Mechanisms for DNA Damage Recognition by Alkyladenine DNA Glycosylase

by

Adam Z. Thelen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biological Chemistry)
in The University of Michigan
2021

Doctoral Committee:

Associate Professor Patrick J. O'Brien, Chair
Associate Professor Bruce Palfey
Professor Lyle Simmons
Associate Professor Raymond Trievel
Professor Nils G. Walter

Adam Z. Thelen

azthelen@umich.edu

ORCID iD: 0000-0001-8406-4164

ACKNOWLEDGMENTS

First, I would like to thank my mentor Pat O'Brien for his constant guidance and expertise, as well as the open and collaborative lab environment he has cultivated at the University of Michigan. From the normal challenges of graduate school to an unprecedented global pandemic, his support and understanding have been unwavering.

I would also like to thank the other members of the O'Brien lab, both past and present, for their support, criticism, and friendship over the years. In particular, I would like to thank Michael Baldwin for his dedicated management of the lab and willingness to answer any question, no matter how obvious or complex. I would also like to thank Jenna Hendershot for her work characterizing damage recognition by AAG, setting the stage for much of my work. Finally, I would like to thank Mengshu He, a wonderful undergraduate student who persevered through years of challenges and incremental optimizations to help develop the DNA bending assay used in this thesis.

I would also like to give my thanks to Yan Zhang and Zhonggang Hou from the Zhang lab for their collaboration during and following my sabbatical in their lab. Their instruction and assistance in human cell culture made the entire HAP1 cell project possible. An enormous thanks to all the other members of the Zhang lab for their flexibility and generosity in making this collaboration so smooth.

None of this work would have been possible without the ongoing support of my family and friends. I would like to thank my parents Bob and Chris Thelen as well as my sister Hannah for their continual encouragement. And finally, my deepest thanks to my wife Sarah for her patience, sacrifice and inspiration all these years. Your support has carried me from application to graduation.

TABLE OF CONTENTS

List of Figures

List of Tables

List of Appendices

List of Abbreviations

A: Adenine

AAG: Alkyladenine DNA glycosylase (also known as MPG)

AlkA: 3-methyladenine DNA glycosylase II

AP: Apurinic/Apyrimidinic

APE1: Human AP endonuclease 1

BER: Base excision repair

BSA: Bovine serum albumin

C: Cytosine

DMEM: Dulbecco's modified eagle medium

dRP: deoxyribosephosphate

DRR: Direct reversal repair

DTT: Dithiothreitol

EDTA: Ethylenediamineteracetic acid

EGTA: Ethylene glycol bis-(2-aminoethyl ether)-N,N,N',N'-tetraacetic acid

ES cells: Embryonic stem cells

εA: 1,N^6-ethenoadenine

εC: 3,N^4-ethenocytosine

εG: 1,N^2-ethenoguanine

FAM: 6-fluorescein

FBS: Fetal bovine serum

FRET: Fluorescence resonance energy transfer (or Förster resonance energy transfer)

G: Guanine

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HMG1: High mobility group 1 protein

HR: Homologous repair

Hx: Hypoxanthine

ICL: Interstrand cross-link

IMDM: Iscove's modified Dulbecco media

MAG1: Yeast alkylation glycosylase

MES: 2-(N-morpholino)ethanesulfonic acid

MMR: Mismatch repair

MMS: Methyl methanesulfonate

MPG: Methylpurine DNA glycosylase (also known as AAG)

MutS: Mutator S protein

NER: Nucleotide excision repair

NHEJ: Non-homologous end joining

NMR: Nuclear magnetic resonance

OGG1: 8-oxoguanine DNA glycosylase

PAGE: Polyacrylamide gel electrophoresis

PBS: Phosphate-buffered saline

PDB: Protein data bank

S.D.: Standard deviation

siRNA: small interfering ribonucleic acid

SMUG1: single-strand-selective monofunctional uracil DNA glycosylase

T: Thymine

TDG: Thymine DNA glycosylase

UDG: Uracil DNA glycosylase

UV: Ultraviolet

WT: Wild-type

3meA: 3-methyladenine

53BP1: p53 binding protein 1

7meG: 7-methylguanine

ABSTRACT

Alkyladenine DNA glycosylase (AAG) initiates the base excision repair of a diverse range of alkylated and deaminated DNA lesions in human cells. These sites of damage are removed from the genome through hydrolytic cleavage of the N-glycosidic bond to produce a shared abasic site repair intermediate. Lesions recognized by AAG include 1,N^6-ethenoadenine, hypoxanthine, and the charged lesion 3-methyladenine. To recognize such a diverse set of lesions from among a vast excess of undamaged bases, AAG must be highly efficient in searching for targets while maintaining strong selectivity to prevent excision of undamaged bases. AAG binds nonspecifically to DNA and rapidly scans the genome in search of targets, but it remains unclear how AAG identifies compatible sites of damage for specific binding. Once a lesion is found, AAG utilizes a nucleotide flipping mechanism to rotate the lesion out of the duplex and into the active site for excision. This provides another selectivity filter for discrimination between damaged and undamaged nucleobases. To gain a better understanding of the substrates recognized by AAG, as well as the active site features important for specificity, we characterized the kinetics for the excision of the controversial target lesion 1,N^2-ethenoguanine (εG). We found that εG is recognized and excised much less efficiently than the primary substrates of AAG. We also identified a key active site residue (N169) involved in discrimination against εG as a substrate. Collectively, these results suggest that poor recognition by AAG would prevent repair of this lesion in cells, which instead likely rely on the expanded repertoire of one of the AlkB homologs. Glycosylases such as AAG have long been speculated to use DNA bending to identify sites of damage during searching, as bending is observed in crystal structures of glycosylases bound to both lesions and undamaged DNA. We developed a stopped flow fluorescence resonance energy transfer (FRET) assay to measure the timing of DNA bending by AAG to place this step within the established AAG mechanism. We found that DNA bending and base flipping occur on the same timescale, after the lesion is already bound by AAG. This

precludes DNA bending from contributing to searching by AAG. However, structural differences between glycosylases could lead to distinct strategies for DNA searching and lesion recognition. This and previous work point to the role of DNA intercalation by a conserved β-hairpin of AAG, which could serve to directly sense altered base pairing, a common feature of the best AAG substrates. Finally, we sought to explore the activity of AAG in the HAP1 cell line, an increasingly popular near-haploid cell line with great potential for use in the study of DNA repair. We determined that AAG does not contribute to the sensitivity of HAP1 cells to the alkylating agent methyl methanesulfonate. This lack of a phenotype contrasts with previously characterized cell lines, suggesting alkylation repair or tolerance is specifically perturbed in HAP1 cells. More work is needed to understand why HAP1 does not benefit from the repair activity of AAG, however, this work makes it clear that other cell models are better suited to the study of AAG function. Together, these results provide a deeper understanding of the mechanisms of searching and specificity used by AAG to recognize DNA damage, and the scope of tools available to study this process.

CHAPTER 1

Introduction to DNA Damage Recognition by AAG

DNA damage and repair

DNA is the hereditary material for all known cellular life. As such, maintenance and replication of the genome are the foremost priorities for all forms of life from single-celled bacteria to complex multicellular organisms including humans. DNA damage can arise from many natural sources, such as from exposure to energetic radiation like UV light or from reactions with dangerous exogenous compounds [1]. In addition, DNA damage can arise from within cells, through exposure of the genome to dangerous metabolic intermediates or through errors in replication and repair [2]. The consequences of these damage sites vary widely but can include cytotoxicity and mutation, the latter of which can lead to genome instability and cancer [1, 3].

A vast array of overlapping DNA damage repair pathways exists to provide a robust and efficient response to any type of DNA damage. The nucleotide excision repair (NER) pathway removes segments of DNA containing lesions that distort base-pairing in the duplex, such as bulky adducts and intra-strand crosslinks [4, 5]. Base excision repair (BER) is primarily responsible for cleavage of small, single-nucleotide sites of DNA damage arising through oxidation alkylation, deamination, and misincorporation [6]. The direct reversal repair pathway (DRR) targets similarly small alkyl modifications and photoproducts, but chemically reverses those sites without removing the damaged base [7]. Common base substitution mismatches and insertion/deletion mismatches during replication are resolved through the mismatch repair pathway (MMR) [8].

There are also numerous pathways for resolving breaks and linkages in DNA. Homologous recombination (HR) and non-homologous end-joining (NHEJ) are the primary pathways that mediate repair of double-strand breaks under different

circumstances [9]. There are additional pathways that contribute to double strand break repair including alternative end-joining and single-strand annealing. Interstrand cross-link (ICL) repair encompasses multiple pathways involved in the removal of covalent linkages between the opposing strands of DNA. ICL repair can occur independent of DNA replication, but has been most thoroughly characterized during replication where the stalled replication fork can be resolved by unhooking the ICL linkage and utilizing a combination of translesion synthesis, NER, BER, and HR factors to repair both copies of DNA [10-12].

Every DNA repair pathway must grapple with the fundamental challenge of how to locate and recognize sites of damage for efficient and accurate repair. Single-nucleotide sites of damage such as those repaired by BER and direct repair are thought to be the most common form of DNA damage. Estimates place the occurrence of these small lesions at roughly 10,000 sites per day per cell in humans [1]. While these lesions are numerous when viewing an organism as a whole, each type of damage is incredibly rare when compared to the billions of bases in the human genome. This presents a pair of challenges to DNA repair pathways. First, the site of damage needs to be physically located from among a vast excess of undamaged sites. Second, the repair enzymes need to have sufficient specificity to distinguish a genuine site of damage from an undamaged base. This thesis seeks to understand the strategies used for DNA damage recognition in the human base excision repair pathway.

Base excision repair

The BER pathway is conserved across every domain of life. Through BER, a wide range of small lesions are cleaved from DNA and replaced using the opposite strand as a template (Fig. 1-1) [6]. Glycosylases initiate the base excision repair pathway by binding a lesion and cleaving the damaged nucleobase at the *N*-glycosidic bond [13]. The simplest monofunctional glycosylases use hydrolysis of the *N*-glycosidic bond to release the damaged nucleobase from the duplex. This produces as an abasic apurinic/apyrimidinic (AP) site, which can be recognized by an AP-endonuclease. Abasic

sites are a dangerous form of DNA damage that can be both mutagenic and cytotoxic if left unresolved, so progress through the pathway is critical [14, 15].

The primary AP endonuclease responsible for BER in humans is called APE1. Once an AP endonuclease like APE1 identifies an abasic site, the phosphodiester backbone is cleaved 5' to the former lesion site, producing a 3' OH group and a 5' deoxyribosephosphate (dRP) group on the ends of DNA. This single-strand break in DNA is quickly detected for further processing, to minimize the potential for further breaks in DNA. In humans, the dRP group is removed by the dRP lyase activity of polymerase β, while prokaryotes rely on other dRPase proteins [16]. The polymerase is then able to extend from the 3' OH to incorporate a properly paired nucleotide into the DNA [17]. The remaining nick is sealed by a DNA ligase. In mammals, Ligase 1 and Ligase 3 are both thought to be essential for BER [18, 19].

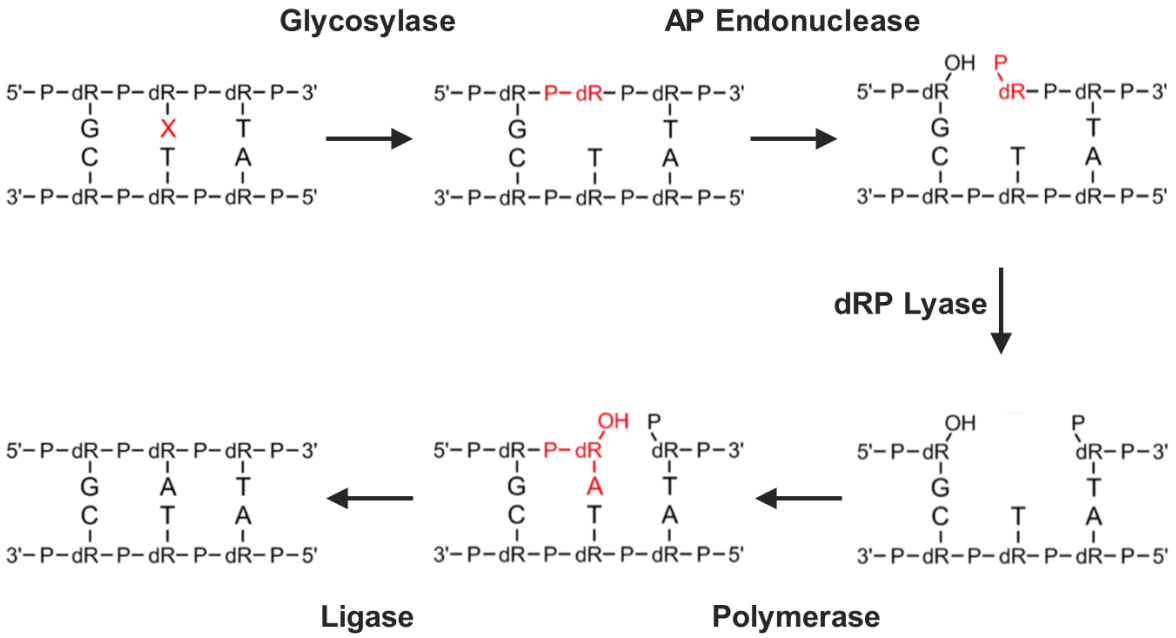


Fig. 1-1. Short-patch base excision repair. A schematic is shown for a generalized short-patch base excision repair pathway with a monofunctional DNA glycosylase. The glycosylase hydrolyzes the *N*-glycosidic bond of the damage base (red X) to produce an abasic site. An AP endonuclease nicks the backbone 5' of the abasic site, producing a hanging dRP group. In humans, polymerase β removes the dRP group before extending the 3' end of the nick to match the opposing strand. Finally, a ligase seals the nick.

Several notable variations of the BER pathway exist, starting with the glycosylase step. While monofunctional glycosylases use an activated water to hydrolyze the *N*-glycosidic bond, bifunctional glycosylase enzymes attack the lesion with an amine group, forming an enzyme-DNA Schiff base intermediate [20]. After elimination, this produces a nicked product rather than an abasic site, which must be processed by alternative factors. Variations in the pathway also occur after the glycosylase step. In certain circumstances in the cell, the polymerase can instigate “long-patch” BER by extending the 5’ DNA fragment past the lesion site, producing a long flap ending with the dRP. This product must then be resolved by a flap endonuclease for ligation to occur [21].

Alkyladenine DNA glycosylase

In human cells there are a total of 11 known glycosylase genes. These include 5 monofunctional and 6 bifunctional glycosylases and span 4 distinct structural families: helix-hairpin-helix, uracil DNA glycosylase, helix-2-turn-helix, and alkyladenine DNA glycosylase. These families are thought to have evolved glycosylase activity independently but have converged on similar mechanistic strategies [22]. Some glycosylases are highly specialized for the recognition of a small selection of lesions, such as the highly specific recognition of uracil by UDG [23, 24]. Other glycosylases like AAG recognize a broad spectrum of diverse lesions [25]. The damage specificity of each glycosylase varies, but many glycosylases have redundant activity to provide cells with robust coverage for important types of damage.

The monofunctional mammalian alkyladenine DNA glycosylase (AAG), also known as methylpurine DNA glycosylase (MPG), is one of the better characterized examples of a DNA repair glycosylase with a broad substrate range [25-28]. AAG is the sole human protein in the eponymous alkyladenine DNA glycosylase structural family. While AAG homologs are widespread in eukaryotes, they are not found in fungi and very few AAG homologs have been discovered in prokaryotes, with the most notable exception being the bacteria *Bacillus subtilis* [29, 30]. In human cells, AAG is the sole glycosylase responsible for the repair of alkylation damage. AAG recognizes a wide array of alkylated bases including 3-methyladenine, 7-methylguanine, and 1,N^6-ethenoadenine, as well as

the deamination product hypoxanthine (Fig. 1-2) [25]. Smaller pyrimidine lesions such as 3,N^4-ethenocytosine are also recognized by AAG, though they are not excised [31, 32].

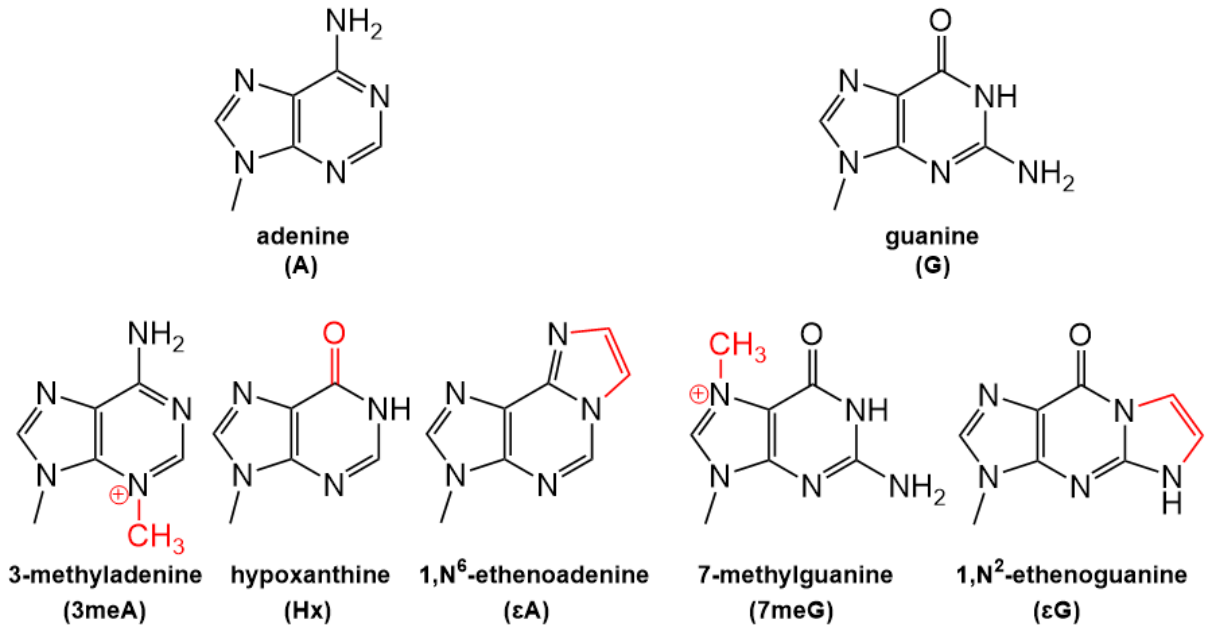


Fig. 1-2. Lesions recognized by human AAG. The structures of the undamaged purine bases adenine and thymine are shown. Below each base are selected derivative lesions excised by human AAG. These lesions have diverse structures and variable charges, with adducts or modifications shown in red.

The role of AAG in cellular repair of alkylation damage has been investigated using mouse models. AAG knockout mice are viable and show no clear phenotype under normal conditions, but they are sensitized to alkylating agents like methyl methanesulfonate and show slower repair of alkylation damage [33, 34]. Although AAG is expected to be protective, overexpression of AAG has also been shown to sensitize mammalian cells to alkylation damage under a variety of circumstances [35-37]. This is thought to be the result of the accumulation of dangerous repair intermediates following AAG-catalyzed cleavage. Thus, proper regulation of AAG is critical for mediating the mammalian response to alkylation damage.

AAG is a monomeric protein composed of 298 amino acids. The C-terminal domain (~220 amino acids) of AAG is highly conserved and contains the enzyme's DNA binding interface and active site. The poorly conserved N-terminal domain of AAG is a flexible

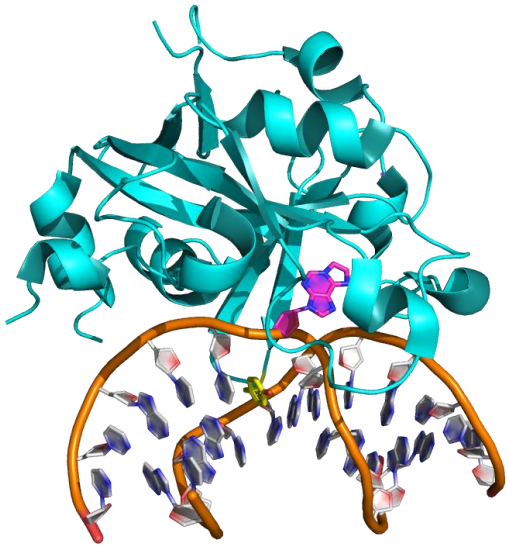


Fig. 1-3. Crystal structure of human AAG bound to εA-containing DNA. AAG bound specifically at a site of εA damage (magenta). The lesion is flipped into the AAG active site and the backbone of DNA is bent by roughly 20° at the site of binding. AAG residue Tyr162 (yellow) from the β3β4 hairpin is intercalated into the DNA duplex in the gap left by the flipped-out base. Image produced using PyMOL (http://www.pymol.org) with coordinates from PDB structure 1F4R.

region that does not strongly impact the substrate selectivity or the rate of base excision by the enzyme but may impact nonspecific DNA binding [27, 38]. Multiple crystal structures of AAG have been solved, including a structure bound specifically at a lesion site (Fig. 1-3) [26, 39]. These structures identify a large positively charged binding surface for nonspecific interaction with the negatively charged DNA backbone. AAG is also shown to utilize base flipping to engage an extrahelical εA lesion.

The minimal kinetic mechanism of AAG has been characterized using the intrinsic fluorescence of the lesion 1,N^6-ethenoadenine (εA) and relevant AAG variants (Fig. 1-4) [28, 38]. This overall strategy for lesion recognition appears to be shared among a number of monofunctional glycosylases including UDG and AlkA [40].

AAG first binds nonspecifically to DNA. Under normal cellular conditions, sites of damage will be vastly outnumbered by undamaged sites, and therefore initial binding will typically occur at an undamaged nucleotide. The enzyme then quickly searches for a site of damage. Early during the searching process, AAG intercalates the β3β4 hairpin into the minor groove of DNA [41]. Once a lesion has been identified, AAG rapidly and

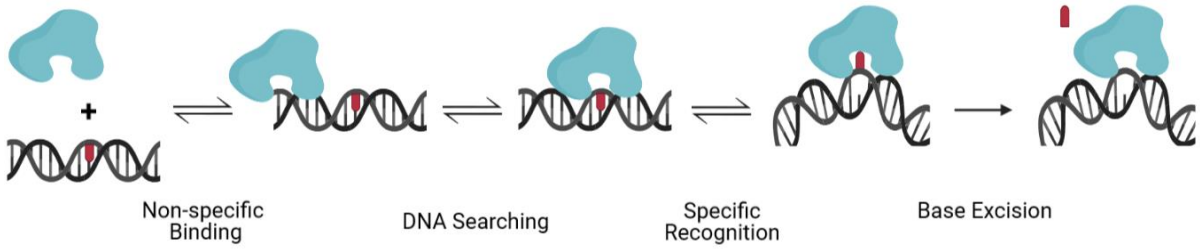


Fig. 1-4. Minimal mechanism for recognition of DNA damage by AAG. AAG binds nonspecifically to any available site on DNA. Searching is initiated quickly, along with early intercalation of a hairpin into the minor groove of DNA. Once a lesion (red) is found, AAG forms an initial recognition complex. Finally, the lesion is flipped into the AAG active site, DNA is bent, and the Tyr162 residue intercalates into the gap in DNA to form the specific recognition complex. This specific complex is then prepared for hydrolysis of the *N*-glycosidic bond. Illustrated with BioRender.com.

reversibly forms an initial recognition complex with the lesion partially unstacked from the duplex. From this initial recognition complex, the lesion flips entirely into the enzyme active site, forming the specific recognition complex for cleavage. While the lesion is flipped into the active site, AAG intercalates a Tyr side chain into the gap left in DNA. *N*-glycosidic bond cleavage is catalyzed via an activated water molecule bound in the active site.

Facilitated diffusion by AAG

Searching the genome solely through diffusion is an incredibly inefficient way to interrogate the vast number of DNA sites in a cell. To improve the efficiency of DNA searching, many DNA-binding proteins employ a series of strategies collectively known as facilitated diffusion [42-46]. Facilitated diffusion describes several approaches utilizing nonspecific electrostatic contacts between the protein and DNA to transfer between proximal DNA sites without fully dissociating into solution. These mechanisms include sliding, hopping, and intersegmental transfer (Fig. 1-5). Sliding refers to 2-dimensional movement of the protein along DNA while maintaining constant contact with the deoxyribose-phosphate backbone. Hopping describes short dissociation of the protein

from DNA, but then rapid reassociation at a nearby site. Finally, intersegmental transfer refers to dissociation and reassociation of the protein at DNA sites that are distant in 2D sequence but close in 3D space.

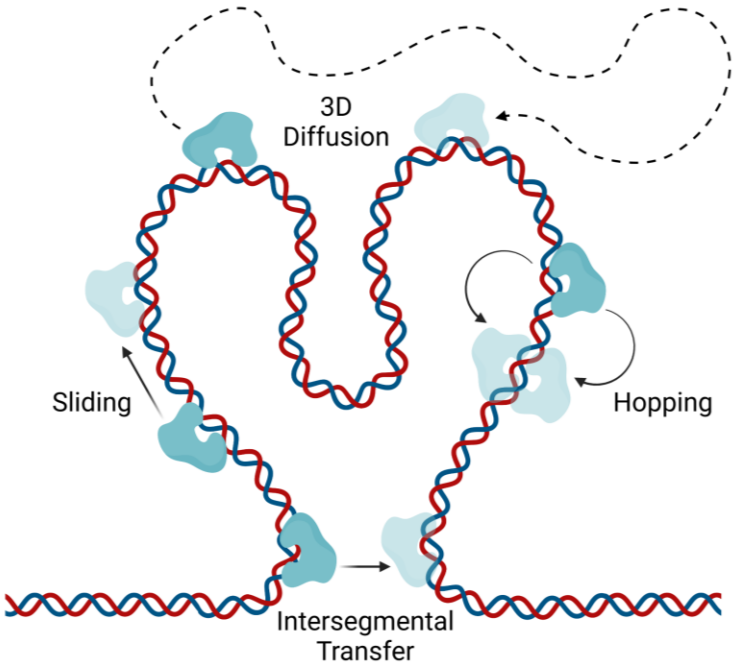


Fig. 1-5. Mechanisms of facilitated diffusion. Facilitated diffusion encompasses multiple mechanisms for DNA searching that enable searching of many DNA sites after binding an initial site. Shown are three distinct mechanisms of facilitated diffusion: sliding, hopping, and intersegmental transfer. These strategies allow more efficient DNA searching than simple 3-dimensional diffusion. Illustrated with BioRender.com.

AAG utilizes several modes of facilitated diffusion to conduct an efficient search of the genome [38, 47, 48]. After binding nonspecifically to DNA, AAG redundantly searches nearby sites for damage. This short-range search includes hopping between opposing DNA strands as well as hopping over obstructions on the DNA, an important feature for searching the heavily occupied human genome [49]. It remains unclear if AAG is capable of sliding or if short-range searching is exclusively performed by hopping. These short movements allow AAG to closely inspect a given DNA region for the presence of damage before moving on to a new stretch of DNA via intersegmental transfer or diffusion once the first region has been sufficiently scanned. A searching mechanism prominently

featuring hopping between DNA sites agrees with observations of uracil DNA glycosylase, suggesting glycosylases may share searching strategies despite their structural differences [50, 51].

While searching by AAG has been extensively explored with short oligonucleotides, less is known about searching in the context of a packed cell nucleus. Knockout of the primary yeast alkylation glycosylase, MAG1, and supplementation with AAG provides a simple model to examine facilitated diffusion in eukaryotic cells [52]. In this study it was found that facilitated diffusion was important for cellular repair when AAG is present at physiological levels, whereas facilitated diffusion is dispensable when AAG is overexpressed to a high level. This suggests that chromatin remodeling is not the rate limiting step for accessing sites of DNA damage, and furthermore that facilitated diffusion allows for efficient DNA repair with a lower steady-state level of repair glycosylase. It remains an open question as to how chromatin is searched in the larger and more complex human nucleus, and it could be informative to employ separation-of-function variants of AAG that are deficient in facilitated diffusion.

Base flipping and side chain intercalation by AAG

To excise lesions from DNA, glycosylases must first access the *N*-glycosidic bond from within the DNA duplex. The human glycosylases in all 4 structural families accomplish this through base flipping [26, 53-55]. Upon binding a compatible lesion, AAG rotates the damage base roughly 180° out of the duplex and into the active site. This disrupts the base stacking of the lesion as well as any Watson-Crick hydrogen bonding with the opposing strand. This step also provides another layer of substrate selectivity by AAG. The ease of flipping a base out of the DNA duplex can contribute to the overall favorability of base excision.

In crystal structures of AAG bound to a flipped-out base, the side chain of Tyr162 is intercalated into the DNA at the gap left by the flipped-out base. Two primary models exist to explain the mechanism of base flipping and side chain intercalation: pushing and plugging. In a pushing model, the protein side chain is intercalated to destabilize the DNA duplex, promoting flipping of the targeted base. In contrast, in a plugging model the

protein side chain moves into position after the nucleotide is flipped out to prevent it from unflipping. These models are not mutually exclusive, and there is experimental evidence that Tyr162 functions in both roles. Mutation of AAG intercalating residue Tyr162 to either alanine or phenylalanine accelerates the rate of εA flipping to different extents, indicating this residue does participate in the flipping step [56]. It was proposed that Tyr162 is playing a role as a gatekeeper in slowing the rate of flipping relative to the smaller alanine and phenylalanine side chains. The plugging role of Tyr162 is clearly established, as the rate of unflipping is slowest for the native Tyr162 and similar Phe162 and becomes greatly accelerated in the Ala162 variant.

Crystal structures of AAG bound to a flipped-out lesion reveal bending of DNA at the site of AAG binding, a feature shared among many glycosylases [26, 54, 57-60]. The role of this DNA bending is not well understood, but several potential models have been proposed for different glycosylases. The most popular model suggests that DNA is bent during the searching process to identify sites of duplex instability due to base modification [59, 61]. The bent site would then be primed for flipping and specific recognition. However, bending could also play a role in the relief of strain generated by rotation of the lesion into the active site, or in the precise positioning of the lesion in the active site for hydrolysis.

Active site discrimination by AAG

Discriminating between substrates and non-substrates is a significant challenge for AAG due to the broad range of lesions it repairs. Highly-specific glycosylases like UDG have an active site tailored to interact with one preferred substrate [24]. By contrast, the active site of AAG must accommodate substrates with varying shapes and charges, all while excluding undamaged bases (Fig. 1-2). The active site of AAG utilizes a number of structural features to achieve this balance of broad recognition and discrimination.

For purine substrates, active site residues His136 and Asn169 play a large role in governing the substrate selectivity of base flipping (Fig. 1-6A & B) [25]. The aromatic side chain of His136 stacks neatly against the face of the flipped-out lesion. However, the primary contribution of His136 to specificity comes not from the side chain, but from the amide backbone. The amide group of H136 provides hydrogen bond stabilization for

bases with a 6-oxo group, like Hx, as well as the N6 nitrogen of εA. The backbone amide is predicted to also clash with the exocyclic amine group of undamaged A, enabling discrimination against that undamaged base. Residue Asn169 closely contacts the Watson-Crick bonding edge of the flipped-out lesion, making it ideal for constraining lesion structure. The side chain amide of Asn169 extends close enough to clash with bases containing a bulky N2 group, such as the exocyclic amine of undamaged G. Indeed, mutation of the Asn169 residue to Ala or Ser leads to significant rate enhancement for excision of all nucleobases, including G [25, 62]. Residues Tyr127 and Tyr159 complete the structural definition of the active site, providing stacking interactions on either side of purine lesions.

The final layer of substrate selectivity of AAG comes from the *N*-glycosidic bond hydrolysis step. AAG relies on general acid-base catalysis for base excision [27]. Residue Glu125 serves as the general base that activates a bound water for nucleophilic attack (Fig. 1-6B). Due to the shared *N*-glycosidic bond structure of all relevant lesions, Glu125 is positioned to excise any lesion successfully flipped into the active site. A general acid is used to protonate uncharged nucleobases, possibly at the N7 site, to stabilize the

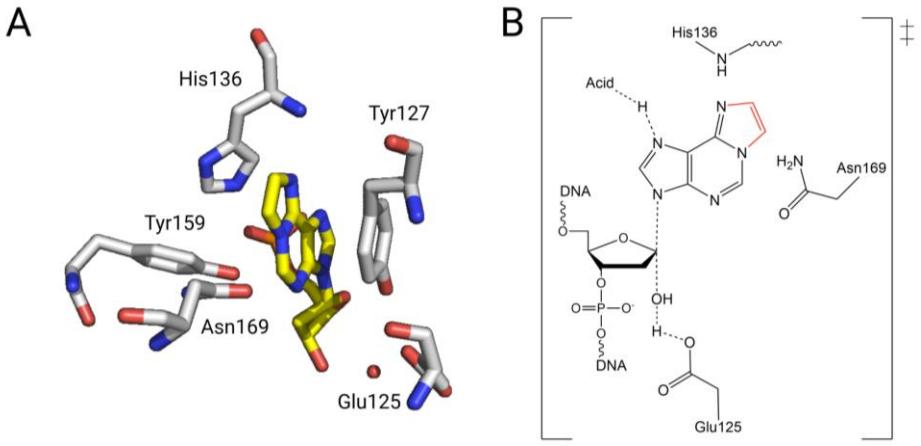


Fig. 1-6. Active site structure and recognition of DNA damage. (A) Diagram of an εA lesion (yellow) surrounded by the active site pocket residues of human AAG. The activated water molecule (red sphere) is bound near the catalytic Glu125 and *N*-glycosidic bond. Structure rendered in PyMOL (http://www.pymol.com) using the coordinates from 1F4R in the PDB. (B) Structure of the transition state for AAG cleavage of εA. Glu125 serves as the general base, while the general acid remains unknown. Residues His136 and Asn169 are positioned for substrate selectivity.

leaving group for base excision. The identity of this acid remains unknown, and mutagenic studies have suggested the possibility that there is more than one pathway to protonation, or an as-yet-undetected conformational change to bring the general acid into position [27]. General acid catalysis appears to play a defining role in preventing excision of small pyrimidine lesions like ethenocytosine (εC) that would otherwise fit into an active site designed for purines. AAG binds εC with relatively strong affinity, but the lack of a properly positioned N7 site for protonation prevents base excision from progressing [32].

Work described in this thesis

In this thesis we explore the strategies used by human AAG to search the genome for sites of DNA damage and to discriminate between potential DNA substrates. We also examine cellular models to study the interrelationship between AAG-initiated repair of alkylation damage and other DNA repair pathways.

In Chapter 2 we examine the excision of controversial AAG substrate 1,N^2-ethenoguanine (εG). Using single turnover and multiple turnover kinetic assays, we compare the efficiency of εG excision to the excision of the well-characterized substrate εA, which arises under similar circumstances in the human cell. εG was excised more slowly than εA in every tested condition, and direct competition between εG and undamaged DNA highlighted the poor recognition of εG. Lastly, mutation of AAG active site residue Asn169 was found to strongly impact the εG excision activity of AAG. This suggests that the lack of AAG activity toward εG is likely tied to active site discrimination against the N2 group of undamaged G.

In Chapter 3, we investigate the role of DNA bending in the recognition of DNA damage by human AAG. DNA bending is widely speculated to contribute to lesion searching by AAG and other glycosylases. To establish the role of DNA bending we developed a FRET assay to measure time-resolved changes in DNA bending to compare with the established kinetics of base flipping. The kinetics of DNA bending were found to consistently align with base flipping. Analysis of AAG active site variants and DNA with flexible linkers showed that alteration of the rates of both flipping and bending failed to

decouple bending and flipping. This places DNA bending after the initial binding of the lesion in the AAG mechanism, ruling out DNA bending as a searching mechanism.

In Chapter 4, we scrutinize the use of the popular human near-haploid cell line HAP1 as a model for DNA repair in ordinary diploid human cells. Knockout of AAG in mammalian cells, including the human HeLa cell line, has been shown to result in increased sensitivity to DNA alkylating agents such as methyl methanesulfonate (MMS). Unexpectedly, we found that HAP1 cells lacking AAG show no change in sensitivity to MMS treatment. Using cell extracts, it was shown that AAG and downstream BER enzymes are active. Thus, alkylation repair and tolerance mechanisms in HAP1 cells must deviate significantly from other characterized human cells.

Our results provide a better understanding of both the mechanisms of DNA damage recognition and the tools available to interrogate human DNA repair.

REFERENCES

1. Lindahl, T., *Instability and decay of the primary structure of DNA.* Nature, 1993. **362**(6422): p. 709-15.

2. Marnett, L.J., J.N. Riggins, and J.D. West, *Endogenous generation of reactive oxidants and electrophiles and their reactions with DNA and protein.* J Clin Invest, 2003. **111**(5): p. 583-93.

3. Hoeijmakers, J.H., *Genome maintenance mechanisms for preventing cancer.* Nature, 2001. **411**(6835): p. 366-74.

4. Novarina, D., et al., *Mind the gap: keeping UV lesions in check.* DNA Repair (Amst), 2011. **10**(7): p. 751-9.

5. Marteijn, J.A., et al., *Understanding nucleotide excision repair and its roles in cancer and ageing.* Nat Rev Mol Cell Biol, 2014. **15**(7): p. 465-81.

6. Robertson, A.B., et al., *DNA repair in mammalian cells: Base excision repair: the long and short of it.* Cell Mol Life Sci, 2009. **66**(6): p. 981-93.

7. Yi, C. and C. He, *DNA repair by reversal of DNA damage.* Cold Spring Harb Perspect Biol, 2013. **5**(1): p. a012575.

8. Kunkel, T.A. and D.A. Erie, *DNA mismatch repair.* Annu Rev Biochem, 2005. **74**: p. 681-710.

9. Scully, R., et al., *DNA double-strand break repair-pathway choice in somatic mammalian cells.* Nature Reviews Molecular Cell Biology, 2019.

10. Semlow, D.R. and J.C. Walter, *Mechanisms of Vertebrate DNA Interstrand Cross-Link Repair.* Annu Rev Biochem, 2021. **90**: p. 107-135.

11. Legerski, R.J., *Repair of DNA interstrand cross-links during S phase of the mammalian cell cycle.* Environ Mol Mutagen, 2010. **51**(6): p. 540-51.

12. Noll, D.M., T.M. Mason, and P.S. Miller, *Formation and repair of interstrand cross-links in DNA.* Chem Rev, 2006. **106**(2): p. 277-301.

13. Krokan, H.E., R. Standal, and G. Slupphaug, *DNA glycosylases in the base excision repair of DNA.* Biochem J, 1997. **325 (Pt 1)**: p. 1-16.

14. Hosfield, D.J., et al., *DNA damage recognition and repair pathway coordination revealed by the structural biochemistry of DNA repair enzymes.* Prog Nucleic Acid Res Mol Biol, 2001. **68**: p. 315-47.

15. Lindahl, T. and S. Ljungquist, *Apurinic and apyrimidinic sites in DNA.* Basic Life Sci, 1975. **5A**: p. 31-8.

16. Matsumoto, Y. and K. Kim, *Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair.* Science, 1995. **269**(5224): p. 699-702.

17. Sobol, R.W., et al., *Requirement of mammalian DNA polymerase-beta in base-excision repair.* Nature, 1996. **379**(6561): p. 183-6.

18. Simsek, D., et al., *Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair.* Nature, 2011. **471**(7337): p. 245-8.

19. Gao, Y., et al., *DNA ligase III is critical for mtDNA integrity but not Xrcc1-mediated nuclear DNA repair.* Nature, 2011. **471**(7337): p. 240-4.

20. Krokan, H.E. and M. Bjoras, *Base excision repair.* Cold Spring Harb Perspect Biol, 2013. **5**(4): p. a012583.

21. Klungland, A. and T. Lindahl, *Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1).* EMBO J, 1997. **16**(11): p. 3341-8.

22. O'Brien, P.J., *Catalytic promiscuity and the divergent evolution of DNA repair enzymes.* Chem Rev, 2006. **106**(2): p. 720-52.

23. Slupphaug, G., et al., *A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA.* Nature, 1996. **384**(6604): p. 87-92.

24. Stivers, J.T. and A.C. Drohat, *Uracil DNA glycosylase: insights from a master catalyst.* Arch Biochem Biophys, 2001. **396**(1): p. 1-9.

25. 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.

26. 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.

27. 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. 12418-29.

28. 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.

29. Aamodt, R.M., et al., *The Bacillus subtilis counterpart of the mammalian 3-methyladenine DNA glycosylase has hypoxanthine and 1,N6-ethenoadenine as preferred substrates.* J Biol Chem, 2004. **279**(14): p. 13601-6.

30. Morohoshi, F., K. Hayashi, and N. Munkata, *Bacillus subtilis alkA gene encoding inducible 3-methyladenine DNA glycosylase is adjacent to the ada operon.* J Bacteriol, 1993. **175**(18): p. 6010-7.

31. Fu, D. and L.D. Samson, *Direct repair of 3,N(4)-ethenocytosine by the human ALKBH2 dioxygenase is blocked by the AAG/MPG glycosylase.* DNA Repair (Amst), 2012. **11**(1): p. 46-52.

32. Lingaraju, G.M., et al., *Structural basis for the inhibition of human alkyladenine DNA glycosylase (AAG) by 3,N4-ethenocytosine-containing DNA.* J Biol Chem, 2011. **286**(15): p. 13205-13.

33. Elder, R.H., et al., *Alkylpurine-DNA-N-glycosylase knockout mice show increased susceptibility to induction of mutations by methyl methanesulfonate.* Mol Cell Biol, 1998. **18**(10): p. 5828-37.

34. Engelward, B.P., et al., *Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase.* Proc Natl Acad Sci U S A, 1997. **94**(24): p. 13087-92.

35. Rinne, M.L., et al., *N-methylpurine DNA glycosylase overexpression increases alkylation sensitivity by rapidly removing non-toxic 7-methylguanine adducts.* Nucleic Acids Res, 2005. **33**(9): p. 2859-67.

36. Rinne, M., D. Caldwell, and M.R. Kelley, *Transient adenoviral N-methylpurine DNA glycosylase overexpression imparts chemotherapeutic sensitivity to human breast cancer cells.* Mol Cancer Ther, 2004. **3**(8): p. 955-67.

37. Coquerelle, T., J. Dosch, and B. Kaina, *Overexpression of N-methylpurine-DNA glycosylase in Chinese hamster ovary cells renders them more sensitive to the production of chromosomal aberrations by methylating agents--a case of imbalanced DNA repair.* Mutat Res, 1995. **336**(1): p. 9-17.

38. Hedglin, M. and P.J. O'Brien, *Human alkyladenine DNA glycosylase employs a processive search for DNA damage.* Biochemistry, 2008. **47**(44): p. 11434-45.

39. 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.

40. Taylor, E.L. and P.J. O'Brien, *Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by AlkA.* Biochemistry, 2015. **54**(3): p. 898-908.

41. Hendershot, J.M. and P.J. O'Brien, *Search for DNA damage by human alkyladenine DNA glycosylase involves early intercalation by an aromatic residue.* J Biol Chem, 2017. **292**(39): p. 16070-16080.

42. Berg, O.G., R.B. Winter, and P.H. von Hippel, *Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory.* Biochemistry, 1981. **20**(24): p. 6929-48.

43. Winter, R.B., O.G. Berg, and P.H. von Hippel, *Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The Escherichia coli lac repressor--operator interaction: kinetic measurements and conclusions.* Biochemistry, 1981. **20**(24): p. 6961-77.

44. Winter, R.B. and P.H. von Hippel, *Diffusion-driven mechanisms of protein translocation on nucleic acids. 2. The Escherichia coli repressor--operator interaction: equilibrium measurements.* Biochemistry, 1981. **20**(24): p. 6948-60.

45. Halford, S.E. and J.F. Marko, *How do site-specific DNA-binding proteins find their targets?* Nucleic Acids Res, 2004. **32**(10): p. 3040-52.

46. Stanford, N.P., et al., *One- and three-dimensional pathways for proteins to reach specific DNA sites.* EMBO J, 2000. **19**(23): p. 6546-57.

47. Hedglin, M., Y. Zhang, and P.J. O'Brien, *Isolating contributions from intersegmental transfer to DNA searching by alkyladenine DNA glycosylase.* J Biol Chem, 2013. **288**(34): p. 24550-9.

48. Hedglin, M., Y. Zhang, and P.J. O'Brien, *Probing the DNA structural requirements for facilitated diffusion.* Biochemistry, 2015. **54**(2): p. 557-66.

49. Hedglin, M. and P.J. O'Brien, *Hopping enables a DNA repair glycosylase to search both strands and bypass a bound protein.* ACS Chem Biol, 2010. **5**(4): p. 427-36.

50. Esadze, A., et al., *Measurement of nanoscale DNA translocation by uracil DNA glycosylase in human cells.* Nucleic Acids Res, 2017. **45**(21): p. 12413-12424.

51. Porecha, R.H. and J.T. Stivers, *Uracil DNA glycosylase uses DNA hopping and short-range sliding to trap extrahelical uracils.* Proc Natl Acad Sci U S A, 2008. **105**(31): p. 10791-6.

52. Zhang, Y. and P.J. O'Brien, *Repair of Alkylation Damage in Eukaryotic Chromatin Depends on Searching Ability of Alkyladenine DNA Glycosylase.* ACS Chem Biol, 2015. **10**(11): p. 2606-15.

53. Stivers, J.T., *Extrahelical damaged base recognition by DNA glycosylase enzymes.* Chemistry, 2008. **14**(3): p. 786-93.

54. 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. 859-66.

55. Zhu, C., et al., *Tautomerization-dependent recognition and excision of oxidation damage in base-excision DNA repair.* Proc Natl Acad Sci U S A, 2016. **113**(28): p. 7792-7.

56. Hendershot, J.M. and P.J. O'Brien, *Critical role of DNA intercalation in enzyme-catalyzed nucleotide flipping.* Nucleic Acids Res, 2014. **42**(20): p. 12681-90.

57. Hollis, T., A. Lau, and T. Ellenberger, *Structural studies of human alkyladenine glycosylase and E. coli 3-methyladenine glycosylase.* Mutat Res, 2000. **460**(3-4): p. 201-10.

58. Hollis, T., Y. Ichikawa, and T. Ellenberger, *DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, Escherichia coli AlkA.* EMBO J, 2000. **19**(4): p. 758-66.

59. Parikh, S.S., et al., *Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA.* EMBO J, 1998. **17**(17): p. 5214-26.

60. Maiti, A., et al., *Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition.* Proc Natl Acad Sci U S A, 2008. **105**(26): p. 8890-5.

61. Chen, L., et al., *Direct visualization of a DNA glycosylase searching for damage.* Chem Biol, 2002. **9**(3): p. 345-50.

62. Connor, E.E. and M.D. Wyatt, *Active-site clashes prevent the human 3-methyladenine DNA glycosylase from improperly removing bases.* Chem Biol, 2002. **9**(9): p. 1033-41.

CHAPTER 2

Recognition of 1,N^2-ethenoguanine by AAG Is Restricted by a Conserved Active Site Residue [1]

ABSTRACT

The adenine, cytosine, and guanine bases of DNA are susceptible to alkylation by the aldehyde products of lipid peroxidation and by the metabolic byproducts of vinyl chloride pollutants. The resulting adducts spontaneously cyclize to form harmful etheno lesions. Cells employ a variety of DNA repair pathways to protect themselves from these pro-mutagenic modifications. Human alkyladenine DNA glycosylase (AAG) is thought to initiate base excision repair of both 1,N^6-ethenoadenine (εA) and 1,N^2-ethenoguanine (εG). However, it is not clear how AAG might accommodate εG in an active site that is complementary to εA. This prompted a thorough investigation of AAG-catalyzed excision of εG from several relevant contexts. Using single-turnover and multiple-turnover kinetic analyses, we found that εG in its natural εG•C context is very poorly recognized relative to εA•T. Bulged and mispaired εG contexts, which can form during DNA replication, were similarly poor substrates for AAG. Furthermore, AAG could not recognize an εG site in competition with excess undamaged DNA sites. Guided by previous structural studies, we hypothesized that Asn-169, a conserved residue in the AAG active site pocket, contributes to discrimination against εG. Consistent with this model, the N169S variant of AAG was 7-fold more active for excision of εG as compared to the wild-type (WT) enzyme. Taken together, these findings suggest that εG is not a primary substrate of AAG, and that current models for etheno lesion repair in humans should be revised. We propose that other repair and tolerance mechanisms operate in the case of εG lesions.

---

[1]Reproduced with permission from Thelen, A.T. and O'Brien, P.J. (2020) Recognition of 1,N2-ethenoguanine by alkyladenine DNA glycosylase is restricted by a conserved active-site residue, *J. Biol. Chem. 295(6),* 1685-1693.

INTRODUCTION

All cellular life contends with the challenge of DNA alkylation damage. DNA bases are alkylated by both endogenous and exogenous compounds, and the failure to repair these base lesions results in a variety of deleterious consequences ranging from point mutations to the stalling of DNA replication or transcription [1]. Etheno lesions represent a subset of exocyclic alkylation adducts that can arise through alkylation of purine and pyrimidine bases of genomic DNA (Fig. 2-1) [2]. The etheno lesions 1,N^6-ethenoadenine (εA), N^2,3-ethenocytosine (εC), 1,N^2-ethenoguanine (εG), and N^2,3-ethenoguanine (N^2,3-εG) are naturally formed through exposure to the reactive aldehyde products of lipid peroxidation and subsequent ring closure [3, 4]. Notably, etheno lesions have also been shown to arise due to reactions with chloroacetaldehyde and other metabolic byproducts of the common industrial compound vinyl chloride [5]. *In vitro* DNA replication assays have demonstrated the miscoding properties of etheno lesions as well as the propensity for them to halt replication [6-8]. Consistent with these deleterious effects on DNA replication, etheno lesions have cytotoxic effects on mammalian cells [9].

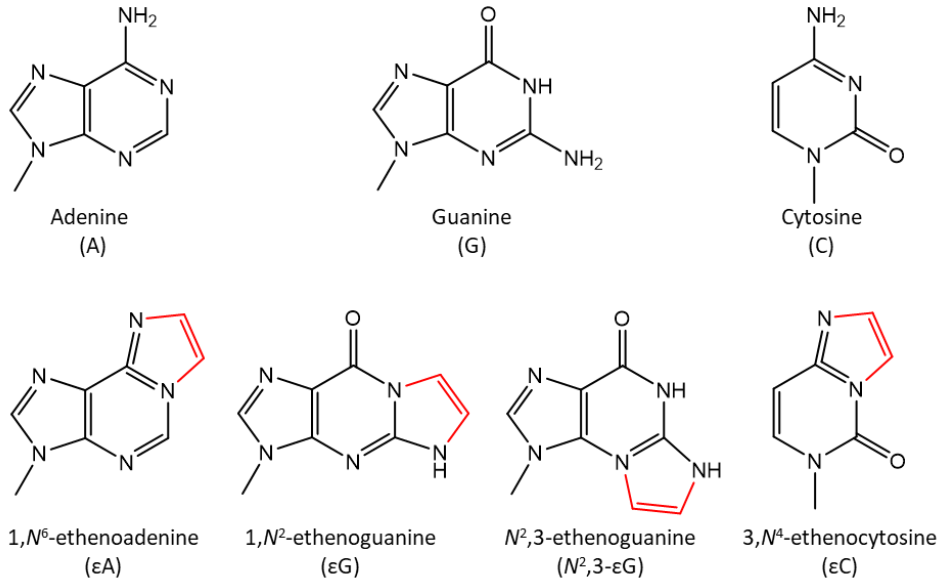


Figure 2-1. Diverse structures of the etheno DNA lesions. The four etheno lesions are depicted below their undamaged forms.

The base excision repair (BER) pathway exists in all domains of life, and it is responsible for removing and replacing diverse single nucleotide lesions such as those which arise through DNA alkylation, oxidation, and deamination [10]. BER is initiated by a DNA glycosylase that searches DNA to locate specific sites of DNA damage and catalyzes the excision of the base lesion to generate an abasic site. Subsequent action by an AP endonuclease, dRP lyase, DNA polymerase and DNA ligase nicks the backbone at the abasic site, removes the deoxy sugar, inserts an undamaged base to complement the opposing strand, and ligates the nick.

Alkyladenine DNA glycosylase (AAG) utilizes a nucleotide flipping mechanism to target a broad range of alkylated DNA lesions [11-13]. Upon specific binding of the enzyme to a target lesion, the lesion is rotated out of the duplex and into the enzyme active site for cleavage. AAG was initially proposed to repair all the etheno adducts which can arise in human cells [14]. εA is a well-characterized substrate of human AAG, and it is recognized with high affinity by the enzyme *in vitro* [11, 15]. AAG also recognizes and binds to sites of εC damage, however it is unable to catalyze cleavage of the *N*-glycosidic bond [16]. Rather, the human glycosylases TDG and SMUG1, along with the direct reversal protein AlkBH2, have been shown to repair εC [17-19].

Compared to εC and εA, less is known about the repair of εG. *In vitro* reactions between DNA and aldehydes derived from lipid peroxidation favor the production of εG over N^2,3-εG, indicating that εG might be the more relevant natural lesion [20-22]. The opposite appears to be true of damage originating from exposure to vinyl chloride and its byproducts [20, 23]. εG causes both replication blocks and a mixture of G → T and G → C transversion mutations in mammalian cells [6, 7]. The lesion also blocks transcription by enzymes such as human RNA polymerase II [24]. The exact frequency of εG lesions in the human genome is not firmly established, however the harmful effects of unrepaired εG combined with the natural origin of the lesion suggest that there must be a means of repairing the lesion.

AAG was previously investigated for glycosylase activity toward εG [14, 25, 26]. While these studies conclude that AAG catalyzes the excision of εG, the kinetics of that excision vary widely between reports [25, 26]. The relevance of the N-terminal domain of AAG for εG recognition is also contentious. The N-terminal 79 residues of human AAG form a

flexible region that is not conserved among different species. One study concluded that truncation of the N-terminus of human AAG eliminated activity toward εG [25], whereas a subsequent study reported similar activity of full-length and N-terminally truncated AAG toward εG [26]. However, this second study reported a very low percentage (6%) of εG could be excised. Studies employing different base lesion substrates concluded that the truncation of the N-terminus reduces the searching efficiency of AAG but does not significantly affect the rate constant for *N*-glycosidic bond cleavage *in vitro* [27, 28].

In this work we characterized the single- and multiple-turnover kinetics of the AAG-catalyzed excision of εG from a variety of DNA duplexes. Comparison of the kinetic constants for excision of εG and εA, and the results of direct competition experiments, together demonstrate that εG is a very poor substrate of AAG. We investigated the structural origins for this substrate specificity and found that mutation of N169 to a smaller side chain allows significantly increased activity toward εG, presumably because the bulky etheno adduct can be better accommodated in the active site. Although we have verified the claims that AAG is capable of catalyzing excision of εG *in vitro*, our work argues against this being a *bona fide* physiological substrate and leads to the prediction that other DNA repair pathways are responsible for εG repair *in vivo*.

RESULTS AND DISCUSSION

Multiple turnover kinetics of AAG-catalyzed excision of εG

To examine the kinetics of εG excision by human AAG, we expressed and purified Δ80 AAG from *E. coli* for *in vitro* glycosylase assays. We used steady-state kinetics to characterize the efficiency of AAG-catalyzed excision of εG from a 25mer DNA duplex containing a central εG•C target site (see Experimental Procedures). Reaction conditions were selected to maximize the activity of AAG while maintaining enzyme stability over a long time-course, and linear initial rates were observed up to 10% product formation with no evidence of a pre-steady state burst (Fig. 2-2A). Despite the optimized conditions for AAG activity, the reaction proceeded slowly with a k_{cat} of 0.0048 min^{-1} (Fig. 2-2B). The

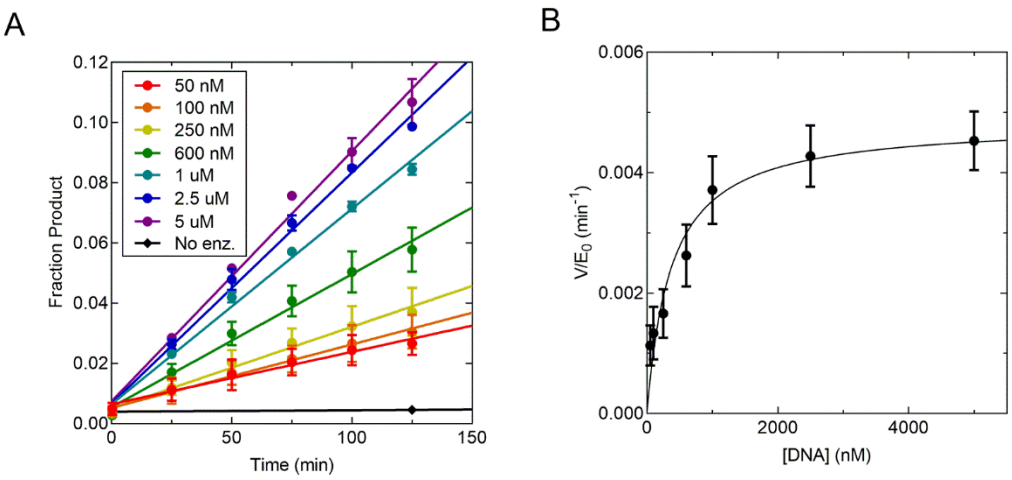


Figure 2-2. Multiple turnover excision of εG by AAG. (A) Time course of the Δ80 AAG-catalyzed excision of εG from a 25mer DNA substrate with an opposing C base. Concentrations of DNA are listed in the legend, with DNA in 50-fold excess of enzyme for each reaction. Error bars represent the S.D. of 2 trials. (B) Dependence of εG excision rate on DNA concentration, fit to a hyperbola with k_{cat} = 0.0048 min^{-1} and K_M = 370 nM^{-1}. Points represent the average of 4 replicates, and error bars represent the S.D.

catalytic efficiency of the reaction, given by k_{cat}/K_M, was measured to be 1.5x10^4 $M^{-1}s^{-1}$. In comparison, the catalytic efficiency of AAG for excision of εA, a well-characterized substrate, has been estimated to be 3 orders of magnitude higher under the same conditions [11]. This discrepancy is conspicuous given the similar origin and the mutagenicity of both lesions.

To more directly compare the efficiencies for the excision of εG and εA, the lesion-containing oligonucleotides were competed in the same glycosylase assay mixture. The εG lesion was paired with a complementary C, whereas the εA lesion was paired with a complementary G. The εA•G pairing is the least efficient base pairing for εA excision, potentially allowing for the relatively slow εG excision to compete [29]. To distinguish the εA and εG DNA substrates and products on a gel, the εG lesion was incorporated into a 25mer DNA sequence while the εA lesion was incorporated in a 19mer DNA sequence. We controlled for the effect of DNA length on the kinetics of glycosylase activity by performing a competition assay between εA incorporated into the 19mer and 25mer sequence. No substantial preference between the 19mer and 25mer was observed (Fig. A1). The formation of the fluorescently labeled 12mer and 9mer products from excision

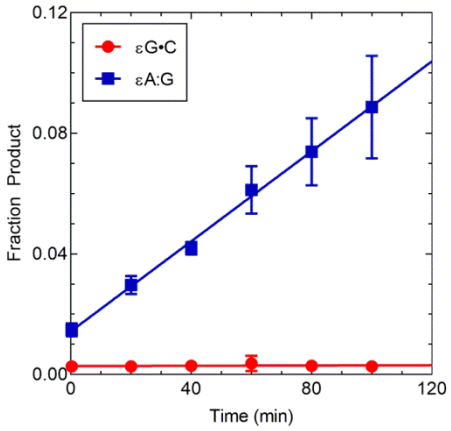


Figure 2-3. Competition of εG and εA for AAG-catalyzed excision. Reactions contained 2 nM Δ80 AAG, 2 μM εG in the 25mer DNA sequence with an opposing C, and 200 nM εA in the 19mer DNA sequence with an opposing G. Lines of best fit were calculated using linear regression. Points represent the average of 4 replicates, and error bars represent the S.D.

of εG and εA, respectively, was monitored to obtain initial rates for the excision of both substrates. At a ratio of 10:1 εG to εA, no detectable εG product was formed during the initial period of the εA reaction (Fig. 2-3). As 0.5% of the εG excision product could have been reliably detected, we conservatively estimate that εA•G is preferred over εG•C by at least 200-fold. These data further support the conclusion that εG is poorly recognized relative to primary substrates of AAG.

AAG-catalyzed single turnover excision of εG

Single turnover glycosylase reactions were performed to measure the rate of εG and εA excision catalyzed by AAG. These reactions report on all the steps preceding and including *N*-glycosidic bond cleavage, but subsequent steps such as product release are excluded. To directly compare to previous results, an optimal pH of 6.1 was used [28], and AAG was up to 3-fold more active for excision of εG at pH 6.1 as compared to pH 7.5 (Fig. A2). Each single turnover reaction of εG excision proceeded to ~85% completion, indicating the presence of a small quantity of nonreactive species in the εG-DNA (Fig. 2-4A & B). As observed in the multiple turnover experiments, AAG-catalyzed excision of εG proceeded more slowly than excision of εA. Furthermore, the concentration dependence of Δ80 AAG exhibited biphasic behavior (Fig. 2-4C; red squares).

In the classic model for single turnover kinetics, the relationship between k_{obs} and the concentration of enzyme should fit to a hyperbola. However, a reduction in the rate constant for excision of εG was observed at elevated concentrations of AAG (Fig. 2-4C). A similar inhibitory effect was previously reported for the *E. coli* 3-methyladenine DNA glycosylase (AlkA), but this is the first instance of such behavior from AAG [30]. Using the inhibition model developed for AlkA, described in Fig. A4, a K_i value of 680 nM was determined for Δ80 AAG. These concentrations of AAG far exceed cellular conditions, making such enzyme crowding unlikely in a cellular context. However, the inability of AAG to distinguish εG from undamaged sites could have serious physiological implications.

To assess the relative affinity of AAG for undamaged and damaged sites, single turnover reactions of AAG with εG and εA were challenged with varying concentrations of undamaged 25mer DNA oligonucleotides containing a central A•T pair. The excision

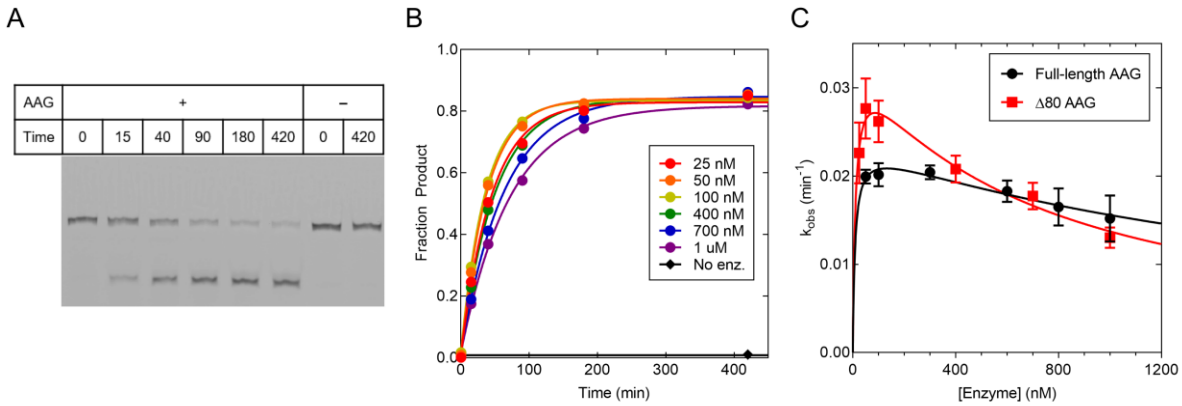


Figure 2-4. Single turnover excision of εG by AAG. (A) Representative gel for the AAG-catalyzed single turnover excision of εG. Reactions contained 20 nM 25mer DNA with εG complemented by a C. Enzyme concentration was varied between reactions. Pictured are reactions with or without 100 nM Δ80 AAG. (B) Time courses for εG excision by Δ80 AAG. Enzyme concentrations are indicated in the legend. (C) The single turnover rate constants for εG excision catalyzed by both full-length and Δ80 AAG are shown plotted by enzyme concentration. Reactions contained 20 nM εG in the 25mer DNA sequence. The values were fit to the multivalent interference model detailed in Fig S4. Points for full-length AAG represent the mean of 4 replicates, while points for 80 AAG represent the mean of 6 replicates. Error bars represent the S.D. For Δ80 AAG the best fit values are k_{max} = 0.034 min^{-1}, $K_{1/2}$ = 11 nM, and K_i = 680 nM. For full-length AAG the values are k_{max} = 0.024 min^{-1}, $K_{1/2}$ = 9 nM, and K_i = 2.0 μM.

of εA was unaffected by the presence of undamaged DNA, up to at least 80-fold excess over the damaged DNA, which was the highest concentration tested (Fig. 2-5). In contrast, the excision of εG was strongly inhibited by the presence of undamaged DNA, with an IC_{50} of 400 nM. This experiment indicates that εG•C is preferred by a factor of 300-fold relative to binding to undamaged sites, which is much less than the lower limit of >30,000 preference for εA•T. The 300-fold preference for εG•C, relative to a typical undamaged DNA site, is unlikely to be sufficient for repair in the cell where there is a vast excess of undamaged relative to damaged sites.

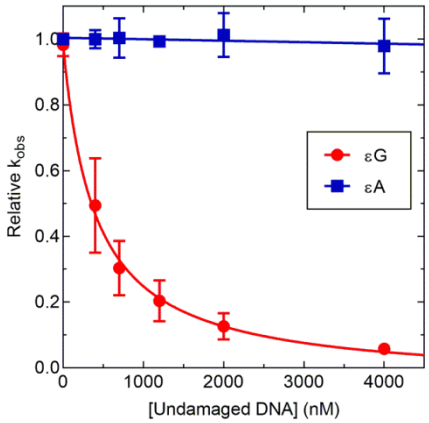


Figure 2-5. Competition of AAG-catalyzed excision of εG and εA by undamaged DNA. Reactions containing 100 nM Δ80 AAG and 50 nM of lesion-containing substrate were incubated with increasing concentrations of undamaged 25mer DNA. The loss of εG excision activity was fit with Equation 5, producing an IC_{50} of 400 nM. Points represent the mean of 4 replicates, and error bars represent the S.D.

The N-terminal region of AAG is not necessary for the excision of εG

Previous studies have demonstrated that the catalytic domain of AAG is more stable than the full-length protein and has similar rates of *N*-glycosidic bond cleavage with many different substrates [27, 28]. Despite this, it has been reported that εG cannot be excised by N-terminal truncations of AAG [25]. Our multiple turnover and single turnover glycosylase assays indicate the contrary, that Δ80 AAG is active toward εG under a variety of conditions. However, differences may still exist between the two protein variants. To compare the activity of full length and truncated AAG, we repeated the single turnover glycosylase assays with full-length AAG (Fig. 2-4C; black circles). Both full-

length and Δ80 AAG were able to reach similar endpoints with comparable k_{max} values. Similar to the truncated protein, the $K_{1/2}$ for εG excision by full-length AAG was too low to accurately measure. The full-length protein also displayed inhibition at high enzyme concentrations, although to a lesser degree than the truncated protein. These results support the model that the N-terminal region of AAG is unnecessary for catalytic activity.

AAG recognizes εG poorly in other relevant DNA contexts

The oligonucleotides used in the preceding assays represent the expected context for the alkylation of a G•C pair to form an εG lesion. However, other pairings can occur during replication of the damaged template [6, 7]. We characterized the single turnover kinetics of εG•T excision and found rates similar to those of εG•C excision (Table 2-1). AAG exhibited similar enzyme concentration-dependent inhibition for excision of εG from εG•T as was observed for εG•C (Fig. A3). The similar glycosylase activity with either opposing pyrimidine is consistent with the similar multiple turnover rates of excision that were previously reported [25]. These results demonstrate that the inhibition of εG excision by excess AAG molecules is not exclusive to a single base complement.

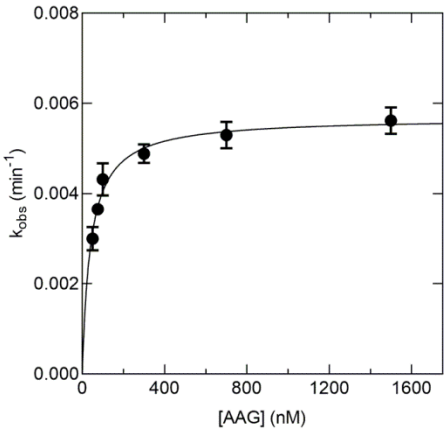


Figure 2-6. Single turnover excision of εG from a bulged substrate. The observed single turnover rate constants for the excision of bulged εG were plotted against varying concentrations of Δ80 AAG. A k_{max} value of 0.0057 min^{-1} was determined, with a $K_{1/2}$ value of 40 nM. Reactions contained 20 nM 25mer bulge DNA with the εG lesion. Points represent the mean of 4 replicates, and error bars indicate the S.D.

During replication of εG, DNA polymerases can slip to generate a -1 frameshift [31, 32]. The slippage event places the εG lesion into a bulged context without a complementary base. Previously it was shown that AAG excises εA and Hx from a bulged context at comparable rates to base-paired contexts [33, 34]. To assess the ability of AAG to recognize εG from a bulged conformation, single turnover glycosylase assays were performed with a 25mer bulged substrate (Fig. 2-6). Notably, the bulged lesion appeared to lack the concentration-dependent inhibitory effect observed with εG base pairs and the data could be readily fit to a typical hyperbolic concentration dependence. It is not clear why there was not an inhibitory effect at high concentrations of AAG, but it is possible that the presence of the bulge structure disrupts the competing nonspecific DNA binding sites which allows for better equilibration between the specific lesion site and the competing nonspecific sites. However, the bulged εG substrate was excised with a 10-fold lower k_{max} value than was observed for the duplex εG•C substrate (Table 2-1). This observation suggests that εG is inefficiently recognized in the bulged context, whereas AAG readily recognizes other base lesions in the same bulged context [33, 34]. An NMR structure of the εG bulge DNA [35] shows that the unpaired εG can be accommodated within the DNA duplex. It appears that this stable structure limits the ability of AAG to gain access and flip out the lesion, as compared to εG mispairs. Although AAG is capable of excising εG from a bulged context *in vitro*, it is clear that this is not a favorable context for AAG-initiated repair of εG.

Asn169 of AAG limits the rate of εG excision

Crystal structures of AAG bound to a flipped-out εA lesion revealed how this active site pocket can accommodate base lesions and exclude undamaged bases [13]. The side chain of N169 defines one surface of the active site, closely contacting the N^1 face of the εA lesion (Fig. 2-7A). This residue plays a role in blocking the binding of undamaged guanine with its N^2-amino group [11, 36], leading us to consider whether this side chain also contacts the 1,N^2-etheno ring of εG. We used site-directed mutagenesis to generate variants of Δ80 AAG with either a serine or an alanine residue in place of N169, and subsequently determined the impact of these mutations on the single turnover excision kinetics of εA and εG.

Table 2-1

Maximum rate constants for the AAG-catalyzed single turnover excision of εG[a]

Lesion	k_{max} (min^{-1})
εG•C	0.034 ± 0.005
εG•T	0.032[b]
εG bulge	0.0057 ± 0.0003
εA•T	0.23 ± 0.015

[a]Values are reported for Δ80 AAG at 37 °C in a buffer containing 50 mM NaMES (pH 6.1), 10% glycerol, 0.1 mg/mL BSA, 1 mM DTT, 1 mM EDTA, and 100 mM ionic strength adjusted with NaCl. Each value represents the average of ≥3 replicates. Reactions were fit to the biphasic model detailed in Figure S4 and the maximum single turnover rate constant k_{max} for the first phase in each reaction are listed. [b]The maximum rate constant was calculated from the average of 2 replicates (Fig. A3).

Table 2-2

Maximum rate constants for the single turnover excision of εG and εA by WT and mutant AAG [a]

AAG Variant	k_{max} (min^{-1})	
	εG•C[a]	εA•T[b]
WT	0.034 ± 0.005	0.23 ± 0.015
N169S	0.25 ± 0.004	0.58 ± 0.034
N169A	0.041 ± 0.006	1.4 ± 0.11

[a]Values are reported for Δ80 AAG at 37°C in a buffer containing 50 mM NaMES (pH 6.1), 10% glycerol, 0.1 mg/mL BSA, 1 mM DTT, 1 mM EDTA, and 100 mM ionic strength adjusted with NaCl. Each value represents the average of ≥3 replicates. [a]Reactions with εG•C were fit to the biphasic model detailed in Figure S4 and the maximum single turnover rate constant k_{max} for the first phase in each reaction are listed. [b]Reactions with εA•T were determined from exponential fits at saturating enzyme concentration of 1 μM.

The N169A mutation caused an increase in the k_{max} value for εG excision from 0.034 min^{-1} to 0.041 min^{-1} and the inhibition at high concentration of enzyme was much less prominent as compared to the behavior of the WT enzyme (Fig. 2-7B; Table 2-2). This increase in the excision rate of the N169A mutant suggests that N169 interferes with the excision of εG to some degree. Surprisingly, the k_{max} for εA excision was also substantially increased by the N169A mutation (Fig. 2-7C). This likely reflects a removal of a deleterious interaction between the Asn169 side chain and the εA lesion or could result from rearrangements in the Ala variant that create more favorable interactions with the substrate.

The N169S mutation is a more conservative change to the active site structure, as it maintains hydrogen bonding capability while shortening the side chain length to expand the binding pocket. The N169S variant displayed a dramatic elevation in the k_{max} values for excision of both εG and εA (Fig. 2-7C). However, the N169S mutant enhanced εG excision by more than 7-fold, while only increasing εA excision by 2-fold (Fig. 2-7D). The greater activity of N169S relative to N169A suggests the possibility of positive hydrogen bonding interactions for the serine side chain which were lacking in the alanine substitution. The N169S mutant also showed little to no enzyme concentration-dependent inhibition of εG excision, even up to concentrations several times higher than those tested for the WT enzyme. This absence of detectable inhibition can be explained by the previously proposed inhibition model, whereby the stronger recognition of εG by the N169S variant would enable the lesion to compete more favorably with undamaged sites for binding and excision. The observation that εG excision by AAG is improved by mutation of N169 is consistent with the model that this side chain clashes with the εG lesion and contributes to its inefficient excision by AAG.

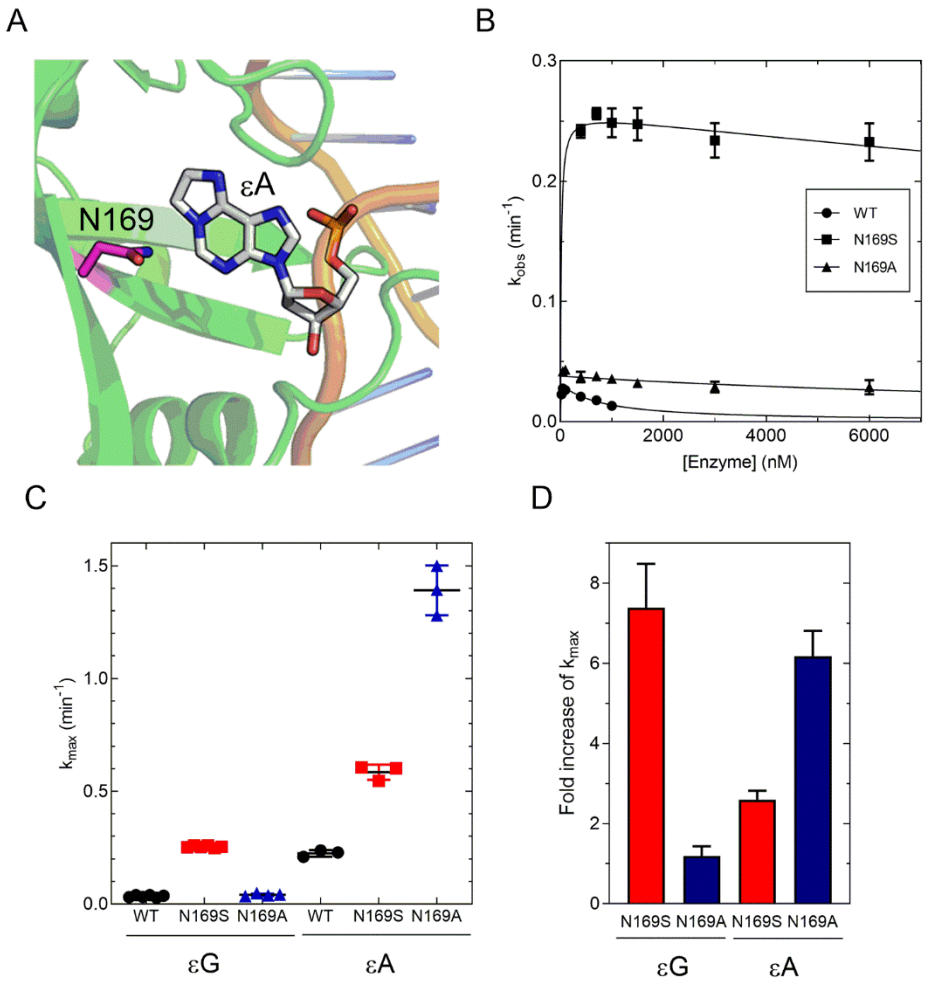


Figure 2-7. Asn169 of AAG restricts excision of εG. (A) In crystal structures of AAG bound to εA, the side chain of N169 is in close proximity to N^1 and C^2 of the nucleobase. If εG were to bind in the same conformation its exocyclic ring is predicted to clash with the side chain of Asn169. (B) Single turnover rate constants for the excision of εG, catalyzed by variants of Δ80 AAG. Reactions contained 20 nM εG in the 25mer DNA sequence with a C complement. Rate constants for WT are replotted from Fig. 2-4C, while data for N169S and N169A represent the average of 4 replicates. Error bars represent the standard deviation. (C) The maximum single turnover rate constants from panel B are plotted alongside values for εA excision. (D) The fold increase in the maximum single turnover rate constant due to each AAG mutation. Each value is calculated by dividing the values of k_{max} for the variant enzyme by the k_{max} for the WT enzyme. Error bars represent the propagated error of both measured values.

CONCLUSIONS

The exocyclic ring structure of the mutagenic and cytotoxic lesion εG presents a unique challenge for recognition by the DNA base excision repair pathway. It has previously been postulated that base excision repair initiated by AAG is the preferred mechanism for repair of both εA and εG [14, 25, 26]. Herein, we demonstrated that εG is excised with much lower efficiency than other primary substrates of AAG under both single- and multiple-turnover conditions. We also provide the first examination of competition between undamaged and εG-containing DNA, demonstrating that AAG has a difficult time recognizing εG sites.

Our findings highlight some of the limitations of single-turnover kinetic approach to studying DNA repair glycosylases with defined homogenous substrates. Whereas experiments with simple substrates are indispensable for quantitative analysis and the dissection of individual reaction steps, these assays neglect the impact of relevant cellular factors such as excess undamaged DNA. We infer that inhibition of εG excision at higher concentration of AAG protein is indicative of relatively poor lesion recognition, such that nonspecific binding modes compete with the lesion-specific binding mode, and this has now been observed for both AAG and *E. coli* AlkA [30]. We demonstrated that nonspecific competitor DNA competes effectively for binding of AAG to the εG lesion, suggesting that it would be difficult for AAG to effectively locate these lesions in the nucleus. The packaging of DNA into nucleosome core particles presents another potential challenge for the repair of εG in cells, as nucleosomes have been shown to restrict AAG and other glycosylases from accessing sites of DNA damage [37, 38].

To provide a physical explanation for the discrimination by AAG against εG, we have shown that residue Asn169 limits the ability of AAG to excise εG from DNA. This is consistent with the model that Asn169 plays a crucial role in governing the selectivity of the enzyme against substrates with a functional group at the C2 position [13, 36]. Mutation of Asn169 results in a substantial increase in the excision of undamaged G from mispairs, a promutagenic change that could offset the benefit of a more versatile active site [36, 39, 40]. The inability of AAG to efficiently catalyze the excision of εG may reflect a tradeoff

for greater specificity and discrimination against undamaged G nucleotides that are present in great excess within the genome [11, 36].

Considering the dangers posed to the cell by unrepaired εG lesions, the inability of AAG to efficiently excise εG suggests that other repair pathways are likely to bear primary responsibility for protecting the genome against this particular lesion. Previous *in vitro* studies of the human homologs of AlkB, which catalyze oxidative dealkylation of certain alkylated bases, demonstrates that ALKBH2, but not ALKBH3, is capable of recognizing and repairing εG in duplex context [41]. Although ALKBH2 is a strong candidate for physiological repair of εG•C, it is not known if this enzyme is able to capture rare εG lesions from amongst the excess of undamaged sites and it is not known how εG might be recognized in post-replicative repair. Although εG has not been specifically investigated, genetic studies of mice lacking either AAG, ALKBH2, or ALKBH3 show increased sensitivity to induced colitis [42]. Strong synergy was observed when all three genes were knocked out, demonstrating redundancy in the repair of DNA alkylation damage *in vivo* [42]. More research is needed to decipher the complexities of substrate specificity in mammalian alkylation repair.

EXPERIMENTAL PROCEDURES

Preparation of DNA

Undamaged oligonucleotides were synthesized by Integrated DNA Technologies, and lesion-containing oligonucleotides were synthesized by the W.M. Keck Facility at Yale University. The lesion-containing strand of each oligonucleotide was labeled at the 5' end with 6-fluorescein (FAM). Oligonucleotides were purified via denaturing polyacrylamide gel electrophoresis, and the concentrations were determined by the theoretical extinction coefficient at 260 nm as described previously [27].

```
25mer          5′ – (FAM) CGATAGCATCCTXCCTTCTCTCCAT
               3′ –       GCTATCGTAGGAYGGAAGAGAGGTA

19mer          5′ –     (FAM) TAGCATCCTXCCTTCTCTC
               3′ –           ATCGTAGGAYGGAAGAGAG

25mer bulge    5′ – (FAM) CGATAGCATCCTXCCTTCTCTCCAT
               3′ –       GCTATCGTAGGA-GGAAGAGAGGTA

X: εG, εA, Hx   Y: C, T
```

Scheme 1

Preparation of enzymes

The catalytic domain of AAG, (Δ80 AAG) was expressed and purified from *E. coli* C41(DE3) as described previously [28]. The construct for N169S was previously described [11] and the N169A mutation was generated by site-directed mutagenesis and confirmed by sequencing both strands of the open reading frame. These two variant AAG proteins were expressed and purified using the same methods. Full-length AAG was expressed and purified as described previously [43]. The concentration of each AAG variant was initially estimated using the UV absorbance and the active concentration of each enzyme was established through analysis of the burst kinetics for excision of hypoxanthine (Hx) as described previously [27].

Multiple turnover glycosylase assay

Reactions were performed at 37°C in reaction buffer containing 50 mM NaHEPES (pH 7.5), 10% glycerol, 0.1 mg/mL BSA, 1 mM DTT, 1 mM EDTA, and 100 mM ionic strength adjusted with NaCl. The DNA concentration was kept at a 50:1 ratio relative to the AAG concentration to ensure multiple turnover conditions. Aliquots were removed from the reactions at various time points and were quenched in an equal volume of 0.4 M NaOH to reach 0.2 M NaOH final concentration. The quenched aliquots were heated at 70°C for 12 minutes to cleave abasic sites, and then were mixed 1:2 with loading buffer (90% formamide, 10 mM EDTA, 0.025% bromphenol blue, 0.025% xylene cyanol). For time courses lasting longer than 24 hours, quenched samples were stored at 4 °C for no more than 12 hours before being heated and mixed with loading buffer. The samples were run out on 20% polyacrylamide gels containing 6.6 M urea and were scanned with an Amersham Typhoon 5 Biomolecular Imager (GE Healthcare Life Sciences). The samples were excited at 488 nm and the emission of fluorescein was measured with a 525BP20 filter. The bands on the gel were quantified using ImageQuant TL (GE Healthcare). The fraction product [product / (substrate + product)] was calculated for each lane, and the steady state formation of product was fit with linear regression. The change in observed reaction velocity at varying DNA concentrations was fit to the Michaelis-Menten equation (1).

$$V_{obs} = \frac{k_{cat}[E][S]}{K_M + [S]} \quad (1)$$

V_{obs} represents the observed initial reaction velocity, k_{cat} the steady state rate constant, [E] the concentration of enzyme, [S] the concentration of substrate, and K_M the Michaelis constant, equal to the concentration of DNA at the half-maximal velocity.

Single turnover glycosylase assay

To achieve single turnover conditions, glycosylase assays were performed with 10–20 nM DNA and 50 nM to 6 μM enzyme in reaction buffer containing 50 mM NaMES (pH 6.1), 10% glycerol, 0.1 mg/mL BSA, 1 mM DTT, 1 mM EDTA, and 100 mM ionic strength adjusted with NaCl. All reactions were performed at 37 °C. Aliquots were quenched and quantified as described above. Reactions were fit to a single exponential according to equation (2).

$$\text{Fraction Product} = A\left(1 - e^{-k_{obs}t}\right) + c \quad (2)$$

A represents the amplitude, k_{obs} the observed single turnover rate constant, t the reaction time, and c the starting amount of abasic DNA. The dependence of the single turnover rate constant, k_{obs}, on enzyme concentration was fit by a hyperbola according to equation (3) in which k_{max} represents the maximum k_{obs} value and $K_{1/2}$ represents the concentration at which enzyme is half saturating.

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

For reactions demonstrating enzyme-dependent inhibition, a multivalent inhibitory model was applied, in which K_i is the binding constant for the inhibitory complex (equation 4).

$$k_{obs} = \frac{k_{max}[E]}{K_{1/2} + [E]\left(1 + \frac{[E]}{K_i}\right)} \quad (4)$$

This model has been used previously for the nonspecific binding of another DNA glycosylase to multiple DNA sites [30]**.** For the titration of undamaged DNA, the IC_{50} was calculated using equation (5), where k_{obs} is the observed rate constant, k_{unin} is the rate constant without undamaged DNA inhibitor, and I is the concentration of undamaged inhibitor DNA.

$$k_{obs} = \frac{k_{unin}[I]}{IC_{50}+[I]} \qquad (5)$$

Acknowledgments – We thank Erin Taylor for preliminary observations and purified oligonucleotides, Michael Baldwin for purified full-length AAG, Karina Kangas for purified Δ80 AAG, and members of the O'Brien lab for helpful discussion.

REFERENCES

1. Lindahl, T., *Instability and decay of the primary structure of DNA.* Nature, 1993. **362**(6422): p. 709-15.

2. Beranek, D.T., *Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents.* Mutat Res, 1990. **231**(1): p. 11-30.

3. Gros, L., A.A. Ishchenko, and M. Saparbaev, *Enzymology of repair of etheno-adducts.* Mutat Res, 2003. **531**(1-2): p. 219-29.

4. Kochetkov, N.K., V.N. Shibaev, and A.A. Kost, *New reaction of adenine and cytosine derivatives, potentially useful for nucleic acids modification.* Tetrahedron Letters, 1971. **12**(22): p. 1993-1996.

5. Guengerich, F.P., *Roles of the vinyl chloride oxidation products 1-chlorooxirane and 2-chloroacetaldehyde in the in vitro formation of etheno adducts of nucleic acid bases [corrected].* Chem Res Toxicol, 1992. **5**(1): p. 2-5.

6. Langouet, S., M. Muller, and F.P. Guengerich, *Misincorporation of dNTPs opposite 1,N2-ethenoguanine and 5,6,7,9-tetrahydro-7-hydroxy-9-oxoimidazo[1,2-a]purine in oligonucleotides by Escherichia coli polymerases I exo- and II exo-, T7 polymerase exo-, human immunodeficiency virus-1 reverse transcriptase, and rat polymerase beta.* Biochemistry, 1997. **36**(20): p. 6069-79.

7. Choi, J.Y., et al., *Translesion synthesis across 1,N2-ethenoguanine by human DNA polymerases.* Chem Res Toxicol, 2006. **19**(6): p. 879-86.

8. Langouet, S., et al., *Misincorporation of nucleotides opposite five-membered exocyclic ring guanine derivatives by escherichia coli polymerases in vitro and in vivo: 1,N2-ethenoguanine, 5,6,7,9-tetrahydro-9-oxoimidazo[1, 2-a]purine, and 5,6,7,9-tetrahydro-7-hydroxy-9-oxoimidazo[1, 2-a]purine.* Biochemistry, 1998. **37**(15): p. 5184-93.

9. Akasaka, S. and F.P. Guengerich, *Mutagenicity of site-specifically located 1,N2-ethenoguanine in Chinese hamster ovary cell chromosomal DNA.* Chem Res Toxicol, 1999. **12**(6): p. 501-7.

10. Krokan, H.E. and M. Bjoras, *Base excision repair.* Cold Spring Harb Perspect Biol, 2013. **5**(4): p. a012583.

11. 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.

12. 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.

13. 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.

14. Dosanjh, M.K., et al., *All four known cyclic adducts formed in DNA by the vinyl chloride metabolite chloroacetaldehyde are released by a human DNA glycosylase.* Proc Natl Acad Sci U S A, 1994. **91**(3): p. 1024-8.

15. Rydberg, B., M.K. Dosanjh, and B. Singer, *Human cells contain protein specifically binding to a single 1,N6-ethenoadenine in a DNA fragment.* Proc Natl Acad Sci U S A, 1991. **88**(15): p. 6839-42.

16. Lingaraju, G.M., et al., *Structural basis for the inhibition of human alkyladenine DNA glycosylase (AAG) by 3,N4-ethenocytosine-containing DNA.* J Biol Chem, 2011. **286**(15): p. 13205-13.

17. Fu, D. and L.D. Samson, *Direct repair of 3,N(4)-ethenocytosine by the human ALKBH2 dioxygenase is blocked by the AAG/MPG glycosylase.* DNA Repair (Amst), 2012. **11**(1): p. 46-52.

18. Kavli, B., et al., *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. 39926-36.

19. Goto, M., et al., *Human DNA glycosylase enzyme TDG repairs thymine mispaired with exocyclic etheno-DNA adducts.* Free Radic Biol Med, 2014. **76**: p. 136-46.

20. Morinello, E.J., et al., *Simultaneous quantitation of N(2),3-ethenoguanine and 1,N(2)-ethenoguanine with an immunoaffinity/gas chromatography/high-resolution mass spectrometry assay.* Chem Res Toxicol, 2001. **14**(3): p. 327-34.

21. Sattsangi, P.D., N.J. Leonard, and C.R. Frihart, *1,N2-Ethenoguanine and N2,3-ethenoguanine. Synthesis and comparison of the electronic spectral properties of these linear and angular triheterocycles related to the Y bases.* The Journal of Organic Chemistry, 1977. **42**(20): p. 3292-3296.

22. Sodum, R.S. and F.L. Chung, *1,N2-ethenodeoxyguanosine as a potential marker for DNA adduct formation by trans-4-hydroxy-2-nonenal.* Cancer Res, 1988. **48**(2): p. 320-3.

23. Muller, M., et al., *Analysis of 1,N2-ethenoguanine and 5,6,7,9-tetrahydro-7-hydroxy-9-oxoimidazo[1,2-a]purine in DNA treated with 2-chlorooxirane by high performance liquid chromatography/electrospray mass spectrometry and*

comparison of amounts to other DNA adducts. Chem Res Toxicol, 1997. **10**(2): p. 242-7.

24. Dimitri, A., et al., *Transcription processing at 1,N2-ethenoguanine by human RNA polymerase II and bacteriophage T7 RNA polymerase.* J Mol Biol, 2008. **375**(2): p. 353-66.

25. Saparbaev, M., et al., *1,N(2)-ethenoguanine, a mutagenic DNA adduct, is a primary substrate of Escherichia coli mismatch-specific uracil-DNA glycosylase and human alkylpurine-DNA-N-glycosylase.* J Biol Chem, 2002. **277**(30): p. 26987-93.

26. 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.

27. Hedglin, M. and P.J. O'Brien, *Human alkyladenine DNA glycosylase employs a processive search for DNA damage.* Biochemistry, 2008. **47**(44): p. 11434-45.

28. 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. 12418-29.

29. Biswas, T., et al., *Binding of specific DNA base-pair mismatches by N-methylpurine-DNA glycosylase and its implication in initial damage recognition.* J Mol Biol, 2002. **320**(3): p. 503-13.

30. Zhao, B. and P.J. O'Brien, *Kinetic mechanism for the excision of hypoxanthine by Escherichia coli AlkA and evidence for binding to DNA ends.* Biochemistry, 2011. **50**(20): p. 4350-9.

31. Chang, S.C., et al., *Next-generation sequencing reveals the biological significance of the N(2),3-ethenoguanine lesion in vivo.* Nucleic Acids Res, 2015. **43**(11): p. 5489-500.

32. Zang, H., et al., *DNA adduct bypass polymerization by Sulfolobus solfataricus DNA polymerase Dpo4: analysis and crystal structures of multiple base pair substitution and frameshift products with the adduct 1,N2-ethenoguanine.* J Biol Chem, 2005. **280**(33): p. 29750-64.

33. 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. 17742-3.

34. Lyons, D.M. and P.J. O'Brien, *Human base excision repair creates a bias toward -1 frameshift mutations.* J Biol Chem, 2010. **285**(33): p. 25203-12.

35. Shanmugam, G., et al., *Structure of the 1,N(2)-etheno-2'-deoxyguanosine lesion in the 3'-G(epsilon dG)T-5' sequence opposite a one-base deletion.* Biochemistry, 2010. **49**(12): p. 2615-26.

36. Connor, E.E. and M.D. Wyatt, *Active-site clashes prevent the human 3-methyladenine DNA glycosylase from improperly removing bases.* Chem Biol, 2002. **9**(9): p. 1033-41.

37. Kennedy, E.E., P.J. Caffrey, and S. Delaney, *Initiating base excision repair in chromatin.* DNA Repair (Amst), 2018. **71**: p. 87-92.

38. Kennedy, E.E., C. Li, and S. Delaney, *Global Repair Profile of Human Alkyladenine DNA Glycosylase on Nucleosomes Reveals DNA Packaging Effects.* ACS Chem Biol, 2019. **14**(8): p. 1687-1692.

39. Eyler, D.E., et al., *Mechanisms of glycosylase induced genomic instability.* PLoS One, 2017. **12**(3): p. e0174041.

40. Connor, E.E., J.J. Wilson, and M.D. Wyatt, *Effects of substrate specificity on initiating the base excision repair of N-methylpurines by variant human 3-methyladenine DNA glycosylases.* Chem Res Toxicol, 2005. **18**(1): p. 87-94.

41. Zdzalik, D., et al., *Differential repair of etheno-DNA adducts by bacterial and human AlkB proteins.* DNA Repair (Amst), 2015. **30**: p. 1-10.

42. Calvo, J.A., et al., *DNA repair is indispensable for survival after acute inflammation.* J Clin Invest, 2012. **122**(7): p. 2680-9.

43. Baldwin, M.R. and P.J. O'Brien, *Nonspecific DNA binding and coordination of the first two steps of base excision repair.* Biochemistry, 2010. **49**(36): p. 7879-91.

CHAPTER 3

DNA Bending by AAG Is Associated With Base Flipping Rather Than DNA Searching

ABSTRACT

Alkyladenine DNA glycosylase (AAG) initiates the DNA base excision repair pathway by hydrolytically excising alkylated and deaminated nucleobases from the genome. AAG employs a nucleotide flipping mechanism to access damaged sites, rotating the damaged nucleotide into the active site and introducing a bend into DNA. This strategy is common across glycosylase structural families. DNA bending by glycosylases has been proposed to contribute to genome searching based on observations of bending in the absence of DNA damage. To elucidate the role of DNA bending in the recognition of damaged DNA by AAG, we characterized the kinetics of bending by AAG using stopped-flow fluorescence resonance energy transfer (FRET) with end-labeled oligonucleotides. These results were compared with the kinetics of binding and nucleotide flipping measured using the fluorescent lesion 1,N^6-ethenoadenosine. To further explore the relationship between bending and flipping, we mutated AAG binding pocket residues Y127 and Y159 and introduced flexibility into the DNA with spacers on the opposing strand. The rate constants for DNA bending increased and decreased in tandem with base flipping, suggesting that DNA bending does not limit the rate of base flipping. These results demonstrate that the initial recognition of the lesion occurs prior to DNA bending and DNA bending occurs on the same time scale as nucleotide flipping. We conclude that DNA bending does not contribute to DNA searching and instead accompanies and potentially facilitates nucleotide flipping. This suggests DNA glycosylases may have evolved distinct mechanisms for identifying sites of damage.

INTRODUCTION

The individual nucleobases of DNA are subject to modification through oxidation, alkylation, and deamination, leading to numerous forms of DNA damage. These DNA lesions can arise through normal cellular processes as well as through exposure to reactive exogenous factors [3]. Modifications to the structure of DNA bases are usually deleterious, causing mutation or cell death. The base excision repair (BER) pathway is conserved across all domains of life as the front-line defense against damaged nucleobases [4]. Estimates of the number of sites repaired through this pathway reach as high as 10,000 sites per day per cell and primarily include small single-nucleotide modifications [3, 5]. To identify and repair small lesions from among the vast excess of undamaged bases, the enzymes involved must be both highly efficient at locating sites of damage and highly specific for the targeted lesions.

Glycosylases initiate BER by catalyzing the hydrolytic cleavage of damaged nucleotides at the *N*-glycosidic bond to release the base lesion. The resulting abasic site is further processed by an AP endonuclease to produce a single-strand break. A DNA polymerase and DNA ligase then replace the excised nucleotide to pair with the undamaged opposing strand. Mammalian cells contain at least 11 different glycosylases, which together recognize a wide spectrum of DNA damage. Glycosylases can be separated into 4 primary structural families: helix-hairpin-helix, helix-2-turn-helix, alkyladenine DNA glycosylase, and uracil DNA glycosylase. Despite their structural differences, glycosylases nevertheless share many strategies for accessing sites of damage. All known mammalian glycosylase families use a nucleotide flipping mechanism to reach lesions within the DNA duplex. Rotating a lesion out of the duplex and into the active site provides access to the *N*-glycosidic bond as well as the opportunity to discriminate between lesions based on the structure of the active site pocket. Many glycosylases also intercalate one or more amino acid side chains into the DNA duplex to fill the gap left by the flipped-out base.

Crystal structures of multiple glycosylases in the flipped-out recognition complex show a bend in DNA at the site of glycosylase binding [1, 6-12]. These structures include

glycosylases from all 4 primary structural families and from both eukaryotes and prokaryotes, suggesting an integral role for bending in the activity of these enzymes. The degree of DNA bending in glycosylase crystal structures varies widely. Human OGG1 has the highest published bend angle at 70°, while human AAG bends DNA by only 20°. It is not clear from these structures alone whether these differences in bend angle are due to genuine mechanistic differences or simply the varying crystallization conditions and packing in these structures. However, atomic force microscopy studies have also characterized a range of DNA bending angles for different enzymes, suggesting some variation between glycosylases [13, 14].

Despite the widespread conservation of DNA bending by glycosylases, the purpose of DNA bending by glycosylases remains poorly understood. It has been proposed that bending plays a role in the DNA searching process [13-15]. Damaged DNA may be easier to bend because disruption of base pairing or base stacking increases the intrinsic flexibility of the DNA duplex. Nonspecific recognition of these sites of increased flexibility would allow glycosylases to narrow their search for compatible sites of damage without the need to flip every undamaged nucleotide, and it could lower the barrier to nucleotide flipping. This model of DNA searching through bending has been proposed for proteins in other DNA repair pathways, such as the MutS enzymes which initiate mismatch repair [16]. Evidence for this type of searching by glycosylases remains limited, however the observation by atomic force microscopy that glycosylases can bend undamaged DNA would be consistent with DNA bending as an early step in the search for damage [13, 14, 17].

Human alkyladenine DNA glycosylase (AAG) is a monomeric enzyme with a broad specificity to excise alkylated and deaminated purines such as 1N^6-ethenoadenine (εA), 3-methyladenine, and hypoxanthine. The intrinsic fluorescence of the εA lesion makes AAG an ideal candidate to dissect the mechanistic contributions of DNA bending to lesion recognition. The minimal kinetic mechanism of AAG has been previously characterized using εA fluorescence [18]. AAG binds nonspecifically to DNA then initiates a search for a site of compatible damage. This searching can involve hopping between nearby sites to navigate around obstructions [19-21]. During the DNA search, it is proposed that the

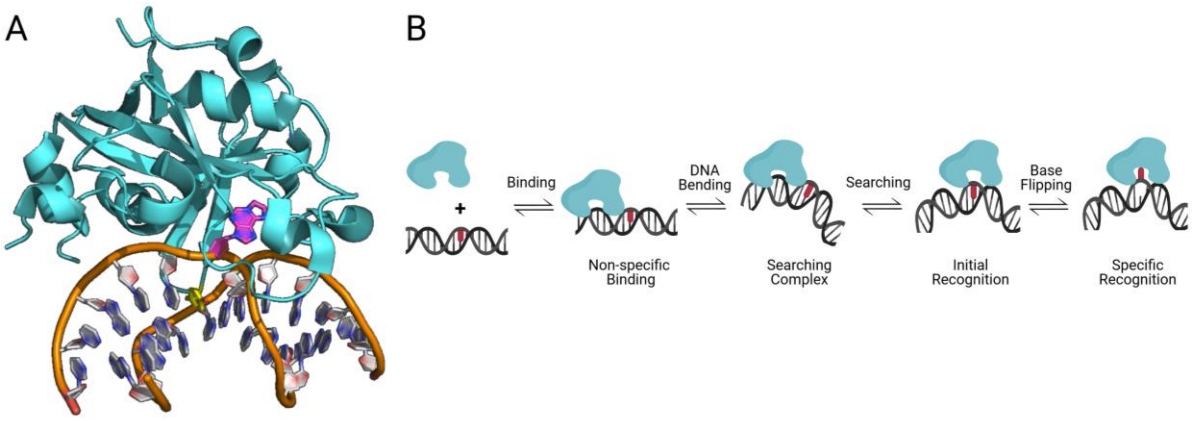


Figure 3-1. Model for DNA bending by AAG. (A) Crystal structure of human AAG (E125Q) bound to DNA with a flipped-out εA lesion (magenta). Coordinates taken from the PDB (1EWN) [1] and rendered using PyMOL (http://www.pymol.com). The backbone of the lesion strand is bent by approximately 20° at the site of AAG binding. (B) Proposed mechanism for lesion recognition by AAG, using DNA bending to detect sites of damage. AAG binds nonspecifically to DNA, bends the DNA backbone, and searches rapidly for a compatible site of damage. Once a site of duplex instability is found, AAG pauses and flips the lesion (red) into the active site for cleavage.

β3β4 hairpin and Y162 side chain intercalate into the minor groove of DNA [22]. Once the protein identifies a site of damage, a reversible initial recognition complex is formed that is marked by partial unstacking of the εA lesion [18]. The εA lesion is then fully flipped into the active site while the intercalating Y162 residue fills the gap in DNA left by the flipped-out base (Fig. 3-1A) [2]. This specific recognition complex proceeds to *N*-glycosidic bond cleavage. Placing the timing of DNA bending within this established mechanism would clarify the role of DNA bending and evaluate the commonly proposed model that DNA bending contributes to DNA searching (Fig. 3-1B).

Large bends in the backbone of DNA can be detected by a variety of techniques such as electrophoretic mobility, atomic force microscopy, and fluorescence resonance energy transfer (FRET). Among these strategies, FRET is the ideal approach for examining the kinetics of DNA bending due to the excellent time resolution of stopped-flow fluorescence measurements. Previous studies have successfully employed FRET to detect equilibrium DNA bending by human glycosylase enzymes using end-labeled oligonucleotides [13]. Stopped-flow FRET measurements have also been used to

characterize the kinetics of DNA bending by diverse DNA-binding proteins like TATA binding protein, HMG1, and T7 RNA polymerase [23-25].

In this study we used transient kinetic approaches to demonstrate that bending occurs subsequent to lesion recognition in a variety of contexts. We perturbed base flipping by mutating Y127 and Y159 in the active site pocket and we altered the flexibility of the DNA by introducing spacers on the opposing strand. The results are consistent with synchronous bending and flipping processes, but we cannot rule out the possibility that flipping is rate limiting and rapid bending follows. Therefore, the bending of DNA by AAG is not a searching strategy, but rather an adaption to nucleotide flipping that stabilizes the specific recognition complex. These results suggest that AAG has evolved to use DNA intercalation by a β-hairpin loop to search DNA and recognize lesions with disrupted base pairing. Structural comparisons suggests that other glycosylases may have independently evolved a similar strategy, whereas others may use a bending mechanism like that of MutS to identify damaged sites.

EXPERIMENTAL PROCEDURES

Preparation of oligonucleotides

Undamaged and unlabeled oligos were synthesized by Integrated DNA Technologies. Damaged oligos and fluorescently labeled oligos were synthesized by the W.M. Keck Facility at Yale University. Oligos with a Cy5 fluorescent tag were labeled on the 5’ end of the lesion-containing strand, while Cy3 was placed at the 5’ end of the undamaged complementary strand (Fig. 3-2A). Oligos were further purified using denaturing PAGE, then extracted and desalted on a C18 column (Sep-Pak). The concentration of each oligo was determined using the theoretical extinction coefficient at 260 nm. The extinction coefficients for labeled and damaged oligos were calculated by using the coefficient for an unlabeled strand containing an A in place of the lesion, then subtracting 9400 $M^{-1}cm^{-1}$ for the εA lesion, 5000 $M^{-1}cm^{-1}$ for Cy3, and 10000 $M^{-1}cm^{-1}$ for Cy5.

Preparation of enzymes

The AAG catalytic domain (Δ80 AAG) was expressed in *E. coli* C41(DE3) cells and purified as described previously [26]. The constructs for the Y127W and Y159W variants have each been described previously [27]. The AAG variants were expressed and purified using the same methods as the WT. A preliminary concentration for each variant was established using UV absorbance. The active concentration of each enzyme was determined by titrating the enzyme against 200 nM DNA containing an εA damage site. The quenching of εA fluorescence at 405 nm was plotted against the concentration of AAG to establish an equivalence point [27].

Steady-state FRET assay

To measure steady state changes in FRET upon binding of AAG, AAG was titrated against fluorescently labeled DNA. Fluorescence measurements were performed in a HORIBA FluoroMax-3 fluorometer with 10 mm quartz cuvettes. Binding titrations were performed at 25 °C in reaction buffer consisting of 50 mM NaHEPES (pH 7.5), 100 mM NaCl, 1 mM EDTA, 1 mM DTT, and 10% glycerol. Each sample was measured twice, first by exciting at 530 nm and recording fluorescence at 560 nm and 660 nm, then by exciting at 630 nm and recording fluorescence at 660 nm. AAG was titrated by successively adding 1-5 mL volumes of concentrated AAG to the 2.2 mL total reaction volume. The total binding time was kept under 15 min to prevent significant *N*-glycosidic bond cleavage. The FRET efficiency of each sample was calculated with the indicated corrections (eq 1) [28]. F represents the fluorescence of doubly-labeled DNA, D represents DNA with only a Cy3 donor, and A represents DNA with only a Cy5 acceptor. The subscript DD indicates excitation of the donor and emission of the donor, DA represents excitation of the donor and emission of the acceptor, and AA represents excitation of the acceptor and emission of the acceptor. DD_0 and AA_0 represent the fluorescence of these oligos with no enzyme added. The AAG concentration dependence of FRET efficiency was fit with segmental linear regression in GraphPad Prism.

$$FRET = \frac{A_{adj}}{A_{adj} + D_{adj}}$$

$$D_{adj} = F_{DD} * \frac{D_{DD\,0}}{D_{DD}}$$

$$A_{adj} = \left(F_{DA} - \left(F_{DD} * \frac{D_{DA}}{D_{DD}}\right) - \left(F_{AA} * \frac{A_{DA}}{A_{AA}}\right)\right) * \frac{A_{AA\,0}}{A_{AA}}$$

Stopped-flow fluorescence kinetics

Pre-steady state fluorescence experiments were performed using a Hi-Tech SF-61DX2 Stopped-Flow System and Kinetic Studio software (TgK Scientific). The fluorescence of εA was measured by exciting samples at 313 nm with a WG360 long-pass emission filter. For oligos with a Cy3 and Cy5 label, samples were excited at 547 nm with an RPE660LP long-pass emission filter. For reactions using singly-labeled oligos, Cy3 was excited at 547 nm and measured with a WG590 long-pass emission filter, while Cy5 was excited at 577 nm and measured with the RPE660LP filter. All reactions were performed at 25 °C in buffer containing 50 mM NaMES (pH 6.5), 100 mM NaCl, 1 mM EDTA, and 1 mM DTT. Reactions were run for 1-20 s, with a 3 ms dead time. A total of 5 shots were performed and averaged for each reaction. Three independent experiments were individually analyzed and the rate constants were averaged.

Reaction models illustrated with BioRender.com.

RESULTS

DNA bending by AAG can be detected by FRET

To interrogate the timing of DNA bending by AAG, we developed a FRET assay using a fluorescently labeled oligonucleotide substrate to report on changes in DNA backbone angle. The 19mer substrate was designed with a centrally placed εA lesion and fluorescent dyes on the 5’ ends of each strand (Fig. 3-2A). The FRET acceptor Cy5 was placed on the 5’ end of the lesion-containing strand and the compatible FRET donor Cy3 was placed on the 5’ end of the complement strand. DNA bending is expected to increase

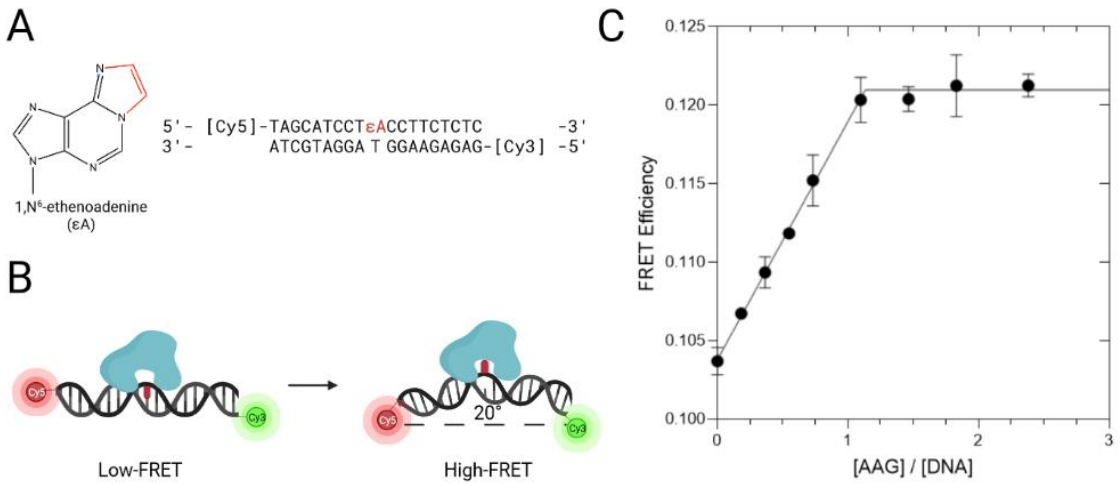


Figure 3-2. FRET assay for detection of DNA bending by AAG. (A) Structure of the εA lesion and sequence of the 19mer FRET substrate used for the detection of DNA bending. (B) Diagram of the FRET assay to measure DNA bending. Binding of AAG to DNA induces an expected ~20° bend, bringing the Cy3 FRET donor and Cy5 FRET acceptor in closer proximity to produce a high-FRET state. (C) Steady state FRET bending titration. Increasing concentrations of AAG were titrated against 100 nM DNA containing the Cy3 FRET donor, Cy5 FRET acceptor, and Cy3+Cy5 FRET pair. FRET efficiency increased linearly up to stochiometric DNA:AAG, with no impact of additional AAG molecules on FRET efficiency.

the FRET efficiency by decreasing the distance between the donor and acceptor fluorophores (Fig. 3-2B).

To validate the ability of AAG to induce a bend in DNA and produce a detectible change in FRET efficiency, steady state measurements of FRET were performed with increasing concentrations of AAG in a fluorometer. As previously noted, cleavage of εA from the duplex occurs slowly relative to binding under these conditions and therefore the FRET efficiency of the substrate-complex could be determined [27]. Measurements were performed with both the Cy3/Cy5-labeled FRET substrate as well as singly-labeled oligos bearing only Cy3 or Cy5 to control for indirect excitation and emission of the dyes. FRET efficiency increased linearly with the titration of AAG up to 1 equivalent of protein (Fig. 3-2C). Beyond stoichiometric AAG, no further change in FRET efficiency was observed, indicating that the increase in FRET efficiency reflects a bending signal from a single specific AAG binding event and is not affected by additional nonspecific binding at higher concentration of AAG.

We performed stopped-flow fluorescence measurements to monitor the binding of AAG to an εA-containing duplex. Rapid mixing of AAG with a slight excess of the doubly-labeled FRET substrate resulted in 4 distinct fluorescence changes (Fig. 3-3). These 4 phases were not dependent on the DNA sequence context, as alteration of the bases directly up and downstream of the lesion produced the same fluorescence phases (Fig. B2). The initial increase in fluorescence intensity was too rapid to capture. When the

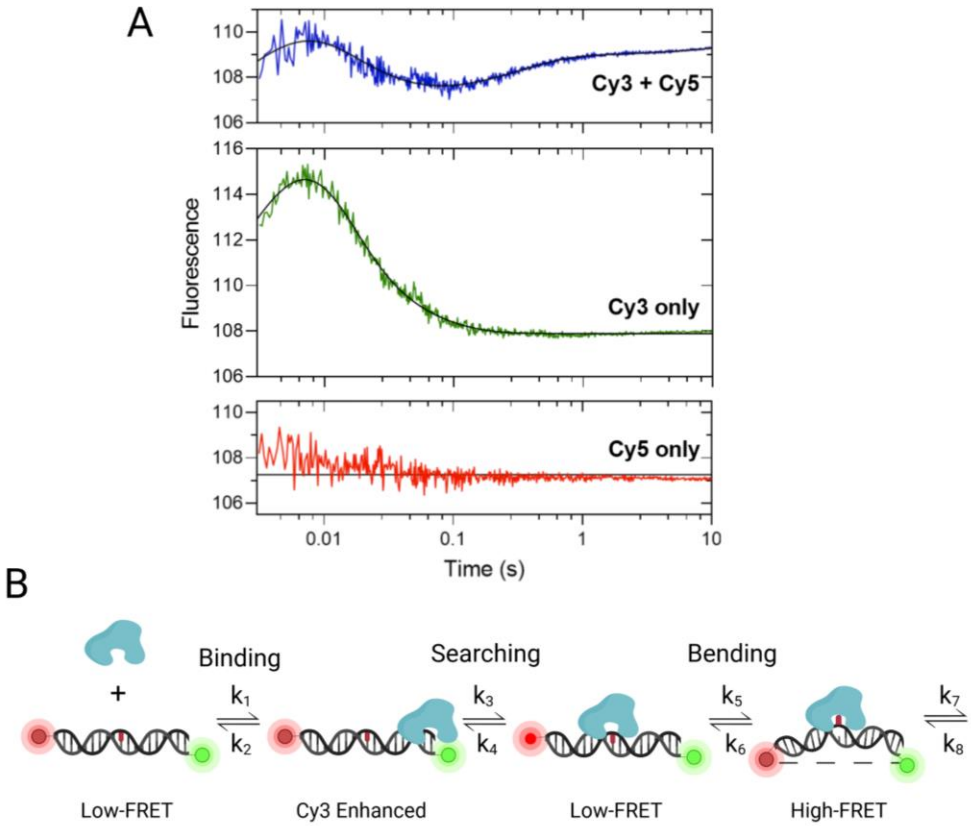


Figure 3-3. Stopped-flow FRET assay for DNA bending. (A) Representative stopped-flow fluorescence traces of 90 nM AAG mixed with 100 nM DNA containing a Cy3 FRET donor (green), a Cy5 FRET acceptor (red), or both (blue). The Cy3+Cy5 trace was fit to the 4-step model described in (B) using Berkeley Madonna (Fig. B4 A-C), producing the effectively irreversible forward rate constants k_1= 3.3x10^8 M^{-1}s^{-1}, k_3=310 s^{-1}, and k_5=3.9 s^{-1}. The Cy3 only trace lacked the k_5 phase and was fit to the forward rate constants k_1=5.1x10^8 M^{-1}s^{-1} and k_3=260 s^{-1}, roughly matching the k_1 and k_3 values with both fluorophores present. Early fluctuations in the Cy5 only trace may correspond to the same k_1 and k_3 phases as the Cy3 only trace, but could not be fit convincingly due to relatively low signal.

assay was repeated with εA-DNA containing either the Cy3 donor only or the Cy5 acceptor only, a large increase in Cy3 fluorescence and a modest increase in Cy5 fluorescence was observed on the same timescale (Fig. 3-3). This demonstrates that the increase in fluorescence is due to direct interaction with AAG, which is further supported by the observation that similar increases in fluorescence were obtained when AAG was mixed with undamaged DNA (Fig. B1). Modeling this phase as limited by binding of AAG to the DNA produces a rate constant of 3.3×10^8 $M^{-1}s^{-1}$, consistent with previously reported values for nonspecific binding by AAG under similar conditions [18, 27].

The initial increase in fluorescence was followed by a decrease in fluorescence intensity on the millisecond time scale. Unlike the first phase, this second fluorescence change was dependent on the presence of damage on the DNA (Fig. B1). However, mixing of AAG with the Cy3-only substrate produced a decrease in Cy3 fluorescence on the same millisecond time scale (Fig. 3-3). Due to the damage-dependent loss of fluorescence, this phase must correspond to AAG successfully locating the target lesion and transitioning away from the nonspecific binding sites near the ends of the oligonucleotide duplex. Modeling this step as a unimolecular search for the lesion site produces a rate constant of 310 s^{-1}.

The third phase of FRET change occurred only with damaged DNA substrates and required both the Cy3 donor and Cy5 acceptor (Fig. 3-3). This fluorescence change is therefore consistent with DNA bending decreasing the average distance between the two fluorophores. The third fluorescence phase occurred with a rate constant of 3.9 s^{-1}, which is similar to the expected rate constant for base-flipping. The fourth fluorescence change was comparatively small, modeled here with a rate constant of 0.04 s^{-1}. This final phase could mark a small change in bending long after formation of the specific recognition complex, but prior to cleavage of the target lesion. However, we cannot rule it out as an artifact due to small changes in the fluorescence of the Cy3-only DNA and undamaged Cy3/Cy5 DNA observed on the same time scale (Fig. 3-3, Fig B1).

Comparison of DNA bending to nucleotide flipping

Previous studies demonstrate that the intrinsic fluorescence of the εA lesion reports on both the formation of the initial recognition complex with AAG as well as the

flipping of the lesion into the active site to form the specific recognition complex [18]. To prevent interference from the brighter Cy3 and Cy5 dyes, we prepared the 19mer εA DNA duplex without either FRET dye. This oligo is identical in sequence context to a previously characterized 25mer oligo [18].

Rapid mixing of εA-DNA with increasing excess concentrations of AAG produced 3 transient fluorescence changes. The initial increase in εA fluorescence has been previously demonstrated to correspond to the partial unstacking of the lesion as AAG binds at the site of damage. This is followed by strong quenching of εA fluorescence as the base is flipped into the enzyme active site. As expected, the rate of flipping was independent of AAG concentration (Fig. 3-4). To obtain the clearest measurement of flipping, experiments were repeated in triplicate with 300-500nM DNA and excess AAG. A unimolecular rate constant of 4.4 s^{-1} was established for flipping (Table 3-1), which is comparable to the rate constant of 3.6 s^{-1} reported for the 25mer εA-DNA under similar conditions [22]. A slow, final small quenching of fluorescence was observed that could reflect a subsequent rearrangement, however similar slow changes in εA fluorescence

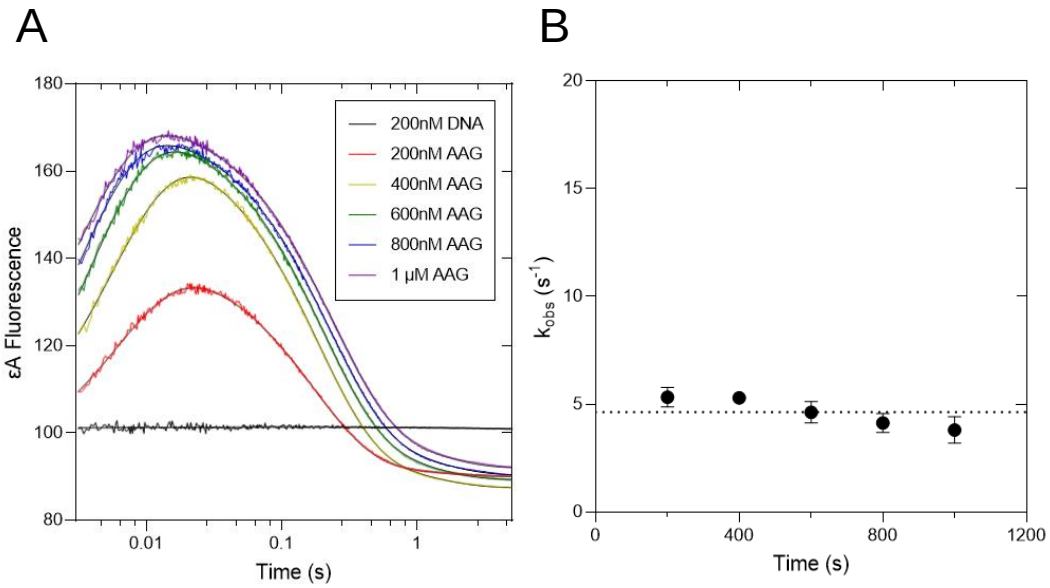


Figure 3-4. Flipping of εA occurs on the same timescale as DNA bending. (A) Representative stopped-flow fluorescence traces of 200 nM unlabeled εA-DNA mixed with varying concentrations of AAG. Reactions were excited at 313 nm with a 360 nm longpass filter, and the resulting traces were fit to the sum of 3 exponentials. (B) Concentration dependence of the rate constants for the second phase (base flipping) in Figure 3-4A.

were sometimes seen in control experiments lacking AAG and therefore this phase was treated as an artifact rather than a genuine εA rearrangement.

To obtain the clearest DNA bending signal for quantitative comparison with base flipping, the FRET DNA bending assay described in Fig. 3-3 was repeated with 100 nM doubly-labeled DNA and 50-80 nM AAG. To limit the complications of interactions between AAG and 5' fluorophores, a slight excess of DNA was maintained. Traces were fit to the 4-step model described in Fig. 3-3B. As with base flipping, DNA bending was fit to a unimolecular step with a rate constant of 4.4 s^{-1} (Table 3-1). This rate constant for bending is a close match for the rate constant established for base flipping, suggesting these measurements of DNA bending and base flipping report on the same step. In the simplest model this would suggest a concerted mechanism of DNA bending and base flipping, however we cannot rule out models where DNA bending or base flipping are rate limiting for the opposing step.

Impact of AAG mutations on kinetics of DNA bending and nucleotide flipping

The AAG lesion binding pocket is defined in large part by a handful of aromatic residues that interface directly with the flipped-out target lesion. Crystal structures reveal that Y127 stacks with one face of the flipped εA lesion and Y159 makes a T-shaped junction with the opposite face of the damaged base (Fig. 3-5A) [1]. These active site residues govern the rate of base flipping by AAG, as the Y127W variant slows the rate of flipping and the Y159W variant increases the rate of flipping [27].

To assess the relationship between DNA bending and nucleotide flipping, we performed stopped-flow mixing experiments with these variants to determine the kinetic parameters for both flipping and bending. If DNA bending were to restrict nucleotide flipping, then we expected to see that the two steps would become decoupled if we changed the nucleotide flipping step by decreasing the size of the active site pocket. Binding reactions were performed with unlabeled εA-DNA and a slight excess of AAG to maximize the εA flipping signal. Consistent with the prior results, the WT flipping rate constant of 4.4 s^{-1} was decreased to 1.7 s^{-1} in the Y127W variant (Fig. 3-5B, Table 3-1) [27]. The opposite was observed with the Y159W mutation, which increased the flipping rate constant to 27.4 s^{-1} (Fig. 3-5B, Table 3-1).

The kinetics of DNA bending for Y127W and Y159W was measured with a slight excess of FRET substrate over AAG to minimize interactions between excess AAG and the 5' FRET labels. Single fluorophore controls were also performed to validate the FRET signal for each variant (Fig. B3). The Y127W mutant showed a reduced bending rate constant of 1.5 s^{-1}, while Y159W increased the rate constant to 32.8 s^{-1}. (Fig. 3-5C, Table 3-1). The rate constants for DNA bending closely parallel the altered rate constants for nucleotide flipping by each variant, supporting a concerted or flipping-limited model for DNA bending. The accelerated flipping of the Y159W variant in particular provides strong evidence against the possibility that DNA bending limits the rate of base flipping, as an increase in the rate of flipping should not also accelerate a rate-limiting DNA bending step.

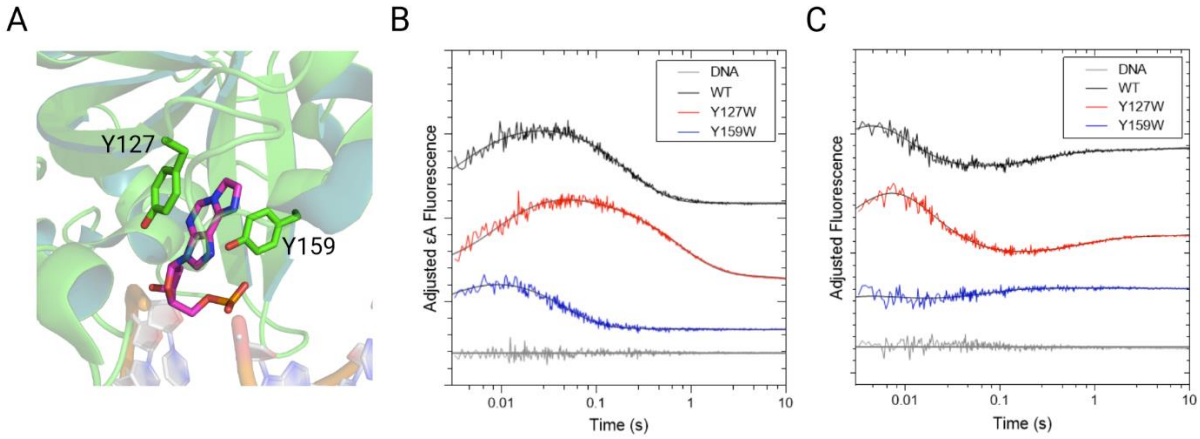


Figure 3-5. Mutation of AAG active site residues alters both base flipping and DNA bending. (A) View of the AAG lesion-binding pocket highlighting the position of the Y127 and Y159 residues that interact with the flipped-out εA lesion (rendered with PyMOL using 1EWN). (B) Representative stopped-flow traces of εA fluorescence after mixing damaged DNA with AAG variants. Each reaction contained 300 nM εA DNA and a small excess of 350 nM of the indicated AAG variant. Reactions were modeled in Berkeley Madonna as described in Fig. B5 A-B and the rate constants for base flipping are listed in Table 3-1. (C) Representative stopped-flow fluorescence traces of FRET DNA mixed with AAG variants. Reactions were performed with 100 nM FRET DNA and 80 nM AAG, with fluorescence measured using a 660 nm long-pass emission filter. Reactions were modeled in Berkeley Madonna as described in Fig. B4 A-C and the rate constants for DNA bending are listed in Table 3-1.

Table 3-1. Kinetic parameters for base flipping and DNA bending by AAG variants[a]

AAG Variant	Lesion Context	k_{flip} (s^{-1}) [b]	k_{bend} (s^{-1}) [c]
WT	εA • T	4.4 ± 0.5	4.2 ± 0.7
Y127W	εA • T	1.7 ± 0.3	1.5 ± 0.1
Y159W	εA • T	27.4 ± 2.9	32.8 ±16.7
WT	εA • Spacer C3	2.2 ± 0.2	1.8 ± 0.04
WT	εA • Spacer 9	9.8 ± 0.4	12.9 ± 5.1

[a]Reactions were performed at 25°C in buffer containing 50 mM MES (pH 6.5), 100 mM NaCl, 1 mM EDTA, and 1 mM DTT. Each value represents the average of 3 measurements ± standard deviation.

[b]Measurements of k_{flip} were performed with 300-500 nM DNA and excess AAG. Reactions were fit in Berkeley Madonna as described in Fig. B5 A-B.

[c]Measurements of k_{bend} were performed with 100 nM DNA and 50-100 nM AAG. Reactions were fit in Berkeley Madonna as described in Fig. B4 A-C.

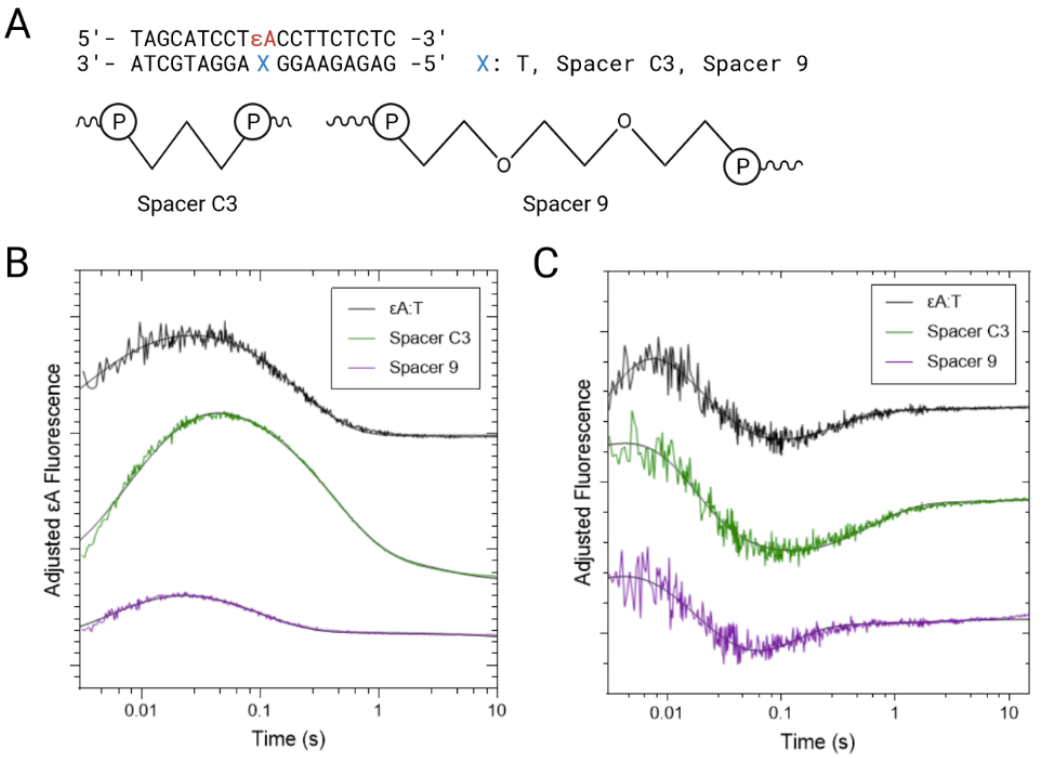


Figure 3-6. DNA linkers perturb kinetics of flipping and DNA bending by AAG. (A) Sequence of the 19mer DNA substrate used to measure the effect of linkers opposite the εA lesion. The structures of each spacer are shown. (B) Representative stopped-flow traces of εA fluorescence with varying opposing strands. Each reaction contained 300 nM εA DNA with the listed complement and a small excess of 350 nM AAG. Reactions were modeled in Berkeley Madonna as described in Fig. B5 A-B and the rate constants for base flipping are listed in Table 3-1. (C) Representative stopped-flow fluorescence traces of FRET DNA mixed with AAG. Reactions were performed with 100 nM FRET DNA and 80 nM AAG, with fluorescence measured using a 660 nm long-pass emission filter. Reactions were modeled in Berkeley Madonna as described in Fig. B4 A-C and the rate constants for DNA bending are listed in Table 3-1.

DNA spacers alter both base flipping and DNA bending by AAG

We reasoned that increased flexibility of the DNA could alter the kinetics of DNA bending. Therefore, we introduced either a propyl spacer (spacer C3) or a triethylene glycol spacer (spacer 9) in place of the opposing T (Figure 3-6A) and tested their impact upon both nucleotide flipping and DNA bending. The results are compared to the εA•T context in Figure 3-6B for εA-fluorescence and in Figure 3-6C for FRET bending signal.

Incorporation of Spacer C3 opposite of the lesion led to a 2-fold decrease in the rate constant for flipping relative to an opposing thymine (Fig. 3-6B, Table 3-1). Spacer 9 had the opposite effect, increasing the rate constant for flipping by over 2-fold (Fig. 3-6B, Table 3-1). Just as we observed with nucleotide flipping, Spacer C3 reduced the rate of bending by roughly 2-fold while Spacer 9 increased the rate of bending by 3-fold (Fig. 3-6C, Table 3-1). The consistent alignment of rate constants for nucleotide flipping and DNA bending agrees with the same measurements made with AAG variants. Although the Spacer C3 is expected to be more flexible than deoxyribose, its nominal spacing of the adjacent nucleotides is the same as found in DNA and it may cause similar resistance to the compression that must occur during DNA bending. In contrast, the triethylene glycol spacer is expected to be much more flexible which correlates with increased rate constant for both bending and flipping.

DISCUSSION

Timing of DNA bending

We report the development of a FRET assay for the kinetic characterization of DNA bending by human AAG. DNA bending has been widely speculated to play a role in genome searching by AAG, but evidence for bending by AAG has previously been limited to snapshots from X-ray crystal structures and atomic force microscopy studies. To elucidate the role of DNA bending we have placed DNA bending into the existing kinetic framework of the AAG mechanism (Fig. 3-7).

The timing of DNA bending relative to base flipping provides the clearest insight into the role of bending. We show that the initial recognition complex forms prior to DNA bending, providing strong evidence that bending is not necessary for lesion recognition. To address the temporal resolution between bending and flipping, we measured both rate constants in a variety of contexts. We were able to manipulate the rate constants for both bending and flipping by altering the DNA context to alter DNA flexibility and by altering the lesion recognition pocket by mutating active site tyrosine residues into larger tryptophan residues. The close agreement in rate constants for bending and flipping

across a range of proteins and substrates strongly suggests that bending does not precede flipping. We cannot distinguish between models where base flipping precedes and limits DNA bending and models where DNA bending and nucleotide flipping occur in tandem (Fig. 3-7). However, both a late DNA bending step and a concurrent bending/flip step preclude the possibility that DNA bending is used to search the genome for sites of damage. In both models the role of DNA bending is to accommodate the strain induced by nucleotide flipping.

Alternative strategies for searching DNA

Our results implicate DNA bending as a late step in the stabilization of the specific recognition complex. The finding that DNA bending does not contribute to the DNA searching mechanism of AAG raises the question of how AAG can efficiently identify diverse sites of damage that are mostly hidden within the DNA duplex. A tyrosine residue from the β3β4 hairpin of AAG intercalates into the minor groove of DNA and captures the lesion in the flipped-out conformation [2]. Side chain intercalation at the lesion site is a feature shared by all base flipping glycosylases, though the identity of the residue(s) involved varies from enzyme to enzyme. Mutation of the intercalating tyrosine162 of AAG

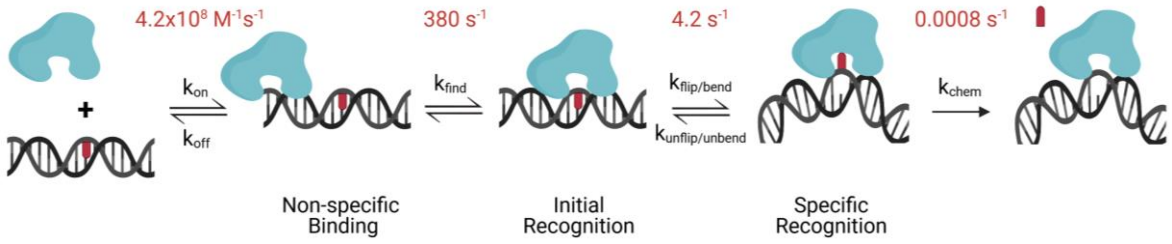


Figure 3-7. Minimal mechanism for recognition and excision of εA by AAG. AAG binds nonspecifically to DNA (k_{on}) and initiates a rapid search for compatible sites of damage. Once a lesion is identified (k_{find}), AAG pauses and flips the damaged base into the active site. The DNA backbone is bent simultaneous with flipping or shortly thereafter, represented here as the singular rate $k_{flip/unflip}$. Finally, the lesion is excised by cleaving the *N*-glycosidic bond, generating an abasic site (k_{chem}). Rate constants for the first 3 forward steps were determined by fitting FRET measurements in Berkeley Madonna to a series of effectively irreversible steps (Fig. B4). The rate constant for k_{chem} was determined in *Hendershot 2014* using a 25mer oligo under identical buffer conditions [2].

to tryptophan provides evidence that this side chain intercalates into the minor groove long before flipping of the lesion site, and the properties of the intercalating residue strongly govern the rate of DNA searching [22].

Taken together, these studies outline a cohesive model for lesion recognition. First, AAG binds to DNA nonspecifically. In our assays this initial binding involves association with the fluorescent dyes on the ends of our short DNA oligos. AAG then quickly initiates its search of the DNA and intercalates the β3β4 hairpin into the minor groove to probe for sites with unusual base structure or hydrogen-bonding properties. Upon identification of a site of damage, the enzyme pauses its search until the damaged base can be rotated from the duplex. This results in a small bend in the DNA backbone, and the Y162 residue fully plugs the gap left by the rotated base. This produces the specific recognition complex favorable for *N*-glycosidic bond cleavage.

This model provides multiple layers of selectivity for identifying sites of damage, as AAG must recognize instability in the minor groove via the intercalating β-hairpin, the lesion must be able to be flipped out of the duplex, the lesion must properly fit in the AAG active site, and the lesion must be compatible with the acid/base chemistry used for hydrolysis of the *N*-glycosidic bond. The flexibility of the DNA backbone at the lesion site could also contribute to specificity, though this warrants further investigation.

DNA bending plays different roles among glycosylases

The 4 primary families of glycosylase enzymes are thought to have evolved separately over time [29]. Despite separate origins, glycosylases from all 4 families share an overall mechanistic strategy involving base flipping, distortion of the DNA backbone, and intercalation of side chains into the minor groove of DNA. The convergence of these striking systems would seem to suggest that each step plays an identical, essential role in recognizing sites of DNA damage. This makes it even more surprising that DNA bending by AAG does not appear to play a role in the enzyme's strategy for DNA searching, in contrast with previous reports from other glycosylases.

Uracil DNA glycosylase (UDG) from *E. coli* excises uracil from a variety of DNA contexts and is among the best-studied enzymes in the uracil DNA glycosylase structural family. Kinetic characterization of UDG indicates that DNA bending precedes base

flipping during the recognition of sites of uracil damage [30]. Another glycosylase from the uracil glycosylase family, thymine DNA glycosylase (TDG), has also been proposed to bend DNA prior to flipping [31]. Outside of the uracil glycosylase family, the helix-hairpin-helix family protein 8-oxoguanine DNA glycosylase (OGG1) has been speculated to bend DNA prior to flipping based on AFM observations of OGG1 associated with bent DNA lacking any damage [17].

In human cells, AAG is the sole glycosylase with the eponymous AAG fold. As such, it remains to be seen whether the timing and function of DNA bending diverges significantly between the major glycosylase enzyme folds, or whether there are variations enzyme-to-enzyme. AAG is notable for recognizing an incredibly broad range of damaged substrates, in stark contrast to other glycosylases like UDG that cleave a specialized subset of DNA lesions. Further studies will be necessary to dissect trends in the DNA bending mechanisms of the numerous glycosylases present in human cells.

REFERENCES

1. 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.

2. Hendershot, J.M. and P.J. O'Brien, *Critical role of DNA intercalation in enzyme-catalyzed nucleotide flipping.* Nucleic Acids Res, 2014. **42**(20): p. 12681-90.

3. Lindahl, T., *Instability and decay of the primary structure of DNA.* Nature, 1993. **362**(6422): p. 709-15.

4. Robertson, A.B., et al., *DNA repair in mammalian cells: Base excision repair: the long and short of it.* Cell Mol Life Sci, 2009. **66**(6): p. 981-93.

5. Lindahl, T. and B. Nyberg, *Rate of depurination of native deoxyribonucleic acid.* Biochemistry, 1972. **11**(19): p. 3610-8.

6. Hollis, T., Y. Ichikawa, and T. Ellenberger, *DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, Escherichia coli AlkA.* EMBO J, 2000. **19**(4): p. 758-66.

7. Hollis, T., A. Lau, and T. Ellenberger, *Structural studies of human alkyladenine glycosylase and E. coli 3-methyladenine glycosylase.* Mutat Res, 2000. **460**(3-4): p. 201-10.

8. Parikh, S.S., et al., *Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA.* EMBO J, 1998. **17**(17): p. 5214-26.

9. Lingaraju, G.M., et al., *Structural basis for the inhibition of human alkyladenine DNA glycosylase (AAG) by 3,N4-ethenocytosine-containing DNA.* J Biol Chem, 2011. **286**(15): p. 13205-13.

10. Maiti, A., et al., *Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition.* Proc Natl Acad Sci U S A, 2008. **105**(26): p. 8890-5.

11. 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. 859-66.

12. Fromme, J.C. and G.L. Verdine, *Structure of a trapped endonuclease III-DNA covalent intermediate.* EMBO J, 2003. **22**(13): p. 3461-71.

13. Bangalore, D.M., et al., *Automated AFM analysis of DNA bending reveals initial lesion sensing strategies of DNA glycosylases.* Sci Rep, 2020. **10**(1): p. 15484.

14. Buechner, C.N., et al., *Lesion search and recognition by thymine DNA glycosylase revealed by single molecule imaging.* Nucleic Acids Res, 2015. **43**(5): p. 2716-29.

15. Friedman, J.I. and J.T. Stivers, *Detection of damaged DNA bases by DNA glycosylase enzymes.* Biochemistry, 2010. **49**(24): p. 4957-67.

16. Wang, H., et al., *DNA bending and unbending by MutS govern mismatch recognition and specificity.* Proc Natl Acad Sci U S A, 2003. **100**(25): p. 14822-7.

17. Chen, L., et al., *Direct visualization of a DNA glycosylase searching for damage.* Chem Biol, 2002. **9**(3): p. 345-50.

18. 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.

19. Hedglin, M. and P.J. O'Brien, *Human alkyladenine DNA glycosylase employs a processive search for DNA damage.* Biochemistry, 2008. **47**(44): p. 11434-45.

20. Hedglin, M. and P.J. O'Brien, *Hopping enables a DNA repair glycosylase to search both strands and bypass a bound protein.* ACS Chem Biol, 2010. **5**(4): p. 427-36.

21. Hedglin, M., Y. Zhang, and P.J. O'Brien, *Probing the DNA structural requirements for facilitated diffusion.* Biochemistry, 2015. **54**(2): p. 557-66.

22. Hendershot, J.M. and P.J. O'Brien, *Search for DNA damage by human alkyladenine DNA glycosylase involves early intercalation by an aromatic residue.* J Biol Chem, 2017. **292**(39): p. 16070-16080.

23. Parkhurst, K.M., M. Brenowitz, and L.J. Parkhurst, *Simultaneous binding and bending of promoter DNA by the TATA binding protein: real time kinetic measurements.* Biochemistry, 1996. **35**(23): p. 7459-65.

24. Jamieson, E.R., et al., *Structural and kinetic studies of a cisplatin-modified DNA icosamer binding to HMG1 domain B.* J Biol Chem, 1999. **274**(18): p. 12346-54.

25. Tang, G.Q. and S.S. Patel, *Rapid binding of T7 RNA polymerase is followed by simultaneous bending and opening of the promoter DNA.* Biochemistry, 2006. **45**(15): p. 4947-56.

26. 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. 12418-29.

27. Hendershot, J.M., A.E. Wolfe, and P.J. O'Brien, *Substitution of active site tyrosines with tryptophan alters the free energy for nucleotide flipping by human alkyladenine DNA glycosylase.* Biochemistry, 2011. **50**(11): p. 1864-74.

28. Blair, R.H., J.A. Goodrich, and J.F. Kugel, *Using FRET to monitor protein-induced DNA bending: the TBP-TATA complex as a model system.* Methods Mol Biol, 2013. **977**: p. 203-15.

29. O'Brien, P.J., *Catalytic promiscuity and the divergent evolution of DNA repair enzymes.* Chem Rev, 2006. **106**(2): p. 720-52.

30. 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.

31. Dodd, T., et al., *Uncovering universal rules governing the selectivity of the archetypal DNA glycosylase TDG.* Proc Natl Acad Sci U S A, 2018. **115**(23): p. 5974-5979.

CHAPTER 4

Human AAG Does Not Contribute to Alkylation Sensitivity in HAP1 Cells [1]

INTRODUCTION

Haploid cells are an attractive platform for any cellular work requiring genome editing. Gene deletions or modifications in diploid cells can often be masked by a second unaltered copy of a gene, complicating phenotypic analysis. The presence of a single copy of each gene in haploid cells dramatically simplifies the process of gene deletion and mutation relative. Yeast are a popular model for eukaryotic cell studies, as species like *Saccharomyces cerevisiae* allow for genome manipulation during the haploid stage of their life cycle [1, 2]. Human cells are almost exclusively diploid, which has historically made genome editing a challenge. Furthermore, many widely used cell lines are polyploid in all or some of their chromosomes, as polyploidy is a common feature of cancer cells [3]. Human haploid cells normally occur only during the gamete stage. However, some rare tumor cells have also been identified as haploid, opening the door for proliferation of haploid human cells [4, 5]. These haploid human cell lines have the potential to provide the convenience of haploid genome manipulation while preserving the context of human cell biology.

The HAP1 cell line is a commercially available near-haploid human cell line derived from the KBM-7 male chronic myeloid leukemia cell line [6]. HAP1 cells contain only 23 chromosomes, but a small duplication of chromosome 15 present on chromosome 19 prevents the line from being fully haploid [7]. Unlike its parental cell lines, HAP1 cells are adherent and have fibroblast-like morphology. While HAP1 cells are a convenient model for exploring the effects of gene knockouts, relatively little work has been done to

[1]Human cell culture work and alkylation sensitivity assays were performed in the Zhang lab with assistance and oversight from Yan Zhang and Zhonggang Hou.

characterize the unique biology of these irregular haploid cells. Genome maintenance and DNA repair are recognized as key features of human cell biology, however very few studies have explored these processes in HAP1 cells [8-10].

Alkyladenine DNA glycosylase (AAG) is one of roughly a dozen human glycosylases responsible for initiating the base excision DNA repair pathway (BER). AAG is notable as the only human glycosylase capable of recognizing lesions formed through DNA alkylation. AAG scans the genome for sites of damage and upon recognizing a lesion cleaves the *N*-glycosidic bond connecting the modified nucleobase to the DNA backbone, generating an abasic site. The abasic site can then be nicked by an AP endonuclease and a new nucleotide can be inserted and ligated to match the opposing strand [11].

The expression of AAG must be carefully controlled for effective protection against DNA alkylation. Both knockout of AAG and overexpression of AAG in mammalian cells have been shown to increase sensitivity to alkylating agents like methyl methanesulfonate (MMS) [12-15]. Knockout of AAG straightforwardly halts the repair of many alkylated lesions, leading to mutations and cell death. Conversely, overexpression of AAG can lead to the accumulation of dangerous repair intermediates such as abasic sites that can be toxic or mutagenic [16]. The former result has been reproduced in human cancer cells, with siRNA knockdown of AAG leading to increased MMS sensitivity in HeLa cells [17].

Alkylation is a common form of DNA damage. Compounds capable of simple alkylation reactions like methylation are present throughout nature as well as inside the human cell [18, 19]. For example, the ubiquitous methylating cofactor S-adenosylmethionine (SAM) is capable of methylating multiple DNA nucleobases at nitrogen and oxygen sites. To simulate these alkylating reactions in controlled studies, monofunctional alkylating agents such as MMS are often used. MMS is a broad-spectrum methylating agent that produces a large variety of methylated nucleobases. Among the major products in DNA are 7-methylguanine and 3-methyladenine, as well as other products including 7-methyladenine, 3-methylguanine, and several cytosine and thymine derivatives (Fig. 4-1) [20].

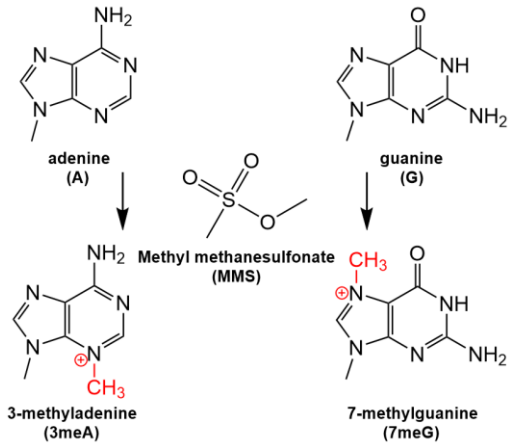


Figure 4-1. Alkylation damage generated by MMS. Chemical structure of MMS and two of its DNA methylation products. MMS methylates a wide range of adenine and guanine lesions, as well as non-DNA targets. Among these products, 3meA and 7meG are established targets of human AAG.

To assess the relevance of AAG in the alkylation sensitivity of HAP1 cells, we first established the activity of AAG in the HAP1 cell line. We then compared the MMS sensitivity of HAP1, HAP1 ΔAAG, and HeLa cells and found no significant difference in the alkylation sensitivity of HAP1 WT and HAP1 ΔAAG cells. This surprising finding indicates that AAG activity does not govern alkylation sensitivity in HAP1 like has been previously observed in HeLa cells [17]. Both HAP1 cell lines showed slightly increased sensitivity to MMS relative to HeLa, indicating that the haploid cells may be increasingly susceptible to loss of function mutation due to alkylation or may lack proper expression of protective factors. Finally, we reconstituted the remainder of the BER pathway in HAP1 cell extracts, confirming that the whole BER pathway is active in HAP1 cells. These results suggest that haploid cell lines like HAP1 may differ from other human cell lines in their response to alkylation damage.

MATERIALS AND METHODS

Cell culture

HAP1 WT and HAP1 AAG cells were purchased from Horizon Discovery (WT: C859, ΔAAG: HZGHC001537c002). All HAP1 cells were cultured in Iscove's Modified Dulbecco's Medium (IMDM; GIBCO) supplemented with 10% FBS (Corning). HeLa cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM; GIBCO) supplemented with 10% FBS (Corning). Both cell lines were cultured at 37°C and 5% CO_2 in a humidified incubator. Media was changed daily and cells were split every 3-4 days using TrypLE Express (GIBCO). HAP1 cells were passaged for no more than 2 weeks before usage, to minimize the documented shift to diploid cells [21].

Antibodies

Antibodies used for western blotting include polyclonal rabbit anti-MPG (1:1000, Sigma HPA006531), polyclonal anti-rabbit IgG-HRP (1:5000, Promega W4011), monoclonal mouse anti-GAPDH (1:1000, Santa Cruz 6C5), and polyclonal anti-mouse IgG-HRP (1:5000, Promega W4021).

Preparation of oligonucleotides

Undamaged 25mer was synthesized by Integrated DNA Technologies. All modified 25mer and 35mer oligos were synthesized by the W.M. Keck Facility at Yale University. Oligos were purified via denaturing PAGE. Bands were extracted and desalted using a C18 column (Sep-Pak). Oligo concentrations were determined using the theoretical extinction coefficient at 260 nm.

5'-FAM-ATGGA-s-GCTAGGATGCCTICCTTCTCTCCATTATCG-s-C3-3'
5'- C3-CGAT-s-AATGGAGAGAAGGAGGCATCCTAGCTCCAT-s-C3-3'
5'-FAM-CGATAGCATCCTXCCTTCTCTCCAT-3'
5'-ATGGAGAGAAGGTAGGATGCTATCG-3'
X: I, A
s: phosphorothioate link

Cell extract preparation

HAP1 and HeLa cell extracts were prepared via sonication. Cells were grown to near 100% confluence then detached from the plate with TrypLE Express, washed with PBS, and suspended in Lysis Buffer (50 mM NaMES pH 6.5, 250 mM potassium acetate, 1 mM EDTA, 1 mM EGTA, 1% NP-40). After 30 min incubation on ice, samples were sonicated twice using a Q800R3 Sonicator (QSonica) with settings: 75% amplitude, 30 seconds, 10 seconds ON, 10 seconds OFF. Cells were pelleted and further clarified using an UltraFree-MC Millipore centrifugal filtration tube for 10 min at 14,000xg. Protein concentrations for each extract were determined using BioRad Protein Assay, with concentrations ranging between 1-10 mg/mL.

Alkylation sensitivity assay

Cells were seeded to approximately 20,000 cells per well and incubated overnight. A fresh 1% solution of MMS was prepared in each cell's respective media then serially diluted for each desired MMS concentration. To initiate the assay, an appropriate amount of each MMS dilution was added to each well. Cells were incubated at 37°C and 5% CO_2 with MMS for 1 hour. Untreated cells were seeded and incubated in parallel. After incubation, each well was washed with PBS then returned to media with Pen/Strep antibiotic. Cells were incubated for 24 hours, then counted via hemocytometer. Each reaction was performed in triplicate and averaged. The average cell count for each cell type and MMS concentration was normalized to the average count for untreated cells. Each experiment was repeated starting with 3 separate cell seedings. Data shown represents the average and standard deviation of these 3 replicates.

Glycosylase activity assay

Reactions were performed at 37°C in Glycosylase Reaction Buffer designed for optimal AAG activity. The buffer contained 50 mM NaMES (pH 6.1), 0.1 mg/mL of BSA, 1 mM DTT, 1 mM EDTA, and 100 mM ionic strength adjusted with NaCl. Each reaction contained 10 nM FAM-labeled 25mer DNA and 1 mg/mL of cell extract. Aliquots were removed from each reaction periodically and quenched in an equal volume of 0.4 M

NaOH. The quenched samples were heated at 70°C for 12 min, then mixed 1:2 with loading buffer (90% formamide, 10 mM EDTA, 0.025% bromphenol blue, 0.025% xylene cyanol) and loaded on a 6.6 M urea, 20% polyacrylamide gel. Gels were scanned using an Amersham Biosciences Typhoon 5 Biomolecular Imager (GE Healthcare Life Sciences). The samples were excited at 488 nm with a 525 nm band-pass filter. Quantification was performed in ImageQuant TL (GE Healthcare). The fraction product (product / (substrate + product)) was determined for each sample and the initial rates of product formation were fit via linear regression.

BER reconstitution assay

Reactions were performed in BER buffer containing 50 mM NaHEPES (pH 7.5), 5 mM $MgCl_2$, 4 mM ATP, 25 μM dNTP's, and 100 mM NaCl. Each sample contained 10 nM FAM-labeled 35mer DNA with a bulged I damage site, 0.4 mg/mL cell extract, and 100 nM undamaged 25mer with a central A:T pair. Aliquots from each reaction were quenched and quantified as described above. For this assay, the fraction product gives the quantity of the non-bulged 34mer oligo as a fraction of the total fluorescence in each lane.

RESULTS AND DISCUSSION

To investigate the relevance of AAG to alkylation repair in HAP1 cells, we obtained HAP1 WT and HAP1 ΔAAG cells from Horizon Discovery. Western blots were performed to verify the presence of AAG in the HAP1 WT cells and absence in the HAP1 ΔAAG cells (Fig. 4-2). Protein was detected in both samples of HAP1 WT cells with molecular weight consistent with the 33 kD AAG. The same band was not present in HAP1 ΔAAG cells, suggesting AAG is properly expressed in HAP1 WT cells and knocked out successfully in the appropriate cell line. The varying intensities of the bands for AAG in each HAP1 WT cell sample could indicate some variability in AAG expression in HAP1 cells.

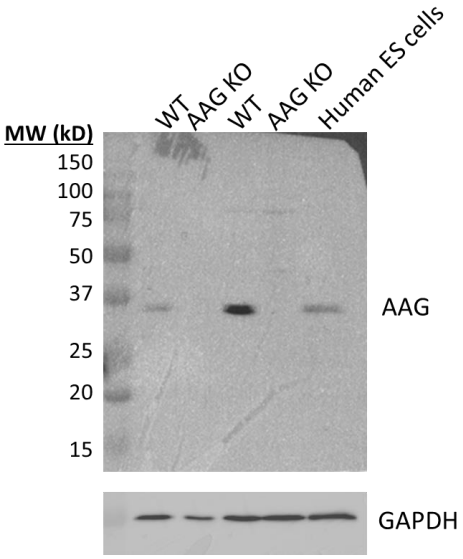


Figure 4-2. Verification of AAG expression in HAP1 cells. Western blot analysis of AAG and GAPDH in HAP1 WT, HAP1 ΔAAG, and ES cells. HAP1 cells were blotted in duplicate, with a single lane of human embryonic stem (ES) cells to confirm activity of the rabbit anti-AAG primary antibody. The expected molecular weight of AAG is 33 kD.

A cell adherence assay was developed to measure the alkylation sensitivity of HAP1 cells. HeLa and HAP1 cells are both adherent, so cells were plated at roughly 20,000 cells per well and were given 24 hours to adhere before alkylation exposure. Each well was incubated in MMS for 1 hour to generate alkylation damage. The MMS was then washed away, and cells were given 24 hours in media to recover from the damage. After 24 hours had elapsed, each well was washed to remove any cells released from the surface of the plate. The remaining adherent cells were thus the cells with alkylation repair or tolerance mechanisms capable of maintaining cellular function enough to adhere to the plate. Cell counts were normalized to the number of cells present in wells lacking any MMS treatment. Exposure to 1 mM MMS was sufficient to cause a >90% reduction in cell count, validating cell adhesion as a robust reporter for alkylation damage (Fig. 4-3).

Based on previous work with siRNA knockdown of AAG in HeLa cells, it was expected that AAG should play a substantial role in determining the alkylation sensitivity of HAP1 cells under strong alkylating conditions like those tested in the adherence assay [17]. MMS is a general methylating agent capable of producing numerous methylated lesions. However, the primary product of MMS, 7-methylguanine, is a well-characterized substrate of AAG [20]. In contrast with this expectation, we observed no substantial difference in the cell survival of HAP1 WT and HAP1 ΔAAG cells exposed to high concentrations of MMS (Fig. 4-3). This lack of a phenotype for AAG knockout could be explained by failure of alkylation repair factors in HAP1 WT cells, or by the presence of

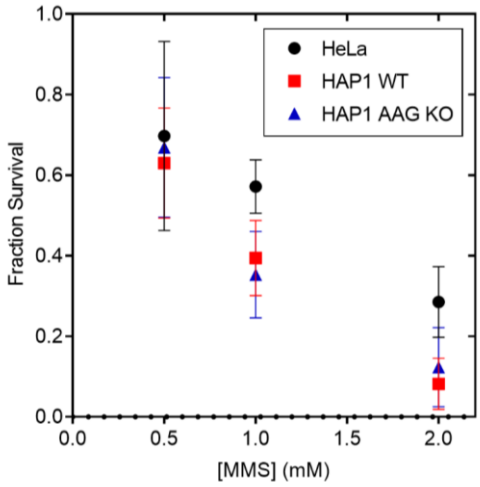


Figure 4-3. Loss of AAG does not sensitize HAP1 cells to alkylation damage. Alkylation sensitivity assay of HeLa, HAP1 WT, and HAP1 ΔAAG cells. Cells were incubated 1 hour with the indicated concentration of MMS, then given 24 hours to recover before washed for counting. Each point represents the average of 3 experiments ± S.D. and each experiment includes the average of 3 identically seeded wells.

unique alkylation repair and tolerance mechanisms in HAP1 cells that offset the loss of AAG activity. Both HAP1 cell lines appeared to be slightly sensitized to MMS damage relative to HeLa cells, suggesting that in either scenario HAP1 cells are unable to fully match the alkylation repair and tolerance of HeLa cells (Fig. 4-3).

To verify the activity of AAG, cell extracts were prepared from both HAP1 cell lines as well as HeLa cells. AAG is the sole glycosylase capable of removing hypoxanthine (Hx) in mammalian cells [22, 23]. As such, a standard gel-based glycosylase activity assay was performed to quantify excision of Hx from a fluorescently labeled 25mer oligo in each extract. Reaction conditions were selected to optimize AAG activity in the extract. Hypoxanthine excision activity was observed in both the HeLa and HAP1 WT extracts, but not the HAP1 ΔAAG extract (Fig. 4-4A). The initial Hx excision rates of 98 fmol/min·mg for HeLa and 110 fmol/min·mg for HAP1 WT cell extracts were very similar, suggesting

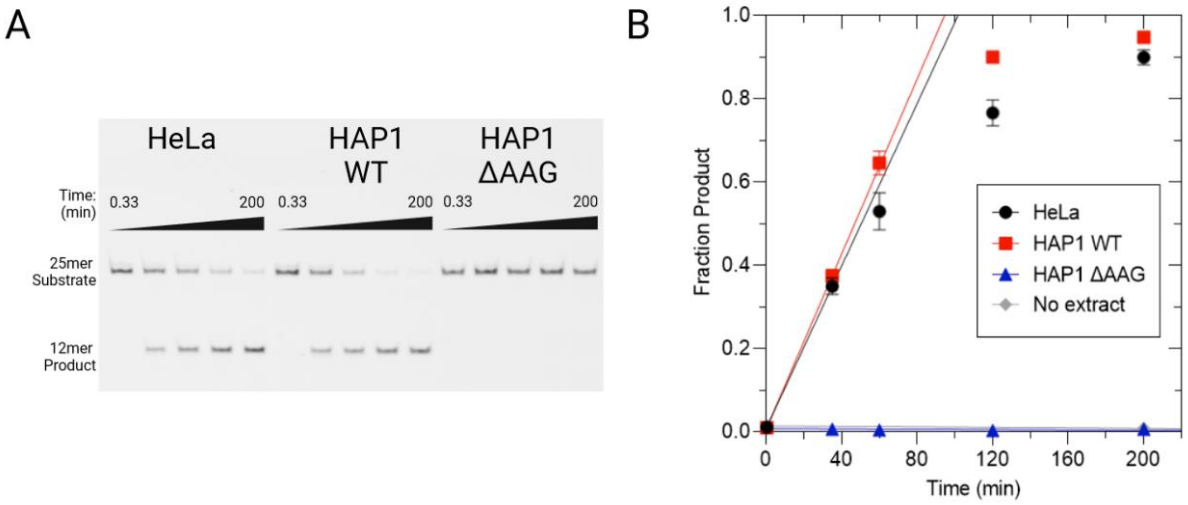


Figure 4-4. Verification of AAG activity in HAP1 cells. (A) Representative gel of an AAG activity assay in HeLa and HAP1 cell extracts. Reactions contained 10 nM 25mer DNA with a central I:T damage site and 1 mg/mL cell extract. (B) Quantified time course for the excision of hypoxanthine by AAG in cell extracts. Each point represents the average of 3 trials ± S.D. Initial rates were fit via linear regression, producing the observed velocities HeLa: v_{init} = 98 fmol/min·mg and HAP1 WT: v_{init} = 110 fmol/min·mg.

comparable levels of AAG activity in each cell line (Fig. 4-4B). These results indicate that any dysfunction in alkylation repair must be separate from the activity of AAG.

Following base excision by AAG, DNA lesions are normally processed by AP endonuclease APE1, polymerase β, and either DNA ligase 1 or DNA ligase 3 [11, 24]. Loss of activity at any of these steps could explain the absence of sensitivity to AAG knockout in HAP1 cells. To evaluate the presence and activity of the entire BER pathway in HAP1 cells, *in vitro* BER assays were performed with cell extracts. A 35mer bulged DNA substrate was chosen to allow for clear distinction between the 35mer starting oligo, a 17mer intermediate produced via cleavage by AP endonuclease, and the final repaired 34mer ligation product (Fig. 4-5A). The reaction buffer contained EDTA to reduce the activity of metal-dependent nucleases, and C3 spacer molecules were included on both the 5' and 3' end of each strand to reduce the activity of cellular exonucleases. Due to the absence of AAG in the HAP1 ΔAAG cell line, no intermediate or product was expected for that cell line, making it an effective baseline.

Both HeLa and HAP1 WT cell extracts showed clear formation of 17mer DNA intermediate, indicating successful cleavage of the abasic site by APE1 (Fig. 4-5B). In addition to this product, an 18 bp band was observed in all 3 extracts. This is likely a result of extension of the nicked 17mer by Polβ, leading to 1 bp strand displacement, as previously observed in HeLa cell extracts [25]. The formation of 34mer product was slightly obscured by the presence of a 5% background of 34mer independent of extract (Fig. 4-5C). However, both the HeLa and HAP1 WT extracts showed accumulation of additional 34mer product corresponding to completion of the BER pathway. While this assay does not rule out the possibility of differences in overall BER efficiency between

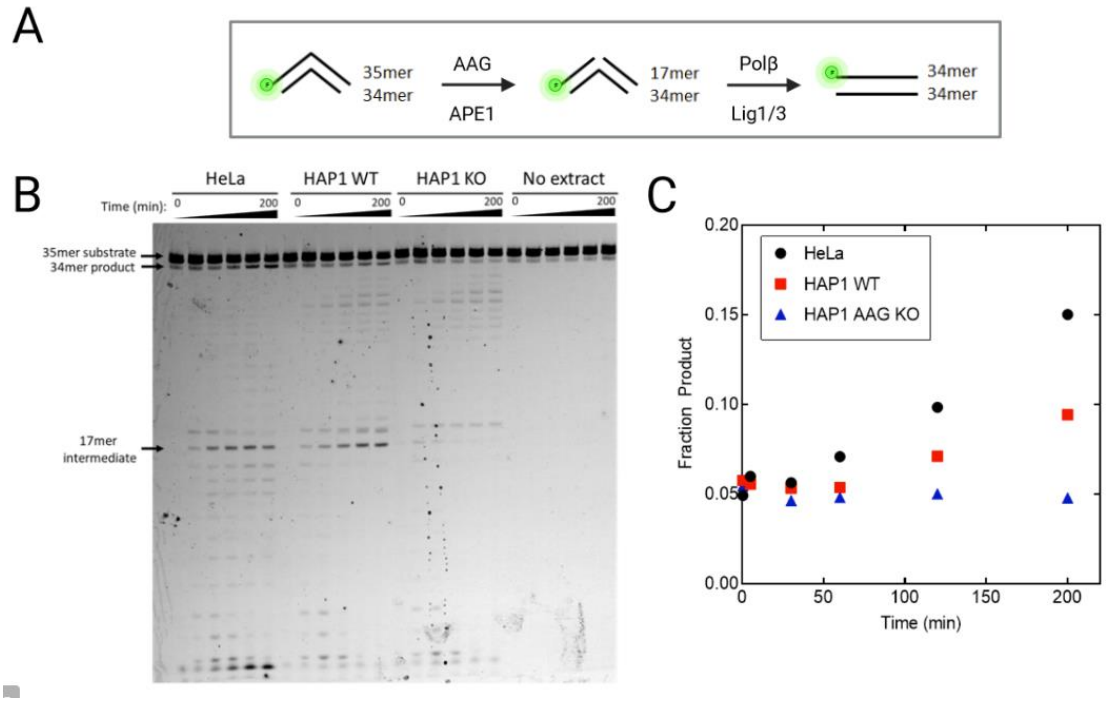


Figure 4-5. The BER pathway is active in HAP1 cells. (A) Diagram of BER reconstitution assay in human cell extract. Hypoxanthine is excised from the 35mer bulged substrate by any active AAG in the extract to produce an abasic site, which is cleaved by APE1. The 17mer nicked intermediate is extended by a polymerase as necessary to match the opposing strand, and a ligase seals the nick to produce a 34mer product with no bulge. (B) Example gel of BER reactions showing the formation of the 35mer substrate, 17mer intermediate, and 34mer product. Each reaction contained 10 nM substrate, 0.4 mg/mL extract, and 100 nM undamaged DNA to compete for nonspecific DNA binding factors. (C) Quantification of the gel in B. Both HeLa and HAP1 show the slow formation of 34mer product, indicating completion of the BER pathway.

HeLa and HAP1 cells, it provides clear evidence that the BER pathway is active in HAP1 cells.

Taken together, these results demonstrate that HAP1 cells differ from HeLa cells and other mammalian cell lines in their response to alkylation damage. While AAG plays a large role in counteracting alkylation damage in HeLa cells, no change in alkylation sensitivity is observed upon deletion of AAG in HAP1 cells. Given that both AAG and the downstream BER pathway enzymes are active in HAP1 cells, this lack of phenotype may arise due to altered efficiency of BER or dysfunction in other repair processes adjacent to BER. Both single and double strand break repair are points of potential impairment, as the activity of APE1 produces single-strand breaks in DNA at each site of DNA damage. There are also several cellular factors capable of forming DNA-protein-crosslinks with abasic sites in DNA, which could be altered to interfere with alkylation repair.

Regardless of the cause, the unique alkylation response of HAP1 cells highlights the risk of using haploid cell lines as model systems for studying the carefully regulated process of DNA repair. The HAP1 cell line appears to be a poor context for studying the function of AAG, and HAP1 may have other deficiencies associated with BER.

REFERENCES

1. Forsburg, S.L., *The art and design of genetic screens: yeast.* Nat Rev Genet, 2001. **2**(9): p. 659-68.

2. Nielsen, J., *Yeast Systems Biology: Model Organism and Cell Factory.* Biotechnol J, 2019. **14**(9): p. e1800421.

3. Davoli, T. and T. de Lange, *The causes and consequences of polyploidy in normal development and cancer.* Annu Rev Cell Dev Biol, 2011. **27**: p. 585-610.

4. Andersson, B.S., et al., *KBM-7, a human myeloid leukemia cell line with double Philadelphia chromosomes lacking normal c-ABL and BCR transcripts.* Leukemia, 1995. **9**(12): p. 2100-8.

5. Sandberg, A.A., N. Wake, and S. Kohno, *Chromosomes and causation of human cancer and leukemia. XLVII. severe hypodiploidy and chromosome conglomerations in ALL.* Cancer Genet Cytogenet, 1982. **5**(4): p. 293-307.

6. Carette, J.E., et al., *Ebola virus entry requires the cholesterol transporter Niemann-Pick C1.* Nature, 2011. **477**(7364): p. 340-3.

7. Essletzbichler, P., et al., *Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line.* Genome Res, 2014. **24**(12): p. 2059-65.

8. Xing, M. and V. Oksenych, *Genetic interaction between DNA repair factors PAXX, XLF, XRCC4 and DNA-PKcs in human cells.* FEBS Open Bio, 2019. **9**(7): p. 1315-1326.

9. Cui, J., A. Gizzi, and J.T. Stivers, *Deoxyuridine in DNA has an inhibitory and promutagenic effect on RNA transcription by diverse RNA polymerases.* Nucleic Acids Res, 2019. **47**(8): p. 4153-4168.

10. Esadze, A., et al., *Measurement of nanoscale DNA translocation by uracil DNA glycosylase in human cells.* Nucleic Acids Res, 2017. **45**(21): p. 12413-12424.

11. Robertson, A.B., et al., *DNA repair in mammalian cells: Base excision repair: the long and short of it.* Cell Mol Life Sci, 2009. **66**(6): p. 981-93.

12. Elder, R.H., et al., *Alkylpurine-DNA-N-glycosylase knockout mice show increased susceptibility to induction of mutations by methyl methanesulfonate.* Mol Cell Biol, 1998. **18**(10): p. 5828-37.

13. Engelward, B.P., et al., *Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase.* Proc Natl Acad Sci U S A, 1997. **94**(24): p. 13087-92.

14. Rinne, M.L., et al., *N-methylpurine DNA glycosylase overexpression increases alkylation sensitivity by rapidly removing non-toxic 7-methylguanine adducts.* Nucleic Acids Res, 2005. **33**(9): p. 2859-67.

15. Rinne, M., D. Caldwell, and M.R. Kelley, *Transient adenoviral N-methylpurine DNA glycosylase overexpression imparts chemotherapeutic sensitivity to human breast cancer cells.* Mol Cancer Ther, 2004. **3**(8): p. 955-67.

16. Loeb, L.A. and B.D. Preston, *Mutagenesis by apurinic/apyrimidinic sites.* Annu Rev Genet, 1986. **20**: p. 201-30.

17. Paik, J., et al., *Sensitization of human carcinoma cells to alkylating agents by small interfering RNA suppression of 3-alkyladenine-DNA glycosylase.* Cancer Res, 2005. **65**(22): p. 10472-7.

18. Lindahl, T., *Instability and decay of the primary structure of DNA.* Nature, 1993. **362**(6422): p. 709-15.

19. Sedgwick, B., *Repairing DNA-methylation damage.* Nat Rev Mol Cell Biol, 2004. **5**(2): p. 148-57.

20. Beranek, D.T., *Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents.* Mutat Res, 1990. **231**(1): p. 11-30.

21. Olbrich, T., et al., *A p53-dependent response limits the viability of mammalian haploid cells.* Proc Natl Acad Sci U S A, 2017. **114**(35): p. 9367-9372.

22. Hang, B., et al., *Targeted deletion of alkylpurine-DNA-N-glycosylase in mice eliminates repair of 1,N6-ethenoadenine and hypoxanthine but not of 3,N4-ethenocytosine or 8-oxoguanine.* Proc Natl Acad Sci U S A, 1997. **94**(24): p. 12869-74.

23. Smith, S.A. and B.P. Engelward, *In vivo repair of methylation damage in Aag 3-methyladenine DNA glycosylase null mouse cells.* Nucleic Acids Res, 2000. **28**(17): p. 3294-300.

24. Krokan, H.E. and M. Bjoras, *Base excision repair.* Cold Spring Harb Perspect Biol, 2013. **5**(4): p. a012583.

25. Lyons, D.M. and P.J. O'Brien, *Human base excision repair creates a bias toward -1 frameshift mutations.* J Biol Chem, 2010. **285**(33): p. 25203-12.

CHAPTER 5

Conclusions and Future Directions

Multi-layered damage recognition

In this thesis we have expanded our knowledge of the mechanisms used by the monomeric DNA repair protein human alkyladenine DNA glycosylase (AAG) to recognize diverse sites of damage in the genome. As the sole alkylation repair glycosylase in humans, AAG must recognize a broad spectrum of lesions while maintaining strong discrimination against undamaged bases [1]. These lesions include bases methylated at multiple sites, etheno adducts, and the deaminated base hypoxanthine. Not only do these bases vary in structure, but they also vary in charge. The remarkable specificity of AAG is accomplished through a series of selectivity filters each designed to accommodate or discriminate against certain nucleobase structures.

A great deal is known about the AAG mechanism for lesion recognition and catalysis (Fig. 5-1) [2-6]. In the established minimal mechanism, AAG binds nonspecifically to DNA, initiates a rapid search for sites of damage, binds at the lesion site in an initial recognition complex, flips the target base out of the duplex into the active site, and then cleaves the *N*-glycosidic bond. This mechanism presents several opportunities for discrimination between different bases, including the initial recognition, base flipping, active site accommodation, and alignment for chemistry. Our work has provided insight into several of these layers of specificity, furthering our understanding of how AAG is able to identify sites of damage and exclude undamaged bases.

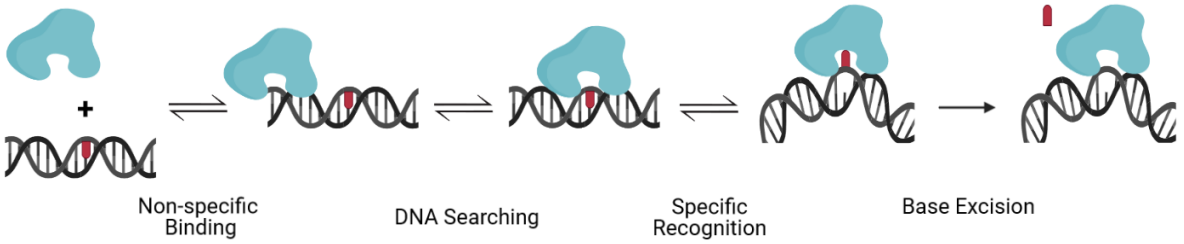


Fig. 5-1. Minimal mechanism of AAG. Overview of the AAG mechanism for recognizing and removing sites of DNA damage. AAG (cyan) binds nonspecifically at any available site of DNA and systematically scans DNA for sites of damage (red) by probing the minor groove using the β3β4 hairpin. Upon identifying a lesion, AAG forms an initial recognition complex with the lesion partially unstacked. The lesion is then rotated fully out of the duplex and into the AAG active site. At approximately the same time as base flipping, the backbone of DNA is bent. Finally, hydrolysis of the *N*-glycosidic bond can occur, releasing the damaged nucleobase and producing an abasic site. Illustrated with BioRender.com.

Initial recognition

DNA bending has long been speculated to play a role in DNA damage recognition by glycosylases, including AAG. Bending of the backbone of DNA presents an elegant and intuitive means to identify sites of damage from within the duplex based on the loss of duplex stability. However, we examined the kinetics of DNA bending using stopped flow FRET and revealed that DNA bending occurs later in the AAG mechanism, after the site of damage has already been bound in the initial recognition complex (Chapter 3). Many glycosylases share a bent specific recognition complex with damaged DNA [7-11]. But while the bent intermediates may look similar, it remains possible that different glycosylases may utilize bending for distinct purposes. DNA bending has already been implicated as a mechanism for searching in uracil DNA glycosylase, the best-characterized member of the UDG structural family [12]. As such, initial searching strategies may vary among glycosylases based on enzyme fold or target lesion.

For AAG, DNA searching appears to primarily involve the intercalation of the β3β4 hairpin into the minor groove of DNA, rather than DNA bending. Replacement of hairpin residue Tyr162 with tryptophan and examination of tryptophan fluorescence indicates this

hairpin intercalates into the minor groove of DNA early in the searching process [6]. This positions the hairpin ideally for assessment of base pairing within the duplex without the need to flip each base, though the exact interactions between AAG and the damaged site remain open for further exploration.

Base flipping and DNA bending

The next layer of specificity comes from the flipping of the damaged base out of the duplex and into the enzyme active site. Crystal structures of AAG show the enzyme bound to a flipped-out lesion while the backbone of the surrounding DNA is bent by roughly 20°. In our stopped flow FRET studies, we determined that flipping of the damaged base and bending of the DNA backbone occur on the same timescale (Chapter 3). Furthermore, modification of the AAG active site and alteration of the flexibility of the backbone of DNA do not decouple these two steps, suggesting base flipping and DNA bending are tightly coordinated (Chapter 3). Because the damaged base must be flipped into the active site for excision, this provides another opportunity for the enzyme to discriminate between nucleobase targets. It is not well understood how AAG selects which bases to flip, but factors like hydrogen bonding capacity and the flexibility of the DNA backbone may contribute to AAG's ability to quickly flip and bend the DNA for specific lesion recognition.

Active site discrimination

The AAG active site is carefully constructed to accommodate a broad spectrum of purine lesions while rejecting undamaged bases. According to crystal structures, the active site residue Asn169 sits close to the C2 position of a flipped-out purine nucleobase. Consequently, mutation of this residue has been shown to loosen discrimination against the binding of undamaged G through removal of the steric hindrance against the bulky N2 group of G [13]. However, undamaged G is not the only nucleobase to contain a large group at the N2 position that could clash with N169. The lesion 1,N^2-ethenoguanine (εG)

has been proposed as a primary substrate of AAG despite its structural similarities to undamaged G [14, 15].

To evaluate εG as a substrate, we characterized both the single and multiple turnover kinetics of εG excision by AAG (Chapter 2). We found that εG is excised with much lower catalytic efficiency than other primary substrates of AAG such as 1,N^6-ethenoadenine (εA). Furthermore, excess undamaged sites of DNA were found to strongly inhibit εG excision. As with undamaged G, mutation of Asn169 to Ser resulted in a large increase in the single turnover rate constant for the excision of εG, establishing Asn169 as a major player in discrimination against a variety of lesions containing a bulky N2 adduct. This strategy of providing a relatively open active site, but introducing steric hinderance to binding of undamaged purines, allows for a broad substrate range. However, it prevents the enzyme from effectively repairing some important alkyl lesions such as εG. Therefore, multiple DNA repair pathways are needed to cover the diverse spectrum of adducts that can be formed in DNA. These pathways include direct repair pathways. Humans have two oxidative dealkylases known as ALKBH2 and ALKBH3, however it is not clear how these and other repair pathways cooperate with AAG to provide protection against DNA alkylation.

Alkylation response of HAP1 cells

In addition to exploring the different steps involved in DNA damage recognition, we also investigated alkylation repair in the increasingly popular near-haploid HAP1 human cancer cell line [16]. Multiple recent studies have attempted to examine the reality of DNA searching by glycosylases in the context of a packed eukaryotic nucleus [17, 18]. As such, the HAP1 cell line has become an appealing system for exploring DNA repair. Unfortunately, DNA repair is a tightly controlled process and very little work has gone into characterizing the integrity of these DNA repair systems in such an unstable cellular environment. We discovered that AAG does not contribute to cellular survival of severe alkylation damage from the general methylating agent methyl methanesulfonate (MMS). This phenotype directly contradicts previous observations in the HeLa human cancer cell line as well as other mammalian cell contexts [19-21]. Therefore, we conclude that the

HAP1 cell line differs from other human cell lines in its repair and tolerance of alkylation damage. Further study of HAP1 may provide insight into currently unrecognized alkylation repair and tolerance pathways. Conversely, future cellular studies of human AAG should employ cell lines that are more representative.

FUTURE DIRECTIONS

Contrasting DNA bending by different glycosylases

In this thesis we have characterized DNA bending by AAG. Contrary to popular models, DNA bending by AAG does not occur during genome searching but instead accompanies base flipping in the formation of the specific recognition complex between AAG and its target lesion. To assess the role of DNA bending, we developed a powerful stopped flow FRET assay capable of measuring rate constants associated with DNA bending in a variety of contexts (Chapter 3).

While AAG does not appear to bend DNA during the searching process, it remains possible that other glycosylases could utilize DNA bending to identify sites of weakness in the duplex corresponding to damaged bases. The stopped flow FRET assay developed to measure DNA bending by AAG could be modified to assess the DNA bending rates of other glycosylases in order to determine the timing of DNA bending for each enzyme. The Cy3 and Cy5 fluorophores on the 5' ends of the FRET DNA allow for robust measurement of DNA bending regardless of the identity of the bound protein. However, the central lesion poses an important challenge for mechanistic analysis of other glycosylases. The εA lesion used in our studies is both strongly recognized by AAG and naturally fluorescent for measurement of base flipping [1]. Identifying lesions that match those characteristics for each glycosylase would be extremely difficult.

The *E. coli* protein 3-methyladenine DNA glycosylase II (AlkA) represents an ideal candidate for ongoing measurement of DNA bending by a glycosylase protein. AlkA is one of 2 glycosylases in *E. coli* tasked with removing sites of alkylation damage [22]. As such, AlkA shares many target lesions with AAG, including the critical εA lesion [23]. The

overall mechanistic strategy of the enzymes is also shared with AAG [24]. AlkA binds nonspecifically to DNA, scans until a site of damage is located, flips the lesion out of the duplex, and hydrolyzes the *N*-glycosidic bond to release the lesion [25]. AlkA has also been crystallized in complex with bent DNA, inducing an even sharper bend than that observed with AAG (Fig. 5-2) [26]. Most importantly, AlkA is a member of the helix-hairpin-helix family of glycosylases, which have a unique structural fold and are thought to have evolved separately from AAG and other glycosylase families [27]. Clarification of the role of DNA bending in the AlkA mechanism would provide valuable insight into the strategies shared among glycosylases and the importance of structural fold and target lesions in governing those mechanisms.

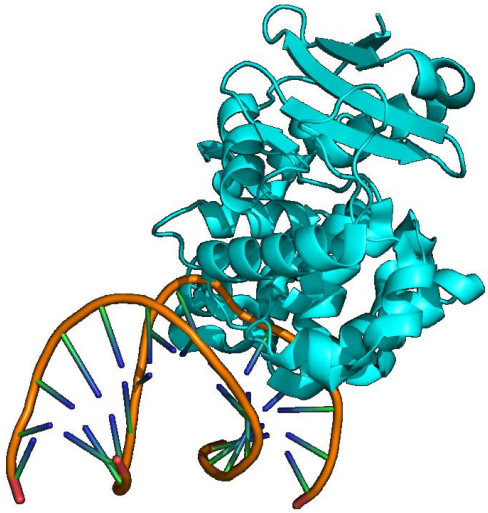


Fig. 5-2. Crystal structure of DNA bending by AlkA. The *E. coli* glycosylase AlkA (cyan) is shown bound to DNA containing a 1-azaribose abasic site. The DNA backbone is bent by 66° at the site of AlkA binding, one of the largest DNA bend angles observed in a glycosylase crystal structure. Image produced in PyMOL using coordinates from PDB structure 1DIZ.

Specific recognition of diverse lesions by AAG

The kinetic mechanism of AAG has largely been characterized using the intrinsic fluorescence of the εA lesion [3]. εA is ideal for kinetic analysis, as its intrinsic fluorescence near 400 nm is sensitive to its environment and fluctuates during several

key steps of the AAG mechanism, including the formation of both the initial and specific recognition complexes [3]. However, this focus on εA ignores potential differences in the recognition of diverse target lesions. It also prohibits examination of the impact of different lesion features on the rates of key steps like base flipping. Unfortunately, designing a system to report on lesion recognition and base flipping without the intrinsic fluorescence of the lesion itself has proven very challenging. Previous attempts to develop such an assay using tryptophan reporters have either failed to detect base flipping (Y162W variant) [6] or directly perturbed the kinetics of base flipping (Y127W, Y159W) [4].

The stopped flow FRET assay described in this thesis provides a new avenue for exploring the excision of non-fluorescent lesions. During the characterization of DNA bending, we established that the Cy3 and Cy5 fluorophores report on several steps of the AAG mechanism outside of bending alone (Chapter 3). Fluorescence of the Cy3 donor fluorophore increases significantly upon binding of AAG, corresponding to initial interaction between the fluorophore and the nonspecifically bound glycosylase near the end of the oligo. This initial binding signal then declines as AAG identifies its target for excision in the middle of the oligo and departs from the DNA ends. At that point, a final FRET increase occurs, representing both base flipping and DNA bending. This system has the potential to report on almost every step of the AAG mechanism, allowing deeper characterization of key AAG substrates such as hypoxanthine, 3-methyladenine, and 7-methylguanine.

DNA alkylation repair by HAP1 cells

Our examination of alkylation repair in HAP1 cells has focused on biochemical assays of base excision repair in cell extracts and cell survival assays with the alkylating agent MMS (Chapter 4). These techniques have provided insight into the overall alkylation repair phenotype of HAP1 cells. However, the exact role of AAG in the alkylation repair of HAP1 cells remains unclear. The observation that AAG knock-out cell lines show no change in MMS sensitivity could indicate that HAP1 cells have developed alternative means for repairing or tolerating alkylation damage. Alternatively, this result could be interpreted to mean the opposite: that both HAP1 WT and HAP1 ΔAAG cells have

compromised alkylation repair due to factors outside of AAG activity. To gain a better understanding of alkylation repair in HAP1 cells, more direct measurements of DNA damage accumulation and repair are necessary. Several strategies are available for measurement of DNA damage in human cells.

Genomic DNA can be fragmented in a damage-dependent manner to quantify the total amount of DNA damage within cells. After treatment of cells with MMS, extraction of DNA from the cells and treatment with AAG and APE1 *in vitro* allows for the generation of single-strand breaks at sites of AAG-compatible damage. Separation of genomic fragments via alkaline agarose gel electrophoresis has the potential to provide time-resolved quantification of DNA alkylation damage in HAP1 cells [28]. This technique has previously been used to estimate AAG-compatible sites of damage in yeast cells treated with MMS [17]. If HAP1 cells have developed AAG-independent mechanisms for repairing MMS alkylation damage, then genomic fragments would be expected to lengthen over time as cells reduce the number lesions available for conversion into single strand breaks by AAG and APE1. If HAP1 cells are deficient in alkylation repair, then the short DNA fragments formed through alkylation damage should persist longer than in other human cell lines.

The activity of DNA double-strand break repair is also commonly used to report on the DNA damage status of human cells. The phosphorylation of histone H2AX to form γH2AX is a well-characterized response to double-strand break damage [29]. The phosphorylation status of H2AX residue Ser139 can be monitored via immunoblotting, immunofluorescent microscopy, or fluorescence activated cell sorting for robust quantification of DNA damage [30]. MMS produces a time-dependent increase in γH2AX foci, making γH2AX an appealing reporter for repair of MMS damage in HAP1 cells [31, 32]. Other potential reporters sensitive to MMS damage include Rad51/Rad52 and p53 binding protein 1 (53BP1) [33, 34].

In closing, we have gained valuable insights into the mechanisms used by AAG to identify sites of DNA damage and the biochemical tools available for the study of human DNA alkylation repair. While AAG is only one of nearly a dozen known human glycosylases, its unique structural fold and impressively broad substrate range clearly

necessitate novel and finely tuned strategies for identifying target lesions. As our understanding of these mechanistic features continues to grow, this knowledge will inevitably inform our understanding not just of other glycosylases, but other DNA repair proteins.

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. 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. 12418-29.

3. 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.

4. Hendershot, J.M., A.E. Wolfe, and P.J. O'Brien, *Substitution of active site tyrosines with tryptophan alters the free energy for nucleotide flipping by human alkyladenine DNA glycosylase.* Biochemistry, 2011. **50**(11): p. 1864-74.

5. Hendershot, J.M. and P.J. O'Brien, *Critical role of DNA intercalation in enzyme-catalyzed nucleotide flipping.* Nucleic Acids Res, 2014. **42**(20): p. 12681-90.

6. Hendershot, J.M. and P.J. O'Brien, *Search for DNA damage by human alkyladenine DNA glycosylase involves early intercalation by an aromatic residue.* J Biol Chem, 2017. **292**(39): p. 16070-16080.

7. Maiti, A., et al., *Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition.* Proc Natl Acad Sci U S A, 2008. **105**(26): p. 8890-5.

8. Hollis, T., Y. Ichikawa, and T. Ellenberger, *DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, Escherichia coli AlkA.* EMBO J, 2000. **19**(4): p. 758-66.

9. Parikh, S.S., et al., *Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA.* EMBO J, 1998. **17**(17): p. 5214-26.

10. 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. 859-66.

11. Fromme, J.C. and G.L. Verdine, *Structure of a trapped endonuclease III-DNA covalent intermediate.* EMBO J, 2003. **22**(13): p. 3461-71.

12. 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.

13. Connor, E.E. and M.D. Wyatt, *Active-site clashes prevent the human 3-methyladenine DNA glycosylase from improperly removing bases.* Chem Biol, 2002. **9**(9): p. 1033-41.

14. Saparbaev, M., et al., *1,N(2)-ethenoguanine, a mutagenic DNA adduct, is a primary substrate of Escherichia coli mismatch-specific uracil-DNA glycosylase and human alkylpurine-DNA-N-glycosylase.* J Biol Chem, 2002. **277**(30): p. 26987-93.

15. 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.

16. Carette, J.E., et al., *Ebola virus entry requires the cholesterol transporter Niemann-Pick C1.* Nature, 2011. **477**(7364): p. 340-3.

17. Zhang, Y. and P.J. O'Brien, *Repair of Alkylation Damage in Eukaryotic Chromatin Depends on Searching Ability of Alkyladenine DNA Glycosylase.* ACS Chem Biol, 2015. **10**(11): p. 2606-15.

18. Esadze, A., et al., *Measurement of nanoscale DNA translocation by uracil DNA glycosylase in human cells.* Nucleic Acids Res, 2017. **45**(21): p. 12413-12424.

19. Paik, J., et al., *Sensitization of human carcinoma cells to alkylating agents by small interfering RNA suppression of 3-alkyladenine-DNA glycosylase.* Cancer Res, 2005. **65**(22): p. 10472-7.

20. Elder, R.H., et al., *Alkylpurine-DNA-N-glycosylase knockout mice show increased susceptibility to induction of mutations by methyl methanesulfonate.* Mol Cell Biol, 1998. **18**(10): p. 5828-37.

21. Engelward, B.P., et al., *Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase.* Proc Natl Acad Sci U S A, 1997. **94**(24): p. 13087-92.

22. Thomas, L., C.H. Yang, and D.A. Goldthwait, *Two DNA glycosylases in Escherichia coli which release primarily 3-methyladenine.* Biochemistry, 1982. **21**(6): p. 1162-9.

23. Saparbaev, M., K. Kleibl, and J. Laval, *Escherichia coli, Saccharomyces cerevisiae, rat and human 3-methyladenine DNA glycosylases repair 1,N6-ethenoadenine when present in DNA.* Nucleic Acids Res, 1995. **23**(18): p. 3750-5.

24. O'Brien, P.J. and T. Ellenberger, *The Escherichia coli 3-methyladenine DNA glycosylase AlkA has a remarkably versatile active site.* J Biol Chem, 2004. **279**(26): p. 26876-84.

25. Taylor, E.L. and P.J. O'Brien, *Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by AlkA.* Biochemistry, 2015. **54**(3): p. 898-908.

26. Hollis, T., A. Lau, and T. Ellenberger, *Structural studies of human alkyladenine glycosylase and E. coli 3-methyladenine glycosylase.* Mutat Res, 2000. **460**(3-4): p. 201-10.

27. O'Brien, P.J., *Catalytic promiscuity and the divergent evolution of DNA repair enzymes.* Chem Rev, 2006. **106**(2): p. 720-52.

28. Sutherland, B.M., P.V. Bennett, and J.C. Sutherland, *DNA damage quantitation by alkaline gel electrophoresis.* Methods Mol Biol, 1999. **113**: p. 183-202.

29. Mah, L.J., A. El-Osta, and T.C. Karagiannis, *gammaH2AX: a sensitive molecular marker of DNA damage and repair.* Leukemia, 2010. **24**(4): p. 679-86.

30. Ivashkevich, A., et al., *Use of the gamma-H2AX assay to monitor DNA damage and repair in translational cancer research.* Cancer Lett, 2012. **327**(1-2): p. 123-33.

31. Staszewski, O., T. Nikolova, and B. Kaina, *Kinetics of gamma-H2AX focus formation upon treatment of cells with UV light and alkylating agents.* Environ Mol Mutagen, 2008. **49**(9): p. 734-40.

32. Zhou, C., et al., *DNA damage evaluated by gammaH2AX foci formation by a selective group of chemical/physical stressors.* Mutat Res, 2006. **604**(1-2): p. 8-18.

33. Liu, Y. and N. Maizels, *Coordinated response of mammalian Rad51 and Rad52 to DNA damage.* EMBO Rep, 2000. **1**(1): p. 85-90.

34. Rappold, I., et al., *Tumor suppressor p53 binding protein 1 (53BP1) is involved in DNA damage-signaling pathways.* J Cell Biol, 2001. **153**(3): p. 613-20.

APPENDIX A

Additional Data Figures to Support Chapter 2

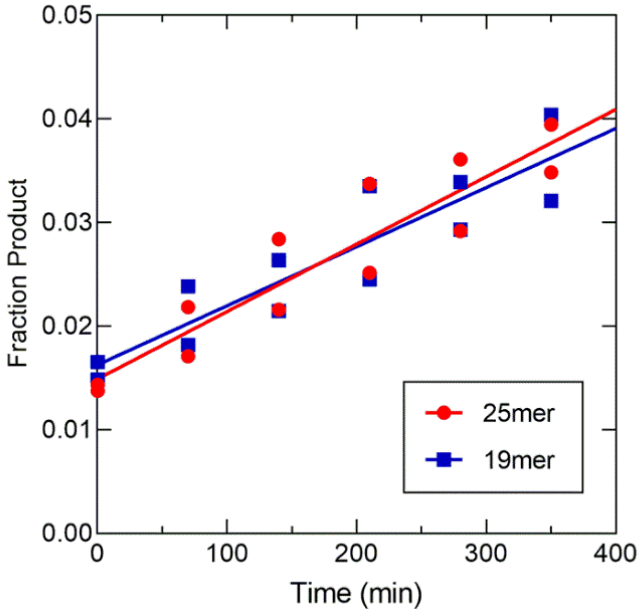


Figure A1. Impact of DNA length on the excision rate of εA. Direct competition between εA in 25mer and 19mer sequences for excision by Δ80 AAG. Reactions were performed under multiple turnover conditions, with 1 μM 25mer DNA, 1 μM 19mer DNA, and 10 nM AAG. A 1.1-fold preference for 25mer over 19mer was observed. Two replicates are pictured for each DNA length.

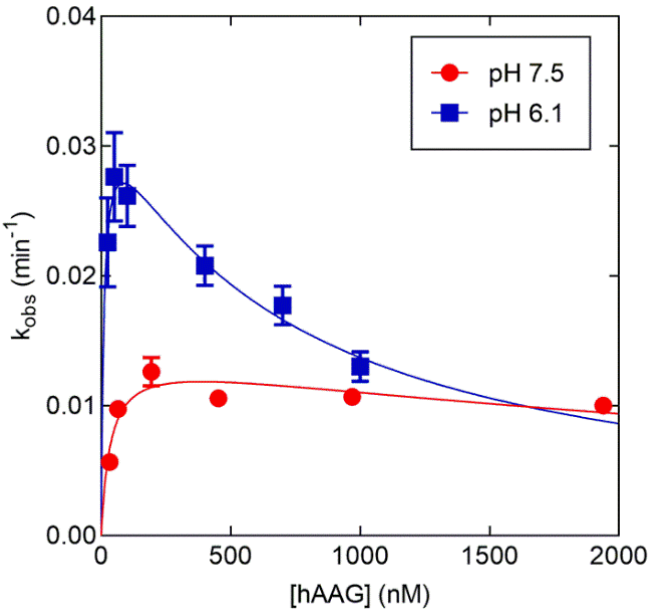


Figure A2. Effect of pH on the AAG-catalyzed excision of εG. Single turnover reactions of 20 nM 25mer εG•T duplex with increasing concentrations of Δ80 AAG were measured at pH 7.5 and 6.1. The data at pH 6.1 are replotted from Figure 2-4C for comparison and represent the average of 6 replicates, while points at pH 7.5 represent the average of 2 replicates, and the error bars reflect the S.D.

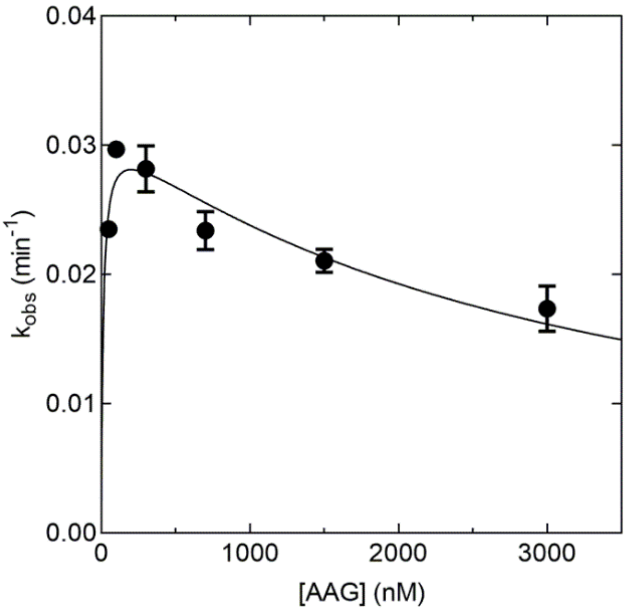


Figure A3. AAG-catalyzed excision of εG with a T complement. Enzyme concentration dependence of the single turnover rate of AAG-catalyzed excision of εG from an εG•T mismatch. Reactions contained 20 nM of the 25mer εG•T duplex and varying concentrations of Δ80 AAG. The best fit values are k_{max} = 0.032 min^{-1}, $K_{1/2}$ = 13 nM, and K_i = 3.1 μM. Each point represents the average of 2 replicates.

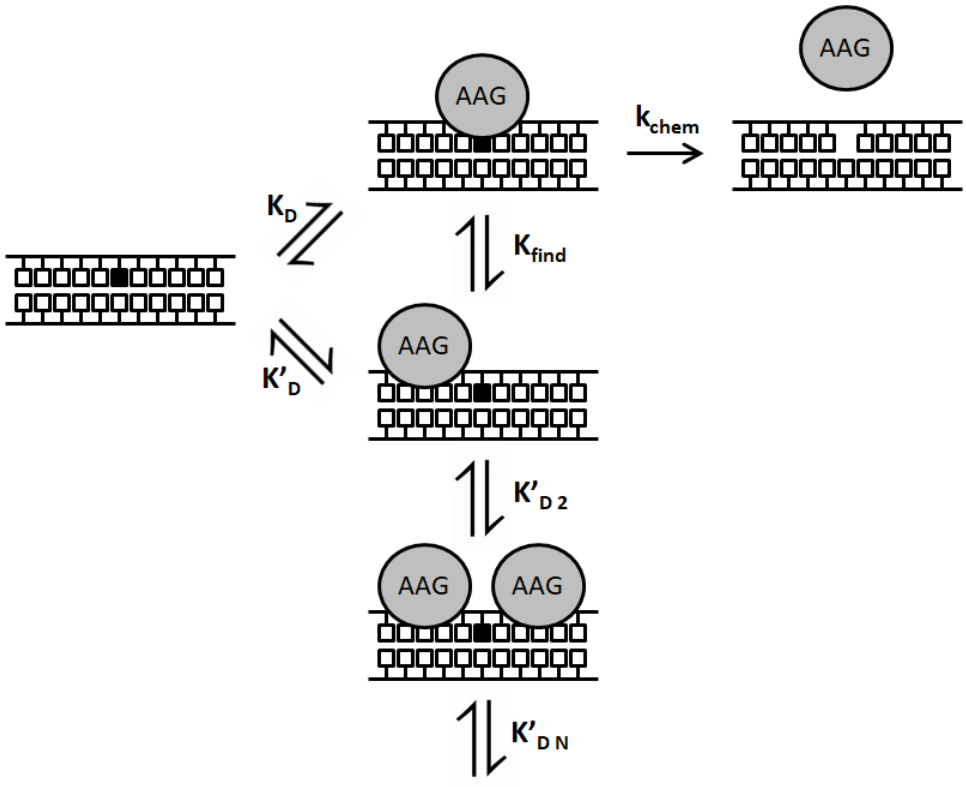


Figure A4. Model for inhibition by nonspecific binding of multiple AAG molecules. Single turnover glycosylase activity assays in Figures 2-4 and A3 reveal an inhibitory effect of excess AAG. An analogous effect was previously observed for the DNA glycosylase AlkA. In that case, the inhibition was explained via competition between specific and nonspecific DNA binding sites. The proposed model for AAG inhibition is the same, with nonspecific binding occluding the damaged site and restricting specific binding. The binding constant (KD) represents specific binding of AAG to the lesion, whereas K′D through K′D N represent the binding of proteins to nonspecific sites which limit and ultimately preclude specific binding. The rate constant k_{chem} represents N-glycosidic bond cleavage.

APPENDIX B

Additional Data Figures to Support Chapter 3

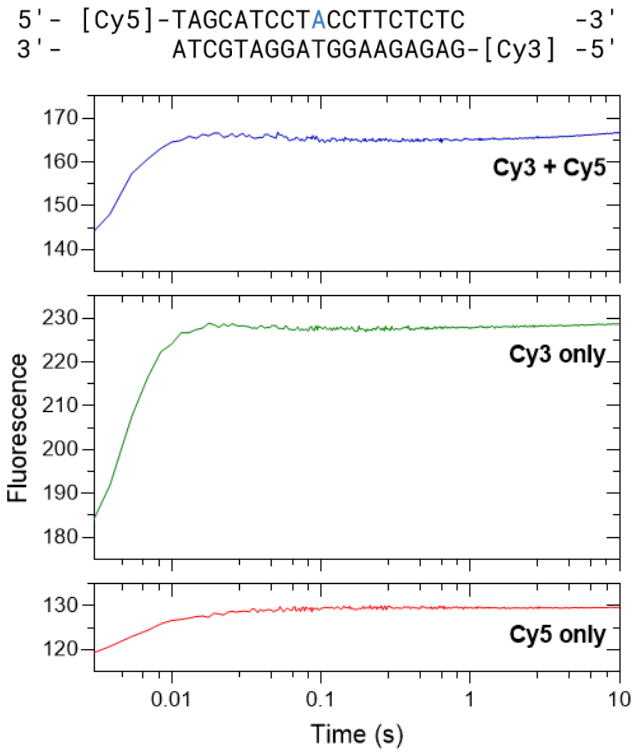


Figure B1. Stopped Flow FRET assay with undamaged DNA. Stopped flow fluorescence traces of 500nM AAG mixed with the indicated 100nM undamaged DNA. Oligos were labeled with a Cy3 FRET donor (green), a Cy5 FRET acceptor (red), or both (blue). Reactions containing Cy3 were excited at 547 nM, while the Cy5-only reaction was excited at 577 nM. A 660 nM long-pass emission filter was used for reactions with the Cy5 label, while a 590 nM long-pass emission filter was used for the Cy3-only reaction. All 3 reactions showed an initial increase in fluorescence comparable in time scale to the first phase of Figure 3-3.

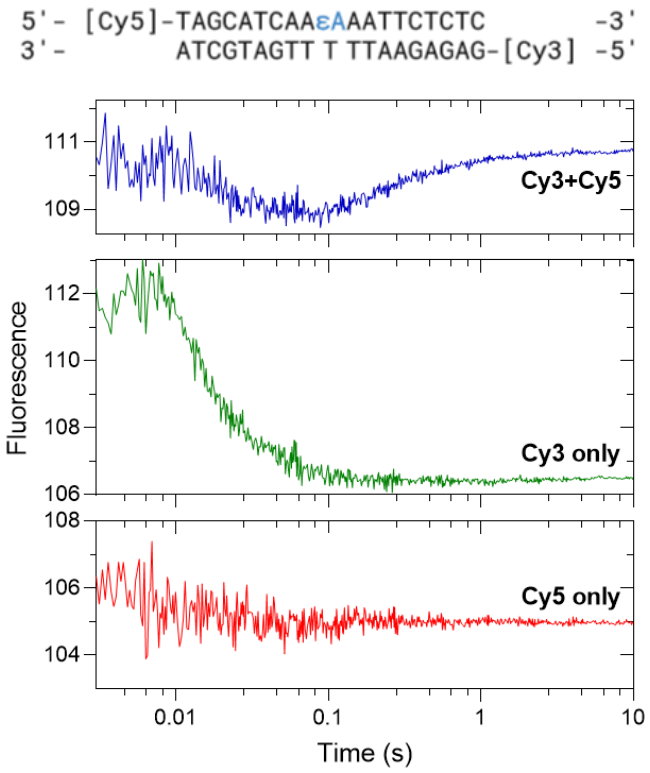


Figure B2. DNA bending FRET signal with a different DNA sequence context. Stopped flow fluorescence traces of 80 nM AAG mixed with 100 nM DNA oligonucleotides containing a Cy3 FRET donor, a Cy5 FRET acceptor, or both. The Cy3 donor was excited at 547 nm with a 590 nm long-pass emission filter. The Cy5 acceptor was excited at 577 nm with a 660 nm long-pass emission filter. Doubly-labeled DNA was excited at 547 nm with a 660 nm long-pass filter.

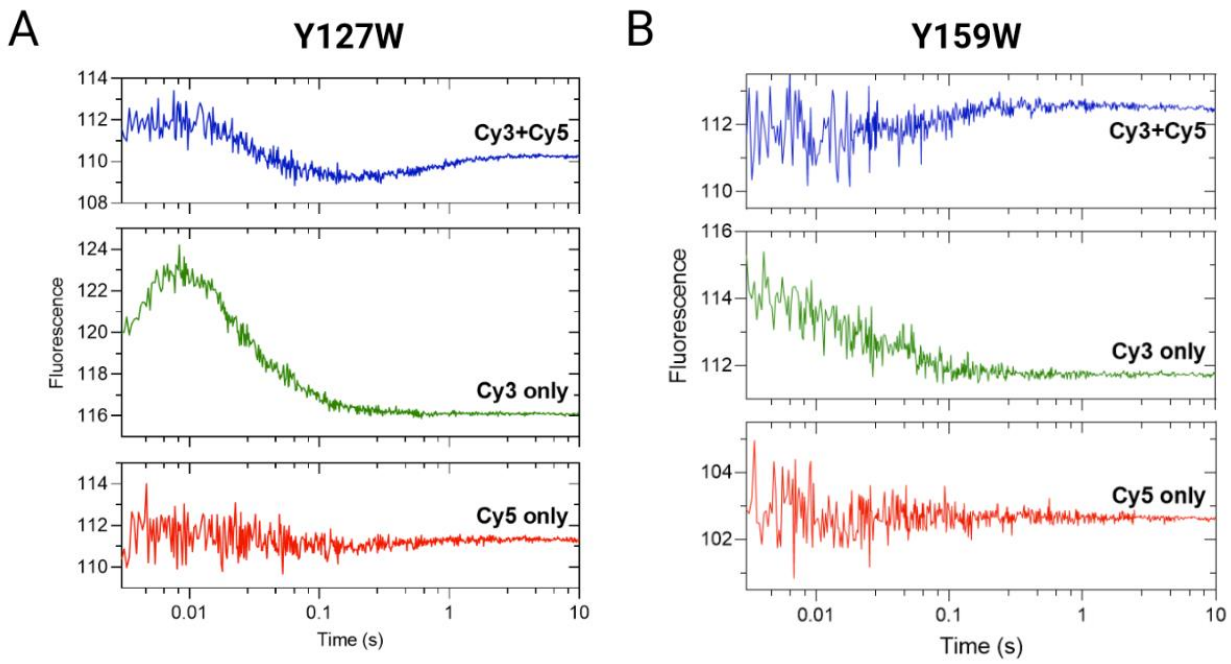
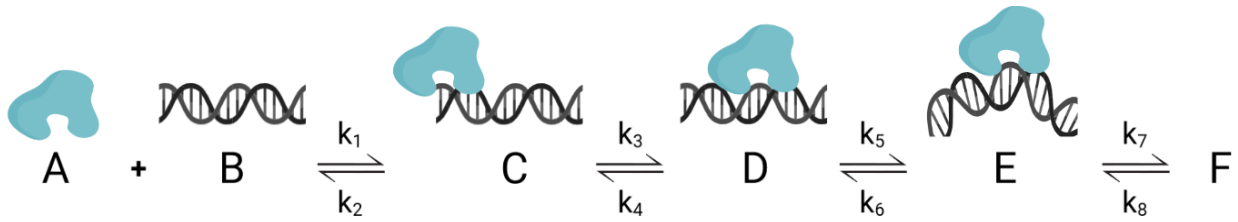


Figure B3. Single-fluorophore interactions of AAG variants. (A) Stopped flow fluorescence traces of 100nM DNA oligonucleotides containing a Cy3 FRET donor, a Cy5 FRET acceptor, or both mixed with 80nM AAG (A) Y127W, (B) Y159W. Donor excited at 547nm with a 590nm longpass filter. Acceptor excited at 577nm with a 660 longpass filter. Doubly-labeled DNA excited at 547nm with a 660nm longpass filter.

Figure B4. Example of Berkeley Madonna script for fitting of DNA bending. (A) Reaction scheme for the binding of AAG to DNA. Variables A-F represent each reaction species, with A = free AAG, B = free DNA, C = nonspecific AAG binding at DNA ends, D = searching complex, E = specific recognition complex, F = slow FRET increase. (B) Rate constants from Berkeley Madonna fits. Each value represents the average of 3 replicates ± S.D. The reverse rate constants for each step fit to very low values, indicating the reaction scheme could be simplified to 4 irreversible steps. (C) Representative script from Berkeley Madonna. To fit the overall fluorescence, Z, of the reaction, fluorescence coefficients were assigned to each DNA species (FB, FC, FD, FE, FF) with a floating baseline, H.

A



B

AAG Variant		WT	Y127W	Y159W	WT	WT
Lesion Context		εA•T	εA•T	εA•T	εA•Spacer C3	εA•Spacer 9
$k_{1\ (on)}$	$(M^{-1}s^{-1})$	$4.2x10^{8} \pm 2.5x10^{8}$	$3.0x10^{8} \pm 9.4x10^{7}$	$4.2x10^{8} \pm 2.2x10^{8}$	$5.9x10^{8} \pm 1.9x10^{8}$	$3.9x10^{8} \pm 1.7x10^{8}$
$k_{2\ (off)}$	(s^{-1})	$9.7x10^{-11} \pm 8.8x10^{-11}$	$9.7x10^{-11} \pm 8.3x10^{-11}$	$3.8x10^{-10} \pm 4.1x10^{-10}$	$4.0x10^{-10} \pm 6.5x10^{-10}$	$1.2x10^{-10} \pm 1.6x10^{-10}$
$k_{3\ (find)}$		380 ± 96	220 ± 100	950 ± 360	290 ± 160	310 ± 98
k_4		$1.7x10^{-7} \pm 2.0x10^{-7}$	1.3 ± 2.2	$3.2x10^{-7} \pm 2.1x10^{-7}$	$7.8x10^{-8} \pm 1.3x10^{-7}$	$8.8x10^{-7} \pm 8.8x10^{-7}$
$k_{5\ (bend)}$		4.2 ± 0.7	1.5 ± 0.1	32.8 ± 16.7	1.8 ± 0.04	12.9 ± 5.1
$k_{6\ (unbend)}$		$8.2x10^{-4} \pm 7.3x10^{-4}$	$6.3x10^{-5} \pm 1.1x10^{-4}$	$3.8x10^{-3} \pm 3.5x10^{-3}$	$1.6x10^{-3} \pm 2.2x10^{-3}$	$6.8x10^{-3} \pm 7.0x10^{-3}$
$k_{7\ (slow)}$		0.016 ± 0.023	0.033 ± 0.029	0.85 ± 1.1	0.054 ± 0.027	0.13 ± 0.19
k_8		$1.6x10^{-4} \pm 4.6x10^{-5}$	$1.7x10^{-5} \pm 1.5x10^{-5}$	$1.8x10^{-4} \pm 2.0x10^{-4}$	$4.7x10^{-5} \pm 7.8x10^{-5}$	$4.0x10^{-5} \pm 4.3x10^{-5}$

C

```
METHOD RK4
STARTTIME = 0
STOPTIME=15
DT = 0.0001

Z = FB*B + FC*C + FD*D + FE*E + FF*F + H

d/dt (A) = k2*C - k1*A*B
d/dt (B) = k2*C - k1*A*B
d/dt (C) = k1*A*B + k4*D - k2*C - k3*C
d/dt (D) = k3*C + k6*E - k4*D - k5*D
d/dt (E) = k5*D + k8*F - k6*E - k7*E
d/dt (F) = k7*E - k8*F
```

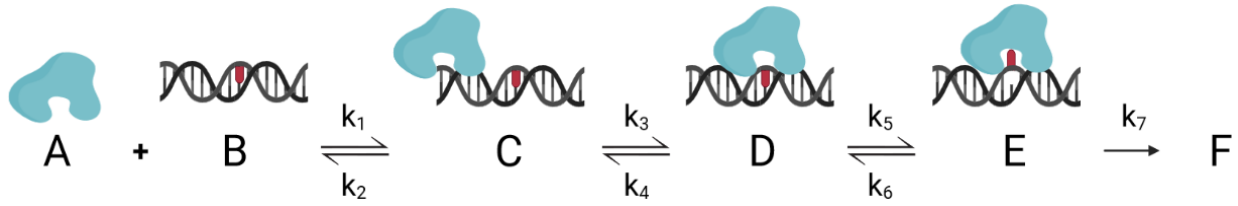
init A = 0.00000009
init B = 0.00000010
init C = 0
init D = 0
init E = 0
init F = 0
H = 107.39015

FB = 5.517106E6
FC = 1.3925E9
FD = 4.4873E7
FE = 5.5339E7
FF = 7.1326E7

k1 = 2.3168E8
k2 = 1.7096E-10
k3 = 377.71242
k4 = 3.8645E-7
k5 = 4.83605
k6 = 3.0182E-6
k7 = 0.00479
k8 = 1.6121E-4

Figure B5. Example of Berkeley Madonna script for fitting base flipping. (A) Reaction scheme for the flipping of eA by AAG. Variables A-F represent each reaction species, with A = free AAG, B = free DNA, C = nonspecific AAG binding, D = initial recognition complex, E = specific recognition complex, F = slow artifact. (B) Table of rate constants for 3 reactions of WT AAG with εA•T DNA. The reverse rate constants fit to very low values, so the reaction scheme could be simplified to 4 irreversible steps. Both binding and finding, k_1 and k_3, could not be averaged across reactions with different excess AAG concentrations, as both steps have been demonstrated to vary with excess AAG. Base flipping, k_5, is independent of AAG concentration, as shown in Fig. 3-4. (C) Representative script from Berkeley Madonna. To fit the overall fluorescence, Z, of the reaction, fluorescence coefficients were assigned to each DNA species (FBC, FD, FE, FF) with a floating baseline, H.

A



B

[AAG WT]		1 μM	350 nM	1 μM
[εA•T DNA]		500 nM	300 nM	200 nM
$k_{1\ (on)}$	$(M^{-1}s^{-1})$	3.2×10^{8}	7.9×10^{8}	3.8×10^{8}
$k_{2\ (off)}$	(s^{-1})	1.8×10^{-9}	1.7×10^{-10}	6.0×10^{-10}
$k_{3\ (find)}$		3700	680	8030
k_4		2.4×10^{-7}	2.8×10^{-7}	2.0×10^{-6}
$k_{5\ (flip)}$		5.7	4.8	4.0
$k_{6\ (unflip)}$		1.8×10^{-17}	2.5×10^{-4}	0.063
$k_{7\ (slow)}$		0.94	0.16	0.30

C

```
METHOD RK4

STARTTIME = 0
STOPTIME=10
DT = 0.0001

Z = FBC*B + FBC*C + FD*D + FE*E + FF*F + H

d/dt (A) = k2*C - k1*A*B
d/dt (B) = k2*C - k1*A*B
d/dt (C) = k1*A*B + k4*D - k2*C - k3*C
d/dt (D) = k3*C + k6*E - k4*D - k5*D
d/dt (E) = k5*D - k6*E - k7*E
```

d/dt (F) = k7*E

init A = 0.00000035
init B = 0.00000030
init C = 0
init D = 0
init E = 0
init F = 0
H = 132.70572

FBC = 8.4226E7
FD = 2.2336E8
FE = 5.7013E7
FF = 4.7639E7

k1 = 8.0075E8
k2 = 6.9911E-11
k3 = 580.03186
k4 = 2.5703E-7
k5 = 4.7494
k6 = 5.5045E-8
k7 = 0.1