Mechanism of Nucleotide Flipping by Human Alkyladenine DNA Glycosylase

by

Jenna M. Hendershot

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biological Chemistry)
in the University of Michigan
2014

Doctoral Committee:

Associate Professor Patrick J. O'Brien, Chair
Professor Hashim M. Al-Hashimi, Duke University
Professor Carol A. Fierke
Associate Professor Bruce A. Palfey
Associate Professor Raymond C. Trievel

Table of Contents

List of Figures

List of Tables

List of Appendices

List of Abbreviations

A: adenine (nucleobase)

AAG: alkyladenine DNA glycosylase, also known as methylpurine DNA glycosylase (MPG) and 3-methyladenine DNA glycosylase

AlkA: 3-methyladenine DNA glycosylase II

AP: apurinic/apyrimidinic

BER: base excision repair

BM: Berkeley Madonna

C: cytosine (nucleobase)

DHU: 5,6-dihydro-2'-deoxyuridine

dRP: deoxyribosephosphate

DTT: dithiothreitol

εA: 1,N^6-ethenoadenine

εC: 3,N^4-ethenocytosine

EDTA: ethylenediaminetetraacetic acid

FAM: 6-aminofluorescein

fapyA: 4,6-diamino-5-formamidopyrimidine

fapyG: 2,6-diamino-4-hydroxy-5-formamidopyrimidine

Fpg: formamidopyrimidine DNA glycosylase

HR: homologous recombination

Hx: hypoxanthine (the base moiety of inosine)

I: inosine

MagIII: 3-methyladenine DNA glycosylase III

MBD4: methyl-CpG binding domain protein 4

MMR: mismatch repair

MutM: formamidopyrimidine DNA glycosylase

MutY: adenine DNA glycosylase

NaHEPES: sodium N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonate)

NaMES: sodium 2-(N-morpholino)ethanesulfonate; PAGE, polyacrylamide gel electrophoresis

NEIL1: nei endonuclease VIII-like I

NER: nucleotide excision repair

NHEJ: non-homologous end joining

NMR: nuclear magnetic resonance

Nth1: nth endonuclease II - like 1

Ogg1: 8-oxoguanine DNA glycosylase

rRNA: ribosomal RNA

SD: standard deviation

SMUG1: single-strand-selective monofunctional uracil DNA glycosylase

T: thymine (nucleobase)

TDG: thymine DNA glycosylase

Tg: thymine glycol

U: uracil (nucleobase)

UDG: uracil DNA glycosylase

UV: ultraviolet

WT: wild-type Δ80 AAG, N-terminal truncation mutant of AAG in which the first 79 amino acids have been deleted

3MeA: 3-methyladenine

5-OHC: 5-hydroxycytosine

5-OH-meU: 5-(hydroxymethyl)-uracil

5-OHU: 5-hydoxyuridine

7MeG: 7-methylguanine

8-oxoG: 8-oxoguanine

Abstract

Nucleotide flipping is the phenomenon whereby a nucleotide in DNA is rotated 180° out of the double helix into an extrahelical position. This event is critical to many biological processes in which enzymes need access to a specific nucleotide, but the mechanism remains poorly understood. Human alkyladenine DNA glycosylase (AAG) is one of many enzymes that use nucleotide flipping to engage substrates. AAG initiates the base excision repair pathway by locating sites of damage and catalyzing the hydrolysis of the *N*-glycosidic bond to release the damaged base. Previous crystal structures of AAG in complex with DNA revealed the DNA is bent where a tyrosine (Y162) intercalates into the space left by the flipped-out damaged nucleotide. The kinetic mechanism of binding, nucleotide flipping, and base excision for wild-type and mutant forms of AAG was determined. Mutation of active site tyrosines to tryptophans revealed a new intermediate that provides selectivity between damaged and undamaged sites prior to nucleotide flipping. Additionally, the functional contributions of the intercalating residue (Y162) to each of the steps in the AAG catalytic mechanism were determined. This residue acts as a selectivity filter and slows nucleotide flipping. Mutating the intercalating tyrosine to tryptophan revealed that intercalation occurs early in the search for damage. Similar intercalating interactions are observed in all nucleotide flipping enzymes, but

the identity of this residue varies. Therefore, a structure/function study of AAG was performed in which the intercalating tyrosine was mutated to numerous amino acids of varying sizes. The results showed that the rate constants for nucleotide flipping and unflipping are faster for non-aromatic mutants, and an aromatic residue is required to stabilize the initial recognition complex and the stable recognition complex. Together, these results provide new insight into the mechanism of AAG, and have implications for other enzymes that use nucleotide flipping for DNA repair or epigenetic DNA modification.

CHAPTER 1

Introduction

Nucleotide Flipping

To recognize and repair DNA, enzymes must overcome the problem of gaining access to bases that are normally embedded in the duplex. One solution to this problem is a process called base or nucleotide flipping that involves the complete 180° rotation of a damaged base into an enzyme active site where its repair can begin (Figure 1.1; (1)). Nucleotide flipping was discovered in 1995 when the co-crystal structure of cytosine-5 DNA methyltransferase, HhaI, bound to DNA was determined (3). In this ternary structure of M. HhaI in complex with the DNA substrate and the reaction product, S-adenosyl-L-homocysteine, the target cytosine had rotated 180° out of the helix and into the enzyme active site. Due to overall similarities in the primary sequences of all cytosine-5 DNA methyltransferases (4), it seemed reasonable that M. HhaI was not an isolated example of nucleotide flipping. It was reassuring when additional structures of different methyltransferases showed the same phenomenon. Since then, nucleotide flipping has been shown for a variety of other enzymes including: DNA glycosylases (2), endonucleases (5, 6), non-catalytic transcription factors (7), and 5-methylcytosine recognition domains (8, 9).

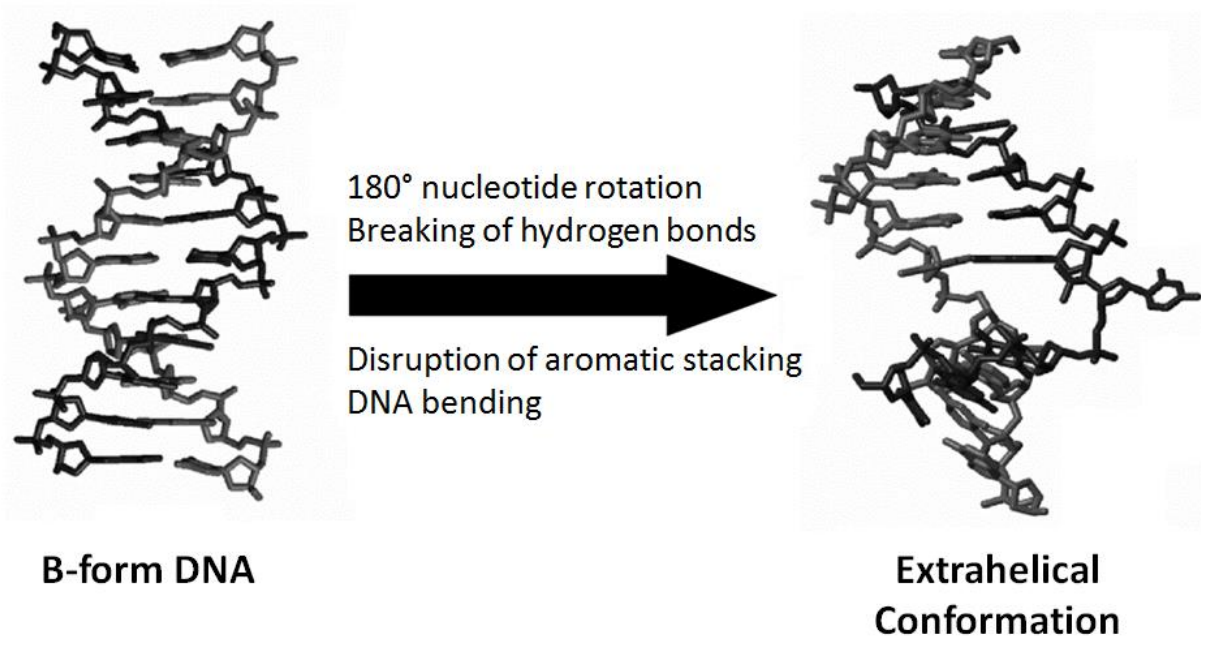


Figure 1.1. Nucleotide flipping. Many enzymes use nucleotide flipping to access nucleobases in duplex DNA. During this process, a nucleotide is rotated 180° into the enzyme active site. Watson-Crick hydrogen bonds are broken, aromatic stacking is disrupted, and the DNA is bent. Enzymes must pay for these unfavorable energetic costs through the use of favorable binding energy. Figure adapted from (2).

Not surprisingly, this process is not limited to DNA; as RNA nucleotide flipping has also been observed (10-13). In one specific example, RNA nucleotide flipping occurs when aminoglycoside antibiotics bind to the 16S ribosomal RNA (rRNA) A site. Upon drug binding, two conserved adenine residues (at positions 1492 and 1493 in *Escherichia coli* 16S rRNA) flip out of the helical stack toward the minor groove, where they interact with and stabilize the minihelix that is formed by the mRNA codon and the tRNA anticodon during translation (14-17).

Overall, three forms of nucleotide flipping have been observed in crystal structures: (**1**) only the target nucleotide is flipped from the DNA duplex (18), (**2**) only the nucleotide across from the target nucleotide is flipped (19), or (**3**) double flipping in which both the target nucleotide and the opposing nucleotide are flipped from the duplex (20). In all of these cases,

Watson-Crick hydrogen bonds are broken, aromatic stacking interactions between DNA bases are disrupted, and the DNA is bent (1, 21). Crystal structures allow us to observe the final extrahelical state, and we can often determine the extent of DNA bending. However, the timing of DNA bending remains unknown. We would like to know whether the DNA bends prior to, simultaneous with, or after nucleotide flipping. Even though nucleotide flipping is a common feature of repair enzymes, little is known about the fundamental mechanisms by which nucleotide flipping is accomplished because it is difficult to observe rapid conformational changes that occur on the millisecond time scale.

DNA Damage and Repair

Although DNA is a remarkably stable molecule, it is nevertheless subject to modifications by a variety of reactive intracellular and environmental agents. Out of the ~12,000,000,000 nucleotides in the genome, there are estimated to be ~10,000 lesions per cell per day (22). If not repaired, DNA damage can cause mutations, alter gene expression, affect epigenetic profiles, and even cause cell death (22, 23). Fortunately, a variety of DNA repair pathways exist to remove the different forms of damage from the genome. For example, the nucleotide excision repair (NER) pathway is responsible for removing helix distorting lesions such as bulky adducts and crosslinked nucleotides (24, 25). Mismatch repair (MMR) can be coupled with DNA replication to repair unpaired and mismatched nucleotides that result from polymerase errors during replication (26). Homologous recombination (HR) and non-homologous end joining (NHEJ) are responsible for repairing double-stranded breaks (27). Overall, each DNA repair pathway specializes in repairing distinct classes of DNA damage.

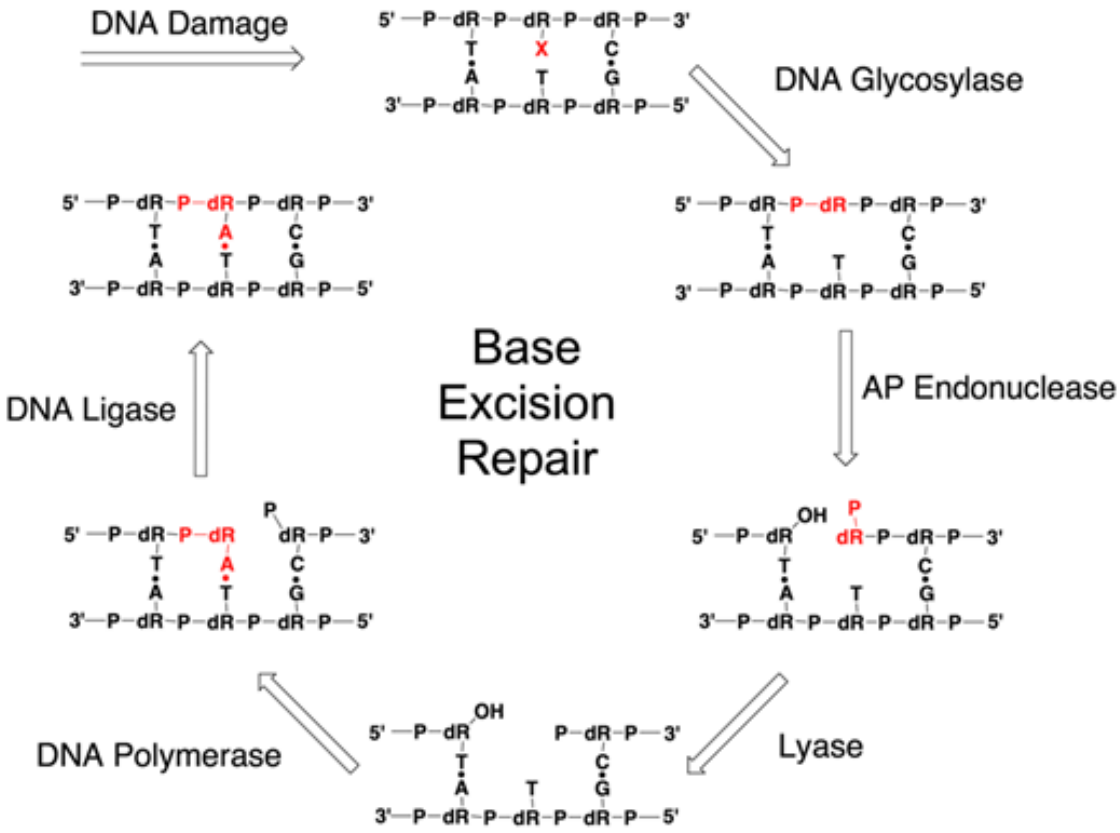


Figure 1.2. Base excision repair pathway initiated by a monofunctional DNA glycosylase. DNA damage (red X) is recognized and removed by a DNA glycosylase. An AP Endonuclease binds to the abasic site, and cleaves 5' of the deoxyribose. This leaves a nick in the DNA strand. The dRP lyase activity of DNA polymerase β removes the deoxyribose, and DNA polymerase fills the gap. Finally, DNA ligase seals the nick in the DNA, and repair is completed.

Most individual damaged nucleobases are repaired by the base excision repair (BER) pathway (Figure 1.2). This pathway is initiated by a DNA glycosylase, an enzyme that continuously searches the genome for a wide variety of damaged bases. Upon locating a site of damage, the glycosylase catalyzes the hydrolysis of the *N*-glycosidic bond to release the damaged base (28). The abasic-site containing DNA product is highly mutagenic and can be lethal if left exposed within cells. Therefore, it is critical that this intermediate is further processed, and the repair pathway is completed (29). An abasic-site-specific endonuclease

cleaves the phosphodiester backbone 5' to the abasic site, leaving a 3' OH and a 5' deoxyribosephosphate (dRP) group. The dRP lyase activity of DNA polymerase β releases the dRP group, and uses the intact template strand to insert the correct nucleotide into the DNA. Finally, DNA ligase completes the pathway by sealing the nick.

DNA Glycosylases

Humans have eleven different glycosylases to remove the broad spectrum of damaged nucleotides (23). These eleven glycosylases are divided into four structural superfamilies: helix-hairpin-helix, uracil DNA glycosylase, fapyG DNA glycosylase, and alkyladenine DNA glycosylase (Table 1.1). While evidence suggests the different structural families evolved independently, all glycosylases recognize and remove damaged bases from the genome (30). Each glycosylase is faced with the task of searching the genome for rare sites of damage. Upon finding a lesion, the DNA glycosylase uses nucleotide flipping to rotate the damaged nucleotide 180° into the enzyme active site. Both DNA bending and DNA intercalation are important for this step.

Structural Family	Glycosylase	Selected Substrates
Helix-hairpin-Helix	Nth1	Tg, fapyA, DHU, 5-OHU, 5-OHC
	OGG1	8-oxoG, fapyG
	MutY	A opposite 8-oxoG
	MBD4	T and U opposite G
Uracil DNA Glycosylase	UDG	Uracil
	SMUG1	Uracil, 5-OH-meU
	TDG	T, U, εC
FapyG DNA Glycosylase	NEIL1	Tg, fapyG, DHU, 5-OHU, fapyA, 8-oxoG
	NEIL2	Tg, fapyG, DHU, 5-OHU, fapyA, 8-oxoG
	NEIL3	Tg, fapyG, 5-OHU, fapyA
Alkyladenine DNA Glycosylase	AAG	3MeA, 7MeG, εA, Hx

Table 1.1. Humans have eleven different DNA glycosylases from four distinct structural families. While some DNA glycosylases, like UDG, have narrow substrate specificity (recognize one or a few structurally related lesions), other DNA glycosylases, like AAG, have broad substrate specificity (recognize a variety of structurally diverse lesions).

All glycosylases that use nucleotide flipping insert one or more intercalating residues into the space left by the flipped-out nucleotide. Once the damaged nucleotide is positioned in the active site, the DNA glycosylase catalyzes the hydrolysis of the *N*-glycosidic bond. In addition to hydrolysis, some glycosylase are bifunctional and have additional AP lyase activity (31). Of the eleven glycosylases, six are bifunctional (Nth1, OGG1, MutY, NEIL1, NEIL2, and NEIL3) (32).

While the overall mechanism for excision of a damaged nucleotide is largely conserved within each glycosylase family, the substrate specificity of each glycosylase varies (Table 1.1). Some DNA glycosylases, like uracil DNA glycosylase (UDG), have narrow substrate specificity. UDG only excises uracil from DNA. This is achieved by a sterically constrained active site that occludes all other nucleobases (33). In comparison, alkyladenine DNA glycosylase (AAG) removes a diverse group of purine lesions that result from alkylation (3-methyladenine, 3-methylguanine, and 7-methylguanine) and reactive oxygen and nitrogen species (hypoxanthine (Hx), 1,N^6-ethenoadenine (εA), and 1,N^2-ethenoguanine) (34-38). AAG uses steric clashes with the exocyclic amines of G and A to discriminate against the excision of undamaged bases. The smaller pyrimidines are too small to be excluded from the active site, but they are not excised because general acid catalysis is optimized for protonation of the larger purine rings (39). However, as a consequence of this broad specificity, AAG has reduced catalytic efficiency compared to UDG. Overall, DNA glycosylases can afford to sacrifice speed for increased selectivity, because sites of damage are rare (40).

Human Alkyladenine DNA Glycosylase (AAG)

AAG is a particularly attractive system to delve into the biophysical and biochemical mechanism of multi-step DNA recognition. Crystal structures of AAG in complex with DNA revealed this

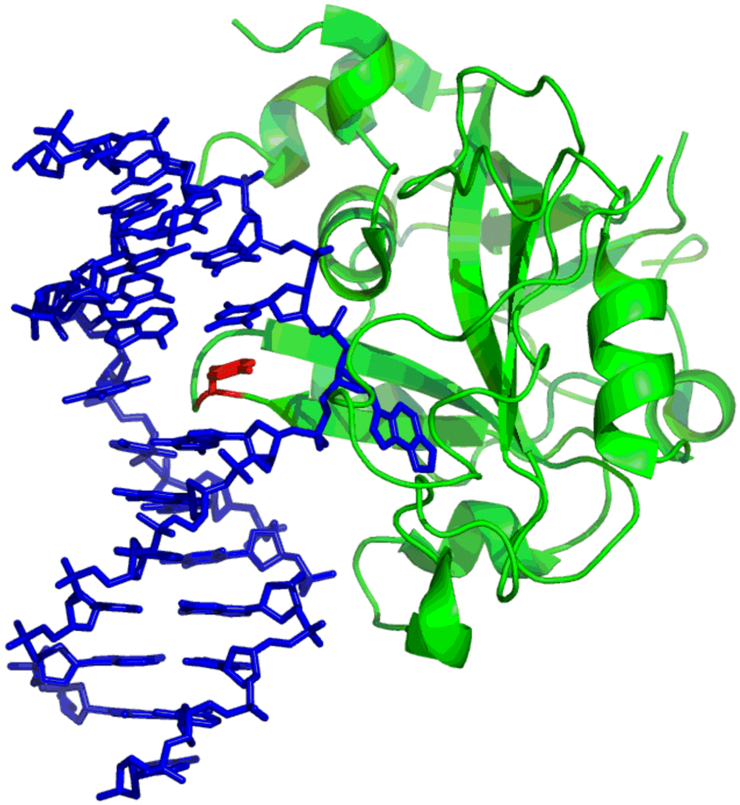


Figure 1.3. Crystal structure of human alkyladenine DNA glycosylase (E125Q) bound to εA-containing DNA. The εA base is flipped into the active site and stacks between two tyrosine residues (Y127 and Y159). The intercalating residue (Y162) is shown in red. Image was rendered with Pymol (http://www.pymol.org) using coordinates from the PDB (1EWN).

enzyme uses nucleotide flipping to rotate lesioned bases out of the DNA duplex and into its active site (Figure 1.3; (18, 41)). The expression of AAG confers increased cell survival in the presence of exogenous alkylating agents (42, 43); however, its glycosylase activity has been shown to be responsible for retinal degeneration (apoptosis) in mice when exposed to DNA alkylating agents (44), and the overexpression of AAG leads to enhanced sensitivity to DNA alkylating agents in other contexts (45, 46). Furthermore, increased levels of AAG have been linked to microsatellite instability (47, 48). Thus, it is important that cells have the correct balance of AAG activity.

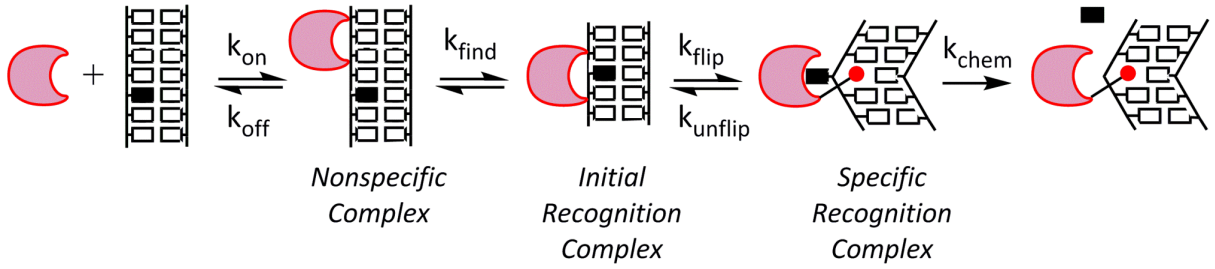


Figure 1.4. Minimal mechanism for the initiation of base excision repair. AAG (crescent) binds to nonspecific DNA and rapidly searches for DNA damage. This searching process is described by the macroscopic rate constant k_{find}. Once a lesion (solid rectangle) is encountered in an initial recognition complex, it can be flipped into the enzyme active site to form the specific recognition complex. In this specific complex, Y162 (red circle) intercalates into the DNA where it takes the place of the extrahelical lesion. AAG catalyzes hydrolysis of the *N*-glycosidic bond (k_{chem}) from this complex. The overall equilibrium constant for nucleotide flipping (K_{flip}) is equal to k_{flip}/k_{unflip}.

One of the lesions that is most efficiently excised by AAG is the naturally fluorescent 1,N^6-ethenoadenine (εA), that is formed by lipid oxidation or exposure to exogenous alkylating agents (37, 49). The minimal kinetic mechanism for multi-step recognition of εA damage has been reported (Figure 1.4; (40, 50)). Given the excess of undamaged DNA, initial binding (k_{on}) of AAG is not at the site of damage. AAG uses nonspecific binding and facilitated diffusion (51) to search for sites of damage (k_{find}). Once an εA is found, AAG forms an initial recognition complex that is accompanied by changes in the stacking of the lesion base (40). From this intermediate, AAG catalyzes nucleotide flipping (k_{flip}) to form a more stable specific recognition complex that positions the substrate for *N*-glycosidic bond cleavage (k_{chem}; (18)). This multi-step recognition provides multiple opportunities for discrimination between damaged and undamaged nucleotides, and similar mechanisms have been proposed for independently evolved DNA repair glycosylases, including 8-oxoguanine DNA glycosylases and uracil DNA glycosylases (52-55).

Monitoring Nucleotide Flipping

While much is known about nucleotide flipping of εA by AAG, the rate constants for nucleotide flipping of non-fluorescent substrates remain unknown. Tryptophan fluorescence has been shown to be an excellent tool to follow structural changes in a protein. The catalytic domain of AAG (residues 80-298) contains three tryptophan and six tyrosine residues. The three native tryptophan residues are all found on the back surface, opposite the protein active site and the DNA binding interface. The location of the three tryptophan residues suggested that they would not change their environment upon DNA binding. Five of the tyrosine residues line either the binding pocket or the DNA-protein interface, but the absorbance of tyrosine is very close to the absorbance of DNA, and it is unlikely that they would serve as useful reporters to monitor nucleotide flipping of non-fluorescent lesions. Therefore, in Chapter 2, tryptophan residues were introduced into the binding pocket of AAG, substituting either Y127 or Y159, the two tyrosine residues that sandwich the bound lesion in the crystal structures (18, 41). While substitution of tyrosine residues with tryptophan at positions 127 and 159 in the AAG active site had minimal effects on the excision of εA, there were much larger effects on the rate and equilibrium constant for nucleotide flipping. These results are consistent with the simple model that bulky tryptophan residues in the active site destabilize the flipped-out conformation and accelerate the rate constant for unflipping. In spite of these differences, stopped-flow fluorescence experiments show that these mutant proteins exhibit the same two-state binding model as was observed for wild-type AAG. These observations indicate that the εA lesion and tryptophan side chains located in the AAG active site both report on the initial recognition complex and on the specific recognition complex.

DNA Intercalation

It is proposed that both DNA bending and DNA intercalation are important for nucleotide flipping (1, 21). Crystal structures of DNA glycosylases in complex with DNA revealed the lesioned base has rotated 180° out of the DNA duplex and into the enzyme active site (18, 33, 41, 56, 57). In these structures, the DNA is bent where an intercalating or 'wedge' residue fills the space left by the flipped-out damaged nucleotide. Intercalating interactions are observed in almost all nucleotide flipping enzymes, but the identity of the intercalating amino acid varies and the function of this residue remains unknown (2, 58). For example, AAG uses Tyr-162, AlkA uses Leu-125, UDG uses Leu-191, OGG1 uses Asn-149, MutM uses Met-77, and TDG uses Arg-275 (59).

The two models for how DNA intercalation could contribute to nucleotide flipping have been described as "pushing" and "plugging" (60-62). In the pushing model, DNA intercalation destabilizes the DNA duplex to accelerate nucleotide flipping. In the plugging model, DNA intercalation occurs subsequent to nucleotide flipping and the intercalating residue prevents the nucleotide from returning to the duplex, effectively slowing the rate of unflipping. These two mechanisms are not mutually exclusive, and in both proposed models the DNA duplex undergoes localized structural perturbations.

The functional contributions of the intercalating residue (Y162) to each of the steps in the AAG catalytic mechanism are addressed in Chapter 3. To determine the role of the intercalating tyrosine, this residue was mutated to phenylalanine (Y162F) and alanine (Y162A). Remarkably, the Y162A mutation increases the rate of nucleotide flipping by 50-fold relative to wild-type AAG. This mutant also exhibits faster unflipping, resulting in a 140-fold reduction in

the equilibrium constant for formation of the specific recognition complex. In contrast, the kinetic parameters for the Y162F and wild-type enzymes are very similar, suggesting that the hydroxyl of tyrosine is not necessary for function. These results establish that DNA intercalation contributes to the specific recognition of DNA damage by acting as a plug to stabilize the specific recognition complex. Tyrosine 162 slows the rate of nucleotide flipping by stabilizing the initial recognition complex. Therefore, our new model suggests that DNA intercalation contributes to the specificity of AAG, but the increased specificity comes at the cost of reduced speed.

In Chapter 4, site-directed mutagenesis and transient kinetic approaches were used to determine the timing of tyrosine intercalation for human AAG. The intercalating tyrosine residue was mutated to tryptophan to use tryptophan fluorescence as a tool to follow structural changes in the protein upon DNA binding. Excision of the naturally fluorescent lesion εA by Y162W is identical to wild-type, and the equilibrium constant for flipping remains highly favorable. Even though the Y162W mutant has not been found in nature, the results indicate that Y162W is a conservative mutation. This raised the possibility that tryptophan fluorescence could be used to probe DNA binding and nucleotide flipping steps. Using stopped-flow fluorescence, changes in both εA and tryptophan fluorescence were monitored. Surprisingly, the tryptophan fluorescence of Y162W does not report on nucleotide flipping. The rapid quenching of Y162W tryptophan fluorescence upon binding damaged or undamaged DNA suggests that Y162 intercalation occurs early in the search for DNA damage.

Almost all DNA glycosylases employ nucleotide flipping to access a damaged base. Many glycosylases use a bulky side chain (Asn, Arg, Leu, or Tyr) to fill the space left by the flipped-out

nucleotide, although an exception includes 3-methyladenine DNA glycosylase III (MagIII), which has a Gly at this position (63). To investigate how the size of the intercalating side chain affects nucleotide flipping, seven Y162 variants of AAG were made and individual steps of the kinetic mechanism were characterized in Chapter 5. The structure/function study suggests formation of the initial recognition complex, and nucleotide flipping are both affected by the size of this residue. The rate constant for nucleotide flipping is 4-9-fold faster for non-aromatic mutants and Y162H compared to wild-type. The non-aromatic mutants also exhibit faster unflipping, resulting in an unfavorable equilibrium constant for nucleotide flipping. The Y162H mutant is unique because it has a fast rate constant for nucleotide flipping, however like the other aromatic mutants, the Y162H mutant stabilizes the specific recognition complex. Overall, AAG requires an aromatic intercalating residue to stabilize the initial recognition complex and the specific recognition complex.

Conclusions

Collectively, these studies provide a foundation for understanding the mechanism of nucleotide flipping and the functional role of the intercalating residue. AAG has been a particularly attractive system to delve into the biophysical and biochemical mechanism of multi-step DNA recognition. As many different enzymes must employ similar strategies to discriminate between specific and nonspecific sites, lessons from AAG may be more widely applicable. Future work is needed to measure bending of DNA to address the temporal relationship of bending and flipping. In addition, nuclear magnetic resonance (NMR) should be used to obtain information about the initial recognition complex to help us understand how damaged sites are recognized.

References

1. Roberts RJ & Cheng X (1998) Base flipping. *Annu Rev Biochem* 67:181-198.

2. Stivers JT (2008) Extrahelical damaged base recognition by DNA glycosylase enzymes. *Chemistry* 14(3):786-793.

3. Klimasauskas S, Kumar S, Roberts RJ, & Cheng X (1994) HhaI methyltransferase flips its target base out of the DNA helix. *Cell* 76(2):357-369.

4. Kumar S*, et al.* (1994) The DNA (cytosine-5) methyltransferases. *Nucleic acids research* 22(1):1-10.

5. Horton JR*, et al.* (2006) DNA nicking by HinP1I endonuclease: bending, base flipping and minor groove expansion. *Nucleic acids research* 34(3):939-948.

6. Tsutakawa SE*, et al.* (2011) Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell* 145(2):198-211.

7. Schubot FD*, et al.* (2001) Crystal structure of the transcription factor sc-mtTFB offers insights into mitochondrial transcription. *Protein science : a publication of the Protein Society* 10(10):1980-1988.

8. Hashimoto H*, et al.* (2008) The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature* 455(7214):826-829.

9. Arita K, Ariyoshi M, Tochio H, Nakamura Y, & Shirakawa M (2008) Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature* 455(7214):818-821.

10. Hoang C & Ferre-D'Amare AR (2001) Cocrystal structure of a tRNA Psi55 pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme. *Cell* 107(7):929-939.

11. Rould MA, Perona JJ, Soll D, & Steitz TA (1989) Structure of E. coli glutaminyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 A resolution. *Science* 246(4934):1135-1142.

12. Rupert PB & Ferre-D'Amare AR (2001) Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature* 410(6830):780-786.

13. Yang X, Gerczei T, Glover LT, & Correll CC (2001) Crystal structures of restrictocin-inhibitor complexes with implications for RNA recognition and base flipping. *Nature structural biology* 8(11):968-973.

14. Kaul M, Barbieri CM, & Pilch DS (2004) Fluorescence-based approach for detecting and characterizing antibiotic-induced conformational changes in ribosomal RNA: comparing

aminoglycoside binding to prokaryotic and eukaryotic ribosomal RNA sequences. *Journal of the American Chemical Society* 126(11):3447-3453.

15. Kaul M, Barbieri CM, & Pilch DS (2005) Defining the basis for the specificity of aminoglycoside-rRNA recognition: a comparative study of drug binding to the A sites of Escherichia coli and human rRNA. *Journal of molecular biology* 346(1):119-134.

16. Kaul M, Barbieri CM, & Pilch DS (2006) Aminoglycoside-induced reduction in nucleotide mobility at the ribosomal RNA A-site as a potentially key determinant of antibacterial activity. *Journal of the American Chemical Society* 128(4):1261-1271.

17. Barbieri CM, Kaul M, & Pilch DS (2007) Use of 2-aminopurine as a fluorescent tool for characterizing antibiotic recognition of the bacterial rRNA A-site. *Tetrahedron* 63(17):3567-6574.

18. Lau AY, Wyatt MD, Glassner BJ, Samson LD, & Ellenberger T (2000) Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG. *Proceedings of the National Academy of Sciences of the United States of America* 97(25):13573-13578.

19. Vassylyev DG*, et al.* (1995) Atomic model of a pyrimidine dimer excision repair enzyme complexed with a DNA substrate: structural basis for damaged DNA recognition. *Cell* 83(5):773-782.

20. Bernards AS, Miller JK, Bao KK, & Wong I (2002) Flipping duplex DNA inside out: a double base-flipping reaction mechanism by Escherichia coli MutY adenine glycosylase. *The Journal of biological chemistry* 277(23):20960-20964.

21. Stivers JT (2004) Site-specific DNA damage recognition by enzyme-induced base flipping. *Progress in nucleic acid research and molecular biology* 77:37-65.

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

23. Robertson AB, Klungland A, Rognes T, & Leiros I (2009) DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cell Mol Life Sci* 66(6):981-993.

24. Diderich K, Alanazi M, & Hoeijmakers JH (2011) Premature aging and cancer in nucleotide excision repair-disorders. *DNA repair* 10(7):772-780.

25. Novarina D, Amara F, Lazzaro F, Plevani P, & Muzi-Falconi M (2011) Mind the gap: keeping UV lesions in check. *DNA repair* 10(7):751-759.

26. Lindahl T & Wood RD (1999) Quality control by DNA repair. *Science* 286(5446):1897-1905.

27. Jackson SP (2002) Sensing and repairing DNA double-strand breaks. *Carcinogenesis* 23(5):687-696.

28. Ye N, Holmquist GP, & O'Connor TR (1998) Heterogeneous repair of N-methylpurines at the nucleotide level in normal human cells. *Journal of molecular biology* 284(2):269-285.

29. Hosfield DJ*, et al.* (2001) DNA damage recognition and repair pathway coordination revealed by the structural biochemistry of DNA repair enzymes. *Progress in nucleic acid research and molecular biology* 68:315-347.

30. O'Brien PJ & Herschlag D (1999) Catalytic promiscuity and the evolution of new enzymatic activities. *Chemistry & biology* 6(4):R91-R105.

31. O'Connor TR & Laval J (1989) Physical association of the 2,6-diamino-4-hydroxy-5N-formamidopyrimidine-DNA glycosylase of Escherichia coli and an activity nicking DNA at apurinic/apyrimidinic sites. *Proceedings of the National Academy of Sciences of the United States of America* 86(14):5222-5226.

32. Svilar D, Goellner EM, Almeida KH, & Sobol RW (2011) Base excision repair and lesion-dependent subpathways for repair of oxidative DNA damage. *Antioxidants & redox signaling* 14(12):2491-2507.

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

34. Engelward BP*, et al.* (1997) Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. *Proceedings of the National Academy of Sciences of the United States of America* 94(24):13087-13092.

35. O'Brien PJ & Ellenberger T (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase. *The Journal of biological chemistry* 279(11):9750-9757.

36. Hitchcock TM*, et al.* (2004) Oxanine DNA glycosylase activity from Mammalian alkyladenine glycosylase. *The Journal of biological chemistry* 279(37):38177-38183.

37. Gros L, Ishchenko AA, & Saparbaev M (2003) Enzymology of repair of etheno-adducts. *Mutation research* 531(1-2):219-229.

38. Lee CY*, et al.* (2009) Recognition and processing of a new repertoire of DNA substrates by human 3-methyladenine DNA glycosylase (AAG). *Biochemistry* 48(9):1850-1861.

39. O'Brien PJ & Ellenberger T (2003) Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines. *Biochemistry* 42(42):12418-12429.

40. Wolfe AE & O'Brien PJ (2009) Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by human alkyladenine DNA glycosylase. *Biochemistry* 48(48):11357-11369.

41. Lau AY, Scharer OD, Samson L, Verdine GL, & Ellenberger T (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. *Cell* 95(2):249-258.

42. Guo HH, Choe J, & Loeb LA (2004) Protein tolerance to random amino acid change. *Proceedings of the National Academy of Sciences of the United States of America* 101(25):9205-9210.

43. Glassner BJ, Rasmussen LJ, Najarian MT, Posnick LM, & Samson LD (1998) Generation of a strong mutator phenotype in yeast by imbalanced base excision repair. *Proceedings of the National Academy of Sciences of the United States of America* 95(17):9997-10002.

44. Meira LB*, et al.* (2009) Aag-initiated base excision repair drives alkylation-induced retinal degeneration in mice. *Proceedings of the National Academy of Sciences of the United States of America* 106(3):888-893.

45. Rinne ML, He Y, Pachkowski BF, Nakamura J, & Kelley MR (2005) N-methylpurine DNA glycosylase overexpression increases alkylation sensitivity by rapidly removing non-toxic 7-methylguanine adducts. *Nucleic acids research* 33(9):2859-2867.

46. Trivedi RN*, et al.* (2008) Human methyl purine DNA glycosylase and DNA polymerase beta expression collectively predict sensitivity to temozolomide. *Mol Pharmacol* 74(2):505-516.

47. Hofseth LJ*, et al.* (2003) The adaptive imbalance in base excision-repair enzymes generates microsatellite instability in chronic inflammation. *J Clin Invest* 112(12):1887-1894.

48. Klapacz J*, et al.* (2010) Frameshift mutagenesis and microsatellite instability induced by human alkyladenine DNA glycosylase. *Mol Cell* 37(6):843-853.

49. Pandya GA & Moriya M (1996) 1,N6-ethenodeoxyadenosine, a DNA adduct highly mutagenic in mammalian cells. *Biochemistry* 35(35):11487-11492.

50. Hendershot JM, Wolfe AE, & O'Brien PJ (2011) Substitution of active site tyrosines with tryptophan alters the free energy for nucleotide flipping by human alkyladenine DNA glycosylase. *Biochemistry* 50(11):1864-1874.

51. Hedglin M & O'Brien PJ (2008) Human alkyladenine DNA glycosylase employs a processive search for DNA damage. *Biochemistry* 47(44):11434-11445.

52. Banerjee A, Santos WL, & Verdine GL (2006) Structure of a DNA glycosylase searching for lesions. *Science* 311(5764):1153-1157.

53. Banerjee A, Yang W, Karplus M, & Verdine GL (2005) Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. *Nature* 434(7033):612-618.

54. Parker JB*, et al.* (2007) Enzymatic capture of an extrahelical thymine in the search for uracil in DNA. *Nature* 449(7161):433-437.

55. Kuznetsov NA*, et al.* (2007) Pre-steady-state kinetic study of substrate specificity of Escherichia coli formamidopyrimidine--DNA glycosylase. *Biochemistry* 46(2):424-435.

56. Maiti A, Morgan MT, Pozharski E, & Drohat AC (2008) Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition. *Proceedings of the National Academy of Sciences of the United States of America* 105(26):8890-8895.

57. Bruner SD, Norman DP, & Verdine GL (2000) Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature* 403(6772):859-866.

58. Brooks SC, Adhikary S, Rubinson EH, & Eichman BF (2013) Recent advances in the structural mechanisms of DNA glycosylases. *Biochimica et biophysica acta* 1834(1):247-271.

59. Stivers JT & Jiang YL (2003) A mechanistic perspective on the chemistry of DNA repair glycosylases. *Chem Rev* 103(7):2729-2759.

60. Jiang YL & Stivers JT (2001) Reconstructing the substrate for uracil DNA glycosylase: tracking the transmission of binding energy in catalysis. *Biochemistry* 40(25):7710-7719.

61. Jiang YL & Stivers JT (2002) Mutational analysis of the base-flipping mechanism of uracil DNA glycosylase. *Biochemistry* 41(37):11236-11247.

62. Parikh SS*, et al.* (1998) Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *The EMBO journal* 17(17):5214-5226.

63. Eichman BF, O'Rourke EJ, Radicella JP, & Ellenberger T (2003) Crystal structures of 3-methyladenine DNA glycosylase MagIII and the recognition of alkylated bases. *The EMBO journal* 22(19):4898-4909.

CHAPTER 2

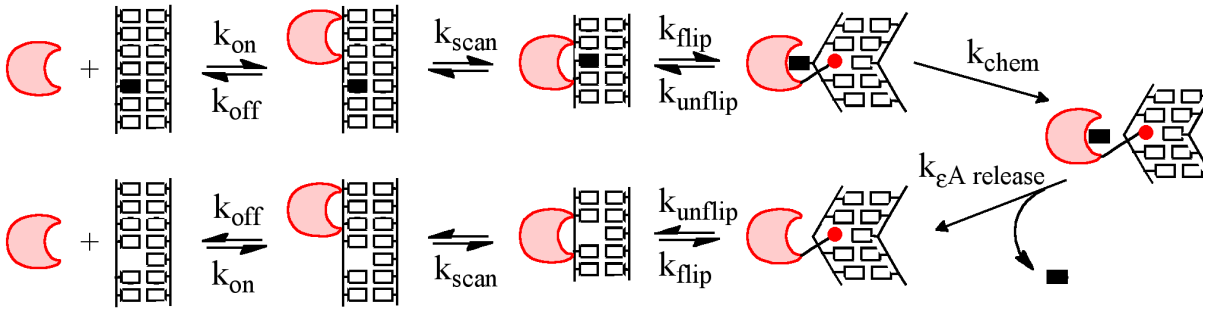
Substitution of Active Site Tyrosines with Tryptophan Alters the Free Energy for Nucleotide Flipping by Human Alkyladenine DNA Glycosylase[1,2]

Human alkyladenine DNA glycosylase (AAG) recognizes a wide variety of structurally diverse deaminated and alkylated purine lesions. This enzyme initiates the base excision DNA repair pathway by catalyzing the hydrolysis of the *N*-glycosidic bond and releasing the lesion base. Subsequent action of an abasic-site specific endonuclease, a 5'-deoxyribose phosphate lyase, a DNA polymerase, and a DNA ligase are required to restore the correct DNA sequence, using the intact strand as a template. AAG is the only human glycosylase known to recognize a wide variety of alkylated and deaminated purines (1-4). The expression of AAG confers increased cell survival in the presence of exogenous alkylating agents (5, 6), however its glycosylase activity has been shown to be responsible for retinal degeneration (apoptosis) in mice (7) and the overexpression of AAG leads to enhanced sensitivity to DNA alkylating agents in certain contexts (8, 9). Furthermore, increased levels of AAG have been linked to

[1]Reproduced with permission from Hendershot, J.M., Wolfe, A.E., and O'Brien, P.J. (2011) Substitution of active site tyrosines with tryptophan alters the free energy for nucleotide flipping by human alkyladenine DNA glycosylase, *Biochemistry 50*, 1864-1874. Copyright 2011 American Chemical Society. [2]All authors participated in planning the experiments. AEW developed the fluorescence-based assays for εA and tryptophan, JMH collected all of the data and analyzed the experiments, PJO and JMH wrote the manuscript.

microsatellite instability (10, 11). Thus, it is important that cells have the correct balance of AAG activity.

Crystal structures of AAG in complex with DNA revealed that this enzyme uses nucleotide flipping to gain access to the *N*-glycosidic bond of a damaged nucleotide (12, 13). We previously used the intrinsic fluorescence of εA to characterize the minimal kinetic mechanism for the recognition, flipping, and excision of this lesion by wild-type AAG (Scheme 2.1; (14)).



Scheme 2.1

Rapid mixing experiments indicated that an initial recognition complex rapidly and reversibly formed in which the εA lesion is partially unstacked. We hypothesized that the further change in the fluorescence of the εA lesion is indicative of the capture of the lesion in the active site of AAG. In order to test this hypothesis we sought to observe conformational changes in the protein directly, as similar approaches have been informative for other DNA glycosylases (15-18).

The catalytic domain of AAG (residues 80-298) contains six tyrosine and three tryptophan residues. The three native tryptophan residues are all found near the surface on the opposite side of the protein from the active site and the DNA binding interface, suggesting that they would not change their environment upon DNA binding. Five of the tyrosine residues line

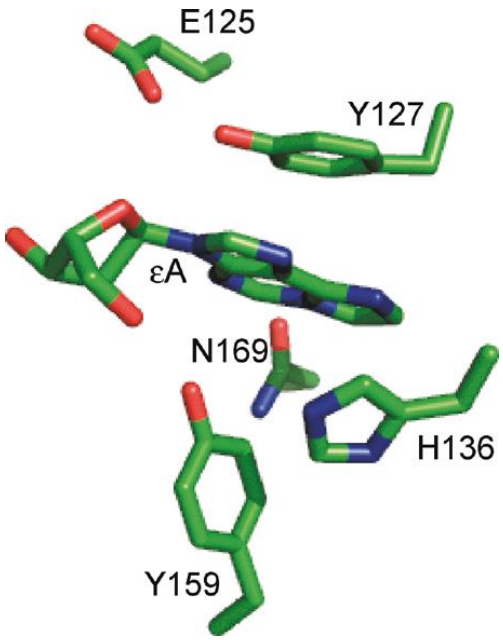


Figure 2.1. Active site contacts with the flipped-out εA from the published crystal structure of AAG bound to εA-DNA. The coordinates are from 1EWN (13) and the image was rendered with Pymol (http://www.pymol.org). Only the immediate residues surrounding the εA base are shown.

either the binding pocket or the DNA-protein interface, but the absorbance of tyrosine is very close to the absorbance of DNA and it is unlikely that they would serve as useful reporters. Therefore, we took the approach of introducing tryptophan residues into the binding pocket of AAG, substituting either Y127 or Y159, the two tyrosine residues that sandwich the bound lesion in the crystal structures (12, 13). The Y127 side chain stacks on top of the εA lesion, whereas the Y159 side chain makes a T-shaped interaction with the lesion (Figure 2.1). Comprehensive analysis of aromatic interactions in protein structures suggests that T-type interactions are common and presumably contribute similar binding energy as stacked interactions (19). It has been suggested that tyrosine to tryptophan mutations are relatively conservative (20), but these two positions in AAG are invariably found to be tyrosine in organisms ranging from bacteria to vertebrates (21).

We found that substitution of tyrosine residues with tryptophan at position 127 and 159 in the AAG active site had minimal effects on the excision of εA, however, there were much larger effects on the rate and equilibrium constant for nucleotide flipping. These results are consistent with the simple model that bulky tryptophan residues in the active site destabilize the flipped-out conformation and accelerate the rate constant for unflipping. However, it is surprising that the Y159W mutation resulted in a 5-fold faster rate constant for εA flipping. Although flipping is perturbed in these mutants, they follow the same mechanism as the wild-type enzyme with formation of two distinct intermediates prior to *N*-glycosidic bond cleavage. The observation of identical rate constants for time-dependent changes in either εA or tryptophan fluorescence, demonstrates that there are already changes in the active site environment prior to full nucleotide flipping, and that the most strongly quenched intermediate previously observed is indeed the extrahelical specific recognition complex (14).

Materials and Methods

Purification of Wild-type and Mutant AAG Proteins. The catalytic domain of human AAG that lacks the first 79 amino acids was expressed in *E. coli* and purified as previously described (22). The Y127W and Y159W mutants were constructed by site-directed mutagenesis and purified using the same protocol (22). The concentration of AAG was determined by fluorescent titration of εA-DNA, as described below, and the concentration was found to be within 2-fold of the concentration determined from the calculated extinction coefficient of the wild-type or mutant proteins.

Synthesis and Purification of Oligodeoxynucleotides. The 25mer oligonucleotides were synthesized by Integrated DNA Technologies or by the W.M. Keck Facility at Yale University (see

Scheme 2.2 for DNA sequences). The oligonucleotides were desalted using Sephadex G-25 and purified using denaturing polyacrylamide gel electrophoresis as previously described (23). Oligonucleotides for gel-based assays were labeled on the lesion containing strand with a 5'-fluorescein (6-fam) label. The concentrations of the single-stranded oligonucleotides were determined from the absorbance at 260 nm, using the calculated extinction coefficients for all oligonucleotides except those containing εA. For oligonucleotides containing εA, the extinction coefficient was calculated for the same sequence with an A in place of the εA and corrected by subtracting 9400 $M^{-1}cm^{-1}$ to account for the weaker absorbance of εA as compared to A (24). The lesion-containing oligonucleotides were annealed with a 1.2-fold excess of the complement by heating to 90 °C and cooling slowly to 4 °C. Two different sequence contexts were used and these are referred to by the central three nucleotides (i.e., TEC has a central εA lesion flanked by a 5'T and a 3'C; Scheme 2.2).

TXC 5′ –CGATAGCATCCT**X**CCTTCTCTCCAT
3′ –GCTATCGTAGGATGGAAGAGAGGTA

AXA 5′ –CGATAGCATCAA**X**AATTCTCTCCAT
3′ –GCTATCGTAGTTTTTAAGAGAGGTA

X= εA, inosine, or pyrrolidine

Scheme 2.2

Steady-State Fluorescence Measurements. Fluorescence emission spectra were collected with a PTI QuantaMaster fluorometer controlled by FeliX software. For εA fluorescence, an excitation wavelength of 320 nm (6 nm band-pass) was used and the total fluorescence was measured at emission wavelengths from 350-500 nm (6 nm band-pass). Samples (300 μL) of 400 nM εA-containing DNA were prepared in the standard buffer (50 mM NaMES (pH 6.5), 100

mM NaCl, 1 mM EDTA, 1 mM DTT) and spectra were recorded at 25 °C. To determine the steady-state fluorescence of εA-containing DNA bound to AAG, the spectra were recorded within one minute. No significant excision of εA occurs during this time. Three independent titrations were averaged, and the data were fit to a quadratic equation assuming tight binding by AAG (Equation 2.1), in which F_{rel} is the relative fluorescence, A is the fractional quenching, and K_d is the dissociation constant for the εA-DNA.

$$F_{rel} = 1 - \frac{A(K_d + E + DNA) - \sqrt{(K_d + E + DNA)^2 - (4E \times DNA)}}{2E} \quad (2.1)$$

Single-turnover Excision of εA Monitored by Fluorescence. Experiments with mutant and wild-type AAG were performed as previously described (14). Since εA is strongly quenched in duplex DNA and when bound to AAG, we were able to follow the release of εA into solution by the increase in fluorescence. Single-turnover glycosylase assays were performed in the standard buffer at 25 °C with 400 nM DNA duplex and 1.2–2.4 μM AAG. A full emission spectrum was recorded at various times between 2 and 120 minutes. The greatest change in fluorescence occurred at 405 nm, therefore this wavelength was chosen to follow the release of εA into solution. The fluorescence at 405 nm was used to determine the fraction of εA product (Fraction Product = $(F_t - F_0)/F_{max}$; F_t is the fluorescence at time = t, F_0 is the initial fluorescence at time = 0 and F_{max} is the maximal fluorescence change). This normalization gives fraction product values between 0 and 1. These data were fit by a single exponential equation using nonlinear least squares regression with Kaleidagraph (Synergy Software), in which k_{obs} is the rate constant, t is the time, and A is the amplitude (Equation 2.2). In all cases, the same rate constant was measured at two different concentrations of AAG. This demonstrates that AAG

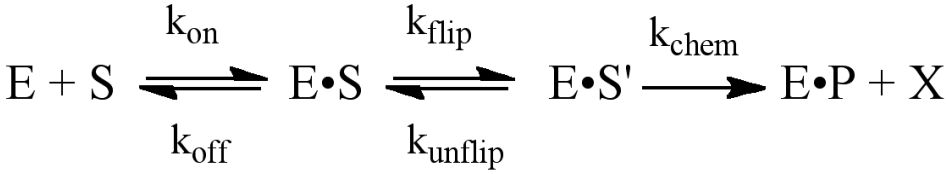
was saturating and that the maximal single-turnover rate constant was determined (see Appendix A). Under these conditions, the observed rate constant is equal to the maximal single-turnover rate constant ($k_{obs} = k_{max}$).

$$\text{Fraction product} = A(1 - e^{-k_{obs}t}) \qquad (2.2)$$

Gel-based Glycosylase Assay. We also measured single-turnover glycosylase activity with the standard glycosylase activity assay that utilizes abasic site cleavage by NaOH followed by DNA separation on a denaturing polyacrylamide gel (22, 23). Fluorescein-labeled DNA substrates (50 nM) containing either εA or inosine (I) were prepared in the standard buffer. The reactions were initiated with the addition of 75 nM – 5 μM AAG and incubated at 25 °C. At various time points, a sample from the reaction was removed and quenched in 2 volumes of 0.3 M NaOH, giving a final hydroxide concentration of 0.2 M. The abasic sites were cleaved by heating at 70 °C for 10 minutes. Samples were mixed with an equal volume of formamide/EDTA loading buffer before loading onto a 15% polyacrylamide gel. Gels were scanned with a Typhoon Imager (GE Trio+ Healthcare) to detect the fluorescein label by exciting at 488 nm and measuring emission with a 520BP40 filter. The gel bands were quantified using ImageQuant TL (GE Healthcare). The data was converted to fraction product [Fraction Product = product / (product + substrate)] and then fit by a single exponential (Equation 2.2). The observed rate constant for the single-turnover reaction was independent of the concentration of AAG, indicating that the maximal rate constant was measured ($k_{obs} = k_{max}$; see Appendix A).

Assignment of Single-turnover Rate Constants. Glycosylases require a nucleotide flipping step to gain access to the labile bond and therefore a saturating single-turnover rate constant could be limited by either nucleotide flipping or by hydrolysis of the *N*-glycosidic bond. This reaction

scheme is shown below (Scheme 2.3) and the contribution of individual rate constants is given in Equation 2.3. For wild-type AAG-catalyzed excision of εA we have previously established that the flipping rate constant (k_{flip}) is much faster than the rate of *N*-glycosidic bond hydrolysis



Scheme 2.3

(k_{chem}). Furthermore, we found that the equilibrium constant for nucleotide flipping (K_{flip}; Equation 2.4) is highly favorable (14). Therefore, for AAG-catalyzed excision of εA the observed single-turnover rate constant is simply k_{chem}. The analysis herein suggests that this is true for the Y127W and Y159W mutants as well. However, for AAG-catalyzed excision of hypoxanthine (Hx), the rates of flipping are unknown and the inverse correlation between base pair stability and excision rate strongly suggests that the equilibrium for nucleotide is unfavorable for this lesion (3, 25). Therefore, the value of k_{max} for wild-type and mutant AAG should be considered an upper limit for the value of k_{chem} (Equation 2.5).

$$k_{max} = k_{chem} \times k_{flip}/(k_{flip}+k_{unflip}) \quad (2.3)$$

$$K_{flip} = k_{flip}/k_{unflip} \quad (2.4)$$

$$k_{max} = k_{chem} (K_{flip}/(1+K_{flip})) \quad (2.5)$$

Stopped-Flow Kinetics. Pre-steady state kinetic experiments were performed on a Hi-Tech SF-61DSX2, controlled by Kinetic Studio (TgK Scientific). The fluorescence of εA was measured using an excitation wavelength of 320 nm and a WG360 long-pass emission filter as previously described (14). The fluorescence of tryptophan was measured using an excitation wavelength

of 295 nm and a 330BP20 band-pass emission filter. Due to the weak fluorescence of εA and tryptophan, and the fast association rate constant, we monitored binding with equimolar concentrations of εA-DNA and protein. Under these conditions the time-dependent changes in fluorescence was fit by Equation 2.6, in which *F* is εA fluorescence as a function of time, *C* is the fluorescence of the free DNA, *Y* and *Z* are the changes in fluorescence of the first and second intermediates relative to the free DNA, E_0 is the starting concentration of AAG and DNA ([AAG] = [DNA]), k_1 is the bimolecular rate constant for binding (k_{on}; Scheme 2.3), $k_{2,obs}$ is the observed unimolecular rate constant for the subsequent flipping step, and *t* is the time (14). For an approach to equilibrium, the observed rate constant is equal to the sum of the rate constants for the forward and reverse flipping steps (Equation 2.7). However, in the case of a favorable equilibrium constant for flipping ($k_{flip} >> k_{unflip}$), the observed rate constant for the forward reaction is approximately equal to the flipping rate constant (Equation 2.8). Dissociation experiments were performed for the mutant proteins to validate this assignment (see below).

$$F = C + Y\left(\frac{E_o^2 k_1 t}{1 + E_o k_1 t}\right) - Z(1 - e^{-k_{2,obs} t}) \qquad (2.6)$$

$$k_{2,obs} = k_{flip} + k_{unflip} \qquad (2.7)$$

$$k_{2,obs} \approx k_{flip} \qquad (2.8)$$

Several different concentrations between 0.125 μM and 1.0 μM (all concentrations are given after mixing) gave the same rate constants, confirming that the concentration chosen was saturating for the mutant and wild-type proteins (see Appendix A and data not shown). At least three traces were averaged together for each concentration in two independent experiments, and the averaged data was fit by Equation 2.6 using Kaleidagraph.

Pulse-Chase Assay to Measure Substrate Dissociation. The macroscopic rate constant for dissociation of wild-type and mutant AAG from εA-containing DNA was measured by pulse-chase in the standard reaction buffer at 25 °C as previously described for wild-type AAG (14). Briefly, in 20 µL reactions, 50 nM fluorescein-labeled TEC DNA was mixed with 75 nM or 150 nM AAG for 20 seconds, and then a chase of 10 µM unlabeled TEC DNA was added. At various time points, a sample from the reaction was removed and analyzed as described for "Gel-based glycosylase assay". Base excision catalyzed by AAG results in fluorescein-labeled product, whereas dissociation releases unreacted fluorescein-labeled substrate. The partitioning between hydrolysis and dissociation can be determined from either the exponential rate constant or by the change in burst amplitude. If AAG dissociates from the labeled DNA before the chemical cleavage step and then binds to the unlabeled DNA, less of the reaction will occur during the single-turnover part of the curve as compared to the same experiment without chase. The data was converted to fraction product and fit by a single exponential (Equation 2.2). The observed rate constant was independent of the concentration of AAG.

According to the two step binding mechanism described in Scheme 2.3, two different partitioning equations can be written (26). All labeled substrate is initially bound, and therefore the fraction of product formed is given by the fraction that goes on to react. This is indicated by Equation 2.9, in which A is the burst amplitude (the fraction of product formed in the burst phase of the experiment), k_{max} is the maximal single-turnover rate constant for formation of product, and $k_{off,obs}$ is the macroscopic rate constant for dissociation from the flipped-out complex. This expression can be rearranged to solve for the desired dissociation rate constant (Equation 2.10). Similarly, for branched pathways, the observed rate constant for the burst

phase of the pulse-chase experiment is given by the sum of the rate constants for the competing pathways, formation of product is given by k_{max} and the macroscopic dissociation of substrate is designated $k_{off,obs}$ (Equation 2.11). Solving for $k_{off,obs}$ gives Equation 2.12.

$$A = \frac{k_{max}}{k_{off,obs} + k_{max}} \quad (2.9)$$

$$k_{off,obs} = \frac{k_{max}}{A} - k_{max} \quad (2.10)$$

$$k_{obs} = k_{off,obs} + k_{max} \quad (2.11)$$

$$k_{off,obs} = k_{obs} - k_{max} \quad (2.12)$$

Control reactions in which no chase was added provided the single-turnover rate constant, k_{max}, and confirmed that these concentrations of AAG were saturating. From these values, the dissociation rate constant, k_{off}, for AAG dissociating from εA-DNA was calculated by two different methods (Equations 2.10 and 2.12). Both methods gave very similar values for $k_{off,obs}$ and we report the average of the results obtained with both methods.

Since the εA-DNA binds in two steps, the observed rate constant for dissociation of substrate ($k_{off,\ obs}$) could be limited by the unflipping rate (k_{unflip}) or dissociation from nonspecific DNA (k_{off}). According to Scheme 2.3, and assuming that the flipped-out complex is stable (i.e., $k_{flip} >> k_{unflip}$), this observed dissociation rate constant can be expressed in terms of the microscopic rate constants (Equation 2.13; (26)). Stopped flow fluorescence suggests that dissociation from the initial AAG•DNA complex is rapid (k_{off} ~30 s^{-1}) relative to the forward flipping rate constant k_{flip} = 2.6 s^{-1}, and therefore the observed rate constant for substrate

dissociation from the εA-DNA•AAG complex is approximately equal to the reverse rate constant for flipping (Equation 2.14).

$$k_{off,obs} = k_{unflip}\left(\frac{k_{off}}{k_{off} + k_{flip}}\right) \quad (2.13)$$

$$k_{off,obs} \approx k_{unflip} \quad (2.14)$$

Direct Measurement of the Macroscopic Rate for Substrate Dissociation. Stopped-flow double-mixing experiments were performed to measure unflipping and dissociation of εA-DNA directly. To rapidly form the flipped-out AAG-DNA complex, 2 μM AEA DNA was mixed with 2.8 μM AAG. After an aging time of one second, 18 μM pyrrolidine-DNA was added as a competitor. The final concentrations after mixing were 500 nM AEA DNA, 700 nM AAG, and 9 μM pyrrolidine-DNA. The fluorescence of εA was measured using an excitation wavelength of 320 nm and a WG360 long-pass emission filter. The reaction was followed for 100 seconds; no significant excision of εA occurs during this time. Three traces were averaged together for each mutant, and the normalized fluorescence was fit by a single exponential (Equation 2.2). As described above for the pulse-chase assay, the observed rate constant for dissociation is described by Equation 2.13 and simplifies to Equation 2.14 under the conditions of $k_{off} >> k_{flip}$.

Free Energy Reaction Profile. The energetic effects of the Y127W and Y159W mutations can be illustrated by constructing a free energy reaction profile for the binding, flipping, and excision of an εA lesion by wild-type and mutant AAG. A standard state of 25 ^{o}C, pH 6.5, 1 nM enzyme, and 1 μM εA-DNA was chosen. The E•S state was arbitrarily set to ΔG=0 for all enzymes and the individual heights of barriers were determined by Equation 2.15, in which *R* is

the gas constant (1.987×10^{-3} kcal/molK), T is the temperature (298 K), k is the rate constant, h is Planck's constant (1.58×10^{-37} kcal s), and k_B is the Boltzmann constant (33×10^{-27} kcal/K).

$$\Delta G^{\ddagger} = -RT\ln(hk/k_B T) \qquad (2.15)$$

Results

Tryptophan Fluorescence of Wild-type AAG

The catalytic domain of AAG (residues 80-298) has almost identical glycosylase activity as the full-length protein (22, 27, 28), but is considerably more stable. Therefore, we used this construct for all of the experiments described below. To gain insight into the steps involved in nucleotide flipping we sought to establish fluorescent probes that could report on ligand binding and conformational changes in AAG. Although it is possible to use the natural

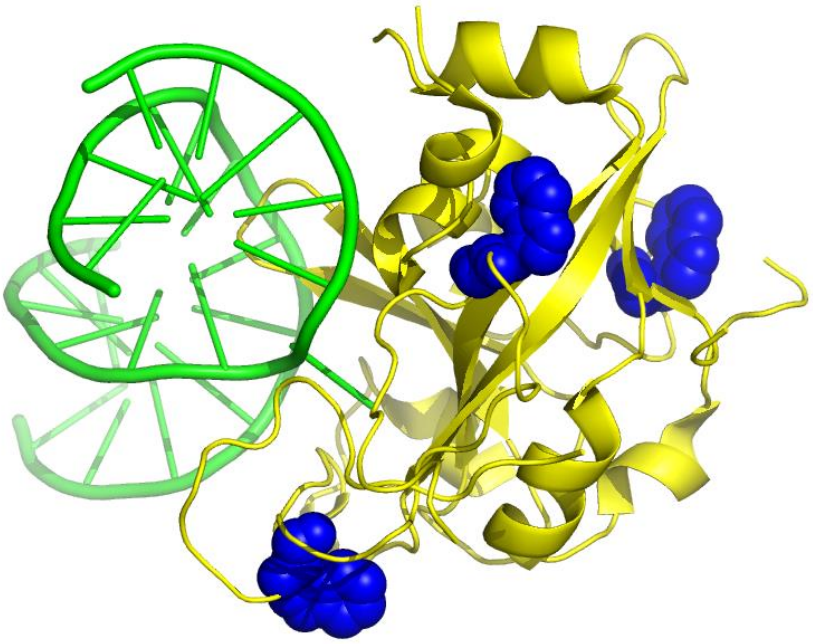


Figure 2.2. Location of tryptophan residues in the catalytic domain of AAG. Cartoon representation of the AAG (yellow) complex with εA-DNA (green) from the PDB (1EWN) was rendered with Pymol (http://www.pymol.org). The three native tryptophan residues are shown in space-filling representation (blue). They are all located near the surface of the protein far away from the DNA-binding surface of AAG.

fluorescence of εA to characterize the interaction of AAG with this alkylative lesion, the majority of substrates recognized by AAG are not fluorescent (14). The catalytic domain of AAG contains three tryptophan residues, raising the possibility that the native protein might undergo changes in its intrinsic fluorescence upon binding to DNA. The crystal structure of AAG bound to εA-DNA shows that all three of these residues are located near the surface on the opposite side of the enzyme from the DNA binding surface (Figure 2.2; (12)). We initially tried to remove these tryptophan residues to simplify the fluorescence spectra when additional tryptophan residues were added to the protein. Each individual tryptophan could be mutated singly to phenylalanine and full glycosylase activity was maintained, however attempts to combine these into a triple mutant resulted in insoluble protein in *E. coli* (data not shown). Before embarking on more extensive mutagenesis studies, we tested whether the native protein undergoes any changes in tryptophan fluorescence upon binding to εA-containing DNA. Whereas binding could be readily detected by changes in the εA fluorescence (excitation at 320 nm and emission at 400 nm), parallel experiments in which tryptophan fluorescence was measured gave no detectable changes in fluorescence (excitation at 295 nm and emission at 330 nm; see Appendix A). This suggests that the native tryptophan residues do not change their environment upon DNA binding and nucleotide flipping. Therefore, the changes in tryptophan fluorescence for the Y127W and Y159W mutants simply reflect changes in the local environment of the active site tryptophan residue.

Interaction of Y127W and Y159W AAG with εA-DNA

We prepared two different mutants of AAG in which one of the active site tyrosines was mutated to tryptophan. These tyrosine residues, Y127 and Y159, interact with the flipped-out

base in the crystal structure of AAG bound to εA-DNA (13). As tyrosine to tryptophan substitutions are often tolerated (20), we investigated the possibility that these mutants might exhibit tryptophan fluorescence that reports on the nucleotide flipping step. This would enable us to correlate changes in the fluorescence of εA-DNA with protein fluorescence. It was previously shown that εA-DNA of the sequence AEA shows a large decrease in fluorescence

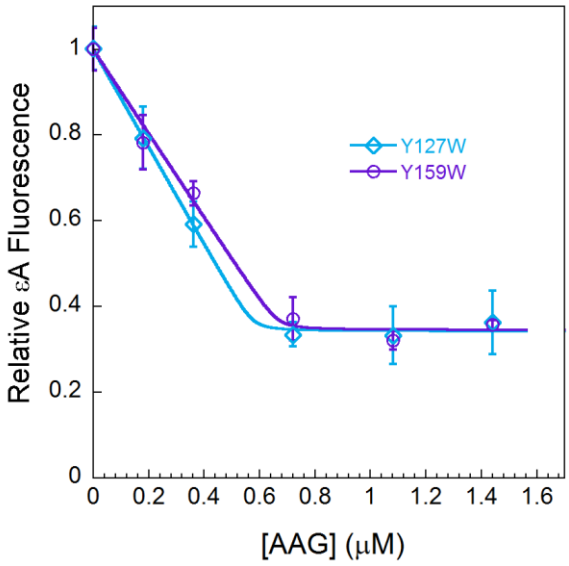


Figure 2.3. Titration of εA-DNA with AAG mutant proteins. Samples of 400 nM εA-DNA were incubated with increasing concentration of either Y127W (◇) or Y159W (○) AAG and the εA fluorescence was recorded. Three independent titrations were performed and the average and standard deviation are shown. The normalized fluorescence was fit by the quadratic equation under conditions of tight binding (Equation 2.1). These fits indicate that the number of εA-binding sites is 70% of the expected value for Y127W and 60% of the expected value for Y159W.

upon binding to wild-type AAG (see Materials and Methods; (14)). Fluorescence titrations were performed with a fixed concentration of εA-DNA and increasing concentration of mutant AAG proteins to determine the concentration of active AAG (Figure 2.3; see Materials and Methods). The active concentration determined by this method was within 60% of the value calculated from the UV-absorbance of the mutant proteins and this corrected concentration was used for all of the experiments described below.

Glycosylase Activity of AAG Mutants Toward εA-DNA

We next measured the single-turnover glycosylase activity of the wild-type and mutant proteins. Since εA is strongly quenched in duplex DNA, and becomes fluorescent upon its release into solution, we monitored the fluorescence of εA-DNA incubated with either wild-type or mutant forms of AAG under single-turnover conditions, with enzyme in excess over substrate and far above the K_d for substrate binding (Figure 2.4). Two different sequence contexts were tested, a heterogenous sequence (TEC) and a polyA sequence (AEA; see Scheme

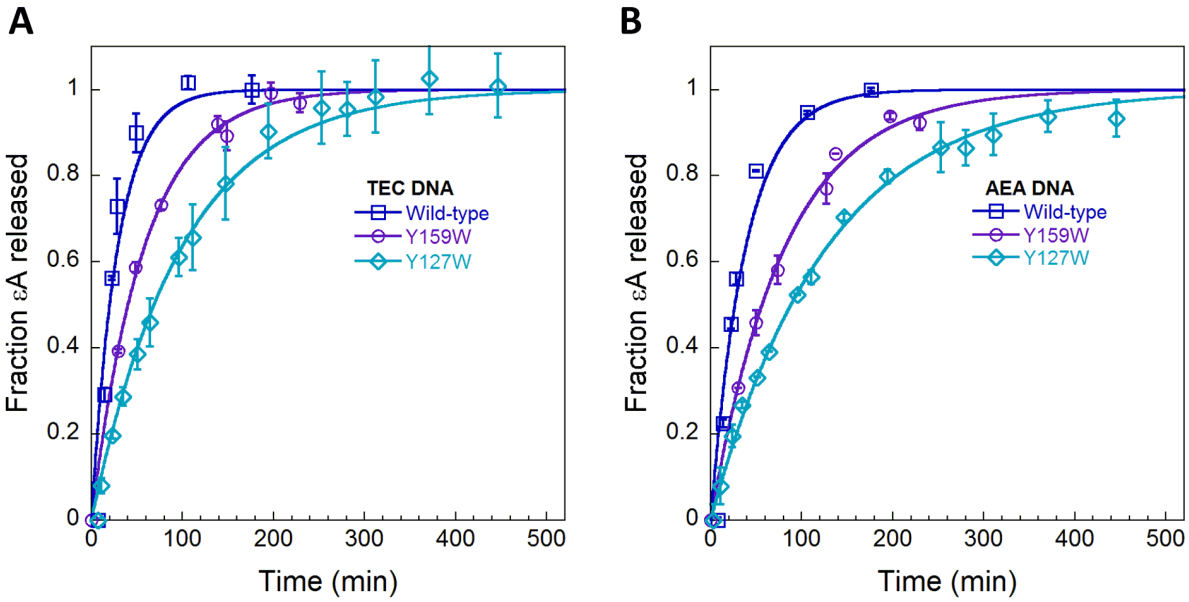


Figure 2.4. Single-turnover excision of εA by wild-type and mutant AAG. The reaction time course for wild-type (□), Y159W (○), and Y127W (◇) AAG are shown for two different sequences, TEC (**A**) and AEA (**B**). The fraction of εA released was determined by measuring the increased fluorescence of εA when it is released into solution (see Materials and Methods; (14)).

2.2; (14)). The analysis of εA excision from the TEC duplex by wild-type AAG gave a single-turnover rate constant of 0.048 min^{-1}, similar to the previously reported value of 0.034 min^{-1} (14). The Y159W mutant is 3-fold slower than wild-type and the Y127W mutant is 4-fold slower (Table 2.1). These effects are similar to the respective 3-fold and 8-fold effects of these

Table 2.1. Kinetic parameters for flipping and excision of damaged bases by wild-type and mutant AAG[a]

	Wild-type[b]	Y159W	Y127W
k_{on} ($M^{-1}s^{-1}$)	2.0×10^{8}	3.1×10^{8}	6.8×10^{7}
k_{flip} (s^{-1})	2.6	13.6	1.1
k_{unflip} (s^{-1})	0.002	0.37	0.12
K_{flip}[c]	1300	40	9
$k_{chem\ \varepsilon A}$ (s^{-1})[d]	8.0×10^{-4}	3.0×10^{-4}	2.0×10^{-4}
$k_{chem\ Hx}$ (s^{-1})[e]	4.9×10^{-2}	2.8×10^{-4}	8.7×10^{-5}

[a]The binding and flipping rate constants were determined from stopped-flow fluorescence using the TEC oligonucleotide. The values shown are from the changes in εA fluorescence, but the tryptophan fluorescence of Y159W and Y127W gave almost identical values (see Appendix A). The standard conditions were 25 °C, 50 mM NaMES (pH 6.5), 100 mM NaCl, 1 mM EDTA, and 1 mM DTT.

[b]The values for wild-type AAG were previously reported (14).

[c]The equilibrium constant for flipping is given by the ratio of the flipping and unflipping rate constants ($K_{flip} = k_{flip}/k_{unflip}$).

[d]The rate constants for single-turnover excision of εA from TEC DNA are from Figure 2.4A, and reflect the rate of *N*-glycosidic bond cleavage.

[e]The rate constants for single-turnover excision of Hx from the same TEC sequence context are from Figure 2.8, and are an underestimate of the rate of AAG-catalyzed *N*-glycosidic bond hydrolysis, because this step is preceded by an unfavorable flipping equilibrium (Equation 2.5).

mutations on the single-turnover excision of εA measured by abasic site formation under different reaction conditions (22). Similar results were obtained for the AEA sequence context with 3-fold and 5-fold decreases in the single-turnover rate constant for Y159W and Y127W, relative to wild-type AAG. The relatively modest decreases in the single-turnover glycosylase activity raised the possibility that these introduced tryptophan residues could serve as useful reporters on the rates of conformational changes within the protein. For wild-type AAG, the maximal single-turnover rate constant is likely to be *N*-glycosidic bond hydrolysis since the equilibrium constant for flipping is highly favorable. However, the maximal single-turnover rate constant could be limited by either the nucleotide flipping step or the *N*-glycosidic bond cleavage step for these mutants (Scheme 2.3).

Stopped-Flow Fluorescence of εA

To evaluate whether the Y127W or Y159W mutants alter the rate constants for DNA binding or nucleotide flipping, we performed rapid mixing experiments with εA-DNA. We used equimolar concentrations of DNA and protein far above the dissociation constant for DNA binding to monitor fluorescence changes in εA, as previously described for wild-type AAG (14). This experiment provides optimal signal for both binding and flipping steps and is necessary given the fast binding of AAG and the weak fluorescence of εA. Representative data are shown in Figure 2.5 (A & C). Both mutants exhibit transient increases and subsequent decreases in εA fluorescence similar to the two-step binding that was observed for wild-type AAG. We previously found that the first phase of the binding reaction is concentration-dependent, and reflects the initial binding of AAG to the site of damage. The second phase of the binding reaction is independent of DNA concentration and corresponds to the flipping of the εA lesion

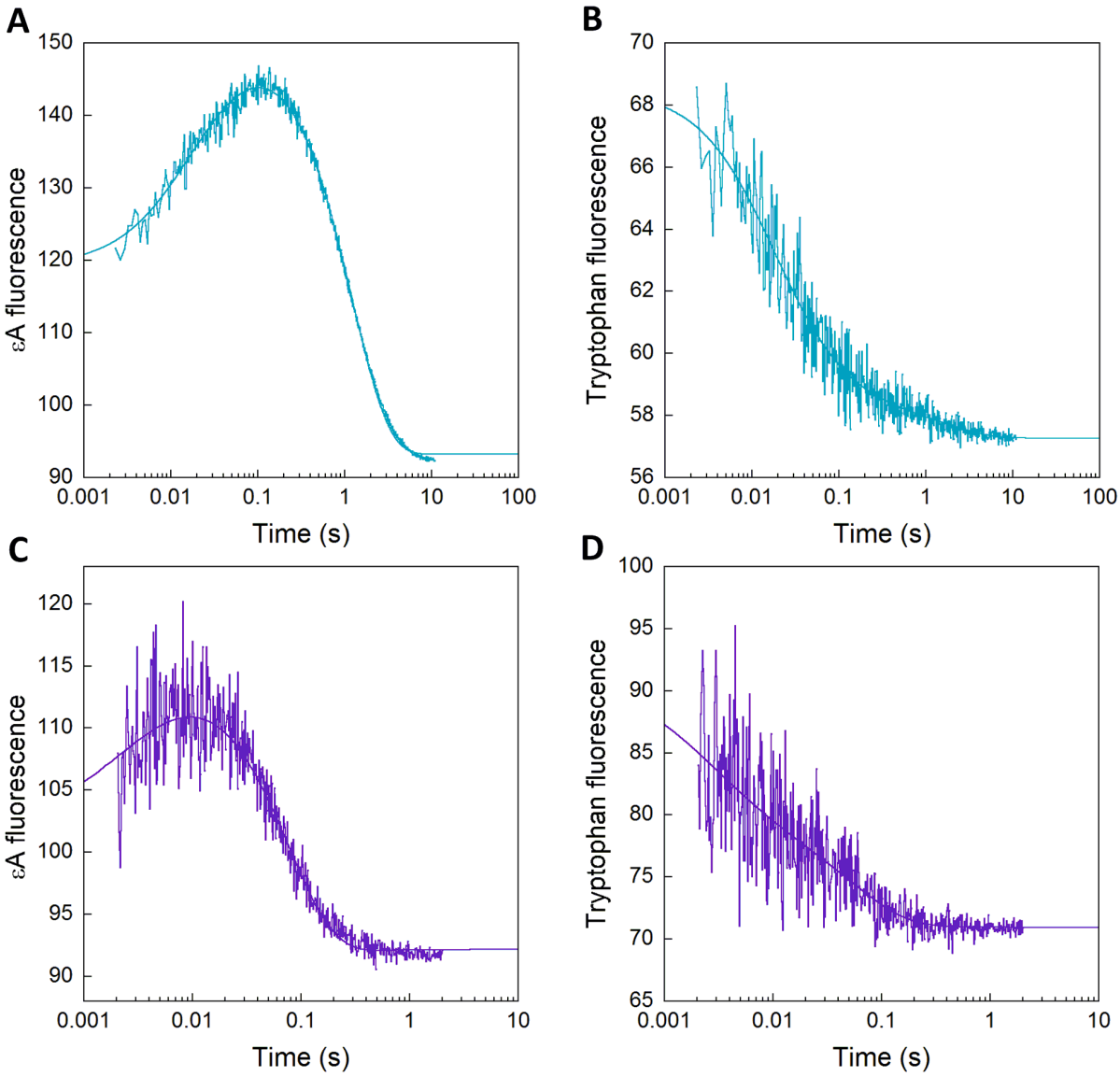


Figure 2.5. **Stopped-flow fluorescence to measure binding and nucleotide flipping by Y127W and Y159W mutants of AAG.** Equimolar mixtures of AAG and εA-DNA were mixed in the stopped-flow (1 μM after mixing) and either εA fluorescence or tryptophan fluorescence was measured as described in the Materials and Methods. (**A**) εA fluorescence for Y127W. (**B**) Tryptophan fluorescence for Y127W. (**C**) εA fluorescence for Y159W. (**D**) Tryptophan fluorescence for Y159W. The data is the average of three independent reactions and the lines indicate the best fits of Equation 2.6. The rate constants are given in Table 2.1.

into the active site pocket (14).

For the Y127W mutant, the initial binding step yielded an observed rate constant that is similar to the rate constant for wild-type AAG under these conditions (Table 2.1). In contrast, the second step is approximately 3-fold slower for the mutant than for wild-type AAG. This observed rate constant is equal to the sum of the rate constants for the flipping and unflipping steps, because it is an approach to equilibrium (Equation 2.7). However, the results from dissociation experiments suggest that the rate constant for flipping is significantly faster than the rate constant for unflipping (see below). This indicates that the observed rate constant for the second phase of the reaction is approximately equal to the microscopic rate constant for flipping (k_{flip}). Therefore, the Y127W mutation decreases the rate constant for flipping by a factor of 3-fold.

The Y159W mutant showed a lower level of εA fluorescence in the transient formation of the initial intermediate. Given this decrease in signal it was more difficult to accurately measure the rate constant for the initial phase of the binding reaction, but the best fits yielded a binding rate constant that is also similar to that of wild-type and the Y127W mutant (Table 2.1). Remarkably, the observed rate constant for the second phase of the reaction (flipping) is more than 5-fold faster for the Y159W mutant than for the wild-type protein. This places an upper limit on the rate constant for flipping, but does not provide sufficient information to establish whether or not this mutation increases the rate constant for nucleotide flipping. This faster observed rate constant could also be explained by a 6000-fold increase in the reverse rate constant for flipping with no change in the forward rate constant for flipping. Therefore,

additional experiments were performed to resolve this ambiguity and to establish that the Y159W mutation does increase the rate constant for nucleotide flipping (see below).

Stopped-Flow Fluorescence of Tryptophan

We also followed the fluorescence of the tryptophan during the rapid mixing experiments, and the results are shown in Figure 2.5 (B & D). Control reactions in which the tryptophan fluorescence was recorded for wild-type AAG mixed with εA-DNA (295 nm excitation, 330/20 nm emission) established that none of the three tryptophan residues present in AAG are sensitive to DNA binding (see Appendix A). When the Y127W and Y159W mutants were mixed with εA-DNA, two phases of decreasing fluorescence were observed (Figure 2.5 B & D). These data were fit by the same equation used to monitor εA fluorescence (Equation 2.6), and yielded almost identical rate constants as were observed for εA fluorescence changes. These experiments reveal that tryptophan at position 127 and 159 in the active site pocket are perturbed in the initial AAG•DNA complex, relative to when AAG is free in solution. The tryptophan quenching in the first phase, corresponding to formation of an initial recognition complex, is coincident with the increased fluorescence of εA, suggesting that the protein changes its local conformation as fast or faster than the DNA conformational change that is monitored by the εA fluorescence. In the second step, tryptophan residues at either position 127 or 159 are quenched in the flipped-out complex with the same rate constant as the εA is quenched. This correlation is consistent with the simple model that the εA lesion interacts with the tyrosine or tryptophan residues at positions 127 and 159 (Figure 2.1).

Pulse-Chase Experiments to Determine the Rate of Substrate Dissociation

As the nucleotide flipping step is an approach to equilibrium, it is important to measure the kinetics of the reaction in the reverse direction. Therefore, we performed pulse-chase experiments in which wild-type or mutant AAG proteins were initially mixed with fluorescently labeled εA-DNA and then chased with excess unlabeled εA-DNA (Figure 2.6). This experiment allows the partitioning between dissociation and base excision to be directly determined (see Materials and Methods). The wild-type protein gave the same results as previously reported, with 30% of the substrate partitioning forward to product and 70% dissociating. The dissociation of AAG from nonspecific DNA is relatively fast, and therefore the observed rate

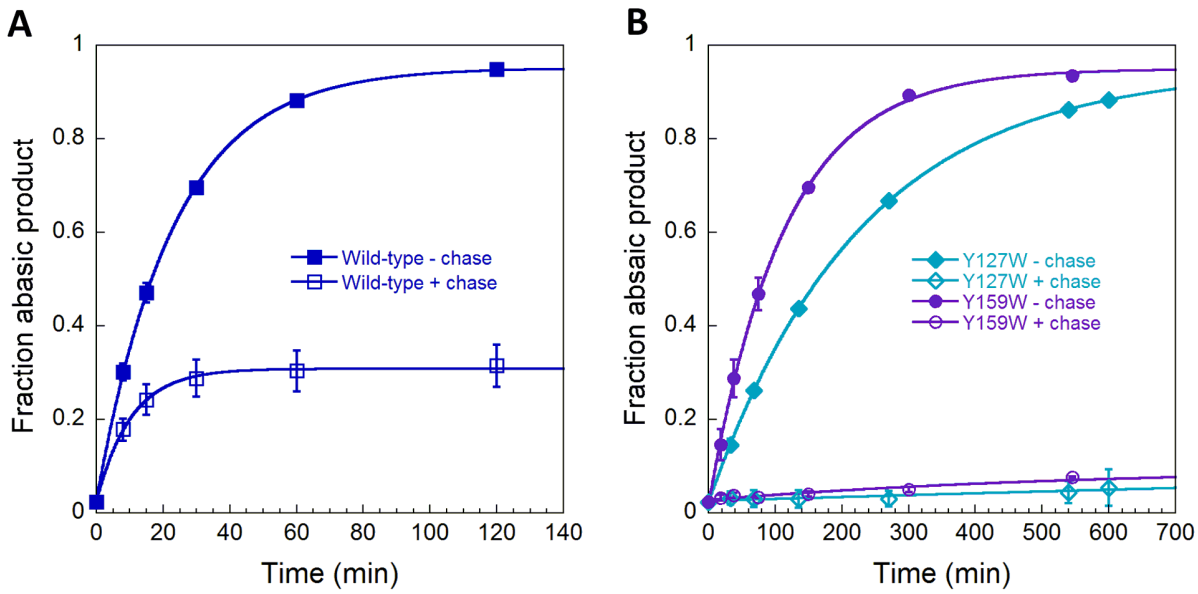


Figure 2.6. Pulse-chase experiments to measure dissociation of εA-DNA. The partitioning between the forward reaction (base excision) and substrate dissociation were measured for wild-type, Y127W, and Y159W AAG. Either wild-type (**A**) or mutant (**B**) AAG was mixed with fluorescein-labeled εA-DNA in the presence (open symbols) or absence (closed symbols) of excess unlabeled εA-DNA as described in the Materials and Methods. The fraction of fluorescent abasic DNA product was quantified using the standard gel-based glycosylase assay. For the wild-type enzyme approximately 30% of the εA-DNA is committed to excision. In contrast, essentially all of the εA-DNA dissociated from both Y127W (◇) and Y159W (○) mutant proteins under these conditions, indicating the k_{unflip} is much greater than k_{chem} (Scheme 2.3).

constant is simply the rate constant for unflipping [$k_{obs} = k_{unflip} = 1.6 \times 10^{-3}$ s^{-1}; (14)]. In contrast, the addition of chase completely blocked the reaction by the Y127W and Y159W mutants (Figure 2.6B). This indicates that the rate of substrate dissociation is much faster than the rate of *N*-glycosidic bond cleavage and places a lower limit on the rate of εA dissociation of 20-fold faster than the rate of *N*-glycosidic bond cleavage, assuming that 5% abasic product could have been detected. However, a much bigger effect on the unflipping rate could have been masked. Therefore, we adopted an alternative approach to measure the rate of substrate dissociation, and to determine the rate constant for the unflipping step with these mutant enzymes.

Double-mixing Experiments to Measure Unflipping and Dissociation of εA-DNA

We used double-mixing experiments to rapidly form the flipped-out AAG-DNA complex with the AEA substrate, and then challenge it with a tight-binding pyrrolidine-DNA competitor. Pyrrolidine is positively charged at the pH of these reactions and binds tightly to AAG as a transition state analog (12, 29). The change in fluorescence is adequately fit by a single exponential increase for both the Y127W and Y159W mutant proteins (Figure 2.7). We interpret these data as a rate-limiting unflipping step, followed by rapid dissociation of the nonspecific DNA. The small deviations from a simple exponential could be explained by a small amount of rebinding of εA-DNA. Consistent with this interpretation, experiments performed with a lower concentration of pyrrolidine-DNA chase resulted in a larger deviation from single exponential behavior (data not shown). The values of 0.12 s^{-1} for Y127W and 0.37 s^{-1} for Y159W are 75-fold and 250-fold faster than the rate of εA-unflipping for the wild-type protein (Table 2.1). Nevertheless, these unflipping rate constants are significantly slower than the observed rate constants for flipping, indicating that the observed rate constant for εA and tryptophan

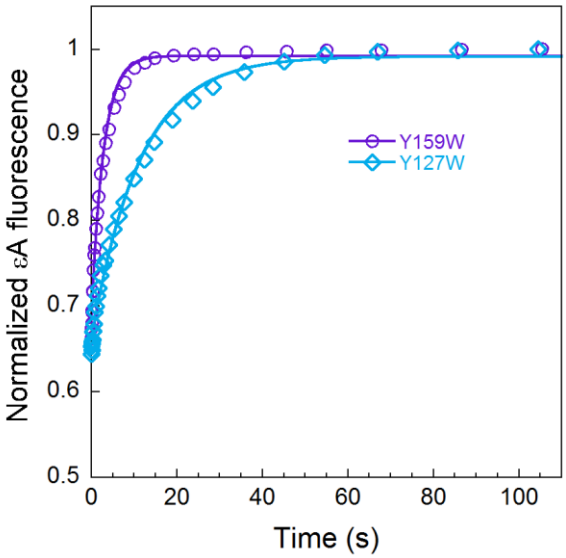


Figure 2.7. Direct measurement of εA-DNA dissociation from Y127W and Y159W AAG. Double-mixing experiments were performed using stopped-flow fluorescence. The AAG•DNA complex was formed, aged for one second, and then chased with an excess of pyrrolidine-DNA as a competitor. The normalized fluorescence was obtained by dividing the observed fluorescence signal by the fluorescence of the sample after the dissociation reaction was complete. The traces reflect the average of three separate mixing experiments and the solid lines depict the best fit to a single exponential. Only one out of every twenty data points is shown for clarity. The observed rate constants for dissociation are 0.12 s^{-1} for Y127W (◇) and 0.37 s^{-1} for Y159W (○) mutant proteins.

quenching that we have assigned to the nucleotide flipping step is approximately equal to the microscopic rate constant for flipping (Equation 2.8). The ratio of the rate constants for flipping and unflipping yields the equilibrium constant for nucleotide flipping (Equation 2.4; Table 2.1).

Single-turnover Excision of Hx

The decreased stability of the flipped-out εA-DNA complex for the Y127W and Y159W mutants suggests that these mutations are likely to have a much larger deleterious effect on the flipping of more stably paired lesions. Therefore, we examined the single-turnover excision of Hx from the most favorable Hx•T context. Wild-type AAG catalyzes the excision of Hx with a rate constant that is 60-fold faster than the single-turnover excision of εA when T is the opposing base (Figure 2.8A; Table 2.1). This is very similar to the 54-fold difference that was reported

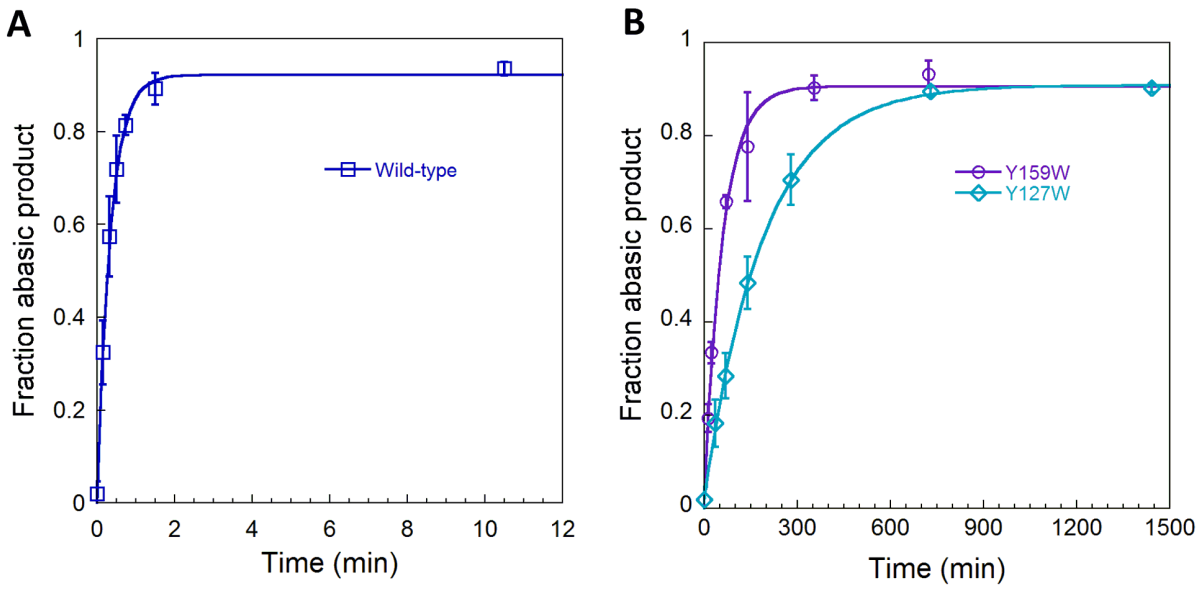


Figure 2.8. Single-turnover excision of Hx. The reaction progress curves are shown for wild-type (**A**), Y159W (○), and Y127W (◇) AAG (**B**). Note that the time scales are different for experiments with the wild-type and mutant proteins. Saturating amount of enzyme (1-5 μM) was always in excess over fluorescein-labeled DNA (50 nM). Reactions were followed in triplicate and the average values are shown; errors bars indicate one standard deviation. The lines indicate the best fit to a single exponential and the rate constants are reported in Table 2.1 (see Materials and Methods for details).

previously at a higher temperature (3). In contrast, the single-turnover rate constants for the AAG mutant proteins are essentially identical for excision of both εA and Hx (Figure 2.8B; Table 2.1). We believe this is a coincidence, because the single-turnover rate constants report on different steps for the excision of εA and Hx. For εA excision, the *N*-glycosidic bond cleavage step is likely to be rate-limiting. In the case of Hx excision, the observed rate constant is significantly smaller than the rate constant for the chemical step because there is an unfavorable equilibrium for nucleotide flipping (Equation 2.5). In the absence of direct measurements of nucleotide flipping for this lesion, it is also possible that the nucleotide flipping step becomes rate-limiting for these mutants. Importantly, the Y127W and Y159W mutations have severe (>100-fold) reductions in the observed rate constant for excision of Hx (Table 2.1).

Discussion

We mutated the two tyrosine residues in the AAG active site pocket to tryptophan in order to gain insight into the mechanism by which AAG recognizes DNA damage. Tryptophan was chosen as a conservative replacement with the intent of being able to directly monitor conformational changes and binding of base lesions in the active site. For both the Y127W and the Y159W mutant proteins, the changes in tryptophan fluorescence were coincident with the changes in εA fluorescence, supporting the previous assignments of the εA fluorescence changes to an initial recognition complex and to an excision-competent extrahelical complex (14). Although these mutants are efficient catalysts for the excision of εA lesions, the full kinetic analysis demonstrates that these mutations severely impair the equilibrium for flipping by AAG. The large perturbation of the equilibrium constant for nucleotide flipping explains why these mutations have a much larger deleterious effect on the excision of other lesions, such as Hx.

Functional Consequences of the Y127W and Y159W Mutations

We performed a pre-steady state kinetic analysis of the interaction of these mutant proteins with εA-DNA under the same conditions that were previously used to characterize wild-type AAG. These experiments provide an unambiguous kinetic characterization of the binding, flipping, and base excision steps (Scheme 2.3). The resulting microscopic rate constants are reported in Table 2.1, and can be used to calculate the free energy profile for the reactions catalyzed by wild-type and mutant forms of AAG (Figure 2.9). From this analysis it is apparent that the substitution of either tyrosine with tryptophan significantly destabilizes the active complex with the flipped-out εA lesion.

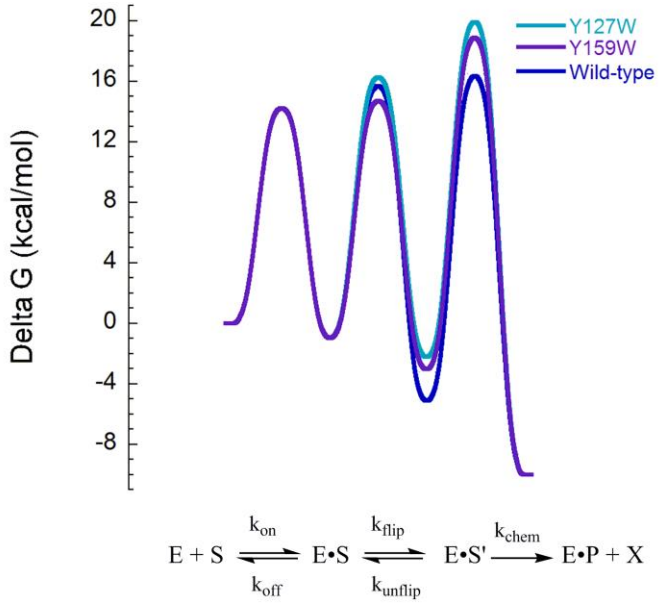


Figure 2.9. Free energy profile to illustrate the effects of the Y127W and Y159W mutations on the AAG-catalyzed excision of εA. The relative barrier heights for wild-type (dark blue), Y127W (light blue), and Y159W (purple) were calculated as described in the Materials and Methods. Whereas there are relatively large effects on the nucleotide flipping step and the stability of the flipped-out nucleotide, similar destabilization of the transition state results in much more modest effects on the rate of *N*-glycosidic bond hydrolysis (k_{chem}).

Perhaps the most remarkable finding is that the Y159W mutation conferred a 5-fold increase in the rate constant for flipping (Table 2.1). This could be due to a favorable interaction with the side chain of W159 that stabilizes the transition state for flipping of εA. The crystal structure of the wild-type protein shows a T-shaped interaction between Y159 and the εA lesion, and it may be possible for the slightly larger tryptophan to contact the lesion before it fully enters the nucleobase binding pocket (Figure 2.1). This faster flipping is countered by a much greater increase in the unflipping rate constant (200-fold; Table 2.1) and the overall equilibrium for nucleotide flipping is perturbed by 30-fold. This destabilization may be due to the loss of the hydroxyl moiety or to the slightly increased side chain volume of tryptophan

relative to tyrosine. In contrast, there is only a subtle effect on the rate constant for εA excision (3-fold; Table 2.1). As the equilibrium constant for flipping remains favorable, this suggests that the Y159W mutation has a very similar destabilizing effect in the ground state of the flipped-out complex as in the transition state for *N*-glycosidic bond cleavage (Figure 2.9).

The Y127W mutation destabilizes the flipped-out complex to a greater extent than the Y159W mutation. Whereas there is a modest 2.4-fold decrease in the rate constant for flipping, the unflipping rate constant is increased by 60-fold, resulting in 140-fold decrease in the stability of the flipped-out complex. The Y127W mutant has only a 4-fold decreased rate constant for εA excision, relative to wild-type, which is very similar to the 3-fold deleterious effect of the Y159W mutation (Table 2.1). We conclude that the increase in the size of Y127 from tyrosine to tryptophan does not greatly alter the transition state for flipping, but destabilizes the flipped-out εA-DNA complex by a similar extent in both the ground state and the transition state for *N*-glycosidic bond cleavage.

Wild-type AAG has a highly favorable equilibrium for flipping of εA, and therefore the Y127W and Y159W mutant proteins retain favorable equilibrium constants for flipping despite the deleterious effects of these substitutions. This results in relatively small effects on the overall rates of εA excision by the Y127W and Y159W mutants. It is likely that the inability of εA to form hydrogen bonds with an opposing base facilitates this highly favorable flipping equilibrium. In contrast to εA, most of the lesions recognized by AAG, including 7-methylguanine and Hx, do not appear to be bound stably in the flipped-out conformation by AAG (3, 25, 30). Consistent with this model, the Y127W and Y159W mutations have much larger deleterious effects of 520- and 140-fold for the excision of Hx from a Hx•T wobble pair (Table

2.1). These defects are larger by a factor of four than the defects in the flipping equilibrium for εA (Table 2.1), suggesting that the tryptophan mutations specifically compromise the positioning of Hx in the transition state for *N*-glycosidic bond cleavage.

Insight Into the Mechanism of AAG-catalyzed Nucleotide Flipping

Although the Y127W and Y159W mutations were more disruptive than expected, the characterization of these mutant proteins provided valuable insight into the mechanism of AAG-catalyzed nucleotide flipping. In spite of the large decreases in the stability of the flipped-out, pre-hydrolysis complex, the stopped-flow fluorescence experiments showed that these mutant proteins exhibit the same two-state binding model as was observed for wild-type AAG. The same rate constants were observed for tryptophan and εA fluorescence with both mutants (Figure 2.5). These observations indicate the εA lesion and tryptophan side chains located in the AAG active site both report on the initial recognition complex and on the fully flipped, active complex. Structural and biochemical studies of two distinct DNA glycosylases, uracil DNA glycosylase (UNG) and 8-oxoguanine DNA glycosylase (OGG1) have suggested that these glycosylases have a pre-binding site for prospective damaged bases that has been referred to as the exosite (31-34). The finding that two different fluorescent reporters in the AAG active site report on an intermediate that forms prior to nucleotide flipping is consistent with a similar initial recognition complex in which a partially flipped εA lesion interacts directly with the tryptophan side chains at the active site. However, we cannot exclude the possibility that other conformational changes in the protein or in the DNA (perhaps accompanying DNA bending) are responsible for the concerted changes in εA and tryptophan fluorescence.

Substitution of tyrosine with tryptophan is expected to be a relatively conservative mutation, since both side chains occupy similar volumes and can make similar stacking interactions. However, positions 127 and 159 in homologs of human AAG are invariably tyrosine. Extensive mutagenesis of AAG and selection for active site variants that conferred resistance to DNA alkylating agents has identified hundreds of allowed mutations, but no active variants at position 127 and only two active variants at position 159 were identified (5, 21). The finding that Y159N and Y159C had detectable activity is consistent with the finding that Y159W has less severe defects on nucleotide flipping than Y127W. Another study sought to identify AAG mutants that increased the frequency of mutations and identified Y127I as conferring a strong mutator phenotype (11).

We were surprised to find that the Y127W mutation has such severe defects in flipping, as a tryptophan side chain would be well-accommodated in the crystal structure of the εA-DNA complex (see Appendix A). Nevertheless, our results and the failure to observe active variants at this position point toward a very precise role in positioning the substrate in the flipped-out, active conformation. It is interesting that the Y127I mutation has been reported to exhibit only a 3-fold decrease in the single-turnover excision of εA, but a much greater decrease in the single-turnover excision of Hx (11). This is similar to the 4-fold defect in excision of εA that we measured for the Y127W mutant. It has been suggested that the mutator phenotype of Y127I is due to tighter binding to undamaged, bulged nucleotides that can occur from polymerase errors in repetitive sequences (11). It will be interesting to test whether the Y127I mutation changes the rate constants for flipping of normal bases, but this will require the development of new fluorescence-based assays.

References

1. Engelward BP*, et al.* (1997) Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. *Proceedings of the National Academy of Sciences of the United States of America* 94(24):13087-13092.

2. Larsen E, Meza TJ, Kleppa L, & Klungland A (2007) Organ and cell specificity of base excision repair mutants in mice. *Mutation research* 614(1-2):56-68.

3. O'Brien PJ & Ellenberger T (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase. *The Journal of biological chemistry* 279(11):9750-9757.

4. Hitchcock TM*, et al.* (2004) Oxanine DNA glycosylase activity from Mammalian alkyladenine glycosylase. *The Journal of biological chemistry* 279(37):38177-38183.

5. Guo HH, Choe J, & Loeb LA (2004) Protein tolerance to random amino acid change. *Proceedings of the National Academy of Sciences of the United States of America* 101(25):9205-9210.

6. Glassner BJ, Rasmussen LJ, Najarian MT, Posnick LM, & Samson LD (1998) Generation of a strong mutator phenotype in yeast by imbalanced base excision repair. *Proceedings of the National Academy of Sciences of the United States of America* 95(17):9997-10002.

7. Meira LB*, et al.* (2009) Aag-initiated base excision repair drives alkylation-induced retinal degeneration in mice. *Proceedings of the National Academy of Sciences of the United States of America* 106(3):888-893.

8. Rinne ML, He Y, Pachkowski BF, Nakamura J, & Kelley MR (2005) N-methylpurine DNA glycosylase overexpression increases alkylation sensitivity by rapidly removing non-toxic 7-methylguanine adducts. *Nucleic acids research* 33(9):2859-2867.

9. Trivedi RN*, et al.* (2008) Human methyl purine DNA glycosylase and DNA polymerase beta expression collectively predict sensitivity to temozolomide. *Mol Pharmacol* 74(2):505-516.

10. Hofseth LJ*, et al.* (2003) The adaptive imbalance in base excision-repair enzymes generates microsatellite instability in chronic inflammation. *J Clin Invest* 112(12):1887-1894.

11. Klapacz J*, et al.* (2010) Frameshift mutagenesis and microsatellite instability induced by human alkyladenine DNA glycosylase. *Mol. Cell* 37(6):843-853.

12. Lau AY, Scharer OD, Samson L, Verdine GL, & Ellenberger T (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. *Cell* 95(2):249-258.

13. Lau AY, Wyatt MD, Glassner BJ, Samson LD, & Ellenberger T (2000) Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG. *Proceedings of the National Academy of Sciences of the United States of America* 97(25):13573-13578.

14. Wolfe AE & O'Brien PJ (2009) Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by human alkyladenine DNA glycosylase. *Biochemistry* 48(48):11357-11369.

15. Kuznetsov NA*, et al.* (2005) Kinetics of substrate recognition and cleavage by human 8-oxoguanine-DNA glycosylase. *Nucleic acids research* 33(12):3919-3931.

16. Kuznetsov NA*, et al.* (2007) Pre-steady-state kinetic study of substrate specificity of Escherichia coli formamidopyrimidine--DNA glycosylase. *Biochemistry* 46(2):424-435.

17. Stivers JT, Pankiewicz KW, & Watanabe KA (1999) Kinetic mechanism of damage site recognition and uracil flipping by Escherichia coli uracil DNA glycosylase. *Biochemistry* 38(3):952-963.

18. Wong I, Lundquist AJ, Bernards AS, & Mosbaugh DW (2002) Presteady-state analysis of a single catalytic turnover by Escherichia coli uracil-DNA glycosylase reveals a "pinch-pull-push" mechanism. *The Journal of biological chemistry* 277(22):19424-19432.

19. Chelli R, Gervasio FL, Procacci P, & Schettino V (2002) Stacking and T-shape competition in aromatic-aromatic amino acid interactions. *Journal of the American Chemical Society* 124(21):6133-6143.

20. Bordo D & Argos P (1991) Suggestions for "safe" residue substitutions in site-directed mutagenesis. *Journal of molecular biology* 217(4):721-729.

21. Chen CY*, et al.* (2008) Substrate binding pocket residues of human alkyladenine-DNA glycosylase critical for methylating agent survival. *DNA repair* 7(10):1731-1745.

22. O'Brien PJ & Ellenberger T (2003) Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines. *Biochemistry* 42(42):12418-12429.

23. Hedglin M & O'Brien PJ (2008) Human alkyladenine DNA glycosylase employs a processive search for DNA damage. *Biochemistry* 47(44):11434-11445.

24. Secrist JA, 3rd, Barrio JR, Leonard NJ, & Weber G (1972) Fluorescent modification of adenosine-containing coenzymes. Biological activities and spectroscopic properties. *Biochemistry* 11(19):3499-3506.

25. Lyons DM & O'Brien PJ (2009) Efficient recognition of an unpaired lesion by a DNA repair glycosylase. *Journal of the American Chemical Society* 131(49):17742-17743.

26. Hsieh J, Walker SC, Fierke CA, & Engelke DR (2009) Pre-tRNA turnover catalyzed by the yeast nuclear RNase P holoenzyme is limited by product release. *RNA* 15(2):224-234.

27. O'Connor TR (1993) Purification and characterization of human 3-methyladenine-DNA glycosylase. *Nucleic acids research* 21(24):5561-5569.

28. Baldwin MR & O'Brien PJ (2009) Human AP endonuclease I stimulates multiple-turnover base excision by alkyladenine DNA glycosylase. *Biochemistry* 48(25):6022-6033.

29. Scharer OD, Nash HM, Jiricny J, Laval J, & Verdine GL (1998) Specific binding of a designed pyrrolidine abasic site analog to multiple DNA glycosylases. *The Journal of biological chemistry* 273(15):8592-8597.

30. Vallur AC, Feller JA, Abner CW, Tran RK, & Bloom LB (2002) Effects of hydrogen bonding within a damaged base pair on the activity of wild type and DNA-intercalating mutants of human alkyladenine DNA glycosylase. *The Journal of biological chemistry* 277(35):31673-31678.

31. Parker JB*, et al.* (2007) Enzymatic capture of an extrahelical thymine in the search for uracil in DNA. *Nature* 449(7161):433-437.

32. Banerjee A, Yang W, Karplus M, & Verdine GL (2005) Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. *Nature* 434(7033):612-618.

33. Banerjee A, Santos WL, & Verdine GL (2006) Structure of a DNA glycosylase searching for lesions. *Science* 311(5764):1153-1157.

34. Friedman JI & Stivers JT (2010) Detection of damaged DNA bases by DNA glycosylase enzymes. *Biochemistry* 49(24):4957-4967.

Appendix A

Additional data figures to support Chapter 2

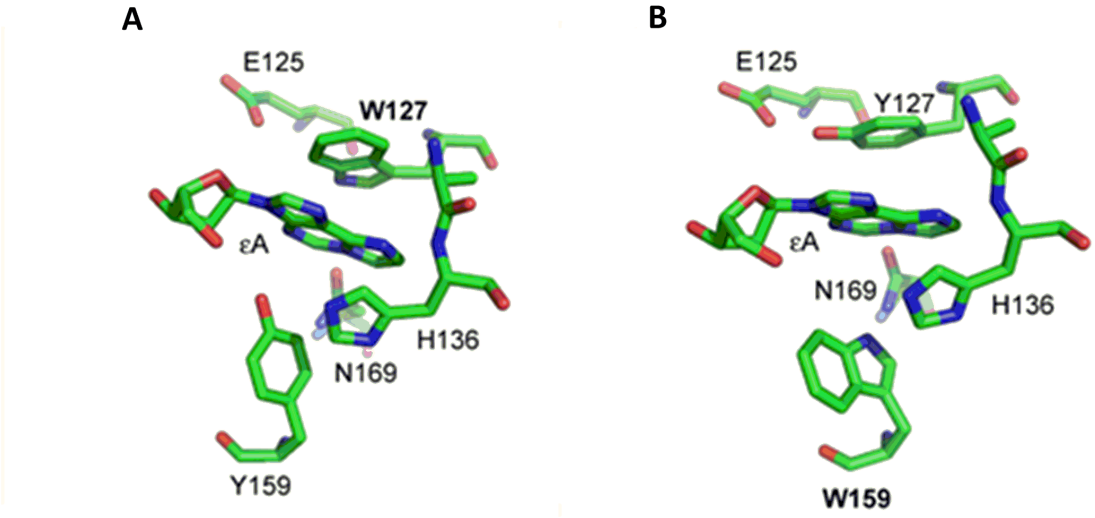


Figure A-1. Substitution of tyrosine with tryptophan can be accommodated in the AAG active site. Tyrosine 127 (**A**) and tyrosine 159 (**B**) were substituted with tryptophan and a rotamer was selected that avoided steric clashes with both DNA and protein side chains (Pymol). No energy minimization or additional modeling was performed, but these exercises suggest that the sligh3tly larger tryptophan side chains could be accommodated in the AAG active site without large-scale changes in other active site residues.

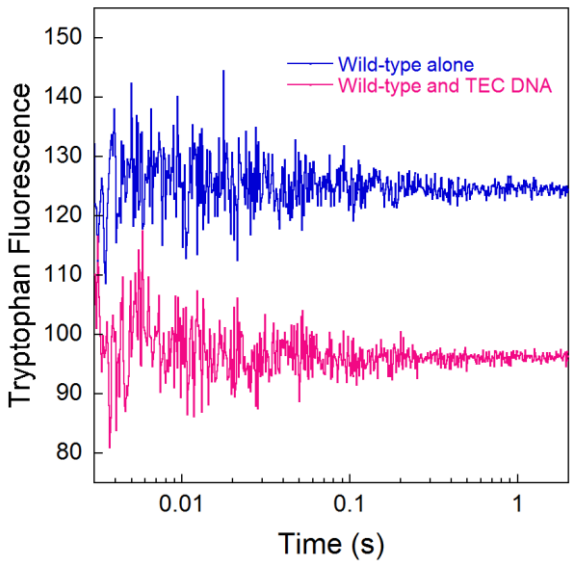


Figure A-2. Tryptophan fluorescence of wild-type AAG does not change upon binding to εA-DNA. Stopped-flow fluorescence was monitored with 1 µM wild-type AAG and 1 µM εA-DNA, exactly as described for the Y159W and Y127W mutants with excision at 295 and a 330 nm (20 nm band-pass) emission filter. The absence of time-dependent decrease in fluorescence indicates that the fluorescence of the native tryptophan residues do not change upon binding to DNA damage and that the changes in the fluorescence of Y127W and Y159W proteins can be ascribed to the new tryptophan residues present in the AAG active site.

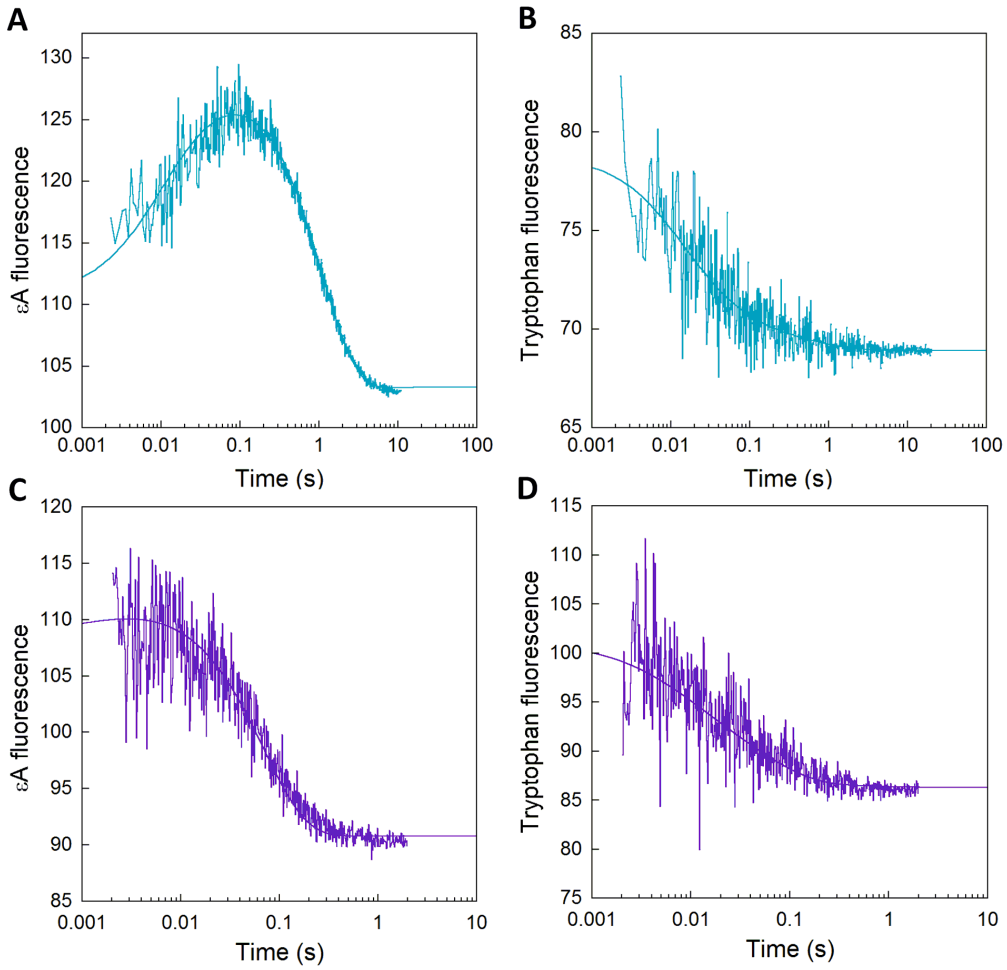
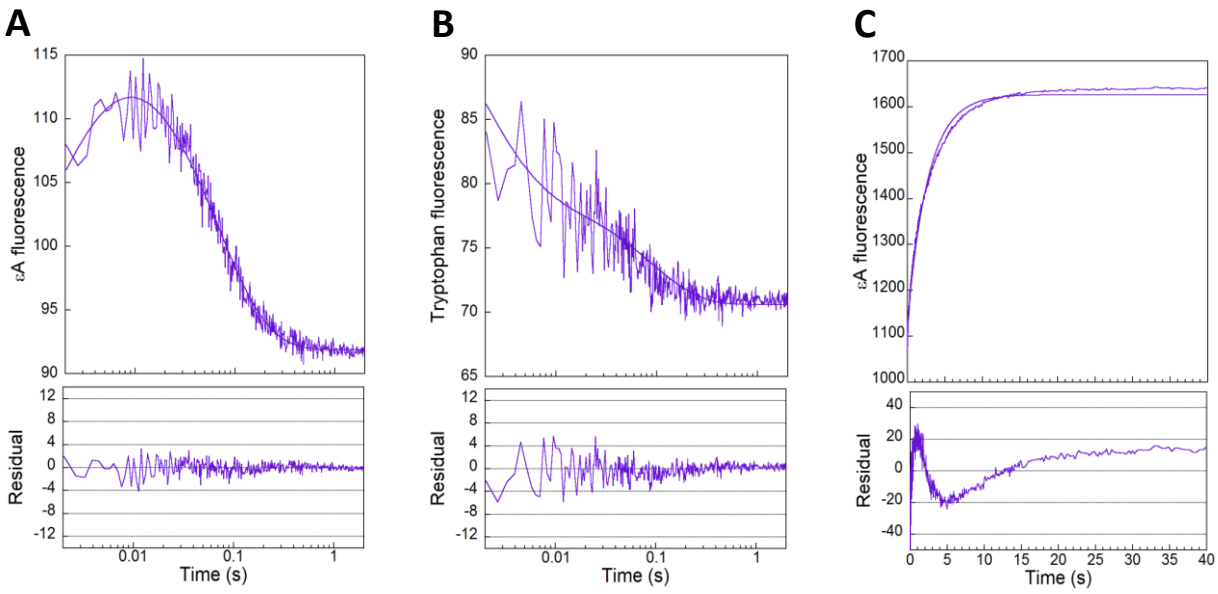


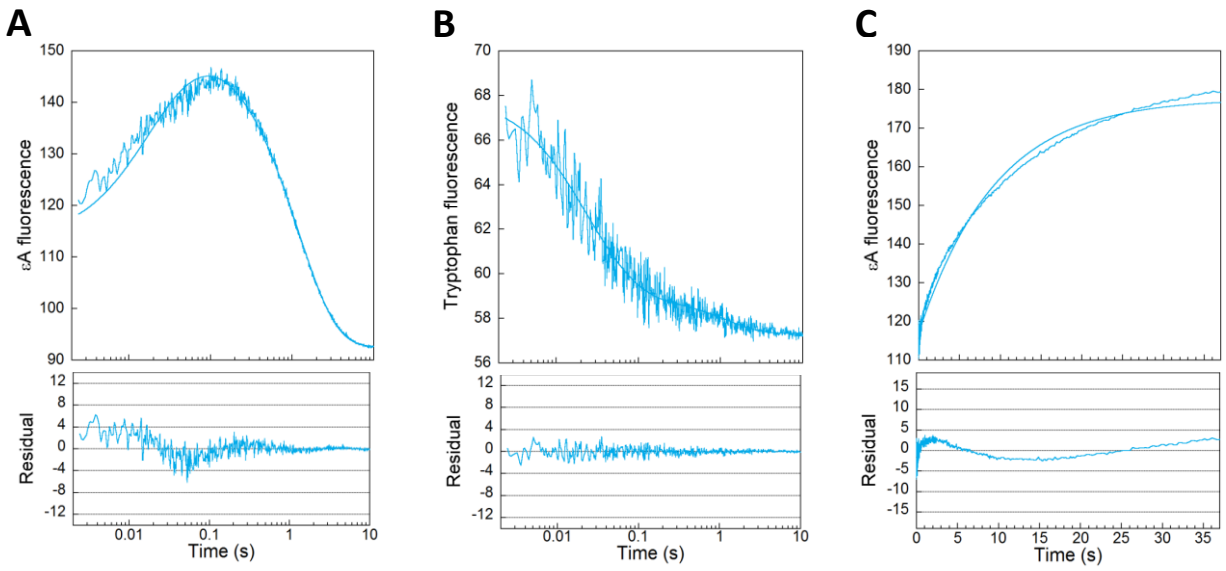
Figure A-3. Stopped-flow fluorescence to measure binding and nucleotide flipping by Y127W and Y159W mutants of AAG. Equimolar mixtures of AAG and εA-DNA were mixed in the stopped-flow exactly as described for Figure 2.5 in the main text, except the concentrations were one half (0.5 μM after mixing). (**A**) εA fluorescence for Y127W. (**B**) Tryptophan fluorescence for Y127W. (**C**) εA fluorescence for Y159W. (**D**) Tryptophan fluorescence for Y159W. The data is the average of three independent reactions and the lines indicate the best fits of Equation 2.6. Several different concentrations between 0.125 μM and 1.0 μM (all concentrations given after mixing) gave the same rate constants, confirming that the concentration chosen was saturating for the mutant and wild-type proteins (data not shown).



D

Y159W	Individual Fits	BM εA Association	BM W Association	BM εA Dissociation
k_{on} ($M^{-1}s^{-1}$)	3.1 x 10^8	2.2 x 10^8	2.8 x 10^8	1.6 x 10^8
k_{off} (s^{-1})	ND	25.2	35.8	10.0
k_{flip} (s^{-1})	13.6	18.7	17.0	16.4
k_{unflip} (s^{-1})	0.37	0.31	0.60	0.56
K_{flip}	40	60	28	29

Figure A-4. Kinetic parameters for DNA binding and nucleotide flipping by Y159W AAG. (**A**) Representative fit for εA association kinetics. (**B**) Representative fit for tryptophan (W) association kinetics. (**C**) Representative fit for εA dissociation kinetics. (**D**) Table summarizing individual fits (Table 2.1 in the main text) and global analysis of curves shown in A-C (using Scheme 2.3 in the main text and the program Berkeley Madonna (BM)).



D

Y127W	Individual Fits	BM εA Association	BM W Association	BM εA Dissociation
k_{on} ($M^{-1}s^{-1}$)	6.8 x 10^7	4.1 x 10^7	3.4 x 10^7	2.6 x 10^7
k_{off} (s^{-1})	ND	3.3	2.1	38.4
k_{flip} (s^{-1})	1.1	0.9	1.0	1.6
k_{unflip} (s^{-1})	0.12	0.10	0.21	0.10
K_{flip}	9	9	5	16

Figure A-5. Kinetic parameters for DNA binding and nucleotide flipping by Y127W AAG. (**A**) Representative fit for εA association kinetics. (**B**) Representative fit for tryptophan association kinetics. (**C**) Representative fit for εA dissociation kinetics. (**D**) Summary table summarizing individual fits (Table 2.1) and global fits shown in part A-C (using Scheme 2.3 in the main text and the program Berkeley Madonna). The rate constants for binding and dissociation are relatively poorly defined, but very similar flipping rate constants are observed in all fits.

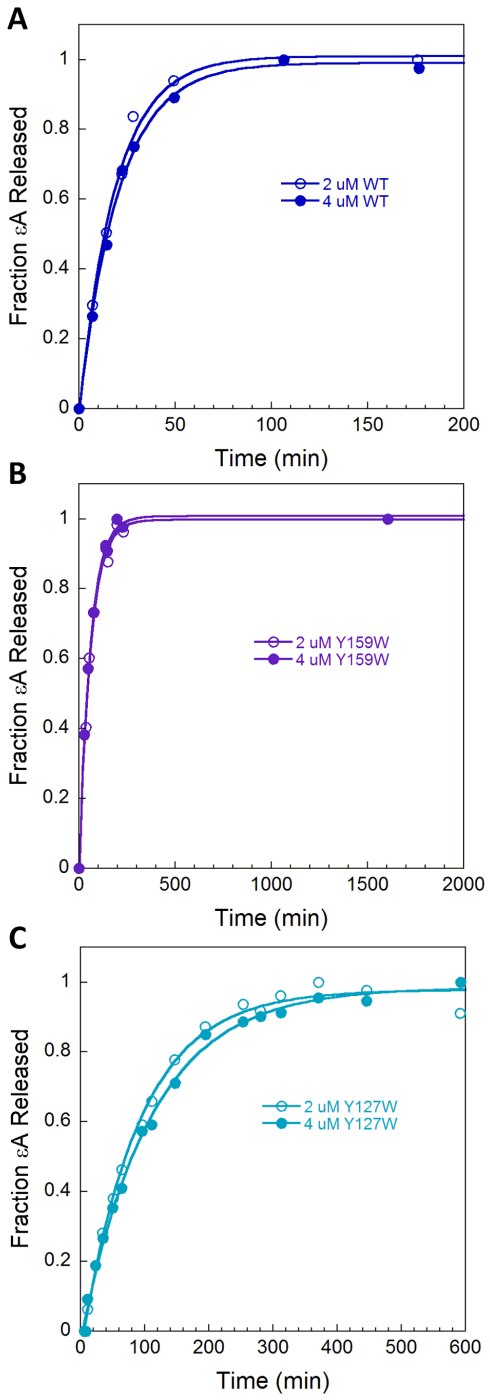


Figure A-6. Single-turnover excision of εA for wild-type (WT) and mutant AAG. Essentially identical rate constants were obtained for excision of εA from TEC DNA (400 nM) and either 2 μM (○) or 4 μM (●) of (**A**) wild-type AAG, (**B**) Y159W, and (**C**) Y127W. The lines indicate the best fit of a single exponential, as described in the main text, and yielded average values of 0.048 min^{-1} for wild-type, 0.019 min^{-1} for Y159W, and 0.012 min^{-1} for Y127W AAG.

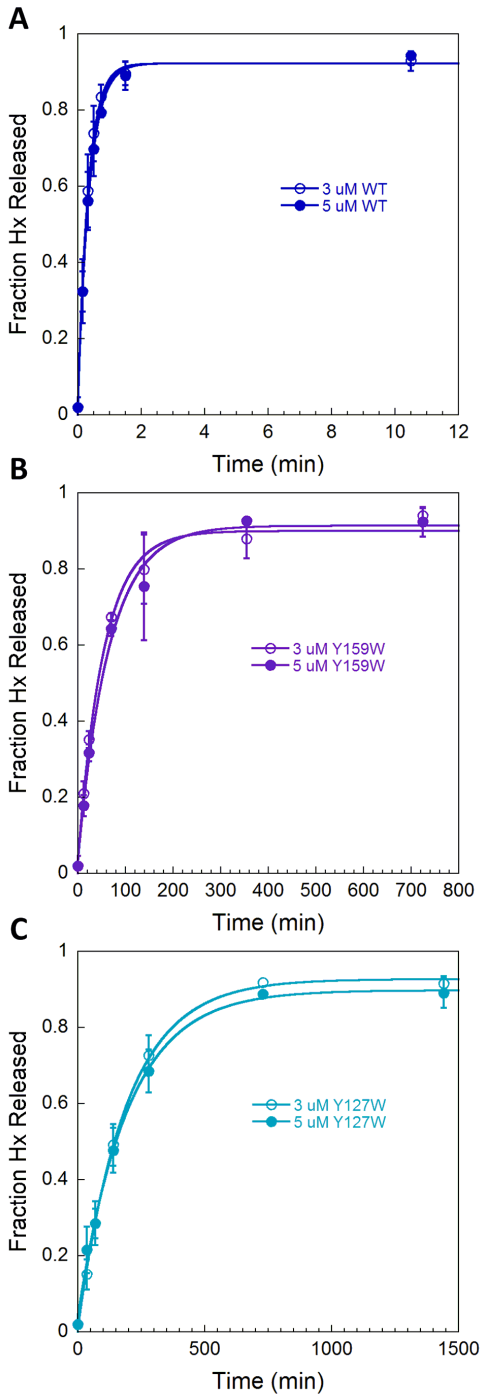


Figure A-7. Single-turnover excision of Hx for wild-type and mutant AAG. Essentially identical rate constants were obtained for excision of Hx from TIC DNA (400 nM) and either 3 µM (○) or 5 µM (●) of wild-type AAG (**A**), Y159W (**B**), and Y127W (**C**). The lines indicate the best fit of a single exponential equation, and yield single- turnover rate constants (k_{max}) of 2.94 min^{-1} for wild-type, 0.017 min^{-1} for Y159W, and 0.0052 min^{-1} for Y127W AAG.

CHAPTER 3

Critical Role of DNA Intercalation in Enzyme-catalyzed Nucleotide Flipping

Although DNA is a remarkably stable molecule, it is nevertheless subject to damage by a variety of reactive intracellular and environmental agents. DNA damage can alter gene expression, affect epigenetic profiles, and even cause cell death (3, 4). To recognize and repair DNA, enzymes must gain access to damaged bases that are normally embedded in the duplex. This is aided by a process called nucleotide flipping that involves the complete 180° rotation of a nucleotide into the enzyme active site (5). It is proposed that both DNA bending and DNA intercalation are important for nucleotide flipping (5, 6). Even though nucleotide flipping is prevalent among DNA repair enzymes, the mechanism is not well understood. In particular, the timing and the energetic contributions provided by DNA intercalation are not known.

Human alkyladenine DNA glycosylase (AAG), also known as methylpurine DNA glycosylase (MPG), is one of many enzymes that use nucleotide flipping to engage substrates. AAG initiates the base excision repair (BER) pathway by locating sites of damage and catalyzing the hydrolysis of the *N*-glycosidic bond to release the damaged base (8). This monomeric repair protein is responsible for removing a diverse set of alkylated and deaminated purine lesions (9-

12). One of the lesions that is most efficiently excised by AAG is 1,N^6-ethenoadenine (εA) that is formed by lipid oxidation or exposure to exogenous alkylating agents (12, 13). The minimal kinetic mechanism for multi-step recognition of εA damage has been reported, including the

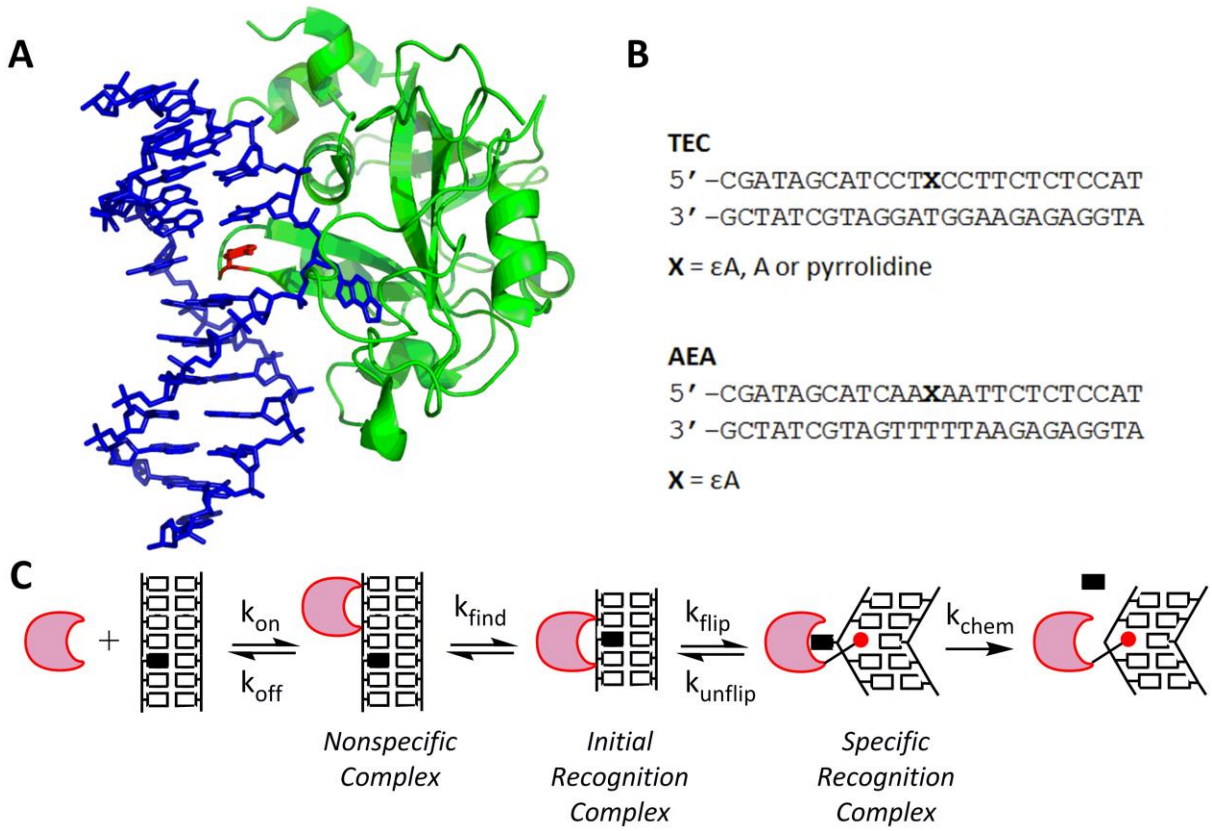


Figure 3.1. Nucleotide flipping by AAG. (**A**) Crystal structure of human alkyladenine DNA glycosylase (E125Q) bound to εA-containing DNA (1, 2). Image was rendered with Pymol (http://www.pymol.org) using coordinates from the PDB (1EWN). The εA base is flipped into the active site, and the intercalating residue (Y162) is shown in red. (**B**) Sequences of the 25mer oligonucleotide substrates. (**C**) Minimal mechanism for the initiation of base excision repair. AAG (crescent) binds to nonspecific DNA and rapidly searches for DNA damage. This searching process is described by the macroscopic rate constant k_{find}. Once a lesion (solid rectangle) is encountered in an initial recognition complex, it can be flipped into the enzyme active site to form the specific recognition complex. In this specific complex, Y162 (red circle) intercalates into the DNA where it takes the place of the extrahelical lesion. AAG catalyzes hydrolysis of the *N*-glycosidic bond (k_{chem}) from this complex.

rate of nucleotide flipping which was directly monitored by fluorescence to follow changes in the intrinsic fluorescence of εA (14, 15). Given the excess of undamaged DNA, initial binding of AAG is not at the site of damage (Figure 3.1C). AAG uses nonspecific binding and facilitated

diffusion (16) to search for sites of damage. Once an εA is found, AAG forms an initial recognition complex that is accompanied by changes in the stacking of the lesion base (14). From this intermediate, AAG catalyzes nucleotide flipping to form a more stable specific recognition complex that positions the substrate for *N*-glycosidic bond cleavage, which releases the damaged base (2). Multi-step recognition provides multiple opportunities for discrimination between damaged and undamaged nucleotides and similar mechanisms have been proposed for independently evolved DNA repair glycosylases, including 8-oxoguanine DNA glycosylases and uracil DNA glycosylases (17-20).

Crystal structures of AAG bound to εA-DNA revealed that the DNA is bent in the specific recognition complex and a tyrosine (Y162) intercalates into the space left by the flipped-out damaged nucleotide (Figure 3.1A; (1, 2)). Similar intercalating interactions are observed in all nucleotide flipping enzymes, but the identity of the intercalating amino acid varies and the function(s) of the intercalating residue remains unknown (21, 22). The two models for how DNA intercalation could contribute to nucleotide flipping have been described as “pushing” and “plugging”(23-25). In the pushing model, DNA intercalation destabilizes the DNA duplex to accelerate nucleotide flipping. In the plugging model, DNA intercalation occurs subsequent to nucleotide flipping and provides a barrier to the return of the nucleotide to the duplex, effectively slowing the rate of unflipping. These two mechanisms are not mutually exclusive, and both models can be addressed by kinetic studies of proteins in which the identity of the intercalating residue has been varied.

We investigated the functional contributions of the intercalating residue to each of the steps in the AAG catalytic mechanism by mutating the intercalating tyrosine residue to

phenylalanine (Y162F) and alanine (Y162A). Remarkably, the Y162A mutation increases the rate of nucleotide flipping by 50-fold relative to wild-type (WT) AAG. This mutant also exhibits faster unflipping, resulting in a 140-fold reduction in the equilibrium constant for formation of the flipped-out complex. In contrast, the kinetic parameters for the Y162F and wild-type enzymes are very similar, suggesting that the hydroxyl of tyrosine is not necessary for function. These results establish that DNA intercalation contributes to the specific recognition of DNA damage by acting as a plug to stabilize the specific recognition complex. DNA intercalation stabilizes the initial recognition complex without providing a push, slowing the rate of nucleotide flipping. Nevertheless, tyrosine 162 plays a critical role in the discrimination between damaged and undamaged nucleotides, by increasing the amount of the extrahelical lesion recognition complex and enabling efficient repair of rare sites of damage.

Materials and Methods

Purification of Wild-type and Mutant AAG Proteins. The catalytic domain of human AAG that lacks the first 79 amino acids was expressed in *E. coli* and purified as previously described (26). The Y162F and Y162A mutants were constructed by site-directed mutagenesis and purified using the same protocol. The concentration of Y162A AAG was determined from the UV absorbance and predicted extinction coefficient and the concentration of wild-type and Y162F AAG were determined by active site burst (16).

Synthesis and Purification of Oligodeoxynucleotides. Oligonucleotides were synthesized by Integrated DNA Technologies or the W.M. Keck Facility at Yale University, were purified and the concentrations determined from the calculated extinction coefficients as previously described (16). Duplexes were annealed with a 1.2-fold excess of the unmodified complement by heating

to 90 °C and cooling slowly to 4 °C. Two different sequence contexts were used and these are referred to by the central three nucleotides (i.e., TEC has a central εA lesion flanked by a 5'T and a 3'C; Figure 3.1B).

Steady-State Fluorescence Measurements. Fluorescence emission spectra were collected with a PTI QuantaMaster fluorometer controlled by FeliX software. For εA fluorescence, an excitation wavelength of 314 nm (6 nm band-pass) was used and the total fluorescence was measured at emission wavelengths from 340-480 nm (6 nm band-pass). Samples (300 µL) of 400 nM εA-containing DNA were prepared in the standard buffer [50 mM NaMES (pH 6.5), 100 mM NaCl, 1 mM EDTA, 1 mM DTT] and spectra were recorded at 25 °C. To determine the steady-state fluorescence of εA-containing DNA bound to AAG, the spectra were recorded within one minute to ensure that no significant excision of εA occurred. Four independent titrations were averaged, and the data were fit by a quadratic equation assuming tight binding by AAG (Equation 3.1), in which F_{rel} is the relative fluorescence, A is the fractional quenching, and K_d is the dissociation constant for the εA-DNA.

$$F_{rel} = 1 - \frac{A(K_d + E + DNA) - \sqrt{(K_d + E + DNA)^2 - (4E \times DNA)}}{2E} \quad (3.1)$$

Gel-based Glycosylase Assay. Single-turnover glycosylase activity was measured with a discontinuous assay that utilizes abasic site cleavage by NaOH followed by DNA separation on a denaturing polyacrylamide gel (14). Fluorescein-labeled (5'-6-fam) DNA substrates (50 nM) containing εA were prepared in the standard buffer. Reactions were initiated with the addition of 75 nM – 600 nM AAG and incubated at 25 °C. At various time points, a sample from the reaction was removed and quenched in 2 volumes of 0.3 M NaOH, giving a final hydroxide

concentration of 0.2 M. The abasic sites were cleaved by heating at 70 °C for 15 minutes. Samples were mixed with an equal volume of formamide/EDTA loading buffer before loading onto a 15% polyacrylamide gel. Gels were scanned with a Typhoon Imager (GE Trio+ Healthcare) to detect the fluorescein label by exciting at 488 nm and measuring emission with a 520BP40 filter. The gel bands were quantified using ImageQuant TL (GE Healthcare). The data were converted to fraction product [Fraction Product = product / (product + substrate)] and then fit by a single exponential (Equation 3.2) using Kaleidagraph (Synergy Software). Single-turnover rates were independent of the concentration of AAG, indicating that the maximal rate constant was measured ($k_{obs} = k_{max}$; see Appendix B).

$$\text{Fraction product} = A(1 - e^{-k_{obs}t}) \qquad (3.2)$$

To measure competition between εA-DNA and undamaged DNA, single-turnover reactions containing 200 nM fluorescein-labeled εA DNA and 0–50 μM undamaged DNA, in which the εA was replaced by a normal A, were initiated with the addition of 300 nM AAG. Data for Y162A AAG was fit by an IC_{50} equation (Equation 3.3); wild-type and Y162F AAG were fit by a line with a slope of zero.

$$\text{Activity} = \frac{1}{1+[I]/IC_{50}} \qquad (3.3)$$

Stopped-Flow Kinetics. Pre-steady state kinetic experiments were performed on a Hi-Tech SF-61DSX2, controlled by Kinetic Studio (TgK Scientific). The fluorescence of εA was measured using an excitation wavelength of 313 nm and a WG360 long-pass emission filter (14, 15). At least three traces were averaged together at each concentration. The traces for wild-type and Y162F AAG were fit by a triple exponential (Equation 3.4) where *F* is the fluorescence as a function of time, *C* is the fluorescence of free DNA, *X, Y,* and *Z* are the changes in fluorescence

of the intermediates, and *t* is the time. The traces for Y162A AAG were fit by a single exponential (Equation 3.2) and k_{obs} is equal to $k_{flip} + k_{unflip}$.

$$F = C - X(1 - e^{-k_{1,obs}t}) - Y(1 - e^{-k_{2,obs}t}) - Z(1 - e^{-k_{2,obs}t}) \qquad (3.4)$$

Observed rate constants were plotted versus concentration and fit by a straight line. $k_{1,obs}$ showed a linear concentration dependence, and the slope is equal to k_{on} ($M^{-1}s^{-1}$). Negative y-intercepts were observed under these conditions. Similar negative intercepts were reported for another DNA binding protein with high affinity for DNA (7). The value of $k_{2,obs}$ was independent of concentration and is equal to $k_{flip} + k_{unflip}$.

Pulse-Chase Assay to Measure Substrate Dissociation. The macroscopic rate constant for dissociation of wild-type and mutant AAG from εA-containing DNA was measured by pulse-chase in the standard reaction buffer at 25 °C as previously described for wild-type AAG (14, 15). In 20 µL reactions, 50 nM fluorescein-labeled TEC DNA was mixed with 300 nM or 600 nM AAG for 20 seconds, and then a chase of 10 µM pyrrolidine-containing DNA was added. Pyrrolidine binds tightly to AAG as a transition state analogue (1, 27). At various time points, a sample from the reaction was removed to evaluate the partitioning of a bound E•S complex between *N*-glycosidic bond hydrolysis and dissociation with the standard glycosylase assay. The fraction of product was fit by a single exponential (Equation 3.2). Control reactions in which no chase was added provided the single-turnover rate constant, k_{max}, and confirmed that these concentrations of AAG were saturating. From these values, the dissociation rate constant, k_{off}, for AAG dissociating from εA-DNA was calculated as previously described (14, 15). The rearranged equation is indicated by Equation 3.5, in which *A* is the burst amplitude (the fraction of product formed in the burst phase of the experiment), k_{max} is the single-turnover rate

constant for formation of product, and $k_{off,obs}$ is the macroscopic rate constant for dissociation from the flipped-out complex.

$$k_{off,obs} = \frac{k_{max}}{A} - k_{max} \qquad (3.5)$$

Since the εA-DNA binds in two steps, the observed rate constant for dissociation of substrate ($k_{off,\,obs}$) could be limited by the unflipping rate (k_{unflip}) or dissociation from nonspecific DNA (k_{off}). According to Figure 3.1C, and assuming that the flipped-out complex is stable (i.e., $k_{flip} >> k_{unflip}$), this observed dissociation rate constant can be expressed in terms of the microscopic rate constants (Equation 3.6; (28)). Given rapid dissociation from nonspecific DNA, the observed rate of dissociation from the εA-DNA•AAG complex is approximately equal to the reverse rate constant for flipping ($k_{off,obs} \approx k_{unflip}$).

$$k_{off,obs} = k_{unflip}\left(\frac{k_{off}}{k_{off} + k_{flip}}\right) \qquad (3.6)$$

Direct Measurement of Substrate Dissociation. Stopped-flow double-mixing experiments were performed to measure unflipping and dissociation of εA-DNA by Y162A. To rapidly form the flipped-out AAG-DNA complex, 2.8 μM AEA DNA was mixed with 2 μM Y162A. After an aging time of one second, 18 μM or 36 μM pyrrolidine-DNA was added as a competitor. The final concentrations after mixing were 700 nM AEA DNA, 500 nM Y162A, and 9 μM or 18 μM pyrrolidine-DNA. The reaction was followed for 10 seconds; no significant excision of εA occurs during this time. Three traces were averaged together and the change in fluorescence was fit by a single exponential (Equation 3.2). As described for the pulse-chase assay, the observed rate constant for dissociation is equal to the rate constant for unflipping.

Results

Binding and Excision of εA by AAG Mutants

We used site-directed mutagenesis to create Y162F and Y162A variants, and characterized their ability to bind and excise εA from synthetic 25mer oligonucleotides containing a central εA•T mismatch. Unless otherwise indicated, the previously characterized TEC sequence context was used (Figure 3.1B). As AAG has a relatively slow rate of base excision, steady-state fluorescence titrations could be used to measure binding of active enzyme to εA-DNA (14). For this experiment it is advantageous to use a DNA sequence with high εA fluorescence (AEA), and the stable binding of the extrahelical εA lesion results in a 5-fold quenching of fluorescence for wild-type AAG (Figure 3.2A). Titrations of Y162A and Y162F AAG showed somewhat less quenching of εA-fluorescence (2–3-fold), but nonetheless indicate tight binding of εA-DNA. The similar 1:1 stoichiometry at the equivalence point validates the protein concentrations that were determined either by UV absorbance or burst analysis.

To evaluate if the mutant proteins retain catalytic activity, single-turnover experiments were performed with enzyme in excess over εA-DNA substrate (Figure 3.2B). For wild-type AAG the single-turnover rate constant reflects the *N*-glycosidic bond hydrolysis step (14). The rate constant for wild-type AAG (0.05 min^{-1}) matches the previously reported value (14, 15). The rate constant for excision of εA by Y162F (0.03 min^{-1}) is similar to wild-type, consistent with reports that the Y162F mutation is well tolerated (29, 30). We were surprised to find that the rate constant for excision of εA is only slightly smaller for Y162A AAG (0.02 min^{-1}), because

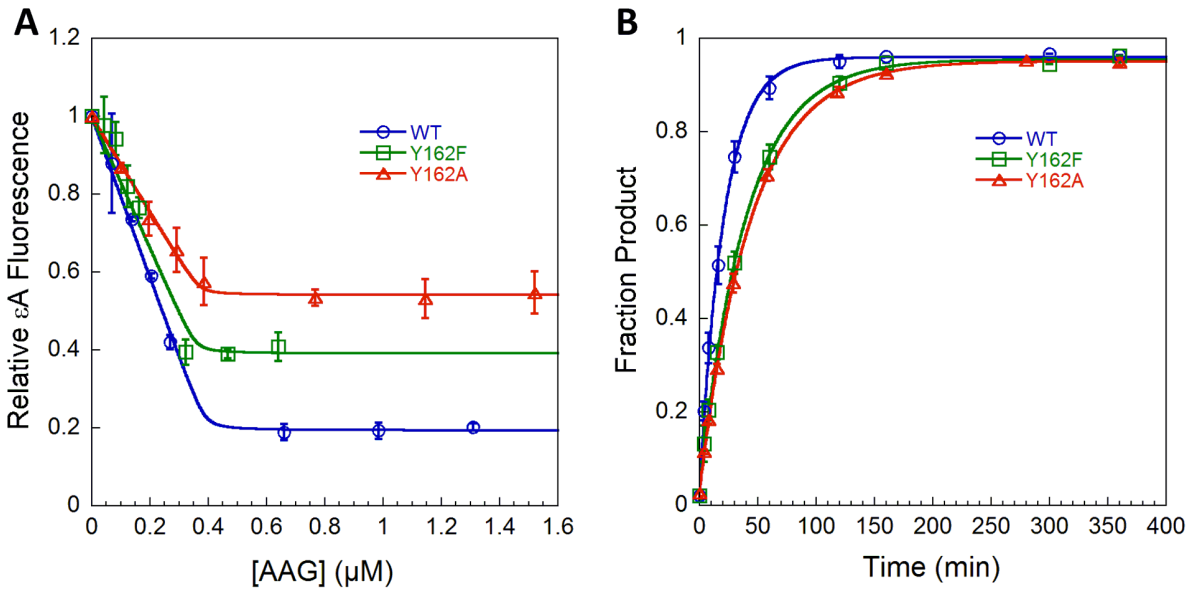


Figure 3.2. Binding and excision of εA by AAG. (**A**) Titration of 0.4 μM εA-DNA with wild-type and mutant AAG proteins. Steady-state fluorescence was monitored within one minute, before excision can occur, and normalized by dividing by the fluorescence of free DNA. Wild-type and mutants show tight binding to εA-DNA (Equation 3.1). (**B**) Single-turnover excision of εA by wild-type and mutant AAG with 50 nM DNA and saturating (300 nM) enzyme was fit by a single exponential (Equation 3.2) to obtain the rate constants for *N*-glycosidic bond cleavage (Table 3.1). Data points are the average ± SD from three independent experiments.

previous work suggested that this mutant may lack catalytic activity (2). The similar rate constants for base excision strongly suggest that the active site is not disrupted by these mutations, but larger effects on recognition and flipping of the damaged base could be masked by the high affinity binding of AAG to εA-DNA. Therefore, it is important to measure individual reaction steps to address the question of how DNA intercalation by Y162 contributes to recognition and flipping of damaged nucleotides.

Stopped-flow Fluorescence to Monitor Binding and Flipping of εA-DNA

We performed rapid mixing experiments of wild-type and mutant enzymes with εA-DNA to determine the rate constants for association and flipping. By improving the sensitivity of the stopped-flow, we were able to make measurements over a wider concentration range than was

previously possible (14, 15). When εA-DNA was mixed with excess wild-type AAG, we observed an initial increase in fluorescence that was followed by a decrease in fluorescence (Figure 3.3A). The first phase is linearly dependent on the enzyme concentration (Figure 3.3C) and this was previously assigned to binding and formation of the initial recognition complex (14). The second phase is independent of concentration and corresponds to nucleotide flipping to form the specific recognition complex (Figure 3.4D). We often observed a much slower third phase, with a small change in amplitude, that may reflect an artifact such as photobleaching. This phase was not observed in all experiments and was not reproducible. When three phases were observed, we fit the traces by a triple exponential (Equation 3.4), because this gives the most consistent fits to the first two phases (see Appendix B for representative rate constants). Experiments with Y162F AAG yielded very similar results as the wild-type enzyme, with a transient increase in εA-fluorescence followed by a decrease in fluorescence (Figure 3.3B). Both the second order rate constant for the binding step (Figure 3.3C), and the observed first order rate constant for formation of the flipped-out specific recognition complex (Figure 3.3D) are 2-fold larger for the Y162F mutant as compared to wild-type (Table 3.1).

We previously noted that the initial recognition complex, in which the εA lesion is partially unstacked, appears to form with the bimolecular rate constant for association (14). As AAG uses facilitated diffusion to search for sites of damage (16), the searching steps must be faster than the association with nonspecific DNA (Figure 3.1C). We hypothesized that this fast searching could be due to multiple proteins simultaneously searching for the εA site under conditions of excess protein. This suggests that the searching time by a single protein would be significantly slower. Therefore, we also performed stopped-flow association experiments in

Table 3.1. Kinetic parameters for recognition and excision of εA[a]

	Wild-type	Y162F	Y162A
k_{on} ($M^{-1}s^{-1}$)	(1.1 ± 0.03) x 10^9	(2.1 ± 0.2) x 10^9	Fast[d]
k_{find} (s^{-1})	116 ± 11	114 ± 4	Fast[d]
k_{flip} (s^{-1})	3.6 ± 0.7 (3.6 ± 0.3)[c]	7.9 ± 1.0 (7.4 ± 0.5)[c]	170 ± 16[e]
k_{unflip} (s^{-1})	(1.6 ± 0.3) x 10^{-3}	(4.6 ± 0.2) x 10^{-3}	10 ± 1[f]
K_{flip}[b]	2300	1700	17
k_{chem} (s^{-1})	(8.0 ± 0.6) x 10^{-4}	(4.3 ± 0.2) x 10^{-4}	(3.8 ± 0.1) x 10^{-4}

[a]Rate constants were determined from changes in εA fluorescence or glycosylase activity using the TEC oligonucleotide, unless otherwise indicated. The standard conditions were 25 ^{o}C, 50 mM NaMES (pH 6.5), 100 mM NaCl, 1 mM EDTA, and 1 mM DTT.

[b]The equilibrium constant for flipping is given by the ratio of the flipping and unflipping rate constants ($K_{flip} = k_{flip}/k_{unflip}$).

[c]Values from stopped-flow experiments with excess protein (values with excess DNA are in parenthesis).

[d]k_{on} and k_{find} are too fast to measure for Y162A.

[e]A value of 66 ± 2 s^{-1} was determined for the AEA oligonucleotide with excess protein (Figure 3.5D).

[f]The AEA oligonucleotide was used, because it gives a larger change in fluorescence (Figure 3.6B).

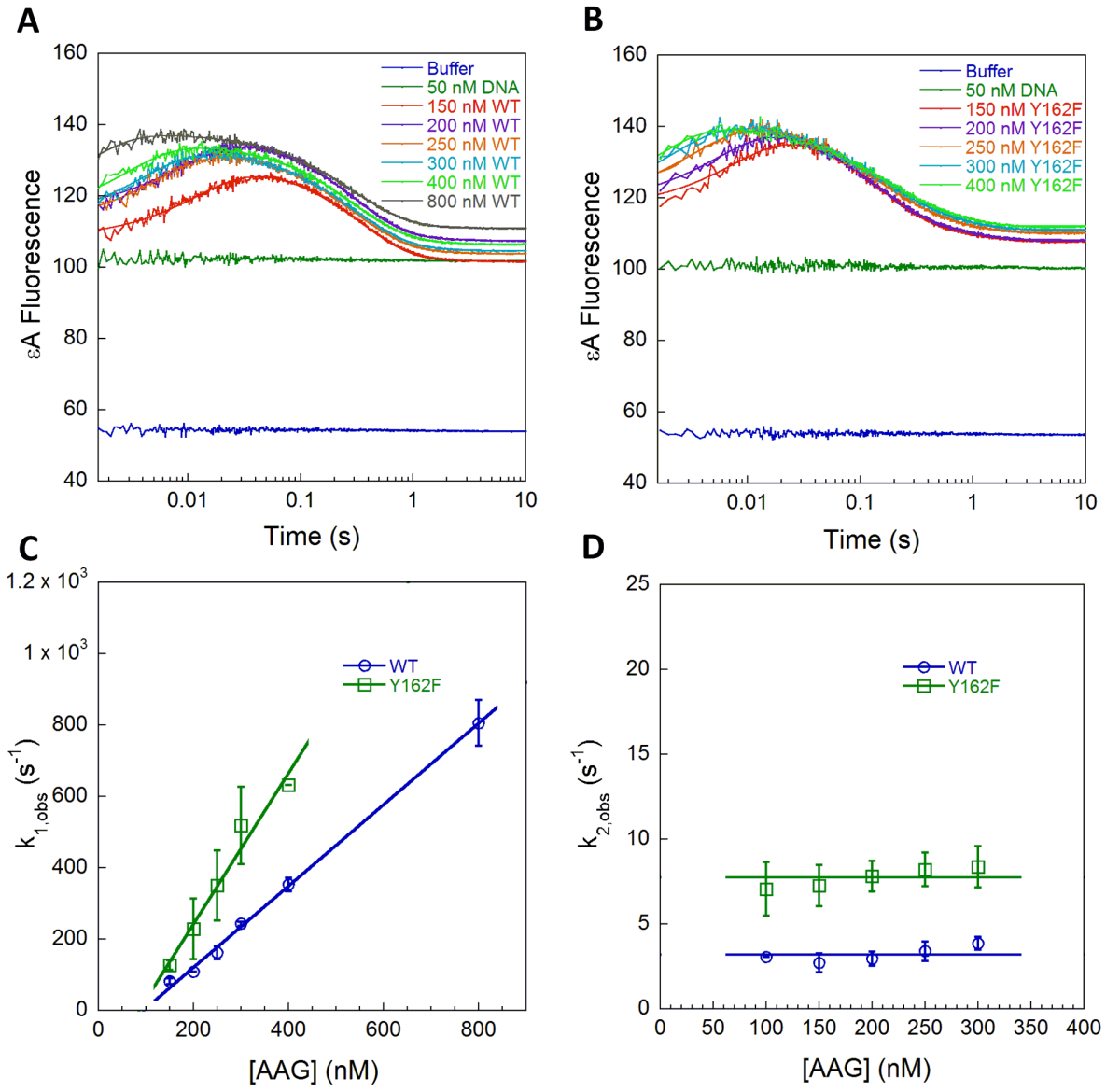


Figure 3.3. **Stopped-flow fluorescence with excess protein to measure binding and nucleotide flipping by wild-type and Y162F AAG.** Representative data from experiments in which 50 nM TEC DNA was mixed with wild-type (**A**) or Y162F (**B**). Traces are the average of three binding reactions and are fit by the sum of exponentials (Equation 3.4). (**C**) The rate constant for the first phase of the binding reaction ($k_{1,obs}$) is dependent on the concentration of AAG and a linear fit yields the bimolecular rate constant for binding (k_{on}) with the values summarized in Table 3.1. It is notable that the intercepts are negative, similar to what has been observed in another case of extremely fast and tight binding (7). (**D**) The rate constants for the second phase of the binding reaction ($k_{2,obs}$) are independent of the concentration of AAG and reflect the sum of the forward and reverse rate constants for nucleotide flipping. Rate constants in C and D are from three independent experiments (average ± SD).

which εA-DNA was in excess over wild-type (Figure 3.4A) or Y162F AAG (Figure 3.4B). Although the fraction of DNA that is bound is much smaller under these conditions, there is sufficient signal to accurately measure protein binding. As was observed for conditions of excess protein, the individual fluorescence traces showed biphasic increase and decrease in fluorescence, but now the first observed rate constant was independent of concentration (Figure 3.4C). We designate this observed rate constant as k_{find}, the macroscopic rate constant for the rapidly reversible searching steps that culminate in the formation of the initial recognition complex (Figure 3.1C). The Y162F mutation did not alter the searching process significantly (Figure 3.4C; k_{find} = 114 s^{-1}). The observed rate constant for the formation of the flipped-out specific recognition complex when there was excess DNA was the same within error as when there was excess protein (Figure 3.4D; Table 3.1). This indicates that excess protein does not interfere with nucleotide flipping.

Association experiments with Y162A AAG yielded surprisingly different fluorescence traces. When excess Y162A AAG was mixed with εA-DNA (TEC), only a single time dependent decrease in fluorescence was observed (Figure 3.5A). Even at low concentrations of protein, the first phase of the reaction was complete in the dead time of the stopped-flow ($k_{on} \geq 10^9$ $M^{-1}s^{-1}$). The exponential decrease in fluorescence, which indicates formation of the flipped-out specific recognition complex, was independent of concentration and much faster than either the wild-type or Y162F AAG (Figure 3.5D, orange symbols). The changes in fluorescence were quite small for the TEC DNA. Therefore we also measured association kinetics for the AEA DNA with excess protein (Figure 3.5B) and with excess DNA (Figure 3.5C). For the Y162A mutant protein, almost identical rate constants were observed whether protein or DNA was in excess (Figure 3.5D, blue

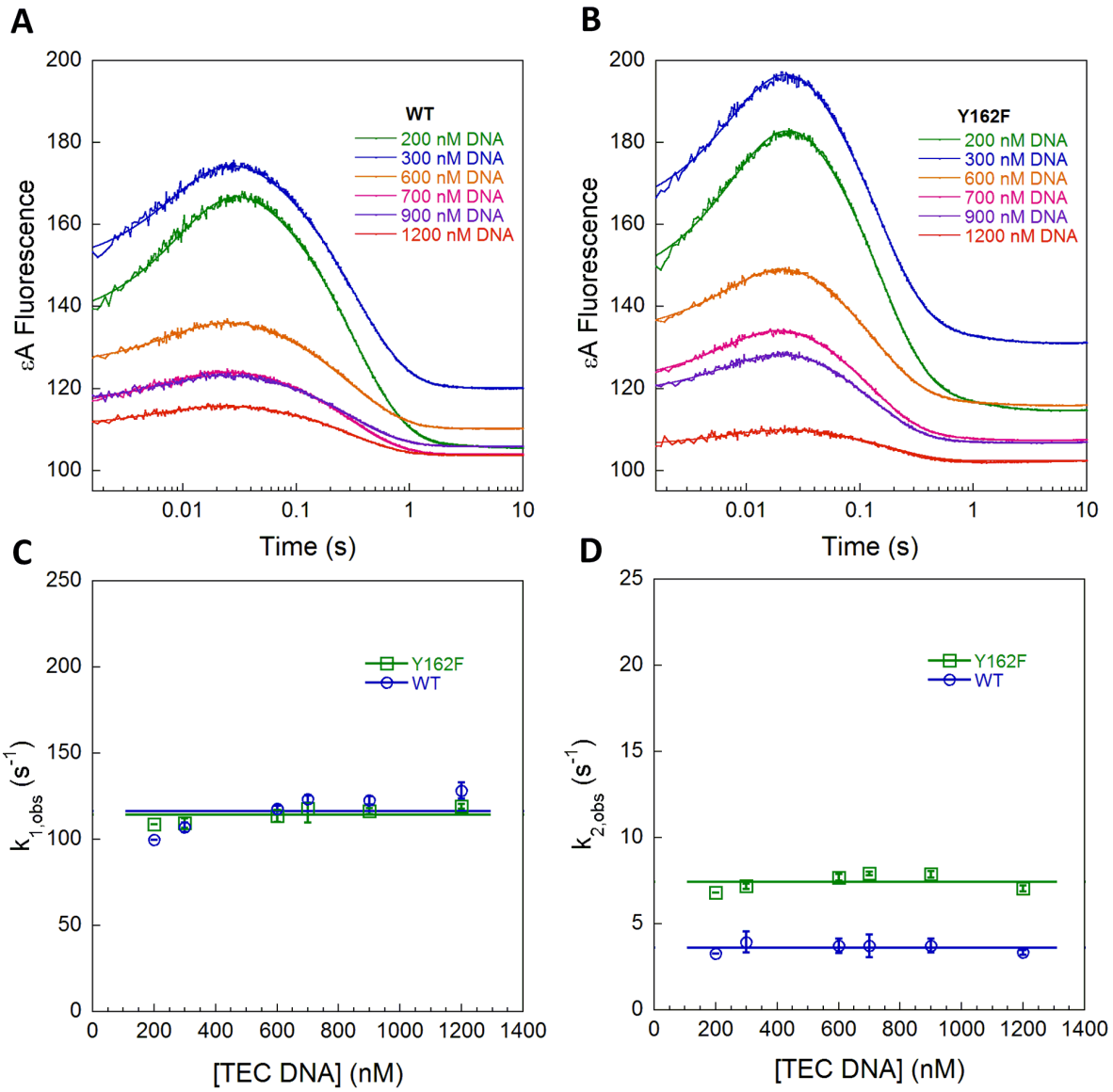


Figure 3.4. **Stopped-flow fluorescence with excess DNA to measure k_{find} for wild-type and Y162F AAG.** Conditions of excess DNA ensure that binding events involve only a single protein. Representative data from experiments in which 100 nM of wild-type (**A**) or Y162F (**B**) protein was mixed with increasing concentrations of TEC DNA. The εA-fluorescence was collected and fit by Equation 3.4 as described for Figure 3.3. (**C**) With excess DNA, the rate constant for the first phase ($k_{1,obs}$) is independent of concentration and is designated k_{find}. (**D**) The rate constant for the second phase ($k_{2,obs}$) is also concentration independent and corresponds to nucleotide flipping. This value is identical to that measured in stopped-flow experiments with excess protein (Table 3.1). Rate constants in C and D are from three independent experiments (average ± SD).

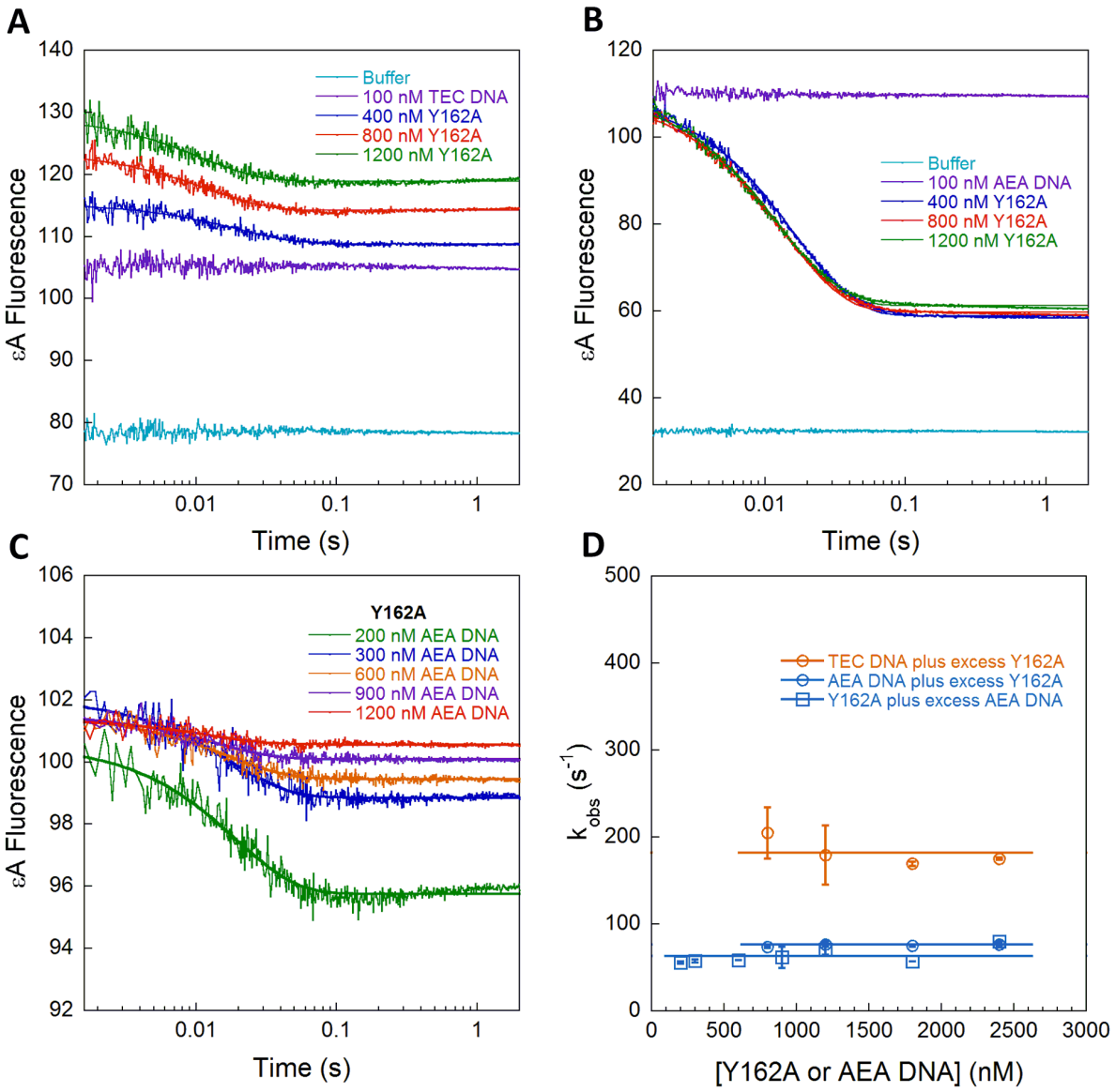


Figure 3.5. **Stopped-flow fluorescence to measure nucleotide flipping by Y162A AAG.** Representative reactions in which 100 nM of TEC (**A**) or AEA (**B**) εA-DNA was mixed with increasing concentrations of excess Y162A protein. (**C**) Representative data for binding of 100 nM Y162A AAG with increasing concentrations of AEA DNA. In each case, the changes in fluorescence were fit by a single exponential. (**D**) Observed rate constants from three independent experiments are plotted as a function of concentration (average ± SD). Remarkably, this rate constant ($k_{obs} = k_{flip} + k_{unflip}$) is much larger for Y162A than for the wild-type enzyme. The initial binding and the formation of the initial recognition complex are not observed, indicating that they are both faster than the nucleotide flipping step.

symbols). These data indicate that Y162A AAG associates with nonspecific DNA and searches significantly faster than the wild-type enzyme, such that the searching step is not rate-limiting under any of the conditions tested. Remarkably, the observed flipping step is also significantly faster for the Y162A mutant than for wild-type AAG (50-fold faster for the TEC DNA; Table 3.1).

Measurement of Unflipping and Dissociation of εA-DNA

As the binding and nucleotide flipping steps are reversible, the observed association is affected by both forward and reverse rates. Therefore, it is necessary to measure the dissociation kinetics to define individual microscopic rate and equilibrium constants. The dissociation of AAG from εA-containing DNA was investigated using the pulse-chase method, whereby the AAG•εA-DNA complex is allowed to form and is subsequently chased with an excess of tight binding pyrrolidine-DNA inhibitor (27). Any AAG that dissociates is bound to the inhibitor, allowing measurement of the partitioning forward (abasic DNA product) and backward (εA-DNA substrate) as previously described (14, 15). For wild-type AAG, 30% of the complex partitions forward to produce product and 70% dissociates (Figure 3.6A). Assuming that dissociation of AAG from nonspecific DNA is fast, these data yield the microscopic rate constant for unflipping (Equation 3.6). More complete analysis is consistent with fast dissociation from nonspecific DNA (see Appendix B). The Y162F mutant showed greater dissociation of the substrate (90% dissociated; Figure 3.6A), indicating a 3-fold faster value of k_{unflip} compared to wild-type (Table 3.1). In contrast, 100% of the Y162A mutant dissociated and no product could be detected in the presence of chase (Figure 3.6A), indicating that unflipping is much faster than *N*-glycosidic bond cleavage. In order to precisely measure the rate of unflipping for the Y162A mutant we next performed a double-mixing experiment in the stopped-flow.

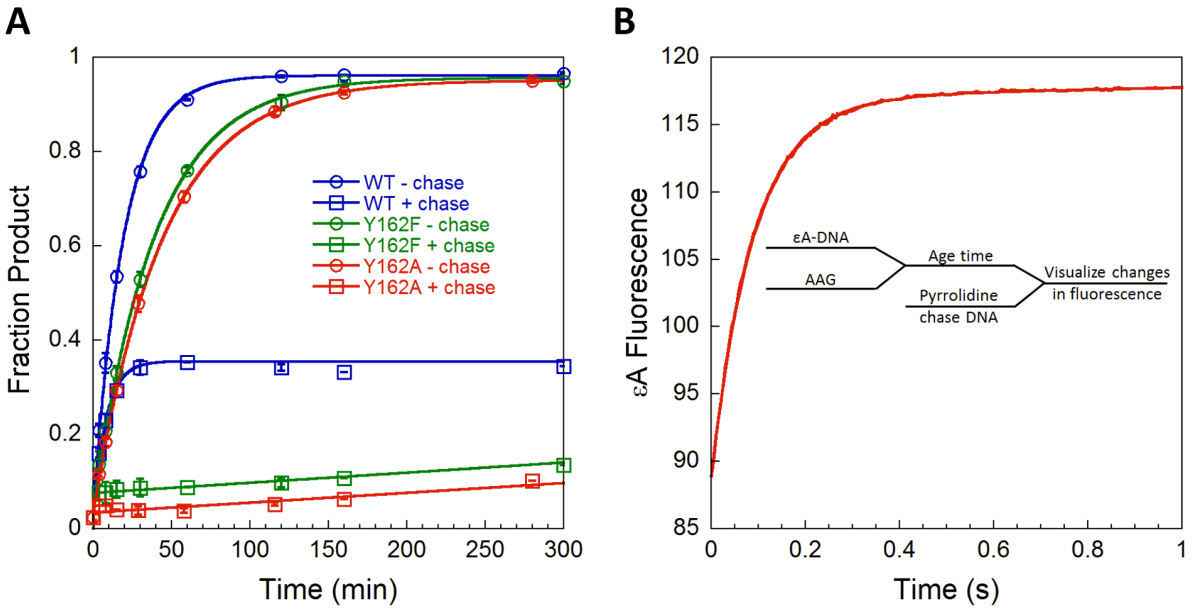


Figure 3.6. Pulse-chase experiments to measure dissociation of εA-DNA. (**A**) The partitioning between the forward reaction (base excision) and substrate dissociation was measured using a gel-based assay. The commitment for wild-type (30%), Y126F (8%), and Y162A (0%) were used to calculate the rates of unflipping (see Materials and Methods). The low commitment for Y162F AAG was reproducible and readily distinguished from the control in which chase and substrate are added at the same time (see Appendix B). (**B**) Double-mixing experiments were performed using stopped-flow fluorescence to monitor the increase in fluorescence upon release of εA-DNA. The AAG•DNA complex was formed, aged for one second, and then chased with an excess of pyrrolidine-DNA as a competitor. The data were fit by a single exponential and the average ± SD from three independent experiments is included in Table 3.1.

In the double-mixing experiment, the flipped-out complex of Y162A AAG with εA-DNA was formed and then challenged with pyrrolidine-DNA competitor (Figure 3.6B; inset scheme). The time-dependent increase in εA fluorescence is fit by a single exponential, which corresponds to the rate-limiting unflipping step. This experiment reveals that the rate of εA unflipping for Y162A AAG (k_{unflip} = 10.3 s^{-1}) is 6,000-fold faster than that of wild-type AAG (Table 3.1). Although the Y162A mutation causes only a modest defect in glycosylase activity towards εA, it imparts a dramatic effect on nucleotide unflipping. This observation supports the model that tyrosine intercalation acts as a plug to block the unflipping of εA.

With the rate constants for unflipping in hand, the rate and equilibrium constants for nucleotide flipping could be determined for wild-type and mutant enzymes (Table 3.1). The microscopic rate constant for flipping was calculated from the observed rate constant for formation of the extrahelical complex ($k_{flip} = k_{2,obs} - k_{unflip}$), and the equilibrium constant for nucleotide flipping is the ratio of the two values ($K_{flip} = k_{flip}/k_{unflip}$). The Y162F mutant exhibits a K_{flip} value that is almost identical to wild-type AAG, with 2-fold faster rates for both flipping and unflipping. In contrast, the Y162A mutant has a greatly reduced equilibrium constant for flipping that is 140-fold less favorable than that of the wild-type enzyme. The destabilization of the specific recognition complex is caused by a 6,000-fold increased rate of unflipping that is only partially balanced by a 50-fold increased rate of flipping (Table 3.1). Thus, kinetic characterization supports the model that tyrosine serves as a plug and argues against the pushing mechanism. Instead of pushing, the intercalation of tyrosine stabilizes the initial recognition complex and effectively slows the nucleotide flipping transition to the stable specific recognition complex.

Competition Between εA-DNA and Undamaged DNA

Although Y162A exhibits only a modest decrease in single-turnover glycosylase activity, the decreased stability of the flipped-out complex suggests that this mutant would be unable to efficiently find sites of εA damage when sites of damage are rare. We tested this hypothesis by using undamaged DNA as a competitor. Under single-turnover conditions, with enzyme in excess over the εA-labeled DNA substrate, the single-turnover rate constant for εA excision by wild-type and Y162F AAG was unchanged in the presence of excess undamaged DNA (Figure 3.7). In contrast, Y162A AAG is inhibited by undamaged DNA with an IC_{50} value of 20 μM. This

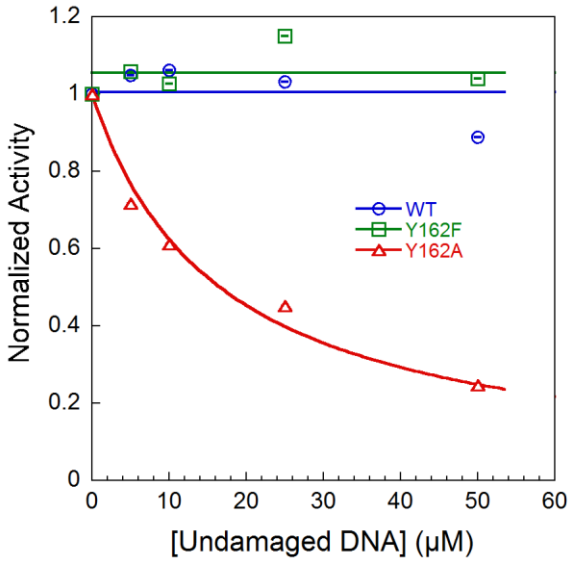


Figure 3.7. Competition between εA-DNA and undamaged DNA. Single-turnover excision of εA by wild-type and mutant AAG was measured as described in Figure 3.2B with 200 nM εA-DNA, 300 nM enzyme, and the indicated concentration of undamaged DNA. The observed rate constants were normalized by dividing by the rate constant in the absence of competitor (average ± SD; three independent experiments). Wild-type and Y162F AAG are unaffected by the presence of excess undamaged DNA. In contrast, Y162A AAG is inhibited by the addition of competitor DNA (IC_{50} = 20 μM).

concentration of undamaged DNA corresponds to 5,000-fold excess of undamaged over damaged nucleotides, which is far below the expected ratio of undamaged to damaged sites in the human genome. Thus, Y162A is less efficient than the wild-type protein at repairing sites of damage when sites of damage are rare.

Discussion

Structures of DNA repair glycosylases in complex with DNA suggest that nucleotide flipping is aided by intercalating residues (1, 5, 21, 25). The prevalence of nucleotide flipping and the dramatic nature of the conformational change that accompanies DNA bending and nucleotide flipping have captured the interest of biologists and chemists alike. However, little is known about the fundamental mechanisms by which nucleotide flipping is accomplished

because it is difficult to observe rapid conformational changes that occur on the millisecond time scale. To understand the contribution of DNA intercalation to the energetic landscape for specific recognition and nucleotide flipping, we performed comprehensive kinetic analysis of mutant forms of AAG in which the conserved intercalating tyrosine (Y162) was mutated. These results provide support for the model that intercalation by an aromatic side chain greatly stabilizes the extrahelical conformation of the specific recognition complex. Unexpectedly, we further identify a role for Y162 in slowing the rate of nucleotide flipping, apparently by stabilizing the initial recognition complex. The new model that emerges suggests that DNA intercalation contributes to the specificity of AAG, but the increased specificity comes at the cost of reduced speed. These results provide new insight into the mechanism of AAG, and have implications for many other enzymes that use nucleotide flipping for DNA repair or epigenetic DNA modification.

Kinetic Mechanism of AAG

Since crystal structures cannot capture the dynamic nature of DNA glycosylases, spectroscopic assays are needed to elucidate reaction mechanisms. Previous work found that AAG and other DNA glycosylases use multi-step recognition to distinguish damaged sites from among an excess of undamaged sites (14, 15, 21, 31). We performed a series of transient kinetic experiments to dissect the individual steps of binding, searching, flipping, and base excision by AAG (Figure 3.1C). Our results with the wild-type enzyme are in agreement with previously reported values (14), but we have gained new insight into the DNA searching process by directly measuring the rate of searching under conditions of excess DNA.

By employing a DNA sequence context in which the εA lesion is quenched (TEC), we observed a transient intermediate in which the εA is less strongly quenched and presumably partially unstacked (Figure 3.3A). Although binding of AAG to the εA lesion appears to be a bimolecular reaction (Figure 3.3C), we favor the model that nonspecific DNA binding and rapid facilitated diffusion allow AAG to locate and bind the damaged nucleotide. Under the conditions of excess protein that are most commonly employed to monitor labeled DNA, simultaneous searching by multiple AAG molecules makes this searching step too fast to detect (14). We have modeled searching by multiple proteins explicitly to show that an additive model, in which two proteins search twice as fast as one protein, is consistent with the experimental data (see Appendix B). To directly measure the searching by a single molecule of AAG we employed conditions of excess DNA (Figure 3.4). Under these conditions the initial recognition complex is formed with a unimolecular rate constant of 120 s^{-1} (Table 3.1). This observed rate constant (k_{find}) is equal to the sum of the rate constants for finding and leaving the site of damage. For wild-type AAG this initial recognition complex is significantly populated, indicating that the rate of finding the εA site is faster than the rate of leaving the site.

Although multiple AAG molecules can bind to a single oligonucleotide, the observed rate constant for nucleotide flipping is identical under conditions of excess DNA or excess protein (Table 3.1). This indicates that multiple proteins do not affect the flipping step and that rate constants obtained in experiments with multiple or single proteins bound can be directly compared. To define the microscopic rate constants for flipping and the equilibrium constant for flipping, the dissociation of the bound εA-DNA must also be measured. Wild-type AAG has a very slow rate of unflipping and an extremely favorable equilibrium constant for stabilizing the

flipped-out nucleotide (Table 3.1). AAG compensates for slow *N*-glycosidic bond cleavage by binding very tightly to εA. We tested the specificity of AAG by employing conditions with 12,500-fold excess of undamaged sites (50 μM undamaged oligonucleotide) and found that the observed rate of εA excision was unaffected (Figure 3.7).

Contributions of DNA Intercalation by Tyr 162

We focused on the role of the intercalating tyrosine, which is universally conserved among eukaryotic AAG homologs (1, 32). We find that both the Y162F and Y162A mutations are well tolerated *in vitro* and do not appear to greatly perturb the structure, because the *N*-glycosidic bond cleavage step is within 2-fold of wild-type AAG. Therefore, the comparison of wild-type and mutant proteins provides valuable insights into the functional contributions of tyrosine to individual microscopic steps in the recognition of εA lesions.

Although the Y162F variant has not been found in nature, the kinetic parameters are almost identical to wild-type (Table 3.1). Association, flipping, and unflipping are approximately 2-fold faster for this mutant, whereas the searching rate constant and the overall equilibrium constant for flipping are essentially identical to wild-type AAG. Therefore, the hydroxyl group of tyrosine is not necessary for any of the microscopic steps associated with finding and flipping out εA lesions. This suggests that aromatic stacking interactions are responsible for the stabilization of the initial recognition complex and that both tyrosine and phenylalanine act as a plug to stabilize the flipped-out complex. Some prokaryotic homologs of AAG have been identified that appear to substitute histidine for tyrosine at the intercalating position (1, 33). It is easy to imagine that histidine could make similar interactions as tyrosine and phenylalanine. Interestingly, Y162 has been identified as the primary site of nitration, which is reported to

decrease glycosylase activity (30). It is possible that nitration of tyrosine serves a biological function that could account for the conservation of this residue. Future *in vivo* experiments with the Y162F mutant, that cannot be nitrated, may be able to shed light on this issue (30).

The comparison of Y162A to wild-type AAG allows the energetic contribution of DNA intercalation to be measured for individual steps in the kinetic mechanism. Both association and searching are significantly faster for Y162A, but were too fast to measure with stopped-flow assays. Parallels can be drawn with recent reports that mutation of intercalating residues to alanine in other DNA glycosylases increases the rate of diffusion (34, 35). Thus, a theme emerges that DNA intercalation slows the macroscopic rate of diffusive searching. Although the Y162A mutant engages εA sites significantly faster than the wild-type enzyme, the destabilization of the specific recognition complex greatly reduces the specificity of the enzyme. This mutant is not able to effectively find sites of damage when challenged by an excess of undamaged bases (Figure 3.7), leading to the prediction that DNA intercalation is essential for *in vivo* DNA repair. Consistent with this conclusion, heterologous expression of Y162A AAG fails to fully protect cells against an exogenous DNA alkylating agent (2).

Conclusions

Structures of AAG bound to εA-DNA (1) raised the possibility that intercalation by Y162 could increase the rate constant for flipping by “pushing” or destabilizing the duplex and decrease the rate constant for unflipping by “plugging” the hole vacated by the flipped-out nucleobase, as has also been suggested for other glycosylases (21, 25). Our study provides a rigorous test of these models. The kinetic parameters for Y162A support the plugging model, with this mutant showing a 6,000-fold increase in the rate of unflipping relative to wild-type AAG. Surprisingly,

the Y162A mutant does not show a reduced rate of flipping, but instead it flips out εA lesions 50-fold faster than the wild-type enzyme. This argues against the pushing model for Y162 and demonstrates that intercalation of Y162 dramatically slows the nucleotide flipping step. The slower rate constant for flipping by wild-type AAG is presumably explained by tight binding of εA in an initial recognition complex. Consistent with this view, the initial recognition complex does not appear to stably form for Y162A (Figure 3.5). Overall, the Y162A mutation greatly destabilizes the specific recognition complex, and we obtain an estimate of 140-fold stabilization for the contribution of tyrosine intercalation to the equilibrium constant for flipping. AAG is a particularly attractive system to delve into the biophysical and biochemical mechanism of multi-step DNA recognition, but it is apparent that many different enzymes must employ similar strategies to discriminate between non-specific and specific sites, and the lessons learned from AAG may be more widely applicable to other DNA modifying enzymes.

References

1. Lau AY, Scharer OD, Samson L, Verdine GL, & Ellenberger T (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. *Cell* 95(2):249-258.

2. Lau AY, Wyatt MD, Glassner BJ, Samson LD, & Ellenberger T (2000) Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG. *Proceedings of the National Academy of Sciences of the United States of America* 97(25):13573-13578.

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

4. Robertson AB, Klungland A, Rognes T, & Leiros I (2009) DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cell Mol Life Sci* 66(6):981-993.

5. Roberts RJ & Cheng X (1998) Base flipping. *Annu Rev Biochem* 67:181-198.

6. Stivers JT (2004) Site-specific DNA damage recognition by enzyme-induced base flipping. *Progress in nucleic acid research and molecular biology* 77:37-65.

7. Kozlov AG & Lohman TM (2002) Stopped-flow studies of the kinetics of single-stranded DNA binding and wrapping around the Escherichia coli SSB tetramer. *Biochemistry* 41(19):6032-6044.

8. Ye N, Holmquist GP, & O'Connor TR (1998) Heterogeneous repair of N-methylpurines at the nucleotide level in normal human cells. *Journal of molecular biology* 284(2):269-285.

9. Engelward BP*, et al.* (1997) Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. *Proceedings of the National Academy of Sciences of the United States of America* 94(24):13087-13092.

10. O'Brien PJ & Ellenberger T (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase. *J Biol Chem* 279(11):9750-9757.

11. Hitchcock TM*, et al.* (2004) Oxanine DNA glycosylase activity from Mammalian alkyladenine glycosylase. *J Biol Chem* 279(37):38177-38183.

12. Gros L, Ishchenko AA, & Saparbaev M (2003) Enzymology of repair of etheno-adducts. *Mutation research* 531(1-2):219-229.

13. Pandya GA & Moriya M (1996) 1,N6-ethenodeoxyadenosine, a DNA adduct highly mutagenic in mammalian cells. *Biochemistry* 35(35):11487-11492.

14. Wolfe AE & O'Brien PJ (2009) Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by human alkyladenine DNA glycosylase. *Biochemistry* 48(48):11357-11369.

15. Hendershot JM, Wolfe AE, & O'Brien PJ (2011) Substitution of active site tyrosines with tryptophan alters the free energy for nucleotide flipping by human alkyladenine DNA glycosylase. *Biochemistry* 50(11):1864-1874.

16. Hedglin M & O'Brien PJ (2008) Human alkyladenine DNA glycosylase employs a processive search for DNA damage. *Biochemistry* 47(44):11434-11445.

17. Banerjee A, Santos WL, & Verdine GL (2006) Structure of a DNA glycosylase searching for lesions. *Science* 311(5764):1153-1157.

18. Banerjee A, Yang W, Karplus M, & Verdine GL (2005) Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. *Nature* 434(7033):612-618.

19. Parker JB*, et al.* (2007) Enzymatic capture of an extrahelical thymine in the search for uracil in DNA. *Nature* 449(7161):433-437.

20. Kuznetsov NA*, et al.* (2007) Pre-steady-state kinetic study of substrate specificity of Escherichia coli formamidopyrimidine--DNA glycosylase. *Biochemistry* 46(2):424-435.

21. Stivers JT (2008) Extrahelical damaged base recognition by DNA glycosylase enzymes. *Chemistry* 14(3):786-793.

22. Brooks SC, Adhikary S, Rubinson EH, & Eichman BF (2013) Recent advances in the structural mechanisms of DNA glycosylases. *Biochimica et biophysica acta* 1834(1):247-271.

23. Jiang YL & Stivers JT (2001) Reconstructing the substrate for uracil DNA glycosylase: tracking the transmission of binding energy in catalysis. *Biochemistry* 40(25):7710-7719.

24. Jiang YL & Stivers JT (2002) Mutational analysis of the base-flipping mechanism of uracil DNA glycosylase. *Biochemistry* 41(37):11236-11247.

25. Parikh SS*, et al.* (1998) Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *The EMBO journal* 17(17):5214-5226.

26. O'Brien PJ & Ellenberger T (2003) Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines. *Biochemistry* 42(42):12418-12429.

27. Scharer OD, Nash HM, Jiricny J, Laval J, & Verdine GL (1998) Specific binding of a designed pyrrolidine abasic site analog to multiple DNA glycosylases. *J Biol Chem* 273(15):8592-8597.

28. Hsieh J, Walker SC, Fierke CA, & Engelke DR (2009) Pre-tRNA turnover catalyzed by the yeast nuclear RNase P holoenzyme is limited by product release. *RNA* 15(2):224-234.

29. Vallur AC, Feller JA, Abner CW, Tran RK, & Bloom LB (2002) Effects of hydrogen bonding within a damaged base pair on the activity of wild type and DNA-intercalating mutants of human alkyladenine DNA glycosylase. *J Biol Chem* 277(35):31673-31678.

30. Jones LE, Jr.*, et al.* (2009) Differential effects of reactive nitrogen species on DNA base excision repair initiated by the alkyladenine DNA glycosylase. *Carcinogenesis* 30(12):2123-2129.

31. Porecha RH & Stivers JT (2008) Uracil DNA glycosylase uses DNA hopping and short-range sliding to trap extrahelical uracils. *Proceedings of the National Academy of Sciences of the United States of America* 105(31):10791-10796.

32. Guo HH, Choe J, & Loeb LA (2004) Protein tolerance to random amino acid change. *Proceedings of the National Academy of Sciences of the United States of America* 101(25):9205-9210.

33. Aamodt RM, Falnes PO, Johansen RF, Seeberg E, & Bjoras M (2004) The Bacillus subtilis counterpart of the mammalian 3-methyladenine DNA glycosylase has hypoxanthine and 1,N6-ethenoadenine as preferred substrates. *The Journal of biological chemistry* 279(14):13601-13606.

34. Dunn AR, Kad NM, Nelson SR, Warshaw DM, & Wallace SS (2011) Single Qdot-labeled glycosylase molecules use a wedge amino acid to probe for lesions while scanning along DNA. *Nucleic acids research* 39(17):7487-7498.

35. Nelson SR, Dunn AR, Kathe SD, Warshaw DM, & Wallace SS (2014) Two glycosylase families diffusively scan DNA using a wedge residue to probe for and identify oxidatively damaged bases. *Proceedings of the National Academy of Sciences of the United States of America* 111(20):E2091-2099.

Appendix B

Additional data figures to support Chapter 3

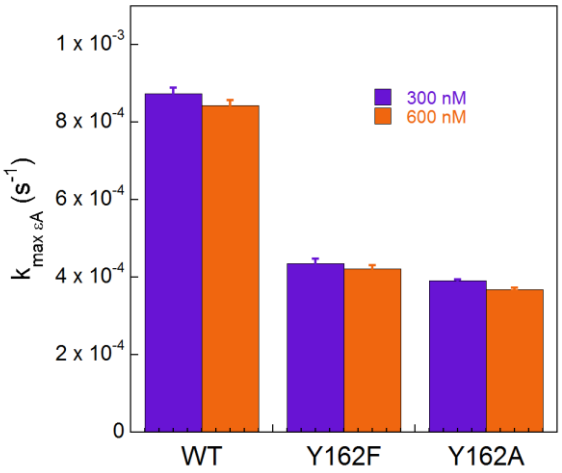


Figure B-1. Single-turnover excision of εA for wild-type and mutant AAG. Essentially identical rate constants were obtained for excision of εA from TEC DNA (50 nM) and either 300 nM or 600 nM AAG. These data indicate that 300 nM enzyme is saturating for all of the AAG variants tested.

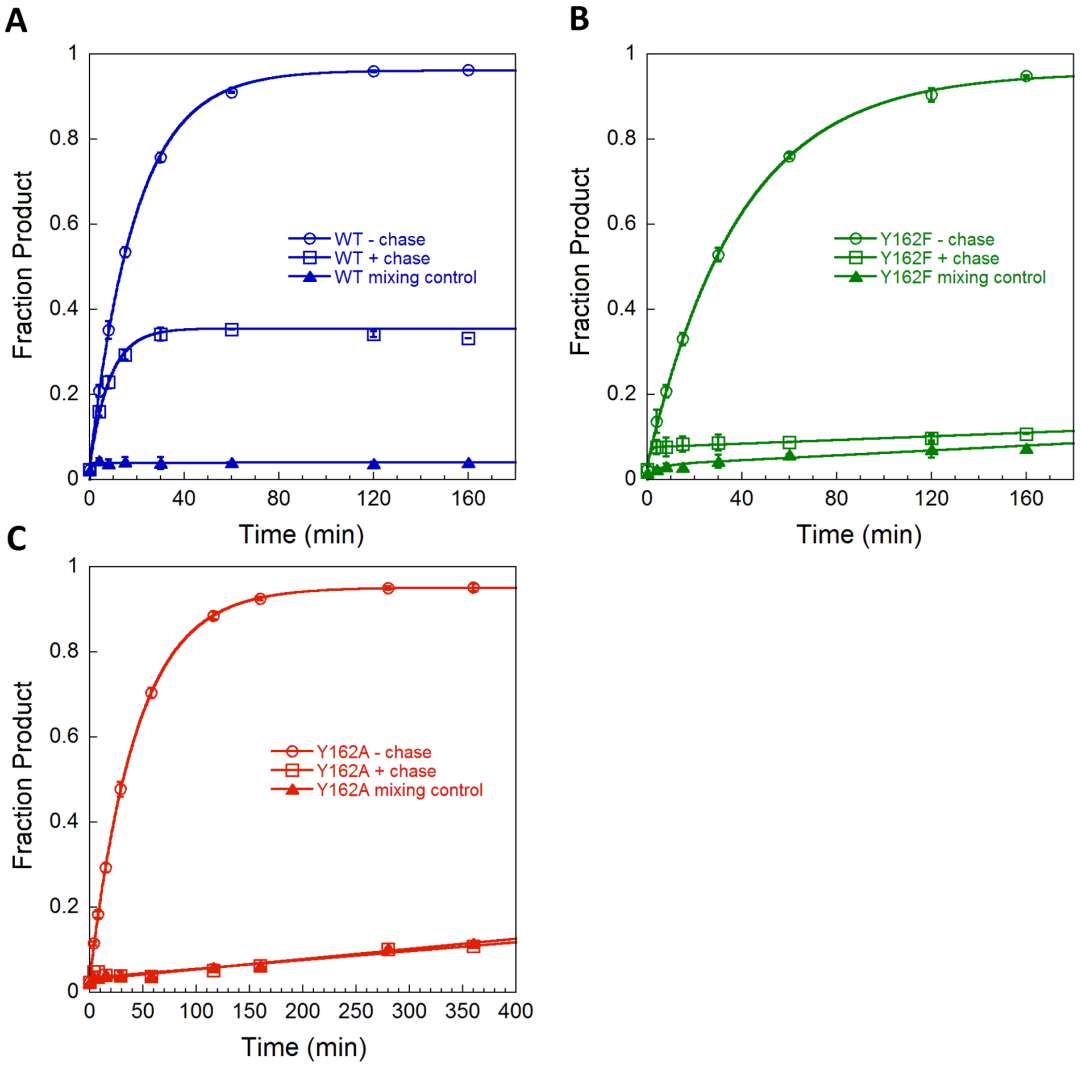
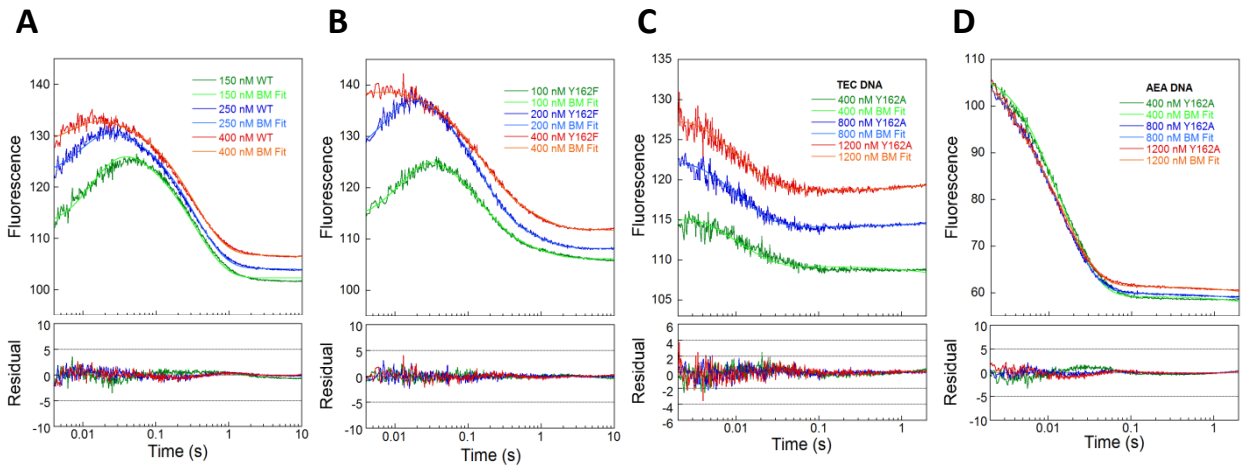


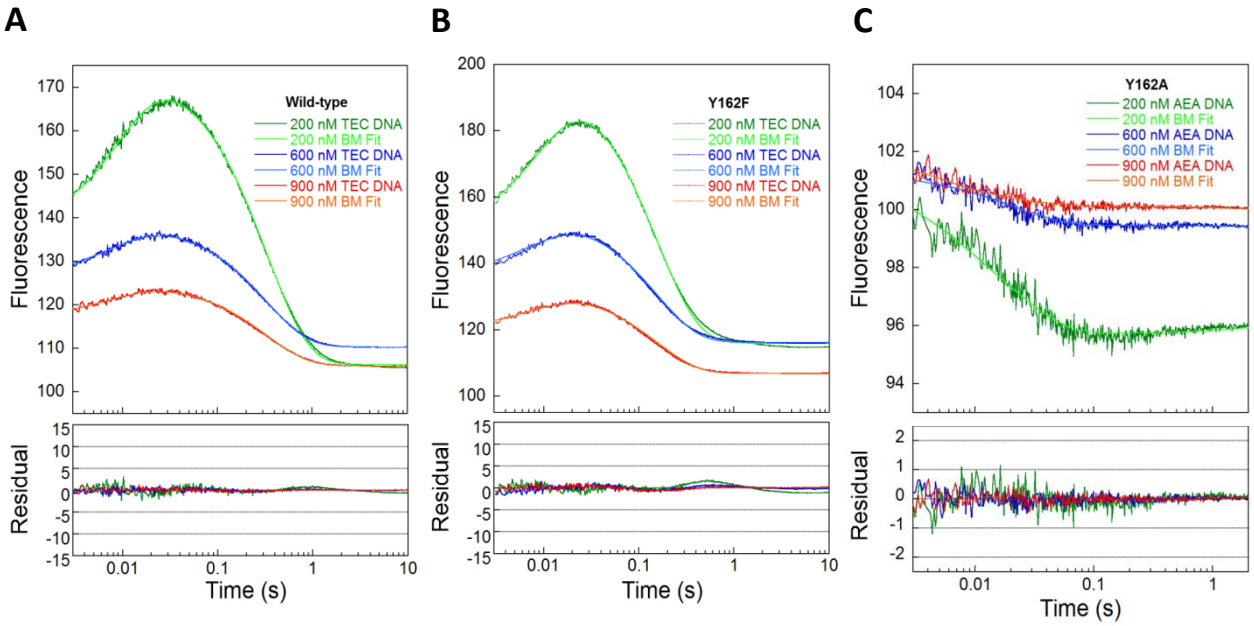
Figure B-2. **Pulse-chase experiment for measuring the partitioning between excision and dissociation of εA-DNA** for wild-type (**A**) Y162F (**B**) and Y162A (**C**) AAG. As a control to test the effectiveness of the unlabeled pyrrolidine-DNA chase, reactions were initiated by adding AAG to a mixture of chase and 5'-fam εA-DNA substrate (triangles). The reactions in the presence (squares) and absence (circles) of chase are from Figure 3.6 in the text. Comparison of the +chase reaction and the chase control for Y162F supports a small, but measurable commitment to excision.



E

Excess Protein	Experimental WT (TEC)	BM WT	Experimental Y162F (TEC)	BM Y162F	Experimental Y162A (TEC)	BM Y162A	Experimental Y162A (AEA)	BM Y162A
$k_1 = k_{on}$	1.1×10^9	1.3×10^9	2.1×10^9	2.4×10^9	n/a	1.8×10^{10}	n/a	1.6×10^9
$k_2 = k_{off}$	n/a	41.4	n/a	62.9	n/a	12.9	n/a	63.4
k_3	n/a	84.2	n/a	93.7	n/a	226.9	n/a	148.5
k_4	n/a	31.3	n/a	15.7	n/a	693.5	n/a	67.6
$k_{find} = k_3 + k_4$	116	115.5	114	109.4	n/a	920.4	n/a	216.1
$k_7 = k_{flip}$	3.6	3.3	7.7	7.5	171.8	171.0	66.0	66.3
$k_8 = k_{unflip}$	0.002	0.0012	0.005	0.005	10.3	10.3	10.3	10.4
$k_9 = k_{artifact}$	0.51 – 1.72	0.012	0.66 – 2.14	1.3	0.03 – 1.8	16.6	1.22 – 12.9	6.5

Figure B-3. **Representative Berkeley Madonna (BM) global fits to excess protein stopped-flow experiments** (**A**) wild-type (**B**) Y162F (**C** and **D**) Y162A AAG. Lines show the best fits obtained with Berkeley Madonna (see Figure B-5), and the plot below shows the residual between the fit and the experimental data. (**E**) Table summarizing individual fits (Table 3.1 in main text) and global analysis of curves shown in A–D.



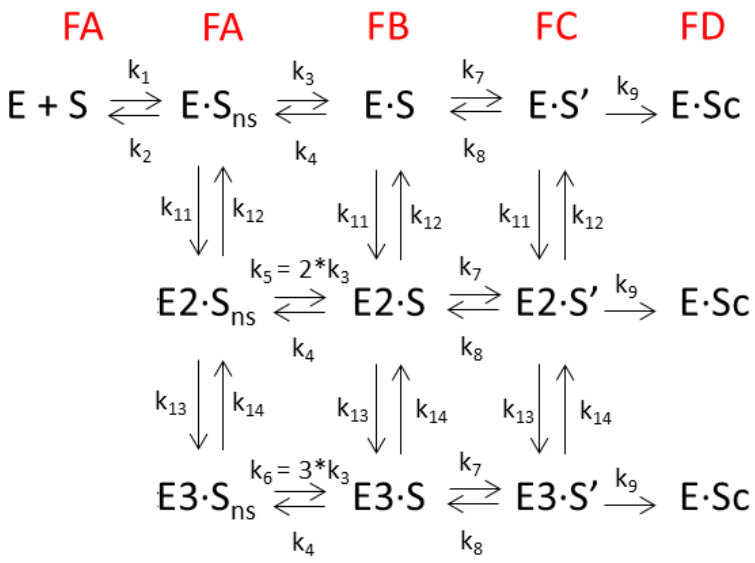
D

Excess DNA	Experimental WT	BM WT (TEC)	Experimental Y162F	BM Y162F (TEC)	Experimental Y162A	BM Y162A (AEA)
$k_1 = k_{on}$	1.1 x 10^9	3.8 x 10^9	2.1 x 10^9	4.3 x 10^9	n/a	1.2 x 10^{10}
$k_2 = k_{off}$	n/a	52.9	n/a	62.3	n/a	65.2
k_3	n/a	104.9	n/a	102.6	n/a	173.4
k_4	n/a	18.3	n/a	16.8	n/a	56.2
$k_{find} = k_3 + k_4$	116	123.2	114	119.4	n/a	229.6
$k_5 = k_{flip}$	3.6	3.6	7.4	7.6	52.6	52.4
$k_6 = k_{unflip}$	0.002	0.0020	0.005	0.005	10.3	10.1
$k_7 = k_{artifact}$	0.74 – 2.98	0.44	1.03 – 2.75	0.77	0.01 – 0.20	0.12

Figure B-4. Representative global fits to stopped-flow experiments with excess DNA and (**A**) wild-type (**B**) Y162F or (**C**) Y162A. Lines show the best fits obtained with Berkeley Madonna (see Figure B-6), and the plot below shows the residual between the fit and the experimental data. (**D**) Table summarizing the exponential fits described in the main text (Table 3.1) and the global analysis of the experimental data at different concentrations of DNA or protein (A–C).

Figure B-5. Example of Berkeley Madonna script for global fits to stopped-flow experiments with excess protein

```
METHOD RK4
STARTTIME = 0
STOPTIME = 9.96
DT = 5e-5

d/dt (E[1..3]) = -k1*E[i]*S[i] + k2*ESns[i] + k12*(E2Sns[i] + E2S[i] + E2S1[i]) + k14*(E3Sns[i] + E3S[i] + E3S1[i])

d/dt (S[1..3]) = -k1*E[i]*S[i] + k2*ESns[i]
d/dt (ESns[1..3]) = k12*E2Sns[i] - (k2 + k3 + k11*E[i])*ESns[i] + k1*E[i]*S[i] + k4*ES[i]
d/dt (E2Sns[1..3]) = k11*E[i]*ESns[i] + k14*E3Sns[i] + k4*E2S[i] - (k12 + k5 + k13*E[i])*E2Sns[i]
d/dt (E3Sns[1..3]) = k13*E[i]*E2Sns[i] + k4*E3S[i] - (k14 + k6)*E3Sns[i]
d/dt (ES[1..3]) = k8*ES1[i] - (k4 + k11*E[i] + k7)*ES[i] + k3*ESns[i] + k12*E2S[i]
d/dt (E2S[1..3]) = k11*E[i]*ES[i] + k5*E2Sns[i] + k14*E3S[i] + k8*E2S1[i] - (k12 + k4 + k13*E[i] + k7)*E2S[i]
d/dt (E3S[1..3]) = k6*E3Sns[i] + k13*E[i]*E2S[i] + k8*E3S1[i] - (k14 + k4 + k7)*E3S[i]
d/dt (ES1[1..3]) = k12*E2S1[i] + k7*ES[i] - (k8 + k11*E[i] + k9)*ES1[i]
d/dt (E2S1[1..3]) = k11*E[i]*ES1[i] + k7*E2S[i] + k14*E3S1[i] - (k8 + k13*E[i] + k9 + k12)*E2S1[i]
d/dt (E3S1[1..3]) = k13*E[i]*E2S1[i] + k7*E3S[i] - (k8 + k14 + k9)*E3S1[i]
d/dt (ESc[1..3]) = k9*ES1[i] + k9*E2S1[i] + k9*E3S1[i]

init S[1..3] = Sub
Sub = 50e-9
init ESns[1..3] = 0
init E2Sns[1..3] = 0
init E3Sns[1..3] = 0
init ES[1..3] = 0
init E2S[1..3] = 0
init E3S[1..3] = 0
init ES1[1..3] = 0
init E2S1[1..3] = 0
init E3S1[1..3] = 0
init ESc[1..3] = 0

init E[1] = E1
init E[2] = E2
init E[3] = E3
E1 = 150e-9
E2 = 250e-9
E3 = 400e-9

k1 = 1.1e9
k2 = 48
k3 = 85
k4 = 30
k5 = 2*k3
```



Scheme for multiple proteins binding, where E = AAG, S = DNA substrate, ESns = nonspecific complex, ES = initial recognition complex, ES' = flipped-out specific recognition complex, ESc = artifact phase. E2 assumes binding of two proteins, while E3 is binding of three proteins. To fit this mechanism to fluorescence data, we also define the fluorescence of each species (FA, FB, FC, and FD).

k6 = 3*k3
k7 = 3.6
k8 = 0.002
k9 = 0.1
k11 = k1
k12 = k2
k13 = k1
k14 = k2

F[1..3] = OS[i] + FA[i]*(S[i] + ESns[i] + E2Sns[i] + E3Sns[i]) + FB[i]*(ES[i] + E2S[i] + E3S[i]) + FC[i]*(ES1[i] + E2S1[i] + E3S1[i]) + FD[i]*ESc[i]

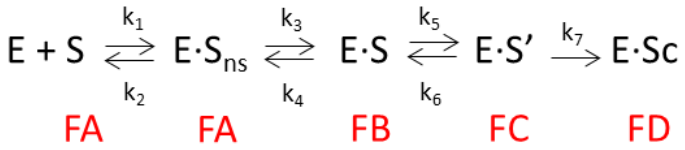
F1 = F[1]
F2 = F[2]
F3 = F[3]
FA[1] = FA1
FA[2] = FA2
FA[3] = FA3
FB[1] = FB1
FB[2] = FB2
FB[3] = FB3
FC[1] = FC1
FC[2] = FC2
FC[3] = FC3
FD[1] = FD1
FD[2] = FD2
FD[3] = FD3

FA1 = 2.0e8
FA2 = 4.0e8
FA3 =5.0e8
FB1 = 6.0e8
FB2 = 6.3e8
FB3 = 6.4e8
FC1 = 2.0e7
FC2 = 5.0e7
FC3 = 9.0e7
FD1 = 5.0e6
FD2 = 1e7
FD3 = 1.5e7

OS[1] = OS1
OS[2] = OS2
OS[3] = OS3
OS1 = 102.3
OS2 = 101.8
OS3 = 102.5

Figure B-6. Example of Berkeley Madonna script for global fits to stopped-flow experiments with excess DNA

```
METHOD RK4
STARTTIME = 0
STOPTIME = 9.96
DT = 5e-5

d/dt (E[1..3]) = -k1*E[i]*S[i] + k2*ESns[i]
d/dt (S[1..3]) = k2*ESns[i] - k1*E[i]*S[i]
d/dt (ESns[1..3]) = k4*ES[i] - (k2 + k3)*ESns[i] + k1*E[i]*S[i]
d/dt (ES1[1..3]) = -k6*ES1[i] + k5*ES[i]  - k7*ES1[i]
d/dt (ES[1..3]) = k6*ES1[i] - (k4 + k5)*ES[i] + k3*ESns[i]
d/dt (ESc[1..3]) = k7*ES1[i]

init E[1..3] = En
En = 100e-9
init ESns[1..3] = 0
init ES[1..3] = 0
init ES1[1..3] = 0
init ESc[1..3] = 0

init S[1] = S1
init S[2] = S2
init S[3] = S3
S1 = 200e-9
S2 = 600e-9
S3 = 900e-9

k1 = 5e9
k2 = 53
k3 = 105
k4 = 18
k5 = 3.6
k6 = 0.002
k7 = 1.1

F[1..3] = OS[i] + FA[i]*S[i] + FA[i]*ESns[i] + FB[i]*ES[i] + FC[i]*ES1[i] + FD[i]*ESc[i]

F1 = F[1]
F2 = F[2]
F3 = F[3]

FA[1] = FA1
FB[1] = FB1
FC[1] = FC1
FD[1] = FD1
FA1 = 3e8
FB1 = 7e8
FC1 = 3e6
FD1 = 1e6
FA[2] = FA2
```



Scheme for binding of a single protein, where E = AAG, S = DNA substrate, ESns = nonspecific complex, ES = initial recognition complex, ES' = flipped-out specific recognition complex, ESc = artifact phase. To fit this mechanism to fluorescence data, we also define the fluorescence of each species (FA, FB, FC, and FD).

FB[2] = FB2
FC[2] = FC2
FD[2] = FD2
FA2 = 2e8
FB2 = 3e8
FC2 = 2e6
FD2 = 1e6
FA[3] = FA3
FB[3] = FB3
FC[3] = FC3
FD[3] = FD3
FA3 = 1e8
FB3 = 2e8
FC3 = 2.5e6
FD3 = 1e6

OS[1] = OS1
OS[2] = OS2
OS[3] = OS3
OS1 = 73
OS2 = 30
OS3 = 20

CHAPTER 4

Evidence for Early Intercalation during the Search for DNA Damage by Human Alkyladenine DNA Glycosylase

DNA is subject to a variety of modifications by reactive intracellular and environmental agents. Failure to repair DNA damage leads to mutations or cell death (1, 2). The base excision repair (BER) pathway is responsible for removing single base lesions from the DNA. This pathway is initiated by a DNA glycosylase that is responsible for finding the damaged site and catalyzing the hydrolysis of the *N*-glycosidic bond. Subsequent action of an abasic-site specific endonuclease, a 5'-deoxyribose phosphate lyase, a DNA repair polymerase, and a DNA ligase are required to restore the correct DNA sequence, using the intact strand as a template.

Human alkyladenine DNA glycosylase (AAG) is a monomeric DNA glycosylase responsible for recognizing a wide variety of structurally diverse deaminated and alkylated purine lesions (3-5). This enzyme gains access to damaged bases that are normally embedded in the duplex through a process called nucleotide or base flipping, which involves the complete 180° rotation of a nucleobase into the enzyme active site (6, 7). Crystal structures of AAG complexed with damaged DNA reveal the DNA is bent where a tyrosine (Y162) intercalates into the space left by the flipped-out damaged nucleotide (Figure 4.1A; (7)). Nucleotide flipping is common in nature,

and it is universally conserved by repair enzymes, however the timing of intercalation remains unknown.

We previously used the intrinsic fluorescence of 1,N^6-ethenoadenine (εA) to characterize the minimal kinetic mechanism for the recognition, flipping, and excision of this common lesion by wild-type AAG (Figure 4.1C; (8, 9)). AAG uses nonspecific binding and facilitated diffusion (10) to search for sites of damage. Stopped-flow experiments indicated that once an εA is found, AAG forms an initial recognition complex in which the εA lesion is partially unstacked (8). From this intermediate, AAG catalyzes nucleotide flipping to form a more stable specific recognition complex in the enzyme active site before *N*-glycosidic bond hydrolysis.

The natural fluorescence of εA allowed for the full kinetic characterization of wild-type AAG, however, the majority of substrates recognized by AAG are not fluorescent. Currently there is no direct method to measure the rate constant of nucleotide flipping for these lesions. Tryptophan fluorescence of proteins is often used as a tool to monitor changes in proteins or DNA, and a tryptophan reporter in AAG would allow measurement of k_{flip} for any lesion. The catalytic domain of AAG contains three tryptophan residues (W243, W270, and W284), raising the possibility that the intrinsic tryptophan fluorescence of wild-type could be used to monitor conformational changes upon binding DNA. However, since the three tryptophan residues are all found near the back surface of the protein, away from the enzyme active site, we predicted their environment would not change upon DNA binding (Figure 4.1A). In the crystal structure of AAG bound to a flipped-out εA lesion, the DNA is bent away from the protein by about 20°, and the intercalating tyrosine-162 is located at the center of the bend (7). The side chain of tyrosine-162 projects from a β-hairpin on the surface of AAG and inserts into

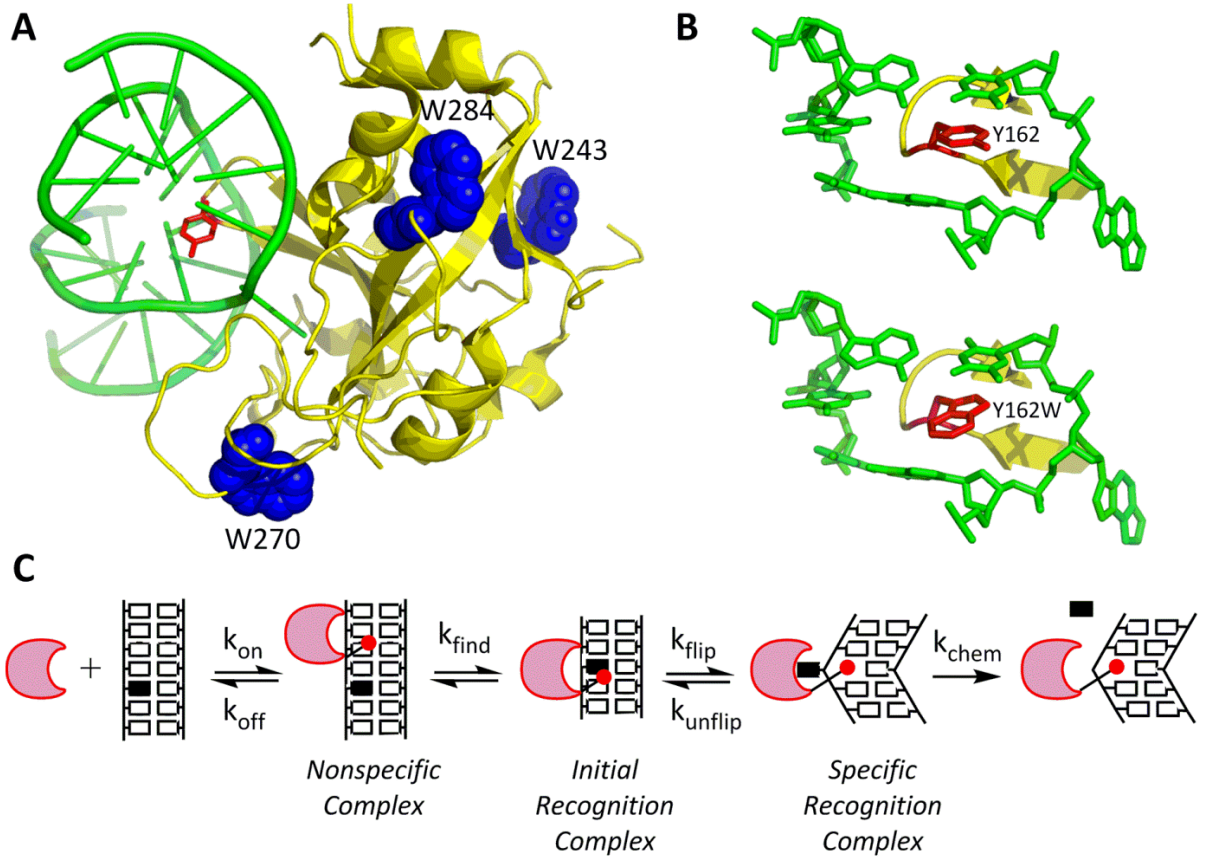


Figure 4.1. AAG uses nucleotide flipping. (**A**) Cartoon representation of the AAG (yellow) complex with εA-DNA (green) from the PDB (1EWN) was rendered with Pymol (http://www.pymol.org). The three native tryptophan residues are shown in space-filling representation (blue). They are all located near the surface of the protein far away from the DNA-binding surface of AAG. The εA base is flipped into the active site, and the intercalating residue (Y162) is shown in red. (**B**) Substitution of tyrosine-162 with tryptophan can be accommodated. Image was rendered with Pymol, and no energy minimization or additional modeling was performed, but this suggests that the slightly larger tryptophan side chain could be accommodated. (**C**) AAG (crescent) binds to nonspecific DNA and rapidly searches for DNA damage. This searching process is described by the macroscopic rate constant k_{find}. Once a lesion (solid rectangle) is encountered in an initial recognition complex, it can be flipped into the enzyme active site to form the specific recognition complex. AAG catalyzes hydrolysis of the *N*-glycosidic bond (k_{chem}) from this complex. As shown by the red ball and stick, the intercalating Y162 residue is important throughout all steps of the kinetic mechanism.

the minor groove DNA, thus filling the space left by the flipped-out nucleotide. Tyrosine-162 was substituted with tryptophan and a rotamer was selected that avoided steric clashes with both DNA and protein side chains (Figure 4.1B). No energy minimization or additional modeling was performed, but this exercise suggests that the slightly larger tryptophan side chain can be accommodated without large-scale changes in other AAG residues. In addition, previous work with Y162 mutants established that phenylalanine at this position is a conservative mutation, and DNA intercalation contributes to specific recognition of DNA damage by acting as a plug to stabilize the specific lesion recognition complex (Chapter 3). Therefore, in this study we mutated the intercalating residue of AAG to tryptophan (Y162W) to potentially generate a fluorescent probe to report on nucleotide flipping.

Substitution of the highly conserved tyrosine with tryptophan at position 162 had minimal effects on the overall kinetic parameters. The slightly bulkier tryptophan residue did not alter the equilibrium constant for flipping, and excision of εA is identical to wild-type. Unexpectedly, upon binding damaged or undamaged DNA, the tryptophan fluorescence of Y162W AAG is rapidly quenched and does not change throughout the timescale of nucleotide flipping. Assignment of Y162W fluorescence was confirmed by characterization of mutation of two of the native tryptophan residues, W270A/W284A. This suggests Y162 intercalates early in the search for DNA damage, and Y162 is critical throughout all steps of the kinetic mechanism.

Materials and Methods

Purification of Wild-type and Mutant AAG Protein. The catalytic domain of human AAG that lacks the first 79 amino acids was expressed in *E. coli* and purified as previously described (11). The Y162W mutant was constructed by site-directed mutagenesis and purified using the same

protocol. The concentration of AAG was determined by fluorescent titration of εA-DNA, as described below, and the concentrations were found to be within 2-fold of the concentration determined from the UV absorbance and the calculated extinction coefficient for the wild-type or mutant protein.

Synthesis and Purification of Oligodeoxynucleotides. The 25mer oligonucleotides were synthesized by Integrated DNA Technologies or by the W.M. Keck Facility at Yale University. The oligonucleotides were desalted using Sephadex G-25 and purified using denaturing polyacrylamide gel electrophoresis as previously described. Oligonucleotides for gel-based assays were labeled on the lesion containing strand with a 5'-fluorescein (6-fam) label. The concentrations of the single stranded oligonucleotides were determined from the absorbance at 260 nm, using the calculated extinction coefficients for all oligonucleotides except those containing εA. For oligonucleotides containing εA, the extinction coefficient was calculated for the same sequence with an A in place of the εA and corrected by subtracting 9400 $M^{-1}cm^{-1}$ to account for the weaker absorbance of εA as compared to A. The lesion-containing oligonucleotides were annealed with a 1.2-fold excess of the complement by heating to 90 °C and cooling slowly to 4 °C. Two different sequence contexts were used and these are referred to by the central three nucleotides (i.e., TEC has a central εA lesion flanked by a 5'T and a 3'C;

TXC 5' –CGATAGCATCCT**X**CCTTCTCTCCAT
3' –GCTATCGTAGGATGGAAGAGAGGTA

AXA 5' –CGATAGCATCAA**X**AATTCTCTCCAT
3' –GCTATCGTAGTTTTTAAGAGAGGTA

X= εA, A, or pyrrolidine

Scheme 4.1

Scheme 4.1). Two DNA sequences were used because the fluorescence is different due to changes in the εA environment (8). The oligonucleotide sequence TEC gave 4-fold lower fluorescence than free εA, consistent with quenching of εA fluorescence even in single-stranded DNA. Surprisingly, the AEA sequence yielded much higher fluorescence, essentially identical to that of free εA. Upon annealing to complement, the fluorescence decreased 1.7-fold for TEC and 1.4-fold for AEA. This is consistent with increased quenching due to base stacking interactions in duplex DNA. It is possible that the lack of nearby G nucleotides can account for the high fluorescence of the AEA sequence. The addition of AAG resulted in a dramatic 5-fold decrease in the fluorescence of the AEA substrate, whereas the TEC substrate only showed slight quenching. The two oligonucleotide sequences exhibit almost identical fluorescence emission when they are bound to AAG, so this strongly suggests that εA is in a similar environment in both complexes (the flipped-out specific recognition complex) (8).

Steady-State Fluorescence Measurements. Fluorescence emission spectra were collected with a PTI QuantaMaster fluorometer controlled by FeliX software. For εA fluorescence, an excitation wavelength of 314 nm (6 nm band-pass) was used and the total fluorescence was measured at emission wavelengths from 340-480 nm (6 nm band-pass). Samples (300 µL) of 400 nM εA-containing DNA were prepared in the standard buffer (50 mM NaMES (pH 6.5), 100 mM NaCl, 1 mM EDTA, 1 mM DTT) and spectra were recorded at 25 °C. To determine the steady-state fluorescence of εA-containing DNA bound to AAG, the spectra were recorded within one minute. No significant excision of εA occurs during this time. Three independent titrations were averaged, and the data were fit to a quadratic equation assuming tight binding

by AAG (Equation 4.1), in which F_{rel} is the relative fluorescence, A is the fractional quenching, and K_d is the dissociation constant for the εA-DNA.

$$F_{rel} = 1 - \frac{A(K_d + E + DNA) - \sqrt{(K_d + E + DNA)^2 - (4E \times DNA)}}{2E} \quad (4.1)$$

Gel-based Glycosylase Assay. We measured single-turnover glycosylase activity with the standard glycosylase activity assay that utilizes abasic site cleavage by NaOH followed by DNA separation on a denaturing polyacramide gel. Fluorescein-labeled DNA substrates (50 nM) containing εA were prepared in the standard buffer. The reactions were initiated with the addition of 75 nM – 1200 nM AAG and incubated at 25 °C. At various time points, a sample from the reaction was removed and quenched in two volumes of 0.3 M NaOH, giving a final hydroxide concentration of 0.2 M. The abasic sites were cleaved by heating at 70 °C for 15 minutes. Samples were mixed with an equal volume of formamide/EDTA loading buffer before loading onto a 15% polyacrylamide gel. Gels were scanned with a Typhoon Imager (GE Trio+ Healthcare) to detect the fluorescein label by exciting at 488 nm and measuring emission with a 520BP40 filter. The gel bands were quantified using ImageQuant TL (GE Healthcare). The data was converted to fraction product [Fraction Product = product / (product + substrate)] and then fit by a single exponential using nonlinear least squares regression with Kaleidagraph (Synergy Software), in which k_{obs} is the rate constant, t is the time, and A is the amplitude (Equation 4.2). In all cases, the same rate constant was measured at two different concentrations of AAG. This demonstrates that AAG was saturating and that the maximal single-turnover rate constant was determined (see Appendix C). Under these conditions, the observed rate constant is equal to the maximal single-turnover rate constant ($k_{obs} = k_{max}$).

$$\text{Fraction product} = A(1 - e^{-k_{obs}t}) \quad (4.2)$$

Stopped-Flow Kinetics. Pre-steady state kinetic experiments were performed on a Hi-Tech SF-61DSX2, controlled by Kinetic Studio (TgK Scientific). The fluorescence of εA was measured using an excitation wavelength of 313 nm and a WG360 long-pass emission filter as previously described (8). The fluorescence of tryptophan was measured using an excitation wavelength of 296 nm and a 330BP20 band-pass emission filter. At least three traces were averaged together at each concentration. The traces for Y162W were fit by a double exponential (Equation 4.3) where *F* is the fluorescence as a function of time, *C* is the fluorescence of free DNA, *X* and *Y* are the changes in fluorescence of the intermediates, and *t* is the time.

$$F = C - X(1 - e^{-k_{1,obs}t}) - Y(1 - e^{-k_{2,obs}t}) \quad (4.3)$$

Observed rate constants were plotted versus concentration and fit to a straight line. $k_{1,obs}$ showed a linear concentration dependence, and the slope is equal to k_{on} ($M^{-1}s^{-1}$), and the y-intercept is equal to k_{off} (s^{-1}). The value of $k_{2,obs}$ was independent of concentration and is equal to $k_{flip} + k_{unflip}$.

Pulse-Chase Assay to Measure Substrate Dissociation. The macroscopic rate constant for dissociation of wild-type and mutant AAG from εA-containing DNA was measured by pulse-chase in the standard reaction buffer at 25 °C as previously described for wild-type AAG (8). Briefly, in 20 µL reactions, 50 nM fluorescein-labeled TEC DNA was mixed with 75 nM - 1200 nM AAG for 20 seconds, and then a chase of 10 µM unlabeled pyrrolidine DNA was added. Pyrrolidine binds tightly to AAG as a transition state analogue (12, 13). At various time points, a sample from the reaction was removed and analyzed as described for “Gel-based glycosylase assay”. Base excision catalyzed by AAG results in fluorescein-labeled product, whereas

dissociation releases unreacted fluorescein-labeled substrate. If AAG dissociates from the labeled DNA before the chemical cleavage step and then binds to the unlabeled DNA, less of the reaction will occur during the single-turnover part of the curve as compared to the same experiment without chase. The data was converted to fraction product and fit by a single exponential (Equation 4.2). The observed rate constant was independent of the concentration of AAG.

According to the two step binding mechanism described in Figure 4.1C, two different partitioning equations can be written (14). All labeled substrate is initially bound, and therefore the fraction of product formed is given by the fraction that goes on to react. This is indicated by Equation 4.4, in which A is the burst amplitude (the fraction of product formed in the burst phase of the experiment), k_{max} is the maximal single-turnover rate constant for formation of product, and $k_{off,obs}$ is the macroscopic rate constant for dissociation from the flipped-out complex. This expression can be rearranged to solve for the desired dissociation rate constant (Equation 4.5). Similarly, for branched pathways, the observed rate constant for the burst phase of the pulse-chase experiment is given by the sum of the rate constants for the competing pathways, formation of product is given by k_{max} and the macroscopic dissociation of substrate is designated $k_{off,obs}$ (Equation 4.6). Solving for $k_{off,obs}$ gives Equation 4.7.

$$A = \frac{k_{max}}{k_{off,obs} + k_{max}} \tag{4.4}$$

$$k_{off,obs} = \frac{k_{max}}{A} - k_{max} \tag{4.5}$$

$$k_{obs} = k_{off,obs} + k_{max} \tag{4.6}$$

$$k_{off,obs} = k_{obs} - k_{max} \quad (4.7)$$

Control reactions in which no chase was added provided the single-turnover rate constant, k_{max}, and confirmed that these concentrations of AAG were saturating. From these values, the dissociation rate constant, k_{off}, for AAG dissociating from εA-DNA was calculated by two different methods (Equations 4.5 and 4.7). Both methods gave similar values for $k_{off,obs}$ and we report the results obtained from Equation 4.5.

Since the εA-DNA binds in two steps, the observed rate constant for dissociation of substrate ($k_{off,\,obs}$) could be limited by the unflipping rate (k_{unflip}) or dissociation from nonspecific DNA (k_{off}). Assuming that the flipped-out complex is stable (i.e., $k_{flip} >> k_{unflip}$), this observed dissociation rate constant can be expressed in terms of the microscopic rate constants (Equation 4.8; (14)) Stopped-flow fluorescence suggests that dissociation from the initial AAG•DNA complex is rapid, and therefore the observed rate constant for substrate dissociation from the εA-DNA•AAG complex is approximately equal to the reverse rate constant for flipping (Equation 4.9).

$$k_{off,obs} = k_{unflip}\left(\frac{k_{off}}{k_{off} + k_{flip}}\right) \quad (4.8)$$

$$k_{off,obs} \approx k_{unflip} \quad (4.9)$$

Results

Binding and Excision of εA by Y162W

We used site-directed mutagenesis to create the Y162W variant, and characterized its ability to bind and excise εA from synthetic 25mer oligonucleotides containing a central εA·T mismatch.

Full characterization of Y162W was necessary to determine whether the tyrosine to tryptophan substitution was conservative. To confirm stable binding of active Y162W to the εA lesion, steady-state fluorescence titrations were used (Figure 4.2A). It was previously shown that εA-DNA of the sequence AEA shows a large decrease in fluorescence upon binding to wild-type AAG (8). The equivalence point of fluorescence titrations with a fixed concentration of εA-DNA and increasing concentrations of protein yields the concentration of active enzyme. The active concentration determined by this method was approximately 80% of the value calculated from the UV absorbance and extinction coefficient. The corrected concentration was used for all experiments described below.

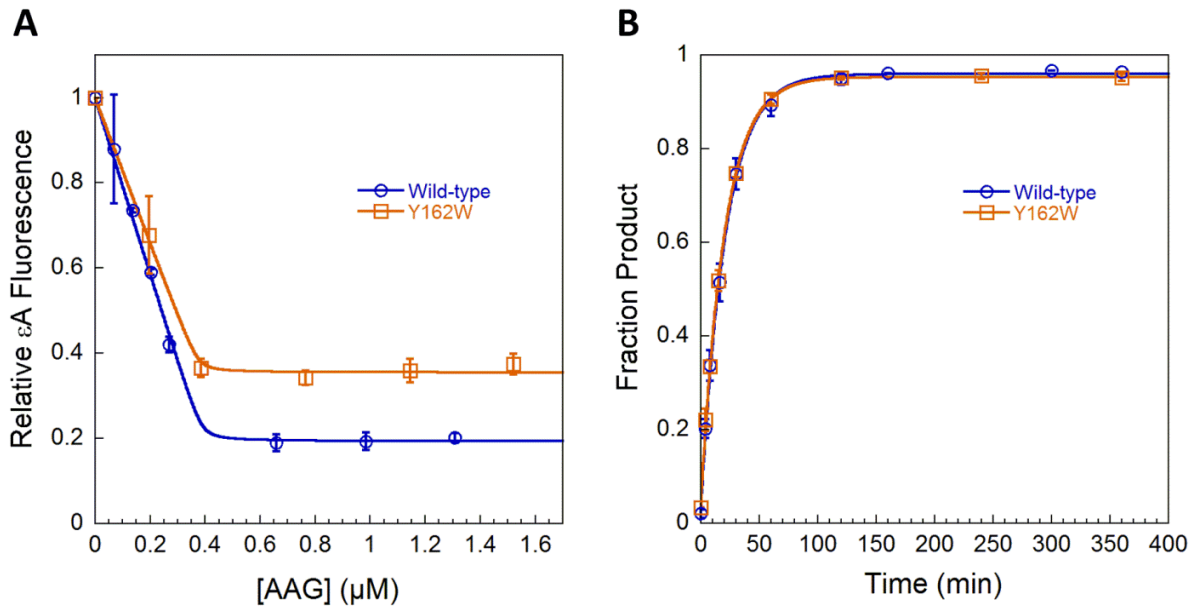


Figure 4.2. Binding and excision of εA by wile-type and Y162W AAG. (**A**) Titration of 0.4 μM AEA DNA with wild-type and Y162W. Normalized fluorescence was fit by the quadratic equation under conditions of tight binding (Equation 4.1). These fits indicate that the number of εA-binding sites is 86% of the expected value for wild-type and 100% of the expected value for Y162W. Titration also shows that both proteins stably bind εA. (**B**) Single-turnover excision of εA by wild-type and mutant AAG with 50 nM TEC DNA and saturating enzyme was fit by a single exponential (Equation 4.2) to obtain the rate constants for *N*-glycosidic bond cleavage (Table 4.1). Data points are an average ± SD from three independent experiments.

Single-turnover experiments with enzyme in excess over εA-DNA substrate were performed to evaluate whether Y162W maintained glycosylase activity (Figure 4.2B). For wild-type AAG, the maximal single-turnover rate constant is *N*-glycosidic bond hydrolysis. The rate constant for excision of εA by Y162W is identical to the previously reported value for wild-type AAG (Chapter 3 and (8)). This provides evidence that the slightly larger tryptophan residue does not disrupt the transition state for *N*-glycosidic bond hydrolysis. However, AAG binds tightly to εA-DNA, so it is possible that the Y162W mutation impairs lesion recognition or nucleotide flipping and this effect is masked by tight εA binding.

Stopped-flow Fluorescence to Monitor Binding and Flipping of εA-DNA

To evaluate whether the Y162W mutant alters the rate constants for DNA binding or nucleotide flipping, rapid mixing experiments were performed to monitor transient changes in εA fluorescence. When a fixed concentration of TEC εA-DNA was mixed with increasing concentrations of Y162W, we observed an initial decrease in fluorescence that was followed by an increase in fluorescence (Figure 4.3A). This was unexpected because wild-type AAG shows the opposite trend under the same conditions (Figure 4.3A; (8)). The concentration of Y162W was varied, and each trace was fit to a double exponential (Figure 4.3B). The first phase is linearly dependent on enzyme concentration (Figure 4.3C) and this is assigned to binding and formation of the initial recognition complex (Chapter 3 and (8)). This initial step yielded an observed rate constant for binding that is nearly identical to wild-type AAG under these conditions (Table 4.1). The concentration independent second phase corresponds to the nucleotide flipping step and formation of the specific lesion recognition complex (Figure 4.3D). The second step is approximately 2-fold faster for the mutant than for wild-type AAG. Since this

Table 4.1. Kinetic parameters for flipping and excision of εA by wild-type and Y162W AAG[a]

	Wild-type[b]	Y162W
k_{on} ($M^{-1}s^{-1}$)	(1.1 ± 0.03) x 10^9	(1.2 ± 0.6) x 10^9
k_{flip} (s^{-1})	3.6 ± 0.7	7.9 ± 0.7
k_{unflip} (s^{-1})	(1.6 ± 0.3) x 10^{-3}	(5.5 ± 0.2) x 10^{-3}
K_{flip}[c]	2300	1400
$k_{max\ \varepsilon A}$ (s^{-1})	(8.0 ± 0.6) x 10^{-4}	(8.3 ± 0.5) x 10^{-4}

[a]All rate constants were determined using the TEC oligonucleotide. The standard conditions were 25 °C, 50 mM NaMES (pH 6.5), 100 mM NaCl, 1 mM EDTA, and 1 mM DTT.

[b]The values for wild-type AAG were previously reported (Chapter 3).

[c]The equilibrium constant for flipping is given by the ratio of the flipping and unflipping rate constants ($K_{flip} = k_{flip}/k_{unflip}$).

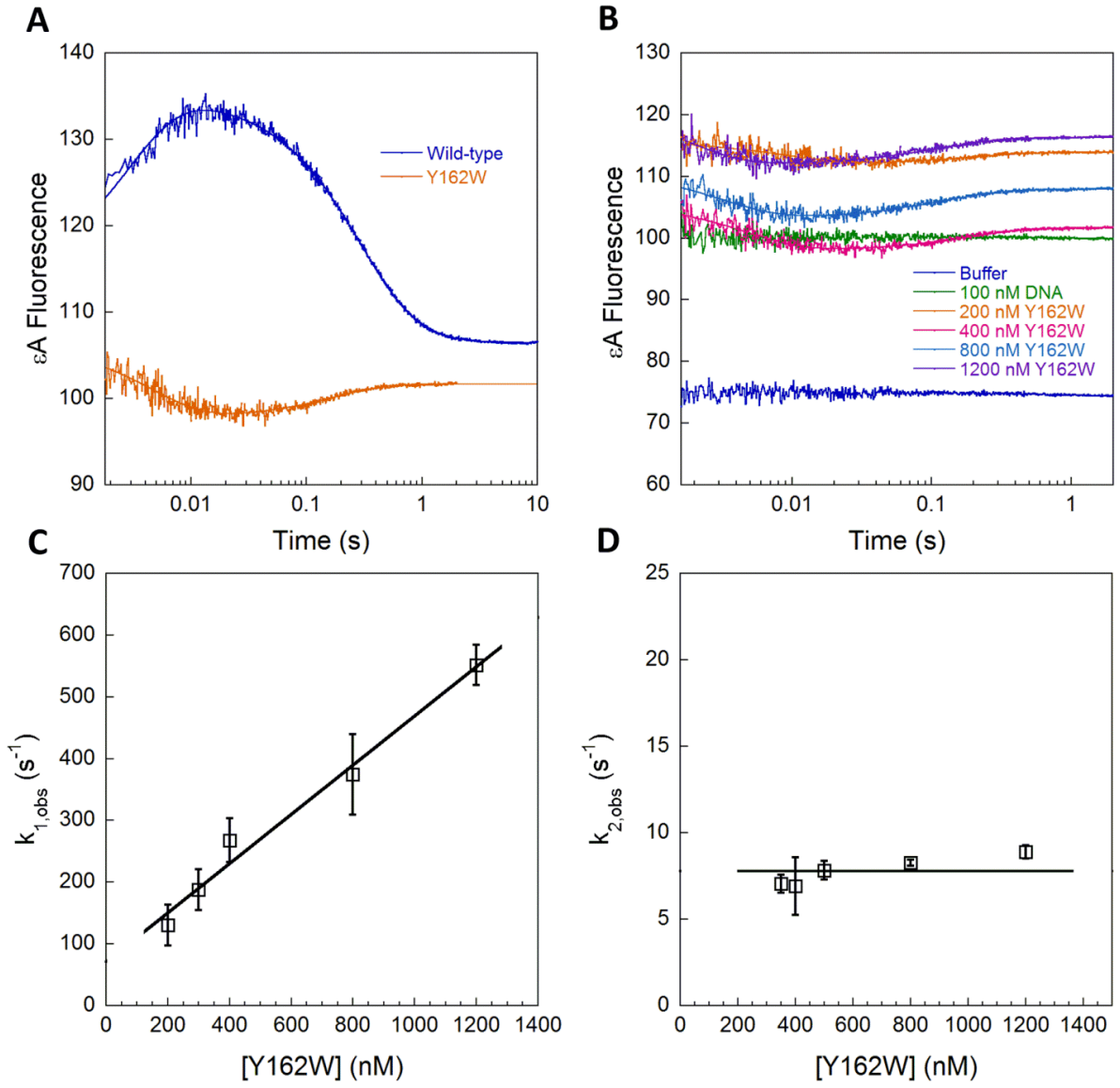


Figure 4.3. Stopped-flow fluorescence with excess protein to measure binding and nucleotide flipping by Y162W. (**A**) Opposite changes in εA fluorescence by wild-type and Y162W (400 nM protein plus 50 nM or 100 nM TEC DNA). (**B**) Representative data from experiments in which 100 nM TEC DNA was mixed with increasing concentrations of Y162W. Traces are the average of three binding reactions and are fit by the sum of exponentials (Equation 4.3). (**C**) The rate constant for the first phase of the binding reaction ($k_{1,obs}$) is dependent on the concentration of AAG and a linear fit yields the bimolecular rate constant for binding (k_{on}) with the values summarized in Table 4.1. (**D**) The rate constants for the second phase of the binding reaction ($k_{2,obs}$) are independent of the concentration of AAG and reflect the sum of the forward and reverse rate constants for nucleotide flipping. Rate constants in C and D are from three independent experiments (average ± SD). Results for wild-type AAG were previously published (8).

observed rate constant for nucleotide flipping is an approach to equilibrium, it is equal to the sum of the rate constants for flipping and unflipping. Therefore, dissociation kinetics must also be monitored in order to determine the macroscopic rate constants for nucleotide flipping.

Pulse-chase Experiment to Measure Unflipping and Dissociation of εA-DNA

We performed a pulse-chase experiment in which wild-type or Y162W AAG are mixed with fluorescently labeled εA-DNA and then chased with an excess of pyrrolidine inhibitor DNA. All AAG that dissociates is immediately bound by the inhibitor, so the partitioning between dissociation and base excision can be directly measured. As previously reported, 70% of wild-type AAG dissociates from the substrate, while 30% partitions to product (8, 9). Assuming that dissociation of AAG from nonspecific DNA is fast, the observed rate constant is simply the rate

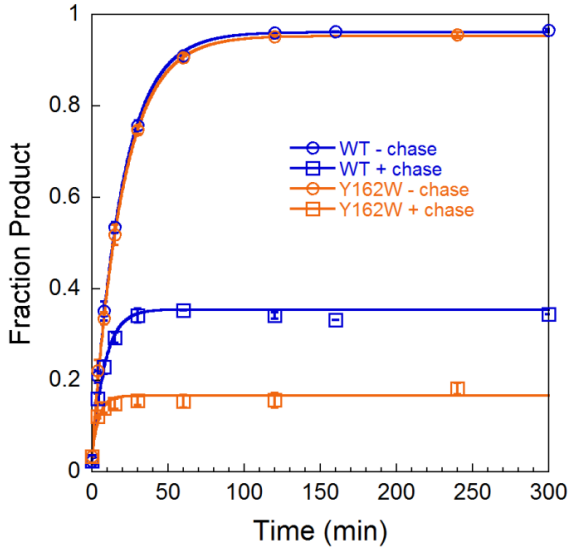


Figure 4.4. Pulse-chase experiments to measure dissociation of εA-DNA. Fluorescein (fam)-labeled DNA (50 nM) was mixed with excess AAG (75 nM - 1200 nM) for 20 seconds, and then 10 μM unlabeled pyrrolidine DNA was added as a chase. The reactions were quenched at the indicated time points, and the fraction of abasic DNA product was determined by alkaline hydrolysis and gel electrophoresis (see Material and Methods for details). The partitioning between the forward reaction (base excision) and substrate dissociation was measured. The commitment for wild-type (30%) and Y162W (13%) were used to calculate the rates of unflipping (see Materials and Methods). The low commitment for Y162W AAG was reproducible and readily distinguished from the control in which chase and substrate are added at the same time (see Appendix C).

constant for unflipping. For the Y162W mutant, 87% of the substrate dissociated (Figure 4.4), indicating a 3-fold faster value for k_{unflip} compared to wild-type. This small amount of product is readily distinguished from the mixing control in which chase and substrate are added together (see Appendix C).

The results from the association and dissociation experiments were used to calculate the rate and equilibrium constants for nucleotide flipping (Table 4.1). The microscopic rate constant for flipping was calculated from the observed rate constant for formation of the extrahelical complex ($k_{flip} = k_{2,obs} - k_{unflip}$). For both wild-type and Y162W, the unflipping rate constants are significantly slower than the observed rate constants for flipping. This indicates that the observed rate constant that was assigned to nucleotide flipping ($k_{2,obs}$) is approximately equal to the microscopic rate constant for flipping (k_{flip}). The equilibrium constant for nucleotide flipping ($K_{flip} = k_{flip}/k_{unflip}$) by Y162W is only 1.6-fold less favorable than wild-type, thus confirming that Y162W is a conservative mutation.

Stopped-flow Fluorescence of Tryptophan

To determine whether the tryptophan fluorescence of the Y162W protein reports on nucleotide flipping, we followed the fluorescence of tryptophan during the rapid mixing experiments. Previous control reactions in which the tryptophan fluorescence was recorded for wild-type AAG mixed with εA-DNA established that none of the three tryptophan residues present in AAG are sensitive to nucleotide flipping (9). Upon mixing a fixed concentration of Y162W with increasing concentrations of εA-DNA, a rapid quenching of tryptophan fluorescence was observed (Figure 4.5A). Surprisingly, the tryptophan fluorescence of Y162W did not change during the timescale of nucleotide flipping. The 20% quenching of tryptophan fluorescence was

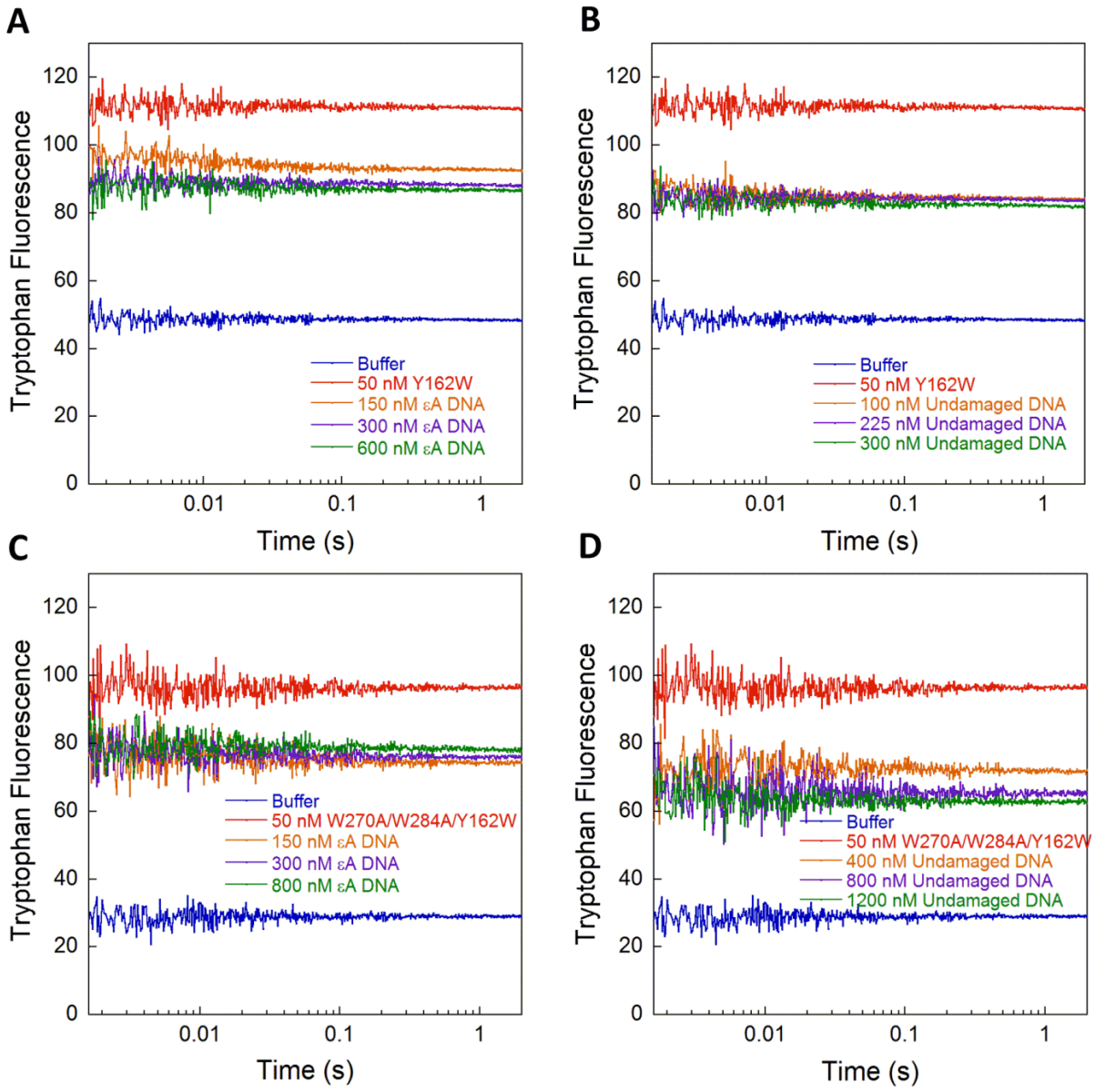


Figure 4.5. Stopped-flow fluorescence to monitor changes in tryptophan fluorescence. Representative data from experiments in which 50 nM Y162W (**A** and **B**) or 50 nM W270A/W284A/Y162W (**C** and **D**) was mixed with increasing concentrations of damaged TEC DNA (A and C) or undamaged TAC DNA (B and D). Samples were excited at 296 nm, and emission was collected with a 330BP20 band-pass filter. Regardless of DNA type, both Y162W and W270A/W284A/Y162W show a rapid 20% decrease in tryptophan fluorescence. The change in fluorescence remains constant, and there is no change in fluorescence on the timescale of nucleotide flipping. Photobleaching was observed for shots longer than two seconds.

highly reproducible and suggests that the environment of Y162W changes rapidly upon binding DNA and is not sensitive to the nucleotide flipping step. The experiment was repeated with undamaged DNA (Figure 4.5B), and interestingly, we still observed the rapid 20% quenching of tryptophan fluorescence. Therefore, the tryptophan fluorescence of Y162W reports on nonspecific DNA binding, and Y162W is not sensitive to formation of either the initial recognition complex, or the specific recognition complex. However, quenching of the other three native tryptophan residues may have masked changes in tryptophan fluorescence from Y162W. Therefore we attempted to minimize the background tryptophan fluorescence by mutating the native tryptophan residues.

Assignment of Y162W Tryptophan Fluorescence

The catalytic domain of AAG contains three tryptophan residues (W243, W270, and W284). All three tryptophan residues are located near the back surface of the protein, away from the enzyme active site (Figure 4.1A). Nevertheless, changes in tryptophan fluorescence were observed upon binding of wild-type AAG to damaged or undamaged DNA in steady-state experiments (Figure 4.6). Therefore, tryptophan mutants were made in an attempt to minimize the background tryptophan fluorescence. Mutation of all three native tryptophan residues resulted in insoluble protein; however both W270 and W284 could be mutated. Two soluble mutants were made, W270A/W284A and W270A/W284A/Y162W. Both mutants stably bind εA and they excise εA only 2-fold slower than wild-type AAG (see Appendix C). Additional experiments showed that W270A/W284A and W270A/W284A/Y162W bind to DNA faster, and the unflipping rate constant is faster than wild-type (see Appendix C). It is not clear why mutating two tryptophan residues on the back side of the protein alters the kinetic

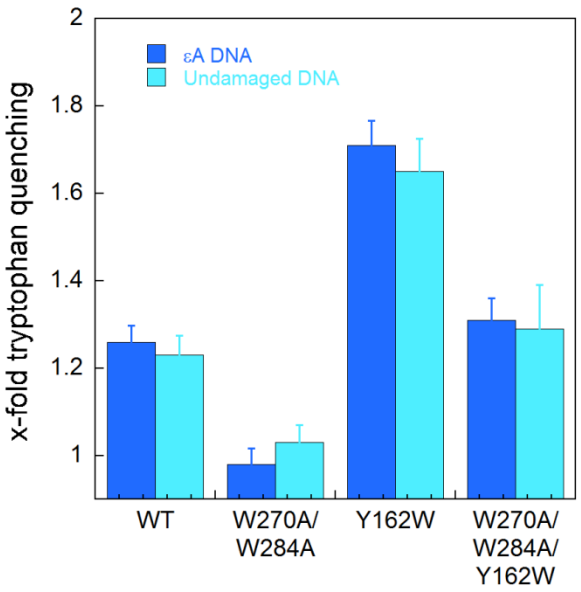


Figure 4.6. Steady-state quenching of AAG tryptophan fluorescence upon binding damaged or undamaged DNA. The three native tryptophan residues in wild-type AAG (W243, W270, and W284) showed quenching of tryptophan fluorescence upon stable binding of damaged or undamaged DNA. Mutation of W270 and W284 to alanine removed all background tryptophan fluorescence. The remaining tryptophan, W243, does not report on DNA binding. Therefore, the tryptophan quenching observed by W270A/W284A/Y162W is due to changes in Y162W tryptophan fluorescence.

parameters, however understanding why these mutations change k_{flip} and k_{unflip} would require mutational analysis which is out of the scope of this work. Nevertheless, we have confidence changes in tryptophan fluorescence can be monitored because both mutants stably bind εA lesions, and excision of εA is similar to wild-type.

Based on steady-state titrations with W270A/W284A, no change in tryptophan fluorescence was observed upon binding damaged or undamaged DNA (Figure 4.6). Therefore, the tryptophan fluorescence of the remaining tryptophan residue, W243, does not change upon binding DNA. The triple mutant, W270A/W284A/Y162W, successfully removed the background tryptophan fluorescence and was soluble under our reaction conditions. The rapid mixing experiment was repeated with a fixed concentration of this mutant and increasing

concentrations of TEC DNA. A similar 20% quenching of tryptophan fluorescence was observed as before (Figure 4.5 C & D). The fast quenching of tryptophan fluorescence was observed for both εA-DNA and undamaged DNA, so we know the Y162W tryptophan fluorescence reports on nonspecific binding. This suggests that Y162 intercalation occurs early in the search for DNA damage.

Discussion

Monitoring Nucleotide Flipping

The rates of DNA binding and nucleotide flipping are critical to understanding how a DNA glycosylase recognizes damaged nucleotides. Several previous studies have used the incorporation of fluorescent reporter groups into DNA, like 2-aminopurine, to investigate the kinetics of DNA binding and nucleotide flipping by other families of DNA glycosylases (15-20). Direct measurement of nucleotide flipping of non-fluorescent lesions is needed because evidence suggests nucleotide flipping provides a barrier to the excision of damaged bases by AAG (5, 21). Knowing the rate of nucleotide flipping for various AAG substrates will help us better understand damage recognition and removal. To generate a tryptophan reporter to directly monitor structural changes in the protein upon binding DNA, we used site-directed mutagenesis to make the Y162W variant. In wild-type AAG, the intercalating residue, Y162, is critical for specific recognition of DNA damage. This residue also acts as a plug to stabilize the specific lesion recognition complex (Chapter 3).

Y162W is a Conservative Mutation

Although tyrosine-162 is highly conserved, the crystal structure raised the possibility that a tryptophan substitution might be tolerated. We used the intrinsic fluorescence of the εA lesion

to determine the kinetic mechanism for the Y162W mutant. Overall this is a very conservative mutation. While both k_{flip} and k_{unflip} are marginally faster for the mutant, the slightly bulkier tryptophan residue did not alter the equilibrium constant for flipping, and excision of εA is identical to wild-type (Table 4.1). This was surprising because the Y162W variant has never been found in nature. Also, extensive mutagenesis of AAG and selection for active site variants that conferred resistance to DNA alkylating agents have identified hundreds of allowed mutations, but no active variants at position 162. However, our results with εA suggest Y162W is similar to wild-type AAG.

Timing of DNA Intercalation

Unexpectedly, the Y162W tryptophan fluorescence does not report on nucleotide flipping. Upon mixing increasing concentrations of damaged or undamaged DNA with a fixed concentration of Y162W, a rapid 20% quenching of tryptophan fluorescence was observed. No changes in fluorescence were observed on the timescale of nucleotide flipping. This suggests that DNA intercalation occurs early in the search for DNA damage. In structural studies of AAG bound to an εC lesion, tyrosine-162 was capable of forming both lower- and higher- affinity interactions with the DNA, possibly playing roles in both a lower-affinity searching process, and a higher affinity recognition process (22). Using tryptophan mutants, we were able to observe changes in fluorescence upon formation of the lower-affinity searching interactions. While this rapid quenching of fluorescence was highly reproducible, we wanted to remove the background fluorescence from the native tryptophan residues (W243, W270, and/or W284) to confirm the result.

Multiple combinations of mutations were tested; however W270A/W284A and W270A/W284A/Y162W were the most soluble. Both tryptophan mutants stably bind εA, and they retain glycosylase activity (see Appendix C). In rapid mixing experiments, when mixed with increasing concentrations of damaged or undamaged DNA, the W270A/W284A/Y162W mutant showed a rapid 20% quenching of tryptophan fluorescence that was similar to what was seen for Y162W AAG (Figure 4.5 C & D). Therefore, the environment of Y162W changes rapidly upon binding DNA, and these results suggest that intercalation occurs early during the search for damage. However we cannot rule out the possibility that other conformational changes result in changes in fluorescence that are too small to detect.

This early role of intercalation appears to be common among independently evolved enzymes. Using a single Qdot-labeled glycosylase molecule, it was previously shown that Fpg glycosylase uses an intercalating phenylalanine residue to test the strength and flexibility of the base pairs, and mutation of the phenylalanine to alanine increased the rate of diffusion (23). The work was expanded to include three bacterial glycosylases in two structural families: Fpg, Nei, and Nth (24). Single molecule studies combined with mutagenesis showed the glycosylases diffuse along DNA, and pause momentarily to interrogate random bases. It is only when a damaged base is recognized, that the enzyme stops and excises it (24). Consistent with our results, this work suggested that the intercalating residue potentially acts as a sensor for DNA damage (23, 24).

Implications

DNA glycosylases are faced with the task of finding rare DNA damage in a sea of undamaged DNA. These enzymes use nonspecific binding and facilitated diffusion to search for sites of

damage (10). We now understand that the intercalating residue plays an important role during this search for damage. By mutating the intercalating tyrosine in AAG to tryptophan, we observed a rapid quenching of tryptophan fluorescence upon binding damaged or undamaged DNA. Strikingly, there were no additional changes in tryptophan fluorescence throughout the timescale of nucleotide flipping. Previous work established that tyrosine 162 plays a critical role in the discrimination between damaged and undamaged nucleotides, and these results suggest that DNA intercalation occurs early in the reaction.

References

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

2. Robertson AB, Klungland A, Rognes T, & Leiros I (2009) DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cell Mol Life Sci* 66(6):981-993.

3. Engelward BP*, et al.* (1997) Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. *Proceedings of the National Academy of Sciences of the United States of America* 94(24):13087-13092.

4. Larsen E, Meza TJ, Kleppa L, & Klungland A (2007) Organ and cell specificity of base excision repair mutants in mice. *Mutation research* 614(1-2):56-68.

5. O'Brien PJ & Ellenberger T (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase. *The Journal of biological chemistry* 279(11):9750-9757.

6. Roberts RJ & Cheng X (1998) Base flipping. *Annu Rev Biochem* 67:181-198.

7. Lau AY, Wyatt MD, Glassner BJ, Samson LD, & Ellenberger T (2000) Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG. *Proceedings of the National Academy of Sciences of the United States of America* 97(25):13573-13578.

8. Wolfe AE & O'Brien PJ (2009) Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by human alkyladenine DNA glycosylase. *Biochemistry* 48(48):11357-11369.

9. Hendershot JM, Wolfe AE, & O'Brien PJ (2011) Substitution of active site tyrosines with tryptophan alters the free energy for nucleotide flipping by human alkyladenine DNA glycosylase. *Biochemistry* 50(11):1864-1874.

10. Hedglin M & O'Brien PJ (2008) Human alkyladenine DNA glycosylase employs a processive search for DNA damage. *Biochemistry* 47(44):11434-11445.

11. O'Brien PJ & Ellenberger T (2003) Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines. *Biochemistry* 42(42):12418-12429.

12. Lau AY, Scharer OD, Samson L, Verdine GL, & Ellenberger T (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. *Cell* 95(2):249-258.

13. Scharer OD, Nash HM, Jiricny J, Laval J, & Verdine GL (1998) Specific binding of a designed pyrrolidine abasic site analog to multiple DNA glycosylases. *The Journal of biological chemistry* 273(15):8592-8597.

14. Hsieh J, Walker SC, Fierke CA, & Engelke DR (2009) Pre-tRNA turnover catalyzed by the yeast nuclear RNase P holoenzyme is limited by product release. *RNA* 15(2):224-234.

15. Stivers JT, Pankiewicz KW, & Watanabe KA (1999) Kinetic mechanism of damage site recognition and uracil flipping by Escherichia coli uracil DNA glycosylase. *Biochemistry* 38(3):952-963.

16. Wong I, Lundquist AJ, Bernards AS, & Mosbaugh DW (2002) Presteady-state analysis of a single catalytic turnover by Escherichia coli uracil-DNA glycosylase reveals a "pinch-pull-push" mechanism. *The Journal of biological chemistry* 277(22):19424-19432.

17. Bellamy SR, Krusong K, & Baldwin GS (2007) A rapid reaction analysis of uracil DNA glycosylase indicates an active mechanism of base flipping. *Nucleic acids research* 35(5):1478-1487.

18. Walker RK, McCullough AK, & Lloyd RS (2006) Uncoupling of nucleotide flipping and DNA bending by the t4 pyrimidine dimer DNA glycosylase. *Biochemistry* 45(47):14192-14200.

19. Kuznetsov NA*, et al.* (2005) Kinetics of substrate recognition and cleavage by human 8-oxoguanine-DNA glycosylase. *Nucleic acids research* 33(12):3919-3931.

20. Kuznetsov NA*, et al.* (2007) Pre-steady-state kinetic study of substrate specificity of Escherichia coli formamidopyrimidine--DNA glycosylase. *Biochemistry* 46(2):424-435.

21. Vallur AC, Feller JA, Abner CW, Tran RK, & Bloom LB (2002) Effects of hydrogen bonding within a damaged base pair on the activity of wild type and DNA-intercalating mutants of human alkyladenine DNA glycosylase. *The Journal of biological chemistry* 277(35):31673-31678.

22. Setser JW, Lingaraju GM, Davis CA, Samson LD, & Drennan CL (2012) Searching for DNA lesions: structural evidence for lower- and higher-affinity DNA binding conformations of human alkyladenine DNA glycosylase. *Biochemistry* 51(1):382-390.

23. Dunn AR, Kad NM, Nelson SR, Warshaw DM, & Wallace SS (2011) Single Qdot-labeled glycosylase molecules use a wedge amino acid to probe for lesions while scanning along DNA. *Nucleic acids research* 39(17):7487-7498.

24. Nelson SR, Dunn AR, Kathe SD, Warshaw DM, & Wallace SS (2014) Two glycosylase families diffusively scan DNA using a wedge residue to probe for and identify oxidatively damaged bases. *Proceedings of the National Academy of Sciences of the United States of America* 111(20):E2091-2099.

Appendix C

Additional data figures and discussion to support Chapter 4

Table C-1. Kinetic parameters for flipping and excision of εA by wild-type and tryptophan mutants[a]

	Wild-type[b]	Y162W	W270A/W284A	W270A/W284A/Y162W
k_{on} ($M^{-1}s^{-1}$)	(1.1 ± 0.03) x 10^9	(1.2 ± 0.6) x 10^9	(7.5 ± 0.07) x 10^8	Fast[c]
k_{flip} (s^{-1})	3.6 ± 0.7	7.9 ± 0.7	9.7 ± 0.3	28.6 ± 0.6
$k_{max\ \varepsilon A}$ (s^{-1})	(8.0 ± 0.6) x 10^{-4}	(8.3 ± 0.5) x 10^{-4}	(2.8 ± 0.1) x 10^{-4}	(3.4 ± 0.2) x 10^{-4}

[a]All rate constants were determined using the TEC oligonucleotide. The standard conditions were 25 ^{o}C, 50 mM NaMES (pH 6.5), 100 mM NaCl, 1 mM EDTA, and 1 mM DTT.

[b]The values for wild-type AAG were previously reported (Chapter 3).

[c]k_{on} is too fast to measure for W270A/W284A/Y162W

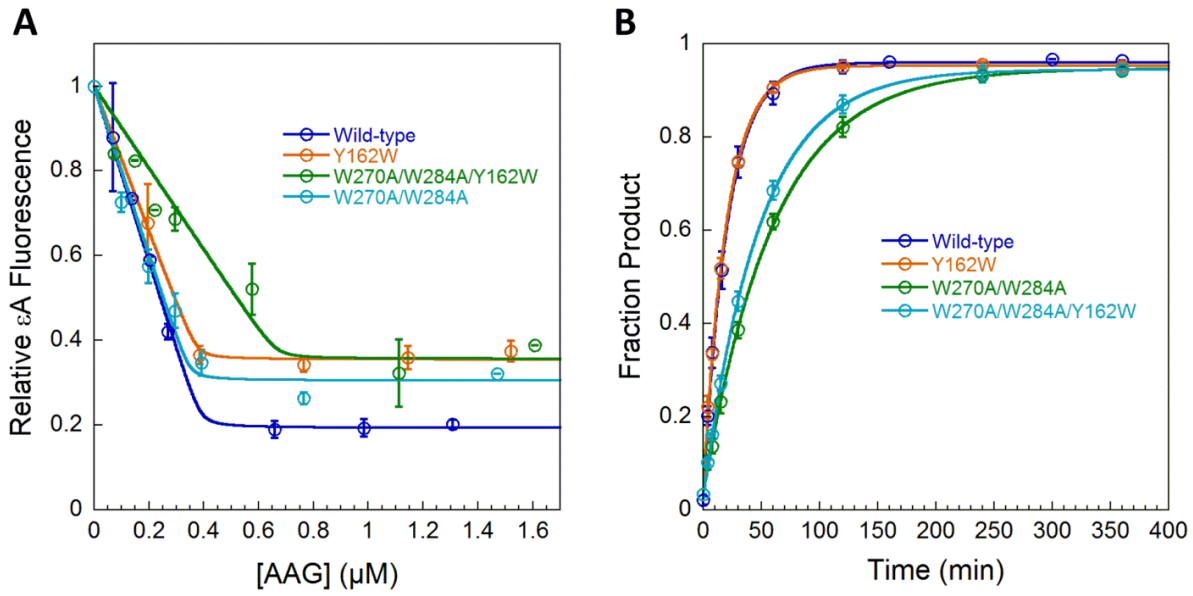


Figure C-1. Binding and excision of εA by tryptophan mutants. (**A**) Titration of 0.4 μM εA-DNA with wild-type and tryptophan mutants. Normalized fluorescence was fit by the quadratic equation under conditions of tight binding. Titration shows that all tryptophan mutants stably bind εA, and the accurate concentration of protein was determined from the equivalence point. The fits indicate that the number of εA-binding sites is 86% of the expected value for wild-type, 100% for Y162W, 80% for W270A/W284A, and 31% for W270A/W284A/Y162W. (**B**) Single-turnover excision of εA by wild-type and mutant AAG with 50 nM DNA and saturating (75 nM - 1200 nM) enzyme was fit by a single exponential (Equation 4.2) to obtain the rate constants for *N*-glycosidic bond cleavage (Table C-1). All tryptophan mutants retained catalytic activity. Data points are an average ± SD from three independent experiments.

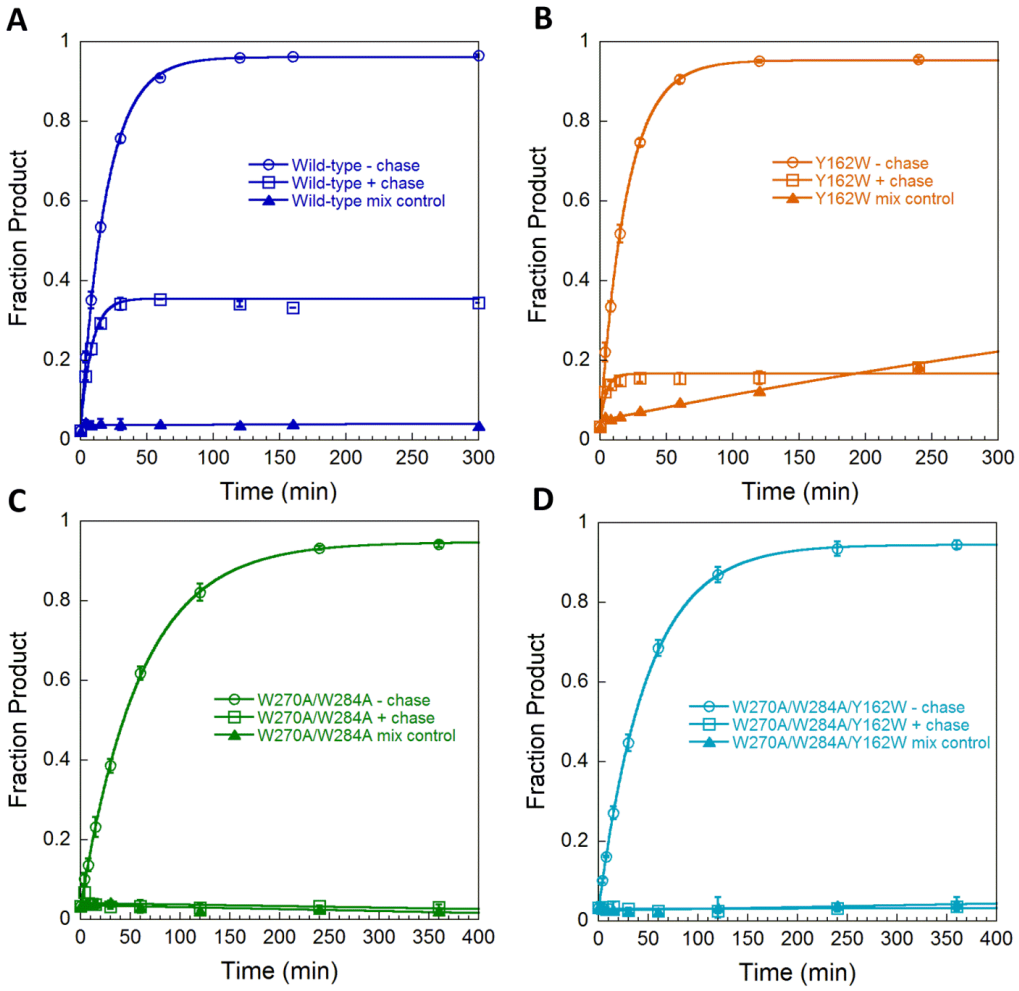


Figure C-2. Pulse-chase experiment for measuring the partitioning between excision and dissociation of εA-DNA for wild-type (**A**) Y162W (**B**) W270A/W284A (**C**) and W270A/W284A/Y162W (**D**) AAG. As a control to test the effectiveness of the unlabeled pyrrolidine-DNA chase, reactions were initiated by adding AAG to a mixture of chase and 5'-fam εA-DNA substrate (triangles). The reactions in the presence (squares) and absence (circles) of chase for wild-type and Y162W are from Figure 4.4 in the text. Comparison of the +chase reaction and the chase control for Y162W supports a small, but measurable commitment to excision. Should be noted this control reaction was only performed once with the Y162W mutant. However, 100% W270A/W284A and W270A/W284A/Y162W dissociated and no product could be detected in the presence of chase. This indicates that unflipping is must faster than *N*-glycosidic bond cleavage. In order to precisely measure the rate of unflipping for these mutants, double-mixing stopped-flow experiments are required.

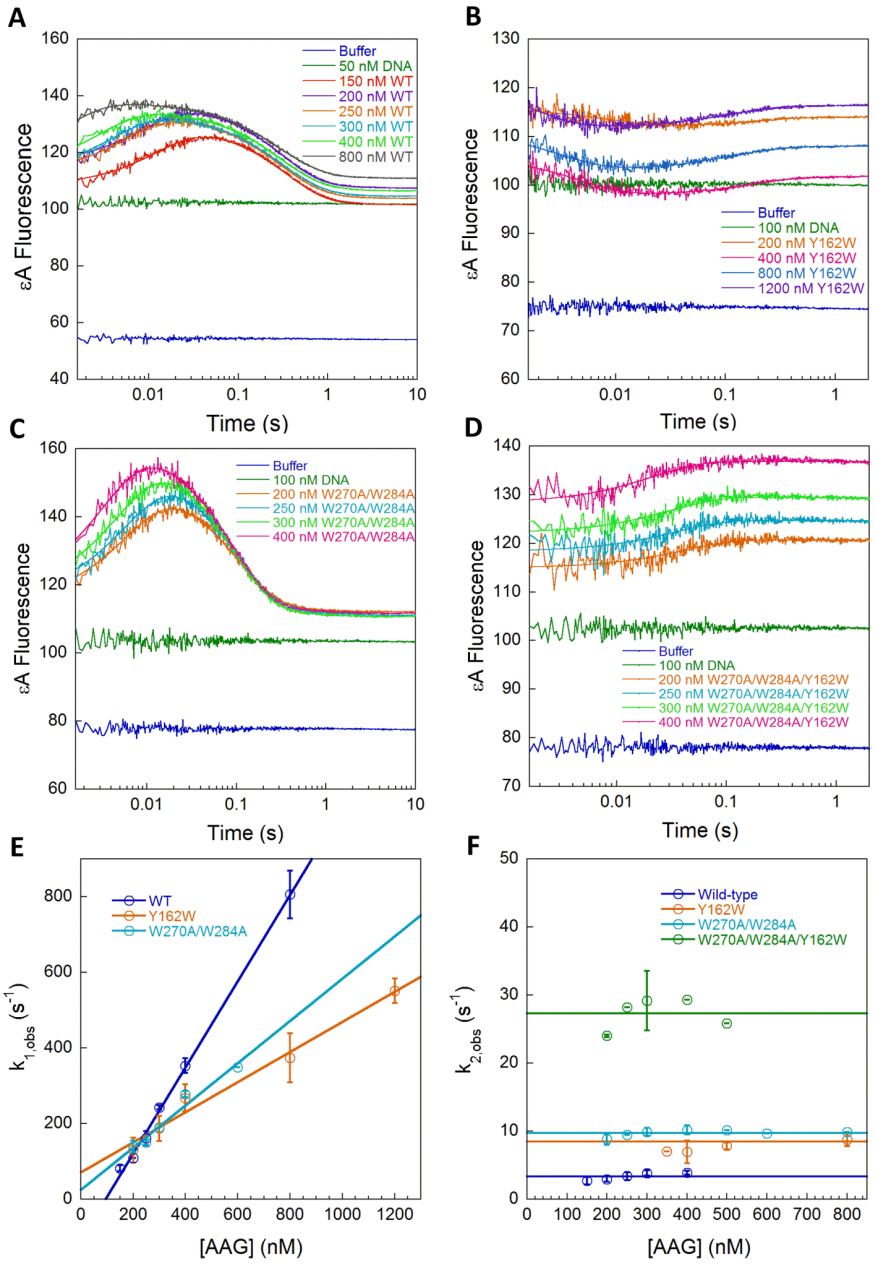


Figure C-3. Stopped-flow fluorescence with excess protein to measure binding and nucleotide flipping by tryptophan mutants. Representative data from experiments in which TEC DNA was mixed with wild-type (**A**; (8)), Y162W (**B**; from Figure 4.3 in main text), W270A/W284A (**C**), and W270A/W284A/Y162W (**D**). Traces are the average of three binding reactions. (**E**) The rate constant for the first phase of the binding reaction ($k_{1,obs}$) is dependent on the concentration of AAG and a linear fit yields the bimolecular rate constant for binding (k_{on}). Binding of W270A/W284A/Y162W occurs during the dead time of the instrument. (D) The rate constants for the second phase of the binding reaction ($k_{2,obs}$) are independent of the concentration of AAG and reflect the sum of the forward and reverse rate constants for nucleotide flipping. Rate constants in C and D are from two independent experiments (average ± SD), and the rate constants are summarized in Table C-1.

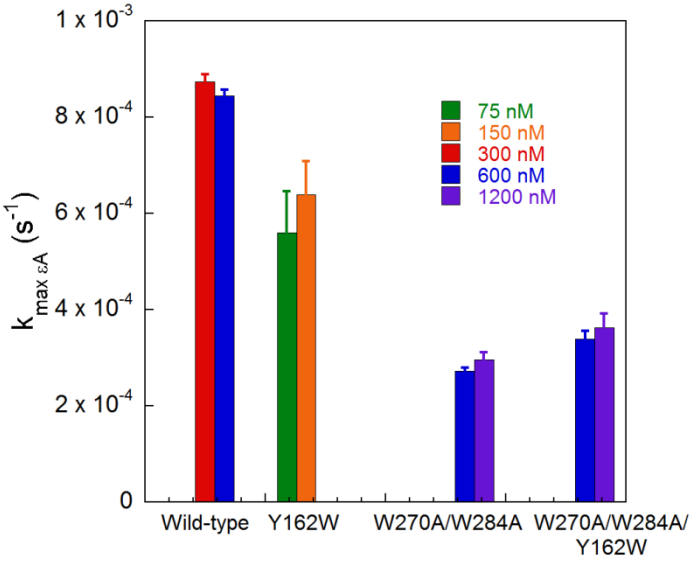


Figure C-4. Single-turnover excision of εA for wild-type and mutant AAG. Essentially identical rate constants were obtained for excision of εA from TEC DNA (50 nM) and excess AAG. These data indicate that the lowest concentration of enzyme is saturating for all of the AAG variants tested.

CHAPTER 5

Structure/Function Analysis of Intercalation and Nucleotide Flipping in Human Alkyladenine DNA Glycosylase[1]

To repair or modify bases embedded in the DNA duplex; enzymes often employ a nucleotide flipping strategy in which the target nucleotide is rotated 180° into the active site (2). It is proposed that both DNA bending and DNA intercalation are important for nucleotide flipping (2, 3). However, the timing of DNA bending is unknown, and the fundamental effects of DNA intercalation remain poorly understood. Structural studies have identified intercalating amino acid side chains in a wide variety of enzymes, but interestingly the identity of the amino acid varies (5).

DNA glycosylases are one type of enzyme that use nucleotide flipping to engage substrates. These enzymes initiate the base excision repair (BER) pathway by locating sites of damage in the genome and catalyzing the hydrolysis of the *N*-glycosidic bond to release the damaged base (6). We took advantage of the fluorescence of 1,N^6-ethenoadenine (εA), a

[1]I planned the experiments and generated mutant plasmids. Karina Kangas purified all the mutant proteins (except for Y162A, Y162F, and Y162W). I performed the pulse-chase experiments for Y162H, Y162I, and Y162A. All other experiments were performed by Karina. We both analyzed the data and I wrote the chapter.

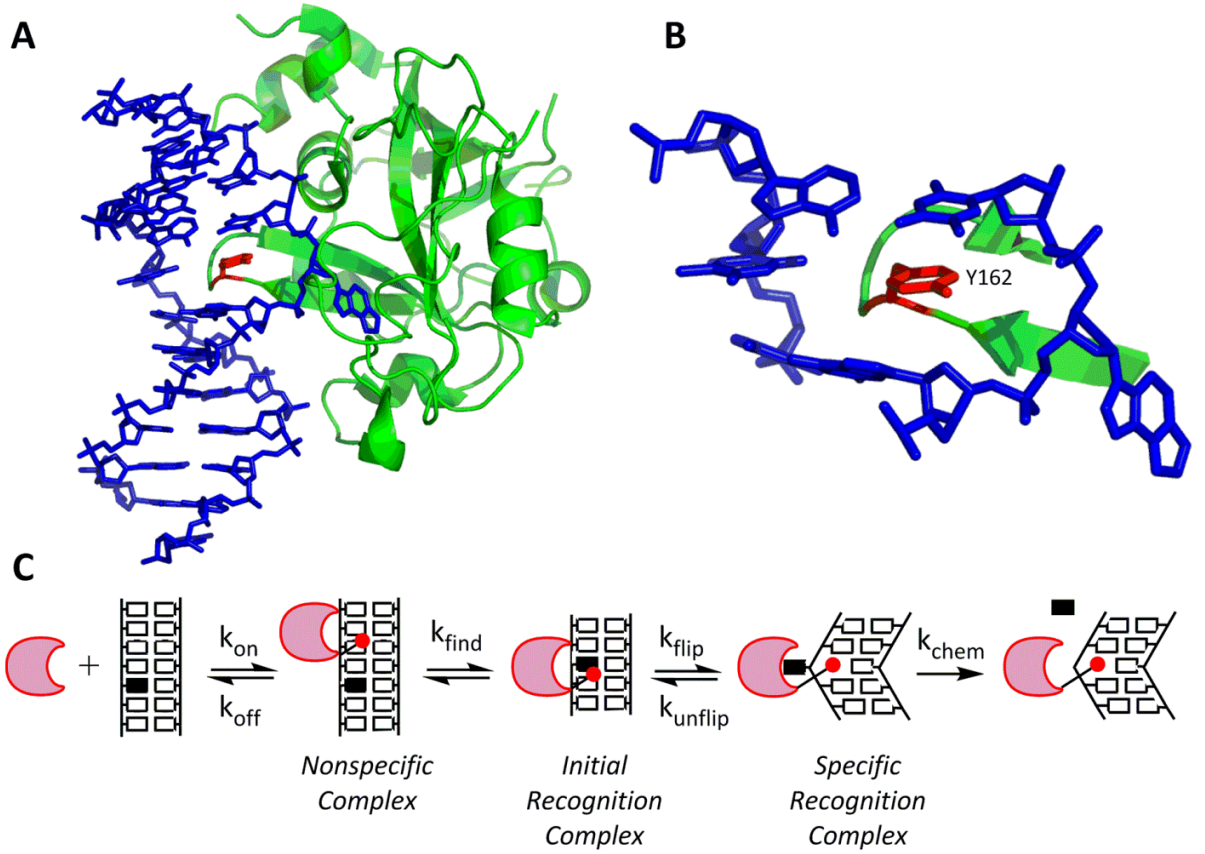


Figure 5.1. Nucleotide flipping and DNA intercalation by human AAG. (**A**) Cartoon representation of the AAG (green) complex with εA-DNA (blue) from the PDB (1EWN) was rendered with Pymol (http://www.pymol.org). The εA base is flipped into the active site, and the intercalating residue (Y162) is shown in red. (**B**) The intercalating tyrosine fills the space left by the flipped-out nucleotide, and stacking interactions are made between the DNA and Y162 after nucleotide flipping. (**C**) AAG (crescent) binds to nonspecific DNA and rapidly searches for DNA damage. This searching process is described by the macroscopic rate constant k_{find}. Once a lesion (solid rectangle) is encountered in an initial recognition complex, it can be flipped into the enzyme active site to form the specific recognition complex. AAG catalyzes hydrolysis of the *N*-glycosidic bond (k_{chem}) from this complex. As shown by the red ball and stick, the intercalating Y162 residue is important throughout all steps of the kinetic mechanism.

common form of DNA damage, to monitor binding and flipping by human alkyladenine DNA glycosylase (AAG) during DNA repair (7, 8). This monomeric repair protein is responsible for removing a diverse set of alkylated and deaminated purine lesions (7, 9-11). The minimal kinetic mechanism for multi-step recognition of εA damage has been reported (4, 12). Given the excess of undamaged DNA, initial binding of AAG is not at the site of damage (Figure 5.1C). AAG uses nonspecific binding (k_{on}) and facilitated diffusion (13) to search for sites of damage (k_{find}). Once an εA is found, AAG forms an initial recognition complex that is accompanied by changes in the stacking of the lesion base (4). From this intermediate, AAG catalyzes nucleotide flipping (k_{flip}) to form a more stable specific recognition complex that positions the substrate for *N*-glycosidic bond cleavage (k_{chem}; (14)). Multi-step recognition provides multiple opportunities for discrimination between damaged and undamaged nucleotides and similar mechanisms have been proposed for independently evolved DNA repair glycosylases (15-18).

In the crystal structure of AAG bound to a flipped-out εA lesion, the DNA is bent away from the protein by about 20°, and the intercalating tyrosine-162 is located at the center of the bend (Figure 5.1A; (14)). The side chain of tyrosine-162 projects from a β-hairpin on the surface of AAG and inserts into the minor groove DNA, thus filling the space left by the flipped-out nucleotide (Figure 5.1B). Recently the functional contributes of the intercalating residue to each of the steps in the AAG catalytic mechanism were determined (Chapter 3). The intercalating tyrosine residue (Y162) was mutated to phenylalanine (Y162F) and alanine (Y162A). Surprisingly, the Y162A mutation was well tolerated *in vitro* and does not appear to greatly perturb the structure, because the *N*-glycosidic bond cleavage is within 2-fold of wild-type (WT) AAG. Remarkably, the Y162A mutation increases the rate of nucleotide flipping by 50-fold

relative to wild-type AAG. This mutant also exhibits faster unflipping, resulting in a 140-fold reduction in the equilibrium constant for formation of the flipped-out complex. In contrast, the kinetic parameters for the Y162F and wild-type enzymes are very similar, suggesting that the hydroxyl of tyrosine is not necessary for function. These findings establish that DNA intercalation contributes to the specific recognition of DNA damage by acting as a plug to stabilize the specific recognition complex. Tyrosine 162 plays a critical role in the discrimination between damaged and undamaged nucleotides, and DNA intercalation slows the macroscopic rate of diffusive searching.

Most DNA glycosylases use a bulky side chain (Asn, Arg, Gln, Leu, or Tyr) to fill the space left by the flipped-out damaged nucleotide, although an exception includes 3-methyladenine DNA glycosylase, which has a Gly at this position (19). Unlike AAG, typically when the intercalating residue of a DNA glycosylase is mutated to a small amino acid, little to no activity is observed (14, 20-25). For example, the intercalating L191 residue in uracil DNA glycosylase (UDG) was mutated to either alanine or glycine (23). Both mutants were unable to excise uracil from single-stranded or duplex DNA. Using a pyrene nucleotide opposite a uracil base, the weak binding affinity and low catalytic activity of L191A and L191G was rescued (26, 27). The pyrene in DNA acts as a wedge to “push” the uracil from the base stack in free DNA, and a plug to prevent reinsertion after nucleotide flipping. This chemical rescue lead to the conclusion that the intercalating residue of UDG plays an early and late role in nucleotide flipping.

We investigated the structure/function relationship of the intercalating residue to each of the steps in the AAG catalytic mechanism by systematically mutating the intercalating tyrosine residue to amino acids of various volumes and properties. Tyrosine 162 was mutated

to tryptophan (W), phenylalanine (F), leucine (L), isoleucine (I), histidine (H), asparagine (N), and alanine (A). The forward and reverse rate constants for nucleotide flipping were measured to investigate whether there is a correlation between the size (volume) of the intercalating residue and k_{flip} or k_{unflip}. Alternatively, it may be critical that the residue is aromatic. Surprisingly, characterization of the seven Y162 AAG mutants resulted in three categories: aromatic (Y162, Y162F, and Y162W), non-aromatic (Y162A, Y162N, Y162L, and Y162L), and Y162H. The rate constant for nucleotide flipping is much faster for the non-aromatic mutants and Y162H compared to wild-type. These mutants are unable to stabilize the initial recognition complex and slow nucleotide flipping like wild-type and the aromatic mutants. The rate constant for unflipping is also much faster for the non-aromatic mutants. The aromatic mutants and Y162H AAG act like a plug, and are able to stabilize the specific recognition complex. Overall, only the aromatic Y162 mutants (Y162, Y162H, Y162F, and Y162W) have a favorable equilibrium constant for nucleotide flipping. No volume dependence of the intercalating residue was observed, as even the bulkier non-aromatic mutants (Y162N, Y162I, and Y162L) were unable to prevent the damaged base from flipping back into the DNA. Overall, the aromatic residues are less mobile, and are able to form favorable stacking interactions that stabilize the initial recognition complex. These larger residues are also able to plug the space left by the flipped-out nucleotide, thus stabilizing the specific recognition complex. The results establish that AAG requires an aromatic residue at position at 162 for optimal activity and specificity.

Materials and Methods

Purification of Wild-type and Mutant AAG Proteins. The catalytic domain of human AAG that lacks the first 79 amino acids was expressed in *E. coli* and purified as previously described (28). The Y162 mutants were constructed by site-directed mutagenesis and purified using the same protocol. The concentration of protein was determined by fluorescent titration of εA-DNA, as described below, and the concentrations were found to be within 5-fold of the concentration determined from the UV absorbance and calculated extinction coefficient.

Synthesis and Purification of Oligodeoxynucleotides. The 25mer oligonucleotides were synthesized by Integrated DNA Technologies or by the W.M. Keck Facility at Yale University. The oligonucleotides were desalted using Sephadex G-25 and purified using denaturing polyacrylamide gel electrophoresis as previously described. Oligonucleotides for gel-based assays were labeled on the lesion containing strand with a 5'-fluorescein (6-fam) label. The concentrations of the single stranded oligonucleotides were determined from the absorbance at 260 nm, using the calculated extinction coefficients for all oligonucleotides except those containing εA. For the oligonucleotide containing εA, the extinction coefficient was calculated for the same sequence with an A in place of the εA and corrected by subtracting 9400 $M^{-1}cm^{-1}$ to account for the weaker absorbance of εA as compared to A. The lesion-containing oligonucleotide was annealed with a 1.2-fold excess of the complement by heating to 90 °C and cooling slowly to 4 °C. One DNA sequence context was used and is referred to by the central three nucleotides (i.e., AEA has a central εA lesion flanked by a 5'A and a 3'A; Scheme 5.1). Two

AXA 5’ –CGATAGCATCAA**X**AATTCTCTCCAT
3’ –GCTATCGTAGTTTTTAAGAGAGGTA

X= εA

Scheme 5.1

different DNA sequences were previously used to determine the minimal kinetic mechanism of AAG because the fluorescence was different for the two sequences (TXC and AXA; (4)). All previously published results primarily utilized TXC DNA (4, 12); however in this work only AXA DNA was used.

Steady-state Fluorescence Measurements. Fluorescence emission spectra were collected with a PTI QuantaMaster fluorometer controlled by FeliX software. For εA fluorescence, an excitation wavelength of 314 nm (6 nm band-pass) was used and the total fluorescence was measured at emission wavelengths from 340-480 nm (6 nm band-pass). Samples (300 µL) of 400 nM εA-containing DNA were prepared in the standard buffer [50 mM NaMES (pH 6.5), 100 mM NaCl, 1 mM EDTA, 1 mM DTT] and spectra were recorded at 25 °C. To determine the steady-state fluorescence of εA-containing DNA bound to AAG, the spectra were recorded within one minute to ensure that no significant excision of εA occurred. Three independent titrations were averaged, and the data were fit by a quadratic equation assuming tight binding by AAG (Equation 5.1), in with F_{rel} is the relative fluorescence, A is the fractional quenching, and K_d is the dissociation constant for the εA-DNA.

$$F_{rel} = 1 - \frac{A(K_d + E + DNA) - \sqrt{(K_d + E + DNA)^2 - (4E \times DNA)}}{2E} \quad (5.1)$$

Gel-based Glycosylase Assay. Single-turnover glycosylase activity was measured with a discontinuous assay that utilizes abasic site cleavage by NaOH followed by DNA separation on a

denaturing polyacrylamide gel (13, 28). Fluorescein-labeled (5'-6-fam) DNA substrates (50 nM) containing εA were prepared in the standard buffer. Reactions were initiated with the addition of 300 nM - 600 nM AAG and incubated at 25 °C. At various time points, a sample from the reaction was removed and quenched in two volumes of 0.3 M NaOH, giving a final hydroxide concentration of 0.2 M. The abasic sites were cleaved by heating at 70 °C for 15 minutes. Samples were mixed with an equal volume of fomamide/EDTA loading buffer before loading onto a 15% polyacrylamide gel. Gels were scanned with a Typhoon Imager (GE Trio+ Healthcare) to detect the fluorescein label by exciting at 488 nm and measuring emission with a 520BP40 filter. The gel bands were quantified using ImageQuant TL (GE Healthcare). The data were converted to fraction product [Fraction Product = product / (product + substrate)] and then fit by a single exponential (Equation 5.2) using Kaleidagraph (Synergy Software). Single-turnover rates were independent of the concentration of AAG, indicating that the maximal rate constant was measured ($k_{obs} = k_{max}$; data not shown).

$$\text{Fraction product} = A(1 - e^{-k_{obs}t}) \qquad (5.2)$$

Stopped-flow Kinetics. Pre-steady state kinetic experiments were performed on a Hi-Tech SF-61DSX2, controlled by Kinetic Studio (TgK Scientific). The fluorescence of εA was measured using an excitation wavelength of 313 nm and a WG360 long-pass emission filter (4). At least three traces were averaged together at each concentration. The traces for wild-type, Y162F, and Y162W were fit by a double exponential (Equation 5.3) where *F* is the fluorescence as a function of time, *C* is the fluorescence of free DNA, *X* and *Y* are the changes in fluorescence of the intermediates, and *t* is the time. The traces for Y162A, Y162N, Y162H, Y162L, and Y162I were fit by a single exponential (Equation 5.2) and k_{obs} is equal to $k_{flip} + k_{unflip}$.

$$F = C - X(1 - e^{-k_{1,obs}t}) - Y(1 - e^{-k_{2,obs}t}) \quad (5.3)$$

For wild-type, Y162F, and Y162W, the observed rate constants were plotted versus concentration and fit by a straight line. $k_{1,obs}$ showed a linear concentration dependence, and the slope is equal to k_{on} ($M^{-1}s^{-1}$), and the y-intercept is equal to k_{off} (s^{-1}). The value of $k_{2,obs}$ was independent of concentration and is equal to $k_{flip} + k_{unflip}$.

Pulse-Chase Assay to Measure Substrate Dissociation. The macroscopic rate constant for dissociation of wild-type and mutant AAG from εA-containing DNA was measured by pulse-chase in the standard reaction buffer at 25 °C as previously described for wild-type AAG (4). Briefly, in 20 µL reactions, 50 nM fluorescein-labeled TEC DNA was mixed with 75 nM or 150 nM AAG for 20 seconds, and then a chase of 10 µM unlabeled TEC DNA was added. At various time points, a sample from the reaction was removed and analyzed as described for "Gel-based glycosylase assay". Base excision catalyzed by AAG results in fluorescein-labeled product, whereas dissociation releases unreacted fluorescein-labeled substrate. The partitioning between hydrolysis and dissociation can be determined from either the exponential rate constant or by the change in burst amplitude. If AAG dissociates from the labeled DNA before the chemical cleavage step and then binds to the unlabeled DNA, less of the reaction will occur during the single-turnover part of the curve as compared to the same experiment without chase. The data was converted to fraction product and fit by a single exponential (Equation 5.2). The observed rate constant was independent of the concentration of AAG.

According to the two step binding mechanism, two different partitioning equations can be written (29). All labeled substrate is initially bound, and therefore the fraction of product formed is given by the fraction that goes on to react. This is indicated by Equation 5.4, in which

A is the burst amplitude (the fraction of product formed in the burst phase of the experiment), k_{max} is the maximal single-turnover rate constant for formation of product, and $k_{off,obs}$ is the macroscopic rate constant for dissociation from the flipped-out complex. This expression can be rearranged to solve for the desired dissociation rate constant (Equation 5.5). Similarly, for branched pathways, the observed rate constant for the burst phase of the pulse-chase experiment is given by the sum of the rate constants for the competing pathways, formation of product is given by k_{max} and the macroscopic dissociation of substrate is designated $k_{off,obs}$ (Equation 5.6). Solving for $k_{off,obs}$ gives Equation 5.7.

$$A = \frac{k_{max}}{k_{off,obs} + k_{max}} \tag{5.4}$$

$$k_{off,obs} = \frac{k_{max}}{A} - k_{max} \tag{5.5}$$

$$k_{obs} = k_{off,obs} + k_{max} \tag{5.6}$$

$$k_{off,obs} = k_{obs} - k_{max} \tag{5.7}$$

Control reactions in which no chase was added provided the single-turnover rate constant, k_{max}, and confirmed that these concentrations of AAG were saturating. From these values, the dissociation rate constant, k_{off}, for AAG dissociating from εA-DNA was calculated by two different methods (Equations 5.5 and 5.7). Both methods gave very similar values for $k_{off,obs}$ and we report the results obtained from Equation 5.5.

Since the εA-DNA binds in two steps, the observed rate constant for dissociation of substrate ($k_{off, obs}$) could be limited by the unflipping rate (k_{unflip}) or dissociation from nonspecific DNA (k_{off}). According to Scheme 2.3, and assuming that the flipped-out complex is stable (i.e.,

$k_{flip} >> k_{unflip}$), this observed dissociation rate constant can be expressed in terms of the microscopic rate constants (Equation 5.8; (29)). Stopped flow fluorescence suggests that dissociation from the initial AAG•DNA complex is rapid, and therefore the observed rate constant for substrate dissociation from the εA-DNA•AAG complex is approximately equal to the reverse rate constant for flipping (Equation 5.9).

$$k_{off,obs} = k_{unflip}\left(\frac{k_{off}}{k_{off} + k_{flip}}\right) \tag{5.8}$$

$$k_{off,obs} \approx k_{unflip} \tag{5.9}$$

Direct Measurement of Substrate Dissociation. Stopped-flow double mixing experiments were performed to measure unflipping and dissociation of εA-DNA by Y162A, Y162N, Y162H, Y162L, Y162I, Y162F, and Y162W. To rapidly form the flipped-out AAG complex, 2.8 μM DNA was mixed with 2 μM AAG. After an aging time of one second, 18 μM or 36 μM pyrrolidine-DNA was added as a competitor. The final concentrations after mixing were 700 nM DNA, 500 nM AAG, and 9 μM or 18 μM pyrrolidine-DNA. The reaction was followed for 1-300 seconds; no significant excision of εA occurs during this time. Three traces were averaged together and the change in fluorescence was fit by a single exponential (Equation 5.2). As described for the gel-based pulse-chase assay, the observed rate constant for dissociation is equal to the rate constant for unflipping.

Results

Binding and Excision of εA by AAG Mutants

We used site-directed mutagenesis to generate seven Y162 AAG variants to test how the volume of the intercalating residue alters nucleotide flipping and excision of εA by AAG. Each mutant was first characterized for its ability to bind and excise εA from a synthetic 25mer oligonucleotide containing a central εA•T mismatch. As AAG has a relatively slow rate of base excision, steady-state fluorescence titrations could be used to measure binding of active enzyme to εA-DNA (4). The AEA DNA sequence used has high εA-fluorescence, and the stable

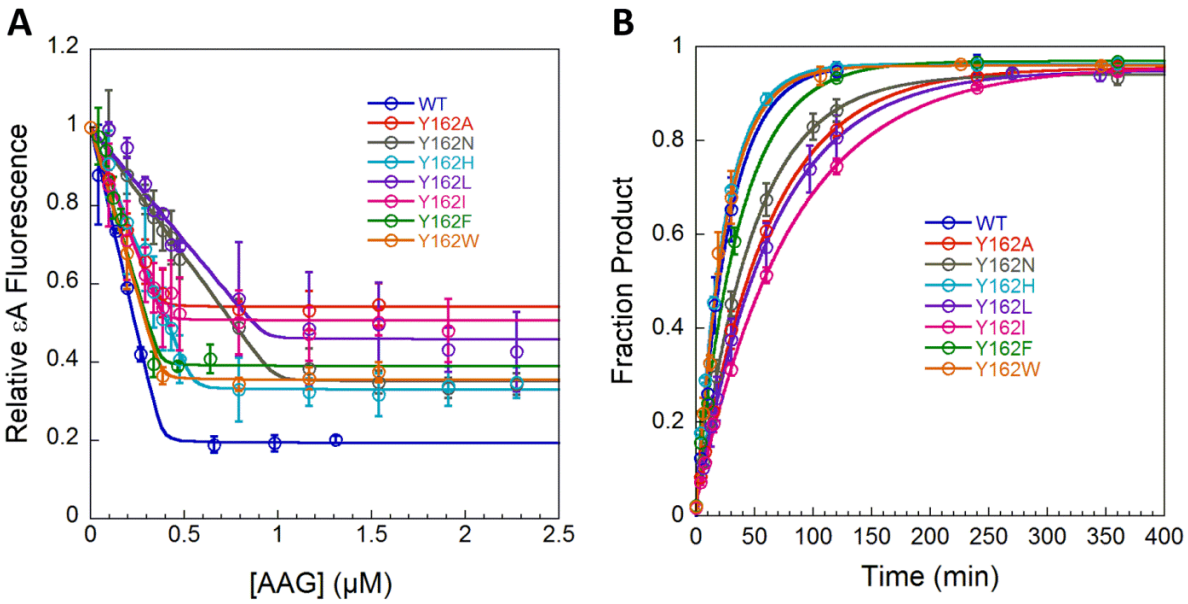


Figure 5.2. Binding and excision of εA by AAG. (**A**) Titration of 0.4 μM AEA DNA with wild-type and Y162 mutants. Steady-state fluorescence was monitored within one minute, before excision can occur. Three independent titrations were performed and the average and standard deviation are shown. The normalized fluorescence was fit by the quadratic equation under conditions of tight binding (Equation 5.1). These fits indicate that the number of εA-binding sites is 86% the expected value for wild-type, 100% for Y162A, 41% for Y162N, 76% for Y162H, 41% for Y162L, 100% for Y162I, 67% for Y162F, and 100% for Y162W. (**B**) Single-turnover excision of εA by wild-type and mutant AAG with 50 nM DNA and saturating enzyme was fit by a single exponential (Equation 5.2) to obtain the rate constants for *N*-glycosidic bond cleavage (Table 5.1). Data points are an average ± SD from three independent experiments.

Table 5.1. Kinetic parameters for flipping and excision of εA by wild-type and Y162 mutants[a]

	Wild-type[b]	Y162F	Y162W	Y162H
k_{flip} (s^{-1})	7.0 ± 0.2	5.2 ± 0.3	7.0 ± 1.1	46.1 ± 1.8
k_{unflip} (s^{-1})	(4.6 ± 0.2) x 10^{-3}	4.6 x 10^{-2} [c]	3.9 x 10^{-2} [c]	0.20[c]
K_{flip}[d]	1520	113	180	231
$k_{chem\ \varepsilon A}$ (s^{-1})	(6.1 ± 0.8) x 10^{-4}	(4.5 ± 0.5) x 10^{-4}	(6.8 ± 1) x 10^{-4}	(7.0 ± 0.1) x 10^{-4}

	Y162A	Y162N	Y162L	Y162I
k_{flip} (s^{-1})	41.1 ± 8.4	26.0 ± 6.9	56.0 ± 2.7	28.9 ± 4.4
k_{unflip} (s^{-1})	10 ± 0.8	2.5 ± 0.08	8.0 ± 0.5	9.6 ± (7.1 x 10^{-5})
K_{flip}[d]	4	10	7	3
$k_{chem\ \varepsilon A}$ (s^{-1})	(2.8 ± 0.2) x 10^{-4}	(3.5 ± 0.4) x 10^{-4}	(2.6 ± 0.4) x 10^{-4}	(2.1 ± 0.1) x 10^{-4}

[a]All rate constants were determined using the AEA oligonucleotide. The standard conditions were 25 ^{o}C, 50 mM NaMES (pH 6.5), 100 mM NaCl, 1 mM EDTA, and 1 mM DTT.

[b]The values for wild-type with AEA DNA (k_{flip} = 7.0 ± 0.2 s^{-1} and k_{chem} = (6.1 ± 0.8) x 10^{-4} s^{-1}) are similar to previously reported values (k_{flip} = 4.3 ± 0.4 s^{-1} and k_{chem} = (5.3 ± 0.2) x 10^{-4} s^{-1}) (4).

[c]Double-mixing stopped-flow experiments were only performed once for these mutants.

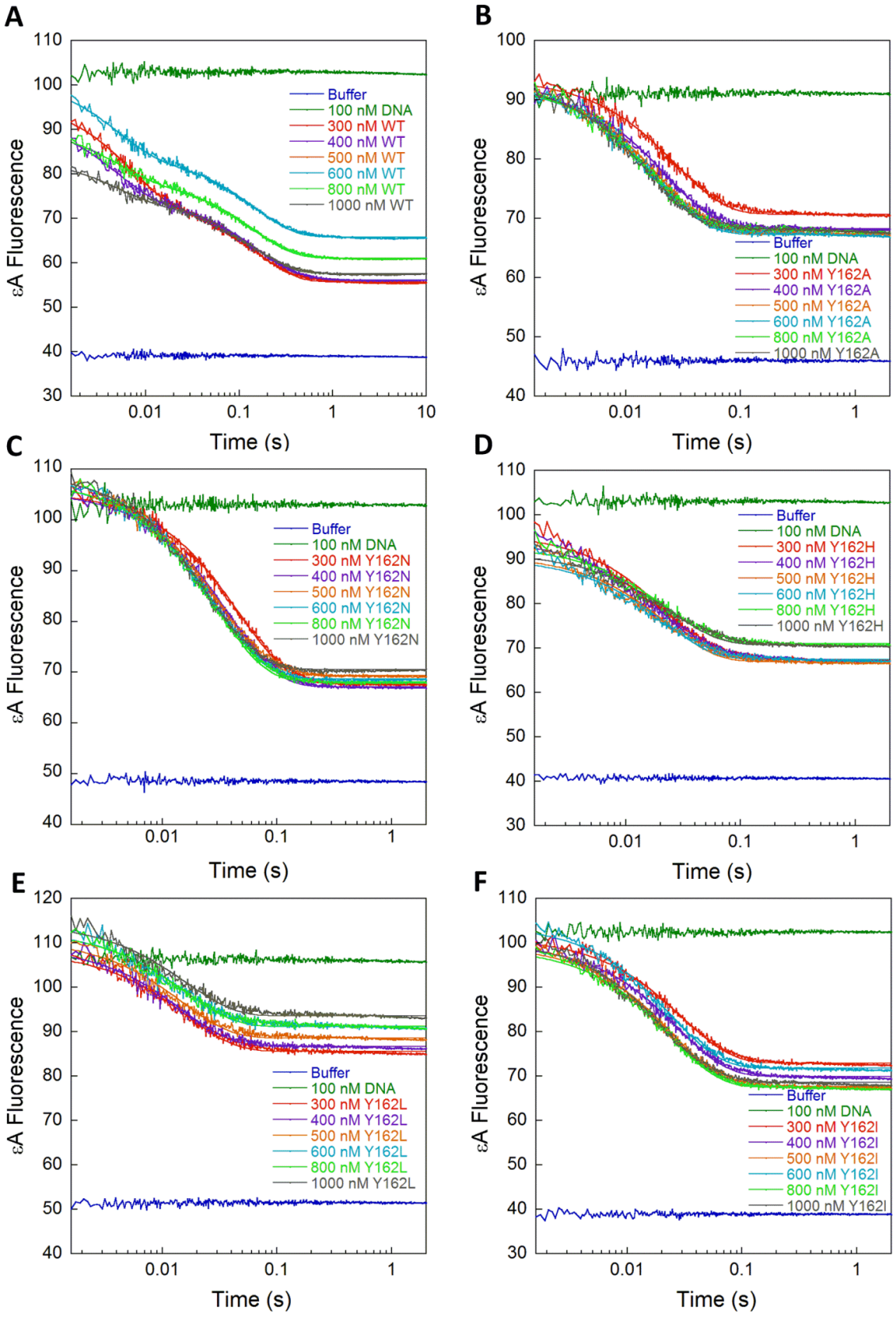
[d]The equilibrium constant for flipping is given by the ratio of the flipping and unflipping rate constants ($K_{flip} = k_{flip}/k_{unflip}$).

binding of the extrahelical εA lesion results in a 5-fold quenching of fluorescence for wild-type AAG (Figure 5.2A). Titrations of the Y162 mutants showed somewhat less quenching of εA-fluorescence (2-3 fold), but nonetheless indicate tight binding of εA-DNA. From these titrations, the active concentration of protein was more accurately determined, and this corrected concentration was used for the remaining experiments.

To evaluate whether the mutant proteins retain catalytic activity, single-turnover experiments were performed with enzyme in excess over the εA-DNA substrate (Figure 5.2B). Previous work established the Y162A mutant retained catalytic activity, which was surprising because it had been suggested that this mutant may lack activity (14). Knowing that the mutant with the smallest amino acid at position Y162 was active; we predicted that all other Y162 mutants, regardless of size or polarity would retain catalytic activity. This was found to be the case, and the rate constant for excision by all of the mutants is within 2.5-fold that of wild-type (Table 5.1). Therefore the transition state for *N*-glycosidic bond hydrolysis is not disrupted by these mutations. This allowed us to measure flipping and unflipping for each mutant.

Stopped-flow Fluorescence to Monitor Binding and Flipping of εA-DNA

Rapid mixing experiments were used to determine the rate constants for association and nucleotide flipping (Figure 5.3). Two phases were observed when excess wild-type AAG was mixed with 100 nM AEA DNA. For this substrate, the εA fluorescence is initially high in the free DNA, is partially quenched in the first intermediate, and is more strongly quenched in the fully flipped-out complex. Similar traces were also observed for Y162F and Y162W, and were fit well by a double exponential (Equation 5.3). The first phase is linearly dependent on the enzyme concentration, and is assigned to binding and formation of the initial recognition complex. The


A
εA Fluorescence
Time (s)
Buffer
100 nM DNA
300 nM WT
400 nM WT
500 nM WT
600 nM WT
800 nM WT
1000 nM WT
B
εA Fluorescence
Time (s)
Buffer
100 nM DNA
300 nM Y162A
400 nM Y162A
500 nM Y162A
600 nM Y162A
800 nM Y162A
1000 nM Y162A
C
εA Fluorescence
Time (s)
Buffer
100 nM DNA
300 nM Y162N
400 nM Y162N
500 nM Y162N
600 nM Y162N
800 nM Y162N
1000 nM Y162N
D
εA Fluorescence
Time (s)
Buffer
100 nM DNA
300 nM Y162H
400 nM Y162H
500 nM Y162H
600 nM Y162H
800 nM Y162H
1000 nM Y162H
E
εA Fluorescence
Time (s)
Buffer
100 nM DNA
300 nM Y162L
400 nM Y162L
500 nM Y162L
600 nM Y162L
800 nM Y162L
1000 nM Y162L
F
εA Fluorescence
Time (s)
Buffer
100 nM DNA
300 nM Y162I
400 nM Y162I
500 nM Y162I
600 nM Y162I
800 nM Y162I
1000 nM Y162I

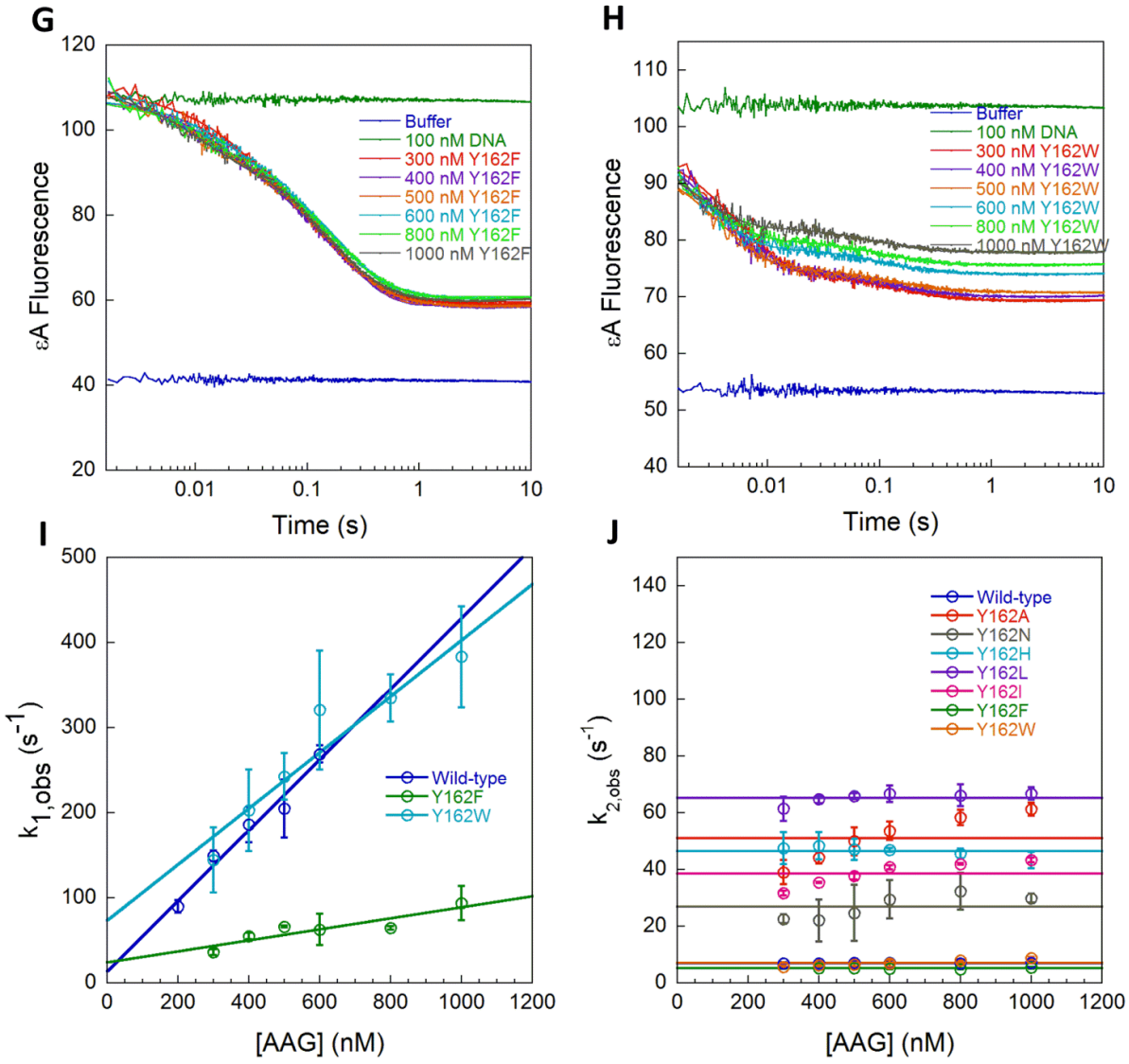


Figure 5.3. Stopped-flow fluorescence with excess protein to measure binding and nucleotide flipping by Y162 mutants. Representative data from experiments in which 100 nM AEA DNA was mixed with increasing concentrations of wild-type (**A**), Y162A (**B**), Y162N (**C**), Y162H (**D**), Y162L (**E**), Y162I (**F**), Y162F (**G**), and Y162W (**H**). Traces are the average of three binding reactions and are fit by either a single or double exponential (Equations 5.2 and 5.3). (**I**) The rate constant for the first phase of the binding reaction ($k_{1,obs}$) is dependent on the concentration of AAG and a linear fit yields the bimolecular rate constant for binding (k_{on}) (**J**) The rate constants for the second phase of the binding reaction ($k_{2,obs}$) are independent of the concentration of AAG and reflect the sum of the forward and reverse rate constants for nucleotide flipping. Rate constants in I and J are from three independent experiments (average ± SD).

second phase is independent of concentration and corresponds to nucleotide flipping to form the specific recognition complex. The observed first order rate constant for formation of the flipped-out specific recognition complex is nearly identical for Y162F, Y162W, and wild-type AAG.

Association experiments with the other Y162 variants yielded different fluorescence traces, and were similar to the previously reported Y162A results (Chapter 3). For Y162A, Y162N, Y162H, Y162L, and Y162I, only a single time dependent decrease in fluorescence was observed (Figure 5.3). Even at low concentrations of protein, the first phase of the reaction was complete in the dead time of the stopped-flow ($k_{on} \geq 10^9$ $M^{-1}s^{-1}$). In each case, the exponential decrease in fluorescence, which indicates formation of the flipped-out specific recognition complex, was independent of concentration and much faster than wild-type AAG. Overall, variants of AAG with a non-aromatic intercalating residue bind nonspecific DNA and search for damage significantly faster than the wild-type enzyme.

Measurement of Unflipping and Dissociation of εA-DNA

To define individual microscopic rate and equilibrium constants it is necessary to measure the dissociation kinetics. As previously described, the dissociation of AAG from εA-containing DNA was investigated using the pulse-chase method, whereby the AAG•εA-DNA complex is allowed to form and is subsequently chased with an excess of tight binding pyrrolidine-DNA inhibitor (12). Any AAG that dissociates is bound to the inhibitor, allowing measurement of the partitioning forward (abasic DNA product) and backward (εA-DNA substrate). Previously reported values for wild-type, Y162F, and Y162W were from experiments with the TEC DNA substrate. Using the AEA substrate, wild-type AAG is the only protein committed to excision of

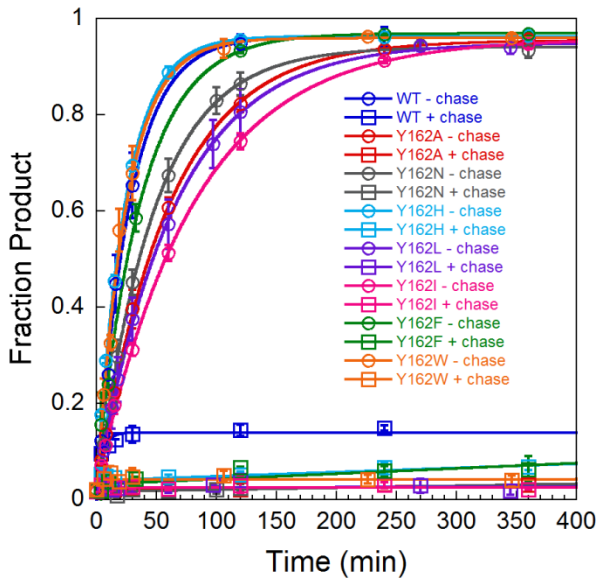


Figure 5.4. Pulse-chase experiments to measure dissociation of εA-DNA. Fluorescein (fam)-labeled DNA (50 nM) was mixed with excess AAG (75 nM - 1200 nM) for 20 seconds, and then 10 μM unlabeled pyrrolidine TXC DNA was added as a chase. The reactions were quenched at the indicated time points, and the fraction of abasic DNA product was determined by alkaline hydrolysis and gel electrophoresis (see Material and Methods for details). The partitioning between the forward reaction (base excision) and substrate dissociation was measured. The commitment for wild-type (15%) was used to calculate the rate of unflipping (see Materials and Methods). 100% of all Y162 mutants dissociated and the +chase reaction matches the control in which chase and substrate are added at the same time (data not shown). Therefore, stopped-flow double-mixing experiments were necessary to measure k_{unflip} for the Y162 mutants.

εA, with 15% of the complex partitioning forward to product and 85% dissociating (Figure 5.4). Assuming the dissociation of AAG from nonspecific DNA is fast, these data yield the microscopic rate constant for unflipping (Equation 5.5). Since 100% of the other Y162 variants dissociated from the DNA, no product could be detected in the presence of chase (Figure 5.4). This indicates that unflipping is much faster than *N*-glycosidic bond cleavage. In order to precisely measure the rate of unflipping for these mutants, we next performed a double-mixing experiment in the stopped-flow.

In the double-mixing experiment, the flipped-out complex of mutant AAG with εA-DNA was rapidly formed and then challenged with pyrrolidine-DNA competitor. The time-dependent

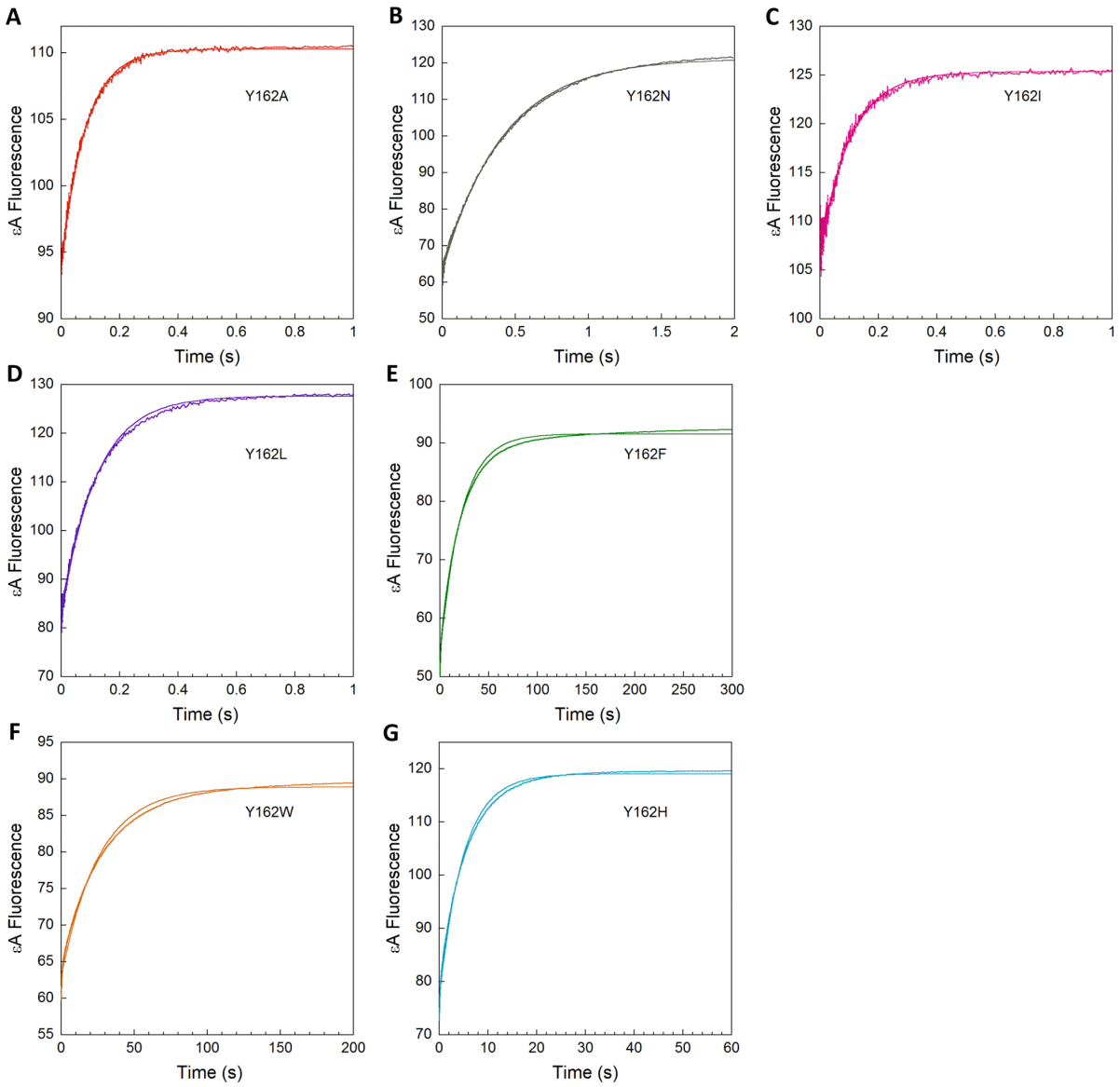


Figure 5.5. Direct measurement of εA-DNA dissociation from Y162 mutants. Double-mixing experiments were performed using stopped-flow fluorescence to monitor the increase in fluorescence upon release of εA-DNA for Y162A (**A**), Y162N (**B**), Y162I (**C**), Y162L (**D**), Y162F (**E**), Y162W (**F**), and Y162H (**G**). The AAG•DNA complex was formed, aged for one second, and then chased with an excess of pyrrolidine-DNA as a competitor. The traces reflect the average of three separate mixing experiments and the solid lines depict the best fit to a single exponential. Rate constants in Table 5.1 are from two independent experiments (average ± SD) unless otherwise stated.

increase in εA fluorescence is fit by a single exponential, which corresponds to the rate-limiting unflipping step. This experiment reveals the rate of εA unflipping is anywhere between 9 to 2,000-fold faster than wild-type, depending on the mutation. Overall, the aromatic residues at position 162 (phenylalanine, tyrosine, histidine, or tryptophan) had slower rates of unflipping compared to the non-aromatic residues (alanine, asparagine, leucine, or isoleucine). This observation supports the model that tyrosine intercalation acts as a plug to block the unflipping of εA.

After determining the rate constants for unflipping, the rate and equilibrium constants for nucleotide flipping could be calculated for wild-type and the Y162 variants (Table 5.1). Even though k_{unflip} is faster for the Y162 mutants compared to wild-type AAG, the unflipping rate constants are significantly slower than the observed rate constants for flipping. This indicates that the observed rate constant that was assigned to nucleotide flipping ($k_{2,obs}$) is approximately equal to the microscopic rate constant for flipping (k_{flip}). The equilibrium constant for nucleotide flipping is the ratio of the two values ($K_{flip} = k_{flip}/k_{unflip}$). For aromatic residues (tyrosine, phenylalanine, histidine, and tryptophan), the equilibrium constant for nucleotide flipping is highly favorable. In contrast, the equilibrium constant for flipping is greatly decreased for proteins with a non-aromatic amino acid at the position of 162.

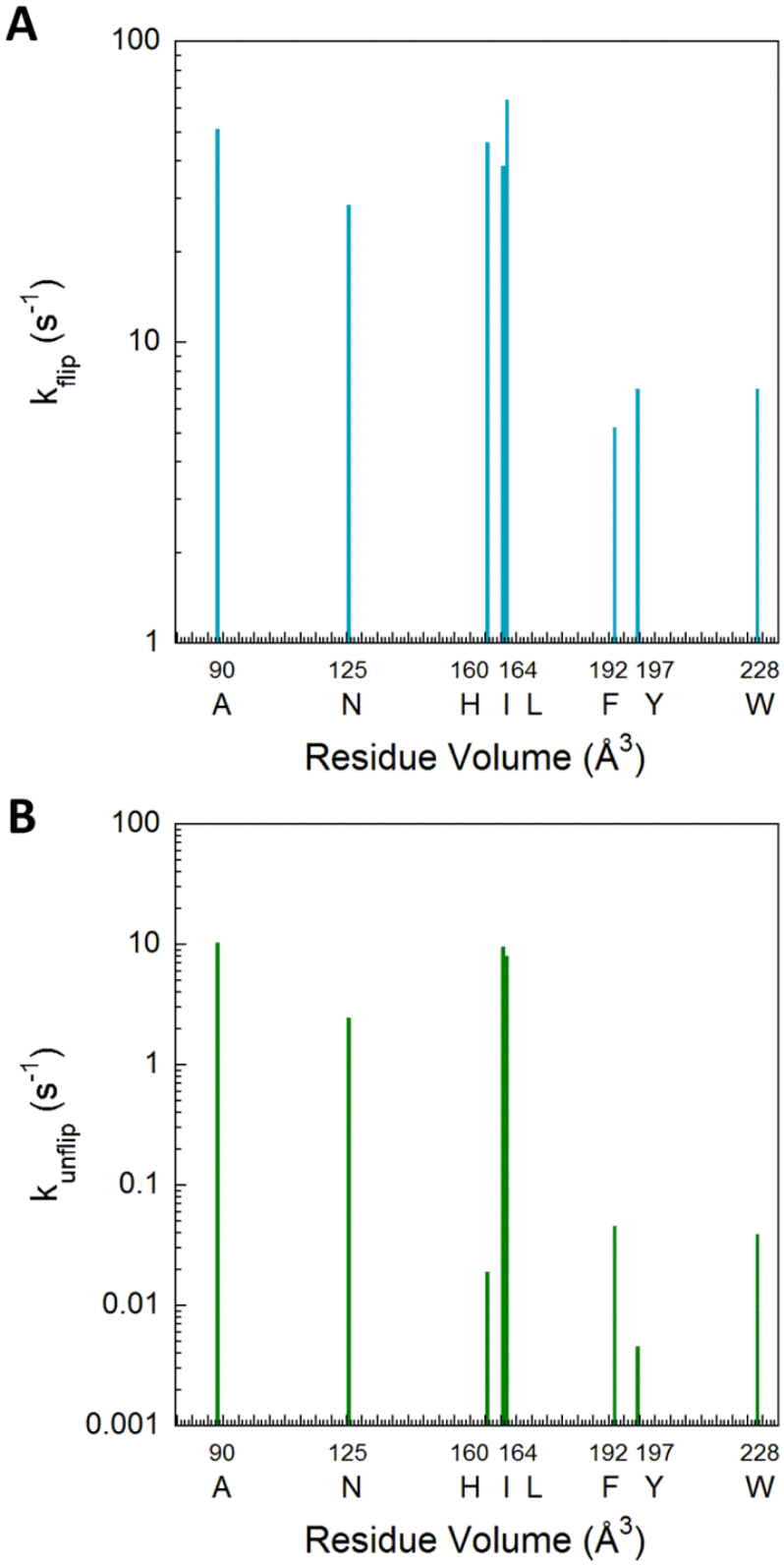
Discussion

Structures of DNA repair glycosylases in complex with DNA suggest that nucleotide flipping is aided by intercalating residues (2, 30-32). The prevalence of nucleotide flipping and the dramatic nature of the conformational change that accompanies DNA bending and nucleotide flipping have captured the interest of scientists for almost twenty years. However, little is

known about the fundamental mechanisms by which nucleotide flipping is accomplished because it is difficult to observe rapid conformational changes that occur on the millisecond time scale. To understand the contribution of DNA intercalation to the energetic landscape for specific recognition and nucleotide flipping, we performed comprehensive kinetic analysis of mutant forms of AAG in which the conserved intercalating tyrosine (Y162) was mutated. These results provide support for the model that intercalation by an aromatic side chain greatly stabilizes the extrahelical conformation of the specific recognition complex. Unexpectedly, we further confirm the role for Y162 in slowing the rate of nucleotide flipping, apparently by stabilizing the initial recognition complex. This structure/function study supports the model that DNA intercalation contributes to the specificity of AAG, but the increased specificity comes at the cost of reduced speed (Chapter 3).

AAG Requires an Aromatic Intercalating Residue

We focused on the role of the intercalating tyrosine (Y162), which is universally conserved among eukaryotic AAG homologs (32, 33). Seven mutations were made at position 162 to determine whether nucleotide flipping and excision by AAG depends on the identity of the intercalating residue. We predicted a volume dependence of the intercalating residue where larger amino acids would serve as a plug and slow k_{unflip}. Instead we found AAG requires an aromatic intercalating residue. No systematic change in the rate constants for flipping or unflipping were observed regardless of the volume or polarity. All non-aromatic residues have a decreased equilibrium constant for nucleotide flipping which may be explained by increased mobility of these residues (Table 5.1 and Figure 5.6). Aromatic residues form favorable pi-pi stacking interactions with nucleobases of DNA which restrict their mobility. This suggests that


A
100
10
1
k_{flip} (s^{-1})
90
125
160
164
192
197
228
A
N
H
I
L
F
Y
W
Residue Volume (Å^3)
B
100
10
1
0.1
0.01
0.001
k_{unflip} (s^{-1})
90
125
160
164
192
197
228
A
N
H
I
L
F
Y
W
Residue Volume (Å^3)

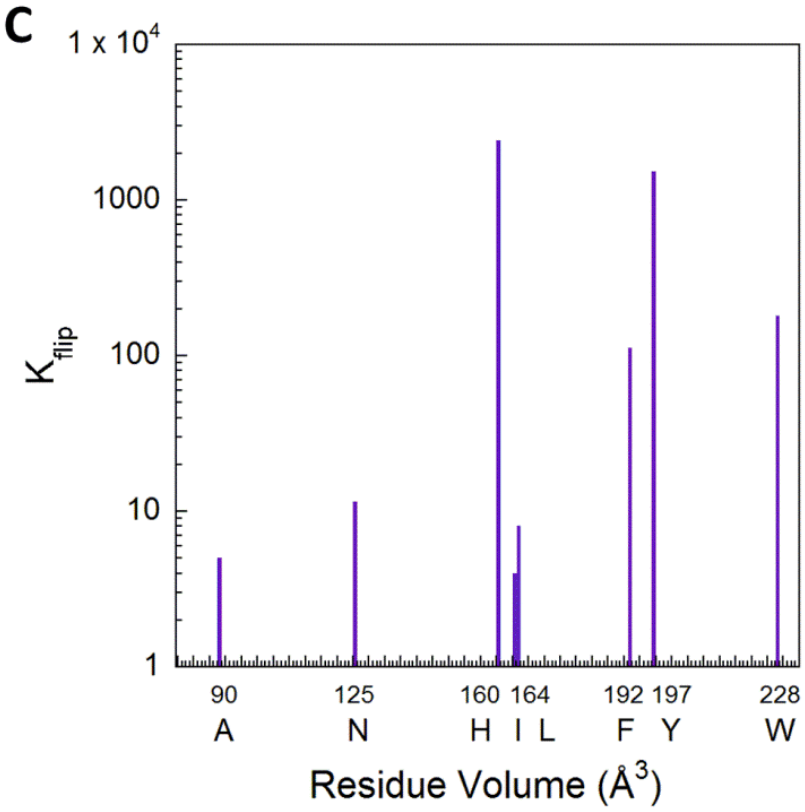


Figure 5.6. Volume dependence of the intercalating residue. (**A**) k_{flip} (**B**) k_{unflip} or (**C**) K_{flip} verses the residue volume (Å^3). The BL- residue volume calculation was used (1). Mutants with an aromatic residue at position 162 have a favorable equilibrium constant for nucleotide flipping because both k_{flip} and k_{unflip} are slow. The opposite trend is true for mutants with a non-aromatic residue.

aromatic stacking interactions are responsible for the stabilization of the initial recognition complex and the specific recognition complex.

All seven Y162 mutations made by site-directed mutagenesis are well tolerated *in vitro* and do not appear to greatly perturb the structure, because the *N*-glycosidic bond cleavage step is within 3-fold of wild-type AAG. Some prokaryotic homologs of AAG have been identified that appear to substitute histidine for tyrosine at the intercalating position (32, 34). Histidine could make similar stacking interactions as tyrosine, phenylalanine, and tryptophan. Kinetic analysis showed that this was the case. While the rate constant for nucleotide flipping is 7-fold faster for the Y162H mutant compared to wild-type, the aromatic Y162H mutant is able to slow the rate constant for unflipping and stabilize the specific recognition complex. Thus the overall equilibrium constant for nucleotide flipping remains very favorable for this mutant.

Binding to nonspecific DNA and searching are significantly faster for all non-aromatic Y162 mutants, and too fast to measure with stopped-flow assays. Parallels can be drawn with recent reports that mutation of intercalating residues to alanine in other DNA glycosylases increases the rate of diffusion (20, 21). Thus, a theme emerges that DNA intercalation slows the macroscopic rate of diffusive searching. Although the non-aromatic Y162 mutants engage εA sites significantly faster than the wild-type enzyme, the destabilization of the specific recognition complex greatly reduces the specificity of the enzyme. Previous work reported Y162A is not able to effectively find sites of damage when challenged by an excess of undamaged bases, leading to the prediction that DNA intercalation is essential for *in vivo* DNA repair (Chapter 3). Consistent with this conclusion, heterologous expression of Y162A AAG fails

to fully protect cells against an exogenous DNA alkylating agent (14). We predict the other non-aromatic Y162 mutants would be deficient for DNA repair *in vivo*.

DNA Intercalation by Other DNA Glycosylases

There are several distinct and evolutionary families of DNA repair glycosylases (35, 36). These enzymes differ in their chemical mechanisms and in the identity of the lesions that they recognize, however they all recognize damaged bases and catalyze cleavage of the *N*-glycosidic bond. To gain access to damaged bases, DNA glycosylases use nucleotide flipping and DNA intercalation. While these enzymes fill the space left by the flipped-out damaged nucleotide with one or more intercalating residues, the identity of the intercalating residues varies.

Fluorescence methods have been used to measure DNA binding and nucleotide flipping by many DNA glycosylases, and results from the literature are summarized in Table 5.2 (4, 18, 37, 38). Since these enzymes belong to different superfamilies and therefore evolved

Table 5.2. Comparison of nucleotide flipping by DNA repair glycosylases.

Enzyme	Superfamily	Substrate Range	Intercalating Residue	Substrate	k_{flip} (s^{-1})	k_{unflip} (s^{-1})	K_{flip}	Reference
UDG	UDG	Narrow	L191	2'-fluoroU·G	1260	38	33	(38)
OGG1	HhH	Narrow	N149	8-oxo-G·C	13	1	10	(18)
Fpg	Fpg	Broad	M73	8-oxo-G·C	12	12	1	(39)
AAG	AAG	Broad	Y162	εA·T	7	0.005	1400	This study and (4)

independently, it is interesting to compare the equilibrium constant for nucleotide flipping and the identity of the intercalating residue. Compared to other DNA glycosylases, the rate constants for both forward and reverse flipping are much slower for AAG. AAG has a very favorable equilibrium constant for flipping of εA. However, biochemical evidence suggests that

εA is unique among the well-characterized AAG substrates, because other lesions such as 7-methylguanosine and inosine appear to have unfavorable equilibrium constants for flipping (10, 39). Most DNA glycosylases do not use an aromatic intercalating residue (Table 5.2), and overall the equilibrium constant for nucleotide flipping is much less favorable than AAG. Interestingly, similar K_{flip} values were obtained for non-aromatic Y162 AAG mutants. It appears that for most glycosylases, the identity of the intercalating residue is less important than its steric bulk, which serves to plug the hole and prevent reinsertion of the damaged base. Based on our results, the aromatic intercalating residue is essential for AAG and increases specificity for εA by stabilizing both the initial recognition complex and the stable recognition complex.

Conclusions

Structures of AAG bound to εA-DNA (32) raised the possibility that intercalation by Y162 could increase the rate constant for flipping by "pushing" or destabilizing the duplex and decrease the rate constant for unflipping by "plugging" the hole vacated by the flipped-out nucleobase, as has also been suggested for other glycosylases (30, 31). Our study provides a rigorous test of these models. The kinetic parameters for small, non-aromatic Y162 mutants support the plugging model, with these mutants showing a 2000-fold increase in the rate of unflipping relative to wild-type AAG. Surprisingly, none of the smaller Y162 mutants show a reduced rate of flipping, but instead all non-aromatic Y162 mutants flip out εA lesions at least 4-fold faster than the wild-type enzyme. This argues against the pushing model for Y162 and demonstrates that intercalation of Y162 dramatically slows the nucleotide flipping step. The slower rate constant for flipping by wild-type AAG is presumably explained by tight binding of εA in an initial recognition complex. Overall, the non-aromatic mutations greatly destabilize the specific

recognition complex, and we obtain an estimate of 380-fold stabilization for the contribution of tyrosine intercalation to the equilibrium constant for flipping. AAG is a particularly attractive system to delve into the biophysical and biochemical mechanism of multi-step DNA recognition, but it is apparent that many different enzymes must employ similar strategies to discriminate between non-specific and specific sites, and the lessons learned from AAG may be more widely applicable to other DNA modifying enzymes.

References

1. Tsai J, Taylor R, Chothia C, & Gerstein M (1999) The packing density in proteins: standard radii and volumes. *Journal of molecular biology* 290(1):253-266.

2. Roberts RJ & Cheng X (1998) Base flipping. *Annu Rev Biochem* 67:181-198.

3. Stivers JT (2004) Site-specific DNA damage recognition by enzyme-induced base flipping. *Progress in nucleic acid research and molecular biology* 77:37-65.

4. Wolfe AE & O'Brien PJ (2009) Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by human alkyladenine DNA glycosylase. *Biochemistry* 48(48):11357-11369.

5. Stivers JT & Jiang YL (2003) A mechanistic perspective on the chemistry of DNA repair glycosylases. *Chem Rev* 103(7):2729-2759.

6. Ye N, Holmquist GP, & O'Connor TR (1998) Heterogeneous repair of N-methylpurines at the nucleotide level in normal human cells. *Journal of molecular biology* 284(2):269-285.

7. Gros L, Ishchenko AA, & Saparbaev M (2003) Enzymology of repair of etheno-adducts. *Mutation research* 531(1-2):219-229.

8. Pandya GA & Moriya M (1996) 1,N6-ethenodeoxyadenosine, a DNA adduct highly mutagenic in mammalian cells. *Biochemistry* 35(35):11487-11492.

9. Engelward BP*, et al.* (1997) Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. *Proceedings of the National Academy of Sciences of the United States of America* 94(24):13087-13092.

10. O'Brien PJ & Ellenberger T (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase. *The Journal of biological chemistry* 279(11):9750-9757.

11. Hitchcock TM*, et al.* (2004) Oxanine DNA glycosylase activity from Mammalian alkyladenine glycosylase. *The Journal of biological chemistry* 279(37):38177-38183.

12. Hendershot JM, Wolfe AE, & O'Brien PJ (2011) Substitution of active site tyrosines with tryptophan alters the free energy for nucleotide flipping by human alkyladenine DNA glycosylase. *Biochemistry* 50(11):1864-1874.

13. Hedglin M & O'Brien PJ (2008) Human alkyladenine DNA glycosylase employs a processive search for DNA damage. *Biochemistry* 47(44):11434-11445.

14. Lau AY, Wyatt MD, Glassner BJ, Samson LD, & Ellenberger T (2000) Molecular basis for discriminating between normal and damaged bases by the human alkyladenine

glycosylase, AAG. *Proceedings of the National Academy of Sciences of the United States of America* 97(25):13573-13578.

15. Banerjee A, Santos WL, & Verdine GL (2006) Structure of a DNA glycosylase searching for lesions. *Science* 311(5764):1153-1157.

16. Banerjee A, Yang W, Karplus M, & Verdine GL (2005) Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. *Nature* 434(7033):612-618.

17. Parker JB*, et al.* (2007) Enzymatic capture of an extrahelical thymine in the search for uracil in DNA. *Nature* 449(7161):433-437.

18. Kuznetsov NA*, et al.* (2007) Pre-steady-state kinetic study of substrate specificity of Escherichia coli formamidopyrimidine--DNA glycosylase. *Biochemistry* 46(2):424-435.

19. Eichman BF, O'Rourke EJ, Radicella JP, & Ellenberger T (2003) Crystal structures of 3-methyladenine DNA glycosylase MagIII and the recognition of alkylated bases. *The EMBO journal* 22(19):4898-4909.

20. Dunn AR, Kad NM, Nelson SR, Warshaw DM, & Wallace SS (2011) Single Qdot-labeled glycosylase molecules use a wedge amino acid to probe for lesions while scanning along DNA. *Nucleic acids research* 39(17):7487-7498.

21. Nelson SR, Dunn AR, Kathe SD, Warshaw DM, & Wallace SS (2014) Two glycosylase families diffusively scan DNA using a wedge residue to probe for and identify oxidatively damaged bases. *Proceedings of the National Academy of Sciences of the United States of America* 111(20):E2091-2099.

22. Koval VV, Kuznetsov NA, Ishchenko AA, Saparbaev MK, & Fedorova OS (2010) Real-time studies of conformational dynamics of the repair enzyme E. coli formamidopyrimidine-DNA glycosylase and its DNA complexes during catalytic cycle. *Mutation research* 685(1-2):3-10.

23. Jiang YL & Stivers JT (2002) Mutational analysis of the base-flipping mechanism of uracil DNA glycosylase. *Biochemistry* 41(37):11236-11247.

24. Handa P, Roy S, & Varshney U (2001) The role of leucine 191 of Escherichia coli uracil DNA glycosylase in the formation of a highly stable complex with the substrate mimic, ugi, and in uracil excision from the synthetic substrates. *The Journal of biological chemistry* 276(20):17324-17331.

25. Vallur AC, Feller JA, Abner CW, Tran RK, & Bloom LB (2002) Effects of hydrogen bonding within a damaged base pair on the activity of wild type and DNA-intercalating mutants of human alkyladenine DNA glycosylase. *The Journal of biological chemistry* 277(35):31673-31678.

26. Jiang YL, Kwon K, & Stivers JT (2001) Turning On uracil-DNA glycosylase using a pyrene nucleotide switch. *The Journal of biological chemistry* 276(45):42347-42354.

27. Jiang YL, Stivers JT, & Song F (2002) Base-flipping mutations of uracil DNA glycosylase: substrate rescue using a pyrene nucleotide wedge. *Biochemistry* 41(37):11248-11254.

28. O'Brien PJ & Ellenberger T (2003) Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines. *Biochemistry* 42(42):12418-12429.

29. Hsieh J, Walker SC, Fierke CA, & Engelke DR (2009) Pre-tRNA turnover catalyzed by the yeast nuclear RNase P holoenzyme is limited by product release. *RNA* 15(2):224-234.

30. Parikh SS*, et al.* (1998) Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *The EMBO journal* 17(17):5214-5226.

31. Stivers JT (2008) Extrahelical damaged base recognition by DNA glycosylase enzymes. *Chemistry* 14(3):786-793.

32. Lau AY, Scharer OD, Samson L, Verdine GL, & Ellenberger T (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. *Cell* 95(2):249-258.

33. Guo HH, Choe J, & Loeb LA (2004) Protein tolerance to random amino acid change. *Proceedings of the National Academy of Sciences of the United States of America* 101(25):9205-9210.

34. Aamodt RM, Falnes PO, Johansen RF, Seeberg E, & Bjoras M (2004) The Bacillus subtilis counterpart of the mammalian 3-methyladenine DNA glycosylase has hypoxanthine and 1,N6-ethenoadenine as preferred substrates. *The Journal of biological chemistry* 279(14):13601-13606.

35. Huffman JL, Sundheim O, & Tainer JA (2005) DNA base damage recognition and removal: new twists and grooves. *Mutation research* 577(1-2):55-76.

36. O'Brien PJ (2006) Catalytic promiscuity and the divergent evolution of DNA repair enzymes. *Chem Rev* 106(2):720-752.

37. Stivers JT, Pankiewicz KW, & Watanabe KA (1999) Kinetic mechanism of damage site recognition and uracil flipping by Escherichia coli uracil DNA glycosylase. *Biochemistry* 38(3):952-963.

38. Kuznetsov NA*, et al.* (2005) Kinetics of substrate recognition and cleavage by human 8-oxoguanine-DNA glycosylase. *Nucleic acids research* 33(12):3919-3931.

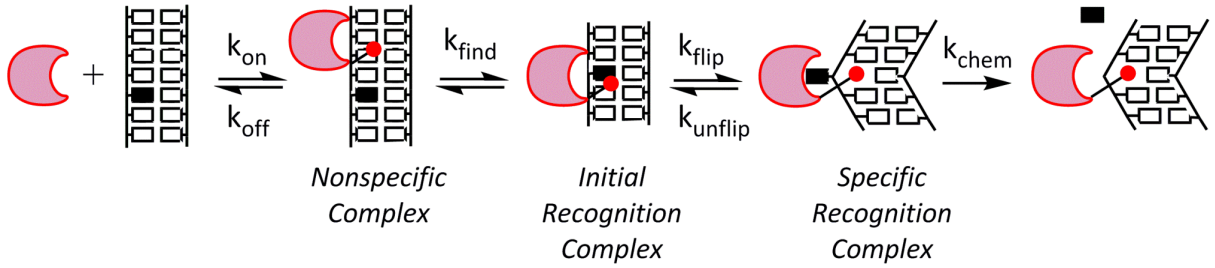
39. Vallur AC, Maher RL, & Bloom LB (2005) The efficiency of hypoxanthine excision by alkyladenine DNA glycosylase is altered by changes in nearest neighbor bases. *DNA repair* 4(10):1088-1098.

CHAPTER 6

Conclusions and Future Directions

Kinetic Mechanism of AAG

Human alkyladenine DNA glycosylase (AAG) is a monomeric repair protein responsible for removing a diverse set of alkylated and oxidized purine bases (1-4). Crystal structures of AAG complexed with damaged DNA revealed the DNA is bent where a tyrosine (Y162) intercalates into the space left by the flipped-out damaged nucleotide (5, 6). Previous work established the kinetic mechanism for AAG (Scheme 6.1; (7)). Briefly, AAG binds nonspecifically to DNA (k_{on}) and uses facilitated diffusion (8) to search for sites of damage. The searching process is described by the macroscopic rate constant k_{find}. This rate constant will depend on the length of the DNA and involves many microscopic searching steps. Once a lesion is found, AAG forms an initial recognition complex that is accompanied by changes in the stacking of the lesion base (7). From this intermediate, AAG catalyzes nucleotide flipping (k_{flip}) to form a more stable specific recognition complex that positions the substrate for *N*-glycosidic bond cleavage (5). In this specific complex the DNA is bent and the damaged nucleotide is flipped 180°. After *N*-glycosidic bond hydrolysis (k_{chem}), the lesion base is released. Dissociation of AAG from the



Scheme 6.1.

abasic product requires rotation of the abasic sugar back into the duplex (k_{unflip}). From the less tightly bound complex, AAG can translocate along the DNA in search for additional sites of damage before ultimately dissociating (k_{off}).

Direct Measurement of k_{find}

In the series of transient kinetic experiments used to dissect the individual steps of binding, flipping, and base excision by AAG, conditions of excess protein were used (7, 9). While this condition is most commonly employed to monitor labeled DNA, simultaneous searching by multiple AAG molecules makes the searching step too fast to detect (7). We modeled searching by multiple proteins explicitly to show that an additive model, in which two proteins search twice as fast as one protein, is consistent with the experimental data; however we were never able to directly measure k_{find} until recently (see Appendix B). We gained new insight into the DNA searching process by directly measuring the rate of searching under conditions of excess DNA. Although the fraction of DNA that is bound is much smaller under these conditions, we were fortunate because there was sufficient signal to accurately measure protein binding. Under these conditions the initial recognition complex is formed with a unimolecular rate constant of 120 s^{-1} for wild-type AAG (Chapter 3). This observed rate constant (k_{find}) is equal to the sum of the rate constants for finding and leaving the site of damage. For wild-type AAG this initial recognition complex is significantly populated, indicating that the rate of finding the εA

site is faster than the rate of leaving the site. Although multiple AAG molecules can bind to a single oligonucleotide, the observed rate constant for nucleotide flipping is identical under conditions of excess DNA or excess protein. This indicates that multiple proteins do not affect the flipping step and that rate constants obtained in experiments with multiple or single proteins bound can be directly compared. This experimental approach can be used with other DNA glycosylases to better understand the searching process.

Functional Insights into Active Site Tyrosines

Much is known about the removal of the naturally fluorescent lesion 1,N^6-ethenoadenine (εA) by AAG. However we sought to observe conformational changes in the protein directly. Similar approaches have been very informative for other DNA repair glycosylases (10-13). In Chapter 2, tryptophan residues were introduced into the binding pocket of AAG with the goal of measuring the rate constant for nucleotide flipping of non-fluorescent lesions. The two tyrosine residues that sandwich the bound lesion in the crystal structures (5, 6), Y127 or Y159, were replaced with tryptophan. While Y127W and Y159W had minimal effects on the excision of εA, there were much larger effects on the rate and equilibrium constant for nucleotide flipping. It has been suggested that a tyrosine to tryptophan mutation should be relatively conservative (14), however both positions are invariantly found to be tyrosine in AAG homologs in organisms ranging from bacteria to plants to vertebrates (15). Extensive mutagenesis of AAG and selection for active site variants that conferred resistance to DNA alkylating agents has identified hundreds of allowed mutations, but no active variants at position 127 have been identified (15, 16). Only two substitutions were observed at position 159 (N and F) in those experiments. This

suggests that these two active site residues are important for AAG to perform its role in DNA repair within the cell.

Therefore, the large effects on the rate and equilibrium constant for nucleotide flipping are consistent with the simple model that bulky tryptophan residues in the active site destabilize the flipped-out conformation and accelerate the rate constant for unflipping. In spite of these differences, stopped-flow fluorescence experiments show that these mutant proteins exhibit the same two-state binding model as was observed for wild-type AAG. These observations indicate that the εA lesion and tryptophan side chains located in the AAG active site both report on the initial recognition complex and on the specific recognition complex. Identical rates were observed from time-dependent changes in both εA and tryptophan fluorescence, demonstrating that there are already changes in the active site environment prior to full nucleotide flipping.

Role of the Intercalating Residue (Y162)

The functional contributions of the intercalating residue (Y162) to each of the steps in the AAG catalytic mechanism were addressed in Chapter 3. The intercalating tyrosine residue was mutated to phenylalanine (Y162F) and alanine (Y162A). Remarkably, the Y162A mutation increases the rate of nucleotide flipping by 50-fold relative to wild-type AAG. This mutant also exhibits faster unflipping, resulting in a 140-fold reduction in the equilibrium constant for formation of the flipped-out complex. In contrast, the kinetic parameters for the Y162F and wild-type enzymes are very similar, suggesting that the hydroxyl of tyrosine is not necessary for function. These results establish that DNA intercalation contributes to the specific recognition of DNA damage by acting as a plug to stabilize the specific recognition complex. Tyrosine 162

plays a critical role in the discrimination of damaged and undamaged nucleotides, by increasing the amount of the extrahelical lesion recognition complex and enhancing efficient repair of rare sites of damage.

Timing of DNA Intercalation

In Chapter 4, site-directed mutagenesis and transient kinetic approaches were used to determine the timing of tyrosine (Y162) intercalation for human AAG. The intercalating tyrosine was mutated to tryptophan since tryptophan fluorescence is an excellent tool to follow structural changes in a protein. Even though this mutation is not found in nature, the results indicate that Y162W is a conservative mutation. Excision of the naturally fluorescent lesion 1,N^6-enthenoadenine (εA) by Y162W is identical to wild-type, and the equilibrium constant for flipping remains highly favorable. Surprisingly, the tryptophan fluorescence of Y162W does not report on nucleotide flipping. Instead, the rapid quenching of Y162W tryptophan fluorescence upon binding damaged or undamaged DNA suggests that Y162 intercalation occurs early in the search for DNA damage.

Important Properties of the Intercalating Residue

Almost all DNA glycosylases that use nucleotide flipping have an intercalating residue to fill the space left by the flipped-out damaged nucleotide, but the identity of the intercalating amino acid varies. Human AAG uses an aromatic tyrosine residue. To investigate how the size (volume) of the intercalating side chain affects nucleotide flipping and base excision, seven Y162 variants of AAG were made and individual steps of the kinetic mechanism were characterized in Chapter 5. All seven Y162 mutants retained catalytic activity, measured by single-turnover excision of εA. This structure/function study shows that the formation of the initial recognition complex

and nucleotide flipping are affected by the size of this residue. Based on the results, the seven mutants fell into three categories: aromatic (Y162, Y162F, and Y162W) non-aromatic (Y162A, Y162N, Y162I, and Y162L), and Y162H. While Y162H is aromatic, the kinetic characterization of this mutant does not fully match the aromatic mutants or the non-aromatic mutants. Like the non-aromatic mutants, the rate constant for nucleotide flipping is much faster for Y162H compared to wild-type. However, the rate constant for unflipping is slow for Y162H, and more similar to the other aromatic mutants. Like the other aromatic mutants, Y162H is able to stabilize the specific recognition complex. Overall, all AAG proteins with an aromatic intercalating residue (Y162, Y162H, Y162F, and Y162W) have a favorable equilibrium constant for flipping. Altering the volume and polarity did not result in a systematic change in the rate constants for flipping and unflipping. The bulkier non-aromatic mutants (Y162N, Y162I, and Y162L) did not act like a plug, and were unable to prevent the damaged base from flipping back into the DNA. Perhaps this is due to increased mobility compared to aromatic side changes, which are expected to be able to make favorable pi-pi interactions with nucleobases of DNA. Therefore, intercalation of the aromatic residue dramatically slows the nucleotide flipping step, and the tyrosine acts as a plug to block the unflipping of εA.

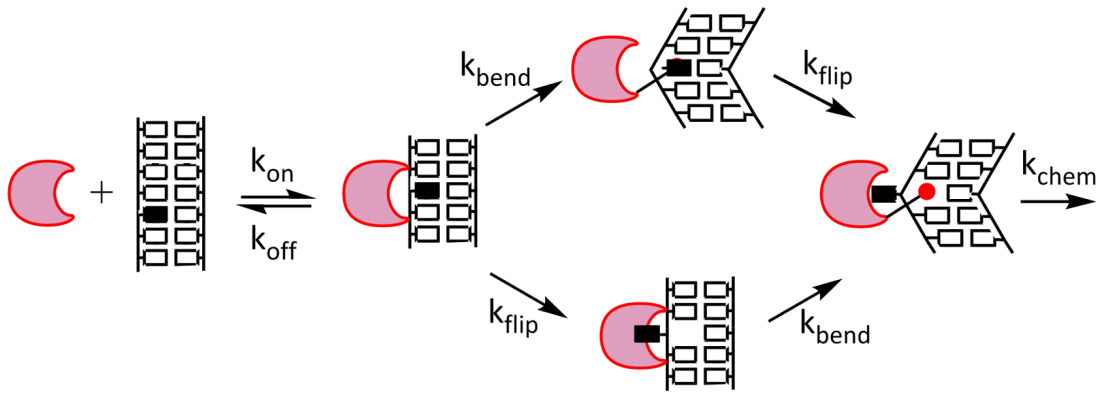
Overall Conclusions

Taken together, the kinetic mechanism of human AAG is optimized for specificity rather than speed, and the intercalating tyrosine residue plays a role throughout the kinetic mechanism. Intercalation occurs early in the search for damage, and tyrosine 162 plays an important role in discriminating between damage and undamaged nucleotides. Favorable aromatic stacking interactions between tyrosine 162 and nucleobases of DNA stabilize both the initial recognition

complex and the stable recognition complex. AAG has been an attractive model system to learn about the biophysical and biochemical mechanism of multi-step DNA recognition, and the lessons learned from AAG may be more widely applicable to other DNA modifying enzymes.

Future Directions

All enzymes that flip DNA bases are also thought to bend DNA at the target site. Crystal structures allow us to observe the final extrahelical state, and we can often determine the extent of DNA bending. However, the timing of DNA bending remains unknown. We would like to know whether the DNA bends prior to, simultaneous with, or after nucleotide flipping (Scheme 6.2). If DNA bending occurs prior to nucleotide flipping, this suggests the ease of bending DNA will contribute to the recognition of damage. If DNA bending is late, then bending can be excluded from the mechanism of damage recognition. Development of an assay to measure DNA bending would be beneficial to address the temporal relationship of bending and flipping. Fluorescence resonance energy transfer (FRET) between donor and acceptor fluorophores is a very sensitive method for measuring DNA bending induced by proteins (17).



Scheme 6.2.

To develop a real-time DNA bending assay for AAG, two orthogonal approaches should be tried since the extent of DNA bending upon AAG binding remains unknown. One set of FRET substrates would contain external fluorophores to observe large bend angles. A second set of substrates would contain nucleotide analogs inside the DNA duplex that should allow for measurement of smaller bend angles. Both types of DNA substrates would contain a centrally located εA lesion. Importantly, the excitation (313 nm) and emission (410 nm) of εA is distinct from the other fluorophores that would be used. For each set of FRET substrates, the donor and acceptor must first be tested individually to be sure AAG binding does not alter the fluorescence.

Upon addition of AAG to the first set of FRET substrates, the fluorescence emission of the acceptor is expected to increase, as DNA bending brings the two DNA ends into closer proximity. If AAG induces small DNA bends, this set of DNA substrates is unlikely to work. To observe 10-30° bend angles, a double-labeled U-shaped DNA construct should be made, or a proximal FRET pair should be used (17). In contrast, internal base analogs will be used in the second set of FRET substrates. These analogs are easy to position close to the site of bending, and they have negligible influence on the normal secondary structure of DNA (18).

Stopped-flow fluorescence can be used to correlate the changes in FRET with changes in εA fluorescence. Previous work established the transient changes in the fluorescence of εA reveal the existence of two distinct complexes that are formed on the millisecond timescale prior to hydrolysis (7, 9). An initial recognition complex is characterized by partial disruption of εA stacking interactions, followed by a stable extrahelical complex in which the εA lesion is strongly quenched by interactions in the AAG active site. Since the excitation and emission of

εA is distinct from the FRET fluorophores, we can follow both signals at the same time. If εA is flipped into the active site prior to DNA bending, we would predict the FRET signal to occur late in the reaction, after the fluorescence of εA is quenched (Scheme 6.2; bottom pathway). However, based on the timing of DNA intercalation determined in Chapter 4, we predict that DNA bending, and therefore FRET, will precede flipping of the εA base (Scheme 6.2; top pathway). Potentially, these experiments may reveal that the AAG-DNA complex is highly dynamic and undergoes conformational transitions between many states with different degrees of DNA bending. At that point it would be necessary to use single-molecule FRET to study rare kinetic or structural changes which are normally lost in the ensemble average. The new methods proposed here to study DNA bending will enable us to examine non-fluorescent lesions and undamaged DNA to learn more about the earliest steps in locating sites of DNA damage. These methods could be widely applicable to the many other proteins that bend DNA.

NMR Studies to Characterize Changes in Conformation and Dynamics of AAG

Previously published crystal structures of AAG bound to damaged DNA help us understand the static arrangement of residues in the final flipped-out specific recognition complex, however, we would like to know the structure and dynamics of AAG free in solution and when bound to undamaged DNA (5). Nuclear magnetic resonance (NMR) is an excellent technique that can be used to understand the molecular motions that occur as AAG binds to a DNA oligonucleotide. Specifically, NMR experiments would allow us to monitor the movement of the intercalating residue (Y162) throughout DNA binding and nucleotide flipping. Combining NMR experiments with the previous kinetic analysis will help us better understand the timing of DNA intercalation and damage recognition by AAG. A protocol for expression of AAG in minimal media was

developed to allow for ^{15}N labeling (*Wolfe Thesis*). The purification protocol was optimized to increase the yield of labeled protein, and preliminary HSQC experiments were done to assess the spectrum of AAG in solution as well as when bound to DNA. Unfortunately AAG was unstable under the experimental NMR conditions used, so optimization is still needed before additional NMR experiments can be performed to probe structural changes and differences in dynamics.

Using the Intrinsic Fluorescence of εA to Probe Structure and Dynamics in Different Sequence Contexts

Since the altered structure and increased dynamics of damaged DNA may serve as a signal to the repair proteins, experiments to study the effects of sequence context on εA fluorescence would help us to understand how DNA repair enzymes can identify damaged bases in DNA. We predict the structure and dynamics of DNA will change for different εA-DNA sequences, and the environment of εA within each sequence will affect the fluorescence. 1,N^6-ethenoadenosine (εA) is a common form of alkylative damage that can be repaired by several different DNA repair pathways (4, 19-21). This lesion is highly mutagenic in mammalian cells, sterically bulky, and does not form hydrogen bonds with the opposing thymine (T). The natural fluorescence of εA provides a probe of the local environment (22).

Two different 25 base pair oligonucleotides, TEC and AEA, each containing a central εA lesion were previously characterized (names refer to the central three nucleotides i.e., TEC has a central εA lesion flanked by a 5'T and a 3'C) (7, 9). While the excitation of both single-stranded oligonucleotides at 320 nm gave a broad fluorescence emission peak centered at 410 nm, the TEC oligonucleotide sequence gave 4-fold lower fluorescence than free εA. This was consistent with quenching of εA fluorescence even in single-stranded DNA. Surprisingly, the

oligonucleotide sequence AEA yielded much higher fluorescence, essentially identical to that of free εA. When appropriate complement was annealed to these sequences, then the fluorescence emission was decreased by a factor of 1.7-fold for TEC and 1.4-fold for AEA. This was consistent with increased quenching due to enhanced base stacking interactions in duplex DNA. It was not clear why these two sequence contexts showed such different quenching, but it is possible that the lack of nearby guanine nucleotides can account for the high fluorescence of the AEA sequence. A previous study of charge transfer between εA and normal bases reported that G is the most efficient quencher of εA, with only modest quenching occurring between εA and C, T, or A bases (23). Properties of the DNA π-stack can range from insulator- to "wire"-like. Therefore, the extent of stacking among εA and the DNA base-pair stack may affect the reaction kinetics and distance dependences of charge transfer mediated by the DNA helix.

Future work is needed to characterize additional DNA sequences to correlate sequence context with εA fluorescence in both single-strand and duplex DNA. Initial experiments showed that quenching of εA fluorescence is reporting on the conformation and dynamics of the εA. Overall, the various sequences tested had large differences in εA fluorescence, but additional DNA sequences should be tested to determine the distance dependence of εA quenching by guanine. In the future, stopped-flow fluorescence should be used to determine the rate constant of nucleotide flipping by AAG for each sequence. The most interesting sequences can be studied using NMR to better understand the structure and dynamics of damaged DNA (24). Potentially the rate of nucleotide flipping is correlated to the dynamics and or conformation of the εA lesion in the different sequence contexts.

References

1. Engelward BP*, et al.* (1997) Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. *Proceedings of the National Academy of Sciences of the United States of America* 94(24):13087-13092.

2. O'Brien PJ & Ellenberger T (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase. *The Journal of biological chemistry* 279(11):9750-9757.

3. Hitchcock TM*, et al.* (2004) Oxanine DNA glycosylase activity from Mammalian alkyladenine glycosylase. *The Journal of biological chemistry* 279(37):38177-38183.

4. Gros L, Ishchenko AA, & Saparbaev M (2003) Enzymology of repair of etheno-adducts. *Mutation research* 531(1-2):219-229.

5. Lau AY, Wyatt MD, Glassner BJ, Samson LD, & Ellenberger T (2000) Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG. *Proceedings of the National Academy of Sciences of the United States of America* 97(25):13573-13578.

6. Lau AY, Scharer OD, Samson L, Verdine GL, & Ellenberger T (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. *Cell* 95(2):249-258.

7. Wolfe AE & O'Brien PJ (2009) Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by human alkyladenine DNA glycosylase. *Biochemistry* 48(48):11357-11369.

8. Hedglin M & O'Brien PJ (2008) Human alkyladenine DNA glycosylase employs a processive search for DNA damage. *Biochemistry* 47(44):11434-11445.

9. Hendershot JM, Wolfe AE, & O'Brien PJ (2011) Substitution of active site tyrosines with tryptophan alters the free energy for nucleotide flipping by human alkyladenine DNA glycosylase. *Biochemistry* 50(11):1864-1874.

10. Kuznetsov NA*, et al.* (2005) Kinetics of substrate recognition and cleavage by human 8-oxoguanine-DNA glycosylase. *Nucleic acids research* 33(12):3919-3931.

11. Kuznetsov NA*, et al.* (2007) Pre-steady-state kinetic study of substrate specificity of Escherichia coli formamidopyrimidine--DNA glycosylase. *Biochemistry* 46(2):424-435.

12. Stivers JT, Pankiewicz KW, & Watanabe KA (1999) Kinetic mechanism of damage site recognition and uracil flipping by Escherichia coli uracil DNA glycosylase. *Biochemistry* 38(3):952-963.

13. Wong I, Lundquist AJ, Bernards AS, & Mosbaugh DW (2002) Presteady-state analysis of a single catalytic turnover by Escherichia coli uracil-DNA glycosylase reveals a "pinch-pull-push" mechanism. *The Journal of biological chemistry* 277(22):19424-19432.

14. Bordo D & Argos P (1991) Suggestions for "safe" residue substitutions in site-directed mutagenesis. *Journal of molecular biology* 217(4):721-729.

15. Chen CY*, et al.* (2008) Substrate binding pocket residues of human alkyladenine-DNA glycosylase critical for methylating agent survival. *DNA repair* 7(10):1731-1745.

16. Guo HH, Choe J, & Loeb LA (2004) Protein tolerance to random amino acid change. *Proceedings of the National Academy of Sciences of the United States of America* 101(25):9205-9210.

17. Dragan AI & Privalov PL (2008) Use of fluorescence resonance energy transfer (FRET) in studying protein-induced DNA bending. *Methods in enzymology* 450:185-199.

18. Wilhelmsson LM (2010) Fluorescent nucleic acid base analogues. *Quarterly reviews of biophysics* 43(2):159-183.

19. Speina E*, et al.* (2003) Decreased repair activities of 1,N(6)-ethenoadenine and 3,N(4)-ethenocytosine in lung adenocarcinoma patients. *Cancer Res* 63(15):4351-4357.

20. Pandya GA & Moriya M (1996) 1,N6-ethenodeoxyadenosine, a DNA adduct highly mutagenic in mammalian cells. *Biochemistry* 35(35):11487-11492.

21. Levine RL*, et al.* (2000) Mutagenesis induced by a single 1,N6-ethenodeoxyadenosine adduct in human cells. *Cancer Res* 60(15):4098-4104.

22. Leonard NJ (1984) Etheno-substituted nucleotides and coenzymes: fluorescence and biological activity. *CRC Crit Rev Biochem* 15(2):125-199.

23. Kelley SO & Barton JK (1999) Electron transfer between bases in double helical DNA. *Science* 283(5400):375-381.

24. Hansen AL, Nikolova EN, Casiano-Negroni A, & Al-Hashimi HM (2009) Extending the range of microsecond-to-millisecond chemical exchange detected in labeled and unlabeled nucleic acids by selective carbon R(1rho) NMR spectroscopy. *Journal of the American Chemical Society* 131(11):3818-3819.