ultimately, persisting breaks can lead to cell destruction by apoptosis (18). Ataxia telangiectasia (AT) is a hereditary genetic disease due to the dysfunction of the ATM dependent DNA damage response, and is characteristic of severe genomic instability and ~100-fold increase in cancer susceptibility (19).

Two distinct pathways exist for repair of double strand breaks, homologous recombination (HR) and non-homologous end-joining (NHEJ). Both pathways work to reunite the ends of the DNA at the break site, however, the relative contribution of these two pathways to break repair is influenced by several cellular processes including the status of the cell cycle (20).

Repair of DNA Helix Distorting Damage

High-energy radiation can induce major structural distortions in the chromosomal DNA through formation of intra- and interstrand crosslinks (21). Xeroderma pigmentosum (XP) is a hereditary disease characterized by enhanced photosensitivity leading to a >1000-fold increased chance in developing UV-induced skin cancer. XP is a result of defects in the nucleotide excision repair (NER) pathway (22). NER is adept at repairing duplex distorting bulky lesions and covalently dimerized nucleotides that are readily induced by non-ionizing UV-radiation (23).

Non-distorting DNA Damage

Smaller, less structurally distorting lesions occur at a high frequency in the genome. It has been estimated that 100,000 lesions are created each day in a human cell as a result of endogenous and exogenous chemicals reacting with DNA (1)[2]. Reactive oxygen species formed during aerobic cellular respiration are thought to be the most significant contributor to the formation of damaged nucleotides. (24). Specifically, hydrogen peroxide, hydroxyl radials, and singlet oxygen can oxidize nucleic acids at several different positions (25). Reports have documented that up to 10,000 8-oxoguanine (8oxoG) lesions are produced per day due to oxidation at the C8 position of guanine (26). Additionally, the loss of exocylic amines via deamination can occur spontaneously and is accelerated in the presence of nitrogen oxides (27, 28). Several studies report an estimated 500 pyrimidines are deaminated per day (1, 29). Specifically, deamination of cytosine creates uracil. 5-methylcytosine, a modification used as an epigenetic marker, induces a higher rate of deamination than cytosine, resulting in formation of thymine.

Alkylating agents transfer one or more carbon atoms to DNA. Common alkylating agents are N-nitroso compounds acquired exogeneously from airborne pollutants like tobacco smoke (30). Additionally, S-adenosylmethionine (SAM) is a vital molecule in many biological processes, but can spontaneously transfer the methyl group to DNA (31).

[2]Unperturbed by chemical agents, DNA degrades at a significant rate, producing an estimated 10,000 spontaneous depurinations in the genome per day (32, 33) The resulting abasic sites are processed by the BER pathway in the same way as glycosylase intiated BER.

Both purine and pyrimidine nucleobases are susceptible to alkylation at several positions, including ring nitrogen atoms and exocyclic oxygen atoms (34). Most frequently, alkylation reactions occur at the N7 position of guanine and account for 60-80% of alkylative modifications. The N3 position of adenine is also susceptible to alkylation, contributing 20-30% of total alkylation modifications to DNA (35), with other ring nitrogen atoms accumulating modifications at much lower rates (34). Ethylation, the addition of two carbon atoms, can occur in significant amounts due to exposure to vinyl chloride. Vinyl chloride is metabolized to 2-chloroacetaldehyde in the liver and can produce several purine adducts, $N^{1,6}$-ethenoadenine, as well as $N^{1,2}$-ethenoguanine and $N^{2,3}$-ethenoguanine (36).

DNA Glycosylase Initiated Repair of Non-distorting Lesions

In general, lesions in the genome must be repaired prior to a replication event to prevent mutations. DNA glycosylases catalyze the removal of damaged nucleotides from the genome. As many damaged nucleotides are not strikingly altered in structure, distinguishing a damaged nucleotide from the enormous excess of undamaged nucleotides poses a unique challenge to glycosylase catalysis. Further, the spectrum of damaged nucleotides varies significantly in shape, size, and charge, providing too much substrate diversity for one active site. Therefore, the human cell employs 11 different glycosylases (Table 1-1). Each glycosylase is responsible for recognizing a single lesion or small subset of damaged nucleotides.

Four different glycosylase structural superfamilies are present in the human glycosylase proteome. There is evidence that the different structural families evolved independently, yet several core features of catalysis are ubiquitously found among all

glycosylases (37). To gain access to a nucleotide in duplex DNA, a glycosylase rotates the nucleotide 180º degrees, flipping the nucleotide into the active site. This mechanism, termed base-flipping, is a common mechanism for accessing a nucleotide buried in the duplex structure of DNA (38). In concert with rotation of the nucleotide out of the DNA duplex, a bend is locally induced in the DNA. DNA bending has been proposed to facilitates nucleotide flipping (39). The final feature of base-flipping is the insertion of an intercalating loop into the gap vacated by the flipped nucleotide.

Structural Family	Glycosylase	Mechanism
Helix-hairpin-Helix	Nth1	Bifunctional
	Ogg1	Bifunctional
	MutY	Bifunctional
	MBD4	Monofunctional
Uracil DNA Glycosylase	UDG	Monofunctional
	SMUG1	Monofunctional
	TDG	Monofunctional
	NEIL1	Bifunctional
FapyG DNA Glycosylase	NEIL2	Bifunctional
	NEIL3	Bifunctional
Alkyladenine DNA Glycosylase	AAG	Monofunctional

Table 1-1. Human glycosylase proteome consists of 11 different glycosylases from four distinct structural families.

Once a damaged nucleotide is bound in the active site, the N-glycosidic bond can be hydrolyzed and the nucleobase liberated, leaving an abasic site. Two general mechanisms for N-glycosidic bond hydrolysis are employed (Figure 1-2). Monofunctional glycosylases activate a water molecule to promote nucleophilic attack on the C1 position of the deoxyribose sugar, catalyzing N-glycosidic bond cleavage by a nuclephilic substitution mechanism (39). An alternative mechanism for N-glycosidic

bond cleavage uses nucleophilic attack from an active site amine on the C1 position of the deoxyribose sugar, forming a covalent Schiff base intermediate (39). As a result of the covalent intermediate, an elimination of the 3'-phosphate at the β position can form a DNA break at the lesion site[3]. Since both lesion excision and breakage of the DNA backbone are required steps for repair of a damaged nucleotide, bifunctional glycosylase activity can contribute the first two steps required for repair of a damaged nucleotide [4].

Although the mechanism for excision of a damaged nucleotide is largely conserved within each glycosylase family, the recognition of a specific lesion or subset of lesions requires unique contributions by each glycosylase. Uracil DNA glycosylase (UNG) removes uracil from DNA with astonishing efficiency (40). Fortuitously, uracil is the smallest nucleobase and specificity for uracil can be achieved by a sterically constrained active site that occludes all other nucleobases (41). In contrast, akyladenine DNA glycosylase[5] removes a diverse array of damaged purines (35). Broad recognition of a range of different lesions is achieved by strategic positioning of active site residues to block exocyclic amines of guanine and adenine. Secondly, as pyrimidines cannot be excluded by an active site large enough to accommodate purine nucleobases, the use of acid/base catalysis to simultaneously protonate the nucleobase in the active site and activate a water molecule for nucleophilic attack is incompatible with

[3]Fapyguanine (Fpg) family of DNA glycosylases utilize a bifunctional mechanism that can perform both a β and δ elimination reaction on a abasic site. [4]Bifunctional glycosylase activity is not obligatory. The covalent Schiff base intermediate can be hydrolyzed to produce only an abasic site, unsuccessfully performing the β−elimination strand cleavage reaction. [5]AAG is also known as methylpurine DNA glycosylase (MPG) and 3MeA DNA glycosylase.

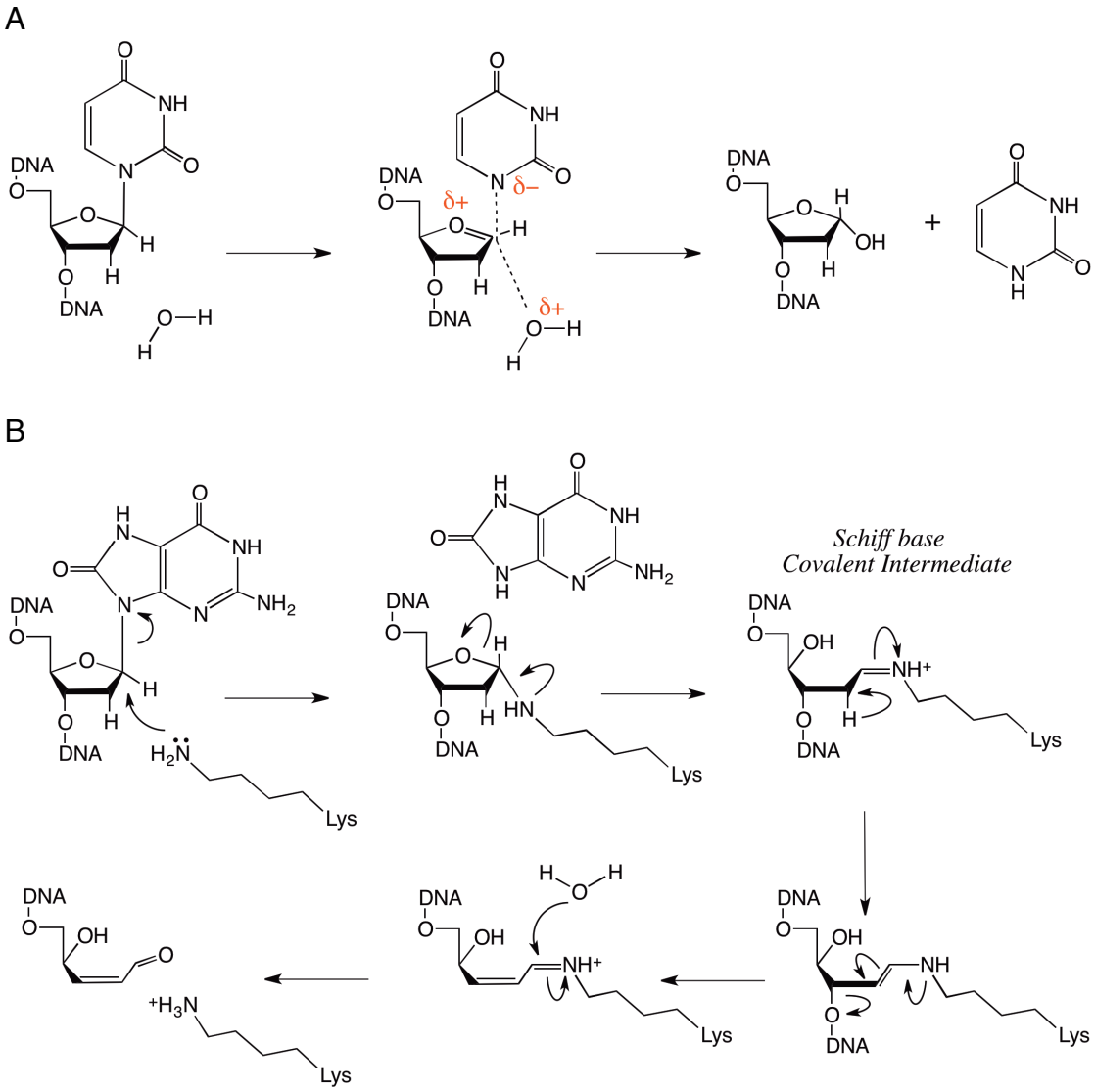

Figure 1-2. Mechanisms of N-glycosidic bond cleavage by a DNA glycosylase. (A) The mechanism for monofunctional glycosylase catalysis utilizes an activated water molecule as the nucleophile in a nucleophilic substitution reaction. The nucleophilic substitution reaction could occur by an associative or a dissociative substitution mechanism. The product of monofunctional glycosylase catalysis is an abasic site, in addition to the liberated nucleobase. (B) Bifunctional glycosylase catalysis utilizes a nucleophilic residue, forming a covalent intermediate upon N-glycosidic bond cleavage. The Schiff base intermediate can be hydrolyzed to dissolve the covalent intermediate, leaving an abasic site product. As illustrated, the Schiff base intermediate can proceed to break the DNA backbone through a β-elimination mechanism. Illustrations adapted from (39, 42).

cleavage of a pyrimidine nucleobase (42). As a consequence of broad substrate specificity, AAG incurs a reduced catalytic efficiency compared to UNG and catalyzes the removal of undamaged nucleotides at a slow rate (42, 43).

Glycosylase catalysis removes a damaged nucleobase from the genome, producing an abasic product. Importantly, an abasic site is still a pro-mutagenic lesion, lacking any information to support DNA templating. Removal of abasic sites requires the activity of the BER pathway (Figure 1-3). Primary processing of abasic sites is performed by apurinic/apyrimidinic (AP)-endonuclease I (APE1). APE1 hydrolyzes the DNA backbone 5' to the abasic site (44). At this point in the repair pathway, the nicked intermediate displays a 3'-OH that is compatible with polymerase extension. Short-patch BER pathway is performed by polymerase β. Polymerase β is non-processive and fills the gap with a new nucleotide. Polymerase β possesses a 5'-deoxyribose phosphate (dRP)-lyase domain that subsequently removes the 5'-dRP moiety, leaving a ligatable nick (45)[6]. Ligase is utilized by the BER pathway to complete repair (46).

[6]Alternatively, long-patch DNA repair proceeds via the activity of a processive polymerase. Polymerase δ or ε can extend the nicked intermediate and incorporate several nucleotides, displacing several nucleotides and creating a flap. Flap endonuclease I (FEN1) removes the flap to allow a final ligation step (47). AP-lyase activity provided by a bifunctional glycosylase can process an abasic site in a different manner than APE1. AP-lyase activity results in a 3'-dR moiety. This moiety can be removed by APE1 phophodiesterase activity.

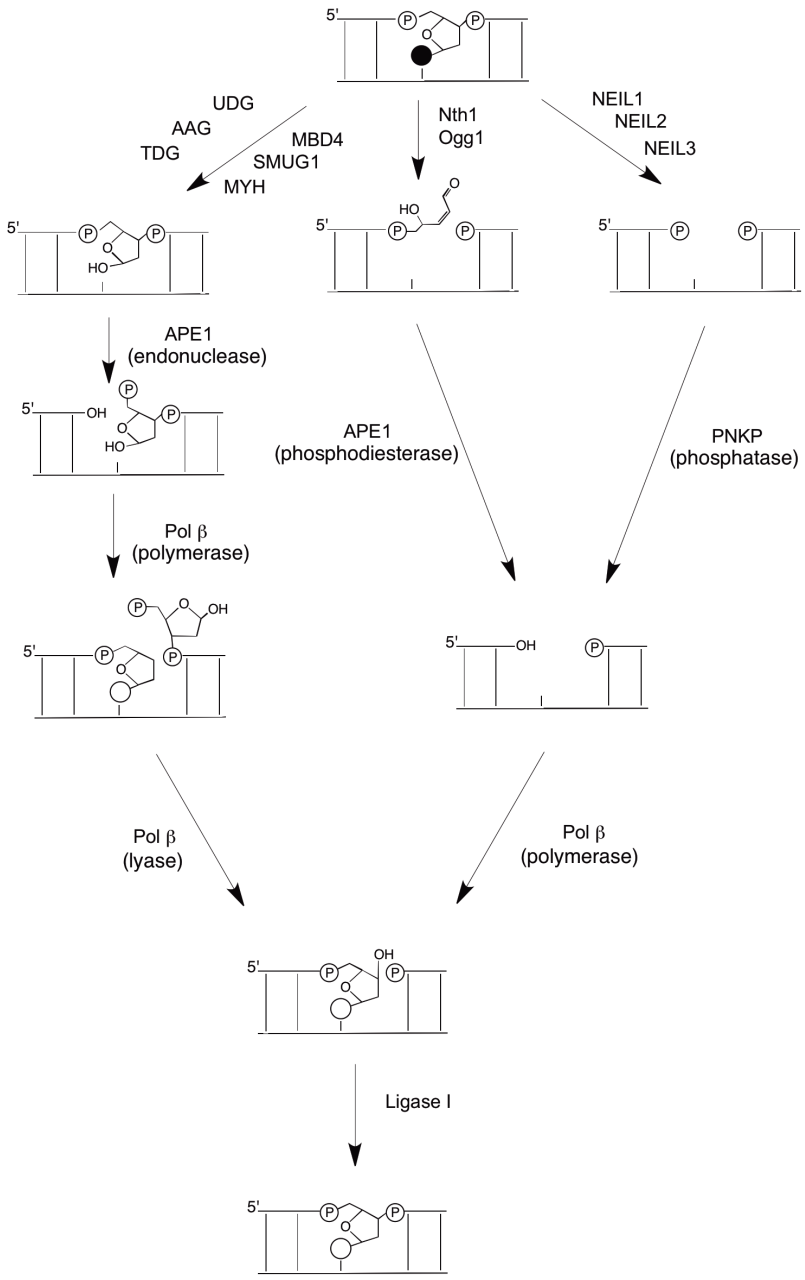

Figure 1-3. The base excision repair pathway is initiated by a glycosylase. Repair is initiated by excision of the lesion (black filled circle) by a glycosylase. The identity of the glycosylase determines the downstream processing. All pathways require nicking of the DNA backbone by bifunctional a glycosylase or endonuclease activity from APE1. A common requirement for all pathways is the incorporation of a new, undamaged nucleotide by polymerase β. Ligation by Ligase I is the ubiquitous final step of BER.

Dysfunction of the BER pathway

As described above, the BER pathway is highly divergent and can utilize several different enzymes to achieve the desired repair outcome. This provides redundant activities at each step in the pathway. Nevertheless, dysfunction of this repair pathway can lead to an overall increase in the mutation rate as the burden of damaged nucleotides are not removed from the genome. A DNA glycosylase initiates the pathway, providing lesion specificity. Knockout mice have been developed for several human glycosylase homologs, the majority maintaining healthy, fertile adults with weak or no phenotype. In fact, these studies lead to the discovery of several novel glycosylases due to redundant activities (48).

Gene-targeted knockout of UNG in mice was expected to display the 50-fold increase in C-G to T-A transition mutations that was observed in the heterologous *S. cerevisae* experiment (49). Surprisingly, no observable phenotypic affects were observed, as the loss of UNG contributed only a 1.5-fold increase in the expected C-G to T-A transition mutations. Back-up repair of uracil lesions was attributed to the glycosylase named single-strand selective monofunctional uracil glycosylase (SMUG1) (50, 51). This was the first observation of redundant, overlapping activity in the glycosylase proteome and has been characterized for several human glycosylases activities. A similar result was reported when the human homolog of endonuclease III (Nth1) knockout was generated in mice. No abnormalities were observed in the mice over a two-year period. This seemingly non-informative result was instrumental in locating redundant repair of several pyrimidine lesions by the Neil family of glycosylases (52). Mice deficient in AAG (*mpg)*

were confirmed to lack glycosylase excision of several lesions that are known substrates for AAG, namely $N^{1,6}$-ethenoadenine and hypoxanthine. Nevertheless, mice were healthy and viable. Only a modest increase in the accumulation of these lesions were observed, but other pathways may exist that don't employ glycosylases including NER and direct repair (53). These mild phenotypes are in sharp contrast to the affects obtained by loss of downstream BER enzymes. Transgenic mice harboring APE1, polymerase β, or ligase I gene knockouts are embryonic lethal (54-56). Redundant BER activities exist for these enzymes, but appear to not be able to cope with the burden of repair intermediates[7].

Failure to remove lesions from the genome can lead to mutations. Nucleotides with small chemical modifications do not significantly distort duplex structure, but can alter the interactions used for templating reactions like replication and transcription (57). For example, an 8oxoG lesion can maintain the Watson-Crick interactions of the canonical G-C basepair, but can also support 8oxoG-A basepairing via a Hoogsteen mode (58). When DNA polymerase encounters 8oxoG, adenine may be incorporated, leading to a permanent mutation upon the subsequent round of replication. Mutation of a G-C to a T-A basepair at the location is termed a transversion mutation (59). Transversion mutations are promoted by several lesions, including uracil. Uracil formed from deamination of cytosine results in a U-G basepair. Replication of a U-G can lead to the

[7]APE1, polymerase β, and ligase I may contribute essential processes during early embroyonic development, contributing to their acute embroyonic lethality.

misincorporation of an adenine and a C-G to T-A transversion (60). Lesions can also block replication and lead to stalling of the replication fork (61). 3-methyladenine does not impair basepairing with thymine. However, the methyl modification blocks DNA polymerization by causing a steric clash on the minor groove of the Watson-Crick basepairing face (62, 63).

A stalled replication fork can be rescued by translesion DNA synthesis (TLS). A class of specialized translesion DNA polymerases exist that may be temporarily substituted for the replicative polymerase at the stalled fork in order effectively bypass an non-canonical or damaged nucleotide (64). Structural and biochemical data suggest that translesion polymerases facilitate lesion bypass by relaxing replication fidelity (65). Low-fidelity replication by translesion bypass polymerases can result in mutations formed by misincorporation of an incorrect nucleotide (66). In addition to instigating miscoding mutations, several translesion bypass polymerases have been documented to increase the rate of insertion and deletion (67) mutations by >1000-fold in comparison to high fidelity replicative polymerases (68).

Frameshift Mutagenesis

Indel mutations are defined by the gain or loss of basepairs from the genome. Indels occurring within the reading frame of a gene are particularly deleterious to protein function due to the requirement that codons are transcribed in a group of three nucleotides, beginning from a defined start position. Therefore, indels of length $3n\pm1$ shift the frame codon sequence out of the intended frame of reference. A frameshift mutation alters the transcription of the gene product and commonly destroys protein function. Additionally, an out of frame gene may reveal an unintended stop codon and

lead to truncation of the gene product. In case of an indel mutation of $3n\pm0$, the frame of reference is maintained but a change in the number of codons occurs. The insertion or deletion of additional codons can affect the protein structure of the gene product and is implicated in several diseases, including Huntington's disease and oculopharyngeal muscular dystrophy (69).

Several mechanisms for indel formation have been suggested, the first described over 40 years ago by Streisinger (70). The foundational concept for indel formation involves misalignment of the primer and template strands within the polymerase active site and generation of unpaired nucleotide(s). Misalignment can be generated by dNTP misincorporation (71). Insertion of an incorrect nucleotide impairs the ability to basepair directly with the opposing template, and subsequently, the incorrect nucleotide may shift to interact with an adjacent nucleotide[8]. This repositioning of the primer in the polymerase active site is termed polymerase slipping, and extension of the repositioned primer leaves one or more bulged, unpaired base(s). In the absence of repair, another round of replication will convert the bulged intermediate into a frameshift mutation and produce a trademark of one daughter cell with a frameshift mutation and one wildtype.

[8]Misalignment of the primer and template strands could occur spontaneously, leading to the presentation of an upstream or downstream nucleotide to the incoming dNTP in the polymerase active site.

As introduced by Streisinger, the classic model for polymerase slipping is promoted in regions of repetitive sequence (Figure 1-4). Consecutively repeating nucleotides, termed homopolymeric tracts, are more adept at fostering a misaligned intermediate due to a localized enrichment of supporting template basepairing partners. Longer homopolymeric tracts allow the misaligned intermediate to propagate away from the extending primer and better escape detection by polymerase proofreading mechanisms. The high propensity of slipping events in homopolymeric tracts can be offset by correcting bulged intermediates.

Repair of Frameshift Mutation Intermediates

A common feature of replicative DNA polymerases is the incorporation of a proofreading domain (72). The proofreading domain performs 3' to 5' exonucleolytic degradation of newly synthesized DNA if an error is detected. Deletion of the exonuclease domain from DNA polymerase highlight the role of polymerase proofreading in extension of misaligned primer-template strands. Studies conducted in *S. cerevisiae* measured the rate of frameshift mutations as a function of hompolymer tract length (73). Results of this study showed that polymerase proofreading activity is most efficient in homopolymeric tracts of eight nucleotides or less, consistent with the ability of longer tracts to sequester a bulged intermediate further away from the DNA polymerase (74).

If a bulged intermediate escapes repair by DNA polymerase proofreading, then mismatch DNA repair pathway (MMR) serves as a backup to detect and repair immediately after replication. MMR provides recognition of a broad range of replication errors, from mismatched bases to small and large bulges produced by polymerase

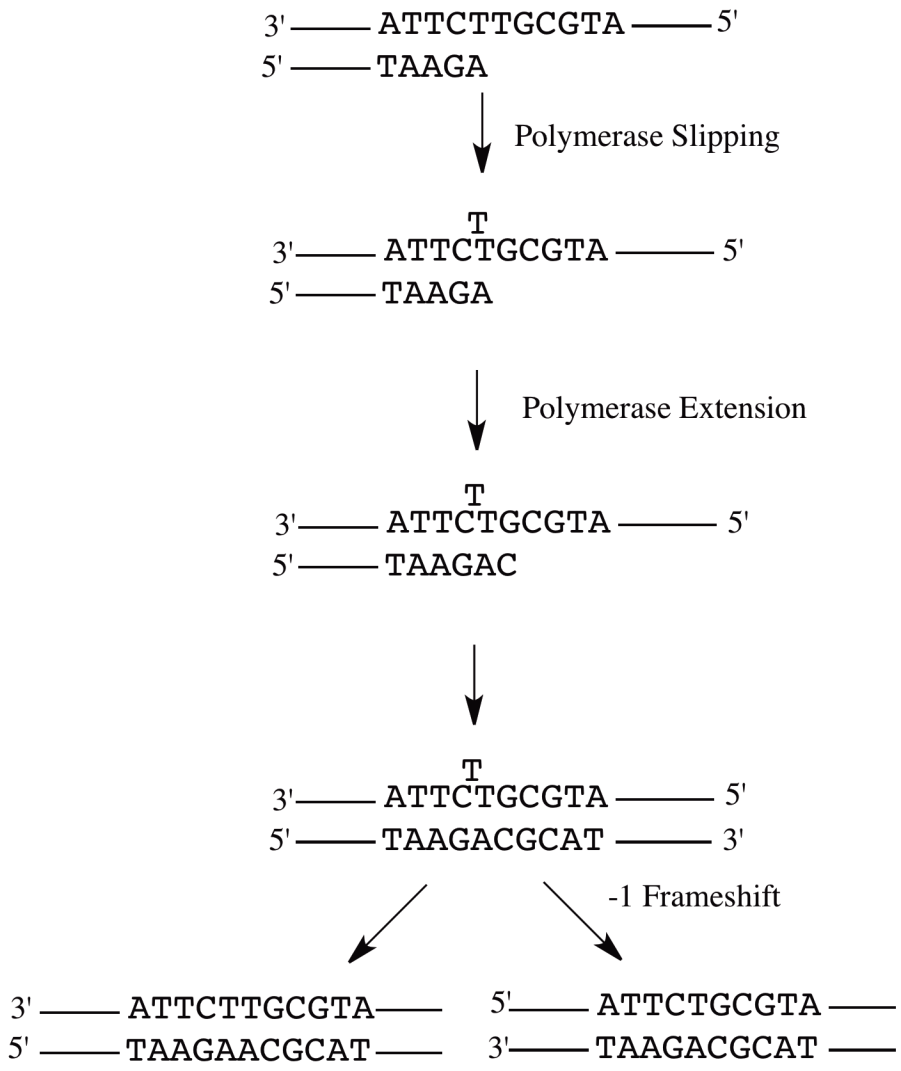

Figure 1-4. Streisinger model for polymerase slipping involves a misalignment of the primer and template strand. Extension of the misaligned intermediate forms a bulged, unpaired nucleotide intermediate. A following round of replication will produce a frameshift mutation. A deletion (-1 frameshift) occurs due a bulge present in the template strand. Similarly, an insertion (+1 frameshift) occurs due to formation of a bulge on the newly synthesized strand.

slipping errors. In humans, two recognition complexes are utilized. MSH2α complex is a heterodimer constructed of MSH2 and MSH6 and confers high affinity for mismatches and small (1-2 nucleotide) bulges (75). Recognition of larger bulges is accomplished by MSH2β, a heterodimer between MSH2 and MLH3 (76). One of these two MSH2 recognition complexes initiate MMR repair by binding to the DNA at the site of the error. After MSH2 binding, the MMR proceeds by selectively nicking the DNA strand with the error and then degrading approximately one kilobase of DNA. Re-synthesis of the degraded region by a replicative polymerase completes the repair pathway (77). A critical feature of MMR is the ability to detect the newly synthesized strand. This feature aids in determining which nucleotide in a mismatch or bulge is incorrect.

Dysfunction of the MMR pathway results in an elevated rate of replication errors, predominantly base substitutions and indels. Remarkably, inactivation of MMR in *S. cerevisiae* increased the rate of frameshift mutations by 10,000-fold in a repetitive, polyadenine tract. As stated above, indels occur with significantly higher frequency in repetitive sequences. Hereditary defects in MMR function predispose individuals for colon cancer. Commonly referred to as hereditary non-polypopsis colon cancer (HNPCC), individuals with this condition have a significant risk of developing early onset colon cancer (78, 79). HNPCC is clinically diagnosed by instability in the length of repetitive sequences, termed microsatellite instability (80).

Intriguingly, under conditions of functional MMR, overexpression of BER enzymes has been documented to result in microsatellite instability[9]. Upregulation of AAG positively correlated with microsatellite instability (81). Further, heterologous overexpression of AAG in *S. cerevisiae* leads to an increase in frameshift mutations at a repetitive sequence with a 30-fold bias toward -1 frameshift mutations. As the induction of microsatellite instability can be attributed to lack of MMR repair activity on bulged intermediates, the mechanism for a glycosylase mediated frameshift mutation is likely to involve competition between AAG and the MMR pathway for repair of a bulged intermediate.

Human BER creates a bias for -1 frameshift mutations

My doctoral work examined the mechanism by which BER contributes to frameshift mutagenesis. We determined the ability of glycosylases to catalyze the excision of a bulged nucleotide in chapter 2. Several different damaged and undamaged nucleotides were presented in a single nucleotide bulge structure in HeLa whole cell extract. Undamaged, bulged nucleotides remained stable against excision. Contrarily, several types of bulged lesions were rapidly excised in HeLa extract. Targeted inhibition confirmed the identity of two glycosylases that were responsible for the bulge excision activity that was observed in HeLa extract, UDG and AAG.

[9]Overexpression of polymerase β was shown to increase the rate of -1 frameshift mutations due to an increase in the propensity for formation of slipped intermediates (82). Polymerase β is a non-processive polymerase. Excess polymerase β was suggested to act in a processive manner, but with increased propensity of forming slipped intermediates (83).

A bulged lesion may arise *in vivo* due to polymerase slippage on a damaged template (70). Alternatively, undamaged bulged bases may be more susceptible to DNA damage. We showed that an undamaged, bulged adenine is at least 20-fold more susceptible to incurring damage than adenine basepaired to thymine. The lack of a base directly opposing the bulged nucleotide increases the dynamics of the bulged nucleotide . The increased dynamics are expected to raise the exposure of a bulged nucleotide to deamination and alkylation damage.

Under physiological conditions, several human cancer cell extracts had the ability to process a bulged lesion to a single nucleotide deletion. Reconstitution of the short patch BER pathway with recombinant human proteins demonstrated that the processing of each enzyme closely follows the canonical pathway for repair of a lesion with an opposing base. Activity of BER is expected to give a bias toward deletion mutation by repairing +1 insertions and making -1 deletions permanent (Figure 1-5).

In order to determine the extent of AAG discrimination against a bulged lesion, the efficiency of AAG excision of a bulged lesion was measured in chapter 3. Compared to a basepaired I-T mismatch, an inosine bulge is excised with a 3-fold greater catalytic efficiency. By comparing inosine excision in several structural contexts, the efficiency of inosine excision was found to correlate inversely with the stability of the inosine interaction with the opposing base. High efficiency excision of a bulged lesion is likely due to a significantly reduced barrier to base-flipping. Therefore, AAG doesn't discriminate against excision of a bulged lesion, suggesting that AAG will be able to compete for bulged species *in vivo*.

We performed similar kinetic analysis to quantify the ability of Nth1 to recognize a bulged lesion in chapter 4. In contrast to the 3-fold increase in affinity for a bulged lesion that was observed by AAG, the relative affinity of Nth1 for a bulged lesion was 20-fold reduced compared to a mismatched lesion. Although, Nth1 discriminates against excision of a bulged lesion, the maximal turnover rate of Nth1 for a 5,6-dihydrouracil (DHU) lesion is the same regardless of structural context. Reconstitution of the BER pathway showed that Nth1 can initiate deletion of a DHU-bulge. Additionally, HeLa nuclear extract possessed the ability to process a DHU-bulge to a single nucleotide deletion on a similar timescale as the repair of a canonical DHU-G. This demonstrates that in spite of a reduction in affinity, there is sufficient glycosylase activity to excise a DHU-bulge.

Bulge excision appears to be a common characteristic of glycosylase catalysis because glycosylases from all four known structural families share this activity (84-86). It is likely that glycoyslases have some flexibility. This flexibility allows the nucleotide to be flipped into the glycosylase active site in spite of structural distortions induced in the DNA by a bulged nucleotide. Our results raise the possibility that other base-flipping enzymes might act on bulged species.

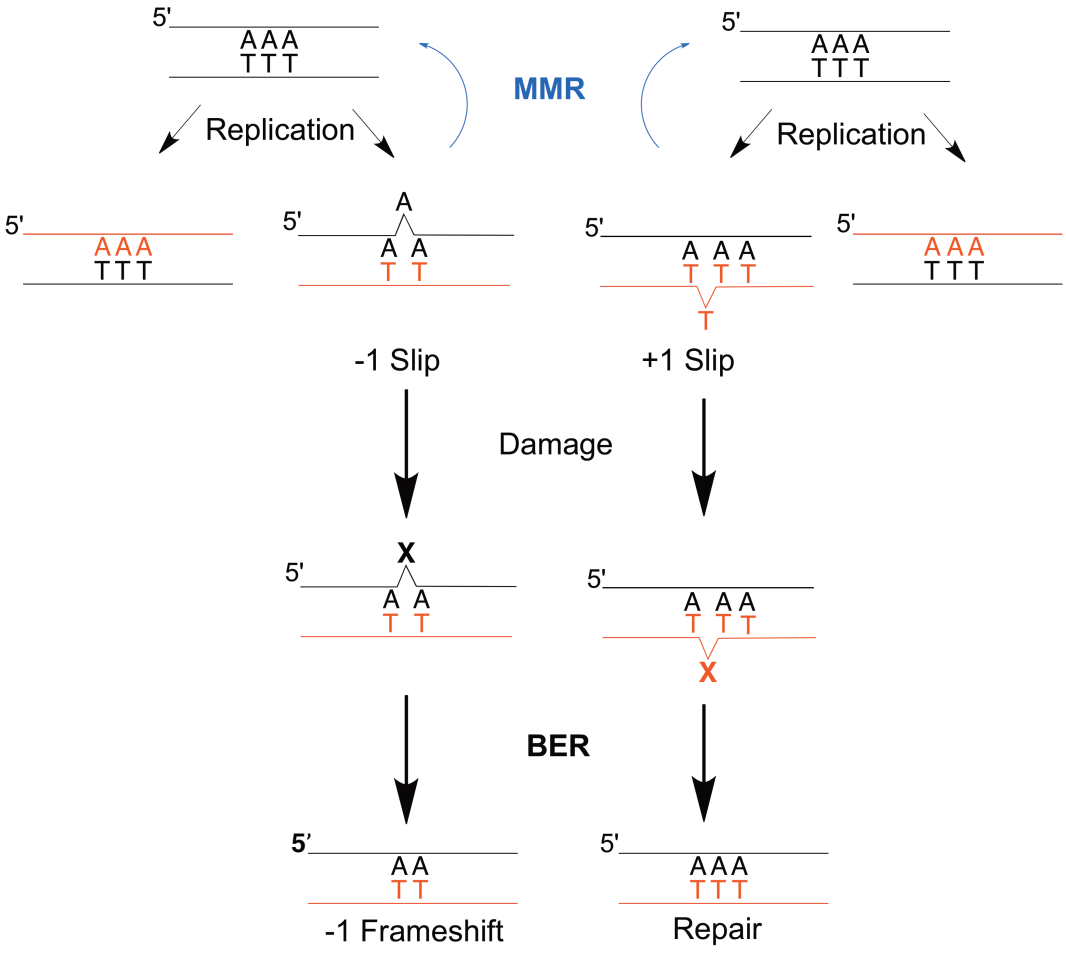

Figure 1-5. BER creates a bias for -1 frameshift mutations. Polymerase slipping can create a bulged, unpaired nucleotide intermediate. A -1 slip is illustrated as skipping over the replication of a nucleotide in reference to the template strand (black), where a +1 slip involves the insertion of an extra nucleotide in the newly synthesized strand (red). A bulged nucleotide is expected to have increased susceptiblilty to alkylation and deamination reactions. Processing of the bulged lesion by the BER pathway would ultimately lead to deletion of the bugled lesion. The BER deletion pathway would be expected to convert -1 slipping events into a -1 frameshift mutation, but repair +1 slipping events to create a bias for -1 frameshift mutations. Replication coupled MMR can provide accurate repair of either a +1 or -1 slipping event.

REFERENCES

1. Lindahl, T., *Instability and decay of the primary structure of DNA*. Nature, 1993. **362**(6422): p. 709-715.
2. Ciccia, A. and S.J. Elledge, *The DNA damage response: making it safe to play with knives*. Mol Cell. **40**(2): p. 179-204.
3. Jacobs, A.L. and P. Schar, *DNA glycosylases: in DNA repair and beyond*. Chromosoma. **121**(1): p. 1-20.
4. Hsu, G.W., M. Ober, T. Carell, and L.S. Beese, *Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase*. Nature, 2004. **431**(7005): p. 217-221.
5. Caporale, L.H., *Chance favors the prepared genome*. Ann N Y Acad Sci, 1999. **870**: p. 1-21.
6. Jackson, A.L. and L.A. Loeb, *On the origin of multiple mutations in human cancers*. Semin Cancer Biol, 1998. **8**(6): p. 421-429.
7. Loeb, L.A., *Mutator phenotype may be required for multistage carcinogenesis*. Cancer Res, 1991. **51**(12): p. 3075-3079.
8. Friedberg, E.C., A. Aguilera, M. Gellert, P.C. Hanawalt, J.B. Hays, A.R. Lehmann, T. Lindahl, N. Lowndes, A. Sarasin, and R.D. Wood, *DNA repair: from molecular mechanism to human disease*. DNA Repair (Amst), 2006. **5**(8): p. 986-996.
9. Nordstrand, L.M., J. Ringvoll, E. Larsen, and A. Klungland, *Genome instability and DNA damage accumulation in gene-targeted mice*. Neuroscience, 2007. **145**(4): p. 1309-1317.
10. Sedgwick, B., P. Robins, and T. Lindahl, *Direct removal of alkylation damage from DNA by AlkB and related DNA dioxygenases*. Methods Enzymol, 2006. **408**: p. 108-120.
11. Lindahl, T. and R.D. Wood, *Quality control by DNA repair*. Science, 1999. **286**(5446): p. 1897-1905.
12. Knoch, J., Y. Kamenisch, C. Kubisch, and M. Berneburg, *Rare hereditary diseases with defects in DNA-repair*. Eur J Dermatol.
13. Imai, K. and H. Yamamoto, *Carcinogenesis and microsatellite instability: the interrelationship between genetics and epigenetics*. Carcinogenesis, 2008. **29**(4): p. 673-680.
14. Banno, K., I. Kisu, M. Yanokura, K. Tsuji, K. Masuda, A. Ueki, Y. Kobayashi, W. Yamagami, H. Nomura, E. Tominaga, N. Susumu, and D. Aoki, *Epimutation and cancer: a new carcinogenic mechanism of Lynch syndrome (Review)*. Int J Oncol.
15. Nowsheen, S., K. Aziz, P.T. Tran, V.G. Gorgoulis, E.S. Yang, and A.G. Georgakilas, *Epigenetic inactivation of DNA repair in breast cancer*. Cancer Lett.
16. Lowe, S.W., E. Cepero, and G. Evan, *Intrinsic tumour suppression*. Nature, 2004. **432**(7015): p. 307-315.

17. Greenman, C., P. Stephens, R. Smith, G.L. Dalgliesh, C. Hunter, G. Bignell, H. Davies, J. Teague, A. Butler, C. Stevens, S. Edkins, S. O'Meara, I. Vastrik, E.E. Schmidt, T. Avis, S. Barthorpe, G. Bhamra, G. Buck, B. Choudhury, J. Clements, J. Cole, E. Dicks, S. Forbes, K. Gray, K. Halliday, R. Harrison, K. Hills, J. Hinton, A. Jenkinson, D. Jones, A. Menzies, T. Mironenko, J. Perry, K. Raine, D. Richardson, R. Shepherd, A. Small, C. Tofts, J. Varian, T. Webb, S. West, S. Widaa, A. Yates, D.P. Cahill, D.N. Louis, P. Goldstraw, A.G. Nicholson, F. Brasseur, L. Looijenga, B.L. Weber, Y.E. Chiew, A. DeFazio, M.F. Greaves, A.R. Green, P. Campbell, E. Birney, D.F. Easton, G. Chenevix-Trench, M.H. Tan, S.K. Khoo, B.T. Teh, S.T. Yuen, S.Y. Leung, R. Wooster, P.A. Futreal, and M.R. Stratton, *Patterns of somatic mutation in human cancer genomes*. Nature, 2007. **446**(7132): p. 153-158.
18. Roos, W.P. and B. Kaina, *DNA damage-induced apoptosis: From specific DNA lesions to the DNA damage response and apoptosis*. Cancer Lett.
19. Zakian, V.A., *ATM-related genes: what do they tell us about functions of the human gene?* Cell, 1995. **82**(5): p. 685-687.
20. Jackson, S.P., *Sensing and repairing DNA double-strand breaks*. Carcinogenesis, 2002. **23**(5): p. 687-696.
21. Masuda, Y. and K. Kamiya, *Molecular nature of radiation injury and DNA repair disorders associated with radiosensitivity*. Int J Hematol. **95**(3): p. 239-245.
22. Diderich, K., M. Alanazi, and J.H. Hoeijmakers, *Premature aging and cancer in nucleotide excision repair-disorders*. DNA Repair (Amst). **10**(7): p. 772-780.
23. Novarina, D., F. Amara, F. Lazzaro, P. Plevani, and M. Muzi-Falconi, *Mind the gap: keeping UV lesions in check*. DNA Repair (Amst). **10**(7): p. 751-759.
24. Klungland, A., I. Rosewell, S. Hollenbach, E. Larsen, G. Daly, B. Epe, E. Seeberg, T. Lindahl, and D.E. Barnes, *Accumulation of premutagenic DNA lesions in mice defective in removal of oxidative base damage*. Proc Natl Acad Sci U S A, 1999. **96**(23): p. 13300-13305.
25. Cooke, M.S., M.D. Evans, M. Dizdaroglu, and J. Lunec, *Oxidative DNA damage: mechanisms, mutation, and disease*. FASEB J, 2003. **17**(10): p. 1195-1214.
26. Fraga, C.G., M.K. Shigenaga, J.W. Park, P. Degan, and B.N. Ames, *Oxidative damage to DNA during aging: 8-hydroxy-2'-deoxyguanosine in rat organ DNA and urine*. Proc Natl Acad Sci U S A, 1990. **87**(12): p. 4533-4537.
27. Akbari, M. and H.E. Krokan, *Cytotoxicity and mutagenicity of endogenous DNA base lesions as potential cause of human aging*. Mech Ageing Dev, 2008. **129**(7-8): p. 353-365.
28. Karran, P. and T. Lindahl, *Hypoxanthine in deoxyribonucleic acid: generation by heat-induced hydrolysis of adenine residues and release in free form by a deoxyribonucleic acid glycosylase from calf thymus*. Biochemistry, 1980. **19**(26): p. 6005-6011.
29. Frederico, L.A., T.A. Kunkel, and B.R. Shaw, *A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy*. Biochemistry, 1990. **29**(10): p. 2532-2537.
30. Mishina, Y., E.M. Duguid, and C. He, *Direct reversal of DNA alkylation damage*. Chem Rev, 2006. **106**(2): p. 215-232.

31. Rydberg, B. and T. Lindahl, *Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction*. EMBO J, 1982. **1**(2): p. 211-216.
32. Overballe-Petersen, S., L. Orlando, and E. Willerslev, *Next-generation sequencing offers new insights into DNA degradation*. Trends Biotechnol. **30**(7): p. 364-368.
33. Lindahl, T. and B. Nyberg, *Rate of depurination of native deoxyribonucleic acid*. Biochemistry, 1972. **11**(19): p. 3610-3618.
34. Fu, D., J.A. Calvo, and L.D. Samson, *Balancing repair and tolerance of DNA damage caused by alkylating agents*. Nat Rev Cancer. **12**(2): p. 104-120.
35. Drablos, F., E. Feyzi, P.A. Aas, C.B. Vaagbo, B. Kavli, M.S. Bratlie, J. Pena-Diaz, M. Otterlei, G. Slupphaug, and H.E. Krokan, *Alkylation damage in DNA and RNA--repair mechanisms and medical significance*. DNA Repair (Amst), 2004. **3**(11): p. 1389-1407.
36. Sambamurti, K., J. Callahan, X. Luo, C.P. Perkins, J.S. Jacobsen, and M.Z. Humayun, *Mechanisms of mutagenesis by a bulky DNA lesion at the guanine N7 position*. Genetics, 1988. **120**(4): p. 863-873.
37. O'Brien, P.J. and D. Herschlag, *Catalytic promiscuity and the evolution of new enzymatic activities*. Chem Biol, 1999. **6**(4): p. R91-R105.
38. Dodson, M.L. and R.S. Lloyd, *Mechanistic comparisons among base excision repair glycosylases*. Free Radic Biol Med, 2002. **32**(8): p. 678-682.
39. Friedman, J.I. and J.T. Stivers, *Detection of damaged DNA bases by DNA glycosylase enzymes*. Biochemistry. **49**(24): p. 4957-4967.
40. Stivers, J.T. and A.C. Drohat, *Uracil DNA glycosylase: insights from a master catalyst*. Arch Biochem Biophys, 2001. **396**(1): p. 1-9.
41. Liu, P., A. Burdzy, and L.C. Sowers, *Substrate recognition by a family of uracil-DNA glycosylases: UNG, MUG, and TDG*. Chem Res Toxicol, 2002. **15**(8): p. 1001-1009.
42. O'Brien, P.J. and T. Ellenberger, *Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase*. J Biol Chem, 2004. **279**(11): p. 9750-9757.
43. Berdal, K.G., R.F. Johansen, and E. Seeberg, *Release of normal bases from intact DNA by a native DNA repair enzyme*. EMBO J, 1998. **17**(2): p. 363-367.
44. Schild, L.J., K.W. Brookman, L.H. Thompson, and D.M. Wilson, 3rd, *Effects of Ape1 overexpression on cellular resistance to DNA-damaging and anticancer agents*. Somat Cell Mol Genet, 1999. **25**(5-6): p. 253-262.
45. Beard, W.A. and S.H. Wilson, *Structural design of a eukaryotic DNA repair polymerase: DNA polymerase beta*. Mutat Res, 2000. **460**(3-4): p. 231-244.
46. Tomkinson, A.E. and D.S. Levin, *Mammalian DNA ligases*. Bioessays, 1997. **19**(10): p. 893-901.
47. Prasad, R., G.L. Dianov, V.A. Bohr, and S.H. Wilson, *FEN1 stimulation of DNA polymerase beta mediates an excision step in mammalian long patch base excision repair*. J Biol Chem, 2000. **275**(6): p. 4460-4466.
48. Wilson, D.M., 3rd and L.H. Thompson, *Life without DNA repair*. Proc Natl Acad Sci U S A, 1997. **94**(24): p. 12754-12757.

49. Impellizzeri, K.J., B. Anderson, and P.M. Burgers, *The spectrum of spontaneous mutations in a Saccharomyces cerevisiae uracil-DNA-glycosylase mutant limits the function of this enzyme to cytosine deamination repair*. J Bacteriol, 1991. **173**(21): p. 6807-6810.
50. Nilsen, H., I. Rosewell, P. Robins, C.F. Skjelbred, S. Andersen, G. Slupphaug, G. Daly, H.E. Krokan, T. Lindahl, and D.E. Barnes, *Uracil-DNA glycosylase (UNG)-deficient mice reveal a primary role of the enzyme during DNA replication*. Mol Cell, 2000. **5**(6): p. 1059-1065.
51. Kavli, B., O. Sundheim, M. Akbari, M. Otterlei, H. Nilsen, F. Skorpen, P.A. Aas, L. Hagen, H.E. Krokan, and G. Slupphaug, *hUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup*. J Biol Chem, 2002. **277**(42): p. 39926-39936.
52. Takao, M., S. Kanno, K. Kobayashi, Q.M. Zhang, S. Yonei, G.T. van der Horst, and A. Yasui, *A back-up glycosylase in Nth1 knock-out mice is a functional Nei (endonuclease VIII) homologue*. J Biol Chem, 2002. **277**(44): p. 42205-42213.
53. Engelward, B.P., G. Weeda, M.D. Wyatt, J.L. Broekhof, J. de Wit, I. Donker, J.M. Allan, B. Gold, J.H. Hoeijmakers, and L.D. Samson, *Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase*. Proc Natl Acad Sci U S A, 1997. **94**(24): p. 13087-13092.
54. Ludwig, D.L., M.A. MacInnes, Y. Takiguchi, P.E. Purtymun, M. Henrie, M. Flannery, J. Meneses, R.A. Pedersen, and D.J. Chen, *A murine AP-endonuclease gene-targeted deficiency with post-implantation embryonic progression and ionizing radiation sensitivity*. Mutat Res, 1998. **409**(1): p. 17-29.
55. Sugo, N., Y. Aratani, Y. Nagashima, Y. Kubota, and H. Koyama, *Neonatal lethality with abnormal neurogenesis in mice deficient in DNA polymerase beta*. EMBO J, 2000. **19**(6): p. 1397-1404.
56. Bentley, D.J., C. Harrison, A.M. Ketchen, N.J. Redhead, K. Samuel, M. Waterfall, J.D. Ansell, and D.W. Melton, *DNA ligase I null mouse cells show normal DNA repair activity but altered DNA replication and reduced genome stability*. J Cell Sci, 2002. **115**(Pt 7): p. 1551-1561.
57. Shibutani, S., *Quantitation of base substitutions and deletions induced by chemical mutagens during DNA synthesis in vitro*. Chem Res Toxicol, 1993. **6**(5): p. 625-629.
58. Bruner, S.D., D.P. Norman, and G.L. Verdine, *Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA*. Nature, 2000. **403**(6772): p. 859-866.
59. Shibutani, S., M. Takeshita, and A.P. Grollman, *Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG*. Nature, 1991. **349**(6308): p. 431-434.
60. Barnes, D.E. and T. Lindahl, *Repair and genetic consequences of endogenous DNA base damage in mammalian cells*. Annu Rev Genet, 2004. **38**: p. 445-476.
61. Mirkin, E.V. and S.M. Mirkin, *Replication fork stalling at natural impediments*. Microbiol Mol Biol Rev, 2007. **71**(1): p. 13-35.
62. Larson, K., J. Sahm, R. Shenkar, and B. Strauss, *Methylation-induced blocks to in vitro DNA replication*. Mutat Res, 1985. **150**(1-2): p. 77-84.

63. Plosky, B.S., E.G. Frank, D.A. Berry, G.P. Vennall, J.P. McDonald, and R. Woodgate, *Eukaryotic Y-family polymerases bypass a 3-methyl-2'-deoxyadenosine analog in vitro and methyl methanesulfonate-induced DNA damage in vivo*. Nucleic Acids Res, 2008. **36**(7): p. 2152-2162.
64. Friedberg, E.C., A.R. Lehmann, and R.P. Fuchs, *Trading places: how do DNA polymerases switch during translesion DNA synthesis?* Mol Cell, 2005. **18**(5): p. 499-505.
65. Zahn, K.E., S.S. Wallace, and S. Doublie, *DNA polymerases provide a canon of strategies for translesion synthesis past oxidatively generated lesions*. Curr Opin Struct Biol. **21**(3): p. 358-369.
66. Broyde, S., L. Wang, O. Rechkoblit, N.E. Geacintov, and D.J. Patel, *Lesion processing: high-fidelity versus lesion-bypass DNA polymerases*. Trends Biochem Sci, 2008. **33**(5): p. 209-219.
67. Orengo, C.A., A.D. Michie, S. Jones, D.T. Jones, M.B. Swindells, and J.M. Thornton, *CATH--a hierarchic classification of protein domain structures*. Structure, 1997. **5**(8): p. 1093-1108.
68. Matsuda, T., K. Bebenek, C. Masutani, I.B. Rogozin, F. Hanaoka, and T.A. Kunkel, *Error rate and specificity of human and murine DNA polymerase eta*. J Mol Biol, 2001. **312**(2): p. 335-346.
69. Davies, J.E. and D.C. Rubinsztein, *Polyalanine and polyserine frameshift products in Huntington's disease*. J Med Genet, 2006. **43**(11): p. 893-896.
70. Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita, E. Terzaghi, and M. Inouye, *Frameshift mutations and the genetic code. This paper is dedicated to Professor Theodosius Dobzhansky on the occasion of his 66th birthday*. Cold Spring Harb Symp Quant Biol, 1966. **31**: p. 77-84.
71. Kunkel, T.A., *Frameshift mutagenesis by eucaryotic DNA polymerases in vitro*. J Biol Chem, 1986. **261**(29): p. 13581-13587.
72. Garcia-Diaz, M., K. Bebenek, J.M. Krahn, L. Blanco, T.A. Kunkel, and L.C. Pedersen, *A structural solution for the DNA polymerase lambda-dependent repair of DNA gaps with minimal homology*. Mol Cell, 2004. **13**(4): p. 561-572.
73. Tran, H.T., J.D. Keen, M. Kricker, M.A. Resnick, and D.A. Gordenin, *Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants*. Mol Cell Biol, 1997. **17**(5): p. 2859-2865.
74. Modrich, P., *Mechanisms and biological effects of mismatch repair*. Annu Rev Genet, 1991. **25**: p. 229-253.
75. Palombo, F., P. Gallinari, I. Iaccarino, T. Lettieri, M. Hughes, A. D'Arrigo, O. Truong, J.J. Hsuan, and J. Jiricny, *GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells*. Science, 1995. **268**(5219): p. 1912-1914.
76. Palombo, F., I. Iaccarino, E. Nakajima, M. Ikejima, T. Shimada, and J. Jiricny, *hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA*. Curr Biol, 1996. **6**(9): p. 1181-1184.
77. Jiricny, J., *The multifaceted mismatch-repair system*. Nat Rev Mol Cell Biol, 2006. **7**(5): p. 335-346.
78. Watson, P. and H.T. Lynch, *Cancer risk in mismatch repair gene mutation carriers*. Fam Cancer, 2001. **1**(1): p. 57-60.

79. Jenkins, M.A., L. Baglietto, J.G. Dowty, C.M. Van Vliet, L. Smith, L.J. Mead, F.A. Macrae, D.J. St John, J.R. Jass, G.G. Giles, J.L. Hopper, and M.C. Southey, *Cancer risks for mismatch repair gene mutation carriers: a population-based early onset case-family study*. Clin Gastroenterol Hepatol, 2006. **4**(4): p. 489-498.
80. Liu, B., R. Parsons, N. Papadopoulos, N.C. Nicolaides, H.T. Lynch, P. Watson, J.R. Jass, M. Dunlop, A. Wyllie, P. Peltomaki, A. de la Chapelle, S.R. Hamilton, B. Vogelstein, and K.W. Kinzler, *Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients*. Nat Med, 1996. **2**(2): p. 169-174.
81. Hofseth, L.J., M.A. Khan, M. Ambrose, O. Nikolayeva, M. Xu-Welliver, M. Kartalou, S.P. Hussain, R.B. Roth, X. Zhou, L.E. Mechanic, I. Zurer, V. Rotter, L.D. Samson, and C.C. Harris, *The adaptive imbalance in base excision-repair enzymes generates microsatellite instability in chronic inflammation*. J Clin Invest, 2003. **112**(12): p. 1887-1894.
82. Chan, K., S. Houlbrook, Q.M. Zhang, M. Harrison, I.D. Hickson, and G.L. Dianov, *Overexpression of DNA polymerase beta results in an increased rate of frameshift mutations during base excision repair*. Mutagenesis, 2007. **22**(3): p. 183-188.
83. Kunkel, T.A., *The mutational specificity of DNA polymerase-beta during in vitro DNA synthesis. Production of frameshift, base substitution, and deletion mutations*. J Biol Chem, 1985. **260**(9): p. 5787-5796.
84. Zhao, X., N. Krishnamurthy, C.J. Burrows, and S.S. David, *Mutation versus repair: NEIL1 removal of hydantoin lesions in single-stranded, bulge, bubble, and duplex DNA contexts*. Biochemistry. **49**(8): p. 1658-1666.
85. Lyons, D.M. and P.J. O'Brien, *Efficient recognition of an unpaired lesion by a DNA repair glycosylase*. J Am Chem Soc, 2009. **131**(49): p. 17742-17743.
86. Lyons, D.M. and P.J. O'Brien, Human base excision repair creates a bias toward -1 frameshift mutations. J Biol Chem. 285(33): p. 25203-25212.

CHAPTER 2

HUMAN BASE EXCISION REPAIR CREATES A BIAS FOR -1 FRAMESHIFT MUTATIONS

The loss or gain of one or more base pairs is one of the most common types of genetic instability (1). Single nucleotide deletions or insertions occur more frequently than larger deletions or insertions and they cause frameshift mutations if they occur within the open reading frame of a gene. Genomic instability is proposed to be a hallmark of carcinogenesis and the loss or alteration of DNA replication or repair pathways is an early step in the progression of cancer (2, 3). Normally the mismatch DNA repair (MMR) pathway suppresses frameshift mutations by performing replication-coupled DNA repair, but many cancer cells inactivate this pathway through mutations or changes in promoter methylation that reduce expression of one or more proteins in this pathway (4-6). However, not all cases of increased frameshift frequency can be attributed to defects in MMR (7, 8). Here we consider the possibility that other DNA repair pathways might also play a role in frameshift mutagenesis.

[1]This research was originally published in Journal of Biological Chemistry. Derek M. Lyons and Patrick J. O'Brien. Human Base Excision Repair Creates a Bias for -1 Frameshift Mutations. *Journal of Biological Chemistry*. 2010; 285(33):25203-12. © the American Society for Biochemistry and Molecular Biology.

Many of the factors that influence the frequency of frameshift mutations are know (9, 10) Streisinger and co-workers (11) proposed that slipping of primer/template pair within a polymerase active site could lead to a misaligned intermediate that would subsequently be extended to generate a bulged intermediate. Subsequent models for frameshift mutations (1, 6), including the effects of damaged templates and DNA intercalators (12, 13), have incorporated this essential feature (Figure 2-1). It is generally assumed that another round of DNA replication is required to make this frameshift event permanent, with one copy that retains the original sequence and another copy that is mutant. Although both -1 and +1 frameshifts occur, human cells show a strong bias (3-6 –fold) toward -1 frameshift mutations (14-16). This bias had been attributed to the intrinsic propensity for some polymerases to make more frequent -1 than +1 errors. For example, yeast polymerase δ generates a greater number of -1 slips (17). Here we have considered the possibility that alternative DNA repair pathways recognize frameshift intermediates and preferentially repair +1 events and/or make -1 events permanent.

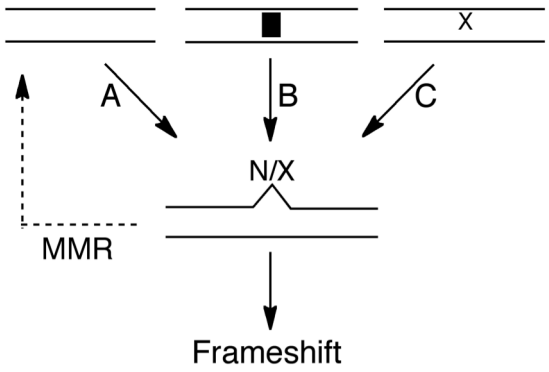

Figure 2-1. Polymerase slipping creates bulged intermediates that give rise to frameshift mutations. *Arrows* indicate DNA replication events. Slipping of polymerases generate a bulged nucleotide (N) that is less likely to be corrected by proofreading if it occurs in a homopolymeric sequence (A) (11). The probability of a polymerase slipping is increased by DNA intercalators (B) (12), or by a damaged template (C) (43,45). Slipping on a damaged nucleotide (X) generates a damaged bulge. MMR can correct these errors, whereas an additional round of replication makes the frameshift permanent.

The base excisions repair (BER) pathway is one DNA repair pathway that might compete against MMR, because the overexpression of BER enzymes result in increased frequency of point mutations and frameshift mutations, especially -1 frameshifts (18-20). The BER pathway repairs the majority of damaged bases that arise from spontaneous hydrolytic, oxidative, or alkylative reactions (for reviews, see Ref, 21-23). This pathway is initiated by glycosylases that locate sites of damage and excise the lesion base. A few of these enzymes also nick the DNA backbone via a lyase mechanism. But most release an apurinic/apyrimidinic (AP) product. The AP site is subsequently hydrolyzed by AP endonuclease 1 (APE1) to create a 3'-hydoxyl and a 5'-deoxyribose phosphate (dRP) group. Commonly the missing nucleotide is incorporated by DNA polymerase b, which also removes the 5'-dRP group via a lyase reaction. Finally, the single stand nick is ligated to complete the BER pathway. This pathway removes the damaged nucleotide and inserts a new nucleotide based upon base pairing with the opposing nucleotide however, if the BER pathway were initiated on a single nucleotide bulge, then it would be expected to delete the bulged nucleotide.

In this study we tested the hypothesis that the human BER pathway can delete bulged nucleotides, by performing DNA repair assays in cell extracts and reconstituting the BER pathway with purified components. The results demonstrates that the BER pathway is active toward alkylated or deaminated nucleotide bulges, and confirm that the deletion of the damaged nucleotide closely follows the well characterized activities of the BER enzymes. Damaged single nucleotide bulges could arise from the slipping of a polymerase on a damaged template, or could result from spontaneous damage of an

undamaged bulge. These observations suggest the BER may compete against MMR and that this competition could contribute to the bias toward single nucleotide deletions.

MATERIALS AND METHODS

Recombinant Proteins

Human BER enzymes were expressed in *Escherichia coli* and purified by affinity and ion exchange chromatography. The affinity tags were removed and purity was judged to be >95% by SDS-PAGE analysis with Coomassie staining. The purification of the catalytic domain of AAG (Δ80), lacking the first 79 amino acids (24), the catalytic domain of DNA ligase I (Δ232), lacking the first 232 amino acids (25), and the full-length APE1 have been previously described (26). Full-length DNA polymerase b was cloned into a modified pET28 vector containing amino-terminal His_6 tag, followed by a TEV cleavage site. After TEV cleavage, the amino terminus is extended by two extra amino acids (GH). It was purified by a similar protocol as previously described (27). Protein concentrations were determined by their estimated extinction coefficients at 280 nm.

Oligonucleotide Substrates

DNA oligonucleotides were synthesized by commercial sources using standard phosphoramidite chemistry and were purified and annealed as described (28). The sequence of 5'-fluorescein (FEM)-labeled oligonucleotide was 5'-FAM-CGATAGCATCCT*X*CCTTCTCTCCAT, in which *X* was either a normal nucleotide or damaged nucleotide. The labeled oligonucleotide or damaged nucleotide. The labeled oligonucleotide was annealed with a 24-mer complement to create a single nucleotide

bulge, or with a 25-mer complement to create a central base pair (X•Y; Appendix Figure A-1).

Glycosylase Assay in Whole Cell Extracts

Glycosylase reactions were monitored at 37ºC in a buffer containing 50 mM NaHEPES (pH 7.5), 100 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, 0.1 mg/mL of bovine serum albumin, and 10% (v/v) glycerol. Reactions (20 μL) contained 10 nM FAM-labeled substrate and 3 μL aliquots were removed at the desired time points and quenched with 2 volumes of 0.3 M NaOH to obtain a final concentration of 0.2 M. AP sites formed from glycosylase reaction were quantitatively converted to DNA breaks by heating at 70ºC for 15 min, followed by the addition of 3.3 volumes of formamide/EDTA loading buffer containing 0.05% (w/v) of bromphenol blue and xylene cyanol FF dyes. Samples were separated via denaturing PAGE and FAM fluorescence was quantified by imaging with a Typhoon trio (GE Healthcare) using 488nm excitation and a 520BP40 emission filter. Flourescence intensity of individual bands was determined with ImageQuantTL (GE Healthcare). The fraction of glycosylase product was determined by dividing the fluorescence intensity of the 12-mer product band by the sum of the 12-mer product and 25-mer substrate bands. The fraction was converted into concentration by multiplying by the total amount of DNA substrate present. Reactions were carried out in triplicate and the initial rates were measured by a linear fit. To confirm the identity of the 12-mer reaction product, we generated the authentic 12-mer from a single turnover reaction with 2 mM AAG and 10 nM I•T DNA. The reaction was quenched with NaOH after 15 min and treated as described above. Whole cell extracts (WCE) were obtained from Active Motif and were stored at -80ºC in lysis buffer (20 mM NaHEPES pH 7.5,

350 mM NaCl, 20% glycerol, 1% Igepal-CA630, 1 mM $MgCl_2$, 0.5 mM EDTA, 0.1 mM EGTA) until immediately before use.

Inhibition of U-bulge glycosylase activity was achieved by addition of 0.02 units of uracil glycosylase inhibitor (New England Biolabs) to 0.4 mg/mL HeLa WCE as described above. Inhibition of the glycosylase activity toward I-bulge and eA bulge bulge DNA (10nM) was investigated by adding 100nM unlabeled 25-mer competitor DNA that was wither damaged (eA•T) or undamaged (A•T). To estimate the amount of AAG present in extracts, the standard glycosylase assay was performed with either 0.2 or 0.4 mg/ml of WCE, 10nM I•T-labeled substrate, and 100nM A•T unlabeled DNA. The glycosylase activity of T47D and HeLa toward the I•T substrate was linearly dependent on the amount of extract added (supplemental Fig. S6).

Alkylation of DNA by Chloroacetaldehyde

A coupled assay was employed to determine the susceptibility of a bulged base to alkylative damage from chloroacetaldehyde. 1 μM A•T or A-bulge oligonucleotide as indicated was incubated at 37˚C in 50 mM Na_2HPO_4 (pH 7.5), and 100 mM NaCl. 100 mM Chloroacetaldehyde (Sigma) was added as indicated. Reactions (4 μL) were incubated for 4 h at 37˚C, followed by the addition of 50 μL TE buffer (10 mM Tris, pH 8.0, 1mM EDTA) and 50 μL of hydrated diethyl ether. The organic layer was removed and a second ether extraction step was performed. DNA was precipitated by the addition of 150 μL of cold ethanol, 0.12M sodium acetate. Samples were incubated for 30 min at -80˚C and centrifuged for 15 min. After removal of the supernatant, samples were washed with 0.5 mL of 95% ethanol and centrifuged for 10 min. The supernatant was removed and the samples were air-dried. The εA lesions that were formed by alkylation of the

DNA were subsequently removed b AAG, which was found to have the same level of activity toward an εA-T mismatch as toward and εA-T bulge (data not shown) Samples were resuspended in 20 μL of AAG reaction buffer, 50mM NaMES (pH 6.1), 100 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, and 0.1 mg/mL bovine serum albumin, and incubated with 2 μM AAG at 37ºC for 30 min. Samples were quenched in NaOH and analyzed by the standard glycosylase assay.

BER Assay in WCE

Assays for the processing of labeled mismatched and bulge DNA in were performed at 37ºC in a 50 mM NaHEPES (pH 7.5), 100 mM NaCl, 5 mM $MgCl_2$, 4nM ATP and 25 μM of each dNTP. Typically 10nM unlabeled (A•T) DNA were incubated with 0.4mg/ml of extract (Active Motif) in a reaction volume of 20 μL. Samples were taken by quenching 3 μL aliquots into an equal colume of formamide with 20nM EDTA, heated to 70ºC for 5 min, and separated by denaturing PAGE. AP sites remained intact during this procedure. The amount of repair that had occurred was determined by quenching reactions with EDTA (10mM final). Recombinant AAG (2 μM) was added and reaction were incubated at 37 ºC for 30 min. Finally, these reaction were quenched in 0.2M NaOH and processed ad described above to cleave abasic sites.

In Vitro BER Reconstitution

Recombinant human proteins were used to reconstitute BER activity on either an I•T mismatch or an I-bulge. Reactions were incubated at 37˚C under that standard conditions of 50 mM HEPES (pH 7.5), 100 mM NaCl, 5 mM $MgCl_2$, 4 mM ATP, and 2.5 μM dNTP. Reactions contained 300 nM DNA substrate with 30 nM of AAG, APE1, polmerase β, and ligase I, as indicated in a total volume of 20 μL. After 60 minutes the

BER reaction was complete, at which point 3 μL aliquots were quenched in formamide/EDTA and analyzed as described above. We used this same protocol to follow the time course for multiple turnover BER activity, except that reactions contained 700nM DNA and the putative physiological concentration (29) of APE (120nM), polymerase β (20nM) , and ligase I (40nM).

RESULTS

Glycosylase-Catalyzed Excision of Bulged Nucleotides in HeLa WCE

Oligonucleotide duplexes were incubated in human WCE to test whether DNA glycosylases can excise unpaired nucleotides (Figure 2-2A). These oligonucleotides were 25 base pairs in length, with the central nucleotide unpaired (i.e. presented as a single nucleotide bulge). In addition to the four natural nucleotides, we also tested some common deaminated or alkylated nucleotides that are know to be present in chromosomal DNA. Deamination of C gives rise to uracil (U), deamination of A yields inosine (I), and alkylation of A can yield 1,N^6-etheno-A (εA). Glycosylases do not require Mg^{2+} ions for activity, therefor we included EDTA in the glycosylase assay buffer, in the absence of Mg^{2+}, most endonuclease and exonuclease activities were blocked and oligonucleotides remained intact for greater the 24h (Figure 2-2B).

There was no detectable glycoslyase activity toward undamaged nucleotides in single nucleotide bulges in HeLa WCE (Figure 2-2B, lanes 1-8). This could be due to extremely low tares of excision of natural nucleotides by human DNA glycosylases or the protection of these sites by other proteins present in the extract. In contrast, damaged nucleotides were rapidly excised (Figure 2-2B, lanes 9-14). Significant excision was detected at 1h, with complete processing of the bulge occurring with in 24h. If reactions

were quenched in formamide, instead of hydroxide, the 12-mer product was not observed. This confirmed that all three damaged nucleotides were excised by monofunctional DNA glycosylase (Appendix Figure A-2). Initial rates of base excision in the HeLa extracts show that the U-bulge was most rapidly removed (Figure 2-2C). The I-bulge was removed at a 4-fold slower rate and the εA-bulge was removed at a 20-fold slower rate than the U-bulge. We next sought to identify the enzymes responsible for this activity on single nucleotide bulges.

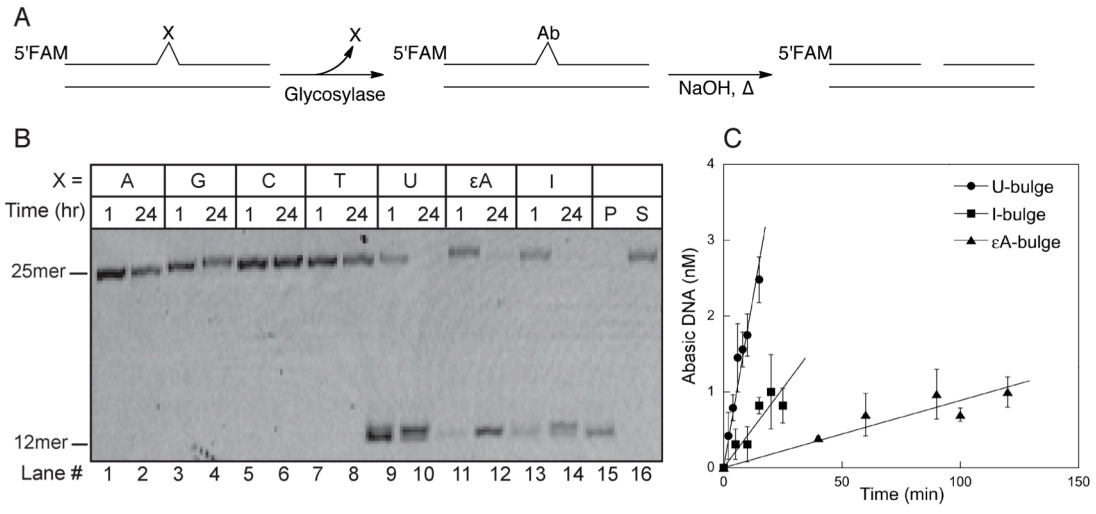

Figure 2-2. Human glycosylases excise alkylated and deaminated bases, but not undamaged bases, from one-nucleotide bulges. (A) DNA glycosylase activity was monitored in HeLa WCE using fluorescently labeled oligonucleotides that contained single nucleotide bugles. AP products were cleaved under alkaline conditions and samples were analyzed by denaturing PAGE. (B) A representative gel is shown. Undamaged (lanes 1-8) and damaged (lanes 9-14) DNA (10nM) was inclubated in 0.4 mg/ml of WCE for the indicated time. An I-T substrate was incubated in the presence or absence of recombinant AAG to provide size standards for the 25-mer intact substrate (S) and the 12-mer glycosylase generated product (P). (C) Initial rate for excision of damaged nucletides present in single nucleotide bulges were determined as described above. The mean standard deviation is shown from 3 independent experiments.

Identification of DNA Glycosylase Acting at Bulged Nucleotides – There are four human glycoslyases that are known to excise U from DNA (30). The most active of these is uracil DNA glycosylase (UDG) (31). UDG is inhibited with high affinity and specificity b the bacteriophage-encoded UDG inhibitory protein (UGI), therefore we tested if UGI would block the excision of a U-bulge in HeLa WCE (32). Complete inhibition was observed, demonstrating the UDG was the glycosylase responsible for acting on a U-bulge.

We expected that the I and εA-specific activities would both be attributed to alkyladenine DNA glycosylase (AAG), because this is the only human enzyme known to excise alkylated and deaminated purines from DNA (33-36). εA is formed by endogenous reactive species resulting from lipid peroxidation (37). Vinyl choloride exposure caused increased levels of εA via the P450 catalyzed production of chloroethylene oxide and chloroacetaldehyde (37, 38). AAG is the only protein found to bind εA•T DNA in cell extracts from human tissues (39, 40) and it binds with very high affinity (41). Therefore, we tested if unlabeled εA•T DNA inhibits the excision of I and εA from single nucleotide bulges. As expected, the addition of 10-fold excess εA•T decreased the excision of εA from the bulge to undetectable levels, whereas undamaged competitor DNA (A•T) had no effect (Figure 2-3B). Similarly, the εA•T competitor completely blocked activity toward and I-containing single nucleotide bulge. This suggests that AAG is responsible for the excision of alkylated and deaminated purines present in a single nucleotide bulge.

Bulges containing damaged nucleotides can be formed when a DNA polymerase encounters a bulky lesion, such as εA, on the template strand. The Y-family polymerases are particularly prone to template slipping. For example, frequent -1 slips are observed during replication of εA –DNA by human polymerases μ, ϰ, and η (43, 44) and during replication of εG-DNA by Dpo4 (45, 46). However, other types of damaged nucleotides are not known to induce polymerase slipping. For these types of damage, which includes U and I, we considered the possibility that they might form subsequent to DNA replication.

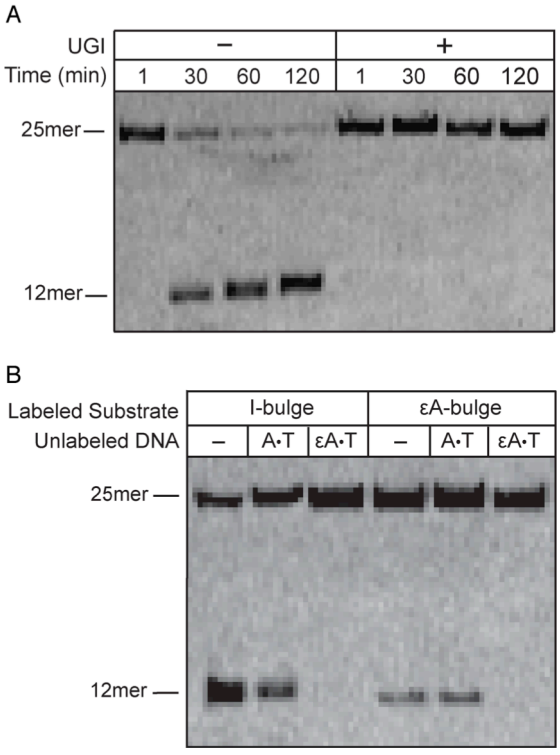

Figure 2-3. Deaminated bulged bases are excised by UDG and AAG. Glycosylase assays in HeLa cell extracts were performed as described in the legend to Figure 2-2. (A) U-bulge DNA was incubated in the presence or absence of UGI (0.02 units). UGI completely blocks U-bulge glycosylase activity, demonstrating that UDG is the enzyme responsible. (B) Fluorescently labeled I-bulge or εA-bulge DNA was incubated with no competitor DNA (-), undamaged DNA (A-T), or damaged DNA (εA-T) for 24 h. AAG is the only human glycosylase known to excise εA lesions, therefore the strong competition by εA-containing DNA suggests that AAG is responsible for both activities.

Unpaired Nucleotides are More Susceptible to Damage

Single nucleotide bulges formed by polymerase slipping events are more exposed than Watson-Crick paired nucleotides (47, 48), suggesting thea bulged nucleotides would also be damaged by oxidation and alkylation. We tested whether a bulged A is more susceptible to alkylative damage than an A•T base pair. As described above, alkylation of DNA by chloroacetaldehde results in the formation of a variety of alkylation adducts, including εA (Fig. 4A). Thus, alkylation of an A-bulge would create a good substrate for AAG.

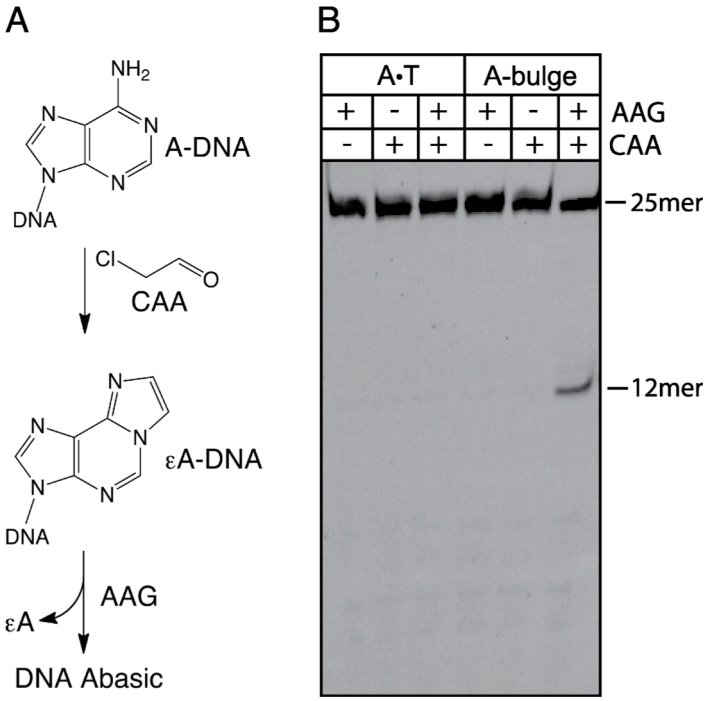

Figure 2-4. Single nucleotide bulges are highly susceptible to alkylation damage. (A) chloroacetaldehyde (CAA) reacts A nucleotides in DNA to produce the alkylation adduct εA. Sites of εA formation were detected by the glycosylase activity of AAG. (B) Stably paired (A-T) or a single nucleotide A-bugle was incubated with chloroacetaldehdye (100mM) for 4 h. After AAG-catalyzed excision of εA lesions alkaline hydrolysis of AP sites, samples were analyzed by denaturing PAGE. Only the A-bulge showed detectable εA formation under these conditions. Quantification of duplicate experiments revealed that 18% of the A-bugle is alkylated. This indicates that the A-bulge is at least 18-fold more reactive than an A-T pair.

We exposed a fully paired duplex (A•T) and a single A-bulge to a high concentration of chloroacetaldehyde (100nM) for 4h. Excess alkylating agent was removed and the DNA was incubated with recombinant AAG in the standard glycosylase assay. Both alkylating agent and AAG were required to get the specific 12-mer product resulting from the glycosylase reaction at this bulged site (Figure 2-4). Additional experiments showed that this alkylation reaction is dependent upon the concentration of chloroacetaldehyde and on the amount of time that DNA was exposed (Appendix Figure B-4). We observed that the bulged A has ≥18-fold increased reactivity to a paired A (Figure 2-4 and Appendix Figure B-4).

Deletion of an I-Bulge in Cell Extracts

Above we demonstrated that DNA glycosylases catalyze the excision of lesions from a single nucleotide bulge to create a bulged AP site. If this intermediate can be processed b the major human APE1, then the BER pathway, including dNTPs, ATP and Mg^{2+}, to HeLa WCE to test if the predicted one nucleotide deletion could be detected with either an I-bulge or a U-bulge substrate. Under these conditions, we observed a variety of BER intermediates that were consistent with the processing of these deaminated single nucleotide bulges by BER (Appendix Figure B-7). However, we observed a substantial level of 3'-5' exonuclease activity in HeLa WCE that degraded in DNA oligonucleotides even when the BER pathway was inhibited. It should not be noted that this exonuclease activity is an artifact of using short oligonucleotides substrates and is not expected to play a role when bulged nucleotides are present in chromosomal DNA. This exonucleolytic activity could be decreased by adding unlabeled competitor DNA (undamaged A•T 25-mer), presumably by providing additional DNA ends that would be

substrates for the hydrolytic enzyme(s). We infer that there was insufficient glycosylase activity for the BER pathway to fully outcompete the degradation pathway(s) in HeLa WCE when short oligonucleotides are used as substrates. Therefore, we investigated several additional human cell lines in which the balance of BER and exonucleolytic degradation might differ.

Many cancer cell lines, especially breast cancers, have increased AAG expression (49, 50). We chose one of these cell lines (T47D), a ductal breast tumor, because AAG is reported to be expressed at a 3-fold higher level relative to HeLa cells (50). We first measured the level of glycosylase activity toward an I•T mismatch in HeLa and T47D WCE and confirmed the 3-fold higher level of AAG activity in T47D (Appendix Figure B-5). We next estimated the repair of I•T and I-bulge DNA in T47D extracts under the BER conditions. The I•T mismatch was processed to build up a nicked DNA intermediate (Figure 2-5A). The presence of the repaired 25-mer product was confirmed by treating it with excess recombinant AAG and alkaline cleavage, which fully cleaved any remaining lesion-containing DNA (Figure 2-5B, lane 2). In the case of the I-bulge, a new 24-mer product was formed (Figure 2-5B, lanes 5-8) that was similarly resistant to cleavage by recombinant AAG (Figure 2-5B, lane 4). This suggests that WCE are capable of catalyzing a single nucleotide deletion with and I-bulge on a similar time scale to repair of an I•T mismatch. When εA•T competitor was added there was no repair of the I•T or I-bulge (Figure 2-5B, lanes 5-8). We also noticed small amounts of polymerase extension products under these conditions. These products were greatly reduced by aphidicolin, an inhibitor of polymerase μ, ϰ, and η, and did not occur if dNTPs were omitted from the reaction (Appendix Figures B-10 and B-11).

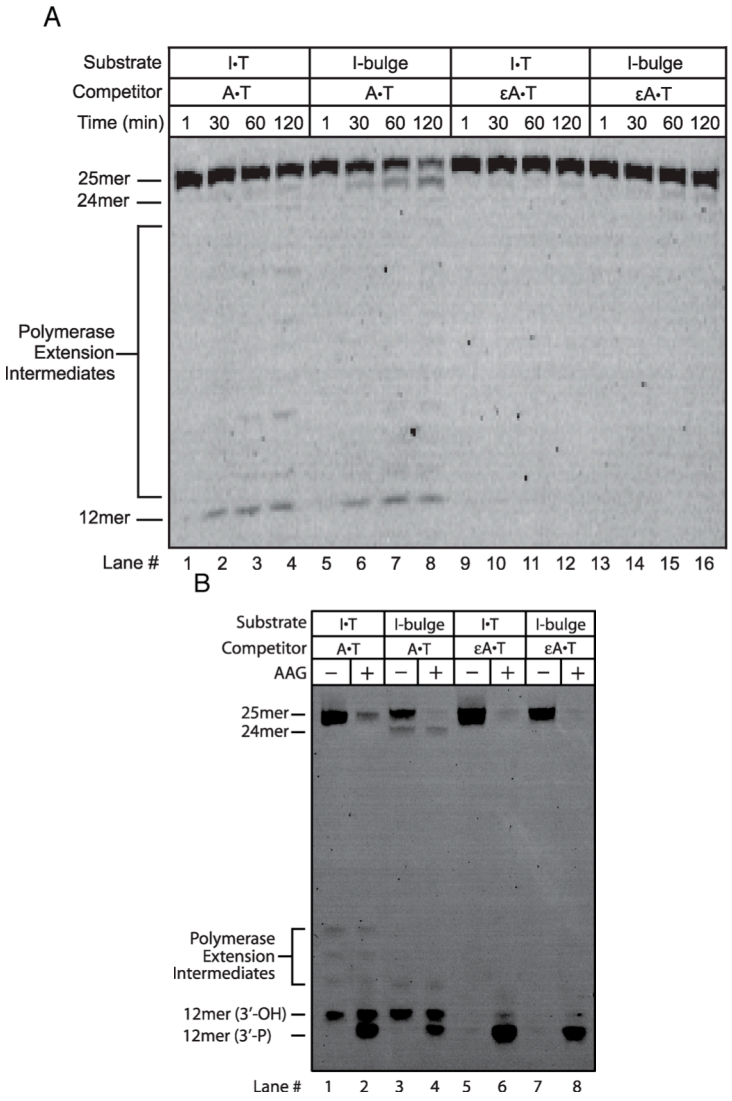

Figure 2-5. Base excision repair assays in T47D WCE demonstrates single nucleotide deletion. Conditions were similar to those in Figure 2-2, but 5 mM $MgCl_2$, 1 mM ATP, and 0.1 mM of the four dNTPs were added to support BER activity. The expected single nucleotide deletion product (24mer) was detected within 30 minutes. The I•T mismatch was processed on the same time scale, as evidence by the build-up and disappearance of BER intermediates. No BER intermediates or deletion products were observed when εA•T-containing unlabeled competitor DNA was included in the incubation.(B) Determination of extent of repair for reactions shown in (A). After 120 min samples were split and either incubated with recombinant AAG (+AAG) to process the remaining I lesions, or immediately quenched in formamide/EDTA (-AAG). Alkaline hydrolysis of the abasic site occurs via β,δ-elimination to generate a 3'-phosphate (3'-P), making it possible to quantify the amount of I lesion present after incubation in WCE. In contrast, the APE1-catalyzed reaction in the extract yields a 3'-hydroxyl (3'-OH). The persistent 25-mer in lane 2 and 24-mer in lane 4 confirms that the I lesion was removed.

This proposed single nucleotide deletion of a bulged lesion is depicted in Figure 2-6, by analogy to the normal BER pathway. Only the short patch pathway is shown, but

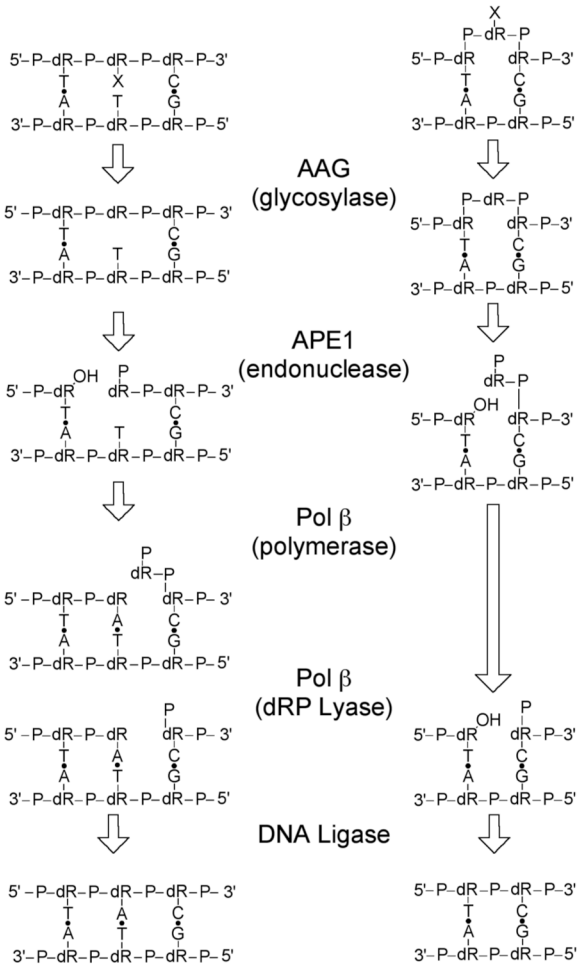

Figure 2-6. Proposed single nucleotide deletion pathway catalyzed by BER enzymes. The BER pathway for AAG-initiated short patch is shown on the left and the proposed pathway for single nucleotide deletion is shown on the right. Note that only the glycosylase and endonuclease reactions differ for the two pathways. After APE1 cleavage the 5'dRP intermediate is chemically identical to the intermediate generated after single nucleotide incorporation in the short-patch BER pathway.

long patch BER with strand displacement, flap cleavage by FEN1, and ligation will also form the single nucleotide deletion. The key steps are initiation of the pathway by a DNA repair glycosylase, as we have shown for AAG and UDG, and the subsequent hydrolysis of the abasic bulge by APE1, as shown below, or by AP lyases (data not shown). To

confirm the proposed pathway and to evaluate the relative rate of single nucleotide deletions we employed in vitro reconstitution using recombinant human proteins.

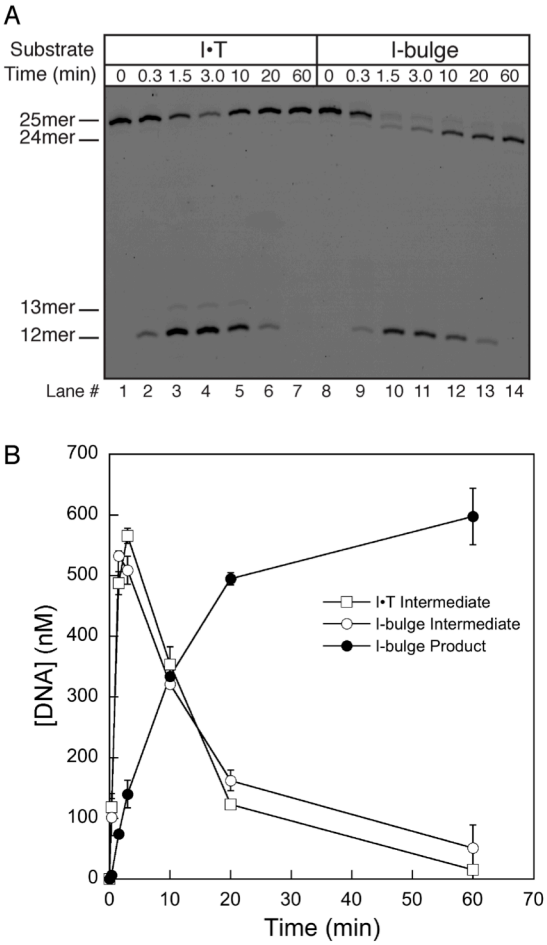

Figure 2-7. Time dependence of repair demonstrates that the single nucleotide deletion pathway occurs at the same rate as the BER pathway. A, multiple turnover reconstituted BER reactions contained 700 nM I mismatch or bulge DNA, 70 nM AAG, 120 nM APE1, 20 nM pol b, and 50 nM DNA Ligase I and the necessary nucleotide and Mg^{2+} cofactors (see Methods for details). Reactions were quenched at the indicated time with formamide/EDTA loading buffer and analyzed by denaturing PAGE. B, the level of the 12mer nicked DNA intermediate was quantified from 3 independent experiments and the mean ± SD is shown.

Reconstitution of Single Nucleotide Deletion by Enzymes of the Short Patch BER Pathway

We examined steady state multiple turnover repair reactions of either an I•T mismatch or an I bulge under the standard BER assay conditions with the putative physiological levels of the human proteins, AAG, APE1, pol b, and DNA ligase I (43). A

representative time course is shown in which reactions were quenched in formamide/EDTA and analyzed by denaturing PAGE (Figure 2-7). An early time point confirms that this quench is sufficient under these conditions (lane 1 & 6). Within 5 minutes a nicked DNA intermediate could be detected that builds up and breaks down, so that by 60 minutes the repair is complete (corresponding to ≥10 turnovers). The quantitative conversion to a single nucleotide deletion product (24mer, lane 10) demonstrates that all enzymes are active and we infer that the 25mer species in the repair of the I•T mismatch is the corrected A•T product. Quantification of the nicked DNA intermediate generated with a mismatch and with a bulge indicates that both substrates are repaired with similar rates (Figure 2-7B).

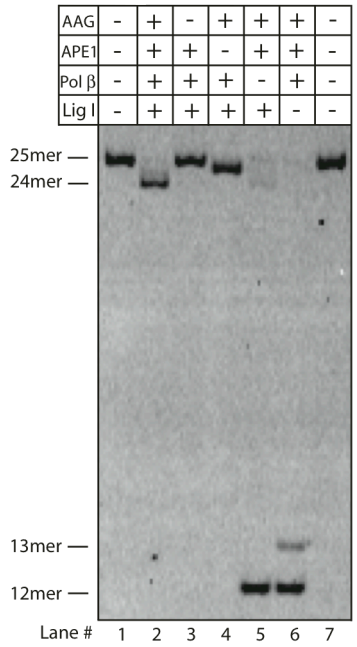

Figure 2-8. In vitro reconstitution of the single nucleotide deletion pathway with recombinant human BER proteins. Recombinant human AAG, APE1, pol β, and DNA lig I were both necessary and sufficient for the single nucleotide deletion in vitro (lane 2), and omission of any single protein gave the expected product (Figure 2-5). Note that the abasic intermediate migrates between the 25mer substrate and the 24mer product (lane 4) and the omission of DNA ligase (lane 5) leads to some strand displacement by DNA pol β.

To confirm that these four proteins are necessary and sufficient for the single nucleotide deletion (Figure 2-7), we omitted each protein in turn (Figure 2-8). After an hour incubation the single nucleotide deletion was complete (lane 2). No detectable reaction occurred in the absence of AAG (lane 3), whereas the omission of APE1 led to the abasic DNA intermediate with slightly increased mobility (lane 4). This is consistent with the extremely weak AP lyase activity of pol β (44). In the absence of pol β, the bulged substrate was processed to a single strand nick, the ligation of which is blocked by the 5'dRP group (lane 5). Finally, omission of DNA ligase I prevented DNA ligation. In the absence of ligase, pol β incorporates an additional nucleotide (lane 6). These experiments demonstrate that the human BER enzymes efficiently catalyze the deletion of a damaged nucleotide bulge, and that this reaction occurs on the same time scale as the well-characterized short patch BER pathway for the repair of damaged nucleotides present as a mismatch (Figure 2-8).

DISCUSSION

Previous reports that imbalanced levels of BER enzymes increase the frequency of frameshift mutations (24,25) led us to test if the BER pathway acts on single nucleotide bulges. Such bulges can occur by a variety of mechanisms (Figure 2-1). Using cell extracts and reconstitution with recombinant proteins, we show that deaminated and alkylated nucleotides present in single nucleotide bulges are good substrates for BER, resulting in rapid deletion of the damaged nucleotide. These observations suggest that competition between repair pathways can affect the frequency of frameshift mutations.

Glycosylases Excise Damaged Nucleotides Present in Single Nucleotide Bulges

We identified UDG and AAG as two glycosylases that efficiently recognize base lesions presented as a single nucleotide bulge. The activity of UDG was not too surprising, because it has robust activity towards U in both single- and double-stranded DNA. In contrast, the activity of AAG towards I in a single stranded substrate is ~10^3-fold lower than towards I in a double stranded substrate. We have used single turnover kinetics to show that AAG has a 3-fold preference for an I bulge over an I•T mismatch (31). These observations demonstrate that the UDG and AAG active sites are very flexible and able to recognize their substrate nucleotides in very different contexts.

Similar to UDG, two human homologs of *E. coli* EndoVIII , NEIL1 and NEIL2, show substantial activity towards oxidized bases in single-stranded DNA (45). NEIL2 has ~5-fold greater activity towards a lesion present in a single nucleotide bulge than the same lesion in single-stranded DNA (46). Other DNA glycosylases show much stronger preference for duplex substrates, including a recently characterized glycosylase from *M. tuberculosis*. In initial experiments this enzyme was shown to excise U from a single nucleotide bulge, albeit at a reduced level of activity relative to U in a mismatch (47). Several glycosylases use specific recognition of the opposing base to ensure high fidelity repair. For example, MutY and its human homolog excises an A preferentially from an A•8-oxoG pair and OGG1 excises 8-oxoG from an 8-oxoG•C pair. On the one hand, this specific recognition may prevent these glycosylases from accepting single nucleotide bulges. On the other hand, a single nucleotide bulge may be more easily flipped by the enzyme.

There was no detectable glycosylase activity towards normal nucleotide bulges in HeLa cell extracts under multiple turnover conditions. However, it has previously been shown that AAG exhibits low levels of activity towards the normal nucleotides A and G, and this activity is greater when they are present in a mismatch (19,21). Under single turnover conditions we found that AAG has similar activity towards undamaged purines as towards purines present in a mismatch (data not shown). Therefore, in MMR deficient cells there is the possibility that AAG could act directly on bulged sites resulting from DNA replication. This would serve to bias frameshift mutations to -1 events, because both daughter cells would be corrected in the case of +1 slipping events and both daughter cells would be mutant in the case of -1 slipping events. A more likely scenario is that spontaneous damage to a bulged nucleotide would generate a good substrate for BER.

Bulged nucleotides are more susceptible to damage

The reactivity of single nucleotide bulges have not been extensively characterized. However, the nucleobases of single stranded DNA are much more reactive those in duplex DNA towards a variety of types of damage. For example, deamination of C to U occurs ~100-fold faster in single stranded DNA than in duplex DNA (48,49). It has also been noted that C mismatches have elevated spontaneous deamination rates relative to C•G pairs, and in some sequences the rates of deamination of the C mismatch approaches that of single stranded DNA (50). Here we have tested the reactivity of a bulged A towards a DNA alkylating agent. Under the conditions tested, chloroacetaldehyde has more than 15-fold greater reactivity towards the single nucleotide bulge relative to a stably paired nucleotide in the same sequence context (Figure 2-4).

The greater reactivity of a single nucleotide bulge towards endogenous and exogenous forms of base damage underscores the importance that polymerase slippage be corrected quickly.

Human BER Excises Damaged Nucleotides Present in Single Nucleotide Bulges

We tested whether the BER pathway can process glycosylase-initiated bulged abasic sites in human cell extracts to effectively delete the damaged nucleotide. In extracts from a human breast cancer cell line (T47D), we observed the expected single nucleotide deletion product (Figure 2-5). The most persistent intermediates are consistent with the short patch BER pathway, whereas some intermediates can be detected that would correspond to the long patch BER pathway. Both pathways give the same single nucleotide deletion. To further test the proposed model, we reconstituted the pathway using recombinant proteins (Figures 2-7 and 2-8). This confirms the proposed pathway (Figure 2-6) and demonstrates that APE1 efficiently processes a bulged abasic site. Abasic site recognition by APE1 involves flipping of the abasic site into the active site pocket where endonucleolytic cleavage 5' of the abasic site can take place (51). APE1 is able to process abasic sites in single stranded DNA (52) and it can also catalyze endonucleolytic cleavage at bulky lesions, such as a-anomeric nucleotides (53). The finding that APE1 acts on a single nucleotide bulge is further evidence for a great flexibility in substrate recognition.

It has recently been demonstrated that Neil1, a bifuctional DNA glycosylase/AP lyase, is able to perform both glycosylase and lyase activities on a single nucleotide bulge (46). In this case the action of polynucleotide kinase 3'phosphatase is required to generate a ligatable nick (46,54).

Competition for Repair of Bulged Nucleotide Intermediates

Our model for BER-mediated frameshift mutations is illustrated in Figure 2-9. In Figure 2-9A, replication across a lesion can allow for polymerase bypass producing a 1-nt slipping error and a damaged, bulged nucleotide product. MMR provides high fidelity repair of bulged intermediates, however, BER activity would resolve the bulge to a 1-nt deletion. One example in which this model is likely to take place is in the replication of abasic sites. If an abasic template is encountered, then it is know that pol β will, with high frequency, slip past. We propose it does this by first generating a bulged abasic site that is subsequently processed by APE1, pol β and either DNA ligase I or DNA ligase III. This could nicely account for the finding that overexpression of pol b gives rise to greater frequency of -1 frameshift events (16).

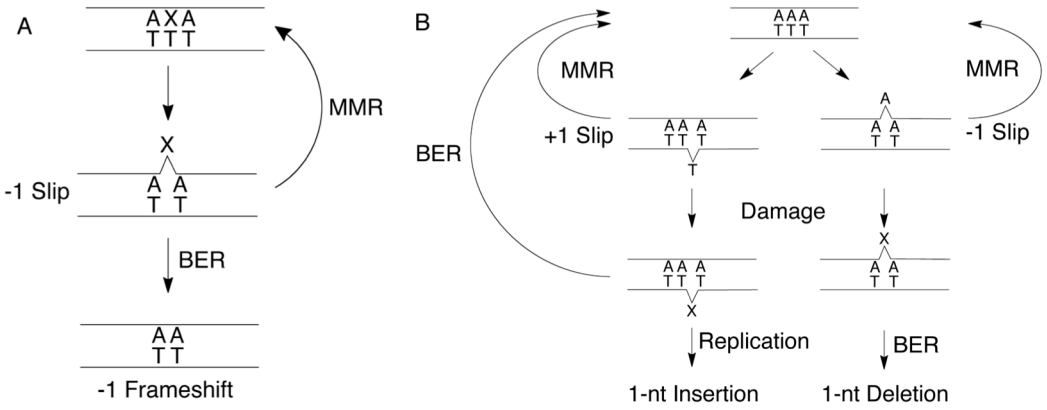

Figure 2-9. Model for competition between MMR and BER in the processing of nascent frameshift mutations. An example is shown in which a polymerase slips on an undamaged homopolymeric run, resulting in a -1 or +1 event (See Figure 2-1 for alternative mechanisms by which bulged intermediates are formed). Subsequent damage allows it to be recognized by BER, which will repair a +1 event but make a -1 event permanent. MMR provides a high fidelity pathway for repairing either -1 or +1 events, but imbalances in MMR or BER is expected to increase the frequency of frameshift mutations.

Alternatively, a polymerase could slip on an undamaged template to give rise to an undamaged bulge. MMR can repair both -1 and +1 slips to prevent these from becoming frameshift mutations. However, if the bulge remains undetected then it is susceptible to DNA damage. We have shown that for at least some types of oxidative and alkylative damage, BER can potentially compete with MMR. This competition will favor BER if there is a high burden of DNA damage or if the levels of repair activities are imbalanced (i.e., depletion of MMR enzymes or overexpression of BER enzymes). In the case of a nascent +1 slip, this cycle of damage and deletion via BER would restore the original DNA sequence. However, a nascent -1 slip would be converted into a -1 frameshift and would no longer be available to another pathway such as MMR.

Consistent with our hypothesis, there is abundant evidence that human cancer cells show a strong preference for -1 versus +1 frameshifts. Most directly, Samson and coworkers demonstrated that overexpression of AAG in yeast leads to a 30-fold increase in the frequency of -1 frameshifts at an A tract site (24). This observation is not easily explained by the established models of low fidelity BER initiated at normal sites, but is fully consistent with the model proposed herein. Intriguingly, several studies found AAG over-expression gives rise to increased mutation rates. AAG expression is commonly up-regulated in breast cancers and previous studies have found that these same cell lines exhibit elevated levels of -1 frameshift mutations, in addition to elevated levels of point mutations. A large body of work suggests that imbalanced BER can lead to increased frequency of point mutations, presumably due to errors in lower fidelity polymerases or two errors in replicating a repair intermediate. Our studies provide a plausible model to account for frameshift mutations under the situation of imbalanced BER.

REFERENCES

1. Loeb, L.A., *A mutator phenotype in cancer.* Cancer Res, 2001. 61(8): p. 3230-3239.
2. Beckman, R.A. and L.A. Loeb, *Efficiency of carcinogenesis with and without a mutator mutation.* Proc Natl Acad Sci U S A, 2006. 103(38): p. 14140-14145.
3. Bebenek, K., J.D. Roberts, and T.A. Kunkel, *The effects of dNTP pool imbalances on frameshift fidelity during DNA replication.* J Biol Chem, 1992. 267(6): p. 3589-3596.
4. Bebenek, K. and T.A. Kunkel, *Frameshift errors initiated by nucleotide misincorporation.* Proc Natl Acad Sci U S A, 1990. 87(13): p. 4946-4950.
5. Shearman, C.W., M.M. Forgette, and L.A. Loeb, *On the fidelity of DNA replication. Mechanisms of misincorporation by intercalating agents.* J Biol Chem, 1983. 258(7): p. 4485-4491.
6. Kunkel, T.A. and D.A. Erie, *DNA Mismatch Repair.* Annu Rev Biochem, 2004.
7. Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita, E. Terzaghi, and M. Inouye, *Frameshift mutations and the genetic code. This paper is dedicated to Professor Theodosius Dobzhansky on the occasion of his 66th birthday.* Cold Spring Harb Symp Quant Biol, 1966. 31: p. 77-84.
8. Colgin, L.M., A.F. Hackmann, M.J. Emond, and R.J. Monnat, Jr., *The unexpected landscape of in vivo somatic mutation in a human epithelial cell lineage.* Proc Natl Acad Sci U S A, 2002. 99(3): p. 1437-1442.
9. Taylor, M.S., C.P. Ponting, and R.R. Copley, *Occurrence and consequences of coding sequence insertions and deletions in Mammalian genomes.* Genome Res, 2004. 14(4): p. 555-566.
10. Frosina, G., *Overexpression of enzymes that repair endogenous damage to DNA.* Eur J Biochem, 2000. 267(8): p. 2135-2149.
11. Fortini, P. and E. Dogliotti, *Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways.* DNA Repair (Amst), 2007. 6(4): p. 398-409.
12. O'Brien, P.J. and T. Ellenberger, *Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines.* Biochemistry, 2003. 42(42): p. 12418-12429.
13. Pascal, J.M., P.J. O'Brien, A.E. Tomkinson, and T. Ellenberger, *Human DNA ligase I completely encircles and partially unwinds nicked DNA.* Nature, 2004. 432(7016): p. 473-478.
14. Baldwin, M.R. and P.J. O'Brien, *Human AP Endonuclease I Stimulates Multiple-Turnover Base Excision by Alkyladenine DNA Glycosylase.* Biochemistry, 2009. 48(25): p. 6022-6033.
15. Werneburg, B.G., J. Ahn, X. Zhong, R.J. Hondal, V.S. Kraynov, and M.D. Tsai, *DNA polymerase beta: pre-steady-state kinetic analysis and roles of arginine-283 in catalysis and fidelity.* Biochemistry, 1996. 35(22): p. 7041-7050.
16. Pascal, J.M., O.V. Tsodikov, G.L. Hura, W. Song, E.A. Cotner, S. Classen, A.E. Tomkinson, J.A. Tainer, and T. Ellenberger, *A Flexible Interface between DNA Ligase and PCNA Supports Conformational Switching and Efficient Ligation of DNA.* Mol Cell, 2006. 24(2): p. 279-291.

17. Lyons, D.M. and P.J. O'Brien, *Efficient recognition of an unpaired lesion by a DNA repair glycosylase.* J Am Chem Soc, 2009. 131(49): p. 17742-17743.
18. Abner, C.W., A.Y. Lau, T. Ellenberger, and L.B. Bloom, *Base excision and DNA binding activities of human alkyladenine DNA glycosylase are sensitive to the base paired with a lesion.* J Biol Chem, 2001. 276(16): p. 13379-13387.
19. Drablos, F., E. Feyzi, P.A. Aas, C.B. Vaagbo, B. Kavli, M.S. Bratlie, J. Pena-Diaz, M. Otterlei, G. Slupphaug, and H.E. Krokan, *Alkylation damage in DNA and RNA--repair mechanisms and medical significance.* DNA Repair (Amst), 2004. 3(11): p. 1389-1407.
20. Kavli, B., O. Sundheim, M. Akbari, M. Otterlei, H. Nilsen, F. Skorpen, P.A. Aas, L. Hagen, H.E. Krokan, and G. Slupphaug, *hUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup.* J Biol Chem, 2002. 277(42): p. 39926-39936.
21. Wolfe, A.E. and P.J. O'Brien, *Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by human alkyladenine DNA glycosylase.* Biochemistry, 2009. 48(48): p. 11357-11369.
22. Patel, D.J., S.A. Kozlowski, L.A. Marky, J.A. Rice, C. Broka, K. Itakura, and K.J. Breslauer, *Extra adenosine stacks into the self-complementary d(CGCAGAATTCGCG) duplex in solution.* Biochemistry, 1982. 21(3): p. 445-451.
23. Joshua-Tor, L., D. Rabinovich, H. Hope, F. Frolow, E. Appella, and J.L. Sussman, *The three-dimensional structure of a DNA duplex containing looped-out bases.* Nature, 1988. 334(6177): p. 82-84.
24. Okochi, E., N. Watanabe, T. Sugimura, and T. Ushijima, *Single nucleotide instability: a wide involvement in human and rat mammary carcinogenesis?* Mutat Res, 2002. 506-507: p. 101-111.
25. Sokhansanj, B.A., G.R. Rodrigue, J.P. Fitch, and D.M. Wilson, 3rd, *A quantitative model of human DNA base excision repair. I. Mechanistic insights.* Nucleic Acids Res, 2002. 30(8): p. 1817-1825.
26. Dou, H., S. Mitra, and T.K. Hazra, *Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2.* J Biol Chem, 2003. 278(50): p. 49679-49684.
27. Srinath, T., S.K. Bharti, and U. Varshney, *Substrate specificities and functional characterization of a thermo-tolerant uracil DNA glycosylase (UdgB) from Mycobacterium tuberculosis.* DNA Repair (Amst), 2007. 6(10): p. 1517-1528.
28. Lindahl, T. and B. Nyberg, *Heat-induced deamination of cytosine residues in deoxyribonucleic acid.* Biochemistry, 1974. 13(16): p. 3405-3410.
29. Frederico, L.A., T.A. Kunkel, and B.R. Shaw, *Cytosine deamination in mismatched base pairs.* Biochemistry, 1993. 32(26): p. 6523-6530.
30. Prasad, R., W.A. Beard, P.R. Strauss, and S.H. Wilson, *Human DNA polymerase beta deoxyribose phosphate lyase. Substrate specificity and catalytic mechanism.* J Biol Chem, 1998. 273(24): p. 15263-15270.
31. Hofseth, L.J., M.A. Khan, M. Ambrose, O. Nikolayeva, M. Xu-Welliver, M. Kartalou, S.P. Hussain, R.B. Roth, X. Zhou, L.E. Mechanic, I. Zurer, V. Rotter, L.D. Samson, and C.C. Harris, *The adaptive imbalance in base excision-repair*

enzymes generates microsatellite instability in chronic inflammation. J Clin Invest, 2003. 112(12): p. 1887-1894.
32. Glassner, B.J., L.J. Rasmussen, M.T. Najarian, L.M. Posnick, and L.D. Samson, *Generation of a strong mutator phenotype in yeast by imbalanced base excision repair.* Proc Natl Acad Sci U S A, 1998. 95(17): p. 9997-10002.
33. Zhao, X., N. Krishnamurthy, C.J. Burrows, and S.S. David, *Mutation versus Repair: NEIL1 Removal of Hydantoin Lesions in Single-Stranded, Bulge, Bubble, and Duplex DNA Contexts.* Biochemistry.
34. Berdal, K.G., R.F. Johansen, and E. Seeberg, *Release of normal bases from intact DNA by a native DNA repair enzyme.* Embo J, 1998. 17(2): p. 363-367.
35. O'Brien, P.J. and T. Ellenberger, *Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase.* J Biol Chem, 2004. 279(11): p. 9750-9757.
36. Frederico, L.A., T.A. Kunkel, and B.R. Shaw, *A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy.* Biochemistry, 1990. 29(10): p. 2532-2537.
37. Mol, C.D., T. Izumi, S. Mitra, and J.A. Tainer, *DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination [corrected].* Nature, 2000. 403(6768): p. 451-456.
38. Marenstein, D.R., D.M. Wilson, 3rd, and G.W. Teebor, *Human AP endonuclease (APE1) demonstrates endonucleolytic activity against AP sites in single-stranded DNA.* DNA Repair (Amst), 2004. 3(5): p. 527-533.
39. Gros, L., A.A. Ishchenko, H. Ide, R.H. Elder, and M.K. Saparbaev, *The major human AP endonuclease (Ape1) is involved in the nucleotide incision repair pathway.* Nucleic Acids Res, 2004. 32(1): p. 73-81.
40. Karimi-Busheri, F., J. Lee, A.E. Tomkinson, and M. Weinfeld, *Repair of DNA strand gaps and nicks containing 3'-phosphate and 5'-hydroxyl termini by purified mammalian enzymes.* Nucleic Acids Res, 1998. 26(19): p. 4395-4400.
41. Chan, K., S. Houlbrook, Q.M. Zhang, M. Harrison, I.D. Hickson, and G.L. Dianov, *Overexpression of DNA polymerase beta results in an increased rate of frameshift mutations during base excision repair.* Mutagenesis, 2007. 22(3): p. 183-188.

APPENDIX A

Additional data figures to support Chapter 2

SUPPORTING RESULTS

```
5'fam-CGATAGCATCCTXCCTTCTCTCCAT
   3'-GCTATCGTAGGAYCCAAGAGAGGTA

                   X
5'fam-CGATAGCATCCT CCTTCTCTCCAT
   3'-GCTATCGTAGGA GGAAGAGAGGTA
```

Figure A-1. Sequence of oligonucleotides that were used in this study. X position is varied between normal and damaged bases and contains a 5'-fluorescein attached by a 6-aminohexyl linker (fam). Y position is varied among normal bases.

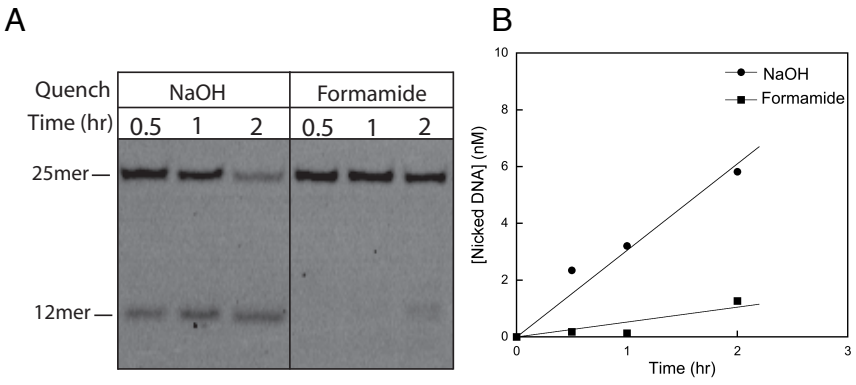

Figure A-2. (A) Monitoring of glycosylase activity on 10nM I-bulge substrate in 0.4 mg/mL HeLa WCE under glycosylase conditions produces a hydroxide cleavable abasic site and the resulting 12mer product. Simultaneous quenching with formamide/EDTA shows that the abasic site largely remains intact under these conditions, producing a small amount of enzymatic cleavage after two hours. This activity is likely due to lyase activity by bifunctional glycosylases. (B) Quantification of nicked intermediate (12mer).

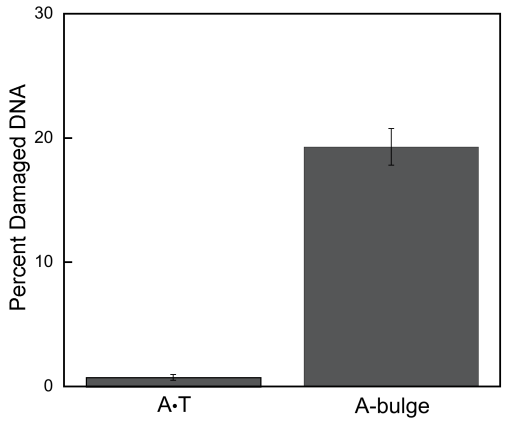

Figure A-3. Quantification of AAG excised CAA alkylation of 100nM DNA at the 12mer position from Figure 4 (lanes 3 and 6). Data is an average of three experiments with error bars representing one standard deviation from the mean.

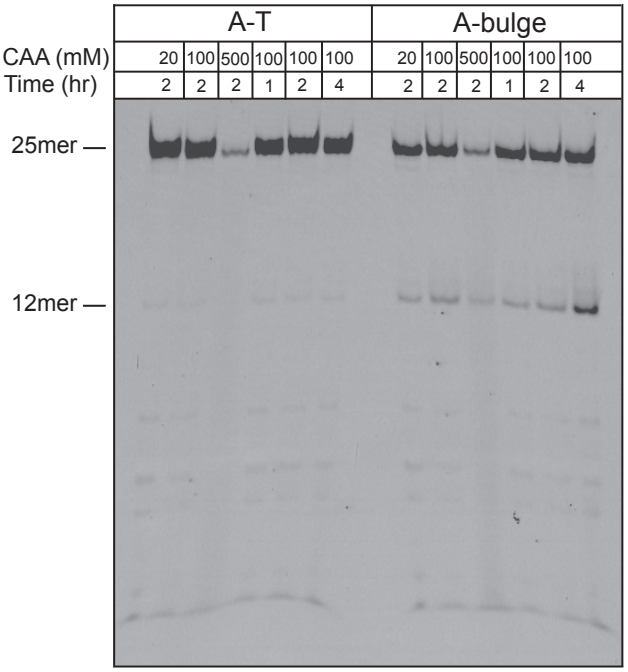

Figure A-4. Concentration and time dependence of chloroacetaldehyde (CAA) alkylation of A•T and A-bulge substrates was performed under similar conditions as Figure 4. 500nM DNA in a 50μL reaction volume was incubated with CAA for the time and concentration of CAA indicated. CAA was extracted twice with hydrated ether without a subsequent ethanol precipitation. 4μL of the CAA damaged DNA was incubated with 2μM AAG for 30 min at 37 °C in a 20μL reaction volume. Samples were analyzed according to the standard glycosylase assay.

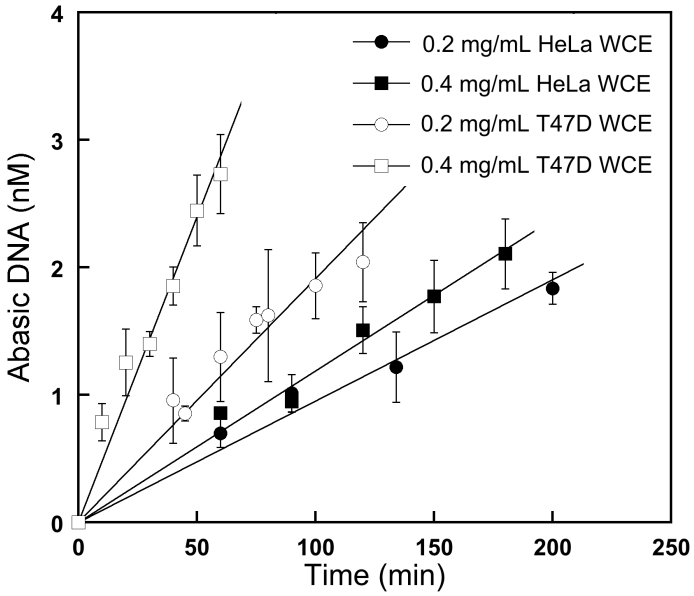

Figure A-5. Initial rates of AAG activity on 10nM I•T substrate under glycosylase conditions at two concentrations of HeLa and T47D WCE as represented in Figure 5. Each point represents the average of three experiments with error bars representing one standard deviation from the mean.

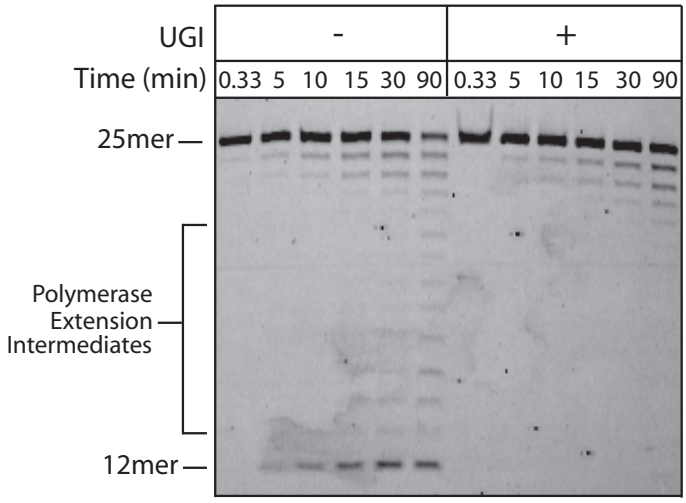

Figure A-6. Base excision repair assay performed in 0.4 mg/mL HeLa WCE with 10nM U-bulge substrate and 100nM unlabeled A•T DNA competitor. Production the expected nicked intermediate (12mer) from U-bulge substrate appears to be further processed by BER to produce polymerase extension intermediates. No BER intermediates were observed when UDG glycosylase activity was inhibited by 0.02 units of UGI, but 3'-5' exonuclease degradation is apparent.

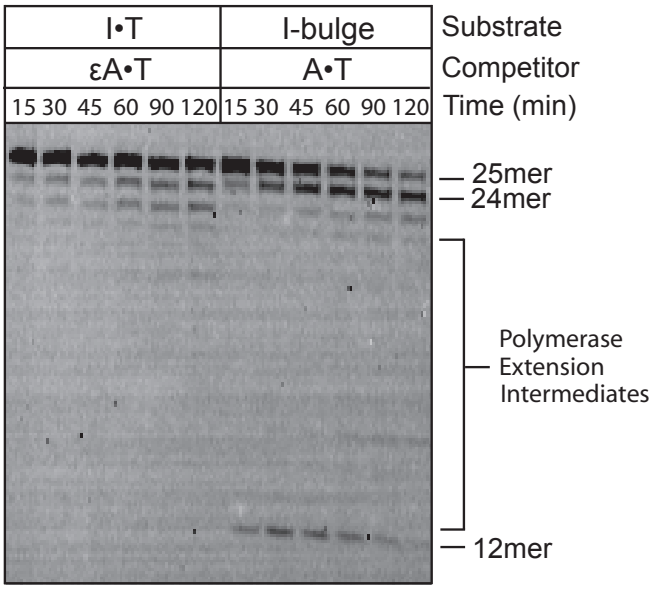

Figure A-7. Base excision repair assay performed in human colon adenocarcinoma nuclear extract. Under same conditions as Figure 6, 0.4 mg/mL HT-29 NE produces the expected nicked intermediate (12mer) that appears to be processed to the deletion product (24mer). No BER intermediates were observed when inhibited by unlabeled εA•T competitor, but 3'-5' exonuclease

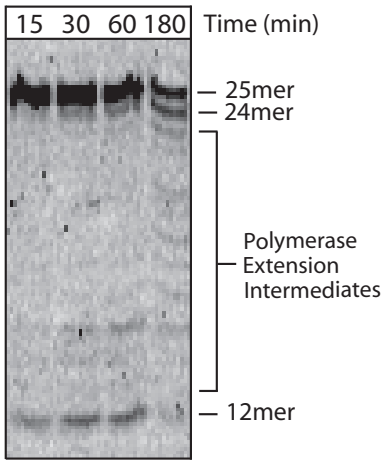

Figure A-8. Base excision repair assay performed in human breast ductal carcinoma nuclear extract, under conditions described for observing BER acitivty. 0.4 mg/mL MCF7 NE produces the expected nicked intermediate (12mer) from I-bulge substrate that is processed to the deletion product (24mer). Reaction contained 10-fold excess unlabeled A•T DNA competitor.

CHAPTER 3

EFFICIENT RECOGNITION OF AN UNPAIRED LESION BY A DNA REPAIR GLYCOSYLASE

The double-helical structure of DNA protects the nucleobases from oxidative and alkylative chemical damage [1]. However, this internal base pairing is also a barrier to the enzymes that recognize and repair DNA damage [2]. Enzymes that modify bases in DNA use nucleotide flipping to rotate the target nucleotide out of the duplex into an active site, but the energetic barrier that must be overcome is not fully understood. Unpaired (bulged) nucleotides are more accessible, but it is not know to what extent they are recognized by nucleotide-flipping enzymes. We have investigated this question with human alkyladenine DNA glycosylase (AAG). AAG recognizes a wide variety of structurally disparate lesions, including deoxyinosine (I), which results from the oxidative deamination of adenosine. AAG catalyzes the hydrolysis of the N-glycosidic bond to release the lesion base and initiate the base excision repair pathway (Figure 3-1A) [3]. We used single-turnover kinetics to demonstrate that AAG is exquisitely sensitive to the structural context of the I lesion. An inverse correlation between duplex stability and catalytic efficiency was observed, indicating that stable pairing is a barrier to recognition

by AAG. Single-stranded DNA is a very poor substrate, but a single-nucleotide bulge is recognized more efficiently than any other context. These results highlight the intrinsic barrier to nucleotide flipping that DNA repair enzymes face and imply that the recognition of DNA damage by AAG is remarkably plastic.

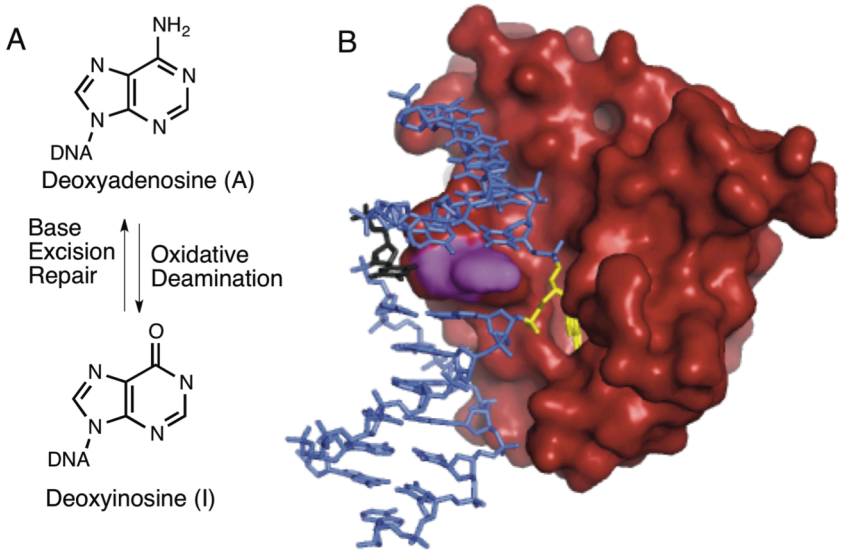

Figure 3-1. Deamination of A to form I and the structure of AAG bound to damaged DNA. (A) Oxidative deamination of A is reversed by AAG and the base excision repair pathway. (B) AAG is red, the intercalating tyrosine (Y162) is magenta, and the DNA is blue, except for the lesion (yellow) that is flipped into the active site and the opposing nucleotide (black). This opposing nucleotide is missing in the bulged substrate. Coordinates are from the Protein Data Bank structure for AAG bound to 1,N6-ethenoA-DNA (PDB entry 1EWN).

Crystal structures of AAG bound to an extrahelical lesion reveal that the DNA is bent, the lesioned base is rotated by ~180° out of the duplex into the active site pocket, and the hole that is left in the DNA is plugged with the phenolic group of Y162 (Figure 3-1B) [4]. Intimate protein-DNA contacts in this static structure seem to preclude the possibility that a bulge could be accommodated. However, the barrier to flipping of a bulged nucleotide is much less than that for a mismatch. Bulged purines favor an

intercalated conformation, but an extrahelical conformation has been observed in a crystal structure [5, 6]. If AAG has sufficient flexibility to encompass a bulged nucleotide, then the decreased barrier to flipping would enable more efficient recognition.

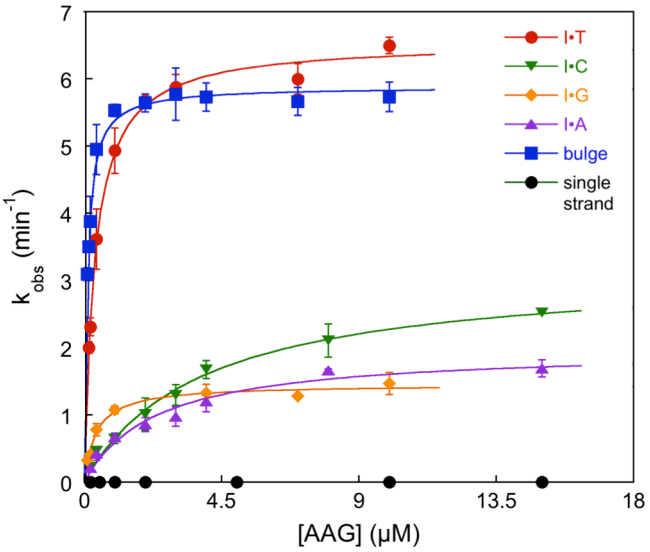

Figure 3-2. Concentration dependence for single-turnover glycosylase activity of AAG. Each data point corresponds to the average and standard deviation from 4-8 individual reactions (see the Supporting Information). Lines indicate the best fits to a hyperbolic concentration dependence: k_{obs}) $k_{max}[AAG]/(K_{1/2} + [AAG])$.

Therefore, we characterized the activity of AAG toward an I lesion in different structural contexts, including a single-nucleotide bulge. Product release is rate-limiting for multiple turnover reactions of AAG, requiring the use of single-turnover methods [7, 8]. The concentration of AAG was varied to determine the maximal single turnover rate constant (k_{max}) and the half-maximal concentration ($K_{1/2}$) for each context (Figure 3-2). Although the individual reaction rates varied by more than 10,000-fold, with half-times that varied from seconds to days, single-exponential fits were excellent in all cases (R^2>0.97; see the Supporting Information for representative data). The rate constant k_{max} reports on the flipping and N-glycosidic bond hydrolysis steps, and the $K_{1/2}$ value

corresponds to the K_d for substrate dissociation.[6] The specificity between competing substrates is given by the ratio $k_{max}/K_{1/2}$, the catalytic efficiency (analogous to k_{cat}/K_M in steady-state kinetics).

opposing base	$K_{1/2}$ (μM)	k_{max} (min^{-1})	$k_{max}/K_{1/2}$ (M^{-1} s^{-1})[b]	relative $k_{max}/K_{1/2}$
bulge (none)	0.08	5.4	1.1×10^6	3
T	0.30	6.3	3.5×10^5	(1)
G	0.41	1.5	6.2×10^4	0.2
C	3.8	3.3	1.4×10^4	0.04
A	2.3	2.0	1.4×10^4	0.04
single strand[c]	0.34	0.0035	1.7×10^2	0.0005

Table 3-1. Kinetic parameters for the AAG-catalyzed hydrolysis of inosine in different structural contexts. [a]Data were obtained at 23C in 50mM NaMES (pH 6.1), 1mM DTT , 0.1 mg/mL BSA. The ionic strength was adjusted to 200mM with NaCl. [b]$k_{max}/K_{1/2}$ is a measure of the catalytic efficiency, analogous to k_{cat}/K_m in steady-state kinetics. [c] This oligonucleotide does not appear to adopt secondary structure, because a single-stranded polyT oligonucleotide gave very similar kinetic parameters (see Appendix B).

The glycosylase activity of AAG is shown in Figure 3-2 and summarized in Table 3-1. The bulge was the best context tested, with a catalytic efficiency that is 3-fold better than that for the biological I•T mismatch. This result indicates that the increased ease in flipping the lesion more than compensates for any unfavorable effects of removing the opposing base (Figure 3-1B). Further work is needed to evaluate whether the bulge might be more easily bent than the mismatch and whether this could contribute to catalytic recognition. Other mismatches were recognized less efficiently than the I•T mismatch. The single-stranded lesion is even more accessible, and several glycosylases efficiently utilize single-stranded substrates; however we found that AAG has ~2000-fold reduced activity toward this substrate.

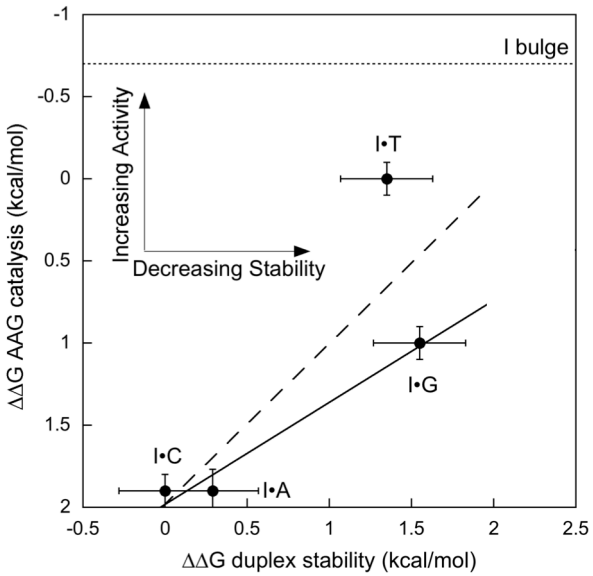

Figure 3-3. Linear free-energy relationship showing an inverse correlation between duplex stability and glycosylase activity. Differences in free energy (ΔΔG) are from ref 7 for duplex stability and from the equation ΔΔG) -RT ln(kmax/K1/2rel) for AAG activity. Linear fits to all of the mismatches (dashed line; slope) -0.97, R2) 0.67) and with the exclusion of the T mismatch (solid line; slope) -0.59, R2) 0.97) are shown. The relative activity toward a bulge is shown as a dotted line.

Since AAG does not make specific contacts with the opposing base, the relationship between base-pair stability and catalytic efficiency for different mismatches can be assessed. Thermodynamic parameters for duplex stability for inosine paired with each of the normal nucleotides are known [9] and serve as a surrogate for base-pair stability. The relative free energies (ΔΔG) for catalytic efficiency of the AAG-catalyzed reaction and for duplex stability are presented in Figure 3-3. A linear fit to all of the mismatches yields a slope of -0.96 (R^2 = 0.67), and omission of the I•T substrate gives a slope of -0.59 (R^2 = 0.97). This is a limited set of data, but the trend is clear. Previous analysis of the thermodynamics of duplex stability and DNA binding by uracil DNA glycosylase found that disruption of base pairing gives greater accessibility and tighter binding [10, 11]. The inverse relationship between duplex stability and efficiency of excision supports the model that AAG and other glycosylases

must overcome the barrier provided by base-pairing interactions. This behavior maximizes the discrimination between damaged and undamaged nucleotides because undamaged nucleotides have more favorable hydrogen-bonding and stacking interactions.

The I•T mismatch deviate by ~1 kcal/mol relative to the other mismatches, raising the possibility that AAG recognizes the I•T wobble pair geometry independent of its effects on duplex stability. Indeed, I•G forms a less stable duplex than I•T and yet it is removed with 5-fold lower catalytic efficiency. Deamination of A in DNA generates an I•T pair, and AAG-initiated repair restores the correct sequence. Incorporation of dIMP or replication of an I lesion predominantly forms an I•C pair [12, 13]. Under these scenarios the activity of AAG would make a permanent mutation. The much lower efficiency of AAG towards I•C leaves open the possibility that another DNA repair pathway corrects replicative events.

By quantifying the energetic differences in the catalytic efficiency of AAG with different mismatches, we obtained strong evidence that base pairing provides a barrier to base excision. The fact that a bulge is recognized with the same efficiency as a mismatch indicates considerable flexibility in DNA recognition. This also has biological ramifications, as DNA polymerases can slip on repetitive or damaged templates to generate bulged structures. Initiation of base excision repair could correct nascent +1 frameshifts but would make -1 frameshifts permanent. This ability of AAG to act on bulged nucleotides may explain the observation that an increased level of AAG expression is correlated with increased frequency of frameshift mutations [14].

MATERIALS AND METHODS

Expression and Purification of Recombinant AAG.

The catalytic fragment of AAG lacking the first 79 amino acids (Δ80) was expressed in *E. coli* and purified as previously described.[1] The truncated protein has a slightly decreased ability to translocate along DNA, relative to the full-length enzyme, but both proteins have identical glycosylase activity towards inosine and 1,N^6-ethenoadenosine lesions.[2,3] The concentration of active enzyme was determined by burst analysis as described below and the concentration of active enzyme was used throughout.

Oligonucleotide Substrates.

The sequences and annealed structures of the DNA oligonucleotides used in this study are given in Figure S1. The lesion-containing 25mer oligonucleotides have a centrally located inosine lesion and includes a 5'-fluorescein (6-fam) label. Complementary strands are unlabeled. The DNA oligonucleotides were synthesized by commercial sources using standard phosphoramidite chemistry, and were purified by denaturing polyacrylamide gel electrophoresis on a 20% polyacrylamide gel (6.6 M urea and Tris-borate buffer; 89 mM tris, 89 mM borate, 2 mM EDTA). Full-length oligonucleotides were excised from the gel, crushed, and soaked overnight in 500 mM NaCl and 1mM EDTA. Desalting was accomplished by reverse phase chromatography (C18 Sep-pak, Waters). Concentrations of single-stranded oligonucleotides were determined by absorbance at 260 nm using the calculated extinction coefficients. For glycosylase assays, oligonucleotides were annealed with a 1.5-fold excess of complementary strand, heated to 90˚C and subsequently cooled to 4˚C over ~15 min.

General Glycosylase Assay.

Samples were quenched with two volumes of 0.3 M NaOH to obtain a final concentration of 0.2 M. Abasic sites were quantitatively converted to DNA breaks by heating at 70˚C for 15 min, followed by the addition of 3.3 volumes of formamide/EDTA loading buffer that contained 0.05% w/v of both bromophenol blue and xylene cyanol FF as tracking dyes. DNA substrate (25mer) and cleaved product (12mer) oligonucleotides were separated on a 20% denaturing polyacrylamide gel containing 6.6 M urea. A typhoon trio imager was used to scan the gels with a 488 nm excitation and 520 nm long pass filter to detect the fluorescein labeled oligonucleotides. The substrate and product

bands were quantified with Image Quant TL (GE Healthcare) and the fraction product formed at each timepoint was calculated by [F = P/(P+S)], in which F is the fraction converted to product, P is the fluorescence of the product, and S is the fluorescence of the intact substrate. Control reactions in which enzyme was omitted revealed that inosine-containing oligonucleotides were stable to this procedure and no product was detected in the absence of glycosylase.

Burst Analysis to Determine Active AAG Concentration.

The active concentration of recombinant AAG was determined by burst analysis as previously described.[2] With a fixed concentration of 1μM of I•T 25mer substrate, which is far above the $K_{d,}$ the concentration of AAG was varied to obtain a burst of 5-20%. The fraction product was determined by polyacrylamide gel electrophoresis, as described above, and converted into concentration of product by multiplying by the concentration of initial substrate. The production of product followed an initial burst, followed by a slow multiple turnover rate. The reaction progress curve was fit by equation #1 using Kaleidagraph, in which P is product, A is the burst amplitude, k_{obs} is the burst rate constant, and V_{obs} is the steady state velocity. The burst amplitude gives the concentration of active enzyme.

$$[P] = A[1-\exp(-k_{obs}t)] + V_{obs}t \qquad (1)$$

Single Turnover Assay for Glycosylase Activity.

Single-turnover glycosylase assays were performed with AAG in excess over the 0.1 μM DNA substrate. The standard conditions were 23°C, 50 mM NaMES (pH 6.0), 1 mM EDTA, 1 mM DTT, 10% (v/v) glycerol, 0.1 mg/mL BSA, and the ionic strength was adjusted to 200 mM with NaCl. Typical reaction volumes were 20 uL, with 3 uL aliquots being removed at the desired time and quenched in NaOH and analyzed as described above. To determine the rate constant, time points were taken over the entire reaction progress curve and the fraction of product was plotted as a function of reaction time. Data analysis was performed using Kaleidagraph. In all cases, the reaction progress curves followed a single exponential and were fit by Equation #2, in which A is the fraction of substrate converted to product at completion ($A \geq 0.94$), k_{obs} is the observed single-

turnover rate constant, and t is the reaction time. Fits were excellent in all cases ($R^2 \geq$ 0.97).

$$F = A[1-\exp(-k_{obs}t)] \quad (2)$$

For extremely fast reactions, with half-lives of less than 20 s, some modifications to this protocol were made. Equal volumes (3 μL) of enzyme and substrate were mixed by hand, with the reaction occurring in the pipet tip, and subsequently quenched in a tube containing 6 μL of 0.3 M NaOH. The concentrations indicated throughout are for the final reaction mixture after mixing substrate and product. With the assistance of a metronome, time points could be accurately and reproducibly quenched at times as fast as 4 s.

For each substrate, we determined the concentration dependence by varying the concentration of AAG under single turnover conditions ([AAG] ≥ [DNA]). The rate constants for 4-8 independent reactions were averaged and the standard deviation was calculated in Excel. Plots of the observed rate constant as a function of the concentration of AAG were fit by a hyperbolic equation analogous to the Michaelis-Menten Equation (Eq. #3), in which the maximal observed rate constant (k_{max}) corresponds to the rate of the reaction at a saturating concentration of enzyme and the $K_{1/2}$ value indicates the concentration at which half of the substrate is bound. The ratio of $k_{max}/K_{1/2}$ is analogous to the steady state rate constant k_{cat}/K_M, which is commonly referred to as catalytic efficiency. Comparison of k_{cat}/K_M values for different substrates takes into account differences in both binding and catalysis.

$$k_{obs} = k_{max}[AAG]/(K_{1/2} + [AAG]) \quad (3)$$

Free Energy Calculations.

A linear free energy relationship correlating the catalytic efficiency of different inosine mismatches with the thermodynamic duplex stability demonstrates an inverse relationship between duplex stability and catalytic efficiency. The ΔΔG value for the relative catalytic efficiencies were calculated according to the equation, ΔΔG = -RT

$\ln(k_{cat}/K_M^{rel})$, in which k_{cat}/K_M^{rel} is the ratio of the k_{cat}/K_M value for the mismatch divided by that of the fastest substrate, the single nucleotide bulge (Table 3-3). The $\Delta\Delta G$ values for duplex stability were calculated from a comprehensive study of the nearest neighbor effects on duplex stability of inosine-containing DNA oligonucleotides, as described in Table 3-2.[4]

REFERENCES

1. Lindahl, T., *Instability and decay of the primary structure of DNA.* Nature, 1993. **362**(6422): p. 709-15.
2. Roberts, R.J. and X. Cheng, *Base flipping.* Annu Rev Biochem, 1998. **67**: p. 181-98.
3. Stivers, J.T. and Y.L. Jiang, *A mechanistic perspective on the chemistry of DNA repair glycosylases.* Chem Rev, 2003. **103**(7): p. 2729-59.
4. Lau, A.Y., et al., *Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG.* Proc Natl Acad Sci U S A, 2000. **97**(25): p. 13573-8.
5. Patel, D.J., et al., *Extra adenosine stacks into the self-complementary d(CGCAGAATTCGCG) duplex in solution.* Biochemistry, 1982. **21**(3): p. 445-51.
6. Joshua-Tor, L., et al., *The three-dimensional structure of a DNA duplex containing looped-out bases.* Nature, 1988. **334**(6177): p. 82-4.
7. Abner, C.W., et al., *Base excision and DNA binding activities of human alkyladenine DNA glycosylase are sensitive to the base paired with a lesion.* J Biol Chem, 2001. **276**(16): p. 13379-87.
8. Baldwin, M.R. and P.J. O'Brien, *Human AP endonuclease 1 stimulates multiple-turnover base excision by alkyladenine DNA glycosylase.* Biochemistry, 2009. **48**(25): p. 6022-33.
9. Watkins, N.E., Jr. and J. SantaLucia, Jr., *Nearest-neighbor thermodynamics of deoxyinosine pairs in DNA duplexes.* Nucleic Acids Res, 2005. **33**(19): p. 6258-67.
10. Krosky, D.J., F. Song, and J.T. Stivers, *The origins of high-affinity enzyme binding to an extrahelical DNA base.* Biochemistry, 2005. **44**(16): p. 5949-59.
11. Krosky, D.J., F.P. Schwarz, and J.T. Stivers, *Linear free energy correlations for enzymatic base flipping: how do damaged base pairs facilitate specific recognition?* Biochemistry, 2004. **43**(14): p. 4188-95.
12. Yasui, M., et al., *Miscoding properties of 2'-deoxyinosine, a nitric oxide-derived DNA Adduct, during translesion synthesis catalyzed by human DNA polymerases.* J Mol Biol, 2008. **377**(4): p. 1015-23.
13. Zhang, H., et al., *Steric and electrostatic effects at the C2 atom substituent influence replication and miscoding of the DNA deamination product deoxyxanthosine and analogs by DNA polymerases.* J Mol Biol, 2009. **392**(2): p. 251-69.
14. Hofseth, L.J., et al., *The adaptive imbalance in base excision-repair enzymes generates microsatellite instability in chronic inflammation.* J Clin Invest, 2003. **112**(12): p. 1887-94.

APPENDIX B

Additional data figures and accompanying discussion to support Chapter 3

SUPPORTING RESULTS AND DISCUSSION

Figure B-1 shows the sequence of the oligonucleotide substrates. A representative gel and accompanying time course are shown for the single turnover glycosylase assay (Figure B-2). Additional representative exponential fits to reactions with rate constants that varied by 4 orders of magnitude (Figure B-3). The concentration dependence of the single turnover reaction for mismatch and bulge substrates from Figure 3-2 in the text have been replotted with more appropriate scales (Figure B-4). The AAG concentration dependence for two single-stranded inosine-containing oligonucleotides is

```
I•T      5′fam-CGATAGCATCCTICCTTCTCTCCAT
            3′-GCTATCGTAGGATGGAAGAGAGGTA

I•C      5′fam-CGATAGCATCCTICCTTCTCTCCAT
            3′-GCTATCGTAGGATGGAAGAGAGGTA

I•A      5′fam-CGATAGCATCCTICCTTCTCTCCAT
            3′-GCTATCGTAGGATGGAAGAGAGGTA

I•G      5′fam-CGATAGCATCCTICCTTCTCTCCAT
            3′-GCTATCGTAGGATGGAAGAGAGGTA
                              I
I bulge  5′fam-CGATAGCATCCTCCTTCTCTCCAT
            3′-GCTATCGTAGGAGGAAGAGAGGTA

I (ss)   5′fam-CGATAGCATCCTICCTTCTCTCCAT

I (ssT)  5′fam-TTTTTTTTTTCTICCTTTTTTTTT
```

Figure B-1. Sequence of oligonucleotides that were used in this study. The inosine-containing strands had a 5'-fluorescein attached by a 6-aminohexyl linker (fam).

displayed in Figure B-5. The similar behavior for two different oligonucleotides, one of which is flanked by poly-thymidine stretches, suggests that the observed reaction took place on a single stranded oligonucleotide. Several previous studies have examined the glycosylase activity of AAG towards single strand and mismatched oligonucleotides substrates. These supporting references are provided and our results are discussed in the context of the other relevant studies.

		–TIC– –AXG–		
	X=C	X=A	X=T	X=G
ΔG TI/AX	-0.46 ± 0.06	0.09 ± 0.06	0.36 ± 0.06	0.76 ± 0.06
ΔG IC/XG	-1.07 ± 0.08	-1.33 ± 0.08	-0.54 ± 0.08	-0.74 ± 0.08
ΔG TIC/AXG	-1.53 ± 0.14	-1.24 ± 0.14	-0.18 ± 0.14	0.02 ± 0.14
ΔΔG	(0) ± 0.28	0.29 ± 0.28	1.35 ± 0.28	1.55 ± 0.28

Table B-1. Calculation of relative duplex stabilities for the inosine mismatches. Duplex stabilities were calculated using nearest neighbor rules from a systematic study of duplex stability for inosine mismatches in different sequence contexts [12]. All values are expressed in units of kcal/mol. The difference in free energy (ΔΔG is calculated by substracting the value of ΔG for the most stable pair (I•C) from the ΔG for the indicated pair and therefore ΔΔG for I•C is defined as zero.

Evaluation of data quality.

Representative time courses are shown for the single turnover glycosylase activity of AAG on the slowest substrates, single stranded DNA, and the fastest substrate, the single nucleotide bulge (Figure B-2). These plots demonstrate that fast and slow reactions all followed the expected single exponential. This indicates that AAG is stable under these assay conditions, and that fast time points could be reliably taken by hand.

Activity of AAG towards mismatches and a single nucleotide bulge.

The single turnover rate constants for the substrates mentioned in the main text and an additional single-stranded substrate are summarized in Table B-2. The best fit of the concentration dependence is reported and the error was estimated by fitting the concentration dependence of all of the independently determined rate constants (from

~50 individual single turnover reactions) with a hyperbolic equation (Eq. #2) using Kaleidagraph.

The full single turnover characterization of inosine-mismatches has not been previously reported for any sequence context, but several previous studies have tested whether the opposing base has an effect on the maximal single turnover rate constant. These are summarized in Table B-3. The maximal rate constants that we report with a fluorescence-based glycosylase assay are within 2-fold of the values reported previously using a ^{32}P-based assay [1].

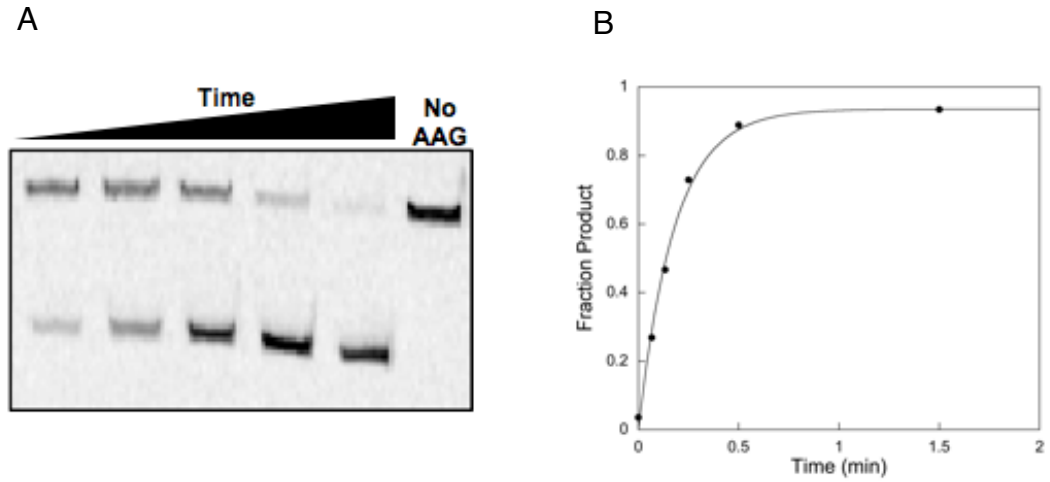

Figure B-2. Representative single-turnover glycosylase assay. (A) The fluorescence scan of a 20% polyacrylamide gel shows a reaction time course with 1μM AAG and 0.1μM I bulged substrate. The 25mer substrate is converted into a 12mer product after AAG-catalyzed N-glycosidic bond cleavage and alkaline cleavage of the abasic site. A control reaction in which AAG was omitted demonstrates that the inosine-containing substrate is stable to the glycosylase assay and provides the background signal expected at the initiation of the reaction. (B) The quantified data is fit to a single exponential curve (Eq. #2) and gives a k_{obs} value of 5.5 min^{-1} with an R^2 value of 0.99. Several independent experiments were performed for each concentration of AAG and the average and standard deviation are reported in Figure B-3.

This suggests that the change in reaction conditions and in labeling method does not have a substantial effect on AAG catalysis. The differences between I•T and I•C range from 2–5 fold for different sequence contexts and different reaction conditions. Other studies have used a fixed concentration of AAG and DNA to report initial rates for glycosylase activity towards inosine mismatches [2-4]. These studies found much larger effects, with the I•T mismatch being acted upon with 10 to 50-fold faster a rate than the I•C mismatch. The similar 24-fold greater catalytic efficiency towards I•T than I•C that we have measured for k_{cat}/K_M suggests that the previously published results also

compared k_{cat}/K_M. This indicates that AAG has a strong preference towards an I•T mismatch, since the competition between substrates for a given enzyme is determined by their relative k_{cat}/K_M values.

Opposing Base	$K_{1/2}$ (μM)	k_{max} (min^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)[b]	Relative k_{cat}/K_M[c]
None (bulge)	0.080 ± 0.008[d]	5.4 ± 0.1	68 ± 7	3.2
T	0.30 ± 0.02	6.3 ± 0.1	21 ± 1.4	(1)
G	0.41 ± 0.05	1.5 ± 0.1	3.7 ± 0.5	0.18
C	3.8 ± 0.6	3.3 ± 0.2	0.87 ± 0.15	0.041
A	2.3 ± 0.5	2.0 ± 0.1	0.87 ± 0.19	0.041
Single strand ss[e]	0.14 ± 0.06[d]	0.0027 ± 0.0002	0.019 ± 0.007	0.00092
Single strand ssT[e]	0.86 ± 0.09	0.0017 ± 0.0001	0.002 ± 0.0002	0.000095

Table B-2. Kinetic parameters for single turnover inosine DNA glycosylase activity of AAG[a]. A simplified form of this table is shown in the text (Table 1). [a]Data were collected at 23 ºC in 50 mM NaMES, pH 6.1, 1 mM EDTA, 1 mM DTT, 0.1 mg/mL BSA and with an ionic strength of 200 mM (adjusted with NaCl). [b]k_{cat}/K_M is calculated from the ratio of $k_{max}/K_{1/2}$, since this process monitors the same steps as k_{cat}/K_M in steady state kinetics. [c]The relative k_{cat}/K_M value was obtained by dividing the k_{cat}/K_M for a given substrate by the k_{cat}/K_M value for the I•T mismatch. [d]These values of $K_{1/2}$ are similar to the concentration of substrate (0.1 mM), so they should be considered upper limits to the true $K_{1/2}$ value. If the $K_{1/2}$ value is significantly lower than the k_{cat}/K_M would also be higher. [e]The two different single stranded sequences that were tested are shown in Figure S1. The errors for the k_{max} and $K_{1/2}$ values were estimated from curve fitting using nonlinear least squares regression to fit a hyperbolic concentration dependence (Equation S3) to all of the individual values of k_{obs} determined for a given substrate (Kaleidagraph), and this error was propagated to determine the error in k_{cat}/K_M.

Since we have analyzed the full concentration dependence, these data allows us to confirm that most of the difference in k_{cat}/K_M is attributed to adifference in binding affinity, with comparably small differences in reaction rate once the substrate is bound (Table B-2). The origin of this large difference in ground state binding is currently unclear. One possible explanation is that the ground state complex is stably flipped (K_{flip} >>1; as has recently been observed for the 1,N^6-ethenoadenosine lesion [5]. In this case, the more stable base pairs will require additional binding energy to reconfigure the substrate and flip out the inosine lesion. However, we cannot rule out the possibility that

the opposing base is directly contacted in an initial recognition complex that differs from the specific lesion recognition complex that has been observed crystallographically and that shows no specific contacts to the opposing base. Stopped-flow binding of 1,N^6-ethenoadenosine to AAG provides evidence that such an intermediate exists for this substrate [5]. Future studies looking directly at the rate and equilibrium constant for flipping of inosine from difference sequence contexts will be required to address this issue.

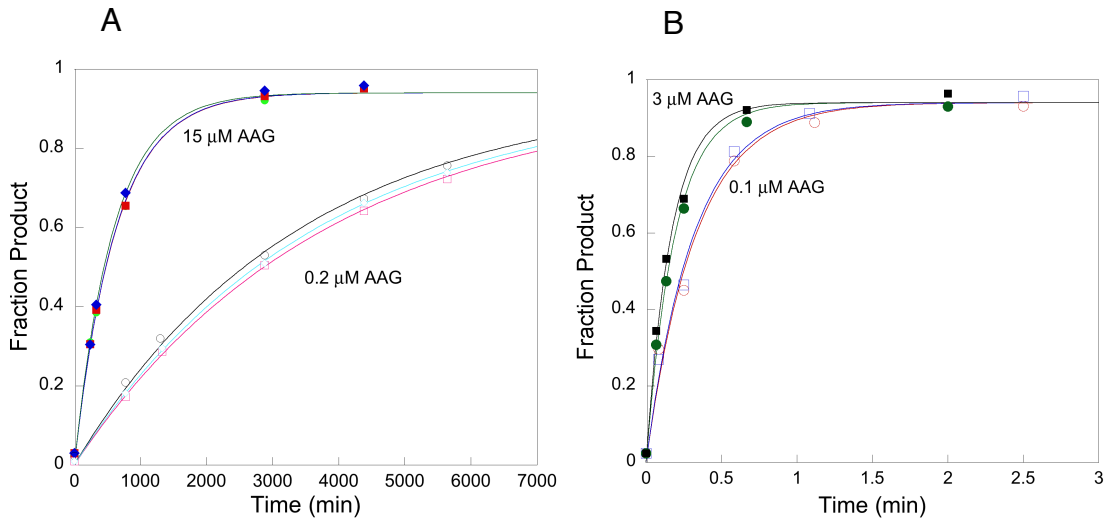

Figure B-3. Representative data for single turnover glycosylase activity with the slowest and fastest substrates. (A) For the thymidine-rich single-stranded substrate time points were taken over >3 days and the triplicate reactions are shown with saturating (15 mM) and sub-saturating (0.2 mM) AAG. (B) For the single nucleotide bulge times points were taken from 5-150 s. Duplicate reactions at saturating (3 mM) and subsaturating (0.1 mM) AAG are shown. For all substrates and all concentrations of AAG the reaction progress curves followed single exponentials and went to a similar endpoint of ≥95% cleaved.

A previous study concluded that an opposing base is required for AAG glycosylase activity, [6] because a lesion that was placed opposite a reduced abasic site (tetrahydrofuran) was not a substrate for AAG. However it is not possible to rule out the possibility that AAG might bind to the reduced abasic site directly since AAG bind with high affinity to this product mimic [6]. The finding that a single nucleotide bulge is an excellent substrate for AAG demonstrates that an opposing base is not required. Therefore, the decrease in glycosylase activity across from an abasic site is due either to

inhibition (binding directly to the abasic site) or to an alternative conformation that is less efficiently recognized by AAG.

Activity of AAG towards single-stranded inosine-containing DNA.

It has previously been reported that AAG has no detectable glycosylase towards inosine in single stranded DNA [7]. This is in contrast to oxanine and 1,N^6-ethenoA for which double stranded DNA is only modestly preferred by AAG [8]. A recent report

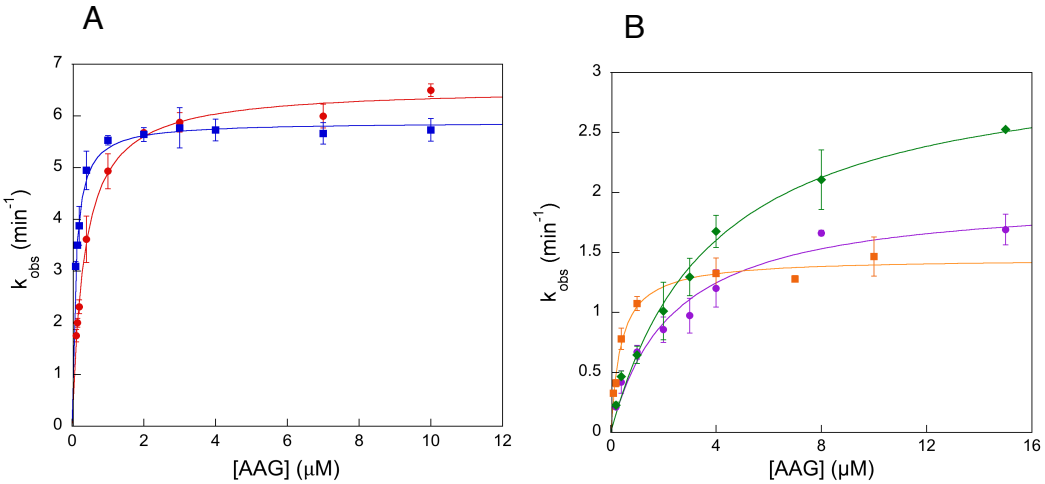

Figure B-4. Concentration dependence for the inosine DNA glycosylase activity of AAG replotted from Figure 3-2. The axes are adjusted to facilitate comparison of individual substrates. (A) The I·T mismatch (□) and the I bulge (□) show very similar single turnover kinetics. (B) AAG shows decreased glycosylase activity towards I·C (□), I·A (□), and I·G mismatches (□) with decreases in k_{max} and increases in $K_{1/2}$ (See Table S2 for the kinetic values and error estimates for the different substrates)

concluded that AAG also has robust activity towards inosine in a single strand, [9] but these reactions did not appear to go to completion and it was noted that the sequence of the oligonucleotide that was used was self-complementary in the region of the inosine lesion raising the possibility that the activity observed was not towards a single stranded lesion. We find that the single stranded oligonucleotide that we have characterized is an extremely poor substrate for AAG, with maximal single turnover rate constant that is 2000-fold slower than that of the I•T mismatch (Table B-2). Similarly, the catalytic efficiency is increased by 1000-fold when the single strand is annealed to its complement to form an I•T mismatch. Although the sequence that we initially used is not predicted to form stable secondary structure, we were concerned that a high concentration of AAG might stabilize an otherwise unstable structure that would allow faster rate of excision. In

this case, these large effects might overestimate the activity of AAG on single-stranded inosine. Therefore, we designed a sequence that was all pyrimidines except for the central inosine lesion, by replacing the flanking nucleotides with thymidines. The single turnover

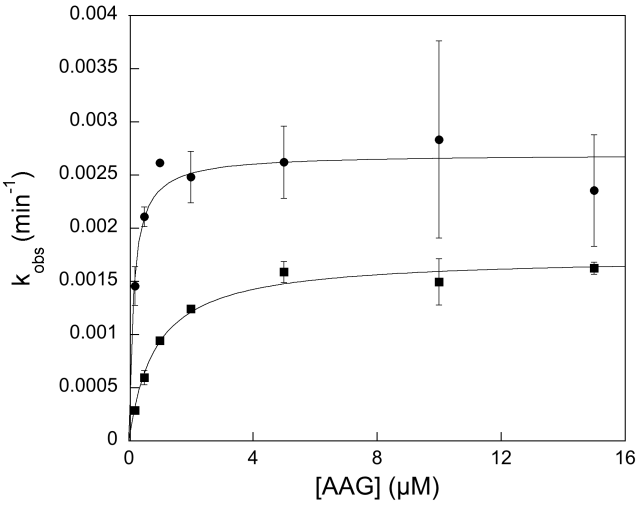

Figure B-5. Single turnover inosine DNA glycosylase activity of AAG for single-stranded substrates. The data for the heterogenous sequence (ss, ●) and the T-rich sequence (ssT, ■) are directly compared. See Figure 1 for the complete oligonucleotide sequences. Each point indicates the average of 3-6 independent determinations and the error bars indicate one standard deviation from the mean. In several cases the error bars are smaller than the symbols that were used. The single turnover rate constants are given in Table B-2.

excision data for this alternative sequence is summarized in Figure B-5 and Table B-2. The less diverse sequence bound 6-fold weaker and had a maximal rate that was ~1.5-fold slower than the heterogenous sequence. This corresponds to a 9-fold decrease in the catalytic efficiency for the thymidine-rich sequence (Table B-2). It is not clear what the origin of this sequence preference is, but it is not simply explained by a difference in secondary structure. If some duplex were formed, an increase in the maximal single turnover rate constant would be expected. These results indicate that inosine present on single-stranded DNA is a very poor substrate for AAG.

A linear free energy relationship for the inosine DNA glycosylase activity of AAG.

The barrier to nucleotide flipping can be evaluated by comparing the activity of the enzyme towards a lesion that is present in different base pairing contexts. The rate constant k_{cat}/K_M monitors the differences in energy between the ground state duplex in solution and the transition state bound to the enzyme with the lesion flipped out. Since

the bond being broken is identical in each case, the transition state is the same for each mismatch. Therefore, differences in reaction rate between different mismatches will be dominated by specific differences between the different base pairs. We have used the available data for the contributions of inosine mismatches to the equilibrium constant for duplex stability. Duplex stability is expected to be sensitive to how well a given mismatch is accommodated in the duplex and therefore serves as a useful surrogate for the base pairing stability of a given mismatch. The relative duplex stability for the sequence context that we used were calculated as described in the Materials and Methods and outlined in Table B-1. The relative k_{cat}/K_M values for AAG-catalyzed glycosylase activity were also converted into changes in free energy ($\Delta\Delta G$) and plotted in the text

		Normalized k_{max}			
Sequence	Reference	I•T	I•G	I•C	I•A
–TIC–	This study	(1)	0.24	0.52	0.32
–TIC–	[1]	(1)	0.14	0.40	0.18
–GIG–	[4]	(1)	0.31	0.31	0.11
–TIG–	[6]	(1)	ND	0.19	ND

Table B-3. Comparison of k_{max} values for the glycosylase activity of AAG towards inosine in different mismatches. The saturating single turnover rate constant, k_{max}, was normalized by dividing the value for a given mismatch by that of the most active mismatch context I•T. ND, not determined.

(Figure 3-3). Since only the natural inosine mismatches were used this is a very limited data set. It will be necessary to look at a larger set of natural and unnatural base pairs in order to obtain sufficient confidence in the exact slope. Nevertheless, the observed slope is sufficiently steep that we believe that the general trend is informative.

The best linear fit gives a slope of -0.96 (R^2=0.67). This strong negative correlation indicates that the differences in duplex stability for the different mismatches are fully experienced in going from that ground state to an extrahelical AAG-bound transition state. This supports the notion that AAG does not make specific contacts with the opposing base [10, 11]. However, this conclusion is tenuous because there is a large

deviation between the two least stable mismatches (I•T and I•G). If I•G deviates due to a negative effect, then the slope would be considerably steeper. If I•T deviates due to a positive effect, then the slope is less steep ($b = -0.6$; $R^2 = 0.97$). Since the biological function of AAG is to recognize and repair I•T mismatches, the natural context for deamination of an adenosine in DNA, it seems plausible that there may be factors beyond duplex destabilization that would lead to more efficient recognition by AAG. Previously it has been suggested that the unique wobble base pair geometry that is formed by an I•T pair could be responsible for the more efficient recognition by AAG [6].

REFERENCES

1. O'Brien, P.J. and T. Ellenberger, *Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase*. J Biol Chem, 2004. **279**(11): p. 9750-7.
2. Saparbaev, M., J.C. Mani, and J. Laval, *Interactions of the human, rat, Saccharomyces cerevisiae and Escherichia coli 3-methyladenine-DNA glycosylases with DNA containing dIMP residues*. Nucleic Acids Res, 2000. **28**(6): p. 1332-9.
3. Wyatt, M.D. and L.D. Samson, *Influence of DNA structure on hypoxanthine and 1,N(6)-ethenoadenine removal by murine 3-methyladenine DNA glycosylase*. Carcinogenesis, 2000. **21**(5): p. 901-8.
4. Asaeda, A., et al., *Substrate specificity of human methylpurine DNA N-glycosylase*. Biochemistry, 2000. **39**(8): p. 1959-65.
5. Wolfe, A.E. and P.J. O'Brien, *Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by human alkyladenine DNA glycosylase*. Biochemistry, 2009. **48**(48): p. 11357-69.
6. Abner, C.W., et al., *Base excision and DNA binding activities of human alkyladenine DNA glycosylase are sensitive to the base paired with a lesion*. J Biol Chem, 2001. **276**(16): p. 13379-87.
7. Saparbaev, M. and J. Laval, *Excision of hypoxanthine from DNA containing dIMP residues by the Escherichia coli, yeast, rat, and human alkylpurine DNA glycosylases*. Proc Natl Acad Sci U S A, 1994. **91**(13): p. 5873-7.
8. Hitchcock, T.M., et al., *Oxanine DNA glycosylase activity from Mammalian alkyladenine glycosylase*. J Biol Chem, 2004. **279**(37): p. 38177-83.
9. Lee, C.Y., et al., *Recognition and processing of a new repertoire of DNA substrates by human 3-methyladenine DNA glycosylase (AAG)*. Biochemistry, 2009. **48**(9): p. 1850-61.
10. Lau, A.Y., et al., *Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision*. Cell, 1998. **95**(2): p. 249-58.

11. Lau, A.Y., et al., *Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG*. Proc Natl Acad Sci U S A, 2000. **97**(25): p. 13573-8.
12. Watkins, N.E., Jr. and J. SantaLucia, Jr., *Nearest-neighbor thermodynamics of deoxyinosine pairs in DNA duplexes*. Nucleic Acids Res, 2005. **33**(19): p. 6258-67.

CHAPTER 4

HUMAN ENDONUCLEASE III (NTH1) INITIATES DELETION OF A BULGED, DAMAGED BASE

Small insertions and deletions (indels) are common mutations that have a high probability of causing a loss of gene function (1). Indels that change the coding region of a gene by integers other than 3 bp are known as frameshift mutations, because they disrupt the normal reading frame and alter the amino acid sequence. The most common indels involve loss or gain of a single nucleotide, but larger indels occur particularly in repetitive regions of the genome. Frameshift mutations predominantly arise from DNA replication errors in which the primer/template pair slips in a polymerase active site and escapes exonucleolytic proofreading to produce a bulged (unpaired) region (2). Several factors are known to increase the frequency of polymerase slippage, including damaged nucleotides, abasic sites, and DNA intercalators (3). Failure to repair a bulged intermediate prior to a subsequent round of DNA replication will produce a single nucleotide deletion or insertion. Most bulged intermediates are recognized and repaired by the mismatch DNA repair pathway (MMR) (4). Defects in the MMR pathway caused by mutations in MMR genes or changes in promoter methylation state that lead to decreased expression are commonly observed in hereditary colon cancer (5-7). Increased rates of insertions and deletions, particularly in repetitive regions of the genome, are considered to be a clinical indicator of defective MMR (8).

Intriguingly, increased rates of frameshift mutations have been reported in the absence of MMR dysfunction (9, 10). Upregulation of enzymes in the base excision DNA repair (BER) pathway has been shown to increase the frequency of frameshift mutations (11-13). The BER pathway is responsible for repairing the majority of base lesions that occur in the genome due to oxidative or alkylative damage (14, 15). A DNA repair glycosylase initiates the BER pathway by recognizing a damaged nucleotide and catalyzing the hydrolysis of the N-glycosidic bond to liberate the damaged nucleobase (16, 17). Human alkyladenine DNA glycosylase (AAG) has been reported to affect frameshift mutations via two distinct mechanisms. The first mechanism involves the high affinity binding of AAG to a bulged nucleotide which blocks access by the MMR pathway, resulting in increased levels of both insertions and deletions. (18). A second mechanism relies on the ability of AAG to catalyze the excision of a bulged lesion (19) and initiate the BER pathway to delete the bulged nucleotide. BER deletion of a bulged nucleotide leads to an immediate -1 frameshift, but prevents +1 frameshift events. This activity would produce a unique spectrum of frameshift mutations, resulting in a bias for -1 frameshift mutations (20). This mechanism may contribute to the known 3-6 fold bias for -1 frameshifts in the human genome (21-23).

In addition to AAG, two other human glycosylases have been shown to have the ability to excise a bulged lesion. The bifunctional glycosylase/lyase Nei-like protein 1 (Neil1) excises oxidized guanine lesions in a bulged confirmation (24) and uracil DNA glyosylase excises bulged uracil lesions (20). These glycosylases represent three of the four structural families of human glycosylases (25). This suggests that it is common for glycosylases to recognize bulged lesions, and that for many forms of DNA damage there

could be a competition between the MMR pathway that detects helix discontinuities and glycosylases that detect distinct modified bases. In the current work we examined a member of the largest superfamily of human DNA glycosylases, the helix-hairpin-helix superfamily (26). We investigated the ability of the human homolog of endonuclease III (Nth1) to excise a bulged lesion and initiate production of single nucleotide deletion by BER.

Nth1 is a bifunctional glycosylase/lyase that recognizes a wide range of oxidatively damaged pyrimidines, including 5,6-dihydrouracil (DHU) that is formed by oxidative damage of cytosine (27). There are crystal structures of prokaryotic homologs of Nth1 (28-30) and there have been numerous biochemical studies examining the activity of Nth1 toward damaged pyrimidines (27, 31, 32). Nth1 is not essential in mice and can be deleted, presumably because other DNA repair glycosylases serve as backup activities (33). Indeed, when both Nth1 and Neil1 are deleted in mice this results in a high incidence of tumors (34).

We compared the rates of Nth1-catalyzed excision of a DHU•G mismatch and single nucleotide bulge and found that Nth1 discriminates against a bulged lesion by a factor of ~20-fold. Nevertheless, Nth1 rapidly excises a DHU lesion from a single nucleotide bulge to initiate deletion of the bulged nucleotide via BER. In human nuclear extracts the rate of deletion of a DHU bulge is comparable to the rate of repair of a DHU•G mismatch. There is now evidence that glycosylases of every known structural family have robust catalytic activity toward bulged, damaged bases, suggesting that it is difficult for base flipping enzymes to rigorously exclude bulged lesions. As bulged lesions can occur during replication of damaged templates or as a result of spontaneous

damaged of an unpaired nucleotide, there is considerable potential for the BER pathway to compete with the MMR pathway for recognition of nascent indels.

MATERIALS AND METHODS

Recombinant Protein

Full-length APE1 (35) and DNA polymerase β were expressed in *Escherichia coli* and purified as previously described (36). Purification of the catalytic domain of DNA ligase I (Δ232), lacking the first 232 amino acids, has been previously described (37). The concentration of APE1, polymerase β, and DNA ligase I were estimated from their UV absorbance using the calculated extinction coefficients at 280 nm. Full-length Nth1 was cloned as an N-terminal 6XHis-sumo fusion and the sequence verified by sequencing. Recombinant protein was produced in Rosetta 2 *E. coli* using autoinduction and cells were lysed in 25 mM potassium phosphate pH 7.4, 10% glycerol, 300 mM NaCl using a cocktail of protease inhibitors (Roche). After purification using NTA-Ni^{2+} agarose and elution with an imidazole gradient, the sumo tag was removed with ULP1, and Nth1 was further purified with cation exchange chromatography using SP sepharose. Peak fractions were pooled and exchanged into storage buffer (50 mM NaHEPES, pH 7.5, 100 mM NaCl, 10% glycerol, 5 mM b-mercaptoethanol) and frozen at -80 ºC. The concentration of Nth1 was determined by active site titration (Figure C-2). Stability of Nth1 at 4ºC was limited, with approximately 8% loss of activity per day (Figure C-3). Therefore, fresh aliquots of Nth1 were thawed from frozen stocks (-80 ºC) immediately prior to kinetic experiments.

Oligonucleotide substrates

DNA oligonucleotides were synthesized by commercial sources and purified by denaturing polyacrylamide gel electrophoresis (PAGE; 6.6M urea and 1x TBE) as described (19). Phosphorothioate containing oligonucleotides were purified in the presence of 0.1mM dithiothreitol (DTT). The sequences of the oligonucleotides are illustrated in Figure C-1. 5'-fluorescein (6-fam) labeled oligonucleotides harbor a centrally located 5,6-dihydrouracil lesion or cytosine nucleotide (Figure C-1). Complementary strands are unlabeled. The labeled oligonucleotide was annealed with a 24-mer complement to create a single nucleotide bulge, or with a 25-mer complement to create a central base pair (X·Y). The 25-mer sequence was extended to a 35-mer to improve stability of nicked BER intermediates at 37°C. Additionally, the ends are decorated with C3-spacer aliphatic chains, bridged by a phosphorothioate linkage to prevent non-specific exonuclease degradation in cell extract. Oligonucleotides were annealed by heating to 37 °C for 5 min and then transferring to an ice bath.

General glycosylase assay

(6-fam)-labeled oligonucleotides (25mer) were incubated with Nth1 at 23°C in glycosylase buffer containing 50mM NaHEPES (pH 7.5), 1mM DTT, 1mM EDTA, 0.1 mg/ml BSA, 10% glycerol and additional NaCl to reach desired ionic strength. Samples were quenched with two volumes of 0.3M NaOH to reach a final concentration of 0.2M. Abasic sites were quantitatively converted to DNA breaks by heating at 70°C for 10 min, followed by the addition of 3.3 volumes formamide/EDTA containing 0.05% w/v of both bromophenol blue and xylene cylanol as tracking dyes. DNA breaks were separated by denaturing PAGE and imaged with a Typhoon trio (GE Healthcare) using 488 nm

excitation and a 520BP40 emission filter. Fluorescence intensity of individual bands was determined with Image-Quant TL (GE Healthcare). The fraction of glycosylase product was determined by dividing the fluorescence intensity of the 12-mer product band by the sum of the 12-mer product and 25-mer substrate bands. The fraction was converted into concentration by multiplying by the total amount of DNA substrate present as previously described (19).

Single turnover glycosylase assay

Single turnover glycosylase assays were performed with Nth1 in at least a 2-fold excess over the DNA substrate. In most cases, 100nM DNA substrate was present, but 20nM DNA was used when Nth1 concentration was 40nM and 80nM. Single turnover rates were measured at 23°C in the glycosylase buffer described above. Single turnover reactions proceed extremely fast, most with half-life of less than 20 s. This required a technique in which equal volumes (3µl) of enzyme and substrate were mixed by hand, with the reaction occurring in the pipette tip. 3µl timepoints were quenched into a tube containing 6µl 0.3mM NaOH as described (19). Timepoints could be taken accurately and reproducibly as fast as 3 s with the aid of a metronome. In all cases, reaction progress curves were fit by a single exponential (eq 1), where A is the fraction of substrate converted to product at completion (A>0.90). k_{obs} is the observed single turnover rate constant and t is the reaction time. $K_{1/2}$ was too low to measure for all substrates expect the DHU-bulge (Figure 4-1). For the DHU-bulge the individual values of $K_{1/2}$ and k_{max} were determined by varying enzyme concentration under single turnover conditions and fitting with a hyperbolic dependence (eq 2). The saturating single turnover rate constant

(k_{max}) was determined at 3μM and 6μM Nth1 concentrations using 0.1μM DHU substrate. Reaction rates were independent of Nth1 concentration in all cases.

$$F = A[1-\exp(-k_{obs}t)] \quad (1)$$

$$k_{obs} = k_{max}[Nth1]/(K_{1/2}+[AAG]) \quad (2)$$

Steady state glycosylase assay

Multiple turnover of Nth1 was measured under glycosylase conditions described above. A linear fit of the initial rate was determined from timepoints less than 20% product formation. Relative specificity for DHU substrates was determined by directly competing two oligonucleotides. Initial velocities of each substrate are proportional to the relative k_{cat}/K_m values as described by eq 3 (35). Typically, a 20μl reaction contained 100nM of each (6-fam)-labeled substrate was incubated with 2nM Nth1. 3μl timepoints were quenched in 6μl 0.3M NaOH and analyzed as above.

$$V_A/V_B = (k_{cat}/K_m)_A[A]/(k_{cat}/K_m)_B[B] \quad (3)$$

In vitro BER reconstitution

Recombinant human proteins were used to reconstitute BER activity on a DHU-bulge. Reactions were incubated at 23°C under BER conditions of 50 mM NaHEPES (pH 7.5), 100 mM NaCl, 5 mM $MgCl_2$, 4 mM ATP, and 2.5 μM of each dNTP. Reactions contained 200 nM DHU-bulge (25mer) substrate with 50 nM Nth1, 120nM APE1, 5nM polymerase β, and 20nM ligase I, in a total volume of 20 μL. 3 μL timepoints were quenched directly into an equal volume of formamide/EDTA loading buffer. Samples were heated for 5 min at 70°C and anaylyzed by denaturing PAGE as above. Abasic sites remained intact during this procedure. Using the same protocol, intermediates in the BER processing of a DHU-bulge are observable by omitting each enzyme in succession.

Reactions were quenched after 60 min in an equal volume of formamide/EDTA loading buffer and analyzed as above.

Reaction of a uracil containing U-bulge (25mer) with *E. Coli* UDG (New England Biolabs) produces a bulged abasic site. Addition of Nth1 with UDG processed the U-bulge to a nicked intermediate, 3'-dR. BER reconstitution reactions with 3'-dR substituted 1mM TCEP in the place of DTT to prevent modification of the 3'-unsaturated sugar (Figure C-9).

Glycosylase assay in HeLa nuclear extract

Glycosylase reactions in HeLa NE were monitored at 37°C in a buffer containing 50 mM NaHEPES (pH 7.5), 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1 mg/ml of bovine serum albumin, and 10% (v/v) glycerol. Typically, 20µl reaction volume contained 10nM 5'-(6-fam)-labeled substrate and 0.4 mg/ml HeLa NE (Santa Cruz Biotechnology). 3µl aliquots were removed at the desired timepoints and quenched with 2 volumes of 0.3M NaOH and analyzed as above.

BER assays in HeLa nuclear extract

Assays for the processing of labeled oligonucleotides in cellular extract were performed at 37°C in 50 mM NaHEPES (pH 7.5), 100mM NaCl, 5mM $MgCl_2$, 4mM ATP, and 25µM of each dNTP as described previously (20). Typically, 10nM (6-fam)-labeled substrate and 100nM unlabeled DNA were incubated with 0.4 mg/ml of HeLa NE in a reaction volume of 20µl. Samples were taken by quenching 3µl aliquots into an equal volume of formamide as above. The amount of repair that had occurred was determined by quenching reactions at indicated time with EDTA (10 mM final). A vast excess of recombinant Nth1 (2µM) was added along with 10nM DHU·G (25mer) as an internal

standard for Nth1 glycosylase activity. After 30 min incubation at 37 °C, these reactions were quenched in 2 volumes of 0.3M NaOH and processed as described above to cleave abasic sites.

RESULTS

In order to characterize the kinetics for Nth1-catalyzed glycosylase activity it is important to know the concentration of active enzyme. As most glycosylases dissociate slowly from their products, it is common to observe burst kinetics in which the transient burst corresponds to the concentration of the active enzyme. Therefore we used burst kinetics to determine the active concentration of Nth1 and this is the concentration that is referred to in all subsequent experiments (Figure C-2).

In the current study we focused on 5,6-dihydrouracil (DHU) as a representative substrate of Nth1 (27). DHU is a ring-saturated pyrimidine lesion formed by a radical reduction and deamination of cytosine (38). We found that DHU is susceptible to non-enzymatic degradation under conditions normally used to anneal DNA and to analyze DNA glycosylase kinetics (see Appendix C; Figures C5-7). However, by annealing at neutral pH and lower temperature we were able to prevent these undesirable side reactions and obtain well-behaved enzymatic kinetics. It is a concern that other studies of DHU may have been affected by its intrinsic reactivity. The spontaneous ring-opening and breakdown of dihydrouridine ribonucleotide has been characterized and a similar mechanism is expected for DHU in DNA (39) (Figure C-4).

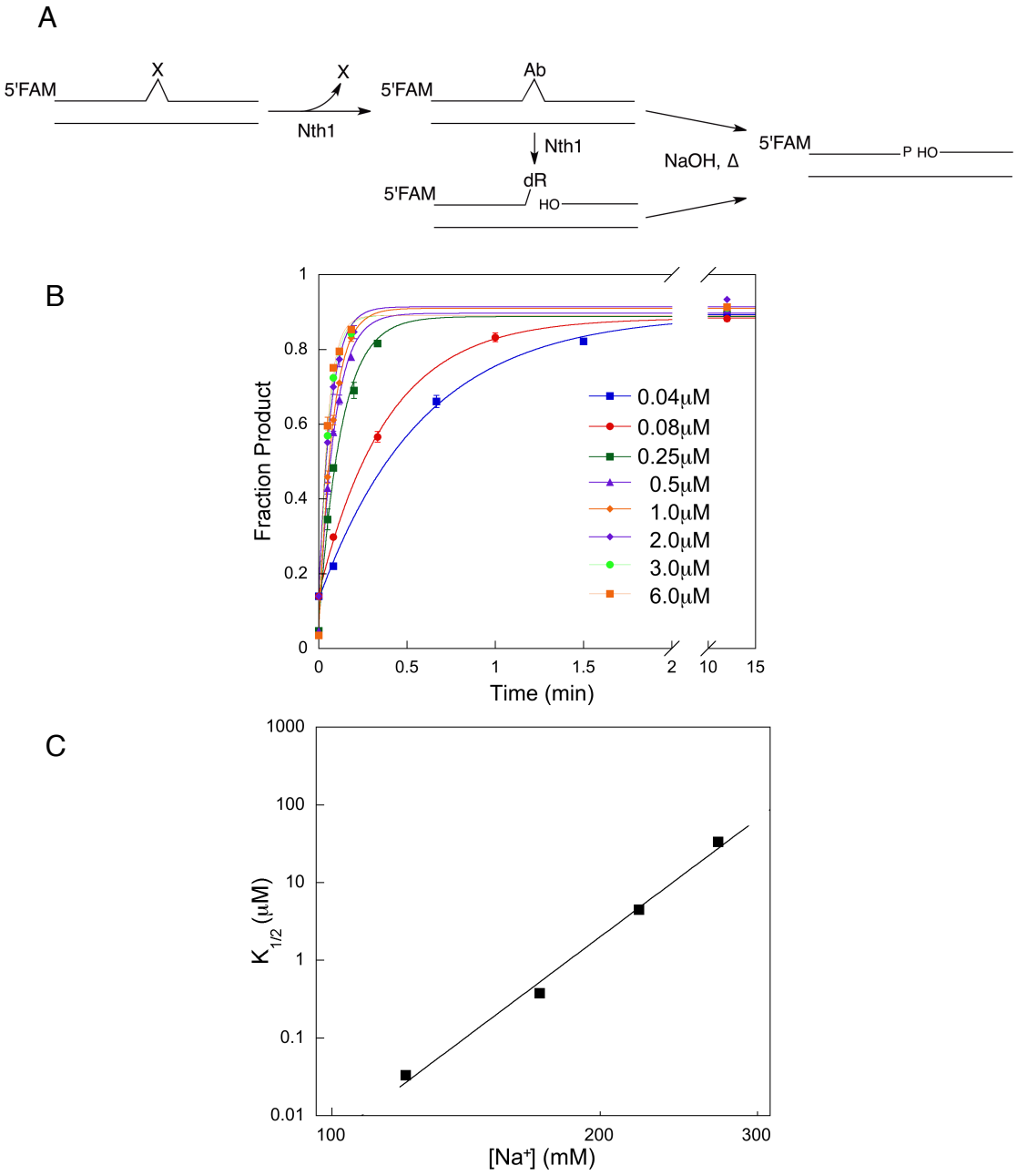

Figure 4-1. Nth1 glycosylase activity under single turnover conditions. (A) Nth1 glycosylase activity was measured using a fluorescently labeled oligonucleotide with a centrally located DHU-bulge. Products of Nth1 catalysis were cleaved under alkaline conditions and samples were analyzed by denaturing PAGE. (B) Representative data for single turnover glycosylase activity with a DHU-bulge at 23°C in 200mM [Na^+], pH 7.5. Each data point corresponds to the average and standard deviation of 2-4 individual reactions. (C) Concentration dependence for Nth1 catalysis of a DHU-bulge at 23°C under increasing [Na^+]. Lines represent the best fits to a hyperbolic concentration dependence: $k_{obs} = k_{max}[Nth1]/(K_{1/2} + [Nth1])$.

Single turnover excision of 5,6-dihydrouracil in different structural contexts

We used single turnover assays, with Nth1 in excess over DNA, to evaluate the ability of Nth1 to recognize and excise a bulged DHU lesion. The single turnover reaction reports on all of the steps up to and including the first irreversible step, and therefore is not affected by slow product dissociation. We found that NaOH (0.2 M final concentration) rapidly quenches Nth1 activity. Subsequent alkaline hydrolysis converts both abasic sites and β-elimination products into a 5'phosphate that provides a quantitative measure of Nth1 glycosylase activity (Figure 4-1A). We observed very rapid excision of DHU from a bulge, and the reactions are complete in under a minute at saturating concentration of Nth1, demonstrating robust activity toward a single nucleotide bulge (Figure 4-1B). The reaction progress curves follow a single exponential (Figure 4-1B), and the plot of the observed single turnover rate as a function of Nth1 follows a hyperbolic dependence on protein concentration (Figure 4-1C). The saturating single turnover rate constant (k_{max}) reaches a value of ~20 min^{-1}.

As most DNA binding enzymes are strongly dependent upon monovalent cation concentration, we repeated the single turnover excision experiments at a range of Na+ concentrations from 150-300 mM and the results are summarized in Figure 4-1C. As the concentration of Na+ is increased it takes significantly higher concentration of Nth1 to saturate.

To compare the catalytic activity of Nth1 on a bulge lesion to a canonical DHU·G mismatch, we also performed single turnover kinetic analysis of this mismatch. In contrast to the bulge, the mismatch is saturated at low nM concentration of Nth1 (Figure 4-2A). Under these conditions of 200 mM NaCl, the $K_{1/2}$ for the DHU•G mismatch is too

low to accurately measure and we can only place a limit of ≤20 nM. Nevertheless, the maximal single turnover rate (k_{max}) is identical for both the mismatch and bulge substrates. We tested the ability of Nth1 to excise DHU from mismatches with the other natural opposing bases and also evaluated to what extent Nth1 can excise DHU from single-stranded DNA. The k_{max} values were determined at saturating concentration of Nth1 and are summarized in Figure 4-2B and Table 4-1. Once saturated, Nth1 shows essentially identical single turnover glycosylase activity toward all duplex contexts that were tested. In contrast, the rate of DHU excision in single strand DNA is decreased by several orders of magnitude (Figure 4-2B). Inefficient excision of a single stranded lesion has been observed for several glycosylases, including AAG (19).

Multiple turnover kinetics to determine the catalytic efficiency of Nth1 toward DHU in different contexts

To better quantify the specificity of Nth1 we also performed multiple turnover kinetics. We measured the initial rates for excision of DHU from the DHU-bulge at a range of DNA concentrations and the hyperbolic substrate concentration dependence is shown in Figure 4-3A. This gives a k_{cat} value of 0.24 min^{-1}, two-orders of magnitude slower than the single turnover rate constant, and a K_M value of 82 nM that is somewhat smaller than the $K_{1/2}$ value of 520 nM. This shows that a step associated with product dissociation limits the multiple turnover excision of DHU from a bulge. The k_{cat}/K_M value is $6.8x10^4$ $M^{-1}s^{-1}$ and this reports on substrate binding and all of the steps up to and including the first irreversible step, presumably N-glycosidic bond cleavage. The tight binding of Nth1 to the DHU•G mismatch precluded measuring the k_{cat}/K_M value directly for this substrate.

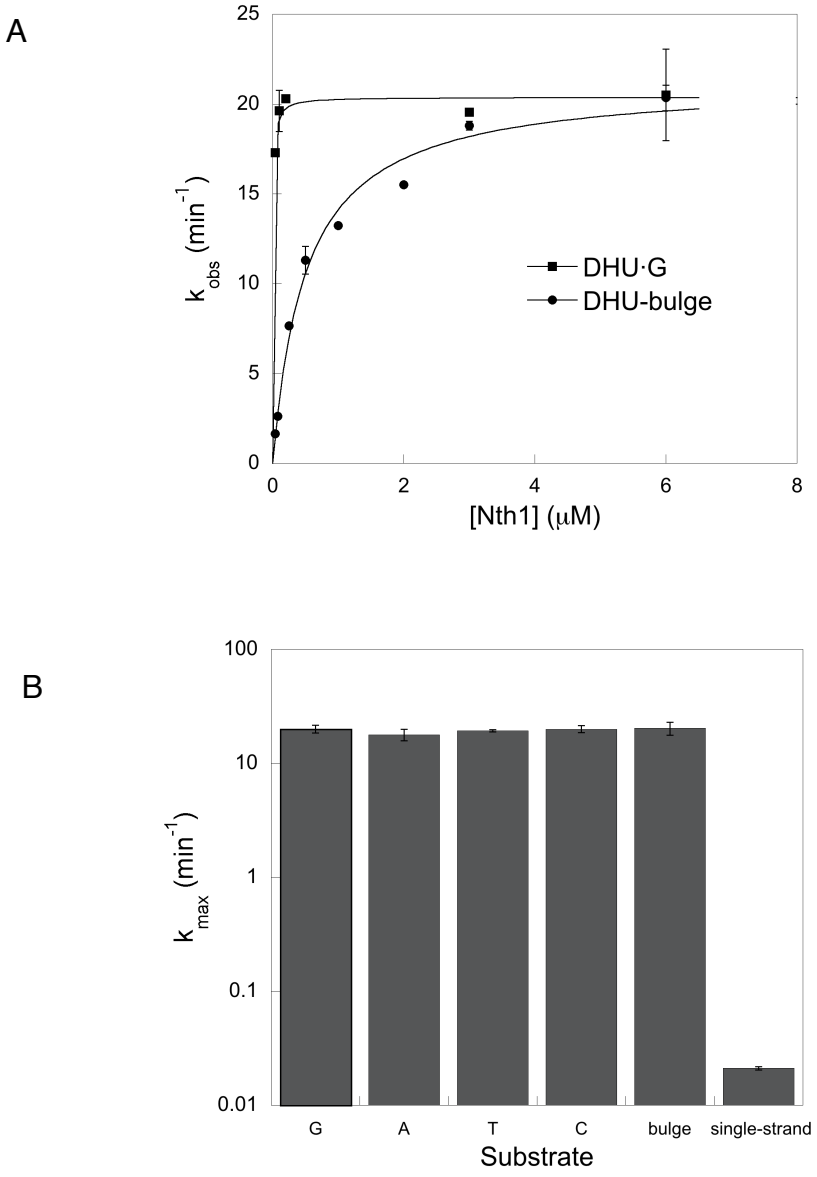

Figure 4-2. Single turnover excision of DHU by Nth1.
(A) Comparison of hyperbolic concentration dependence of DHU-bulge and DHU·G structural contexts at 25°C in 200mM [Na^+]. The $K_{1/2}$ value of DHU·G is at least 12-fold lower than DHU-bulge. (B) Values of k_{max} were measured at saturating Nth1 concentration. Values are average k_{obs} for two independent reactions at 3μM and 6μM [Nth1]. Error bars represent the standard deviation of k_{obs}. k_{max} value for the DHU-bulge substrate was taken from the more complete concentration dependence in part (A).

Therefore, we used direct competition experiments to measure the relative k_{cat}/K_M values for the different opposing base context.

Competition between two substrates binding to a single active site is given by the relative concentrations of the two substrates and their respective k_{cat}/K_M values (Figure 4-3B; Eq 3). For competition experiments, we extended the lesion-containing strand by 6 T at the 5'end so that the substrate and product would differ by 6 nt from our reference 25mer substrate and they could be readily resolved by denaturing PAGE (see Appendix C; Figures C-1 and C-8). Competition between the blunt end 25mer and the 5'-extended oligonucleotide, both with a central DHU•G mismatch, established that the 5' poly T extension does not affect Nth1 specificity (Figure C-8). As an example of this approach, equal concentrations of a DHU•G and DHU•A mismatch were combined and the initial rates of Nth1-catalyzed excision were measured (Figure 4-3C). These data demonstrate that Nth1 has almost a 10-fold preference for the DHU•G context over the DHU•A context under conditions of 200 mM NaCl. Analogous experiments were used to determine the relative k_{cat}/K_M values for all of the different contexts (Figure C-7) and the results are summarized in Table 4-1. It was predicted that DHU•G would be the preferred substrate of Nth1, because DHU is expected to arise from oxidation of C. Surprisingly, Nth1 shows similar efficiencies of excision of DHU opposite G, T, and C (Table 4-1). There is ~20-fold discrimination against DHU•A, ~40-fold discrimination against the DHU bulge, and ~800-fold discrimination against an unpaired DHU lesion.

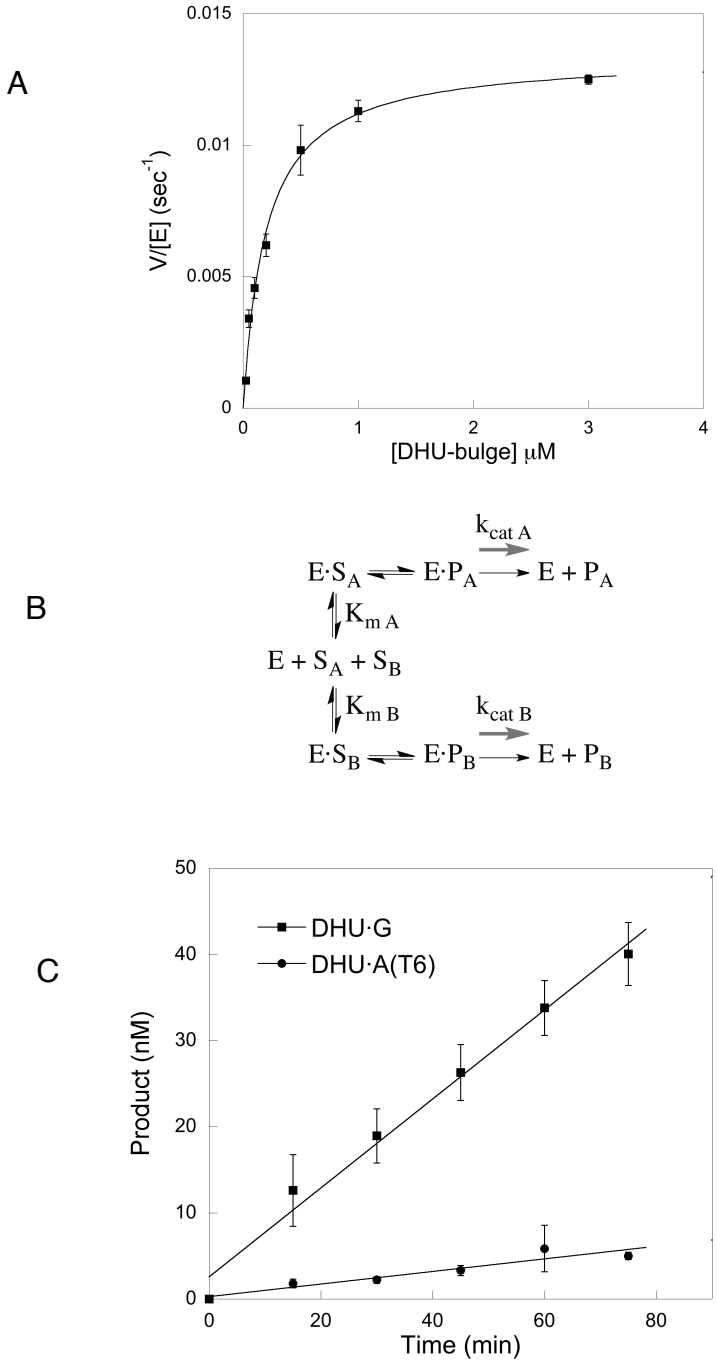

Figure 4-3. Steady-state excision of DHU by Nth1. (A) Steady-state excision of a DHU-bulge by 2nM Nth1 at 200mM [Na^+]. Initial rates were calculated from the first 20% of the reaction. Line indicates best fit to a hyperbolic concentration dependence. (B) Velocity for two substrates binding to a single active site is proportional to the k_{cat}/K_m for each substrate. The DNA substrates are abbreviated S and the product of Nth1 catalysis is labeled P. (C) Representative plot of the steady-state excision of DHU·G and DHU·A(T6). A mixture of 100nM of each DNA substrate was incubated with 2nM Nth1. (D) Ratio of the velocity between DHU·G and DHU·A(T6) ($V_{int\ (G)}/V_{int\ (A)}$) is dependent on [$Na^+$] concentration.

The salt dependence of the single turnover reaction suggested that Nth1 affinity for DNA is strongly sensitive to the salt concentration. Under conditions of very tight DNA binding the competition experiment does not accurately report on the specificity of the enzyme. This is because k_{cat}/K_M can become equal to the rate constant for substrate binding if every binding event produces product. To evaluate whether 200 mM NaCl captures the maximal specificity of Nth1, the DHU•G and DHU•A mismatches were competed at sodium concentrations between 100 and 300 mM (Figure 4-3D). These results suggest that the binding of the DHU•G lesion may not be fully reversible at 200 mM. The catalytic specificity of Nth1 that are reported at 200 mM NaCl underestimates the maximum discrimination that Nth1 is capable of (Table 4-1).

opposing base	k_{max} (s^{-1})	relative k_{cat}/K_m [a]	k_{cat}/K_m ($M^{-1}\,s^{-1}$) [c]
bulge (none)	23.0	0.026	$6.8x10^4$
G	22.0	0.61	$7.7x10^6$
A	19.0	0.064	$1.7x10^5$
T	20.0	0.73	$1.9x10^6$
C	19.0	(1)	$2.6x10^6$
single strand	0.02	0.0012 [b]	$3.1x10^3$

Table 4-1. [a]Relative k_{cat}/K_m values were determined by competition at 200mM [Na^+]. [b]Limit for value as measurement did not eliminate any minor contribution of duplex formed by annealed to excess unlabeled complement oligonucleotide. [c]k_{cat}/K_m was calculated using the experimentally defined k_{cat}/K_m for the DHU-bulge of $6.8x10^4$ $M^{-1}s^{-1}$.

Deletion of a DHU-bulge by BER

We next reconstituted the BER pathway for the complete processing of a DHU bulge. The β−elimination reaction that Nth1 catalyzes is distinct from the monofunctional glycosylase reaction catalyzed by AAG and UDG (19) and from the β,δ-elimination reaction catalyzed by Neil1 (24). The pathways for short-patch BER initiated by Nth1 is illustrated in Figure 4-4A. By analogy, two possible pathways can be drawn for deletion of a single nucleotide bulge (Figure 4-4B). The main difference is that there is no requirement for new DNA synthesis, and therefore it is possible to complete the pathway without a requirement for DNA polymerase β (Figure 4-4B; right branch).

As predicted, the incubation of DHU bulge with BER enzymes Nth1, APE1, polymerase β, DNA ligase I and the required substrates and cofactors resulted in robust single nucleotide deletion (Figure 4-5). Under these conditions the buildup and breakdown of nicked intermediate is observed. To evaluate the proposed pathway (Figure 4-4B), additional experiments were performed in which individual enzymes were omitted from the reaction mixture (Figure 5-6). No reaction occurs in the absence of Nth1 (lane 3). Omission of APE1 (lane 4) results in formation of the 3' unsaturated sugar, that we attribute to the product of the Nth1 lyase reaction. When polymerase b was omitted (lane 5), a single nucleotide deletion product (24mer) was observed, in addition to a nicked intermediate (12mer 3'-OH). As predicted by the model illustrated in Figure 4-4B, two pathways exist for processing of the bulged abasic site resulting from Nth1 glycosylase activity. One pathway for formation of a single nucleotide deletion occurs via

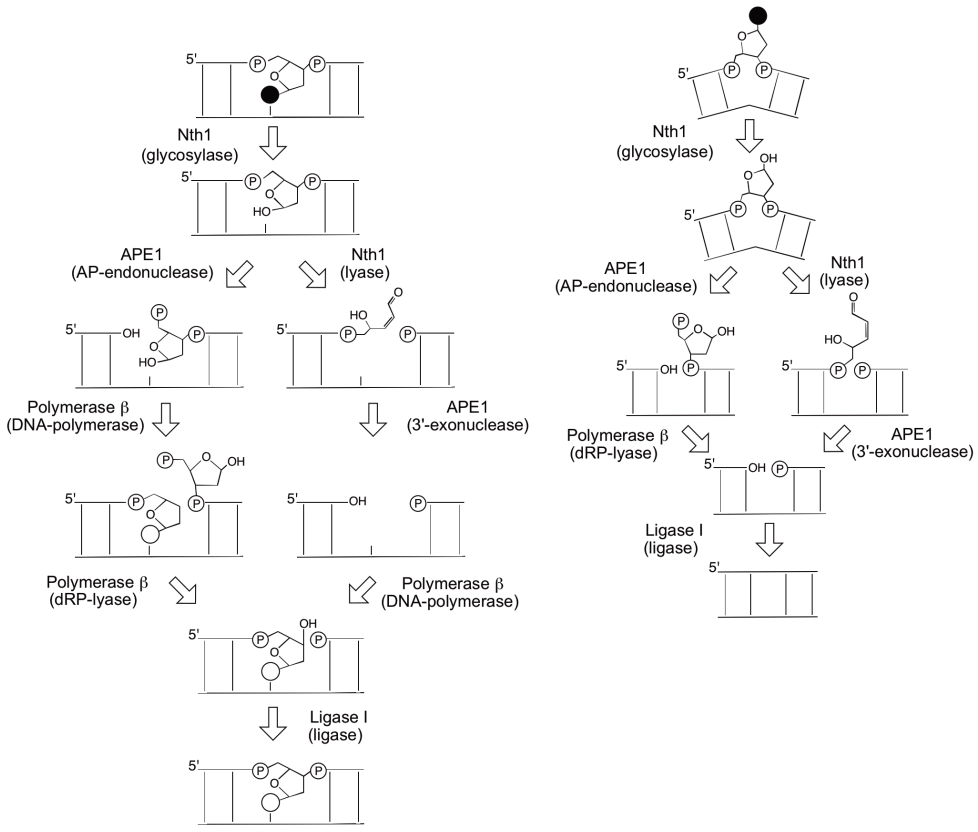

Figure 4-4. Short patch and single nucleotide deletion catalyzed by BER enzymes.
A pathway for short-patch BER is shown on the left and the proposed pathway for single nucleotide deletion is shown on the right. Filled circle represents a damaged nucleotide and an open circle, an undamaged nucleotide. Note that only the glycosylase and endonuclease steps differ for the two pathways. After APE1 cleavage of the 5'-dRP intermediate is chemically identical to the intermediate generated after nacent nucleotide incorporation in the short-patch BER pathway. Nth1 glycosylase activity produces a bulged abasic site. Endonucleolytic cleavage of the abasic site by APE1 results in a 5'-dRP nick with a 3'-OH. Alternatively, Nth1 lyase activity produces a 3'-unsaturated sugar (3'-dR). Bifunctional activity by Nth1 is non-obligatory, but both pathways of abasic site processing support single nucleotide deletion.

bifunctional Nth1-catalyzed β-elimination at the abasic site, followed by APE1 phosphodiesterase activity, and then DNA ligation. If APE1 acts as an AP endonuclease, outcompeting Nth1 at the abasic site, the 5' termini will be blocked by the dRP group in the absence of polymerase b and prevent ligation. Under the conditions tested here, APE1 outcompetes Nth1 for processing of the abasic site, producing and explains the persistent 12mer labeled product. Finally, the omission of DNA ligase (lane 6) resulted in formation of the 3'hydroxyl product of APE1.

We confirmed that APE1 is responsible for the 3' end processing by preparing a nick with a 3'unsaturated sugar and incubating it with different combinations of recombinant proteins (Figure 4-7A). The 3' phosphodiesterase activity of APE1 can be observed in lane 3 by the formation of the 3' hydroxyl, the identity of which is confirmed by its ability to be ligated by DNA ligase I (lane 6). DNA polymerase β has no detectable AP-lyase activity on this substrate (lane 4). Similar experiments were performed with a prepared abasic bulge (Figure 4-7B). Addition of APE1 produces a nicked intermediate (12mer 3'-OH; lane 4) that is distinct from the Nth1 product of abasic site processing (12mer 3'-dR; lane 6). Polymerase β is necessary for production of the single nucleotide deletion product (lane 3), because only the nicked DNA is observed in the absence of polymerase β (lane 5).

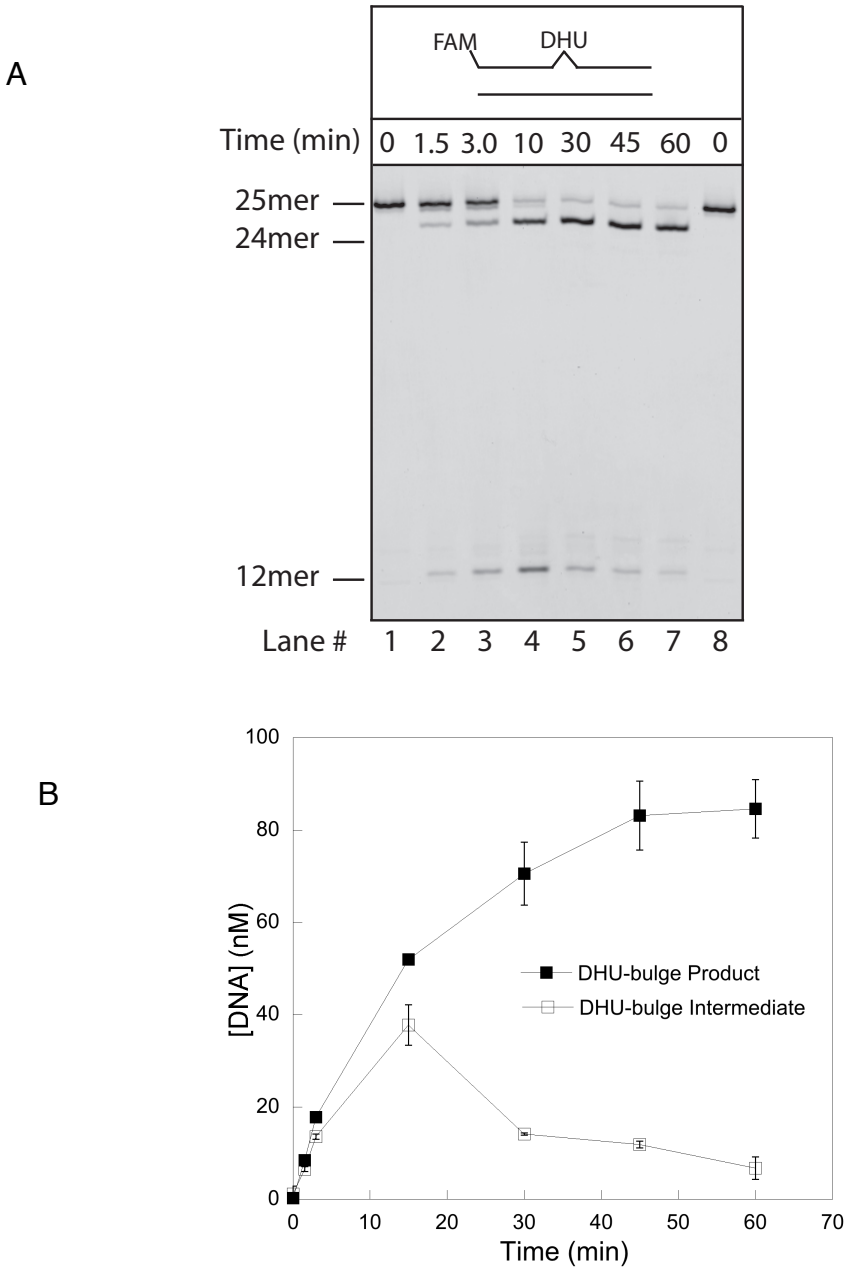

Figure 4-5. *In vitro* reconstitution of the single nucleotide deletion pathway with recombinant BER proteins. (A) Multiple turnover reconstituted BER reactions at 25°C contained 200nM DHU-bulge DNA, 50nM Nth1, 120nM APE1, 5nM Polymerase β, and 20nM DNA ligase I. Mg^{2+} and nucleotide cofactors necessary for BER activity were included (see Methods). Timepoints were quenched at the indicated time with formamide/EDTA loading buffer and analyzed by denaturing PAGE. The transient appearance of nicked intermediate (12mer) precedes the formation of single nucleotide deletion (24mer). (B) Quantification of the formation of the 24-mer deletion product and the formation and breakdown of the 12-mer nicked DNA intermediate. Data points represent the average and standard deviation of three independent experiments.

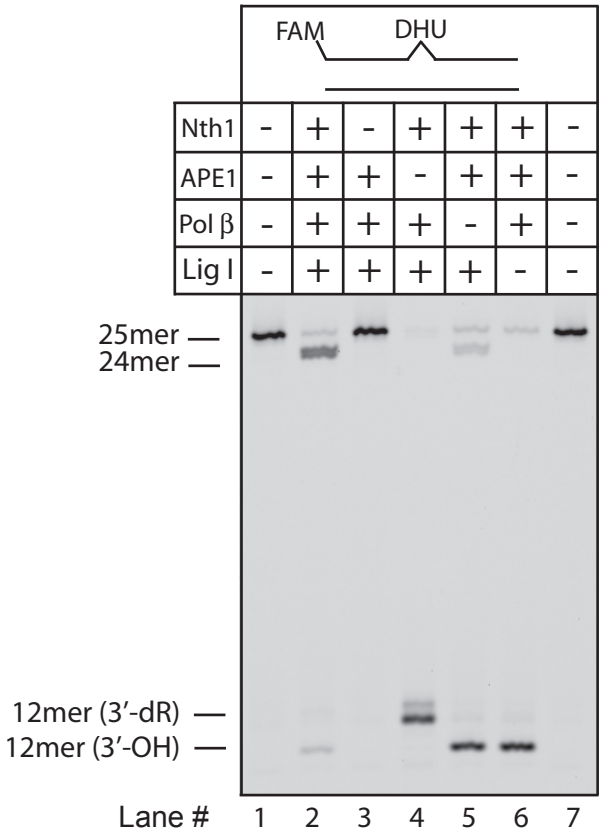

Figure 4-6. *In vitro* reconstitution of the single nucleotide deletion pathway with recombinant BER proteins. Recombinant Nth1, APE1, polymerase β, and DNA ligase I were sufficient for the single nucleotide deletion *in vitro* (lane 2). Omission of any single protein gave the expected product (lanes 3-6). Absence of polymerase β (lane 6) shows that APE1 is responsible for the majority of abasic site processing. The nicked 12mer (3'-OH) is resistant to ligation to due a 5'-dRP. Nth1 AP-lyase activity results in a polymerase β independent formation of the 24mer deletion product.

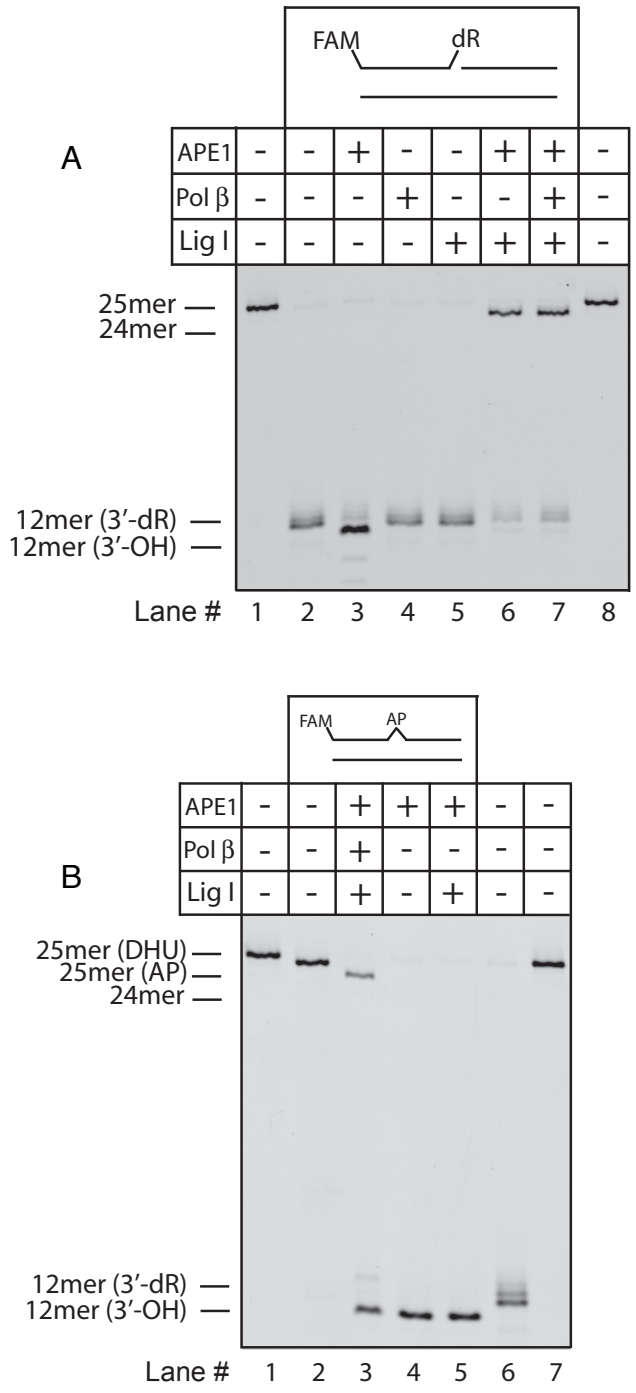

Figure 4-7. Processing of Nth1 products by BER (A) 120nM APE1 was incubated with 200nM nicked DNA at 23°C. APE1 possesses phosphodiesterase activity and catalyzes the removal of the 3'-unsaturated sugar (lane 3). APE1 activity allows formation of deletion product (24mer) independently of polymerase β (lane 6 and 7). (B) AP-bulge intermediate is processed to single nucleotide deletion (lane 3), but requires polymerase β (lane 5).

Glycosylase excision of a DHU lesion in HeLa nuclear extract

The reconstitution reaction with recombinant proteins establishes that the proposed pathways for deletion of bulged, damaged nucleotide can occur. However, it is possible that other proteins might compete with Nth1 for a DHU bulge or some downstream intermediate and result in an outcome other than single nucleotide deletion. Therefore we examined the reactivity of a DHU bulge in human cell extracts. Although we were previously able to observe single nucleotide events with bulged, deaminated lesions in human cell extracts (19), the nonspecific nuclease activities that are present limits the incubation time. Therefore, we sought to prevent degradation by modifying the ends of the oligonucleotide substrates. Similar strategies have been successful in preventing nucleolytic degradation (40). We used phosphorothioate linkages to link propyl alcohol to the ends (Figure C-1). Control experiments demonstrated that the degradation of the protected oligonucleotide duplex was undetectable under our reaction conditions (Figure C-10).

We first examined the ability of HeLa nuclear extract to catalyze excision of DHU in a G mismatch and single nucleotide bulge context. As DNA glycosylases do not require Mg^{2+} for activity, and it is an essential cofactor for APE1, DNA polymerases, and ligases, we omitted Mg^{2+} and added 1 mM EDTA. Both the DHU•G and DHU-bulge were excised on the minute to hour time scale (Figure 4-8). The rate of excision of the DHU-bulge is 2-fold faster than that of the DHU•G mismatch (Figure 4-9B).

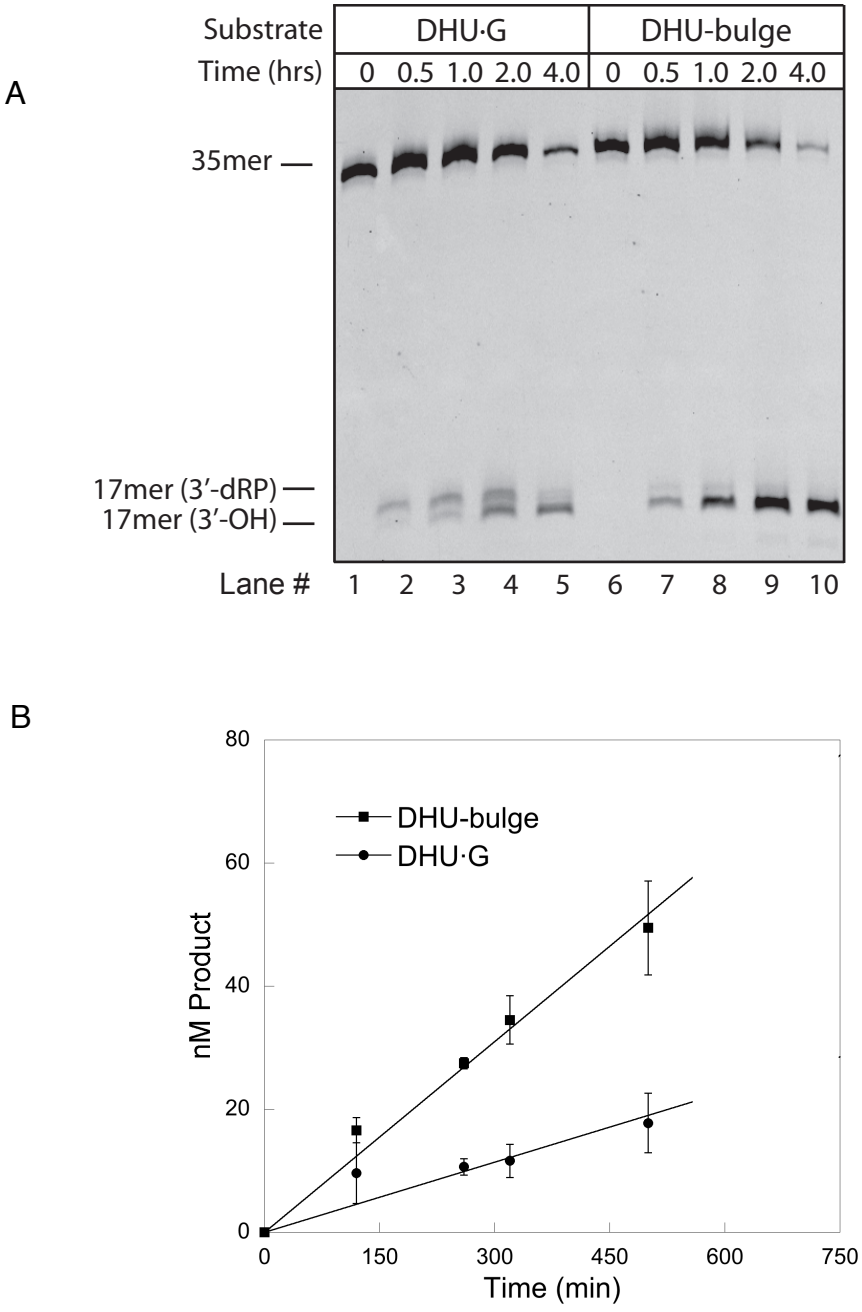

Figure 4-8. Glycosylase activity in HeLa NE. (A) Bifunctional glycosylase activity was observed in HeLa NE as formation of nicked DNA intermediates. A 35mer oligonucleotide (10nM) containing a DHU-bulge was incubated in 0.4 mg/mL HeLa NE at 37°C in absence of Mg^{2+}. Formation of a nicked intermediate requires elimination of the resultant abasic site (17mer). Timepoints were quenched with formamide/EDTA, which preserves abasic sites as 35mer. Bifunctional glycosylase activity is expected to produce a 3'-dR nicked DNA. (B) Quantification of nicked product. For lanes 1-6, the two products (3'-dR and 3'-OH) were not differentiated. Data points represent the average and standard deviation of three independent experiments.

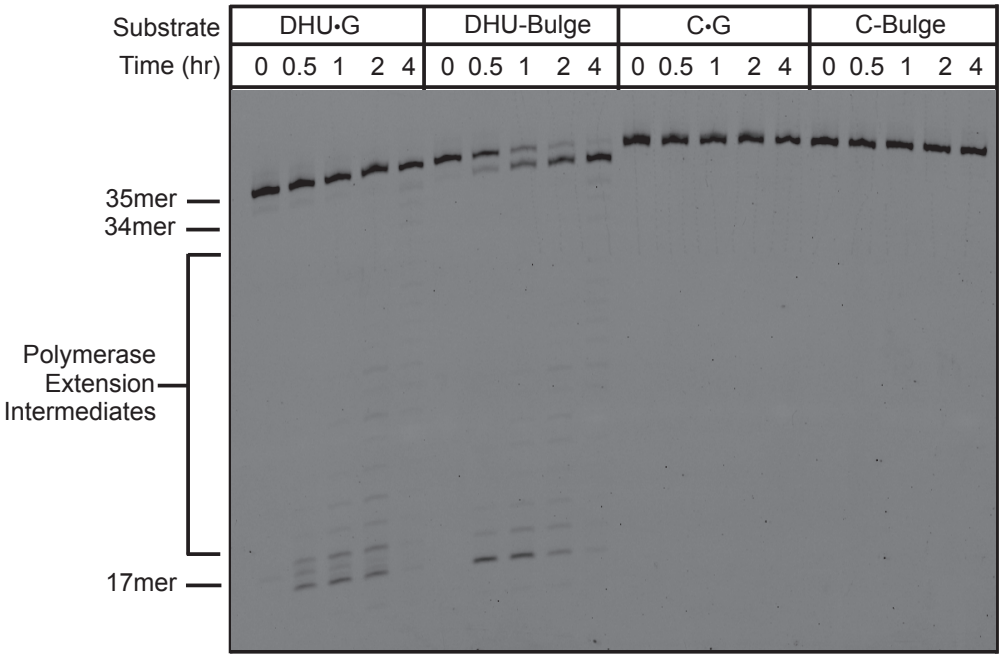

Figure 4-9. Base excision repair assays in HeLa NE demonstrates single nucleotide deletion. Nuclease resistant 35mer oligonucleotide was incubated in 0.4 mg/ml of HeLa NE at 37°C. 5mM $MgCl_2$, 2.4mM ATP, and 25μM of each dNTP were added to support BER activity. Complete processing of a DHU-bulge (35mer) to a deletion product (34mer) occurred within 4 hours at 37°C (lanes 6-10). The DHU·G mismatch was processed on the same timescale (lanes 1-5), as evidenced by the build-up of BER intermediates (17mer). No BER intermediates or products were observed when an undamaged cytosine was substituted in place of the DHU lesion, in either a basepaired or a bulge confirmation (lanes 16-20).

Deletion of a DHU-bulge in HeLa nuclear extract

Mg^{2+} and other cofactors were added to reactions in HeLa nuclear extract to allow for full BER activity. These conditions supported the production of a single nucleotide deletion (34mer) from a DHU-bulge substrate (Figure 4-9). No 24mer was detected with a DHU•G mismatch, because BER results in removal of the DHU, incorporation of a C, and ligation. No detectable reactions were observed on a Watson-Crick duplex (C•G) or on an undamaged C-bulge. In contrast, both the DHU•G substrate and the DHU•G bulge were processed to give a nicked DNA intermediate (17mer) and some aphidicolin-sensitive products that are likely to be the result of long-patch BER (19).

To confirm that the DHU lesion was removed during the processing reactions in the cell extracts, we performed coupled assays similar to those that we reported previously (19). The design of the experiment is illustrated in Figure 4-10A. After incubation from 0-4 hours in HeLa nuclear extract, reactions were quenched by the addition of EDTA to prevent ligation and polymerization, and recombinant Nth1 was added in large excess over the substrate DNA. As an internal standard, a 25mer DHU·G containing oligonucleotide was added at the same time. During this second 20 min incubation any remaining DHU lesion is excised by recombinant Nth1. Reactions were quenched in NaOH and all abasic sites and 3'dR groups were hydrolyzed. Complete processing of the control 25mer confirms that this protocol completely processes any DHU that was present (Figure 4-10B). For the DHU•G substrate, an increasing proportion of the 35mer becomes impervious to Nth1 excision and alkaline cleavage (17mer). This 35mer is the fraction of substrate that has been repaired by BER in the

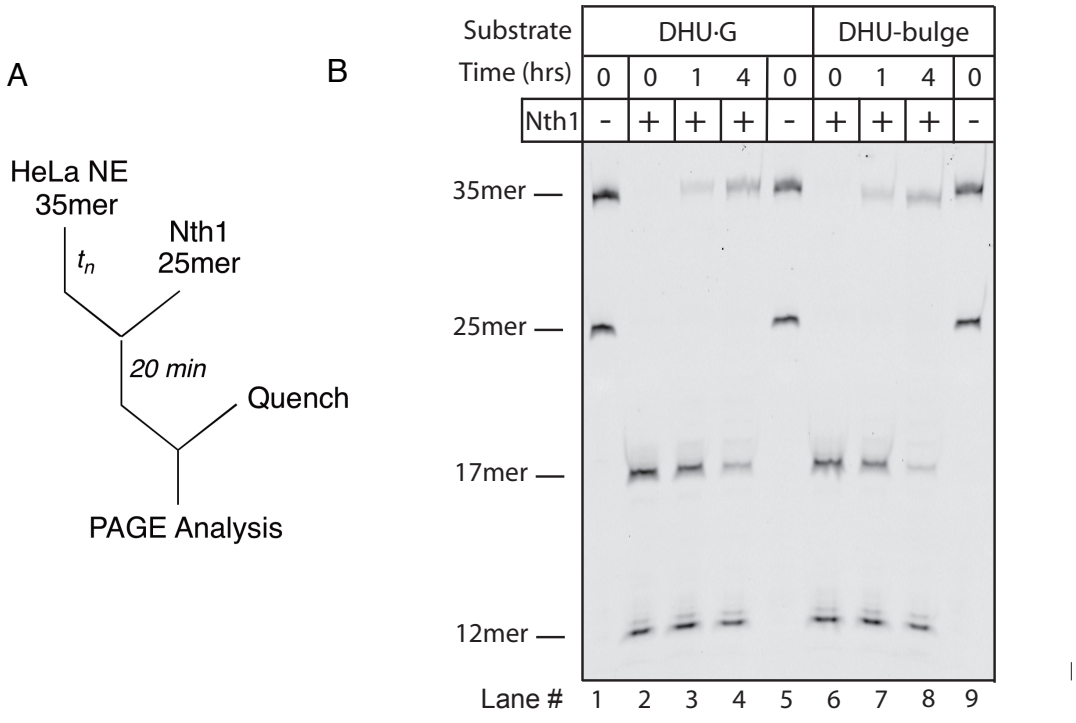

Figure 4-10. Removal of DHU lesion in HeLa NE. (A) 35mer DNA (10nM) was incubated in HeLa NE. At the timepoints indicated, BER processing was quenched with 10mM EDTA. Recombinant Nth1 (100nM) was added along with a DHU·G containing 25mer as an internal control for Nth1 competency. Reaction proceeded for 20min at 37°C before quenching with formamide/EDTA. (B) Removal of DHU lesion from 35mer DNA by BER activity in HeLa NE is apparent as resistance to recombinant Nth1 activity (*+Nth1)*. Increasing persistence of the 35mer over time is significant of DHU removal (lanes 2-4).

elapsed time. Similarly, for the DHU-bulge it is the 34mer deletion product that accumulates over time. The results of this coupled assay were quantified and summarized in Figure 4-10C. Under the BER-permissive conditions both the DHU•G mismatch and the DHU single nucleotide bulge are repaired in HeLa NE at similar rates.

DISCUSSION

Several DNA glycosylases have been found to recognize lesions present in a single nucleotide bulge and this activity can result in deletion of the damaged nucleotide via mechanisms that are closely analogous to BER pathways for the repair of damaged nucleotides present in mismatches (19, 24). It is suggested that these reactions will compete against other DNA repair pathways, including MMR and that the collective activities of glycosylases will generate a bias for deletions over insertions. To evaluate the scope of the competition that is possible, we have examined Nth1, a representative member of the largest superfamily of DNA glycosylases, the helix-hairpin-helix superfamily. In this study, we used kinetic experiments to determine the efficiency of Nth1 to excise a DHU lesion in several structural contexts, including a single nucleotide bulge. Our results show that Nth1 discriminates strongly against a single nucleotide bulge at the initial binding step, but lacks the ability to discriminate once it is bound. Apparently lesion recognition does not limit the overall reactions for processing of bulged or mismatched DHU lesions, because they are processed with similar rates in cell extracts and reconstituted reactions. This study extends the repertoire of enzymes (and types of damage) that can impact the frequency of frameshift mutagenesis.

Specificity of Nth1 for a DHU lesion in different structural contexts

We used direct competition experiments to measure the catalytic specificity of Nth1 toward DHU lesions present in different structural contexts. The specificity of Nth1 is strongly dependent on ionic strength. This ionic strength dependence can be easily understood as a manifestation of a slow rate of dissociation from DNA substrates when the concentration of cations is low. At the extreme, k_{cat}/K_M is limited by the rate of DNA

association because every binding encounter results in successful reaction. This is expected to be a common situation, because many DNA binding proteins use electrostatic interactions to strengthen DNA binding We report the specificity of Nth1 at 200 mM concentration of cations, as a reasonable approximation of the salt concentration found in the cell. Experiments at higher concentration of sodium suggest that this condition underestimates the true catalytic specificity that Nth1 is capable of. Nevertheless, these data provide new insight into the recognition of DNA damage by Nth1.

We found Nth1 has the highest catalytic efficiency for DHU•T, DHU•C, and DHU•G, with 10-fold reduced catalytic recognition of a DHU•A mismatch. Previous studies failed to detect a preference for DHU•G over DHU•A (27). However, these experiments were conducted at a low ionic strength concentration that would allow a significant contribution of non-specific DNA interactions to the overall analysis. Our measurements detected a mere 3-fold preference for DHU-G over DHU-A at an ionic strength of 100mM, whereas, a 100-fold preference was measured at 300mM ionic strength (Figure 4-3).

It is expected that DHU•G is the structural context that occurs most often *in vivo* due to oxidation and deamination of cytosine. In contrast, replication of a DHU lesion is expected to incorporate an opposing A (42). Excision of DHU from a DHU•A pair would lead to C to T transition mutation. Discrimination by Nth1 against A leaves open the possibility for post-replicative repair that could target the A. There is precedent for such a pathway in the repair of 8-oxoguanine (8-oxoG). 8oxoG is initially formed paired to cytosine, but replication of 8oxoG supports incorporation of either cytosine or adenine opposing the 8oxoG lesion (41). 8-oxoguanine DNA glycosylase I (Ogg1) is specific for

8oxoG-C, excising the 8oxoG lesion (42). MutY is specific for 8oxoG-A, excising the adenine and retaining the 8oxoG lesion in order to allow another replication event to incorporate 8oxoG and lead to restoration of the correct G-C basepair (43). It is not known to what extent MutY is able to recognize a DHU•A pair.

There are currently no structures available for human Nth1, but structural analysis of the *Bacillus stearothermophilus* homolog of Nth1 trapped with a reduced Schiff base illustrated that several contacts are established with a guanosine nucleotide that are absent with adenosine (30). There appear to be minor rearrangements in the protein backbone between the structures with a bound A and G that act to condense the protein structure around the guanine. However, it is not clear from these structures how pyrimidines could be accommodated by Nth1.

Nth1 discriminates against excision of a DHU-bulge

A primary objective of this study was to determine if Nth1 is able to excise a bulged lesion. Measured at 200mM ionic strength, Nth1 shows a 20-fold preference for a canonical DHU•G mismatch over a DHU bulge. The bulge is known to distort the DNA backbone. As Nth1 makes extensive contacts with the DNA backbone, is not surprising that Nth1 has a more difficult time engaging a bulge.

Importantly, the maximal single turnover rate for DHU excision was the same for all structural contexts. Single stranded DHU lesion was an exception, incurring substantial reduction in maximal turnover rate. The maximal single turnover rate is measured under saturating concentration of enzyme and reports on all catalytic steps after substrate binding through the first irreversible step. Therefore, we can conclude that the

diminished specificity for DHU-A and DHU-bulge contexts are due to a reduction in affinity.

In contrast, AAG showed a 3-fold higher affinity for a bulged lesion than a mismatched lesion. The catalytic efficiency for AAG excision of a mismatched lesion inversely correlated with the stability of duplex, specifically the interaction of the lesion with the corresponding opposing base (20). AAG makes no observable contacts with the opposing base, therefore, the absence of opposing base interactions is predicted to reduce the energetic barrier to base flipping. As both AAG and Nth1 make extensive contacts along the DNA backbone when engaging a lesion, backbone distortions caused by a bulged lesion may impair binding affinity. AAG may overcome the reduction in affinity by possessing an inherent flexibility in the DNA-protein interface, where Nth1 may lack the same extent of flexibility.

Human BER processes a DHU bulge to a single nucleotide deletion

Nth1 excision of a DHU-bulge initiates the BER pathway, ultimately leading the deletion of the bulged lesion. We reconstituted the short-patch BER pathway *in vitro* with recombinant human proteins to demonstrate that a DHU-bulge is efficiently deleted and each enzyme in the BER pathway functions as expected. We show that Nth1 can act as a monofunctional glycosylase, processing a DHU-bulge to a bulged abasic site. This activity utilizes the AP-endonuclease activity of APE1 and also requires the 5'-dRP lyase activity of polymerase β. Alternatively, long-patch BER could be realized by loading of a replicative polymerase at the nick site. Replicative polymerases have been shown to incorporate several nucleotides, displacing several nucleotides into a flap that is removed

by Flap Endonuclease I (FEN1). Importantly, both long patch and short patch pathways result in deletion of a bulged nucleotide.

Bifunctional glycosylase/lyase activity was also observed for Nth1 catalysis of a DHU-bulge. Bifunctional activity nicks the bulged abasic site by a 3'-β-elimination to generate a 3' block. APE1 can act as a 3'-phosphodiesterase to generate the 3'OH that is required for DNA ligation. This pathway occurs independently of polymerase β, the lyase activity conferred by Nth1. Under the experimental conditions employed, both polymerase β-dependent and independent reactions occur. The balance between these pathways depends on the relative levels of APE1 and Nth1 activity.

DHU-bulge is efficiently converted to a single nucleotide deletion in HeLa nuclear extract

We observed robust processing of a DHU lesion in HeLa nuclear extract. Depletion of divalent metals prevents nuclease processing. We exploited this property to measure the rate of glycosylase excision and found that a DHU-bulge is processed at twice the steady state rate as a DHU-G mismatch.

Previous studies with Nth1 excision of a thymine glycol (Tg) lesion observed that the lower affinity Tg:A substrate displayed a higher steady state turnover rate than Tg:G. The authors proposed that the lower affinity substrate allows for a higher rate of product dissociation (44).

Implications of glycosylase-initiated deletion of bulged nucleotides

The ability of Nth1 to excise a bulged lesion is consistent with reports for other glycosylases (19, 20, 24). Therefore, efficient recognition of a bulged nucleotide appears to be a common for feature of glycosylases. It is possible that the flexible recognition of bulged nucleotides is a consequence of the need to bind to DNA in multiple

conformations to accommodate nucleotide flipping, in which case other base flipping enzymes might also have activity toward bulged DNA.

Glycosylase interaction with a bulged nucleotide is expected to interfere with high-fidelity MMR activity. As detailed in this study, glycosylase excision of a bulged nucleotide can initiate the permanent deletion of the bulged nucleotide by BER. BER deletion pathway would prevent +1 frameshift mutations but make permanent -1 frameshift mutations, creating a bias for -1 frameshift mutations (20).

REFERENCES

1. Kunkel, T.A., *Misalignment-mediated DNA synthesis errors.* Biochemistry, 1990. **29**(35): p. 8003-8011.
2. Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita, E. Terzaghi, and M. Inouye, *Frameshift mutations and the genetic code. This paper is dedicated to Professor Theodosius Dobzhansky on the occasion of his 66th birthday.* Cold Spring Harb Symp Quant Biol, 1966. **31**: p. 77-84.
3. Shearman, C.W., M.M. Forgette, and L.A. Loeb, *On the fidelity of DNA replication. Mechanisms of misincorporation by intercalating agents.* J Biol Chem, 1983. **258**(7): p. 4485-4491.
4. Strauss, B.S., *Frameshift mutation, microsatellites and mismatch repair.* Mutat Res, 1999. **437**(3): p. 195-203.
5. Iyer, R.R., A. Pluciennik, V. Burdett, and P.L. Modrich, *DNA mismatch repair: functions and mechanisms.* Chem Rev, 2006. **106**(2): p. 302-323.
6. Kunkel, T.A. and D.A. Erie, *DNA mismatch repair.* Annu Rev Biochem, 2005. **74**: p. 681-710.
7. Greene, C.N. and S. Jinks-Robertson, *Spontaneous frameshift mutations in Saccharomyces cerevisiae: accumulation during DNA replication and removal by proofreading and mismatch repair activities.* Genetics, 2001. **159**(1): p. 65-75.
8. Drotschmann, K., A.B. Clark, and T.A. Kunkel, *Mutator phenotypes of common polymorphisms and missense mutations in MSH2.* Curr Biol, 1999. **9**(16): p. 907-910.
9. Poynter, J.N., K.D. Siegmund, D.J. Weisenberger, T.I. Long, S.N. Thibodeau, N. Lindor, J. Young, M.A. Jenkins, J.L. Hopper, J.A. Baron, D. Buchanan, G. Casey, A.J. Levine, L. Le Marchand, S. Gallinger, B. Bapat, J.D. Potter, P.A. Newcomb, R.W. Haile, and P.W. Laird, *Molecular characterization of MSI-H colorectal cancer by MLHI promoter methylation, immunohistochemistry, and mismatch*

repair germline mutation screening. Cancer Epidemiol Biomarkers Prev, 2008. **17**(11): p. 3208-3215.
10. Jackson, A.L., R. Chen, and L.A. Loeb, *Induction of microsatellite instability by oxidative DNA damage.* Proc Natl Acad Sci U S A, 1998. **95**(21): p. 12468-12473.
11. Chan, K., S. Houlbrook, Q.M. Zhang, M. Harrison, I.D. Hickson, and G.L. Dianov, *Overexpression of DNA polymerase beta results in an increased rate of frameshift mutations during base excision repair.* Mutagenesis, 2007. **22**(3): p. 183-188.
12. Frosina, G., *Overexpression of enzymes that repair endogenous damage to DNA.* Eur J Biochem, 2000. **267**(8): p. 2135-2149.
13. Hofseth, L.J., M.A. Khan, M. Ambrose, O. Nikolayeva, M. Xu-Welliver, M. Kartalou, S.P. Hussain, R.B. Roth, X. Zhou, L.E. Mechanic, I. Zurer, V. Rotter, L.D. Samson, and C.C. Harris, *The adaptive imbalance in base excision-repair enzymes generates microsatellite instability in chronic inflammation.* J Clin Invest, 2003. **112**(12): p. 1887-1894.
14. Fan, J. and D.M. Wilson, 3rd, *Protein-protein interactions and posttranslational modifications in mammalian base excision repair.* Free Radic Biol Med, 2005. **38**(9): p. 1121-1138.
15. Fortini, P. and E. Dogliotti, *Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways.* DNA Repair (Amst), 2007. **6**(4): p. 398-409.
16. David, S.S. and S.D. Williams, *Chemistry of Glycosylases and Endonucleases Involved in Base-Excision Repair.* Chem Rev, 1998. **98**(3): p. 1221-1262.
17. Stivers, J.T. and Y.L. Jiang, *A mechanistic perspective on the chemistry of DNA repair glycosylases.* Chem Rev, 2003. **103**(7): p. 2729-2759.
18. Klapacz, J., G.M. Lingaraju, H.H. Guo, D. Shah, A. Moar-Shoshani, L.A. Loeb, and L.D. Samson, *Frameshift mutagenesis and microsatellite instability induced by human alkyladenine DNA glycosylase.* Mol Cell. **37**(6): p. 843-853.
19. Lyons, D.M. and P.J. O'Brien, *Efficient recognition of an unpaired lesion by a DNA repair glycosylase.* J Am Chem Soc, 2009. **131**(49): p. 17742-17743.
20. Lyons, D.M. and P.J. O'Brien, *Human base excision repair creates a bias toward -1 frameshift mutations.* J Biol Chem. **285**(33): p. 25203-25212.
21. Colgin, L.M., A.F. Hackmann, M.J. Emond, and R.J. Monnat, Jr., *The unexpected landscape of in vivo somatic mutation in a human epithelial cell lineage.* Proc Natl Acad Sci U S A, 2002. **99**(3): p. 1437-1442.
22. Zhang, Z. and M. Gerstein, *Patterns of nucleotide substitution, insertion and deletion in the human genome inferred from pseudogenes.* Nucleic Acids Res, 2003. **31**(18): p. 5338-5348.
23. Taylor, M.S., C.P. Ponting, and R.R. Copley, *Occurrence and consequences of coding sequence insertions and deletions in Mammalian genomes.* Genome Res, 2004. **14**(4): p. 555-566.
24. Zhao, X., N. Krishnamurthy, C.J. Burrows, and S.S. David, *Mutation versus repair: NEIL1 removal of hydantoin lesions in single-stranded, bulge, bubble, and duplex DNA contexts.* Biochemistry. **49**(8): p. 1658-1666.

25. McCullough, A.K., M.L. Dodson, and R.S. Lloyd, *Initiation of base excision repair: glycosylase mechanisms and structures.* Annu Rev Biochem, 1999. **68**: p. 255-285.
26. Doherty, A.J., L.C. Serpell, and C.P. Ponting, *The helix-hairpin-helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA.* Nucleic Acids Res, 1996. **24**(13): p. 2488-2497.
27. Ikeda, S., T. Biswas, R. Roy, T. Izumi, I. Boldogh, A. Kurosky, A.H. Sarker, S. Seki, and S. Mitra, *Purification and characterization of human NTH1, a homolog of Escherichia coli endonuclease III. Direct identification of Lys-212 as the active nucleophilic residue.* J Biol Chem, 1998. **273**(34): p. 21585-21593.
28. Kuo, C.F., D.E. McRee, R.P. Cunningham, and J.A. Tainer, *Crystallization and crystallographic characterization of the iron-sulfur-containing DNA-repair enzyme endonuclease III from Escherichia coli.* J Mol Biol, 1992. **227**(1): p. 347-351.
29. Thayer, M.M., H. Ahern, D. Xing, R.P. Cunningham, and J.A. Tainer, *Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure.* EMBO J, 1995. **14**(16): p. 4108-4120.
30. Fromme, J.C. and G.L. Verdine, *Structure of a trapped endonuclease III-DNA covalent intermediate.* EMBO J, 2003. **22**(13): p. 3461-3471.
31. Eide, L., L. Luna, E.C. Gustad, P.T. Henderson, J.M. Essigmann, B. Demple, and E. Seeberg, *Human endonuclease III acts preferentially on DNA damage opposite guanine residues in DNA.* Biochemistry, 2001. **40**(22): p. 6653-6659.
32. Katafuchi, A., T. Nakano, A. Masaoka, H. Terato, S. Iwai, F. Hanaoka, and H. Ide, *Differential specificity of human and Escherichia coli endonuclease III and VIII homologues for oxidative base lesions.* J Biol Chem, 2004. **279**(14): p. 14464-14471.
33. Ocampo, M.T., W. Chaung, D.R. Marenstein, M.K. Chan, A. Altamirano, A.K. Basu, R.J. Boorstein, R.P. Cunningham, and G.W. Teebor, *Targeted deletion of mNth1 reveals a novel DNA repair enzyme activity.* Mol Cell Biol, 2002. **22**(17): p. 6111-6121.
34. Chan, M.K., M.T. Ocampo-Hafalla, V. Vartanian, P. Jaruga, G. Kirkali, K.L. Koenig, S. Brown, R.S. Lloyd, M. Dizdaroglu, and G.W. Teebor, *Targeted deletion of the genes encoding NTH1 and NEIL1 DNA N-glycosylases reveals the existence of novel carcinogenic oxidative damage to DNA.* DNA Repair (Amst), 2009. **8**(7): p. 786-794.
35. Baldwin, M.R. and P.J. O'Brien, *Human AP endonuclease 1 stimulates multiple-turnover base excision by alkyladenine DNA glycosylase.* Biochemistry, 2009. **48**(25): p. 6022-6033.
36. Werneburg, B.G., J. Ahn, X. Zhong, R.J. Hondal, V.S. Kraynov, and M.D. Tsai, *DNA polymerase beta: pre-steady-state kinetic analysis and roles of arginine-283 in catalysis and fidelity.* Biochemistry, 1996. **35**(22): p. 7041-7050.
37. Pascal, J.M., P.J. O'Brien, A.E. Tomkinson, and T. Ellenberger, *Human DNA ligase I completely encircles and partially unwinds nicked DNA.* Nature, 2004. **432**(7016): p. 473-478.

38. Peoples, A.R., J. Lee, M. Weinfeld, J.R. Milligan, and W.A. Bernhard, *Yields of damage to C4' deoxyribose and to pyrimidines in pUC18 by the direct effect of ionizing radiation.* Nucleic Acids Res.
39. House, C.H. and S.L. Miller, *Hydrolysis of dihydrouridine and related compounds.* Biochemistry, 1996. **35**(1): p. 315-320.
40. Rahman, S.M., T. Baba, T. Kodama, M.A. Islam, and S. Obika, *Hybridizing ability and nuclease resistance profile of backbone modified cationic phosphorothioate oligonucleotides.* Bioorg Med Chem. **20**(13): p. 4098-4102.
41. Einolf, H.J. and F.P. Guengerich, *Fidelity of nucleotide insertion at 8-oxo-7,8-dihydroguanine by mammalian DNA polymerase delta. Steady-state and pre-steady-state kinetic analysis.* J Biol Chem, 2001. **276**(6): p. 3764-3771.
42. Nash, H.M., R. Lu, W.S. Lane, and G.L. Verdine, *The critical active-site amine of the human 8-oxoguanine DNA glycosylase, hOgg1: direct identification, ablation and chemical reconstitution.* Chem Biol, 1997. **4**(9): p. 693-702.
43. Hashimoto, K., Y. Tominaga, Y. Nakabeppu, and M. Moriya, *Futile short-patch DNA base excision repair of adenine:8-oxoguanine mispair.* Nucleic Acids Res, 2004. **32**(19): p. 5928-5934.
44. Marenstein, D.R., M.K. Chan, A. Altamirano, A.K. Basu, R.J. Boorstein, R.P. Cunningham, and G.W. Teebor, *Substrate specificity of human endonuclease III (hNTH1). Effect of human APE1 on hNTH1 activity.* J Biol Chem, 2003. **278**(11): p. 9005-9012.

APPENDIX C

Additional data figures to support Chapter 4

SUPPORTING RESULTS

Figure C-1. DNA substrates used in this study. The centrally located nucleotide (X) is either a 5,6-dihydrouracil lesion or an undamaged cytosine. DNA labeled with a 5'-fluorescein attached by a 6-aminohexyl linker (6-fam). Duplex DNA is produced by annealing to an unlabeled, complementary strand. The nucleotide at position (Y) is one of four canonical nucleotides or omitted to form a bulged nucleotide at position (X). A six nucleotide polyT extension was affixed (25mer(T6)) to allow simultaneous quantification with 25mer in steady state kinetic assay. Terminal phosphorothioate and propane modifications were incorporated for exonuclease resistance in HeLa cell extract, as well as increased length for thermal stability at 37°C (35mer). A single phosphorothioate linkage is incorporated internally for endonuclease resistance.

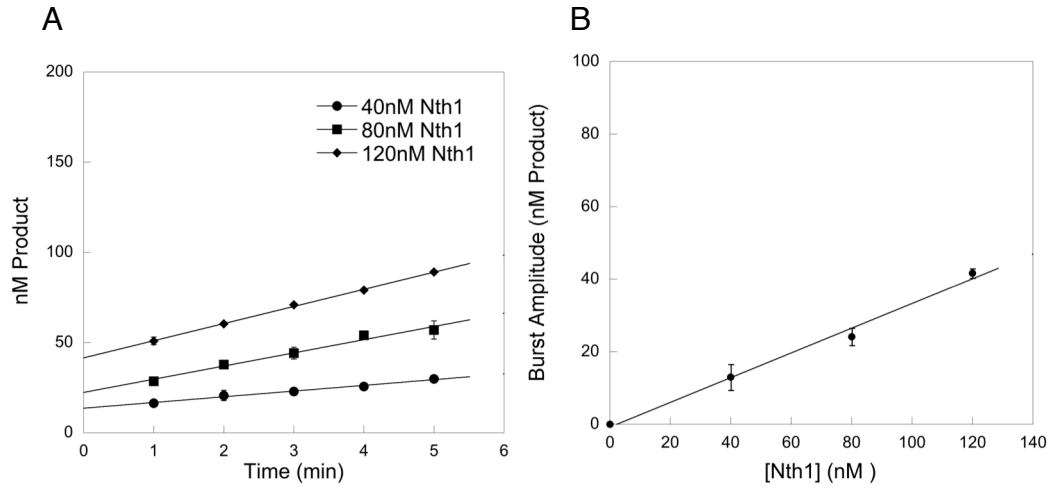

Figure C-2. Active concentration of recombinant Nth1 stock was determined by pre-steady state kinetic measurement under standard glycosylase assay conditions, 150mM [Na^+]. (A) A rapid burst of abasic site produce is formed during the first turnover of 200nM DHU·G, corresponding to the concentration of active Nth1 molecules. A slow, steady state rate of DHU excision follows the burst phase. (B) The y-intercept defines the burst amplitude and increases linearly with Nth1 concentration, at a slope of 0.33. Nth1 concentration was estimated to be 100μM by absorbance A_{280nm} and the active concentration was determined to be 33μM.

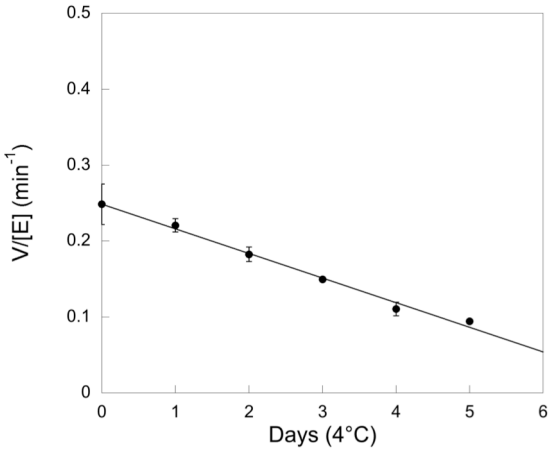

Figure C-3. Loss of Nth1 glycosylase activity at 4°C was quantified by measuring the initial rate of DHU-bulge excision. Separate aliquots of recombinant Nth1 (33μM) were thawed from -80°C stocks and stored at 4°C for the indicated time period. 20μl reactions containing 100nM DHU-bulge at 200mM [Na^+] were incubated with 2nM Nth1 and 3μl aliquots were quenched in 6μl 0.3mM NaOH and processed as previously described. A 7.5% decrease in activity occurred per day at 4°C, leading to a 38% decrease in glycosylase activity after 5 days. To alleviate the affects of Nth1 stability, aliquots of 33μM Nth1 were flash frozen and stored at -80°C. Aliquots were thawed and used immediately for assays.

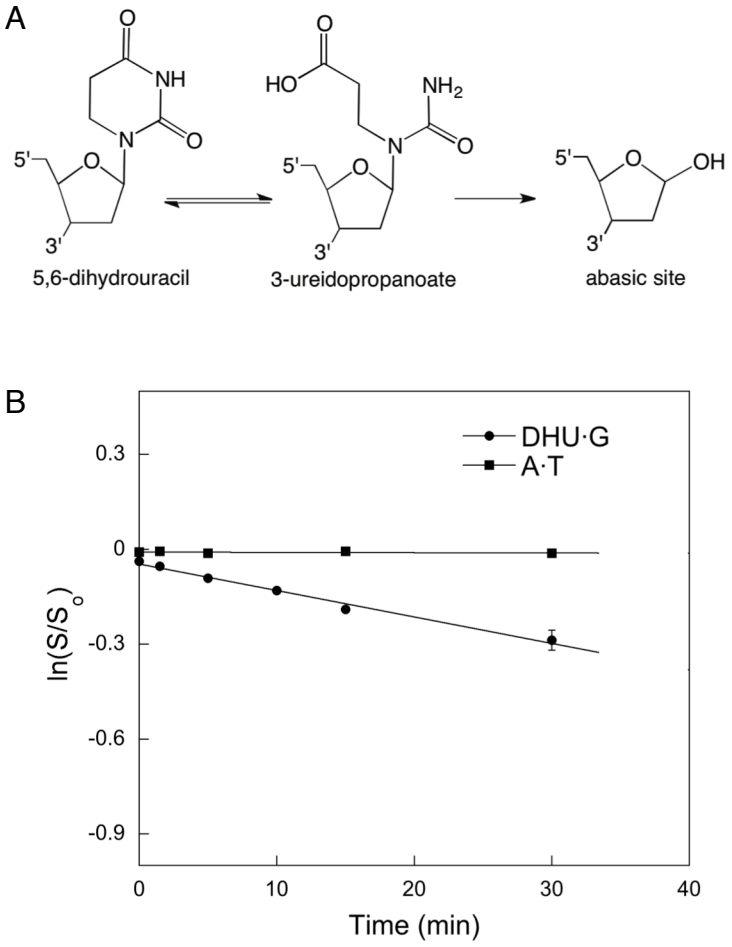

Figure C-4. Non-enzymatic degradation of a DHU nucleotide leads to the formation of an abasic site. (A) Ring-opening of 5,6-dihydrouracil forms a 3-uridopropanoate moiety that is susceptible to *N*-glycosidic bond cleavage resulting in an abasic site product. (B) The rate of abasic site formation from a DHU containing oligonucleotide (25mer) was quantified at 95°C in annealing buffer. Annealing buffer contains 10mM NaMES (pH 6.0) and 40mM NaCl. 1µM 5'-(fam)-labeled DHU containing strand was added with 1.5µM unlabeled complement strand and incubated at 95°C. Timepoints were taken by quenching 3µl aliquots in 6µl 0.3mM NaOH. Heating at 70°C for 10 min in alkaline conditions results in quantitative cleavage of abasic sites resulting from loss of DHU. In comparison, the stability of an oligonucleotide without a DHU lesion (A-T) was stable over entire timecourse. Rate of DHU degradation to abasic site was determined by fitting to a single exponential decay of substrate, eq C1, where S is fraction substrate at time, t. S_o is initial fraction of substrate and k is the rate of substrate decay. The rate of DHU degradation was 0.84% min^{-1} 95°C.

$$\ln(S/S_o) = -kt \qquad \text{(C1)}$$

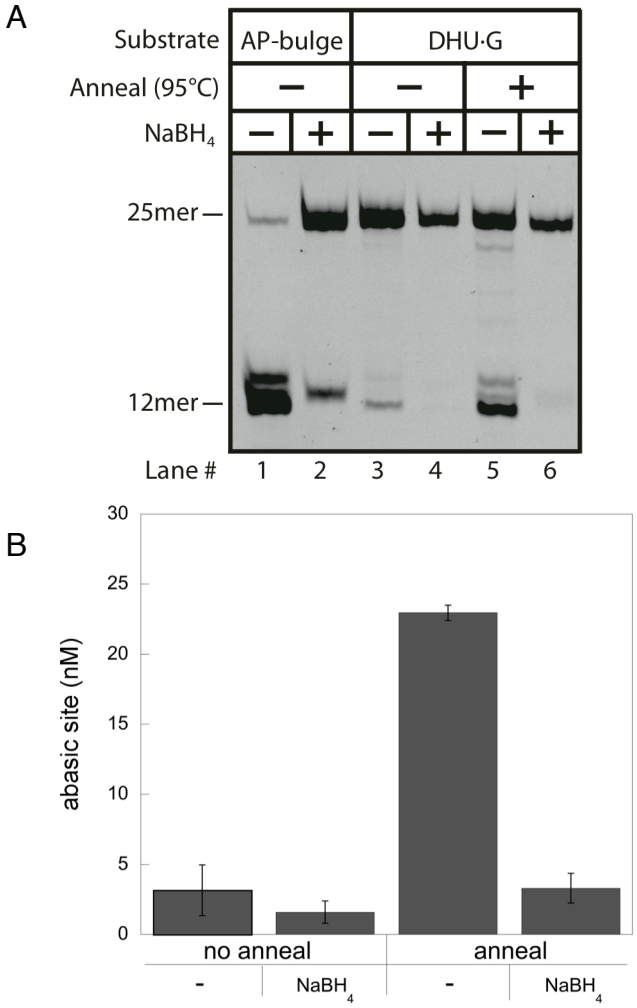

Figure C-5. Abasic site formation due to DHU degradation was quantitated by reduction of abasic sites with $NaBH_4$. A reduced abasic site is protected from NaOH cleavage. Standard protocol for annealing oligonucleotides (25mer) requires melting DNA strands at 95°C for 5 min followed by slow cooling on ice over 30 min. (A) 20μL reactions contained 1μM 5'(6-fam)-labeled forward strand containing a centrally located DHU lesion was incubated with 1.5μM unlabeled complementary oligonucleotide in annealing buffer (pH 6.0). Reactions were subjected to standard annealing protocol and then reduced by addition of 1μL 500mM $NaBH_4$ (in 10mM NaOH solution) on ice for 30 min. $NaBH_4$ reduction was quenched by adding 1μL 85.2mM acetic acid and allowed to sit for 30 min on ice. Finally, 8μL solution of formamide loading buffer (with 40mM NaOH) was added and samples analyzed by denaturing PAGE. Lane 2 demonstrates the stability of a reduced abasic site to cleavage under alkaline conditions. Lane 5 shows significant formation of abasic sites occur during the standard annealing protocol, which are protected by $NaBH_4$ reduction in lane 6. (B) Quantification of three independent reactions illustrates that over 20% of DHU lesions are degraded to abasic sites by the standard annealing protocol.

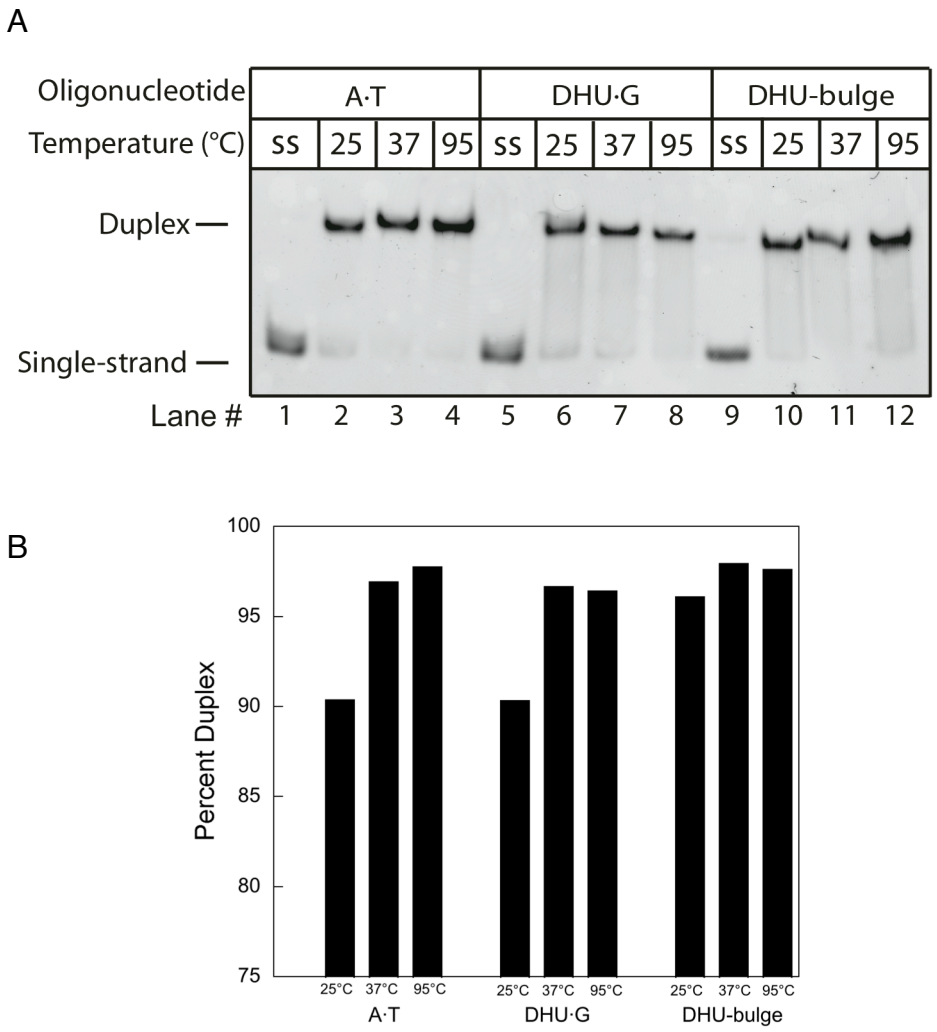

Figure C-6. Annealing of 25mer oligonucleotides is performed by melting DNA at elevated temperature for 5 min, followed by slow cooling over 30 min on ice to promote correct hybridization of duplex DNA structure. (A) DHU is unstable at high temperatures, so the lowest melting temperature was selected that facilitates annealing. (A) 1μM 5'-(6-fam)-labeled oligonucleotide was incubated with 1.5μL unlabeled complement strand in annealing buffer (pH 6.0) and subjected to the standard annealing protocol at various melting temperatures. A non-denaturing native PAGE gel (20% acrylimide) demonstrated that 37°C allowed sufficient annealing of oligonucleotides while lessening degradation of DHU. Single-strand markers did not include complementary strand. (B) Quantification of native gel.

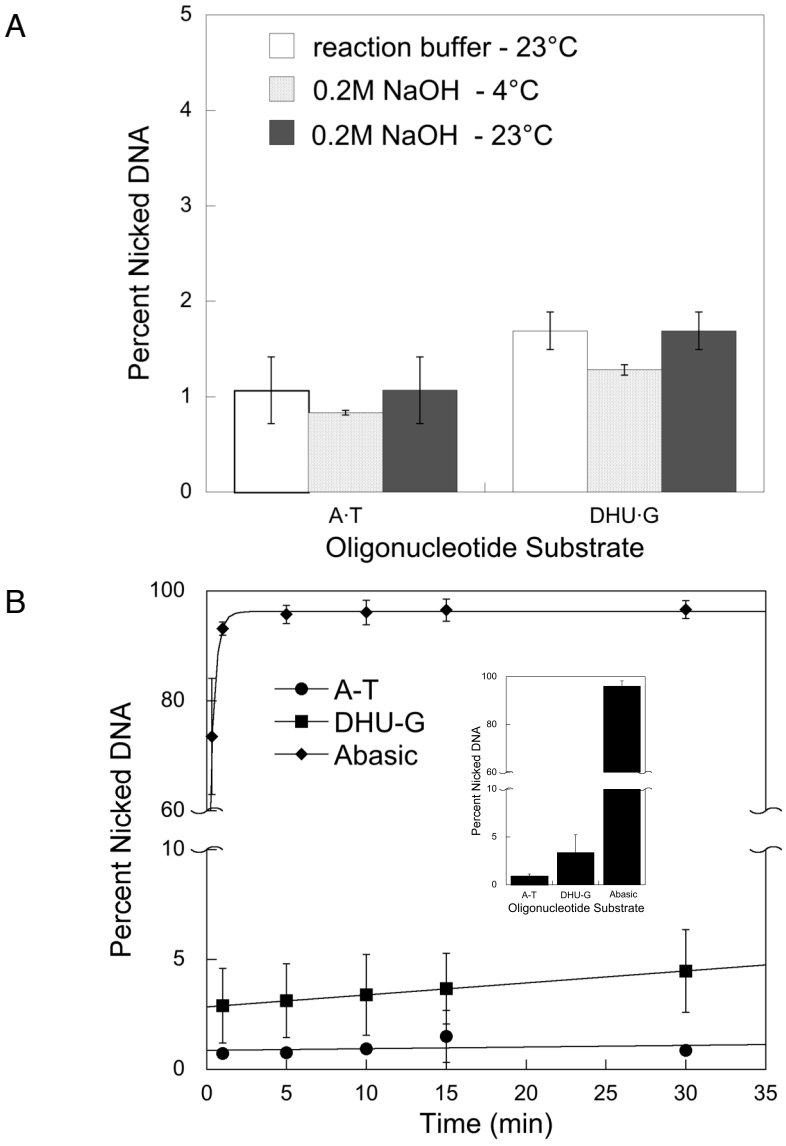

Figure C-7. Determination of DHU stability in reaction conditions and alkaline workup. (A) 1μM of the 25mer oligonucleotide containing a centrally located DHU·G mismatch or undamaged A·T basepair was incubated under conditions utilized for assaying Nth1 activity. After 24 hours, reactions in standard reaction buffer (23°C) were quenched into a final concentration of 0.2M NaOH and heated at 70°C for 10 min to cleave any abasic sites formed during the reaction. Reactions already in 0.2M NaOH were processed the same. Production of nicked DNA is indicative of abasic site formation and was quantified by denaturing PAGE analysis. These results demonstrate that a DHU lesion is stable in reaction conditions. (B) Typically, quantification of glycosylase activity requires cleavage of abasic sites in 0.2M NaOH at 70°C. At 70°C in 0.2M NaOH, abasic sites are formed by from a DHU lesion at a rate of 0.06% min^{-1}. A burst of 2.85% is indicative of abasic sites already present in the DHU·G stock and results in a formation of 4.5% non-enzymatically produced nicked intermediate after 30 min. 10 min at 70°C was chosen as standard protocol for alkaline cleavage of abasic sites. Inset graph shows the percentage of non-enzymatically produced nicked DNA present at 10 min timepoint.

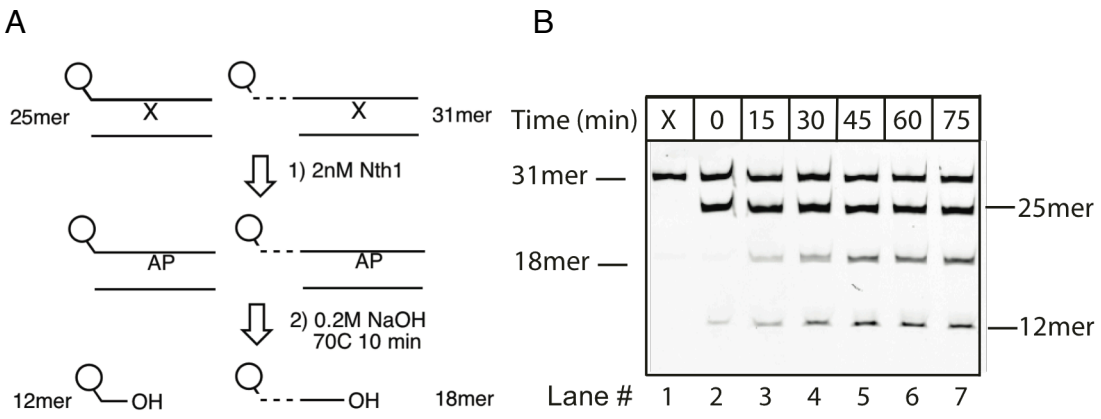

Figure C-8. Relative k_{cat}/K_m values for a DHU lesion in different structural conformations were determined by competition. (A) A 6-nucleotide, polyT tract was attached to the 5'end of the 25mer oligonucleotide (see Fig. 1 for sequences) to allow simultaneous quantitation of two different substrates during steady state turnover. Abasic sites produced by glycosylase activity, as well as nicked intermediates produced by bifunctional AP-lyase activity, are cleaved to a 3'OH nick in alkaline conditions. (B) A typical denaturing polyacrylimide gel showing the timecourse of timecourse of a reaction with DHU·G (25mer) and DHU·T(T6) (31mer). Each oligonucleotide was present at 100nM and excision of DHU was catalyzed by 2nM Nth1.

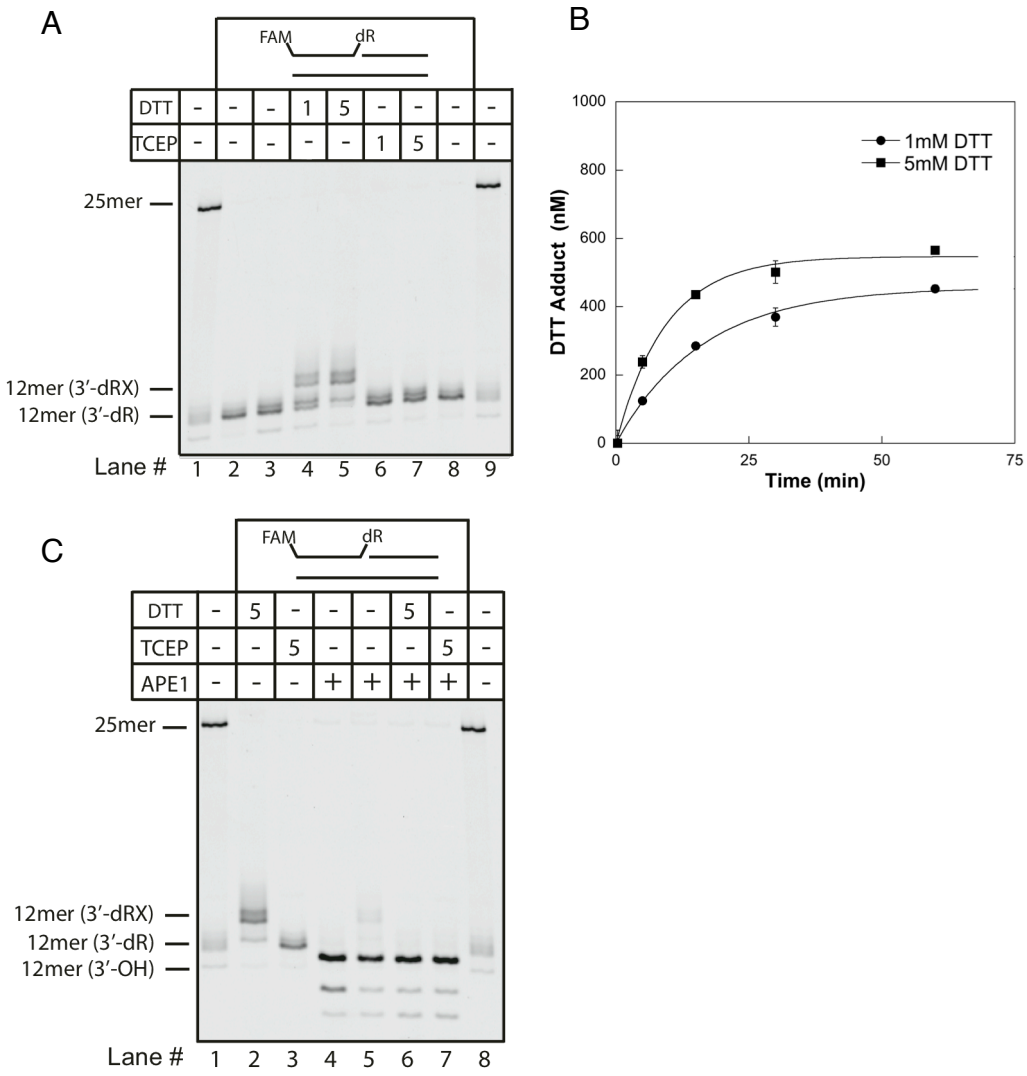

Figure C-9. Bifunctional glycosylase/lyase activity produces a nick with a 3'-unsaturated sugar (3'-dR). Standard buffer conditions include 1mM dithiothreitol (DTT), which reacts with the α,β-unsaturated aldehyde to produce a lower mobility species (3'-dRX) in a bimolecular dependence (lanes 4 and 5). Incubating Nth1 with a DHU-bulge substrate produced a 3' -dR nicked intermediate that was extracted using phenol-chloroform and a buffer exchange column. 1μM 3'-dR intermediate was incubated in standard reaction conditions at 23°C. The reaction was quenched in formamide/EDTA loading buffer at 60 min. Substituting tris,2-carboxyethylphosphine (TCEP) for DTT prevents modification of the 3'dR intermediate (lanes 6 and 7). (B) Quantification of DTT reactivity with 3'-dR moiety at 23°C. Best fit of data to a single exponential with a rate of $0.064min^{-1}$ and $0.11min^{-1}$ for 1mM and 5mM DTT, respectively. (C) Addition of 50nM APE1 (with 5mM $MgCl_2$) demonstrates 3'-exonucleolytic activity. APE1 removed the unsaturated sugar moiety and produced a 3'-OH intermediate within 60 min at 23°C (lane 4). Phosphodiesterase activity is robust enough to remove the DTT adduct (3'-dRX) (lane 6).

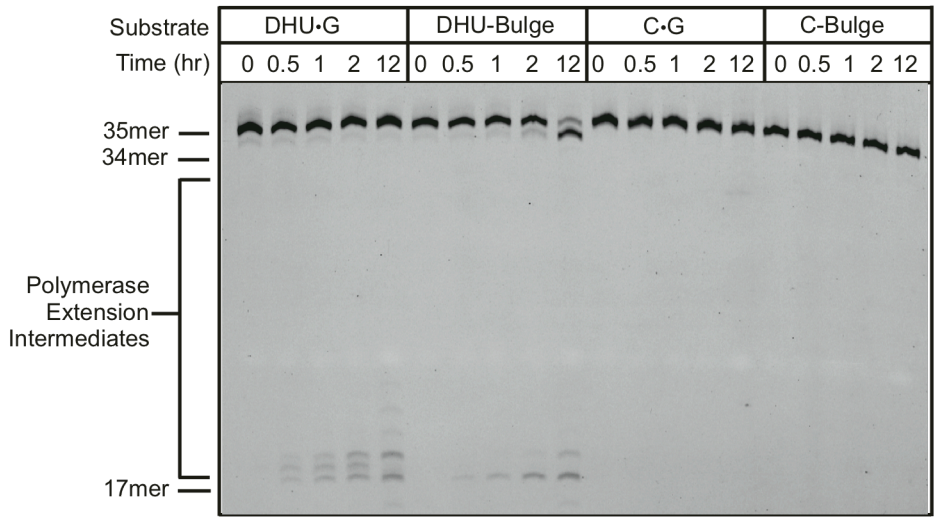

Figure C-10. Incubation of oligonucleotides containing a DHU lesion showed BER activity in HeLa whole cell extract (WCE). Complementary to figure 4-9 from the main text, showing BER processing in HeLa nuclear extract (NE), 10nM oligonucleotide was incubated with 0.4 mg/ml HeLa WCE (Active Motif) at 37°C along with all cofactors necessary for BER activity. Reactions were quenched with formamide/EDTA loading buffer. Resolution by denaturing PAGE displays both formation of nicked BER intermediates (17mer) as well as production of single nucleotide deletion from a DHU-bulge.

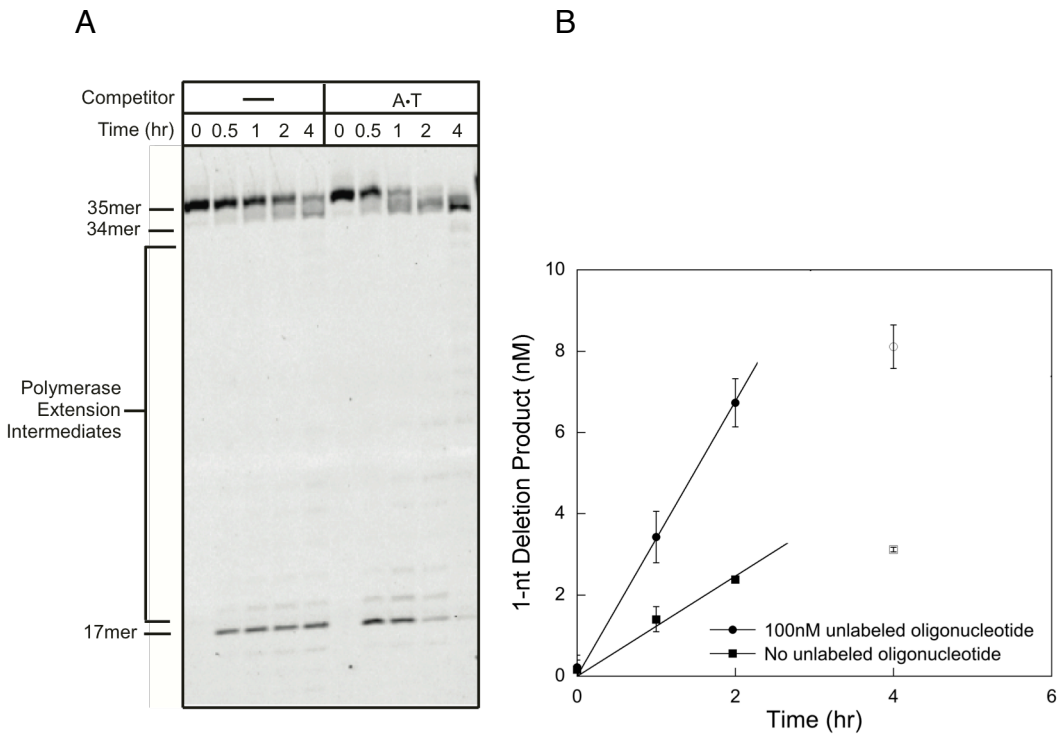

Figure C-11. Addition of 100nM competitor oligonucleotide in addition to 10nM DHU-bulge (35mer) substrate in HeLa NE facilitates more rapid and more complete BER processing of a DHU-bulge (35mer) to a deletion product (34mer). Competitor oligonucleotide is an unlabeled, unmodified 25mer and may serve to occupy some amount of non-specific DNA binding proteins in HeLa NE and allow BER to process the DHU lesion. (B) Quantitation of deletion product (34mer) with and without addition of competitor DNA. Points represent average of duplicate experiments (± standard deviation). Initial rate of deletion formation is increased 3-fold in conjunction with competitor oligonucleotide as shown by a linear initial rate (filled circles). In addition, the end-point of deletion product formation is increased (open-circles).

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

The burden of damaged DNA poses a risk to genomic stability. A diverse array of DNA repair pathways have evolved to both repair damaged DNA, as well as correct replication errors. Insufficient DNA repair activity has been linked to an increased accumulation of mutations, leading to an increased risk of cancer (1, 2). For example, the hereditary or somatic downregulation of mismatch DNA repair (MMR) activity is well-established link to genomic instability (3). Specifically, regions of repetitive sequences, termed microsatellites, become unstable in length due to increased frequency of insertion and deletion mutations. Microsatellite instability (MSI) is a hallmark of MMR deficiency and is used to clinically identify MMR deficient colon cancer (4, 5).

It is known that polymerase slipping errors occur at elevated rates in microsatellite repeats, producing a bulged nucleotide that is an intermediate to a frameshift mutation (6, 7). A polymerase slipping error can insert an extra (+1 slip) or omit (-1 slip) a nucleotide in reference to the template strand during replication. Importantly, replication coupled MMR has the ability to differentiate the newly synthesized strand and provide accurate repair of either +1 or -1 slips. Downregulation of MMR fails to provide adequate repair of bulged intermediates and results in a higher frequency of frameshift mutations (8, 9).

Paradoxically, upregulation of a DNA repair glycosylase, AAG, was correlated to MSI in patients with chronic inflammation of the colon (ulcerative colitis) (10). This observation was supported with studies to quantify the impact of AAG overexpression on frameshift mutation frequency. Overexpression of AAG produced a 6-fold increase in frequency of insertion (+1 frameshift) mutations and a 30-fold increase in deletion (-1 frameshift) mutations. The proposed model accused hyperactive AAG activity for producing an excess of BER intermediates that overwhelmed downstream BER processing, and BER intermediates are known to be mutagenic (11, 12) (Figure 5-1A). Co-overexpression of the downstream BER enzyme, APN1 endonuclease, was predicted to alleviate the increased mutation rate by removing excess BER intermediates caused by AAG. Surprisingly, APN1 did not decrease the frameshift rate (10). These results are in contrast with a previous report documenting that the burden of BER intermediates caused by overexpression the yeast functional homolog of AAG, 3-methyladenine DNA glycosylase (MAG1), was alleviated by co-overexpression of APN1 (13). The inability of APN1 to rescue the increased frameshift mutation rate in the presence of AAG overexpression suggests that MAG1 and AAG induce frameshift mutations by differing mechanisms.

Chapter 2 of this thesis provides an alternative model for AAG induced frameshift mutations. Following the initial observation that several different bulged lesions were excised in HeLa cell extract, reconstitution of the BER pathway *in vitro* demonstrated the ability of BER to convert a bulged lesion into single nucleotide deletion. The distinct feature of the BER deletion pathway is the ability of a glycosylase to excise a bulged

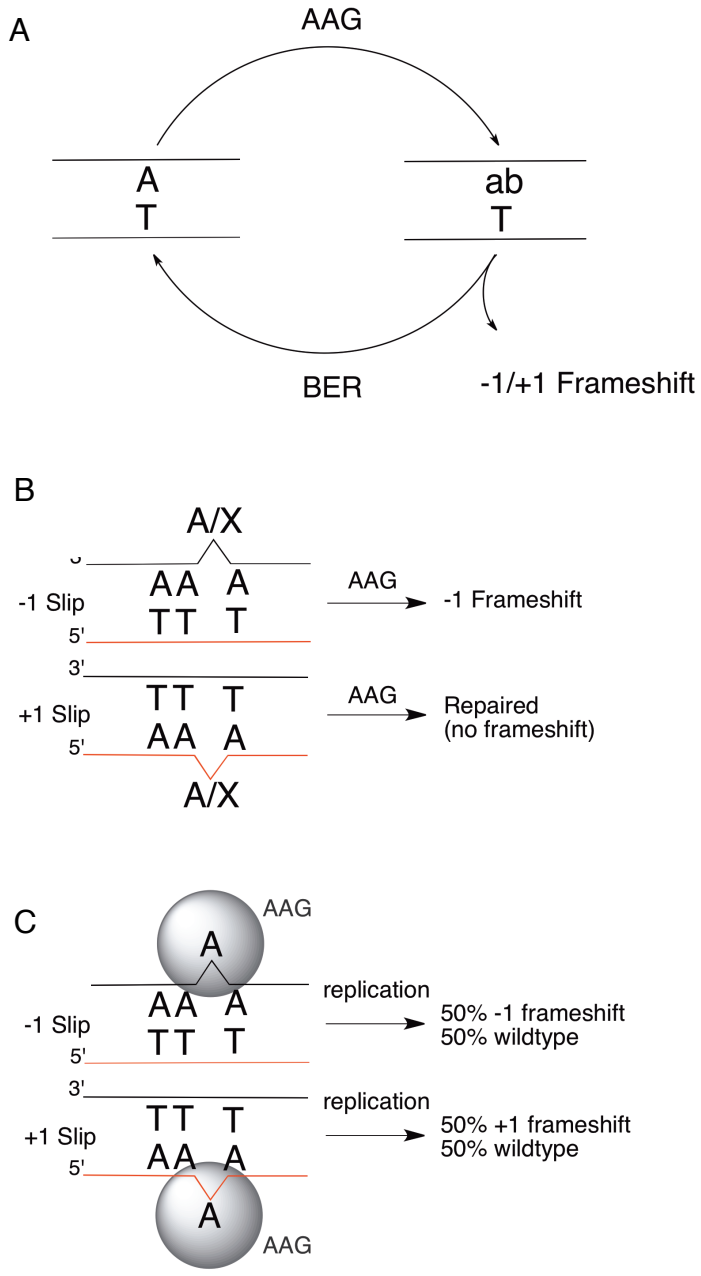

Figure 5-1. Models to account for AAG overexpression induced frameshift mutations. (A) Hyperactive AAG produces excess of abasic sites (AP) in genome that overwhelms downstream BER processing. Unprocessed abasic sites induce frameshift mutations. (B) AAG excises damaged (X) or undamaged bulge to create a bias for -1 frameshift mutations by repairing +1 slip and thereby preventing +1 frameshift mutations. Template strand is indicated in black, newly synthesized strand is red. (C) AAG binds to bulges and prevents repair by MMR. A subsequent round of replication generates one mutant copy and one wildtype copy.

nucleotide. Once the bulged nucleotide is excised, the DNA backbone can be cleaved at the resulting abasic site. The absence of a template base opposing the bulged nucleotide allows the site to be ligated without addition of a new nucleotide, resulting in the deletion of the bulged nucleotide (14).

There are 11 known human glycosylases, each providing recognition of a single or a subset of damaged nucleotides. Therefore, the impact of the BER deletion pathway on the frequency of frameshift mutations would likely be influenced by both the number of glycosylases capable of excising a bulged nucleotide, as well as the identity of the bulged nucleotide. Characterization of AAG in chapter 3 showed the excision of an inosine lesion is 3-fold more efficient in a bulge than in a mismatch. Therefore, AAG has no inherent ability to discriminate against excising a bulged nucleotide. In contrast, a similar characterization of a different human glycosylase, Nth1, in chapter 4 showed a 20-fold discrimination against excision of a bulged lesion. These two examples, in combination with additional reports, provide evidence that excision of a bulged nucleotide may be a common feature of glycosylase catalysis (15, 16). All glycosylases require the target nucleotide to be flipping out of the duplex DNA and into the active site prior to catalyzing nucleotide excision, and a bulged nucleotide may lower the energetic barrier to baseflipping. Nevertheless, distortions in the DNA structure induced by a bulged nucleotide structure does affect glycosylase catalysis to various extents, although it is unclear what steps in glycosylase catalysis are being affected by the bulged nucleotide.

Ultimately, BER deletion of a single nucleotide bulge would produce a bias for -1 frameshift mutations by repairing +1 slipping errors (insertion) and making -1 slipping

errors (deletion) permanent (Figure 5-1B). Therefore, the BER deletion model can account for the bias for -1 frameshifts observed with AAG overexpression, but does not explain increased formation of +1 mutations. A model that does account for formation of both +1 (insertion) and -1 (deletion) frameshift mutations was proposed (17). This model incorporates AAG binding with high affinity to a bulged intermediate, blocking repair activity by MMR. A subsequent round of replication would convert the unrepaired bulged intermediate into a frameshift mutation. Consequentially, AAG would block repair of both +1 and -1 slipping errors and lead to the formation of both +1 and -1 frameshift mutations (Figure 5-1C).

These two models do not need to be mutually exclusive. A combination of both models would elegantly explain the observed affect of AAG overexpression on frameshift mutation rates. The binding of AAG to bulged intermediates without excision would explain the elevated rate both +1 and -1 frameshift mutations due to disruption of MMR activity. The observed bias for -1 frameshift mutations can be accounted for by AAG excising the bulged intermediate and allowing BER to delete the bulged intermediate. The contribution of either pathway to BER mediated frameshift mutations can be addressed with further experiments.

Classically, a genetic reversion assay in *S. cerevisiae* based on auxotrophic selectable marker has been utilized for measuring the frequency of frameshift mutations (18-20). This well-established tool for measuring the frameshift mutations rate utilizes an engineered homopolymeric run into the *lys2* gene. Lys2 is an α-aminoadipate reductase, playing an essential role in biosynthesis of lysine (21). Functional Lys2 can be selected against by addition of α-aminoadipic acid. α-aminoadipic acid is metabolized by Lys2

reductase activity to build toxic levels of the δ-semialdehyde intermediate. Alternatively, Lys2 function is auxotrophic selectable by growth on minimal media lacking lysine (20).

The affect of AAG overexpression has been documented using the Lys2 genetic reversion assay (10). This study reported a 6-fold increase in +1 and a 30-fold increase in -1 frameshift mutations in the presence of AAG overexpression. In order to measure reversion of both +1 and -1 frameshift mutations, two strains are constructed. Both strains are designed to have *lys2* shifted one nucleotide out of frame, reverting to functional Lys2 after incurring a frameshift mutation (20).

Expanding on these results, conditions can be perturbed to enhance the bias for different models of AAG induced frameshift mutations. Inactivating the glycosylase activity of AAG by mutation of an essential glutamate (E125A) disallows processing through the BER deletion pathway. Inactive AAG overexpression resulted in a significant decrease in the frequency of frameshift mutations and reduction in the bias for -1 frameshift mutations (17).

Alternatively, elevated levels of DNA damaging chemicals would be expected to enhance the bias for -1 frameshift mutations by enabling higher flux through the BER deletion pathway. As stated above, the bulged nucleotide can be more susceptible to chemical reactions with DNA damaging agents and damaged nucleotides are excised at a much higher rate than undamaged, bulged nucleotides (14). Chloroacetaldehyde has proven specificity for damaging bulged nucleotides and can be added to yeast growth media to increase the probability of producing a bulged lesion. However, the cellular chloroacetaldehyde concentration must be optimized to increase the probability of

creating a bulged lesion while minimizing adverse damage to the genome and other cellular biomolecules.

Alternatively, a single mutation in AAG has been shown to significantly increase the rate of undamaged guanine excision (22). This mutant (N169S) could be employed to enhance the glycosylase excision pathway in the absence of DNA damaging agents due to excision of undamaged, bulged guanine. Characterization of the N169S mutant showed an inverse correlation with basepairing stability and efficiency of guanine excision, a property that would be expected to aid in selection for bulged guanine excision. Nevertheless, the expression of N169S mutant protein must be carefully controlled to avoid excess excision of basepaired guanine throughout the genome.

An important limitation of measuring frameshift frequency by genetic reversion is that only a single site is tested in a single experiment. Next generation sequencing techniques are an attractive approach to quantify the frameshift mutations over the entire genome. This approach sequences an extremely large number of short lengths of the genome (100 nucleotides) in parallel. The sequences are then mapped to a reference genome. Recently, the technological advances in both the technique and analysis of next generation sequencing will facilitate accurate determination of indel mutations in homopolymeric and other repetitive sequences.

An advantage of whole genome sequencing is the ability to correlate the contribution of genomic features to frameshift mutation frequency. Specifically, nucleosome positioning, epigenetic markings, and gene transcription rates have been characterized across the genome and could utilized to correlate frameshift mutation frequency with features of the genome (23). For example, DNA repair activities are

restricted on nucleosomal DNA (24, 25). Similarly, nucleosome positioning has been shown to influence the exposure of the DNA to damaging agents (26, 27). Taken together, the overall mutation rate at a specific genomic position may be governed by the level of exposure to DNA damaging agents as well as the ability of DNA repair to function efficiently.

Correlation of mutation rates with features of the genome requires consideration of the dynamics of the genome, as many features of the genome are not static. Some changes are transient, for example, a replication event removes histone structures from the DNA (27). Additionally, cell cycle dependent fluctuations of DNA repair activity and exposure of DNA to damaging agents may provide significant, yet transient alterations in mutation rates. The most intriguing influences to frameshift mutation rate, uncovered by whole genome sequencing, can be studied in greater detail by employing the established *in vitro* assays for detecting single nucleotide deletions (14).

In conclusion, these models supply explanation of how the overexpression of DNA glycosylases can lead to increased mutation rates, highlighting that the glycosylase's unique ability to recognize a structurally diverse range of substrates must be carefully balanced to avoid superfluous activity leading to mutations. Downregulation of DNA repair activity is well-established to increase the mutation rate due to mutagenic intermediates remaining in the genome in the absence of repair. Surprisingly, the upregulation of BER repair pathway can lead to an imbalance in repair activities and conditions that reduce the probability of correct repair of frameshift mutation intermediates. Therefore, maintaining the highest level of genome stability requires

carefully balanced levels of DNA repair pathway activity, ensuring that the appropriate pathway is utilized to provide the greatest fidelity of repair.

REFERENCES

1. Cook, P.J., R. Doll, and S.A. Fellingham, *A mathematical model for the age distribution of cancer in man.* Int J Cancer, 1969. **4**(1): p. 93-112.
2. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.
3. Drotschmann, K., A.B. Clark, H.T. Tran, M.A. Resnick, D.A. Gordenin, and T.A. Kunkel, *Mutator phenotypes of yeast strains heterozygous for mutations in the MSH2 gene.* Proc Natl Acad Sci U S A, 1999. **96**(6): p. 2970-2975.
4. Hawn, M.T., A. Umar, J.M. Carethers, G. Marra, T.A. Kunkel, C.R. Boland, and M. Koi, *Evidence for a connection between the mismatch repair system and the G2 cell cycle checkpoint.* Cancer Res, 1995. **55**(17): p. 3721-3725.
5. Drotschmann, K., A.B. Clark, and T.A. Kunkel, *Mutator phenotypes of common polymorphisms and missense mutations in MSH2.* Curr Biol, 1999. **9**(16): p. 907-910.
6. Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita, E. Terzaghi, and M. Inouye, *Frameshift mutations and the genetic code. This paper is dedicated to Professor Theodosius Dobzhansky on the occasion of his 66th birthday.* Cold Spring Harb Symp Quant Biol, 1966. **31**: p. 77-84.
7. Tippin, B., S. Kobayashi, J.G. Bertram, and M.F. Goodman, *To slip or skip, visualizing frameshift mutation dynamics for error-prone DNA polymerases.* J Biol Chem, 2004. **279**(44): p. 45360-45368.
8. Modrich, P., *Mechanisms and biological effects of mismatch repair.* Annu Rev Genet, 1991. **25**: p. 229-253.
9. Palombo, F., P. Gallinari, I. Iaccarino, T. Lettieri, M. Hughes, A. D'Arrigo, O. Truong, J.J. Hsuan, and J. Jiricny, *GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells.* Science, 1995. **268**(5219): p. 1912-1914.
10. Hofseth, L.J., M.A. Khan, M. Ambrose, O. Nikolayeva, M. Xu-Welliver, M. Kartalou, S.P. Hussain, R.B. Roth, X. Zhou, L.E. Mechanic, I. Zurer, V. Rotter, L.D. Samson, and C.C. Harris, *The adaptive imbalance in base excision-repair enzymes generates microsatellite instability in chronic inflammation.* J Clin Invest, 2003. **112**(12): p. 1887-1894.
11. Clauson, C.L., K.J. Oestreich, J.W. Austin, and P.W. Doetsch, *Abasic sites and strand breaks in DNA cause transcriptional mutagenesis in Escherichia coli.* Proc Natl Acad Sci U S A. **107**(8): p. 3657-3662.
12. Mourgues, S., M.E. Lomax, and P. O'Neill, *Base excision repair processing of abasic site/single-strand break lesions within clustered damage sites associated with XRCC1 deficiency.* Nucleic Acids Res, 2007. **35**(22): p. 7676-7687.
13. Begley, T.J., S.A. Jelinsky, and L.D. Samson, *Complex transcriptional responses to macromolecular damaging agents: regulatory responses specific for SN2*

alkylation and the MAG1 gene. Cold Spring Harb Symp Quant Biol, 2000. **65**: p. 383-393.
14. Lyons, D.M. and P.J. O'Brien, *Human base excision repair creates a bias toward -1 frameshift mutations.* J Biol Chem. **285**(33): p. 25203-25212.
15. Lyons, D.M. and P.J. O'Brien, *Efficient recognition of an unpaired lesion by a DNA repair glycosylase.* J Am Chem Soc, 2009. **131**(49): p. 17742-17743.
16. Zhao, X., N. Krishnamurthy, C.J. Burrows, and S.S. David, *Mutation versus repair: NEIL1 removal of hydantoin lesions in single-stranded, bulge, bubble, and duplex DNA contexts.* Biochemistry. **49**(8): p. 1658-1666.
17. Klapacz, J., G.M. Lingaraju, H.H. Guo, D. Shah, A. Moar-Shoshani, L.A. Loeb, and L.D. Samson, *Frameshift mutagenesis and microsatellite instability induced by human alkyladenine DNA glycosylase.* Mol Cell. **37**(6): p. 843-853.
18. Tran, H.T., N.P. Degtyareva, N.N. Koloteva, A. Sugino, H. Masumoto, D.A. Gordenin, and M.A. Resnick, *Replication slippage between distant short repeats in Saccharomyces cerevisiae depends on the direction of replication and the RAD50 and RAD52 genes.* Mol Cell Biol, 1995. **15**(10): p. 5607-5617.
19. Tran, H.T., J.D. Keen, M. Kricker, M.A. Resnick, and D.A. Gordenin, *Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants.* Mol Cell Biol, 1997. **17**(5): p. 2859-2865.
20. Harfe, B.D. and S. Jinks-Robertson, *Removal of frameshift intermediates by mismatch repair proteins in Saccharomyces cerevisiae.* Mol Cell Biol, 1999. **19**(7): p. 4766-4773.
21. Morris, M.E. and S. Jinks-Robertson, *Nucleotide sequence of the LYS2 gene of Saccharomyces cerevisiae: homology to Bacillus brevis tyrocidine synthetase 1.* Gene, 1991. **98**(1): p. 141-145.
22. O'Brien, P.J. and T. Ellenberger, *Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase.* J Biol Chem, 2004. **279**(11): p. 9750-9757.
23. Wei, G., G. Hu, K. Cui, and K. Zhao, *Genome-wide mapping of nucleosome occupancy, histone modifications, and gene expression using next-generation sequencing technology.* Methods Enzymol. **513**: p. 297-313.
24. Prasad, A., S.S. Wallace, and D.S. Pederson, *Initiation of base excision repair of oxidative lesions in nucleosomes by the human, bifunctional DNA glycosylase NTH1.* Mol Cell Biol, 2007. **27**(24): p. 8442-8453.
25. Hinz, J.M., Y. Rodriguez, and M.J. Smerdon, *Rotational dynamics of DNA on the nucleosome surface markedly impact accessibility to a DNA repair enzyme.* Proc Natl Acad Sci U S A. **107**(10): p. 4646-4651.
26. Enright, H.U., W.J. Miller, and R.P. Hebbel, *Nucleosomal histone protein protects DNA from iron-mediated damage.* Nucleic Acids Res, 1992. **20**(13): p. 3341-3346.
27. Lucchini, R., R.E. Wellinger, and J.M. Sogo, *Nucleosome positioning at the replication fork.* EMBO J, 2001. **20**(24): p. 7294-7302.