# Kinetic Mechanism for Binding and Flipping of Damaged Bases By Alkyladenine DNA Glycosylase

**by**

**Abigail E. Wolfe**

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biological Chemistry)
in The University of Michigan
2010

Doctoral Committee:

Assistant Professor Patrick O'Brien, Chair
Professor Carol A. Fierke
Professor Ari Gafni
Associate Professor George A. Garcia
Associate Professor Bruce A. Palfey

To my amazing husband, Britton

# Acknowledgements

There are so many people that I'd like to thank for helping me finish. Through my advisor, Pat O'Brien, I have learned so much about science, research and myself. I appreciate his support and encouragement and patience with me throughout my time in graduate school. I also want to thank my committee members for great suggestions on my project and the opportunity to learn from them in courses and time in their labs. Thanks to the CBI training grant for providing funding for two years of my program. The Biological Chemistry Department was a great place to do my graduate work and I appreciate the work of the staff and faculty in the department. The other students in the lab and department have been great friends and colleagues during my time here. I would like to thank Jenna Tomlinson, a rotation student in the O'Brien lab, who I worked with to collect the data for Chapter 4. Thanks to Evgenia Nikolova and Hashim Al-Hashimi with whom we collaborated on an NMR project, the results of which are in Chapter 5.

I have made so many amazing friends during my time in Michigan. The people from Hope Christian Fellowship have been a wonderful blessing to me as well as my many friends from MCGrads. I will remember many special times with them and appreciate all of their prayers and support.

My family has been an incredible blessing to me as well. My parents, Ed and Leann, gave me a love of learning that led me to pursue my degree and they have been so supportive and encouraging throughout my life. Thanks to Chelsea and Ian who are always eager to listen and celebrate or commiserate, depending on what I needed. I also want to thank the rest of my family who have been such a blessing to me – my grandparents, the Wolfes, the Forstroms, aunts, uncles and cousins.

Finally, I want to thank my amazing husband, Britton. He has been a rock for me through this sometimes difficult journey. Thank you for your math and computer expertise that I always seemed to need. Thank you for reading my papers and listening to my presentations. Thank you so much for your patience and your encouragement. I couldn't have done it without you! I love you!

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

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

AlkA: 3-methyladenine DNA glycosylase

APE1: AP endonuclease 1

BER: base excision repair

BME: β-mercaptoethanol

BSA: bovine serum albumin

D: tetrahydrofuran (a stable abasic site analog)

DTT: dithiothreitol

εA: 1, *N6*-ethenoadenine

εAMP: 5'phospho 2'deoxyethenoAMP

EDTA: ethylenediaminetetraacetic acid

FAM: 6-aminofluorescein

Fpg: formamidopyrimidine DNA glycosylase

HSQC: heteronuclear single quantum coherence

Hx: hypoxanthine (the base moiety of inosine)

I: ionic strength

IPTG: isopropyl-β-D-thiogalactopyranoside

KPi: potassium phosphate

NaHEPES: sodium *N*-(2-hydroxyethyl)piperazine-*N*′-(2-ethanesulfonate)

NaMES: sodium 2-(*N*-morpholino)ethanesulfonate

NMR: nuclear magnetic resonance

PAGE: polyacrylamide gel electrophoresis.

PEI: polyethelenimine

PMSF: phenylmethanesulphonylfluoride

UV: ultraviolet

# Abstract

Alkyladenine DNA glycosylase (AAG) initiates the base excision repair pathway that repairs damage to single bases within DNA. AAG recognizes lesions caused by alkylation and deamination. AAG first locates the site of damage then excises the damaged base, leaving an abasic site in the DNA that is further processed by other repair proteins to complete the pathway. Although many substrates have been identified and there are high resolution structures, our understanding of the AAG mechanism remains incomplete. To investigate this further, the thermodynamics and kinetics of binding and base-flipping, along with structural conformational changes of AAG, were investigated. The thermodynamics of AAG binding to damaged and undamaged DNA was examined using fluorescence anisotropy. Surprisingly, this revealed that multiple proteins could bind with nanomolar affinity to short DNA oligonucleotides, which might be a common phenomenon for DNA repair enzymes. These results reveal the pitfalls of studying DNA binding by fluorescence anisotropy, since nonspecific binding dominates the changes in signal. The kinetic mechanism of the AAG reaction with 1,$N^6$-ethenoadenine (εA)-containing DNA was established, including binding, nucleotide flipping, base excision, and product release steps, by taking advantage of the natural fluorescence of the εA lesion. We observed that the flipping step is fast and the equilibrium for flipping is highly favorable. This kinetic mechanism maximizes specificity between damaged and undamaged bases. To study possible conformational changes in AAG, we took two approaches. First, tyrosine residues in the active site pocket were mutated to tryptophans

to serve as fluorescence reporters. We found Y127W and Y159W mutants had robust activity towards εA. However, a full kinetic characterization revealed that these mutations have large effects on the rates and equilibria for flipping. This suggests these mutants will have limited utility in studying recognition and flipping of other damaged nucleotides. Secondly, preliminary experiments established the feasibility of using NMR to study AAG and provided evidence for extensive conformational changes that take place upon binding to DNA. These studies have provided a mechanistic framework that will facilitate future investigations into the role of conserved residues and the energetic basis for the discrimination between damaged and undamaged DNA.

# Chapter 1
# Introduction

DNA is subject to many different types of damage from both endogenous and exogenous sources such as ionizing radiation, UV radiation, spontaneous deamination, alkylation, oxidation and DNA polymerase errors. These sources of DNA damage can cause many types of DNA damage, including single or double strand breaks, single base lesions, crosslinks and mismatches (*1*). Failure to repair DNA damage can lead to cancer causing mutations or cell death. Since there are only two copies of DNA within each cell and these cannot be replaced, it is essential that cells repair the damage. With so many types of damage occurring in the DNA, cells have several different pathways of repair, including homologous recombination, nonhomologous end joining, base excision repair and mismatch repair (*2*).

The major pathway utilized to repair damage to single bases within the DNA is called the base excision repair (BER) pathway (*3-4*), which is outlined in Figure 1.1. This is the pathway for a monofunctional glycosylase. Other pathways of BER are possible for bifunctional glycosylases that also catalyze β-elimination. The first step in the pathway is for a DNA glycosylase to recognize a site of DNA damage. The glycosylase cleaves at the glycosidic bond, yielding a free base and an abasic site in the DNA. The glycosylase likely initiates this reaction by positioning a water molecule in the active site that attacks at the glycosidic bond, which can be seen in Figure 1.2 (*5*). The abasic site is further

processed by an AP endonuclease, which nicks the DNA strand 5' of the deoxyribose. The deoxyribose is then removed by a dRP lyase activity. In humans, the major dRP lyase activity is provided by the lyase domain of DNA polymerase β. The DNA polymerase then fills in the gap in the DNA, using the intact strand as a template. Finally, a DNA ligase seals the nick in the DNA strand to complete the repair pathway and restore the DNA to its undamaged state.

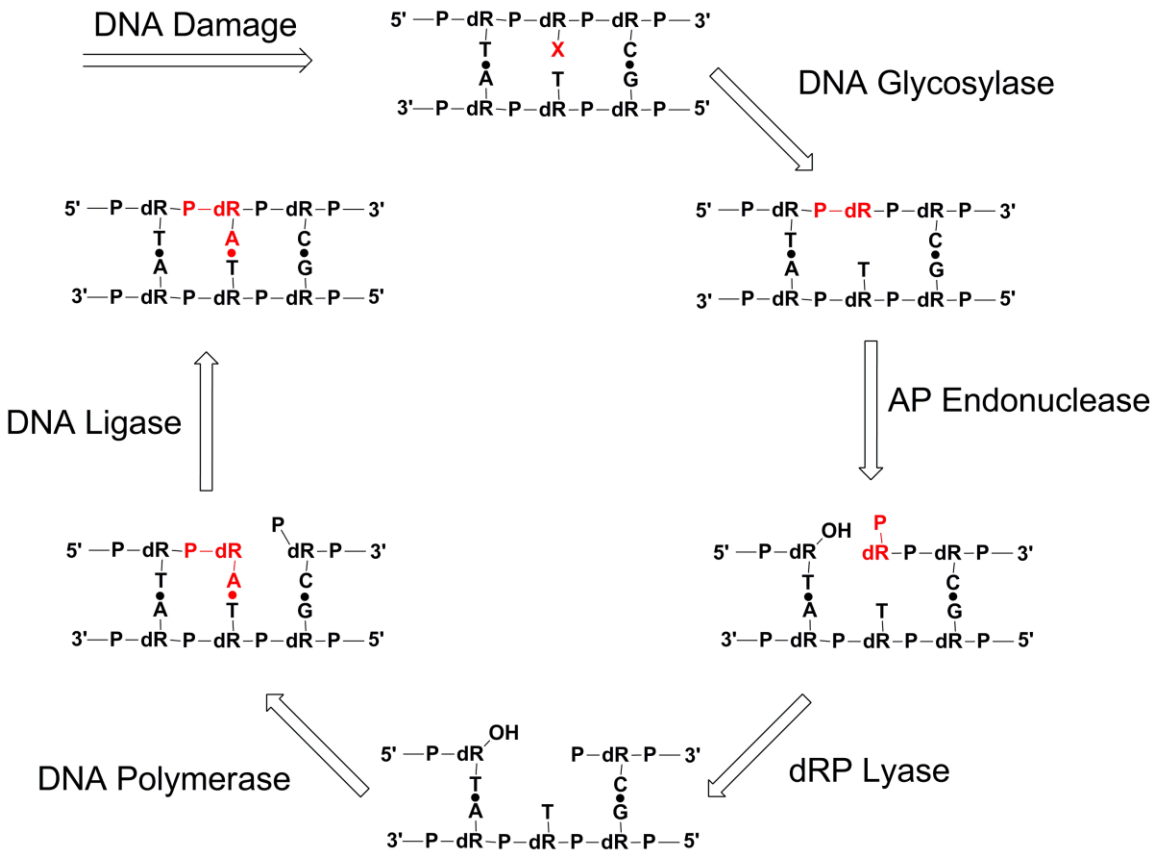



Figure 1.1. Base excision repair pathway initiated by a monofunctional DNA glycosylase. Shown is a single base lesion, X, in red and the intermediates formed during the base excision repair pathway. Once DNA damage is recognized by a DNA glycosylase, it cleaves the glycosidic bond, leaving an abasic site in the DNA. AP Endonuclease binds to the abasic site, cleaving 5' of the deoxyribose, leaving a nick in the DNA strand. The dRP lyase activity of DNA polymerase β removes the deoxyribose then the DNA polymerase fills in the gap, using the intact strand as a template. DNA ligase seals the nick in the DNA, completing the repair.

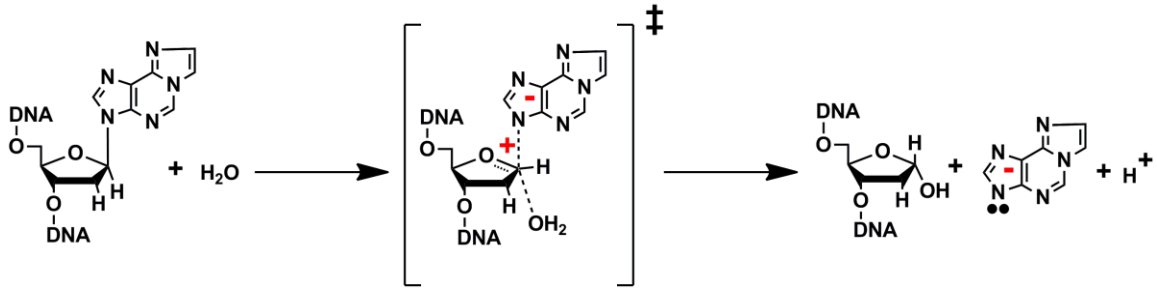



Figure 1.2. DNA glycosylase excision reaction. The glycosylase finds a damaged base within the DNA, here εA is the lesion shown, and the excision reaction is initiated with the attack of the N-glycosidic bond by a catalytic water molecule. The damaged base is released into solution and an abasic site is left in the DNA strand, which is further processed by the BER pathway to complete the repair.

It is important for the rest of the pathway to be completed once it has been initiated by a DNA glycosylase. This is due to the fact that some of the DNA intermediates, such as an abasic site or single strand break, of BER can be more harmful to the cell than the original lesion.

There are ten different known DNA glycosylases in humans (*3*). Some glycosylases are very specific and only recognize one lesion and there are a few that are broadly specific and can recognize and cleave many different lesions. As mentioned before, there are also monofunctional and bifunctional DNA glycosylases.

One of these monofunctional, broadly specific DNA glycosylases is alkyladenine DNA glycosylase (AAG). AAG recognizes many different lesions caused by alkylation and deamination to purines (*6*-*7*), some of which are shown in Figure 1.3. Also shown are the undamaged A and G bases, which are not good substrates for AAG. The lesions that AAG recognizes are very diverse. Some are positively charged, such as 7-methylguanine and 3-methyladenine. Others are large and bulky adducts, such as 1, *N6*- ethenoadenine (εA). Some lesions are of similar size to the undamaged bases, such as purine and

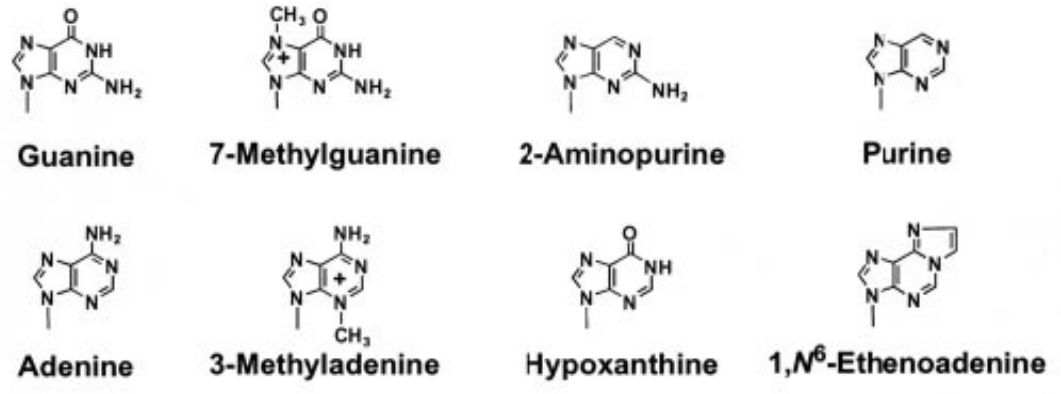



Figure 1.3. Diversity of AAG substrates. On the left are the undamaged bases, guanine and adenine. The rest of the bases are a subset of the lesions that AAG recognizes as substrates.

hypoxanthine. Thus, there is no common feature about the lesions that allows AAG to select for them. Therefore, AAG has methods to select against bases that are not substrates, such as undamaged bases. For the undamaged purines, adenine and guanine, AAG selects against them due to steric clashes in the active site between the 6-amino group of adenine and H136 in the active site or the 2-amino group of guanine and the N169 residue of AAG (*8*). Although pyrimidines are smaller than the substrates cleaved by AAG and could presumably fit into the AAG active site, they are selected against due to the acid-base catalytic mechanism employed by AAG (*5*).

A high resolution crystal structure of AAG bound to a 13 base pair εA-containing DNA oligonucleotide has been previously solved and is shown in Figure 1.4 (*9*). AAG interacts with about 10 base pairs of the DNA. Clearly seen in the structure, AAG employs a base flipping mechanism, rotating the εA 180° out of the duplex and into the active site in order to orient the glycosidic bond in a position for attack by the catalytic water molecule.

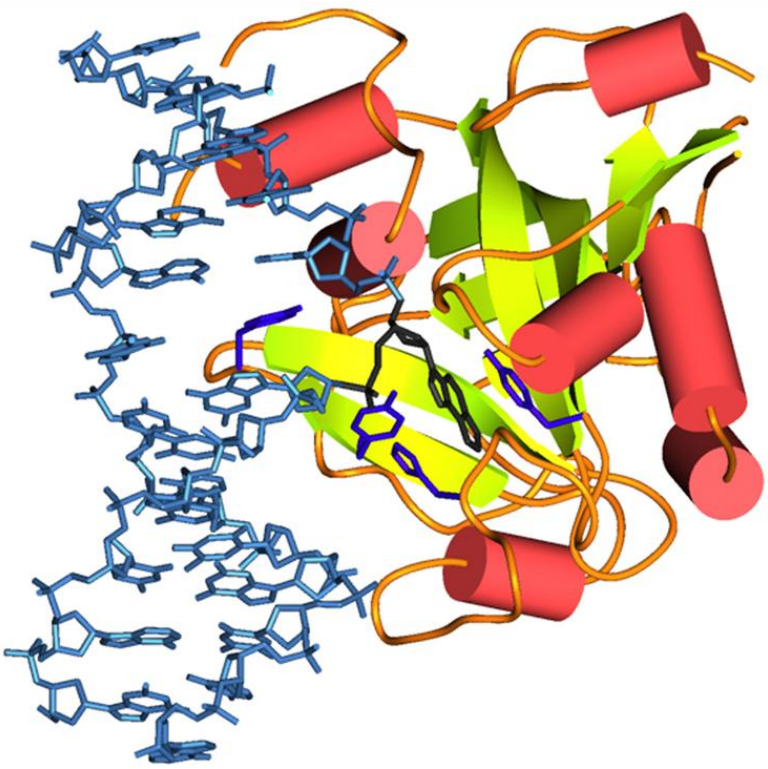


Figure 1.4. Crystal Structure of AAG. AAG is in ribbon representation on the right and the 13mer DNA is on the left in blue stick representation. Key active site residues of the protein are in blue and the εA lesion is in black. AAG employs a base flipping mechanism which can be seen as εA is flipped 180° out of the duplex and into the active site of AAG. (9)

Two tyrosine residues and one histidine residue line the active site pocket and interact with the lesion. Y162 intercalates into the DNA duplex to occupy the space that is left when εA is flipped. Not highlighted in the structure, but located in the active site is the residue E125, which orients the catalytic water molecule so it can attack the glycosidic bond of the substrate.

The crystal structure of AAG bound to a lesion-containing DNA provides a lot of information about the conformational changes that occur during base flipping. To achieve the flipped structure seen here, AAG must undergo at least three conformational changes: the DNA bends by 20°, the εA lesion flips 180° into the active site of AAG and Y162

intercalates into the DNA duplex. However, this structure is static and does not give any information about the kinetics of the AAG reaction.

Much is known about the substrates of AAG and its catalytic mechanism, but little is known about the binding of AAG to DNA. Chapter 2 addresses the thermodynamic parameters and stoichiometry of AAG binding to different DNA oligonucleotides.

The kinetic mechanism of εA excision by AAG is addressed in Chapter 3. Utilizing the fluorescence of εA, the minimum number of steps and the rate constants for each step were determined. This information is extended in Chapter 4, in which the conformational changes in AAG were investigated by substituting active site tyrosines with tryptophan. The fluorescence of the tryptophan residues in AAG was monitored and the fluorescence changes were compared to changes seen in εA fluorescence.

NMR is another powerful way to investigate conformational changes. In Chapter 5, preliminary experiments are described that demonstrate the feasibility of doing NMR dynamics experiments with AAG.

## References


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

2. Kao, J., Rosenstein, B. S., Peters, S., Milano, M. T., and Kron, S. J. (2005) Cellular response to DNA damage, *Ann N Y Acad Sci 1066*, 243-258.

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

4. Almeida, K. H., and Sobol, R. W. (2007) A unified view of base excision repair: Lesion-dependent protein complexes regulated by post-translational modification, *DNA Repair 6*, 695-711.

5. O'Brien, P. J., and Ellenberger, T. (2003) Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines, *Biochemistry 42*, 12418-12429.

6. Wyatt, M. D., Allan, J. M., Lau, A. Y., Ellenberger, T. E., and Samson, L. D. (1999) 3-methyladenine DNA glycosylases: structure, function, and biological importance, *Bioessays 21*, 668-676.

7. McCullough, A. K., Dodson, M. L., and Lloyd, R. S. (1999) Initiation of base excision repair: glycosylase mechanisms and structures, *Annu Rev Biochem 68*, 255-285.

8. O'Brien, P. J., and Ellenberger, T. (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase, *J Biol Chem 279*, 9750-9757.

9. Lau, A. Y., Wyatt, M. D., Glassner, B. J., Samson, L. D., and Ellenberger, T. (2000) Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG, *Proc Natl Acad Sci U S A 97*, 13573-13578.

# Chapter 2

# Alkyladenine DNA Glycosylase Binds with High Density to Single- and Double-Stranded DNA

Alkyladenine DNA glycosylase (AAG) initiates the base excision repair (BER) pathway by finding the site of damage and cleaving at the N-glycosidic bond, yielding a free damaged base and an abasic site in the DNA. The abasic site is further processed by the rest of the BER pathway to complete the repair. Many DNA repair proteins can use nonspecific DNA binding to enable a more efficient one dimensional search of the genome (e.g. see (*1-4*)). Recently, AAG has been shown to use a similar mechanism to slide along the DNA to find a site of damage (*5*).

Although nonspecific binding is important for locating the site of damage, AAG binding to the specific site is important for catalysis to occur. There is little information about the affinity of AAG for damaged and undamaged DNA, so the thermodynamics of AAG binding to DNA was examined.

A useful way to measure binding to DNA is by fluorescence anisotropy (*6-8*). With a fluorescent label on the end of the DNA, the fluorescence anisotropy is expected to increase when a protein binds to the DNA. This is due to an increase in mass, which slows the tumbling in solution. When these experiments were performed with AAG, some interesting and unexpected results were found.

AAG contains an N-terminal domain that is thought to be involved in protein-protein interactions. The N-terminal domain has been shown to make a modest contribution to AAG processivity under physiological conditions *(5)*. For *in vitro* experiments, the full length protein is not very stable and is difficult to work with. A mutant that has truncated the N-terminal domain of AAG, leaving only the catalytic C-terminal domain, has been made (Δ80 AAG). This truncation mutant is much more stable *in vitro* and has been found to have identical catalytic activity to the full length enzyme in a variety of assays (*9*).

Glutamate 125 is in the active site of AAG and positions the catalytic water molecule so that it can attack the N-glycosidic bond. Mutation of this residue to a glutamine abolishes AAG activity (*9-10*). This E125Q AAG mutant is useful for experiments where a catalytically inactive protein is needed.

## Materials and Methods

### Purification of Recombinant Proteins

The catalytic domain of human AAG that lacks the first 80 amino acids was expressed in *E. coli* both with and without the catalytically inactive mutation E125Q, then purified as previously described (*9*). The full length human AAG protein (FL AAG) was expressed in *E. coli* and purified as previously described (*9*). The *E. coli* protein AlkA was expressed in *E. coli* and purified as previously described (*11-12*). The human protein APE1 was expressed in *E. coli* and purified as previously described (*13*).

The concentration of active AAG for the wildtype proteins (FL AAG and WT Δ80 AAG) was determined by burst analysis, using an inosine-containing oligonucleotide substrate, as previously described (*5*). The concentration for the other proteins (E125Q

AAG, AlkA and APE1) was determined by measuring the absorbance at 280 nm and calculated with the respective extinction coefficients ($2.5 \times 10^4$ $M^{-1}$ $cm^{-1}$, $6.7 \times 10^4$ $M^{-1}cm^{-1}$, and $5.6 \times 10^4$ $M^{-1}$ $cm^{-1}$).

**Synthesis and Purification of DNA Oligonucleotides**

The oligonucleotides of lengths 13, 25 and 49 nucleotides were obtained from commercial sources and contained a central lesion that was placed opposite of a T. The sequences are given in Table 2.1. The oligonucleotides were labeled with a 5'-fluorescein (6-fam) label with the duplex oligonucleotides being labeled on the non-lesion containing strands. Oligonucleotides containing only normal deoxynucleotides and the abasic site

Table 2.1. DNA oligonucleotide sequences.

| DNA[a] | Sequence |
|---|---|
| εA·T 49mer | 5′-GACATGATTGCCCGATAGCATCCT**E**CCTTCTCTCCATGCGTCAATTGTC<br>3′-CTGTACTAACGGGCTATCGTAGGATGGAAGAGAGGTACGCAGTTAACAG |
| εA·T 25mer | 5′-CGATAGCATCCT**E**CCTTCTCTCCAT<br>3′-GCTATCGTAGGATGGAAGAGAGGTA |
| A·T 25mer | 5′-CGATAGCATCCT**A**CCTTCTCTCCAT<br>3′-GCTATCGTAGGATGGAAGAGAGGTA |
| D·T 25mer | 5′-CGATAGCATCCT**D**CCTTCTCTCCAT<br>3′-GCTATCGTAGGATGGAAGAGAGGTA |
| εA·T 13mer | 5′- GACATG**E**TTGCCT<br>3′-ACTGTACTAACGG |
| A·T 13mer | 5′- GACATG**A**TTGCCT<br>3′-ACTGTACTAACGG |
| D·T 13mer | 5′- GACATG**D**TTGCCT<br>3′-ACTGTACTAACGG |
| ss εA 25mer | 5′-CGATAGCATCCT**E**CCTTCTCTCCAT |
| ss T 25mer | 5′-ATGGAGAGAAGG**T**AGGATGCTATCG |

[a]The double-stranded DNA oligonucleotides are named based on the central lesion or undamaged base that is opposite a T and the length of the DNA strand. εA is 1, *N6*-ethenoadenine and D is tetrahydrofuran, an abasic site analog. Two single-stranded (ss) DNA oligonucleotides are also shown and named for the central base and the length of the DNA strand.

analog, tetrahydrofuran (D), were synthesized and deprotected with standard phosphoramidite chemistry. The oligonucleotides containing 1, *N6*-ethenoadenine (εA) were synthesized by the W. M. Keck Facility at Yale University (New Haven, CT), or the phosphoramidite was obtained from Glen Research and incorporated using an ABI 394 DNA synthesizer then deprotected using Ultra-Mild chemistry according to the supplier's recommendations. The oligonucleotides were desalted using Sephadex G-25 and purified using denaturing polyacrylamide gel electrophoresis as previously described (*5*, *9*). 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 εA containing oligonucleotides, the extinction coefficient was calculated for the same sequence with an A in place of the εA, then 9400 $M^{-1}cm^{-1}$ was subtracted to account for the weaker absorbance of εA as compared to A (*14*). The lesion-containing oligonucleotides were annealed with a 1.1-fold excess of the complement by heating to 90°C and cooling slowly to 4°C.

**Fluorescence and Fluorescence Anisotropy Binding Experiments**

Fluorescence anisotropy measurements were collected with a PTI QuantaMaster fluorometer controlled by FeliX software. To measure anisotropy, both the parallel ($I_{||} = I_{VV}$) and perpendicular ($I_{\perp} = I_{VH}$) fluorescence intensity with respect to the vertical excitation light must be measured. The fluorometer was set up in a T-format with polarizers set up to let parallel light pass through one side and perpendicular light pass through the other. Thus, both fluorescence intensity measurements could be made simultaneously.

Samples were prepared in a buffer containing 50 mM NaHEPES, pH 7.5, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1 mg/mL BSA. All measurements were recorded at 25°C. Experiments with 10 nM DNA were performed by adding concentrated aliquots of protein to 2.5 mL solutions of FAM-labeled DNA and recording the parallel and perpendicular fluorescence intensity. For experiments with higher concentrations of DNA (100-500 nM), a microcuvette was used and concentrated aliquots of protein were added to 200 μL solutions of FAM-labeled DNA and the parallel and perpendicular fluorescence intensities were recorded. To measure FAM fluorescence intensity, an excitation wavelength of 495 nm (band pass was between 6 and 12 nm) and emission wavelength of 518 nm (band pass was between 6 and 12 nm) were used.

The G-factor was measured by exciting with horizontal light and measuring both the parallel ($I_{HH}$) and perpendicular ($I_{HV}$) light with respect to the excitation light. The G-factor is then calculated ($G=I_{HV}/I_{HH}$). The G-factor is a term that accounts for differences in the two different light paths in a T-format fluorometer.

The total fluorescence intensity was calculated with equation 2.1, where $I_T$ is the total fluorescence intensity, $I_{||} = I_{VV}$ is the parallel fluorescence intensity and $I_{\perp} = I_{VH}$ is the perpendicular fluorescence intensity. The anisotropy values were calculated using equation 2.2, where r is the anisotropy, $I_{||} = I_{VV}$ is the parallel fluorescence intensity, $I_{\perp} = I_{VH}$ is the perpendicular fluorescence intensity and G is the G-factor (*8*).

$$I_T = I_{||} + 2I_{\perp} \qquad \text{(Equation 2.1)}$$

$$r = \frac{I_{||} - G * I_{\perp}}{I_{||} + 2 * G * I_{\perp}} \qquad \text{(Equation 2.2)}$$

For experiments measuring εA fluorescence, an excitation wavelength of 310 nm (4 nm band-pass) and emission wavelength of 410 nm (612 nm band-pass) were used.

When equilibrium binding constants were able to be measured, the data were fit to Equation 2.3, where $F_{bound}$ is the fraction bound, [DNA] is the initial concentration of DNA, [E] is the concentration of AAG, $K_d$ is the dissociation constant. The second phase of the anisotropy data was fit to Equation 2.4, which was Equation 2.3 that was changed so that [E] was replaced with ([E]-C), where C is the nonzero y-intercept, .

$$F_{bound} = \frac{[DNA]+[E]+K_d-\sqrt{([DNA]+[E]+K_d)^2-4[DNA][E]}}{2[DNA]} \quad \text{(Equation 2.3)}$$

$$F_{bound} = \frac{[DNA]+([E]-C)+K_d-\sqrt{([DNA]+([E]-C)+K_d)^2-4[DNA]([E]-C)}}{2[DNA]}$$

(Equation 2.4)

## Results and Discussion

### Anisotropy Shows High Density Binding of AAG to εA-Containing DNA

Experiments were done to compare εA·T 25mer DNA binding by both full length AAG (FL AAG) and the N-terminal truncation mutant (Δ80 AAG) to see if the N-terminal domain of AAG plays a role in AAG binding. When full length AAG (FL AAG) was used, the anisotropy saturates at a high [AAG] (Figure 2.1A). The anisotropy also reaches a much higher value than is expected for a protein the size of AAG binding to a 25mer oligonucleotide (see below). Additionally, the FAM fluorescence seen during the binding experiment quenches at a high [AAG] (Figure 2.1B). These results were

unexpected and suggest that FL AAG is binding at high density to this DNA. On the timescale of this experiment, no appreciable glycosylase activity was expected.

We are using "high density binding" to represent the idea that multiple AAG molecules are binding to a single DNA oligonucleotide. In fact, more AAG molecules can bind than is predicted by the footprint of AAG from the crystal structure (see below). These experiments indicate that AAG is binding at high density due to the large change in anisotropy seen when AAG is bound to the DNA and the high stoichiometry necessary to see the maximal change in anisotropy during these experiments (greater than 1:1).

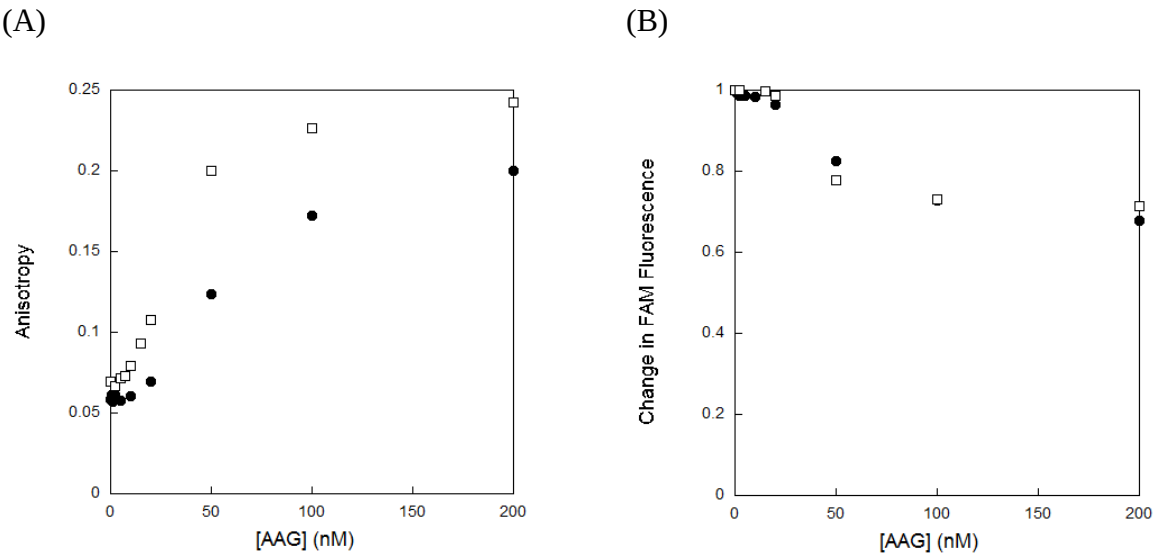



Figure 2.1. Binding of full length and Δ80 AAG to εA-containing DNA. Experiments were performed to compare the effect of the N-terminal domain of AAG, which is truncated for Δ80 AAG, on binding to 10 nM εA·T 25mer. Δ80 WT AAG (●) and full length WT AAG (□) were each added to the DNA in small increments. (A) The anisotropy shows that the curves for each protein reach a high saturating value, which indicates that both Δ80 WT AAG and FL WT AAG bind at high density to DNA. (B) The quenching in FAM fluorescence intensity is also shown. Data shown is from single experiments.

When Δ80 AAG was added to the same DNA, a similar result was seen (Figure 2.1A and B). Although the two experiments do not reach the same final anisotropy value, it takes a greater than 1:1 ratio of AAG: DNA to achieve saturation of the anisotropy

signal. Even without the N-terminal domain, AAG can still bind at high density to DNA. Additionally, the FAM fluorescence quenching curve is very similar to that of FL AAG, which suggests that the two proteins have similar behavior. The rest of the experiments described here use Δ80 AAG since it is more stable *in vitro*.

Since the experiments are performed under equilibrium conditions, it would be advantageous to use a catalytically inactive mutant of AAG to perform these experiments in order to prevent catalysis from occurring. Binding to εA·T 25mer DNA by wildtype Δ80 AAG and E125Q AAG was measured (Figure 2.2). For both proteins, high density binding was seen. However, some differences in the two binding curves were apparent. The final anisotropy value is higher in the experiment with E125Q AAG than with WT AAG. Also, E125Q AAG saturates at a much lower [AAG] than is seen for WT AAG.

(A) (B)

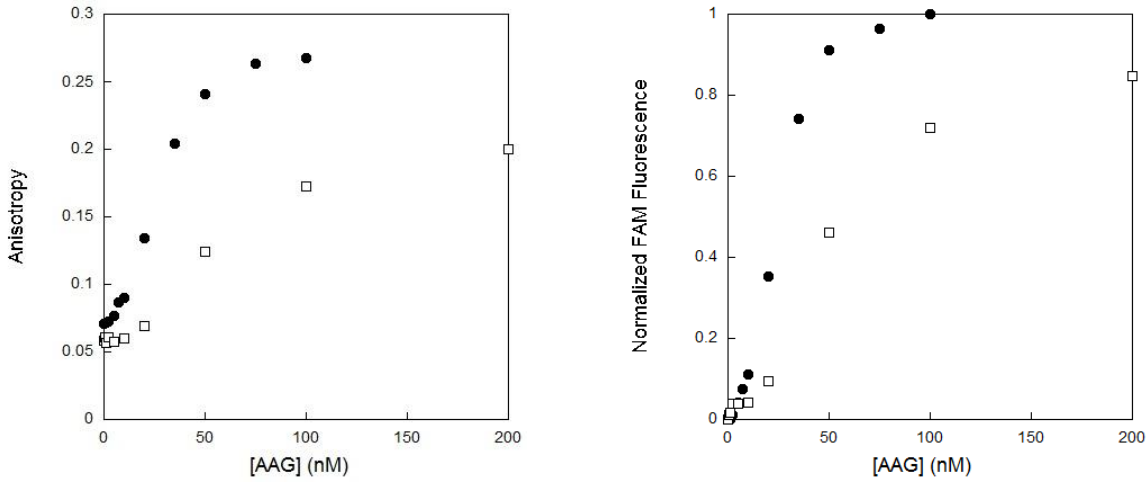



Figure 2.2. Comparison of WT to E125Q AAG. Experiments were performed with 10 nM εA·T 25mer and Δ80 E125Q AAG (●) or Δ80 WT AAG (□). (A) The anisotropy plots both show a high saturating value, indicating that E125Q AAG exhibits a similar behavior of binding at high density to εA·T 25mer as seen with Δ80 WT AAG. This suggests that using the E125Q mutant is a valid substitution to use, which prevents excision by AAG. (B) The change in FAM fluorescence intensity is also shown. Data shown is from single experiments.

This indicates that E125Q AAG binds more tightly to nonspecific sites than WT AAG does. The FAM fluorescence showed similar curves quenching for the two proteins, but FAM quenching saturated at a lower concentration of AAG for the E125Q mutant. Since the E125Q mutant has no glycosylase activity (*9-10*), the remainder of the experiments used this protein.

To test that the high density binding is real and not an artifact, experiments were performed at different concentrations of εA·T 25mer DNA (Figure 2.3). Figure 2.3A and C show the raw data from the experiments, while Figure 2.3B and D show the data after it has been normalized for the difference in DNA concentration. This normalization to equivalents of AAG enables a direct comparison of the three experiments. Figure 2.3D was also normalized in the y-axis, showing the FAM quenching on the same scale. These experiments show the same break in the curve at 1:1 stoichiometry (Figure 2.3B). In all three experiments, a similar linear curve is seen up to a 1:1 ratio of AAG: DNA. The most likely explanation for this is that AAG binds with highest affinity to the lesion site on the DNA and this is saturated at a 1:1 ratio of AAG: DNA. As the concentration of AAG is further increased beyond 1:1, the anisotropy continues to increase, indicating AAG molecules are binding to other, nonspecific, sites on the DNA.

When the high concentration of DNA (500 nM) was used, there is a pronounced separation between the two phases of the curve. For the middle concentration of DNA (100 nM) a separation between the two phases is seen, yet it is less pronounced than for the high concentration of DNA. Because of this distinct separation between the phases in the curves for both of these concentrations of DNA, these curves are similar and suggest a titration is being performed in both cases, where the concentration of DNA is higher

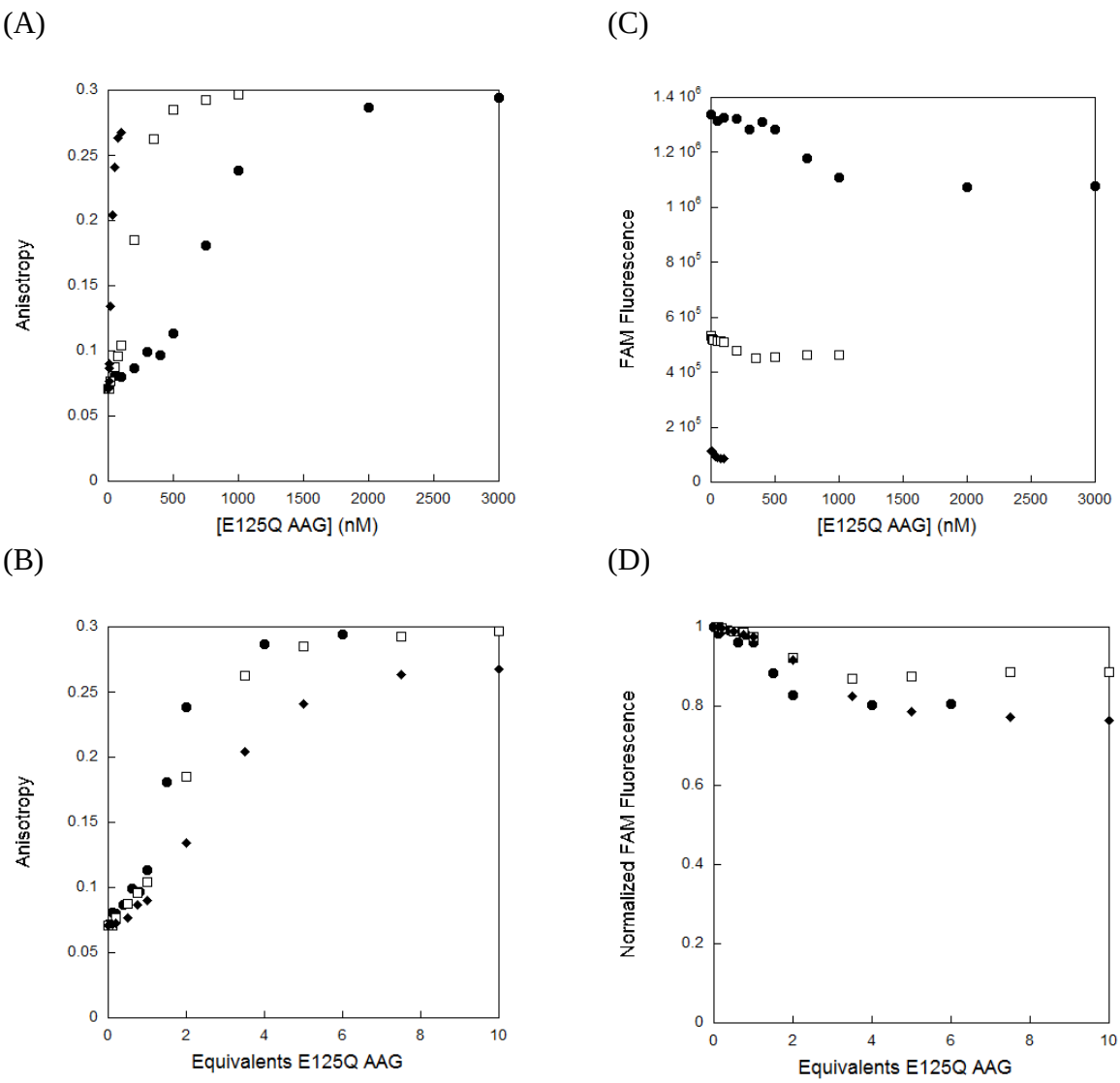



Figure 2.3. Binding of E125Q AAG to εA-containing DNA at different concentrations of DNA. E125Q AAG was added to εA·T 25mer DNA at concentrations of 10 nM (♦), 100 nM (□) and 500 nM (●). The raw data from these experiments are shown in panels (A) and (C), displaying the anisotropy and FAM fluorescence intensity, respectively. Each binding curve is achieving a high maximal anisotropy value. The FAM fluorescence intensity shows a decrease in intensity as more than one equivalent of AAG binds to the DNA oligonucleotide. The data in panels (B) and (D) are both plotted on an x-axis that is [AAG]/ [DNA] to get equivalents of E125Q AAG added, showing anisotropy and the FAM fluorescence intensity normalized to the signal from the DNA alone, respectively. This conversion allows the three curves to be compared directly and shows a similar linear behavior up to a 1:1 ratio of E125Q AAG: DNA. The curves diverge after the 1:1 concentration, with the 500 nM DNA data showing a more pronounced separation between the two phases of the curve. As the [DNA] is decreased, it becomes closer to the $K_d$ and is influenced by equilibrium binding of E125Q AAG, as seen in the 10 nM DNA data. Data shown is from single experiments.

nM), there is not a prominent separation between the two phases of the binding curve, suggesting that equilibrium binding is influencing the nonspecific binding curve. This indicates that the $K_d$ for nonspecific binding is now greater than [DNA]. This data gives that the $K_d$ for AAG binding to nonspecific DNA is between 10 and 100 nM. This result that AAG binds more tightly to the εA lesion than the nonspecific sites on the DNA gives an upper limit for the $K_d$ of AAG binding to εA on the order of 10 nM, which is consistent with previous measurements of ~20 nM (*10, 15*).

The normalized FAM fluorescence from these experiments is seen in Figure 2.3D, which shows very similar curves for all three concentrations of DNA. The FAM fluorescence shows little change up to a 1:1 stoichiometry then quenching can be seen up to a stoichiometry of ~3:1. After this, no further change in FAM fluorescence is seen. It is likely that this is caused by the first AAG molecule binding at the site of damage in the center of the DNA, which does not change the environment of the FAM molecule. When one or more of the additionally bound proteins binds, it is then physically close to the 5' end of the DNA, resulting in increased quenching of the FAM. When the FAM quenching no longer changes at high concentrations of AAG, this suggests that the DNA is fully saturated with AAG molecules.

One of the limitations of these experiments is that the theoretical maximum anisotropy value is 0.4. It is likely that the maximum anisotropy value we can measure in this system is even less than that since there is flexibility in the linker region between the DNA and the FAM label. This limitation exists because the fluorescence lifetime of fluorescein (4 ns) (*16*) restricts the maximum anisotropy signal that can be detected. Since most of the binding curves saturate around an anisotropy value of 0.3, it is possible

that there could be more AAG molecules binding to each DNA. This means that the saturation of the anisotropy signal does not necessarily indicate that the DNA is saturated with AAG molecules. If a fluorophore with a longer lifetime than FAM was used, the experiments could reach a higher anisotropy value, it could reveal a greater ratio of AAG: DNA binding. Since the FAM fluorescence does not have these limitations, it is expected that quenching of the fluorescent signal is indicative of the DNA being fully saturated with AAG molecules.

**Simple 1:1 Binding Seen with APE 1**

Since these results are unusual, we examined the binding of a different protein, Human AP endonuclease 1 (APE1) to a similar DNA. APE1 catalyzes the second step in BER, where it cleaves the DNA backbone at an abasic site. Studies have shown that APE1 is also active on the stable abasic site analog, tetrahydrofuran (D) (*17*). A recent study used fluorescence anisotropy to characterize binding of APE1 to a D-containing DNA (*18*). A $K_d$ value of 11 nM was obtained from these experiments. The paper also showed that APE1 binds at a 1:1 ratio to the D-containing DNA oligonucleotide. APE1 was chosen for comparison in this assay because it binds at a 1:1 ratio and it is of similar size to Δ80 AAG (37 kDa and 24 kDa, respectively) (*19*). Since it is of similar size to AAG, it would be expected to give comparable anisotropy changes when binding to a 25 base pair DNA molecule. The experiment was done with 100 nM DNA, which is far above the $K_d$ for APE1 binding to a 25 base pair oligonucleotide, which means that a titration was performed. The titration was done in the absence of magnesium, which is a required cofactor for APE1 activity. This prevented any cleavage from occurring during the experiment, allowing an equilibrium measurement. Figure 2.4 shows the titration with

D·T 25mer and APE1. It shows a simple 1:1 binding curve for this enzyme since the titration saturates when the concentration of APE1 equals that of the DNA. The final anisotropy value of ~0.1 is what is expected when a protein of this size binds to a 25mer duplex. The FAM fluorescence does not change during the experiment, which suggests that no protein is binding near the end of the DNA. Using APE1 in this titration verifies the experimental technique.

(A) (B)

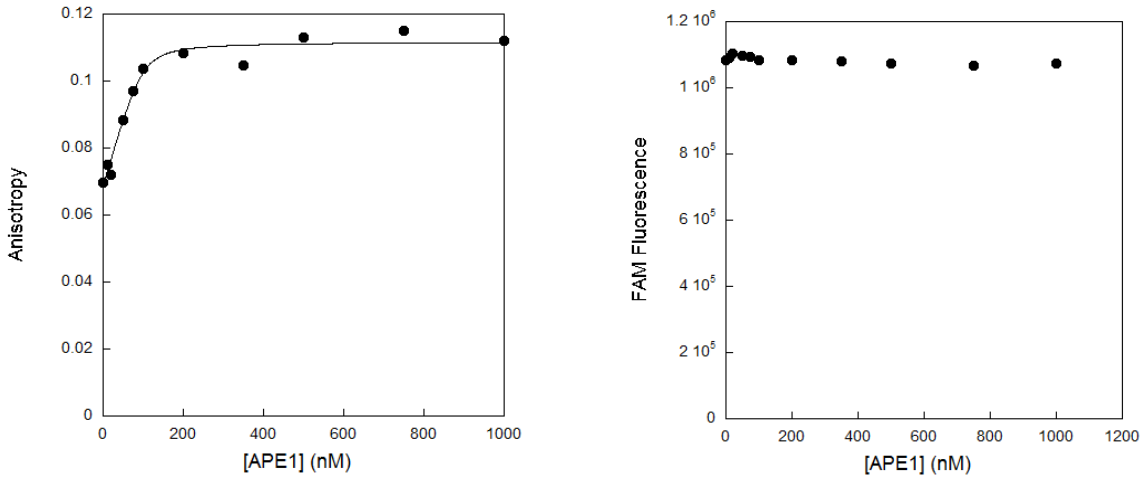



Figure 2.4. Titration of THF-containing DNA with APE1. Shown are the anisotropy (A) and the FAM fluorescence intensity (B) when 100 nM D·T 25mer DNA was titrated with APE1. APE1 shows a well behaved titration curve that represents 1:1 binding. The FAM fluorescence intensity shows no significant change throughout the titration. Data shown is from a single experiment.

**AlkA Shows Complicated Binding**

The 3-methyladenine DNA glycosylase II from *E. coli* (AlkA) also excises εA from DNA. It was used to examine binding by a different glycosylase to εA·T 25mer DNA (Figure 2.5). AlkA shows that a high concentration of protein is needed to reach a saturating anisotropy value. No change in FAM quenching is also seen up to a 1:1 ratio, while modest FAM quenching is seen at greater than 1:1 for this experiment. It will be necessary to repeat this experiment with another concentration of DNA to rule out the

possibility of this being equilibrium binding with a $K_d$ of ~500 nM. More work needs to be done to examine if AlkA can bind at high density to undamaged DNA and also to see if this high density binding is a more common phenomenon among glycosylases than previously thought.

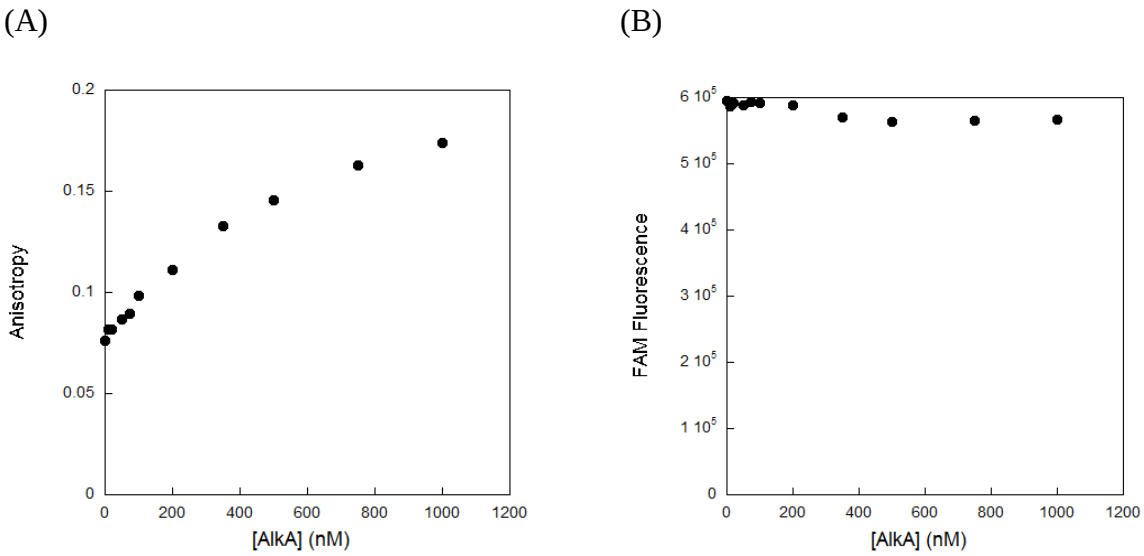



Figure 2.5. Binding of AlkA to εA-containing DNA. WT AlkA was added to 100 nM εA·T 2501 to measure binding. The measured anisotropy (A) and FAM fluorescence intensity (B) are shown. AlkA appears to bind similarly to AAG in that the anisotropy saturates when greater than 1 AlkA molecule is bound. Data shown is from a single experiment.

**AAG Binds Specifically to εA Site**

In Figure 2.6, binding of AAG to damaged (εA-containing) and undamaged DNA oligonucleotides are compared for 25mer (A and B) and 13mer DNA oligonucleotides (C and D). Both experiments show that AAG binds to damaged DNA differently than it does to undamaged DNA. For the 25mer anisotropy data (Figure 2.6A), there is a small change in anisotropy up to a 1:1 ratio of AAG: DNA with the εA·T DNA followed by a large change in anisotropy when more AAG molecules bind. However, with the A·T DNA, the binding curve shows only one phase of binding. This clearly shows that there is a difference in binding to specific and nonspecific DNA molecules. AAG binds with higher

affinity to the specific site than to the nonspecific sites since the specific sites are all occupied when AAG binds with 1:1 stoichiometry. Also, seen in the data is that the anisotropy of the nonspecifically bound complex is at a higher value than that of the specifically bound complex. The FAM fluorescence intensity for the two experiments is very similar (Figure 2.6B) but it can be seen that there is little fluorescence change up to a 1:1 ratio of AAG: DNA followed by a large change as more AAG molecules bind for the experiment with the εA·T DNA, as seen before. In contrast, the experiment with the A·T DNA shows some FAM fluorescence quenching up to a 1:1 stoichiometry.

For the anisotropy data with AAG binding to the 13mer DNA oligonucleotides (Figure 2.6C), there is not as pronounced of a difference between the binding curves. There is, however, a large difference in the FAM fluorescence seen with these different DNA molecules (Figure 2.6D). The FAM fluorescence for the binding of AAG to εA·T 13mer increases during the course of the experiment while that from AAG binding to A·T 13mer quenches slightly. While there is not an obvious explanation for this, it is possible that the proteins binding to each DNA are positioned differently with respect to the FAM on the DNA end which could cause the FAM to be in a different environment depending on the DNA sequence.

**Specific Binding Seen with εA Fluorescence**

The εA lesion is naturally fluorescent and is highly sensitive to its environment (*20*). This makes it possible to watch the εA fluorescence quench when AAG binds to the specific site in εA·T DNA. In Figure 2.7, an experiment was performed in which the εA fluorescence was compared with the anisotropy and FAM fluorescence. Both

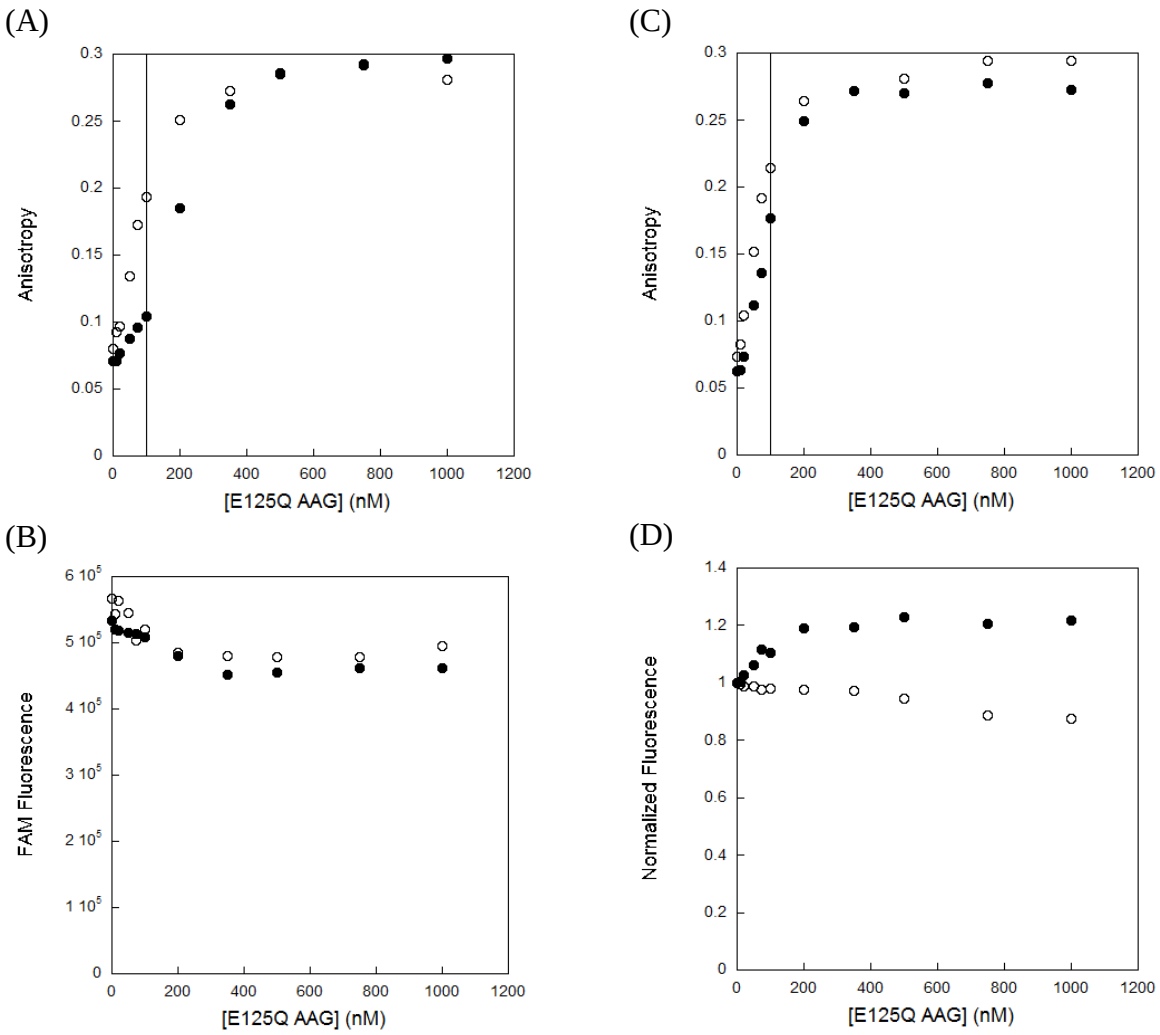



Figure 2.6. Comparison of E125Q AAG binding to damaged and undamaged DNA. Experiments were performed with E125Q AAG binding to 100 nM of either εA·T (●) or A·T (❍) DNA. The anisotropy values (A) and FAM fluorescence values (B) are shown for the 25mer oligonucleotides. The vertical line denotes the concentration of E125Q AAG where a 1:1 stoichiometry is reached. Both binding curves show similar ending anisotropy and FAM fluorescence intensity values, suggesting that a similar number of AAG molecules are binding to each oligonucleotide. A pronounced difference between the two is that the curve from εA·T 25mer shows very little increase in anisotropy up to a 1:1 ratio followed by a large increase in anisotropy when more AAG molecules bind, while the curve from A·T 25mer appears to be a single binding curve that represents multiple AAG molecules binding. This difference can be attributed to AAG having higher affinity for the εA·T site than for nonspecific DNA so that only the specific site is filled until > 1 equivalent of AAG is added. The same data are shown for the 13mer oligonucleotides as well, with anisotropy (C) and normalized FAM fluorescence (D) values. The vertical line denotes the concentration of E125Q AAG where a 1:1 stoichiometry is reached. Little difference is seen from the anisotropy data, while the FAM fluorescence shows much different behavior with an increase in fluorescence with the εA·T DNA and quenching with the A·T DNA. Data shown is from single experiments.

fluorescence curves were fit to Equation 2.3, while the anisotropy, showing two phases, was fit to a line for the first phase then the second phase was fit to Equation 2.4. The data were normalized for easier comparison. There is an obvious difference between the εA

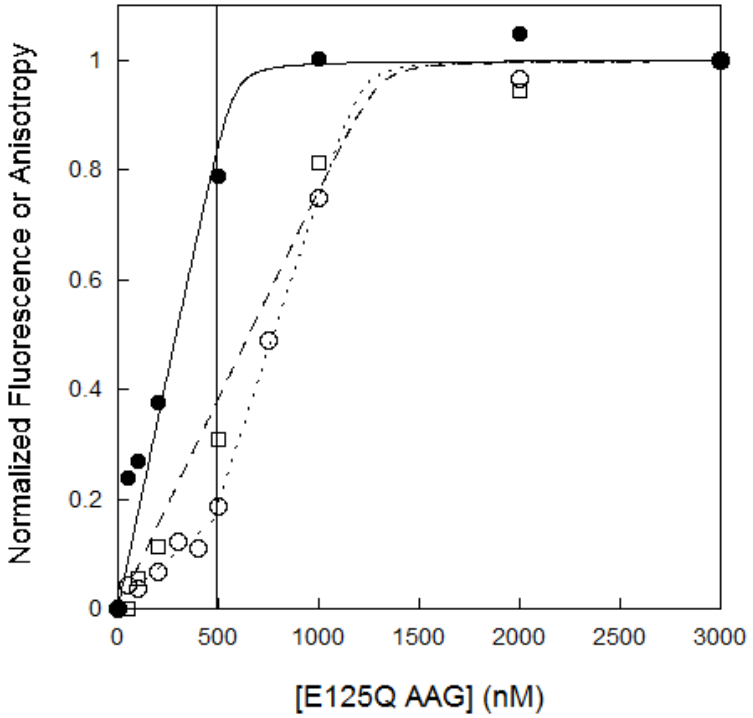



Figure 2.7. Comparison of εA and FAM fluorescence. The binding of E125Q AAG to 500 nM εA·T 25mer was followed by measuring both fluorophores on the DNA as well as the anisotropy. The fluorescence from the εA lesion (●, solid line) shows a binding curve that saturates at 1:1 [E125Q AAG]: [DNA] (shown by vertical solid line), indicating that εA fluorescence quenching is only affected by an AAG molecule binding to the specific εA site. Anisotropy (❍, dotted line) and fluorescence from FAM were also measured (□, dashed line). The anisotropy and FAM fluorescence saturate at ~ 3:1 ratio of protein to DNA, indicating that this signal does not saturate until multiple proteins are bound. Data shown is from single experiments.

fluorescence and the other two measurements. The εA fluorescence clearly shows a binding curve that saturates at a 1:1 ratio of AAG: DNA. Since εA fluorescence saturates at 1:1, it is only reporting on binding of AAG that occurs at the site of damage and is not affected by any additional AAG molecules binding nonspecifically. The anisotropy and FAM fluorescence, as seen above, saturate at a much higher density of AAG, at ~3:1

ratio. While the anisotropy also reports on 1:1 binding of AAG, it also shows large changes when more AAG molecules bind to the DNA. The FAM fluorescence is only changing when the nonspecifically bound AAG molecules bind and these proteins give a large change in that signal.

**More AAG Molecules Can Bind to Longer DNA Oligonucleotides**

Binding experiments were performed with εA·T DNA molecules of increasing length in Figure 2.8. A linear deviation to the binding curve is seen with the 13mer, 25mer and 49mer up to a 1:1 ratio of AAG: DNA, though the curves appear to be different from each other. Higher initial anisotropy values are seen for the longer length DNA oligonucleotides due to the fact that larger macromolecules tumble more slowly in solution, which leads to higher anisotropy values.

Saturation of anisotropy at similar ending values is then seen at higher concentration of AAG for longer DNA oligonucleotides, which indicates that multiple AAG molecules are binding to each DNA. As discussed above, saturation of the anisotropy signal is not necessarily indicative of saturation of the DNA by AAG molecules. This means that it is possible that more AAG molecules can bind to the longer DNA oligonucleotides that cannot be detected by anisotropy. In order to assess the number of AAG molecules binding to the DNA oligonucleotides, the FAM fluorescence data can be analyzed since it is not limited in the way that the anisotropy data is. The results of this analysis can be seen in Table 2.2, which includes data for these and several other DNA oligonucleotides. It can be seen that more AAG molecules can bind to the longer DNA oligonucleotides. This is consistent with the idea that longer DNA molecules will have more nonspecific binding sites that the protein can bind to. Even though the

number of AAG molecules bound to each DNA is different, a similar binding density is seen for all DNA oligonucleotides so that one AAG is bound every 6-10 bases.

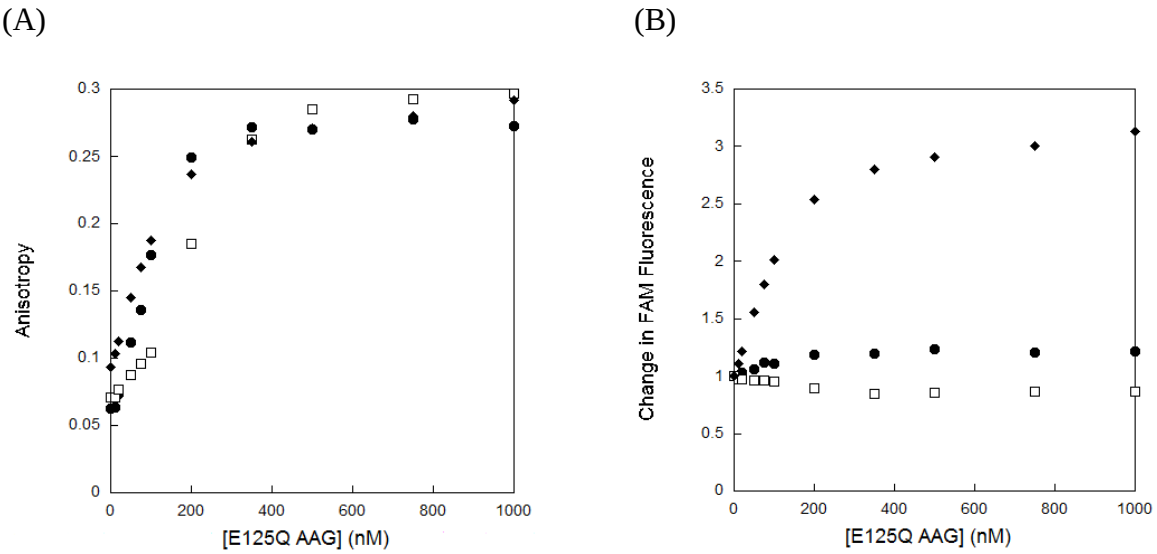



Figure 2.8. Comparison of binding by E125Q AAG to εA-containing DNA oligonucleotides of different length. The binding of E125Q AAG to 100 nM of 13mer (●), 25mer (□) and 49mer (♦) εA·T DNA was measured. (A) The initial anisotropy values are not the same since this value will increase with DNA length. All of the binding curves reach a similar final anisotropy value. Each shows a similar linear increase to a 1:1 ratio of [AAG]: [DNA], indicating the formation of a specific complex, followed by a binding curve for the nonspecifically bound proteins. Saturation is reached at increasing [AAG] as the length of DNA increases. (B) The change in FAM fluorescence is shown for comparison of the DNA fluorescence. Data shown is from single experiments.

Interestingly, the FAM fluorescence intensity for the three different length DNA oligonucleotides changes in different directions and with different magnitudes when it is fully saturated by bound AAG. The 25mer shows a small degree of FAM quenching when fully saturated, while the 13mer shows a small increase in FAM fluorescence. The FAM fluorescence from the 49mer shows a large increase in fluorescence when it is fully bound by AAG. It is not clear why the FAM fluorophores on the DNA molecules of different lengths show such different behavior. If is possible that the proteins bound near the end of the DNA have different interactions with the FAM molecule. Or there could

simply be length dependent or sequence specific differences in the way AAG binds near the end of the DNA.

Table 2.2 shows the number of AAG molecules bound for the different DNA oligonucleotides that were studied above. This gives 1-2 AAG molecules bound to the 13mer DNA oligonucleotides, 3-3.8 AAG molecules bound to the 25mer DNA oligonucleotides and 4.7 AAG molecules bound to the 49mer DNA oligonucleotide. The DNA lengths are doubled at each interval and the number of AAG molecules bound also doubles between the 13mer and 25mer DNA. The number of AAG molecules is less than doubled between the 25mer and 49mer DNA. This could be due to some interaction of

Table 2.2. Comparison of AAG molecules bound to DNA of different lengths.

| DNA | Number of AAG Molecules Bound[a] | Ratio of AAG : DNA Base Pairs[b] |
|---|---|---|
| εA•T 49mer | 4.7 | 1:10 |
| | | |
| A•T 25mer | 3.8 | 1:6 |
| εA•T 25mer | 3.5 | 1:7 |
| D•T 25mer | 3.0 | 1:8 |
| | | |
| A•T 13mer | 0.9 | 1:13 |
| εA•T 13mer | 2.2 | 1:6 |
| D•T 13mer | 1.9 | 1:6 |

[a] The number of AAG molecules bound to one DNA oligonucleotide was obtained by determining the concentration of protein where the FAM fluorescence intensity saturated, which was then divided by the concentration of DNA in the experiment. [b] The ratio of AAG: DNA base pairs was determined by dividing the number of base pairs in the DNA by the number of AAG molecules bound.

AAG with the ends of the 25mer DNA oligonucleotide that requires the AAG molecules to pack more tightly along the DNA, while the 49mer, with greater binding sites, does not require such tight packing of AAG molecules while still achieving high density binding.

Also seen in Table 2.2 is the ratio of AAG: DNA base pairs, which is between 6-8 base pairs for the 25mer DNA oligonucleotides. This appears to be an appropriate estimate for the footprint of AAG binding.

**AAG Footprint Comparison**

This remarkably high binding density appears to be in conflict with the known structural information. Shown above is evidence that the footprint for AAG binding to DNA is 6-8 base pairs. However, the footprint for AAG binding to DNA is predicted to be 10 base pairs from the crystal structure (*21*) and 11-12 base pairs from footprinting assays (*22*). The only way to reconcile these numbers is for the AAG binding sites to be overlapping. This is a feasible model because the crystal structure shows that AAG binds to only one face of the DNA. Thus, two AAG binding sites can overlap if the proteins bind to opposite faces of the DNA. Since the binding sites overlap, this gives AAG binding at a higher density to DNA than was predicted by other footprinting assays. The high density binding by AAG suggests that AAG can sample very small segments of accessible DNA. The data from the 13mer DNA oligonucleotides is the best evidence that AAG is able to overlap binding sites in this way since two AAG molecules bind to a 13mer DNA, giving a footprint of ~6 bases per AAG, which is obviously smaller than the footprint for AAG from the traditional assays.

Additional evidence that AAG binds at high density to DNA is a series of electrophoretic mobility shift assays (EMSA) that have been done with AAG and multiple 25mer DNA oligonucleotides (*10, 23-24*). These assays show a shift in the mobility on a gel between unbound and bound AAG. In many of these assays, an additional band or two can be seen above the bound AAG band, which is consistent with complexes containing multiple molecules of AAG.

**AAG Binds at High Density to Single-Stranded DNA**

AAG is also capable of binding to single-stranded DNA (ssDNA). Binding experiments were performed with AAG and εA-containing and undamaged ssDNA molecules in Figure 2.9, which shows that AAG binds at high density to both. Although saturation was not yet reached in the experiment, it required a very high concentration of AAG (8-10:1 AAG: DNA) to see the anisotropy changes. This means that AAG can bind at a much higher density to ssDNA than to duplex DNA (Table 2.3). This very high density of AAG binding to ssDNA is very similar to the binding density of another DNA repair protein. Recently the crystal structure of RecA, an enzyme involved in homologous repair was solved and it shows that one RecA monomer binds to every three bases of a single-stranded DNA (*25*).

Table 2.3 shows a comparison between the footprint for single- and double-stranded DNA. Clearly, AAG binds with much smaller footprint to single-stranded DNA only binding to 3 bases for each molecule, compared with 6-8 bases per molecule for double-stranded DNA. Perhaps AAG is able to extend the single-stranded DNA so that it can fit one molecule for every 3 bases, since it is unlikely that AAG binding sites can

overlap to the extent they do on duplex DNA because there is only a single strand that AAG can bind to.

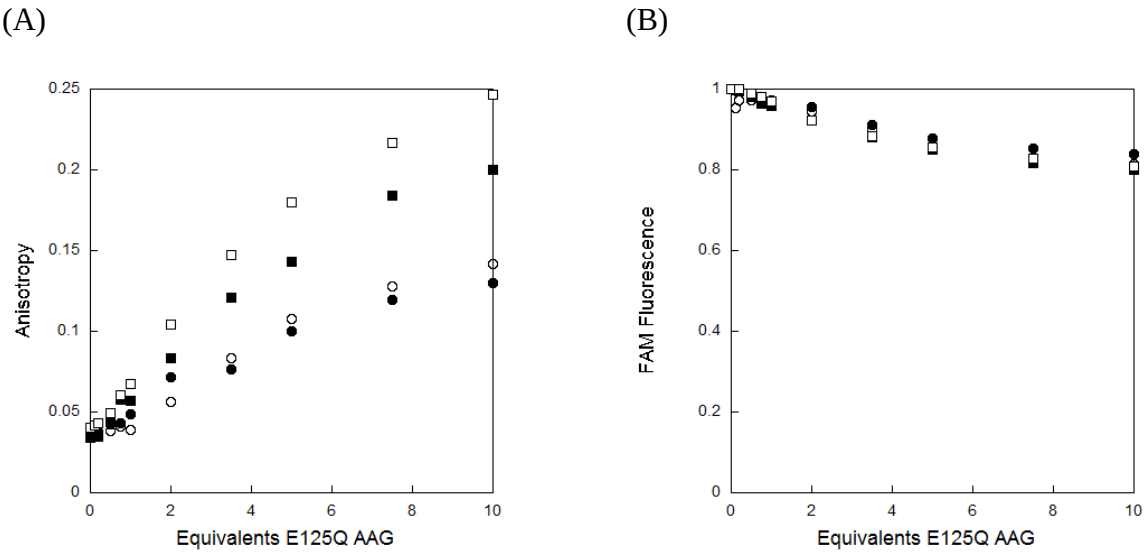



Figure 2.9. E125Q AAG binding to single stranded DNA oligonucleotides. Experiments were performed with both damaged and undamaged single stranded DNA. Both 10 nM (●) and 100 nM (❍) ss εA 25mer as well as 10 nM (■) and 100 nM (□) ss T 25mer were mixed with E125Q AAG. The x-axis was converted to equivalents of E125Q AAG that were added by dividing [AAG] by [DNA] used in the experiment. This allows for an easier comparison between the different experiments. (A) The anisotropy shows a similar behavior for both of the single-stranded DNA oligonucleotides. Much greater concentrations of AAG are required to saturate the signal than is seen with duplex DNA. This indicates that AAG binds at a much higher density to single-stranded DNA oligonucleotides than duplex oligonucleotides. (B) The normalized FAM fluorescence intensity is also shown. Data shown is from single experiments.

Table 2.3. Comparison of the footprint of AAG in single- and double-stranded DNA.

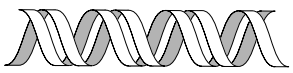
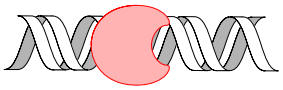
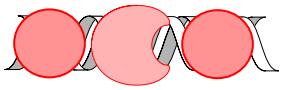
| DNA Type | Bases per AAG Molecule[a] |
|---|---|
| Double-Stranded | ≤ 6-8 |
| Single-Stranded | ≤ 3 |

[a] These values were obtained by dividing the number of bases in the DNA oligonucleotide by the number of AAG molecules that are bound to that DNA. They are upper limits because the number of AAG molecules that we get from the experiments are only lower limits.

### Model for AAG Binding

Our findings are summarized in Table 2.4. When DNA is free in solution (no AAG molecules bound) there is a low level of εA fluorescence quenching, no FAM quenching and low anisotropy values. When a single AAG is bound to an εA containing DNA, it is bound at the specific site which leads to high εA fluorescence quenching, no FAM quenching and an intermediate anisotropy value. As nonspecifically bound AAG molecules bind to the DNA (> 1 AAG bound), there is no further quenching of εA fluorescence. FAM fluorescence shows a high level of quenching and the anisotropy is also high.

Table 2.4. Model for multiple AAG binding.

| |  |  |  |
|---|---|---|---|
| # AAG molecules bound | 0 | 1 | >1 |
| εA quenching | low | high | no further |
| FAM quenching | none | none | high |
| anisotropy | low | intermediate | high |

## Implications

### Complications of Multiple Protein Binding on the Binding Assay

One of the major implications of the findings presented above is that DNA binding proteins can bind multiple molecules to short DNA oligonucleotides. Proteins that behave in this manner make it difficult to distinguish between specific and

nonspecific binding using an assay such as fluorescence anisotropy. Fluorescence anisotropy measurements are still very useful for measuring binding to DNA when the protein binds in a 1:1 ratio to that DNA. It is important to know if a protein binds in a 1:1 ratio or >1:1 ratio to DNA before trying to study binding by using fluorescence anisotropy. Interestingly, some specificity is seen with AAG binding to εA·T DNA, which means that equilibrium binding can be measured at low concentrations of DNA. In this case, however, fluorescence anisotropy is not sensitive enough for these measurements.

**High Density Binding of Other DNA Repair Proteins**

Since DNA repair enzymes are expected to be able to bind nonspecifically to DNA, this high density binding may be a more common phenomenon than previously thought. If this is the case, there could be a more general problem with this method for other DNA binding proteins. Already shown above are two DNA repair enzymes that bind at high density to DNA, AAG and AlkA. There is evidence in the literature that other DNA binding proteins exhibit similar behavior. AlkD, another glycosylase, has been shown to take high concentrations (~10:1 AlkD: 25mer DNA) of protein to saturate an anisotropy signal (*26*). This indicates, as seen in the experiments above, that multiple AlkD proteins are binding to a single DNA. Another example is $O^6$-alkylguanine DNA alkyltransferase (AGT), another DNA repair protein, has been shown to bind with high density (every ~4 base pairs) to DNA with analytical ultracentrifugation techniques (*27*). Probably the most well known example of DNA repair proteins binding at high density to DNA is the RecA family of proteins (including Rad51). These proteins coat both single-

and double-stranded DNA, which helps it perform its role in homologous recombination (*25*).

**Biological Implications of High Density Binding by AAG**

It is unclear whether the high density of AAG binding to DNA has any physiological significance. A single AAG molecule can perform its glycosylase activity, as seen in multiple turnover assays (*5, 13, 28*). Also, single turnover kinetics show that the specifically bound proteins are not affected by nonspecifically bound proteins. This is shown by single turnover experiments done at 1:1 AAG: DNA having exactly the same rate constant as those done with much greater than 1:1 AAG: DNA (*5, 10, 13, 28*). It is possible that AAG may have some other biological function, such as protecting DNA or regulating transcription, which relies on high density binding to DNA. More work is required to determine if there is any other function of AAG within the cell.

# References


1. Carey, D. C., and Strauss, P. R. (1999) Human apurinic/apyrimidinic endonuclease is processive, *Biochemistry 38*, 16553-16560.

2. Higley, M., and Lloyd, R. S. (1993) Processivity of uracil DNA glycosylase, *Mutat Res 294*, 109-116.

3. Francis, A. W., and David, S. S. (2003) Escherichia coli MutY and Fpg utilize a processive mechanism for target location, *Biochemistry 42*, 801-810.

4. Porecha, R. H., and Stivers, J. T. (2008) Uracil DNA glycosylase uses DNA hopping and short-range sliding to trap extrahelical uracils, *Proc Natl Acad Sci U S A 105*, 10791-10796.

5. Hedglin, M., and O'Brien, P. J. (2008) Human alkyladenine DNA glycosylase employs a processive search for DNA damage, *Biochemistry 47*, 11434-11445.

6. Heyduk, T., Ma, Y., Tang, H., and Ebright, R. H. (1996) Fluorescence anisotropy: rapid, quantitative assay for protein-DNA and protein-protein interaction, *Methods Enzymol 274*, 492-503.

7. LeTilly, V., and Royer, C. A. (2002) Fluorescence anisotropy assays implicate protein-protein interactions in regulating trp repressor DNA binding, *Biochemistry 32*, 7753-7758.

8. Jameson, D. M., and Sawyer, W. H. (1995) Fluorescence anisotropy applied to biomolecular interactions, *Methods Enzymol 246*, 283-300.

9. O'Brien, P. J., and Ellenberger, T. (2003) Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines, *Biochemistry 42*, 12418-12429.

10. Abner, C. W., Lau, A. Y., Ellenberger, T., and Bloom, L. B. (2001) Base Excision and DNA Binding Activities of Human Alkyladenine DNA Glycosylase Are Sensitive to the Base Paired with a Lesion, *Journal of Biological Chemistry 276*, 13379-13387.

11. O'Brien, P. J., and Ellenberger, T. (2004) The Escherichia coli 3-methyladenine DNA glycosylase AlkA has a remarkably versatile active site, *J Biol Chem 279*, 26876-26884.

12. Labahn, J., Scharer, O. D., Long, A., Ezaz-Nikpay, K., Verdine, G. L., and Ellenberger, T. E. (1996) Structural basis for the excision repair of alkylation-damaged DNA, *Cell 86*, 321-329.

13. Baldwin, M. R., and O'Brien, P. J. (2009) Human AP endonuclease 1 stimulates multiple-turnover base excision by alkyladenine DNA glycosylase, *Biochemistry 48*, 6022-6033.

14. Secrist, J. A., 3rd, Barrio, J. R., Leonard, N. J., and Weber, G. (1972) Fluorescent modification of adenosine-containing coenzymes. Biological activities and spectroscopic properties, *Biochemistry 11*, 3499-3506.

15. Kartalou, M., Samson, L. D., and Essigmann, J. M. (2000) Cisplatin adducts inhibit 1,N(6)-ethenoadenine repair by interacting with the human 3-methyladenine DNA glycosylase, *Biochemistry 39*, 8032-8038.

16. Klonis, N., and Sawyer, W. H. (1996) Spectral properties of the prototropic forms of fluorescein in aqueous solution, *Journal of Fluorescence 6*, 147-157.

17. Erzberger, J. P., and Wilson Iii, D. M. (1999) The role of Mg2+ and specific amino acid residues in the catalytic reaction of the major human abasic endonuclease: new insights from EDTA-resistant incision of acyclic abasic site analogs and site-directed mutagenesis, *Journal of Molecular Biology 290*, 447-457.

18. Maher, R. L., and Bloom, L. B. (2007) Pre-steady-state kinetic characterization of the AP endonuclease activity of human AP endonuclease 1, *J Biol Chem 282*, 30577-30585.

19. Demple, B., Herman, T., and Chen, D. S. (1991) Cloning and expression of APE, the cDNA encoding the major human apurinic endonuclease: definition of a family of DNA repair enzymes, *Proc Natl Acad Sci U S A 88*, 11450-11454.

20. Leonard, N. J. (1984) Etheno-substituted nucleotides and coenzymes: fluorescence and biological activity, *CRC Crit Rev Biochem 15*, 125-199.

21. Lau, A. Y., Wyatt, M. D., Glassner, B. J., Samson, L. D., and Ellenberger, T. (2000) Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG, *Proc Natl Acad Sci U S A 97*, 13573-13578.

22. Miao, F., Bouziane, M., and O'Connor, T. (1998) Interaction of the recombinant human methylpurine-DNA glycosylase (MPG protein) with oligodeoxyribonucleotides containing either hypoxanthine or abasic sites, *Nucl. Acids Res. 26*, 4034-4041.

23. Vallur, A. C., Feller, J. A., Abner, C. W., Tran, R. K., and Bloom, L. B. (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*, 31673-31678.

24. Vallur, A. C., Maher, R. L., and Bloom, L. B. (2005) The efficiency of hypoxanthine excision by alkyladenine DNA glycosylase is altered by changes in nearest neighbor bases, *DNA Repair (Amst) 4*, 1088-1098.

25. Chen, Z., Yang, H., and Pavletich, N. P. (2008) Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures, *Nature 453*, 489-484.

26. Rubinson, E. H., Metz, A. H., O'Quin, J., and Eichman, B. F. (2008) A new protein architecture for processing alkylation damaged DNA: the crystal structure of DNA glycosylase AlkD, *J Mol Biol 381*, 13-23.

27. Melikishvili, M., Rasimas, J. J., Pegg, A. E., and Fried, M. G. (2008) Interactions of human O(6)-alkylguanine-DNA alkyltransferase (AGT) with short double-stranded DNAs, *Biochemistry 47*, 13754-13763.

28. O'Brien, P. J., and Ellenberger, T. (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase, *J Biol Chem 279*, 9750-9757.

# Chapter 3

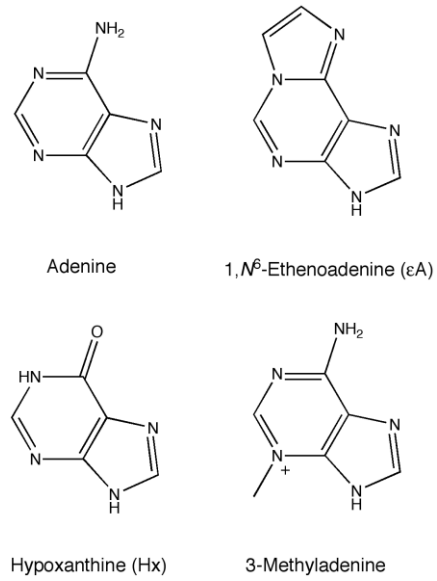
## Kinetic Mechanism for the Flipping and Excision of 1, *N6*-Ethenoadenine by Human Alkyladenine DNA Glycosylase[1]

The nucleobases of DNA are susceptible to spontaneous chemical modification, including oxidation and alkylation. Failure to repair the resulting DNA lesions can result in mutations or in some cases cell death. The majority of simple base lesions are repaired via the base excision repair (BER) pathway (*1*). This pathway is initiated by a DNA repair glycosylase that is responsible for finding the damaged site and catalyzing the hydrolysis of the N-glycosidic bond. In the case of a monofunctional DNA glycosylase, this results in the release of the lesioned base and the formation of an abasic site in the DNA. The abasic site is further processed to nick the DNA backbone and remove the abasic sugar residue. Repair synthesis uses the intact strand as a template to ensure incorporation of the correct nucleotide and the nick is ultimately sealed by DNA ligase.

The human enzyme, alkyladenine DNA glycosylase (AAG), is responsible for repairing a diverse set of alkylated and oxidized purines, including purines alkylated at the N7 and N3 positions, and the deaminated purines hypoxanthine (Hx) and oxanine [(*2*, *3*) and references therein]. Chemical structures of damaged and undamaged adenine bases are shown in Scheme 3.1. AAG also recognizes the bulkier alkylation adduct, 1,$N^6$-

[1] Reproduced with permission from Wolfe, A. E., and O'Brien, P. J. (2009) Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by human alkyladenine DNA glycosylase, *Biochemistry 48*, 11357-11369. Copyright 2009 American Chemical Society

ethenoadenosine (εA) that can be formed from lipid peroxidation products and exposure to exogenous agents such as urethane and vinyl chloride (*4, 5*). This lesion is highly mutagenic in mammalian cells and leads to misincorporation of all possible nucleotides if it is replicated prior to repair (*6, 7*).




Scheme 3.1

AAG is the only known human glycosylase that catalyzes the excision of εA, but an alternative direct reversal pathway has recently been discovered in which εA can be directly oxidized to restore the adenine base by AlkB family members (*8*). This backup pathway may account for the observation that mice lacking AAG remain viable (*9, 10*). Nevertheless, mice lacking AAG have decreased rates of repair of εA and are unable to tolerate the burden of increased levels of DNA alkylation (*11, 12*).

Structural and biochemical studies have provided considerable insight into the mechanism by which AAG recognizes and excises damaged bases and have suggested the minimal mechanism that is illustrated in Figure 3.1. Given the low frequency with

which sites of DNA damage occur, initial binding is most likely to occur at an undamaged site. AAG uses nonspecific binding interactions to scan DNA in the search for sites of damage (*13*). Once a damaged nucleotide is encountered, it must be flipped out of the duplex by 180° into the enzyme active site where lesion recognition and N-glycosidic bond hydrolysis can occur (*14-16*). The excised base is released and the abasic nucleotide must disengage from the active site in the reverse of the flipping step. The nonspecifically bound AAG is able to diffuse away along the duplex in search of additional sites of damage and eventually can be released into solution. The rates of DNA binding and nucleotide flipping have not previously been measured for AAG, but are critical to understanding how the enzyme recognizes a wide variety of damaged nucleotides.

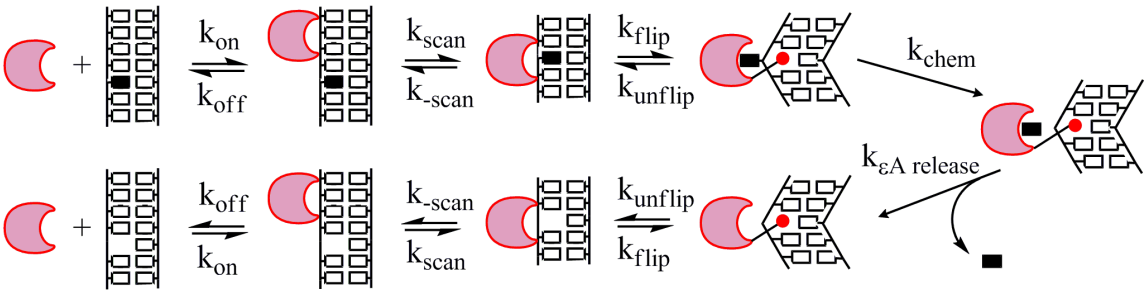



Figure 3.1. Minimal mechanism for AAG-catalyzed nucleotide flipping and base excision. AAG (crescent) binds nonspecifically to DNA and rapidly slides in search of DNA damage. Once a lesion (filled rectangle) is encountered in an initial damage recognition complex it can be flipped out into the active site to form the specific recognition complex. In this specific complex the DNA is bent, the damaged nucleotide is flipped out by 180°, and Y162 (red circle) is intercalated into the DNA to take the place of the extrahelical lesion (*15, 16*). From this specific complex, AAG catalyzes hydrolysis of the N-glycosidic bond ($k_{chem}$). The lesion base is subsequently released ($k_{\varepsilon A\ release}$). The abasic product is an inhibitor of AAG, so presumably the sugar can be stably bound in a flipped out conformation analogous to the abasic pyrrolidine conformation observed by x-ray crystallography (*15*). Dissociation of AAG from the 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 of additional sites of damage and can ultimately dissociate ($k_{off}$).

Several previous studies have used intrinsic protein fluorescence or the incorporation of fluorescent reporter groups into DNA in order to investigate the kinetics of DNA binding and nucleotide flipping by other families of DNA glycosylases (*17-22*). Together with studies of other DNA modifying enzymes, these studies have begun to address the fundamental, common features of nucleotide flipping enzymes (*14*).

There is, however, circumstantial evidence that the requirement for nucleotide flipping provides a barrier to the excision of damaged bases by AAG (*2, 23, 24*). For example, the rate of excision of a given lesion paired with different opposing bases is inversely proportional to the base pairing stability [e.g., 7-methylguanine (7mG) is excised from a 7mG•T mismatch 50-fold faster than it is removed from a stable 7mG•C pair]. Since AAG does not make specific contacts with the opposing base, this observation is best explained by an AAG complex in which the lesion can still hydrogen bond with the opposing base (*2*). In addition, rates of repair vary widely for a given lesion base pair in different local sequence contexts (*25*). The effect of sequence context is most pronounced for excision of Hx (*25, 26*). For example, rates of excision of Hx from different regions of a polyA tract vary by almost two orders of magnitude (*26*). An early report mentioned that excision of εA is also sensitive to its position in a polyA tract (*27*), but subsequent studies found little or no sequence context effects for the excision of εA by AAG (*25, 26*). We were interested in learning whether nucleotide flipping is rapid, as has been observed for other DNA glycosylases (*17, 19-22*), or whether the flipping step might be significantly slower than the base excision step.

Since εA is known to be fluorescent and very sensitive to its local environment (*28*), we investigated the possibility of directly monitoring the dynamics for DNA

binding and nucleotide flipping for this biologically relevant DNA adduct. We combined single turnover and multiple turnover glycosylase assays with stopped-flow fluorescence measurements to characterize the kinetic mechanism for AAG-catalyzed excision of εA. We show that AAG binding to the DNA is fast compared to the other steps and that binding is reversible. We present evidence that nucleotide flipping is rapid and highly favorable and that the excised lesion is rapidly released. These fast conformational changes, with relatively slow base excision are integral to the ability of AAG to discriminate between damaged and undamaged nucleotides.

## Materials and Methods

### Purification of Human Recombinant Protein

The catalytic domain of human AAG that lacks the first 79 amino acids was expressed in *E. coli* and purified as previously described (*29*). The concentration of active AAG was determined by burst analysis, using an inosine-containing oligonucleotide substrate, as previously described (*13*, *30*).

### Synthesis and Purification of Oligodeoxynucleotides

The 25mer oligonucleotides or the constituent phosphoramidites were obtained from commercial sources and contained a central lesion that was placed opposite of a T. The sequences are given in Scheme 3.2. Oligonucleotides containing only normal deoxynucleotides or deoxyinosine were synthesized and deprotected with standard phosphoramidite chemistry, whereas those containing a single εA lesion were synthesized and deprotected using Ultra-Mild chemistry according to the supplier's

recommendations. The εA-containing oligonucleotides were either synthesized by the W.M. Keck Facility at Yale University or the phosphoramidite was obtained from Glen Research and incorporated using an ABI 394 DNA synthesizer. The oligonucleotides

TEC 5′-CGATAGCATCCT**E**CCTTCTCTCCAT
3′-GCTATCGTAGGATGGAAGAGAGGTA

AEA 5′-CGATAGCATCAA**E**AATTCTCTCCAT
3′-GCTATCGTAGTTTTTAAGAGAGGTA

Scheme 3.2

were desalted using Sephadex G-25 and purified using denaturing polyacrylamide gel electrophoresis as previously described (*13*). Oligonucleotides for gel-based assays were labeled on the lesion containing strands 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 εA containing oligonucleotides, 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 (*31*). The lesion-containing oligonucleotides were annealed with a 1.1-fold excess of the complement by heating to 90°C and cooling slowly to 4°C.

**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 (200 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 1 minute. No significant excision of εA occurs during this time.

**Single Turnover Excision of εA Monitored by Fluorescence**

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 410 nm; therefore this wavelength was chosen to follow the release of εA into solution. The fluorescence at 410 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 3.1). The single turnover rate constant was determined at saturating concentration of AAG and this was confirmed by the observation of the same rate constant at two different concentrations of AAG (Figure 3.2). Under these conditions, the observed rate constant is equal to the maximal single turnover rate constant ($k_{obs} = k_{max}$).

$$\text{Fraction product} = A(1 - e^{-k_{obs}t}) \qquad \text{(Equation 3.1)}$$

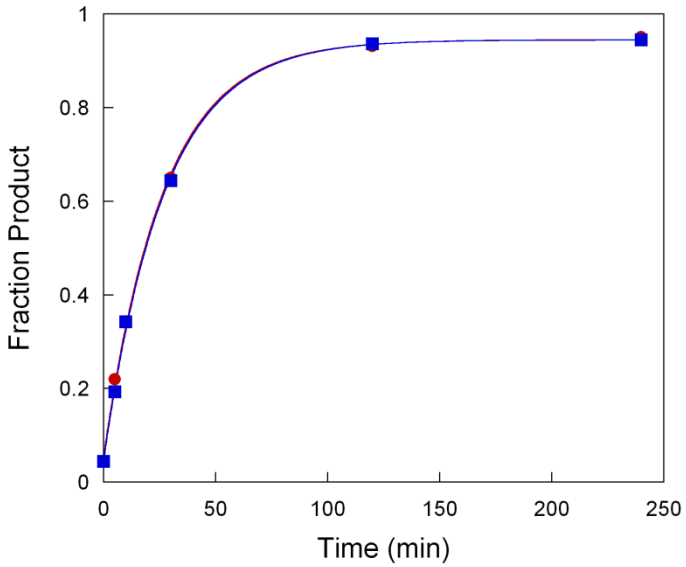



Figure 3.2. Confirmation that the maximal single turnover excision rate was determined. Identical rate constants were obtained ($k_{max}$ = 0.037 $min^{-1}$) for reaction with –TEC- (100 nM) and either 200 nM (red circles) or 400 nM (blue squares) AAG. The fraction abasic DNA product was determined from the gel-based glycosylase assay and the lines indicate the best fit of a single exponential equation (Equation 3.1).

**Glycosylase Activity 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 PAGE gel (*13, 29*). Fluorescein labeled DNA substrates (20-100 nM) were prepared in the standard buffer with 0.1 mg/mL BSA added. The reactions were initiated with the addition of 100-400 nM AAG and incubated at 25°C. At various time points, a sample from the reaction was removed and quenched in 2 equivalents 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 and heated 3 minutes at 70°C before loading onto a 12-

15% polyacrylamide gel. Gels were scanned with a Typhoon Imager (GE 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 3.1). The observed rate constant for the single turnover reaction was independent of the concentration of AAG, indicating that the maximal rate constant was measured ($k_{obs} = k_{max}$; Figure 3.2).

**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 320 nm and a WG360 long-pass emission filter. The concentrations that are given for enzyme and substrate are those obtained after mixing equal volumes in the stopped-flow. The reactions were followed for either 2 or 10 seconds to monitor initial binding events occurring before the base excision step could occur. Binding of AAG to εA containing DNA was measured by mixing a constant concentration of DNA with increasing concentrations of AAG. Experiments were performed with either 0.15 μM or 0.5 μM concentration of DNA after mixing, and final concentrations of AAG up to 8 μM. Three traces were averaged together at each concentration, which revealed two phases. The traces were fit by a double exponential (Equation 3.2), in which F is εA fluorescence as a function of time, C is the fluorescence of the free DNA, Y is the change in fluorescence of the first intermediate, and Z is the change in fluorescence of the second intermediate. The observed rate constants for the first and second steps are designated $k_{1, obs}$ and $k_{2, obs}$ (Equation 3.2).

$$F = C - Y(1 - e^{-k_{1,obs}t}) - Z(1 - e^{-k_{2,obs}t}) \quad \text{(Equation 3.2)}$$

It is possible to relate these observed rate constants to the microscopic rate constants in Scheme 3.3, with a bimolecular binding step followed by a unimolecular conformational change that occurs prior to the chemical step.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E \cdot S \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E \cdot S'$$

Scheme 3.3

The initial binding step appears to be reversible (see below), and therefore the observed rate constant is given by the sum of the association and dissociation steps in Scheme 3.3 ($k_{1,\,obs}$ = [AAG] k1 + $k_{-1}$). It should be noted that this equation applies when [E]>>[S]. Under our conditions the enzyme was in only modest excess (2–7-fold). Nevertheless, the resulting values of $k_1$ are in reasonably good agreement with the values measured by mixing equimolar solutions of enzyme and substrate (see below). We suggest that the unimolecular step corresponds to nucleotide flipping. Since this is an approach to equilibrium, the observed rate constant is equal to the sum of the rate constants for the forward and reverse flipping steps ($k_{2,\,obs} = k_2 + k_{-2}$). Substrate dissociation experiments, measured by pulse-chase suggest that the value of $k_2$ is much greater than $k_{-2}$, indicating that $k_{2,\,obs}$ is approximately equal to $k_2$, the microscopic rate constant for flipping.

Rapid binding and weak fluorescence made it difficult to fully analyze the binding step under conditions of true excess of enzyme over substrate, therefore we also

performed mixing experiments in which a 1:1 ratio of [DNA]: [AAG] was maintained (0.15-1.5 μM). Three traces were averaged together at each concentration and they were fit by equation 3.3, where F is εA fluorescence as a function of time, C is the fluorescence of the free DNA, Y and Z are the change 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, $k_{2,\,obs}$ is the observed unimolecular rate constant for the subsequent flipping step, and t is the time. The expression that corresponds to the initial binding step in equation 3.3 has been previously described (*32*), and the origin of this equation is provided in the Appendix.

$$F = C + Y\left(\frac{E_o^2 k_1 t}{1 + E_o k_1 t}\right) - Z(1 - e^{-k_{2,obs}t}) \qquad \text{(Equation 3.3)}$$

The fits to equation 3.3 were excellent in all cases and the observed rate constants that were obtained were in good agreement with the values of $k_1$ and $k_{2,\,obs}$ that were obtained from the binding experiments in which AAG was in excess over DNA.

**Pulse-Chase Assay**

To measure the rate constant for AAG dissociating from the εA-containing DNA substrate, pulse-chase experiments were performed. These assays were done at 25°C in the standard buffer, except that the concentration of NaCl was varied and 0.1 mg/mL BSA was added. In 20 μL reactions, 100 nM fluorescein-labeled –TEC– DNA was mixed with 200 nM AAG for 20 seconds. A chase of 20 μM unlabeled –TEC– DNA was then mixed in. 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. Samples were taken at the specified times and quenched as described above. The samples were run on sequencing gels and the fraction product was calculated as described above. It was expected that the production of product would follow a simple exponential (Equation 3.1), but under the conditions employed a very slow steady-state reaction was also observed. This was accounted for by Equation 3.4, in which the burst phase gives the amplitude (A) and an observed rate constant ($k_{obs}$). The steady-state phase is dependent upon the steady-state rate constant ($k_{cat}$), the concentration of enzyme (E), and the total concentration of labeled and unlabeled DNA substrate.

$$\text{Fraction product} = A(1 - e^{-k_{obs}t}) + (k_{cat}[E]/[DNA]_{total})t \qquad \text{(Equation 3.4)}$$

According to the two step binding mechanism described in Scheme 3.3, two different partitioning equations can be written (*33*). Since all labeled substrate is initially bound, the fraction of product formed is given by fraction that goes on to react. This is indicated by Equation 3.5, in which A is the burst amplitude (the fraction of product cleaved 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 stable flipped out complex. This expression can be rearranged to solve for the desired dissociation rate constant (Equation 3.6). 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 3.7). Solving for $k_{off, obs}$ give Equation 3.8.

$$A = \frac{k_{max}}{k_{off,obs} + k_{max}} \quad \text{(Equation 3.5)}$$

$$k_{off,obs} = \frac{k_{max}}{A} - k_{max} \quad \text{(Equation 3.6)}$$

$$k_{obs} = k_{off,obs} + k_{max} \quad \text{(Equation 3.7)}$$

$$k_{off,obs} = k_{obs} - k_{max} \quad \text{(Equation 3.8)}$$

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 3.6 and 3.8). Both methods gave essentially identical 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 ($k_{-2}$) or dissociation from nonspecific DNA ($k_{-1}$). According to Scheme 3.3, and assuming that the flipped out complex is stable (i.e., $k_2 >> k_{-2}$), this observed dissociation rate can be expressed in terms of the microscopic rate constants [Equation 3.9; (*33*)]. Stopped-flow fluorescence suggests that dissociation from the initial AAG•DNA complex is rapid ($k_{-1}$ ~30 $s^{-1}$) relative to the forward flipping rate constant $k_2 = 2.6\ s^{-1}$, 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.

$$k_{off,obs} = k_{-2}\left(\frac{k_{-1}}{k_{-1} + k_2}\right) \quad \text{(Equation 3.9)}$$

The pulse chase experiment was performed at high and low ionic strength by varying the concentration of sodium chloride from 50-300 mM. It has previously been shown that the rate constant for dissociation of the abasic DNA product increases as ionic strength is increased, presumably due to the disruption of electrostatic interactions at higher ionic strength (*13*, *34*).

**Steady-State Excision of Hypoxanthine**

The maximal rate constant for excision of Hx ($k_{cat}$) was determined for a 25 base pair oligonucleotide of the same sequence as the –TEC– DNA, but with deoxyinosine in place of the εA. Under multiple turnover conditions, the rate-limiting step is dissociation of the abasic DNA [$k_{cat} = k_{off}$ for abasic DNA; (*34*, *35*)]. The experiments were performed at 25°C and in the standard buffer that also contained 0.1 mg/mL BSA. In 20 μL reactions, 2 μM fluorescein labeled DNA was mixed with 20-50 nM AAG. Aliquots were removed at specific times and quenched as described above. These samples were diluted 10-fold into formamide/EDTA buffer and run on sequencing gels and quantified as described above. The fraction product was converted into the concentration of product ([P] = fraction product × [DNA]). The initial velocity ($V_i$) was determined from a linear fit to the time-dependent formation of abasic product over the first 15% of the reaction. The value of $k_{cat}$ was calculated by dividing the velocity by the concentration of enzyme ($k_{cat} = V_i / [E]$).

# Results

## Fluorescence of εA is a Sensitive Reporter of the Lesion Environment

The feasibility of using the intrinsic fluorescence of εA to monitor the individual steps of the AAG-catalyzed reaction was evaluated for two different 25 base pair oligonucleotides that each contained a central εA lesion. These two sequences will be referred to by the central sequence of the lesion-containing strand (Scheme 3.2). Excitation of both single stranded oligonucleotides at 320 nm gave a broad fluorescence emission peak centered at 410 nm (Figure 3.3A). The oligonucleotide of sequence –TEC- gave ~4-fold lower fluorescence than free εA, consistent with quenching of εA fluorescence even in single-strand DNA. Surprisingly, the oligonucleotide of sequence –AEA- yielded much higher fluorescence, essentially identical to that of free εA. When a 1.1-fold excess of the 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- (Table 3.1). This is consistent with increased quenching due to enhanced base stacking interactions in duplex DNA. Addition of a higher concentration of complement did not further quench the fluorescence (data not shown).

It is not clear why these two sequence contexts show such different quenching, but it is possible that the lack of nearby G nucleotides can account for the high fluorescence of –AEA-. 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 (*36*). The trivial possibility that the

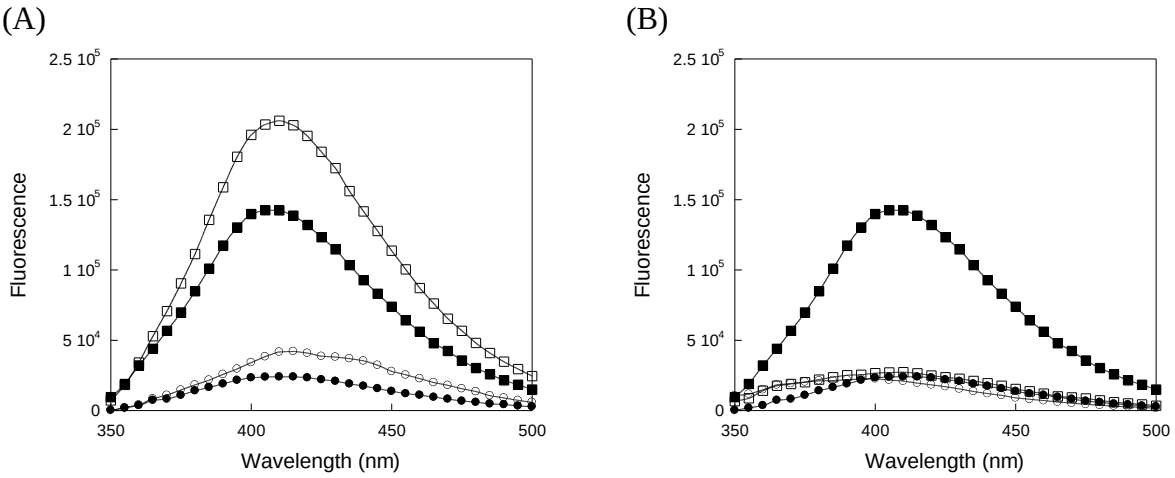



Figure 3.3. Fluorescence of εA-containing DNA in single-strand, duplex, and in complex with AAG. Fluorescence emission spectra of 400 nM DNA were collected with excitation at 320 nm, as described in the Materials and Methods. (A) Both single-strand (○) and duplex –TEC– (●) are compared to single-strand (□) and duplex (■) –AEA– DNA. The εA in both sequences is more fluorescent in the single-strand and becomes quenched in the duplex. (B) Fluorescence emission spectra were collected in the presence of saturating concentration of AAG and compared to the spectra in the absence of AAG. The –TEC– duplex DNA exhibited very similar fluorescence in the absence (●) or presence (○) of AAG. In contrast, the –AEA– duplex DNA, which was strongly fluorescent in the absence of protein (■), was strongly quenched in the presence of AAG (□). The spectra were taken within 1 minute of AAG addition (1 μM), during which time negligible excision can occur.

concentration of oligonucleotides was incorrect could be ruled out by measuring the fluorescence of the εA base that is released upon complete hydrolysis by AAG (see below).[2]

Since the AAG-catalyzed excision of εA is relatively slow under these conditions, it was possible to examine the effect of AAG binding to each oligonucleotide in a conventional fluorometer. The addition of one equivalent of AAG resulted in a dramatic 5-fold decrease in the fluorescence of the –AEA– substrate, whereas that of the –TEC–

[2] Samples that contained 400 nM of either –TEC– or –AEA– duplex (determined using the UV-absorbance and the calculated extinction coefficients) gave identical fluorescence emission spectra upon the complete hydrolysis by AAG (Figure 3.4A and B). This eliminates the trivial possibility that different amounts of DNA were compared. Therefore, we conclude that the –AEA– sequence has significantly higher fluorescence than the –TEC– sequence in both single-strand and when annealed to its complement (Figure 3.3).

substrate was only slightly quenched (Figure 3.3B). Addition of more than one equivalent of AAG had no further effect on either spectrum (data not shown). The two oligonucleotides exhibit almost identical fluorescence emission when they are bound to AAG (Figure 3.3B). This observation strongly suggests that the εA is in a similar environment in both complexes, most likely flipped out of the duplex and bound in the AAG active site. Crystal structures of the AAG•εA-DNA complex indicate potential quenching interactions between the εA lesion and active site tyrosines Y127 and Y159 (*16*). The environment of the protein active site is more strongly quenching than the–AEA– sequence context, but very similar to that of the –TEC– sequence context. The strong quenching of both complexes relative to free εA in solution suggested that a continuous fluorescence-based assay could be used to follow the εA excision activity of AAG.

**Single Turnover Glycosylase Activity of AAG Monitored by εA Fluorescence and by Formation of the Abasic DNA Product**

Hydrolysis of the *N*-glycosidic bond by AAG results in the formation of an abasic site and a free base (Figure 3.1). Early assays of AAG employed alkylated genomic DNA and relied upon directly detecting the released base by HPLC or scintillation counting (*37, 38*). More commonly, gel-based discontinuous assays are used to detect the formation of a single strand break (*27, 29, 39*). These assays can employ either a fluorescent or radioactive label, and the abasic product is converted into a single-strand nick by hydrolyzing the abasic site at alkaline pH. In contrast, the fluorescent properties of εA provide the opportunity for a continuous glycosylase assay. We directly

Table 3.1. Sequence-dependent changes in εA fluorescence and rate constants for binding and excision by AAG[a]

| | Sequence Context | |
|---|---|---|
| | –TEC– | –AEA– |
| ΔF hybridization[b] | 1.7 | 1.4 |
| ΔF binding[b] | 1.1 | 5.2 |
| $k_{on}$ ($M^{-1}s^{-1}$)[c] | $2.1 \pm 0.9 \times 10^{8}$ | $1.5 \pm 0.4 \times 10^{8}$ |
| $k_{off}$ ($s^{-1}$)[d] | 33 ± 15 | 48 ± 15 |
| $k_{flip}$ ($s^{-1}$)[e] | 2.6 ± 0.1 | 4.3 ± 0.4 |
| $k_{unflip}$ ($s^{-1}$)[f] | $2.0 \pm 0.2 \times 10^{-3}$ | ND[k] |
| $K_{flip}$[g] | 1300 | ND[k] |
| $k_{chem}$ ($s^{-1}$)[h] | $5.7 \pm 0.3 \times 10^{-4}$ | $5.3 \pm 0.2 \times 10^{-4}$ |
| $K_d^{ns}$ (nM)[i] | 160 | 320 |
| $K_d^{εA\text{-}DNA}$ (nM)[j] | 0.1 | ND[k] |

[a]Values are reported at 25 °C in a buffer containing 50 mM NaMES, pH 6.5, 100 mM NaCl, 1 mM EDTA, and 1mM DTT. [b]εA fluorescence was determined with 400 nM εA-containing DNA and a 3 mm pathlength cuvet with excitation at 320 nm and emission at 410 nm. The ΔF hybridization is the ratio of the fluorescence of single-stranded DNA divided by the fluorescence of the double-stranded DNA (Figure 3.3A). The ΔF binding to AAG is the ratio of the fluorescence of the double-stranded DNA divided by the fluorescence of the saturated AAG•DNA complex (Figure 3.3B). [c]The association rate constant ($k_{on}$) was determined from the data in Figure 3.7. [d]The dissociation rate constant ($k_{off}$) was determined from the data in Figure 3.6. [e]The rate constant for flipping was determined by stopped-flow (Figure 3.8). As discussed in the text, the second phase of the change in εA fluorescence is simply $k_{flip}$, since $k_{unflip}$ is much slower than $k_{flip}$. Given the similar rate constants for the two sequences, we assumed that the observed rate constant for the –AEA– sequence is also the flipping rate constant. [f]The reverse rate constant for flipping ($k_{unflip}$; Figure 3.1), is determined from the macroscopic rate constant for dissociation of the substrate (Figure 3.9), as described in the text. [g]The equilibrium for nucleotide flipping is defined as the ratio of the forward rate constant for flipping measured by stopped-flow and the reverse rate constant for flipping ($K_{flip} = k_2/k_{-2}$; Scheme 3.3). [h]The rate for the chemical step is simply the rate constant for the single turnover excision of εA (Figure 3.4D), since the equilibrium constant for base flipping is highly favorable and the rate constant for flipping is much faster than the hydrolysis step. The values shown are determined from εA fluorescence. For the –TEC– DNA, a similar value of $6.8 \pm 0.3 \times 10^{-4}$ $s^{-1}$ was determined from the gel-based assay. [i]The dissociation constant for the nonspecific AAG•DNA complex is given by the ratio of the off and on rates measured by stopped-flow ($K_d^{ns} = k_{off}/k_{on}$). [j]The dissociation constant for the specific, flipped-out complex between AAG and εA-DNA is defined by $K_d^{εA\text{-}DNA} = K_d^{ns}/K_{flip}$. [k]ND, not determined.

compared the fluorescence-based assay to the traditional gel-based assay to determine whether εA release is fast or slow and to evaluate the potential of using fluorescence to monitor the glycosylase activity of AAG.

Under single turnover conditions, with enzyme in excess over the DNA substrate, AAG was incubated with oligonucleotide –TEC– and fluorescence emission spectra were collected at the indicated intervals (Figure 3.4A). The εA fluorescence ($\lambda_{max}$ = 410 nm) increased over time, reaching a maximum value at ~2 hours. Under identical conditions, a fluorescein-labeled oligonucleotide of the identical sequence (–TEC–) was incubated with AAG and aliquots were quenched in sodium hydroxide and analyzed by denaturing polyacrylamide gel electrophoresis (Figure 3.4C). The change in fluorescence was converted into the fraction of εA released into solution and the fraction of abasic product was quantified from the gel and these data were plotted together (Figure 3.4D; See Materials and Methods for details). The reaction progress curves were readily fit by a single exponential, and the observed rate constants for εA release ($k_{max}^{eA}$ = 0.036 ± 0.002 $min^{-1}$) and for abasic site formation ($k_{max}$ = 0.038 ± 0.002 $min^{-1}$) were identical within error. A simultaneous fit to both reaction progress curves gave the same value ($k_{max}$ = 0.038 $min^{-1}$; Figure 3.4D). Since the detection of the abasic DNA product by sodium hydroxide quench/hydrolysis does not require dissociation of the εA, the observation of the same rate constant for release of εA into solution is most simply consistent with rate-limiting N-glycosidic bond hydrolysis, followed by rapid release of εA.

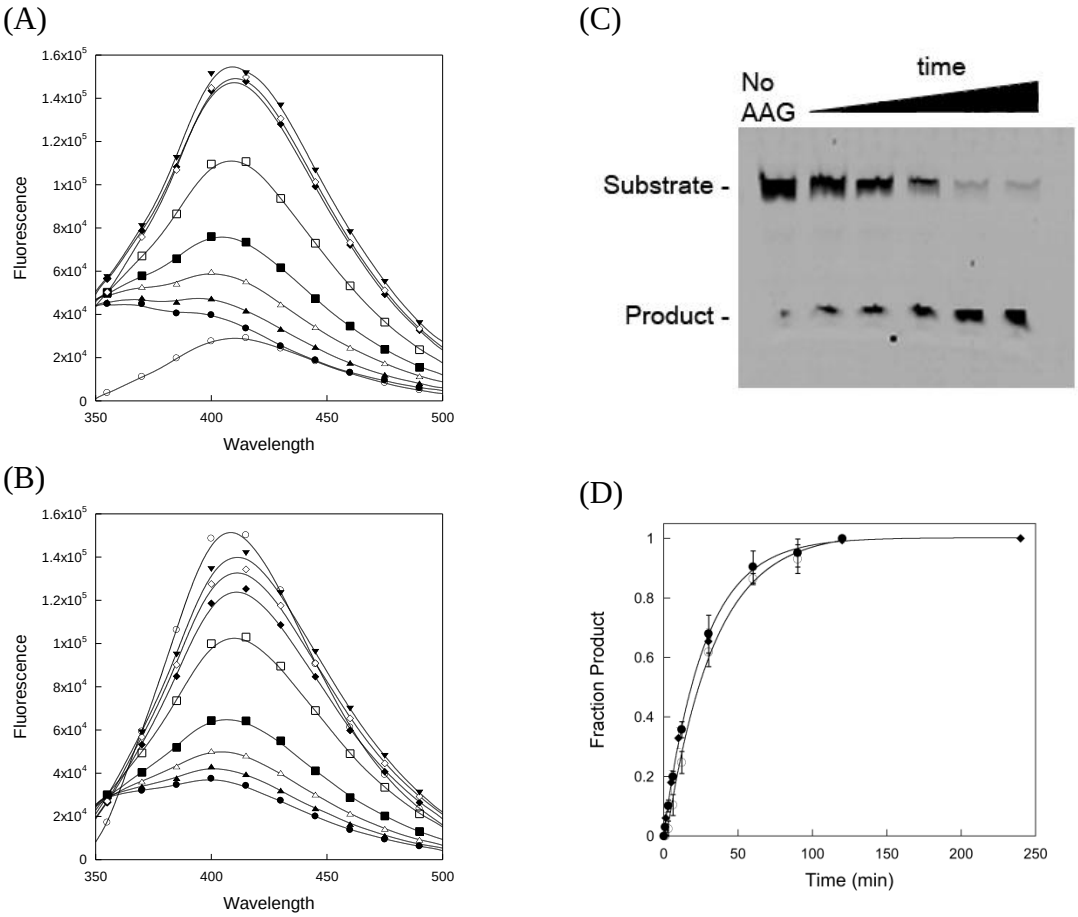



Figure 3.4. Comparison of εA release and N-glycosidic bond cleavage from single turnover assays. Fluorescence emission spectra monitoring the increase in fluorescence that is observed upon single turnover excision of εA from (A) –TEC– and (B) –AEA– DNA catalyzed by AAG. The spectra for samples containing 400 nM duplex DNA were collected prior to the addition of protein (○), and within 1 minute of the addition of 1.2 μM AAG (●). Additional spectra were collected after 3 (▲), 6 (△), 12 (■), 30 (□), 60 (♦), 90 (◇) and 120 (▼) minutes. The high fluorescence of the –AEA– substrate is strongly quenched upon binding to AAG. Representative experiments are shown, but duplicate reactions gave essentially identical results, and the rate of reaction was independent of [AAG]. (C) The gel-based glycosylase assay relies on selective hydroxide-catalyzed cleavage of abasic sites and the different mobility of intact substrate and nicked product to monitor formation of the abasic DNA product. A representative denaturing polyacrylamide gel shows the separation of the fluorescein-labeled substrate (25-mer) and product (12-mer) and a time course for the formation of the abasic DNA product. The first lane shows a no enzyme control. The small amount of nicked DNA (~4%) results from nonenzymatic ring-opening and depurination during synthesis, purification, and the hydroxide heating step. The next 5 lanes indicate samples taken from a reaction containing 100 nM –TEC– DNA and 200 nM AAG that were quenched 2–120 minutes after the initiation of the reaction by the addition of AAG. (D) Quantification of the above results indicates that both assays are limited by a common step, presumably N-glycosidic bond cleavage. The fraction product was calculated from the εA fluorescence assay for –TEC– (●) and –AEA– (○) as well as the gel based excision assay for –TEC– (♦). The error bars indicate the standard deviation for 2-5 independent determinations. The reaction curves for –TEC– monitored by both assays are fit by a single exponential (Equation 3.1; $k_{max}$ = 0.038 $min^{-1}$). The reaction of –AEA– is slightly slower ($k_{max}$ = 0.031 $min^{-1}$).

The single turnover assay for release of εA was also performed for the –AEA– DNA under the same conditions that were used for the –TEC– DNA (Figure 3.4B). The initial fluorescence of this DNA in the absence of AAG is much higher than for the –TEC– sequence, but upon the addition of AAG the fluorescence was strongly quenched. The observed increase in fluorescence was of similar magnitude to that which was observed for the –TEC– substrate, and a similar single turnover rate constant was observed ($k_{max}$ = 0.031 ± 0.002 $min^{-1}$; Figure 3.4D). [1] These data indicate that once AAG is bound, it shows very similar efficiency of excising εA from these two different sequence contexts.

**Association and Nucleotide Flipping Monitored by Stopped-Flow Fluorescence**

The quenching of εA fluorescence observed upon binding of AAG to εA-containing DNA suggested that εA would be a useful reporter to monitor association of AAG and DNA. Therefore, we performed stopped-flow fluorescence experiments to monitor the time-dependent changes in εA fluorescence. Two phases were observed when excess AAG was mixed with 500 nM –TEC– DNA, both of which occur well before the base excision step (Figure 3.5A).

In the first 100 ms, there is a transient increase in fluorescence and this is followed by a decrease in fluorescence that is complete within a few seconds after mixing. The fluorescent traces were best fit by a double exponential equation, indicating a two-step model for DNA binding (Scheme 3.3). As increasing concentrations of AAG were mixed with a fixed concentration of –TEC– DNA, the rate of the first phase increased (Figure 3.5A and Figure 3.6). However, at concentrations above 1 μM AAG this

initial phase occurred in the mixing time of the stopped-flow. The dependence of the first phase of the fluorescence change upon the concentration of reactants indicates that a bimolecular event is being monitored.

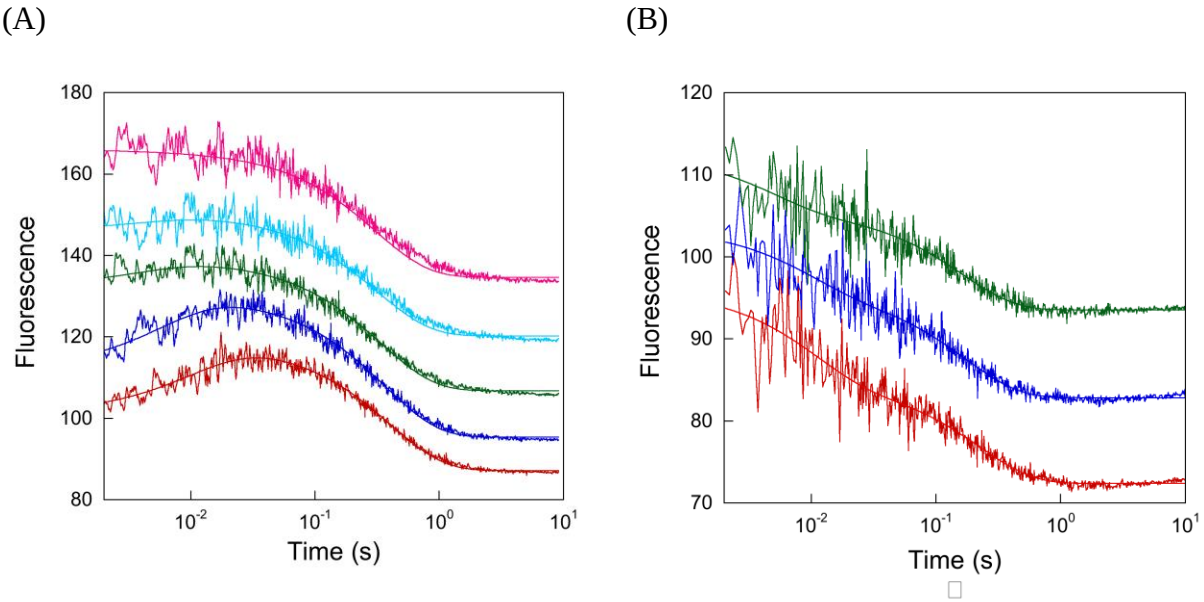



Figure 3.5. Stopped-flow fluorescence measurements of the binding of εA-DNA to AAG with excess protein. Representative data is shown from experiments in which a fixed concentration of DNA was mixed with increasing concentrations of AAG. Samples were excited at 320 nm and emission was collected with a 360 nm long-pass filter (See Materials and Methods for details). Several individual reactions were monitored at each concentration of AAG, and the traces shown are the averages of all reactions. Since the fluorescence changes were superimposed at the different concentrations of AAG, the traces are offset for clarity. (A) The binding of 0.5 μM –TEC– duplex DNA by 0.5 (bottom), 1, 2, 4, and 8 μM AAG (top). (B) The binding of 0.15 μM –AEA– duplex DNA by 0.5 (bottom), 1, and 2 μM (top) AAG. The lines indicate the best fits of equation 3.2 to each binding reaction.

It should be noted that the fit of the two concentrations of AAG with the –TEC– substrate in Figure 3.6 is tenuous, however it was possible to measure binding over a wider concentration range for the more highly fluorescent –AEA– DNA.

Since initial binding is expected to be away from the site of damage, and the bimolecular step gives a detectable change in fluorescence, we interpret this behavior to

reflect a rate-limiting binding step with no change in fluorescence, followed by a rapid sliding step that results in the formation of a complex with a detectable change in fluorescence. We infer that the scanning step is too fast to be directly observed in our experiments (Figure 3.1), with a lower limit of 250 $s^{-1}$ indicated by the data in Figure 3.6.

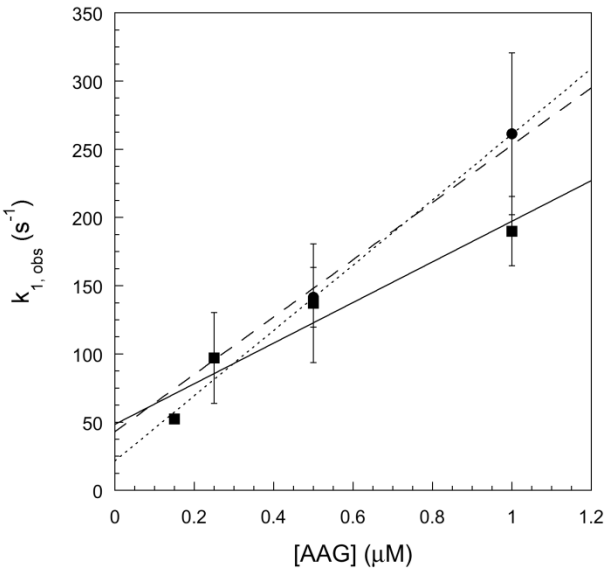



Figure 3.6. Determination of the forward and reverse rate constants for the initial binding of AAG to εA-DNA. The observed rate constant for DNA binding ($k_{1, obs}$) is from the first phase of the binding reactions shown in Figure 3.5. The average and standard deviations are shown for 2-4 determinations. The dissociation rate constant for the initial AAG•DNA complex can be determined from the y-intercept of this plot ($k_{1, obs} = k_{-1} + k_1$ [AAG]; Scheme 3.3). For –AEA– DNA (■), the best linear fit (solid line) gives an intercept of 48 $s^{-1}$ ($k_{-1}$) and a slope of $1.5\times10^8$ $M^{-1}s^{-1}$ ($k_1$). This second order rate constant is identical to the value determined from equimolar binding experiments (Table 3.1 and Figure 3.8). For –TEC– DNA (●), only two concentrations of protein gave reliable data due to the weak fluorescence and more rapid binding. A linear fit (dotted line) of the two data points that were obtained gives an intercept of 22 $s^{-1}$ and a slope of $2.4 \times 10^8$ $M^{-1}s^{-1}$. By forcing the slope of the line to be $2.1\times10^8$ $M^{-1}s^{-1}$ (dashed line), the value that was independently determined from equimolar binding reactions (Table 3.1 and Figure 3.8), an intercept of 43 $s^{-1}$ was obtained. We take the average ($33 \pm 15$ $s^{-1}$) as a rough estimate for the $k_{-1}$ value for this substrate.

The increased fluorescence of this early intermediate is likely to reflect partial unstacking of the εA lesion during the initial damage recognition by AAG. The second phase of the binding reaction was insensitive to the concentration of AAG and we interpret this to be the flipping of the εA lesion into the AAG active site, where it is strongly quenched. For

the –TEC– substrate, this AAG-bound complex happens to have very similar fluorescence to that of free DNA duplex in solution (Figure 3.3B).

We performed analogous stopped-flow binding experiments with the –AEA– substrate (Figure 3.5B). Binding of AAG to this sequence appears to be superficially different, without a transient increase in fluorescence, but is also best explained by a two-step binding mechanism. For this substrate, the fluorescence is initially high in the free DNA, is partially quenched in the first intermediate, and is more strongly quenched in the fully flipped complex. These traces could not be fit well by a single exponential (data not shown), but were well fit by a double exponential (Figure 3.5B). As the concentration of AAG was increased at a fixed concentration of DNA, the rate for the first exponential was proportional to the amount of AAG, and the second phase was not (Figure 3.5 and Figure 3.6 and data not shown). The fits to the second phase of the reaction were excellent in all cases, but the fits to the first phase were more variable due to the very fast binding. As the fluorescence of εA is relatively weak, it was not possible to use significantly lower concentrations of DNA that would have allowed more accurate determination of the association and dissociation rate constants of this fast binding reaction.

Therefore, the stopped-flow experiments were repeated with equimolar amounts of enzyme and substrate, with concentrations ranging from 0.15–1.5 μM (Figure 3.7). This causes the binding reaction to slow as free ligand is consumed, allowing a time-dependent change in fluorescence to be observed even at high initial concentrations Similar experiments have been used to study binding reactions of equimolar protein fragments (*40, 41*). Since the unimolecular flipping step also gives an observed fluorescence change, a second exponential term is also required (Equation 3.3). The

fitting of this equation to the data in Figure 3.7 gave similar association rate constants of ~$2 \times 10^8$ $M^{-1}s^{-1}$ for both sequences (Figure 3.8A; Table 3.1). Experiments carried out across a 10-fold concentration range gave the same rate constant. This better-defined value for the association rate constant ($k_1$; Scheme 3.3) was used to fit the data from the concentration dependence of AAG when protein was in excess and provide an estimate of the dissociation rate constant ($k_{-1} \approx 30$ $s^{-1}$; Figure 3.6 and Table 3.1).

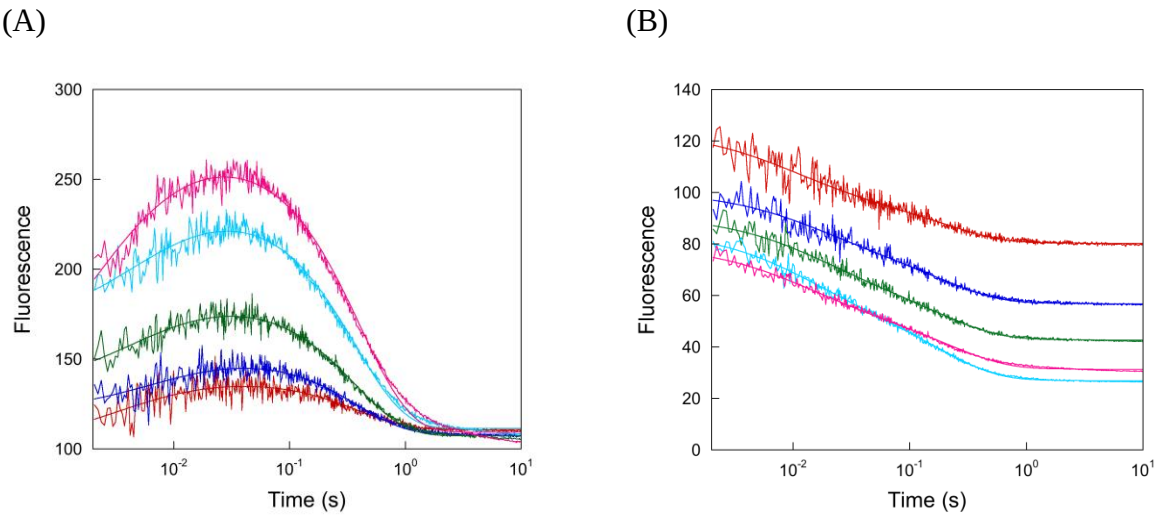



Figure 3.7. Kinetics of DNA binding and nucleotide flipping, determined by stopped-flow fluorescence at equimolar concentrations of AAG and DNA. AAG was rapidly mixed with an equimolar solution of –TEC– (A) or –AEA– (B) DNA to yield final concentrations after mixing of 0.15 (bottom) to 1.5 μM (top). Each trace is the average of 3 independent binding reactions, and the lines indicate the best fits of equation 3.3. For the –TEC– substrate the raw data is plotted, indicating the increased signal observed at higher concentrations of DNA, but for the –AEA– substrate the larger fluorescence change required that the gain be adjusted between different concentrations of DNA and therefore the individual traces are arbitrarily offset.

The second phase of the binding reaction followed a single exponential and the same rate constant was observed across the concentration range tested (Figure 3.8B). The observed rate constant for flipping is an approach to equilibrium and therefore is the sum of the forward and reverse rate constants for flipping ($k_{2,\,obs} = k_2 + k_{-2}$; Scheme 3.3).

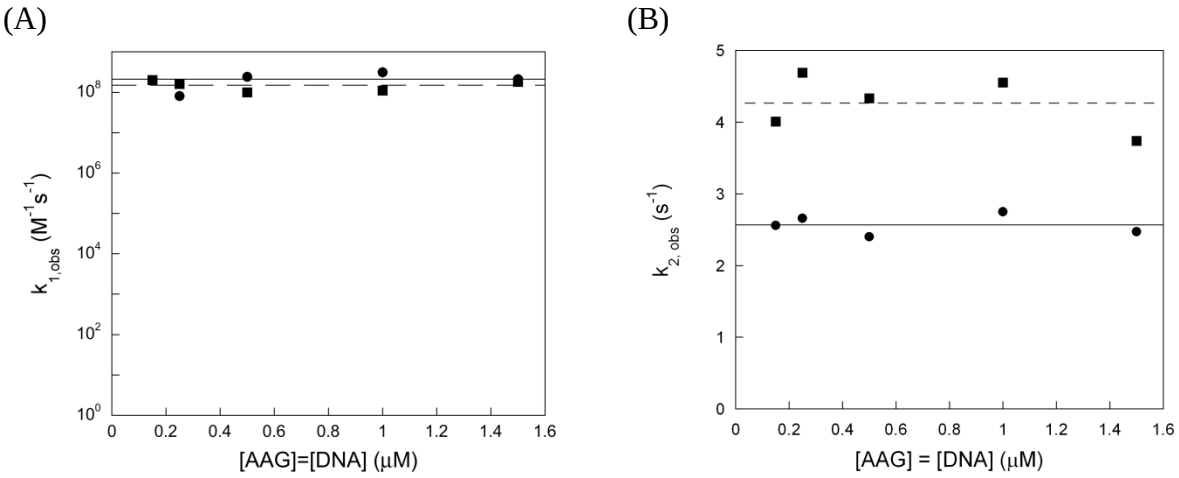



Figure 3.8. The rate constants for binding and nucleotide flipping by AAG in two different sequence contexts. The observed rate constants for –TEC– DNA (●, solid line) and –AEA– DNA (■, dashed line) were taken from the fits of equation 3.3 to the data for equimolar binding reactions (Figure 3.7). (A) The observed bimolecular rate constant for the initial binding of AAG to DNA is very similar for both substrates (~$2 \times 10^8$ $M^{-1}s^{-1}$; Table 3.1). According to the simple two-step binding mechanism depicted in Scheme 3.3, this observed rate constant is simply $k_1$, the microscopic rate constant for binding. (B) The observed rate constant for the second step is independent of concentration, and yields an average value of 4.3 $s^{-1}$ for the –AEA– sequence context compared to a value of 2.6 $s^{-1}$ for the –TEC– sequence context. According to Scheme 3.3, this observed rate constant is the sum of the forward and reverse rate constants for flipping ($k_{2, obs} = k_2 + k_{-2}$). Additional experiments described below indicate that the equilibrium constant for flipping is highly favorable ($k_2 >> k_{-2}$), and therefore the observed rate constant for flipping is simply the forward rate constant for flipping (see text for additional discussion).

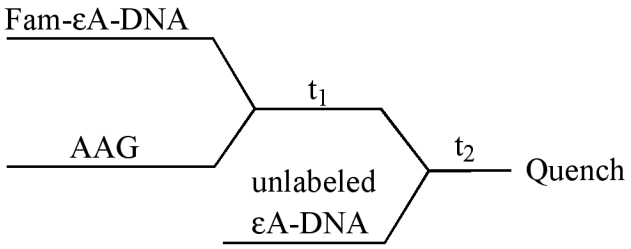
Below we describe a pulse-chase experiment that resolves this ambiguity by establishing that the reverse rate constant for flipping ($k_{-2}$) is much slower than the rate constant measured by the stopped-flow fluorescence, such that the observed rate constant is simply the forward rate constant for flipping ($k_{2, obs} = k_{flip}$). The nucleotide flipping step ($k_{flip}$) is significantly faster for the –AEA– sequence than for the –TEC– sequence (4.3 ± 0.4 $s^{-1}$ versus 2.6 ± 0.1 $s^{-1}$). The difference in the rates of flipping in two different sequence contexts may reflect subtle differences in how AAG engages these different DNA molecules or sequence-dependent changes in the intrinsic motions of DNA. Regardless of the origin of the differences between the two sequence contexts that were

examined, these data establish that the flipping step is much faster than the N-glycosidic bond hydrolysis step (Table 3.1).

**A Pulse-Chase Assay to Determine the Macroscopic Rate of Substrate Dissociation**

In order to gain further insight into the two-step binding mechanism of AAG, we also examined the dissociation of substrate from AAG by a pulse-chase assay (*30*, *33*). In this experiment, fluorescein-labeled –TEC– substrate was briefly incubated with AAG ($t_1$ = 20 s) to give it time to bind, and then was chased with a 100-fold excess of –TEC– substrate lacking the fluorescein label (Figure 3.9A). Samples were taken over the normal single turnover time course to evaluate the partitioning of a bound E•S complex between N-glycosidic bond hydrolysis and dissociation. Figure 3.9B shows the εA-excision reaction in the presence and absence of chase. The addition of chase caused a decrease in the fraction of substrate hydrolyzed and an increase in the observed rate constant, as expected if the forward rate constant for reaction is similar to the rate constant for dissociation (see Methods for details). Under these conditions (I = 120 mM), the macroscopic rate constant for dissociation is 0.12 ± 0.02 $min^{-1}$ (Table 3.1). Since the equilibrium for bound εA-DNA lies towards the fully flipped out complex, either $k_{-1}$ or $k_{-2}$ could in principle limit the observed rate constant for dissociation. However, the rapidly reversible initial binding step exhibits a much faster rate constant ($k_{-1}$ ~ 30 $s^{-1}$; Figure 3.6). Therefore, the observed rate constant for the dissociation of the substrate that was measured by the pulse-chase assay is simply the unflipping rate constant ($k_{-2}$; Scheme 3.3; see the Materials and Methods for additional details).

(A)




(B) (C)

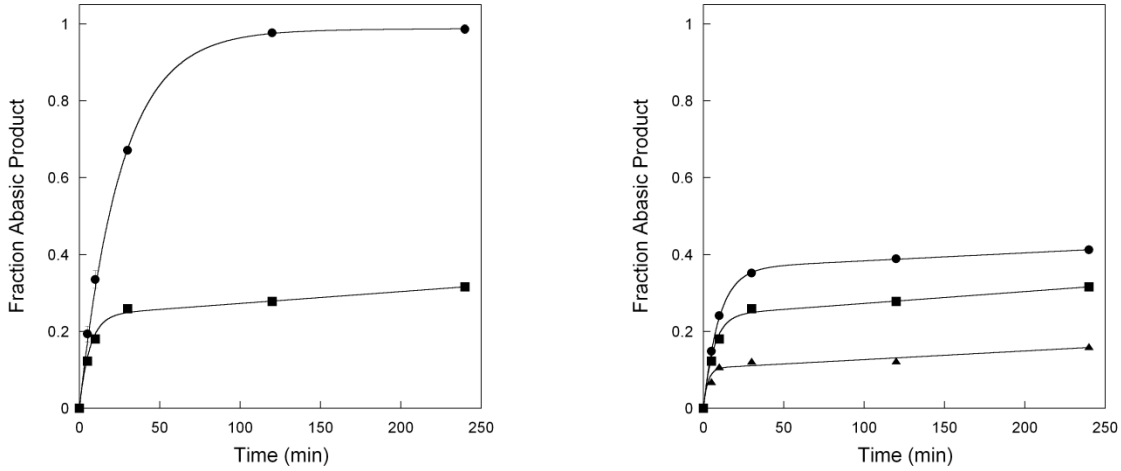



Figure 3.9. Pulse-chase experiment to measure the macroscopic dissociation rates for εA-containing DNA substrate. (A) The experimental design is indicated. Fluorescein (FAM)-labeled DNA (100 nM –TEC–) was mixed with excess AAG (200 nM) for 20 seconds (incubation time, $t_1$) then 20 μM unlabeled –TEC– DNA was added as a chase. The reactions were quenched at the indicated time points ($t_2$), and the fraction of abasic DNA product determined by alkaline hydrolysis and gel electrophoresis (See Materials and Methods for additional details). (B) Single turnover excision of εA from –TEC– DNA by AAG in the absence (●) or presence (■) of chase under the standard reaction conditions. In the absence of chase, the line indicates the best fit to a single exponential (Equation 3.1). In the presence of chase, the line indicates the best fit of an exponential burst followed by a slow steady-state reaction (Equation 3.4). This gives a macroscopic dissociation rate constant of 0.12 ± 0.01 $min^{-1}$. (C) Comparison of pulse chase experiments performed at concentrations of sodium chloride. The data from experiments at (●) 70 mM, (■) 120 mM and (▲) 320 mM ionic strength were fit by equation 3.4 (lines). Analysis of the observed rate constant and the endpoint gave essentially identical macroscopic rate constants and the average and standard deviation of duplicate reactions determined by both methods yielded macroscopic dissociation rate constants of 0.08 ± 0.005, 0.12 ± 0.01, and 0.35 ± 0.05 $min^{-1}$, respectively for 70, 120, and 320 mM ionic strength.

It has previously been observed that the rate of dissociation of AAG from its abasic DNA product is extremely sensitive to the ionic strength, with a 25-fold increase

in the rate of dissociation between 40 and 120 mM ionic strength (*34*). This effect of ionic strength is presumably due to electrostatic interactions between the positively charged protein and the negatively charged DNA that become weaker at higher ionic strength. In order to evaluate whether the unflipping rate constant or the DNA dissociation rate constant is rate-limiting, we varied the ionic strength with the expectation that the dissociation step should exhibit a large ionic strength effect. The results of pulse-chase experiments carried out at 70, 120, and 320 mM ionic strength are shown in Figure 3.9C. The pulse-chase experiment revealed a small decrease in amplitude and concomitant increase in the observed rate constant as the ionic strength is increased. Control reactions established that the single turnover rate constant in the absence of chase is independent of ionic strength under these conditions (data not shown), consistent with previous results (*13*, *34*). Therefore, the modest ~4-fold change in the observed substrate dissociation rate constant between 70 and 320 mM is further evidence that $k_{-2}$ and not $k_{-1}$ is rate limiting for dissociation of εA-DNA. This moderate sensitivity to ionic strength indicates either a small effect of ionic strength on the flipping step, or a small contribution from the rate constant for dissociation from nonspecific DNA.

**Rate Constant for Dissociation of the Abasic DNA Product**

We measured the rate of dissociation of the abasic DNA product to compare it to the rate of εA product release. The steady-state reaction of AAG with –TEC- substrate gave a $k_{cat}$ value that is essentially identical to the single turnover rate constant under these reaction conditions (data not shown), suggesting that the very slow rate of hydrolysis of the N-glycosidic bond is rate limiting for multiple turnover excision of εA

at high ionic strength. Since AAG shows much faster rates of excision of Hx from inosine lesions, we measured the $k_{cat}$ value for multiple turnover on an inosine-containing substrate, in Figure 3.10.

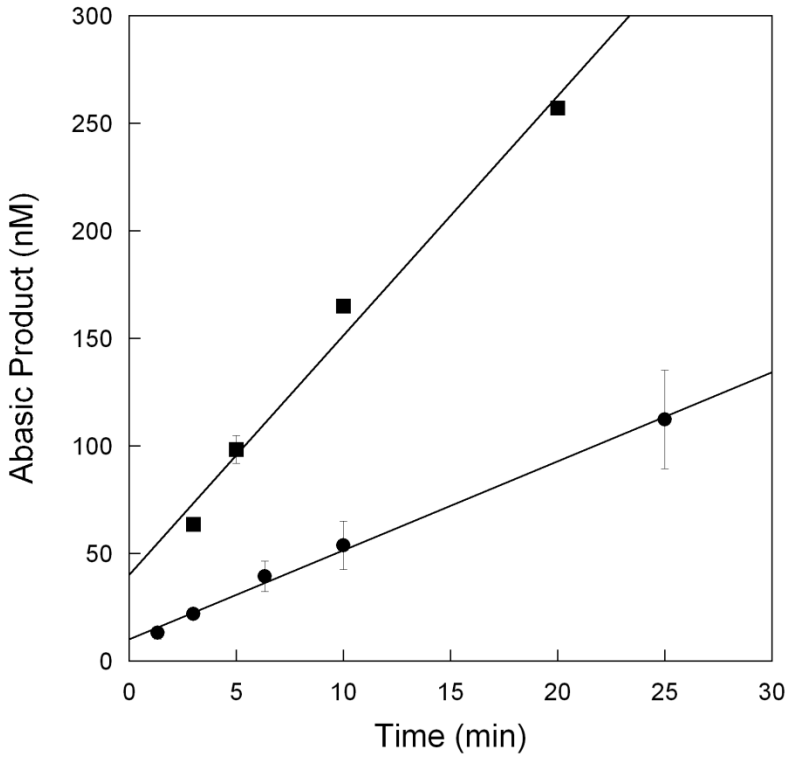



Figure 3.10. Multiple turnover excision of hypoxanthine by AAG. Under the standard conditions, the rate-limiting step is dissociation of the abasic DNA product. Shown are the fits to the first 10% of the reaction for 2 μM –TEC- DNA with 20 nM (●) and 50 nM (■) AAG. These data gave a rate constant for abasic DNA dissociation of 0.24 ± 0.06 $min^{-1}$. This is similar to the value of 0.73 $min^{-1}$ that was measured at 37 $^{o}$C and pH 7.0 (*34*).

It has previously been shown that AAG exhibits burst kinetics for the excision of Hx and that the steady-state rate constant, $k_{cat}$, reflects the rate of dissociation of the abasic DNA product (*34*). We determined that the rate constant for the dissociation of abasic DNA is 0.24 $min^{-1}$ under these reaction conditions, ~6-fold faster than the rate constant for N-glycosidic bond cleavage (Scheme 3.4). This is only slightly faster than the observed rate constant for dissociation of the εA substrate ($k_{off, obs}$ = 0.12 $min^{-1}$),

suggesting that the abasic DNA product is also likely to be stably bound in an extrahelical conformation.

## Discussion

### Minimal Kinetic and Thermodynamic Mechanism for Excision of εA

We have used the intrinsic fluorescence of the εA lesion to establish the kinetic mechanism for the AAG-catalyzed recognition and excision of this alkylated base. The fluorescence of εA is highly sensitive to the local environment of the lesion and it proved to be a sensitive reporter of three distinct steps in the AAG-catalyzed reaction; initial lesion recognition, nucleotide flipping, and εA product release. The rate constants that were determined from changes in εA fluorescence are in excellent agreement with the rate constants measured by the traditional glycosylase assay that detects the abasic DNA product, indicating that εA release is fast and that fluorescence provides a useful assay of AAG activity.

The kinetic mechanism for recognition and excision of εA is depicted for the –TEC– sequence in Scheme 3.4.

$$E + S \underset{2\times10^{3}\ min^{-1}}{\overset{1\times10^{10}\ M^{-1}min^{-1}}{\rightleftharpoons}} E\cdot S \underset{0.1\ min^{-1}}{\overset{160\ min^{-1}}{\rightleftharpoons}} E\cdot S' \xrightarrow{0.04\ min^{-1}} E\cdot P_1\cdot P_2 \xrightarrow[+\ P_1]{>>0.04\ min^{-1}} E\cdot P_2 \overset{0.2\ min^{-1}}{\rightleftharpoons} E + P_2$$

Scheme 3.4.

Initial binding is near the diffusion-limited rate and is rapidly reversible. Given that a detectable change in εA fluorescence occurs as a bimolecular reaction, we cannot

distinguish between the possibilities that initial binding by AAG occurs solely at the εA lesion or that AAG binds at another site but rapidly transfers to the εA lesion. We favor the latter model, since it is known that AAG can use nonspecific DNA binding interactions to translocate between sites (*13*). Since these data require the scanning step(s) to be very rapid ($k_{scan}$; Figure 3.1), this step is not shown in Scheme 3.4. This initial recognition complex is rapidly reversible, with a dissociation rate constant of ~1800 $min^{-1}$. This indicates a relatively weak complex for initial nonspecific DNA binding and we calculate a dissociation constant of ~150 nM from these data ($K_d^{ns}$; Table 3.1). Occasionally, AAG catalyzes full nucleotide flipping with a rate constant of ~160 $min^{-1}$. The extrahelical complex is highly stable, with a calculated equilibrium constant of ~1300 for the flipping step, and this tight binding decreases the dissociation constant for binding of AAG to an εA lesion by several orders of magnitude to approximately 100 pM (Table 3.1). This exceptionally tight binding to the εA site is consistent with previous reports of sub-nanomolar binding by AAG under different experimental conditions (*2, 42*). The crystal structure of AAG bound to εA reveals a snug fit of the εA lesion into the AAG pocket where it accepts a hydrogen bond from the backbone amide of His136 (*16*). This interaction with His136 can also account for the discrimination against normal adenosine nucleotides (*2, 16*). Tight binding of the damaged nucleotide by AAG could protect it from DNA-templated activities such as transcription and replication.

Despite the rapid and highly favorable flipping of the εA lesion, AAG is limited by a very slow rate of N-glycosidic bond cleavage, such that the specific recognition complex is only partially committed to base excision. Release of εA, measured by the increased fluorescence of the free base, is much faster than the rate of N-glycosidic bond

cleavage measured by formation of the abasic DNA product. At 25 $^{o}$C, this rate constant for N-glycosidic bond cleavage is 0.04 min$^{-1}$, 5-fold slower than is observed at 37 $^{o}$C (*2*). Under these conditions, base excision is also slower than release of the abasic DNA product, such that no burst is observed (Table 3.1 and data not shown). Since N-glycosidic bond cleavage for inosine is much faster than for εA, the slow release of the abasic DNA product is rate-limiting for the multiple turnover reaction of AAG with inosine and other rapidly hydrolyzed lesions (*34*). Rate-limiting release of the abasic DNA product provides the opportunity for coordination of the steps of the BER pathway. It has previously been shown that the slow rate of dissociation of AAG from its abasic DNA product can be stimulated by APE1, the enzyme that acts next in the base excision repair pathway (*34, 35, 43*).

**Recognition of εA in a PolyA-Tract**

Earlier studies had investigated the AAG-catalyzed excision of εA from different sequence contexts. One study reported that the excision of εA by AAG is sensitive to the sequence context and that an εA lesion in the center of a polyA-tract is a particularly poor substrate for AAG (*27*). In contrast, a later report examined a number of similar sequence contexts and found that there did not appear to be significant effects of sequence context on the excision of εA (*26*). However, Hx-containing substrates show large sequence context effects hinting at a fundamental difference between the recognition of εA and Hx (*25, 26, 43, 44*). In this study we performed a more thorough kinetic characterization to compare the recognition of εA in the polyA-tract (–AEA–) and in a heterogenous sequence (–TEC–).

Overall, the steps of εA recognition and excision by AAG are very similar for the two DNA substrates (Table 3.1). The rate constants for DNA binding and excision are identical within error, but the observed rate constant for flipping is ~1.7-fold higher for the polyA-tract sequence. Although this is a small change, the rate constant for flipping is independent of DNA and enzyme concentration and is sufficiently well defined by the stopped-flow data to indicate a real difference. This difference in rate could be due to changes in either the structure or the dynamics of different DNA sequences. Nevertheless, since nucleotide flipping is rapid and highly favorable, the difference in flipping rate does not translate into a significant difference in the rate of εA excision from these two sequences. The minimal kinetic and thermodynamic framework (Scheme 3.3) provides a molecular explanation for why AAG does not exhibit strong sequence-dependent differences in the rates of excision of εA, but does show strong effects for other lesions. By extension, it is very likely that other lesions that are recognized by AAG, such as the deaminated base lesion Hx, are not stably flipped and thus changes in stacking interactions have a larger effect on the flipping equilibrium and on the observed rates of excision (*2*, *25*).

**Comparison of AAG to Other Glycosylases**

There are several distinct structural and evolutionary families of DNA repair glycosylases (*45*, *46*). These enzymes differ in their chemical mechanisms and in the identity of the lesions that they recognize, but nonetheless all recognize damaged bases and catalyze the cleavage of the N-glycosidic bond (*47*, *48*). The observation that all of these enzymes that recognize nucleotides in duplex DNA catalyze nucleotide flipping attests to the requirement of the enzyme to gain access to the N-glycosidic bond that is

occluded in duplex DNA (*26*). Fluorescence methods have been used to study the DNA binding and nucleotide flipping steps for a few of these enzymes and the results from the literature are summarized in Table 3.2. Since each of these enzymes belongs to a different

Table 3.2. Comparison of nucleotide flipping by DNA repair glycosylases

| Enzyme | Super-family[a] | Substrate | $k_{flip}$ ($s^{-1}$) | $k_{unflip}$ ($s^{-1}$) | $K_{flip}$ | $k_{chem}$ ($s^{-1}$)[b] | Reference |
|---|---|---|---|---|---|---|---|
| *Narrow substrate range* | | | | | | | |
| UDG | UDG | 2'-fluoroU•G | 1260 | 38 | 33 | N/A[c] | (*17*) |
| | | U•A | 1100 | 370 | 3 | 58 | (*18*) |
| | | U (–AUA–)[d] | 670 | 16 | 41 | 38 | (*19*) |
| | | U (–GUC–)[d] | 48 | 12 | 4 | ND[e] | (*19*) |
| T4 Pdg | HhH | D•2AP[f] | ≥300 | ND[e] | ≥10 | N/A[c] | (*20*) |
| Ogg1 | HhH | 8-oxo-G•C | 13 | 1 | 10 | 0.07 | (*21*) |
| *Broad substrate range* | | | | | | | |
| Fpg | Fpg | 8-oxo-G•C | 12 | 12 | 1 | 0.16 | (*22*) |
| AAG | AAG | εA•T | 3 | 0.002 | 1300 | 0.0006 | This study |

[a]HhH, helix-hairpin-helix superfamily. UDG (uracil DNA glycosylase), Fpg (formamidopyrimidine DNA glycosylase), and AAG are the founding members of distinct enzyme superfamilies. T4 Pdg, pyrimidine dimer DNA glycosylase from phage T4, and Ogg1, human 8-oxoguanine DNA glycosylase. [b]$k_{chem}$ is the rate constant for the first chemical step, hydrolysis of the N-glycosidic bond. [c]N/A, not applicable. [d]Rate constants for binding of UDG to single-stranded DNA with two different sequence contexts. It should be noted that a very different process of nucleotide selection is expected in single strand and duplex DNA, but the values are included here because they are remarkably similar to the values that were independently measured for the enzyme binding to duplex DNA. [e]ND, not determined. [f]The D•2AP (tetrahydrofuran•2-aminopurine) pair mimics an abasic site product.

superfamily and therefore evolved independently, it is interesting to consider whether there are common functional constraints that have dictated the kinetic mechanisms of these enzymes (*46, 49, 50*). One crude classification that can be used is whether the enzyme exhibits a narrow substrate range (recognizing one or a few structurally related lesions) or whether it exhibits a broad substrate range (recognizing a variety of

structurally diverse lesions). The rate constant for flipping varies across two orders of magnitude from 3 $s^{-1}$ to more than 1000 $s^{-1}$. With these limited data, it seems that the enzymes with a narrow substrate range exhibit faster flipping rates than those with a broader substrate range. This could be related to the difficulty of optimizing the rate of flipping of distinct lesions by the enzymes with a broad substrate range, or it could simply reflect the slow rate of N-glycosidic bond cleavage since the flipping step is still much faster than the hydrolysis step. Indeed, the nucleotide flipping step is much faster than the N-glycosidic bond cleavage step for all of these enzymes (Table 3.2).

The equilibrium constant for flipping is close to unity for formamidopyrimidine DNA glycosylase (Fpg), the only enzyme with a broad substrate range that has been previously characterized. It makes sense that an enzyme that can accommodate structurally diverse lesions might not be able to hold all lesions tightly. In contrast, the enzymes that recognize a narrow range of substrates all exhibit favorable equilibrium constants for flipping of ≥10. Interestingly, AAG has an exceptionally broad range of substrates and yet also has the most highly favorable equilibrium constant for flipping that has been measured. 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 by AAG (*2*, *25*).

**Implications of Multi-stage Recognition for the Biological Specificity of AAG**

The findings presented above have several implications for the initiation of base excision repair by AAG. The most obvious is that the kinetic mechanism is optimized for specificity rather than speed. The vast excess of undamaged nucleotides in DNA provides

a challenge to an enzyme that must recognize a broad range of structurally distinct substrates. Two-step binding is rapid and readily reversible at each step, which provides multiple opportunities for discrimination between damaged and undamaged bases. The multiplicative effect of multi-stage recognition allows even small differences between normal and damaged bases to be magnified to achieve the specificity that is required to recognize rare lesions that can occur throughout the genome.

Once the lesion is flipped, rate-limiting N-glycosidic bond cleavage allows for additional discrimination between damaged and undamaged nucleotides. Pyrimidines are selected against by the catalytic mechanism, which is optimized for purine nucleotides, and A and G are selected against by steric clashes with their exocyclic amino groups (*2, 16, 29, 51*). Since alkylated nucleotides that bear positive charge are activated for N-glycosidic bond cleavage, some alkylated purines are removed at biologically relevant levels despite the relatively poor rate enhancement (*52, 53*).

We propose that the highly favorable equilibrium constant for the flipping of εA by AAG is the exception that proves the rule. The opposing base effects that have previously been noted strongly suggest that the equilibrium constant for flipping methylated and deaminated purines is unfavorable (*2, 25*). In spite of this unfavorable equilibrium, many of these other lesions are excised by AAG more rapidly than the εA lesion is excised. The favorable equilibrium constant for εA helps to compensate for the very slow rate of base excision. If εA is not repaired, it is highly mutagenic, and therefore it must be removed from the genome. However, human cells have a relatively long cell cycle and this provides a wide window of time during which repair can be performed prior to DNA replication. Therefore DNA repair enzymes can afford to sacrifice speed

(overall rate of reaction) for increased selectivity (ability to recognize subtle perturbations in the chemical structure of DNA). AAG is able to effectively locate rare sites of damage that are hidden in the sea of undamaged DNA by employing multiple DNA binding steps that are rapidly reversible and each provide specificity towards DNA damage.

# References


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

2. O'Brien, P. J., and Ellenberger, T. (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase, *J Biol Chem 279*, 9750–9757.

3. Hitchcock, T. M., Dong, L., Connor, E. E., Meira, L. B., Samson, L. D., Wyatt, M. D., and Cao, W. (2004) Oxanine DNA glycosylase activity from Mammalian alkyladenine glycosylase, *J Biol Chem 279*, 38177–38183.

4. Gros, L., Ishchenko, A. A., and Saparbaev, M. (2003) Enzymology of repair of etheno-adducts, *Mutat Res 531*, 219–229.

5. Speina, E., Zielinska, M., Barbin, A., Gackowski, D., Kowalewski, J., Graziewicz, M. A., Siedlecki, J. A., Olinski, R., and Tudek, B. (2003) Decreased repair activities of 1,N(6)-ethenoadenine and 3,N(4)-ethenocytosine in lung adenocarcinoma patients, *Cancer Res 63*, 4351–4357.

6. Pandya, G. A., and Moriya, M. (1996) 1,N6-ethenodeoxyadenosine, a DNA adduct highly mutagenic in mammalian cells, *Biochemistry 35*, 11487–11492.

7. Levine, R. L., Yang, I. Y., Hossain, M., Pandya, G. A., Grollman, A. P., and Moriya, M. (2000) Mutagenesis induced by a single 1,N6-ethenodeoxyadenosine adduct in human cells, *Cancer Res 60*, 4098–4104.

8. Ringvoll, J., Moen, M. N., Nordstrand, L. M., Meira, L. B., Pang, B., Bekkelund, A., Dedon, P. C., Bjelland, S., Samson, L. D., Falnes, P. O., and Klungland, A. (2008) AlkB homologue 2-mediated repair of ethenoadenine lesions in mammalian DNA, *Cancer Res 68*, 4142–4149.

9. Elder, R. H., Jansen, J. G., Weeks, R. J., Willington, M. A., Deans, B., Watson, A. J., Mynett, K. J., Bailey, J. A., Cooper, D. P., Rafferty, J. A., Heeran, M. C., Wijnhoven, S. W., van Zeeland, A. A., and Margison, G. P. (1998) Alkylpurine-DNA-N-glycosylase knockout mice show increased susceptibility to induction of mutations by methyl methanesulfonate, *Mol Cell Biol 18*, 5828–5837.

10. Engelward, B. P., Weeda, G., Wyatt, M. D., Broekhof, J. L., de Wit, J., Donker, I., Allan, J. M., Gold, B., Hoeijmakers, J. H., and Samson, L. D. (1997) Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase, *Proc Natl Acad Sci U S A 94*, 13087–13092.

11. Kisby, G. E., Olivas, A., Park, T., Churchwell, M., Doerge, D., Samson, L. D., Gerson, S. L., and Turker, M. S. (2009) DNA repair modulates the vulnerability of the developing brain to alkylating agents, *DNA Repair (Amst) 8*, 400–412.

12. Meira, L. B., Moroski-Erkul, C. A., Green, S. L., Calvo, J. A., Bronson, R. T., Shah, D., and Samson, L. D. (2009) Aag-initiated base excision repair drives alkylation-induced retinal degeneration in mice, *Proc Natl Acad Sci U S A 106*, 888–893.

13. Hedglin, M., and O'Brien, P. J. (2008) Human alkyladenine DNA glycosylase employs a processive search for DNA damage, *Biochemistry 47*, 11434–11445.

14. Roberts, R. J., and Cheng, X. (1998) Base flipping, *Annu Rev Biochem 67*, 181–198.

15. Lau, A. Y., Scharer, O. D., Samson, L., Verdine, G. L., and Ellenberger, T. (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision, *Cell 95*, 249–258.

16. Lau, A. Y., Wyatt, M. D., Glassner, B. J., Samson, L. D., and Ellenberger, T. (2000) Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG, *Proc Natl Acad Sci U S A 97*, 13573–13578.

17. Stivers, J. T., Pankiewicz, K. W., and Watanabe, K. A. (1999) Kinetic mechanism of damage site recognition and uracil flipping by Escherichia coli uracil DNA glycosylase, *Biochemistry 38*, 952–963.

18. Wong, I., Lundquist, A. J., Bernards, A. S., and Mosbaugh, D. W. (2002) Presteady-state analysis of a single catalytic turnover by Escherichia coli uracil-DNA glycosylase reveals a "pinch-pull-push" mechanism, *J Biol Chem 277*, 19424–19432.

19. Bellamy, S. R., Krusong, K., and Baldwin, G. S. (2007) A rapid reaction analysis of uracil DNA glycosylase indicates an active mechanism of base flipping, *Nucleic Acids Res 35*, 1478–1487.

20. Walker, R. K., McCullough, A. K., and Lloyd, R. S. (2006) Uncoupling of nucleotide flipping and DNA bending by the t4 pyrimidine dimer DNA glycosylase, *Biochemistry 45*, 14192–14200.

21. Kuznetsov, N. A., Koval, V. V., Zharkov, D. O., Vorobjev, Y. N., Nevinsky, G. A., Douglas, K. T., and Fedorova, O. S. (2007) Pre-steady-state kinetic study of substrate specificity of Escherichia coli formamidopyrimidine--DNA glycosylase, *Biochemistry 46*, 424–435.

22. Kuznetsov, N. A., Koval, V. V., Zharkov, D. O., Nevinsky, G. A., Douglas, K. T., and Fedorova, O. S. (2005) Kinetics of substrate recognition and cleavage by human 8-oxoguanine-DNA glycosylase, *Nucleic Acids Res 33*, 3919–3931.

23. Vallur, A. C., Feller, J. A., Abner, C. W., Tran, R. K., and Bloom, L. B. (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*, 31673–31678.

24. Biswas, T., Clos, L. J., 2nd, SantaLucia, J., Jr., Mitra, S., and Roy, R. (2002) Binding of specific DNA base-pair mismatches by N-methylpurine-DNA glycosylase and its implication in initial damage recognition, *J Mol Biol 320*, 503–513.

25. Vallur, A. C., Maher, R. L., and Bloom, L. B. (2005) The efficiency of hypoxanthine excision by alkyladenine DNA glycosylase is altered by changes in nearest neighbor bases, *DNA Repair (Amst) 4*, 1088–1098.

26. Wyatt, M. D., and Samson, L. D. (2000) Influence of DNA structure on hypoxanthine and 1,N(6)-ethenoadenine removal by murine 3-methyladenine DNA glycosylase, *Carcinogenesis 21*, 901–908.

27. Hang, B., Sagi, J., and Singer, B. (1998) Correlation between sequence-dependent glycosylase repair and the thermal stability of oligonucleotide duplexes containing 1, N6-ethenoadenine, *J Biol Chem 273*, 33406–33413.

28. Leonard, N. J. (1984) Etheno-substituted nucleotides and coenzymes: fluorescence and biological activity, *CRC Crit Rev Biochem 15*, 125–199.

29. O'Brien, P. J., and Ellenberger, T. (2003) Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines, *Biochemistry 42*, 12418–12429.

30. Rose, I. A., O'Connell, E. L., and Litwin, S. (1974) Determination of the rate of hexokinase-glucose dissociation by the isotope-trapping method, *J Biol Chem 249*, 5163–5168.

31. Secrist, J. A., 3rd, Barrio, J. R., Leonard, N. J., and Weber, G. (1972) Fluorescent modification of adenosine-containing coenzymes. Biological activities and spectroscopic properties, *Biochemistry 11*, 3499–3506.

32. Fersht, A. (1999) *Structure and Mechanism in Protein Science*, pp 200–201, W.H. Freeman and Company, New York.

33. Hsieh, J., Walker, S. C., Fierke, C. A., and Engelke, D. R. (2009) Pre-tRNA turnover catalyzed by the yeast nuclear RNase P holoenzyme is limited by product release, *RNA 15*, 224–234.

34. Baldwin, M. R., and O'Brien, P. J. (2009) Human AP Endonuclease I Stimulates Multiple-Turnover Base Excision by Alkyladenine DNA Glycosylase, *Biochemistry 48*, 6022–6033.

35. Maher, R. L., Vallur, A. C., Feller, J. A., and Bloom, L. B. (2007) Slow base excision by human alkyladenine DNA glycosylase limits the rate of formation of AP sites and AP endonuclease 1 does not stimulate base excision, *DNA Repair (Amst) 6*, 71–81.

36. Kelley, S. O., and Barton, J. K. (1999) Electron transfer between bases in double helical DNA, *Science 283*, 375–381.

37. Riazuddin, S., and Lindahl, T. (1978) Properties of 3-methyladenine-DNA glycosylase from Escherichia coli, *Biochemistry 17*, 2110–2118.

38. O'Connor, T. R. (1993) Purification and characterization of human 3-methyladenine-DNA glycosylase, *Nucleic Acids Res 21*, 5561–5569.

39. Wyatt, M. D., Allan, J. M., Lau, A. Y., Ellenberger, T. E., and Samson, L. D. (1999) 3-methyladenine DNA glycosylases: structure, function, and biological importance, *Bioessays 21*, 668–676.

40. de Prat Gay, G., Ruiz-Sanz, J., and Fersht, A. R. (1994) Generation of a family of protein fragments for structure-folding studies. 2. Kinetics of association of the two chymotrypsin inhibitor-2 fragments, *Biochemistry 33*, 7964–7970.

41. Schreiber, G., and Fersht, A. R. (1993) Interaction of barnase with its polypeptide inhibitor barstar studied by protein engineering, *Biochemistry 32*, 5145–5150.

42. Scharer, O. D., and Verdine, G. L. (1995) A designed inhibitor of base excision repair, *J. Am. Chem. Soc. 117*, 10781–10782.

43. Xia, L., Zheng, L., Lee, H. W., Bates, S. E., Federico, L., Shen, B., and O'Connor, T. R. (2005) Human 3-methyladenine-DNA glycosylase: effect of sequence context on excision, association with PCNA, and stimulation by AP endonuclease, *J Mol Biol 346*, 1259–1274.

44. Abner, C. W., Lau, A. Y., Ellenberger, T., and Bloom, L. B. (2001) Base excision and DNA binding activities of human alkyladenine DNA glycosylase are sensitive to the base paired with a lesion, *J Biol Chem 276*, 13379–13387.

45. Huffman, J. L., Sundheim, O., and Tainer, J. A. (2005) DNA base damage recognition and removal: new twists and grooves, *Mutat Res 577*, 55–76.

46. O'Brien, P. J. (2006) Catalytic promiscuity and the divergent evolution of DNA repair enzymes, *Chem Rev 106*, 720–752.

47. Berti, P. J., and McCann, J. A. (2006) Toward a detailed understanding of base excision repair enzymes: transition state and mechanistic analyses of N-glycoside hydrolysis and N-glycoside transfer, *Chem Rev 106*, 506–555.

48. Stivers, J. T., and Jiang, Y. L. (2003) A mechanistic perspective on the chemistry of DNA repair glycosylases, *Chem Rev 103*, 2729–2759.

49. Gerlt, J. A., and Babbitt, P. C. (2001) Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprafamilies, *Annu Rev Biochem 70*, 209–246.

50. Glasner, M. E., Gerlt, J. A., and Babbitt, P. C. (2006) Evolution of enzyme superfamilies, *Curr Opin Chem Biol 10*, 492–497.

51. Connor, E. E., and Wyatt, M. D. (2002) Active-site clashes prevent the human 3-methyladenine DNA glycosylase from improperly removing bases, *Chem Biol 9*, 1033–1041.

52. O'Brien, P. J., and Ellenberger, T. (2004) The Escherichia coli 3-methyladenine DNA glycosylase AlkA has a remarkably versatile active site, *J Biol Chem 279*, 26876–26884.

53. Berdal, K. G., Johansen, R. F., and Seeberg, E. (1998) Release of normal bases from intact DNA by a native DNA repair enzyme, *Embo J 17*, 363–367.

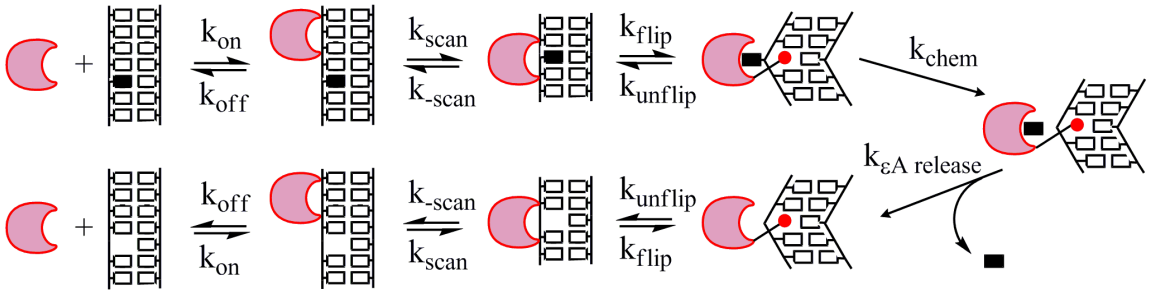
# Chapter 4

# Substitution of Active Site Tyrosines with Tryptophans Alters the Free Energy for Nucleotide Flipping by Alkyladenine DNA Glycosylase[3]

Human alkyladenine DNA glycosylase (AAG) is responsible for recognizing a wide variety of structurally diverse deaminated and alkylated purine lesions (*1-2*). It 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 repair polymerase, and a DNA ligase are required to restore the correct DNA sequence, using the intact strand as a template. 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 (*3-4*). We have recently used the intrinsic fluorescence of εA to characterize the minimal kinetic mechanism for the recognition, flipping, and excision of this lesion by wildtype AAG (*5*). The minimal mechanism from this work is summarized in Scheme 4.1. Rapid mixing experiments indicated that an initial recognition complex rapidly and reversibly formed in which the εA lesion was partially unstacked. We hypothesized that further changes in the fluorescence of the εA lesion was indicative of the flipping step in which the εA was captured in the active site of AAG. In order to test this hypothesis, we sought

[3] Special thanks to Jenna Tomlinson who helped perform the experiments and analyze the data in this chapter.

to observe conformational changes in the protein directly. Similar approaches have been very informative for other DNA repair glycosylases (*6-9*).




Scheme 4.1.

The catalytic domain of AAG contains 3 tryptophan and 6 tyrosine residues. The 3 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 their environment would not change upon DNA binding. Five of the tyrosine residues line either the binding pocket or the DNA-protein interface, but tyrosine absorbance is very close to the absorbance of the DNA and it seemed unlikely that they would serve as useful reporters. Therefore, we introduced 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 (*3-4*). The conformation of Y127 is stacked over the εA lesion, whereas the Y159 phenolic side chain makes a T-shaped interaction with this lesion (Figure 4.1). Comprehensive analysis of aromatic stacking interactions in protein structures suggests that T-type interactions are very common and presumably contribute similar favorable interaction energies (*10*). It has been suggested that a tyrosine to tryptophan mutation should be relatively conservative (*11*). However these two positions are invariantly found to be tyrosine in AAG homologs in organisms ranging from bacteria to plants to vertebrates (*12*).

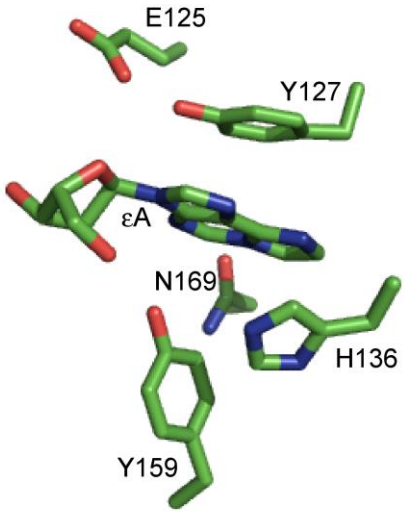



Figure 4.1. Active site contacts with the flipped out εA from the published crystal structure of AAG bound to εA-DNA. The coordinates are from 1EWN (*4*) and the figure was rendered with Pymol. For clarity, only the immediate residues surrounding the εA base are shown.

Substitution of conserved 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 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. Although flipping is perturbed in these mutants, they follow the same mechanism as the wildtype enzyme with formation of two distinct intermediates prior to N-glycosidic bond cleavage. Identical rate constants 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 and that the most strongly quenched intermediate that was previously observed is indeed the extrahelical specific recognition complex (*5*).

# Materials and Methods

## Purification of Wildtype and Mutant AAG

The catalytic domain of human AAG that lacks the first 79 amino acids was expressed in *E. coli* and purified as previously described (*13*). The Y127W and Y159W mutants were constructed by site directed mutagenesis and purified using the same protocol (*13*). 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 wildtype 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, shown in Scheme 4.2.

TEC 5′-CGATAGCATCCT**E**CCTTCTCTCCAT
3′-GCTATCGTAGGATGGAAGAGAGGTA

AEA 5′-CGATAGCATCAA**E**AATTCTCTCCAT
3′-GCTATCGTAGTTTTTAAGAGAGGTA

Scheme 4.2

The oligonucleotides were desalted using Sephadex G-25 and purified using denaturing polyacrylamide gel electrophoresis as previously described (*14*). 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 εA containing oligonucleotides, 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 (*15*). The lesion-containing oligonucleotides were annealed with a 1.1-fold excess of the complement by heating to 90°C and cooling slowly to 4°C.

**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 (200 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 1 minute. No significant excision of εA occurs during this time. The data were fit to a quadratic equation assuming tight binding by AAG (~2 nM $K_d$) (*16-17*).

**Single Turnover Excision of εA Monitored by Fluorescence**

Experiments with mutant and wildtype AAG were performed as previously described (*5*). 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.

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 410 nm. Therefore this wavelength was chosen to follow the release of εA into solution. The fluorescence at 410 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 4.1). The single turnover rate constant was determined at saturating concentration of AAG and this was confirmed by the observation of the same rate constant at two different concentrations of AAG. Under these conditions, the observed rate constant is equal to the maximal single turnover rate constant ($k_{obs} = k_{max}$).

$$\text{Fraction product} = A(1-e^{-k_{obs}t}) \quad (4.1)$$

**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 (*5*). 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 could be fit by Equation 4.2, in which F is εA fluorescence as a function of time, C is the fluorescence of the free DNA, Y and Z are the change 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, $k_{2,\,obs}$ is the observed unimolecular rate constant for the subsequent flipping step, and t is the time (*5, 18*).

$$F = C + Y\left(\frac{E_o^2 k_1 t}{1 + E_o k_1 t}\right) - Z(1 - e^{-k_{2,obs} t}) \qquad (4.2)$$

Several different concentrations between 0.15 and 1.5 μM (all concentrations are given after mixing) gave the same rate constants, confirming that the concentration chosen was saturating for mutant and wildtype proteins. At least three traces were averaged together for each concentration in two independent experiments, and the averaged data were fit by Equation 4.2 using Kaleidagraph.

**Pulse-Chase Assay to Measure Substrate Dissociation**

The macroscopic rate constant for dissociation of wildtype and mutant AAG from εA-containing DNA was measured by pulse-chase in the standard reaction buffer as previously described for wildtype AAG (*5*). Briefly, 50 nM fluorescein-labeled TEC DNA was mixed with 100 nM AAG for 20 seconds, a chase of 200 μM unlabeled TEC DNA was added. At various time points, a sample from the reaction was removed and quenched in 2 volume equivalents 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 and

heated 3 minutes at 70°C before loading onto a 12-15% polyacrylamide gel. Gels were scanned with a Typhoon Imager (GE 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)].

Base excision catalyzed by AAG results in fluorescein-labeled product, whereas dissociation releases uncleaved fluorescein-labeled substrate. The partitioning between hydrolysis and dissociation are can be determined from both the exponential rate constant and by the change in burst amplitude (Equation 4.3 and 4.4). These assays were done at 25°C in the standard buffer, except that the concentration of NaCl was varied and 0.1 mg/mL BSA was added. In 20 µL reactions, 100 nM fluorescein labeled –TEC– DNA was mixed with 200 nM AAG for 20 seconds. A chase of 20 µM unlabeled –TEC– DNA was then mixed in. 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. Samples were taken at the specified times and quenched as described above. The samples were run on sequencing gels and the fraction product was calculated as described above. It was expected that the production of product would follow a simple exponential (Equation 4.1), but under the conditions employed a very slow steady-state reaction was also observed. This was accounted for by Equation 4.4, in which the burst phase gives the amplitude (A) and an observed rate constant ($k_{obs}$). The steady-state phase is dependent upon the steady-state rate constant ($k_{cat}$), the concentration of enzyme (E), and the total concentration of labeled and unlabeled DNA substrate.

$$\text{Fraction product} = A(1 - e^{-k_{obs}t}) + (k_{cat}[E]/[DNA]_{total})t \qquad (4.4)$$

$$E + S \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} E \cdot S \underset{k_{-flip}}{\overset{k_{flip}}{\rightleftharpoons}} E \cdot S' \xrightarrow{k_{chem}} E \cdot P_1 \cdot P_2$$

Scheme 4.3.

According to the two-step binding mechanism described in Scheme 4.3, two different partitioning equations can be written (*19*). Since all labeled substrate is initially bound, the fraction of product formed is given by fraction that goes on to react. This is indicated by Equation 4.5, in which A is the burst amplitude (the fraction of product cleaved 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 stable flipped complex. This expression can be rearranged to solve for the desired dissociation rate constant (Equation 4.6). 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 4.7). Solving for $k_{off,\,obs}$ give Equation 4.8.

$$A = \frac{k_{max}}{k_{off,obs} + k_{max}} \qquad (4.5)$$

$$k_{off,obs} = \frac{k_{max}}{A} - k_{max} \qquad (4.6)$$

$$k_{obs} = k_{off,obs} + k_{max} \qquad (4.7)$$

$$k_{off,obs} = k_{obs} - k_{max} \quad (4.8)$$

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 (Equation 4.6 and 4.8). Both methods gave essentially identical values for $k_{off, obs}$ and we report the average of the results obtained with both methods.

Since the εA-containing DNA binds in two steps, the observed rate constant for dissociation of substrate ($k_{off, obs}$) could be limited by the unflipping rate ($k_{-2}$) or dissociation from nonspecific DNA ($k_{-1}$) or both. According to Scheme 4.3, and assuming that the flipped complex is stable (i.e., $k_2 >> k_{-2}$), this observed dissociation rate can be expressed in terms of the microscopic rate constants [Equation 4.9; (*19*)]. Stopped-flow fluorescence experiments suggests that dissociation from the initial AAG•DNA complex is rapid ($k_{-1}$ ~30 $s^{-1}$) relative to the forward flipping rate constant $k_2$ = 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.

$$k_{off,obs} = k_{-2}\left(\frac{k_{-1}}{k_{-1} + k_2}\right) \quad (4.9)$$

The pulse chase experiment was performed at high and low ionic strength by varying the concentration of sodium chloride from 50-300 mM. It has previously been shown that the rate constant for dissociation of the abasic DNA product increases as ionic

strength is increased, presumably due to the disruption of electrostatic interactions at higher ionic strength (*14, 20*).

## Results

### Tryptophan Fluorescence of Native AAG

Since the catalytic domain of AAG (residues 80-298) has almost identical glycosylase activity as the full-length protein (*13, 20-21*), but is considerably more stable, we have used this construct for all of the experiments described below. In addition to exploring the functional contributions of the invariant tyrosine residues in the AAG active site, we hoped to develop fluorescent probes that could report on ligand binding and conformational changes. The catalytic domain of AAG contains 3 tryptophan residues, raising the possibility that the native protein might undergo changes in its intrinsic fluorescence upon binding to DNA. All 3 of these residues are located near the surface on the opposite side of the enzyme from the DNA binding surface that was identified in crystal structures (*3-4*). We initially tried to remove these tryptophan residues to simplify the fluorescence spectra when additional tryptophan residues were added to the protein. Although each individual tryptophan could be mutated singly to phenylalanine and full glycosylase activity was maintained, 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 would undergo any changes in tryptophan fluorescence upon binding to an 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; data not shown). This strongly suggests that the tryptophan residues do not change their environment upon DNA binding and nucleotide flipping. These observations and the data below support our conclusion that time-dependent changes in the tryptophan fluorescence of mutant proteins with an additional tryptophan simply follow the local environment of the single tryptophan mutation in the active site pocket.

**Titration of AAG Mutants with εA-containing DNA**

We performed titrations with a fixed concentration of εA-containing DNA and increasing concentration of AAG to determine the concentration of AAG. It was previously shown that –AEA- DNA shows a large decrease in fluorescence upon binding to wildtype AAG (See Materials and Methods; (*5*)). The resulting change in fluorescence as a function of AAG concentration was consistent with a simple titration and provided the concentration of AAG (Figure 4.2; See Materials and Methods for details). This concentration of AAG mutants that was determined from these titrations was used throughout as the concentration of active AAG.

**Glycosylase Activity of AAG Mutants**

We next measured the single turnover glycosylase activity of the wildtype 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-containing DNA incubated with either wildtype 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 4.3).

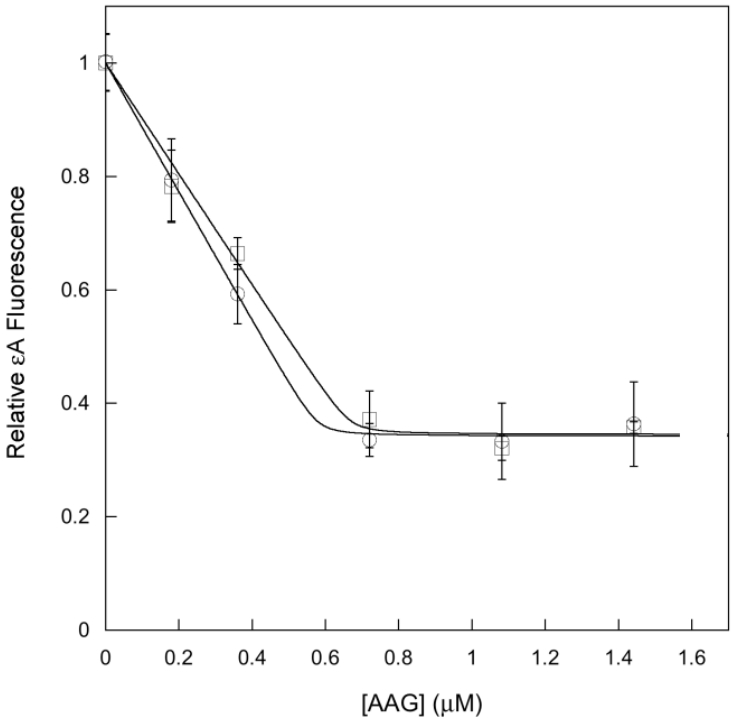



Figure 4.2. Titration of εA-containing DNA to determine the concentration of AAG. Samples of 400 nM –AEA- DNA were titrated with increasing concentration of either Y127W (squares) or Y159W (circles) AAG and the εA fluorescence was recorded. Each curve is an average of 3 independent titrations.

(A) (B)

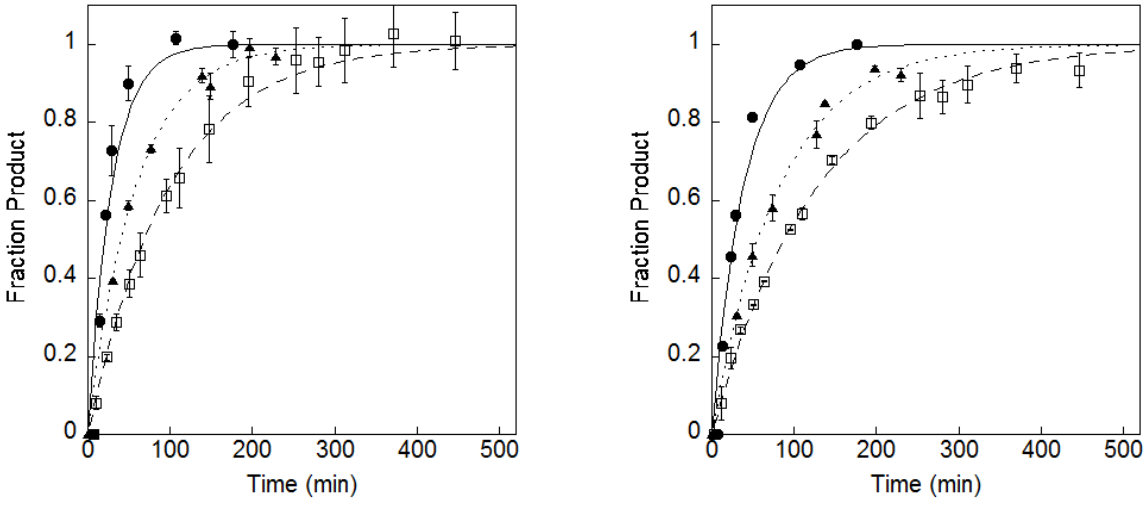



Figure 4.3. Single turnover excision of εA monitored by fluorescence. The reaction time-course for wildtype (●, solid line), Y127W (□, dashed line), and Y159W (▲, dotted line) AAG are shown for two different sequences, -TEC- (A) and –AEA- (B).

For wildtype AAG, a rate constant of 0.048 $min^{-1}$ was obtained, similar to the previously reported value of 0.034 $min^{-1}$ (*5*). The Y159W mutant was 2.7-fold slower than wildtype and the Y127W mutant was 4-fold slower (Table 4.1). These effects are similar to the respective 2.7-fold and 8-fold effects of these mutations on the single turnover excision of εA measured by abasic site formation under different reaction conditions (*13*). These are relatively modest decreases in the single turnover glycosylase activity, which raised the possibility that these introduced tryptophan residues could serve

Table 4.1. Kinetic parameters for flipping and excision of εA by wildtype and mutant AAG.

| | Wildtype AAG | Y159W AAG | Y127W AAG |
|---|---|---|---|
| $k_{on}$ ($M^{-1}s^{-1}$) | $2x10^{8}$ [a] | $7x10^{8}$ | $2x10^{8}$ |
| $k_{flip}$ ($s^{-1}$) [b] | 2.6 [a] | $\leq$ 14.5 | $\leq$ 1 |
| $k_{-flip}$ ($s^{-1}$) [c] | $2x10^{-3}$ [a] | $\geq$ $3x10^{-2}$ | $\geq$ $2x10^{-2}$ |
| $K_{flip}$ [d] | 1300 [a] | $\leq$ 470 | $\leq$ 50 |
| $k_{chem}$ ($s^{-1}$) [e] | $8x10^{-4}$ | $\geq$ $3x10^{-4}$ | $\geq$ $2x10^{-4}$ |

[a]reported in (*5*) [b]These data are from the stopped-flow experiment. They are an upper limit because $k_{obs} = k_{flip} + k_{-flip}$, so $k_{obs}$ could reflect either of the two microscopic flipping rates [c]These data are from the pulse-chase experiment. They are a lower limit because all of the DNA was chased off during the experiment, so it cannot be lower than these values but it could be faster [d]$K_{flip}$ is calculated by dividing $k_{flip}$ by $k_{-flip}$. It is an upper limit due to the limits on the two microscopic rate constants [e]These data are from the glycosylase activity assay. They are a lower limit because the if $k_{chem}$ is the rate limiting step, this is the rate constant. However, but the observed rate constant could also be affected by the flipping step, which means that $k_{chem}$ could be greater than this value if the flipping step was rate limiting

as useful reporters on the rates of conformational changes within the protein. For wildtype 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 4.3).

**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 –TEC-DNA. We chose to use equimolar concentrations of DNA and protein at concentrations far above the dissociation constant for DNA binding, since this proved to be a convenient way to follow the two-step binding of AAG to this DNA. We monitored the fluorescence changes in εA, as previously described for wildtype AAG (*5*), and the data are shown in Figure 4.4A and B. Both mutants exhibit transient increases and subsequent decreases in εA fluorescence similar to the two-step binding that was observed for wildtype AAG.

For the Y127W mutant, the initial phase of the reaction that corresponds to the initial binding step and the observed rate constant is identical within error to the rate constant that was previously measured for wildtype AAG under these conditions (Table 4.1). The second phase of the reaction corresponds to the nucleotide flipping step, and this occurs approximately 3-fold slower than observed for wild-type AAG. For a reversible reaction, this observed rate constant is equal to the sum of the rate constants for the forward and reverse nucleotide flipping steps. We previously established that the forward rate constant for flipping is much greater than the reverse rate constant for flipping for the wild-type AAG, so that $k_{flip} = k_{2, obs} = 3\ s^{-1}$(*5*). As the sum of the forward and reverse rate constants is decreased by a factor of 3, the forward rate constant for flipping must be decreased by at least 3-fold.

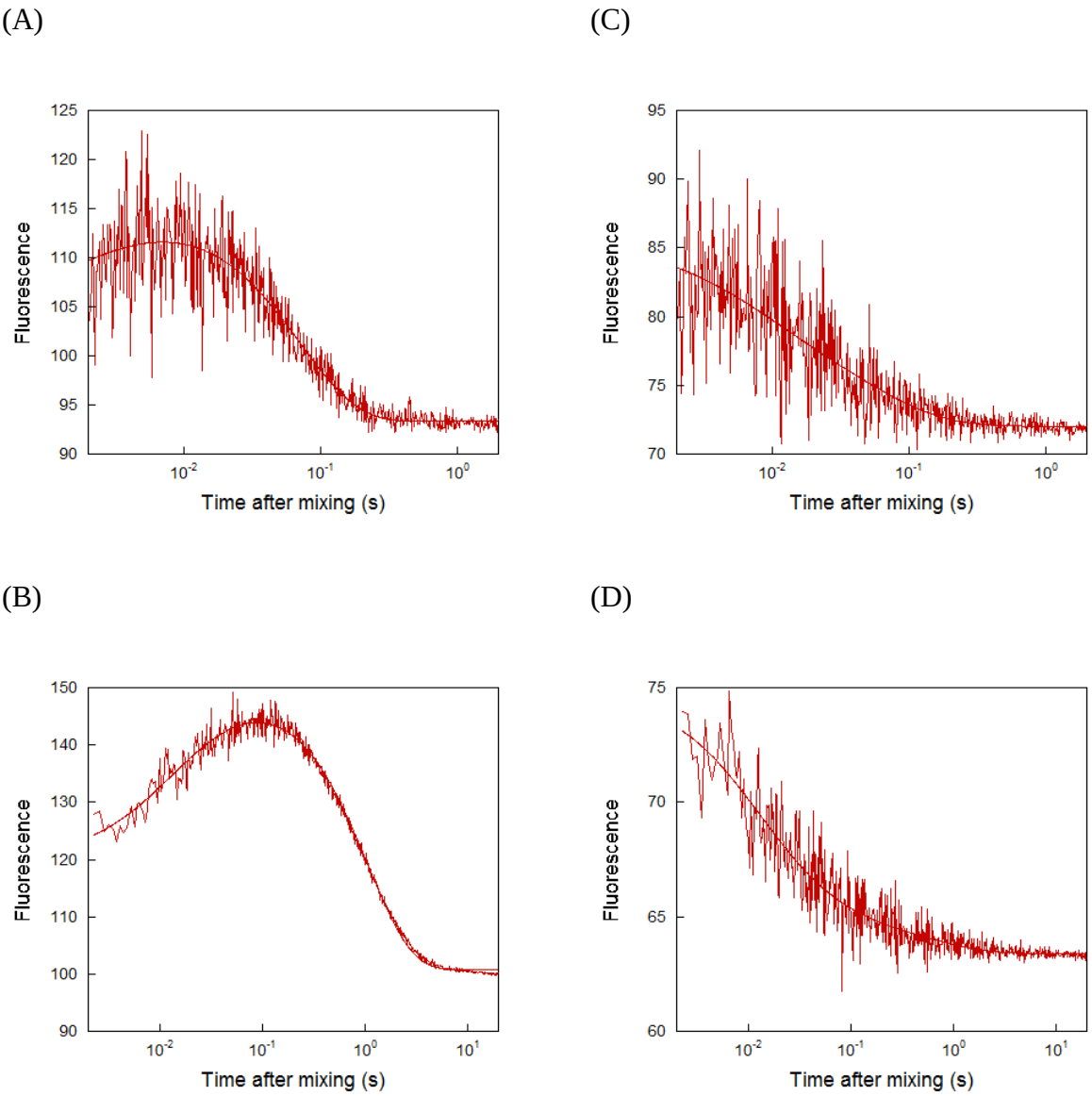



Figure 4.4. Stopped-flow fluorescence to measure binding and nucleotide flipping by Y127W and Y159W mutants of AAG. Equimolar mixtures of AAG and –TEC- DNA were mixed in the stopped-flow instrument and either εA fluorescence or tryptophan fluorescence was measured as described in the methods. (A) εA fluorescence for Y159W. (B) Tryptophan fluorescence for Y159W. (C) εA fluorescence for Y127W. (D) Tryptophan fluorescence for Y127W.

The Y159W mutant showed less εA fluorescence in the transient formation of the initial intermediate. This could indicate either a lower concentration of an intermediate that has similar fluorescent properties as wildtype, or it could indicate that this mutant has a more strongly quenching environment than wildtype. The first rate constant appears to be ~4-fold faster than was observed for the wildtype and Y127W, but this value is poorly defined by these data and the weak εA fluorescence precludes the direct determination of the binding rate constant under conditions of excess protein or excess DNA. Remarkably, the observed rate constant for the second phase of the reaction is more than 5-fold faster for the Y159W mutant than for the wildtype protein (Table 4.1). This places an upper limit on the forward rate constant for flipping of 15 $s^{-1}$, but this data does not establish whether 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. As described below, additional experiments are required to resolve these ambiguities.

**Stopped-Flow Measurement of Tryptophan Fluorescence**

We also followed the fluorescence of the tryptophan during the rapid-mixing experiments and the results are shown in Figure 4.4C and D. Control reactions in which the tryptophan fluorescence was recorded for wildtype AAG that was mixed with εA-containing DNA established that none of the three tryptophans present in AAG were sensitive to DNA binding (data not shown). This is consistent with the location of the three native tryptophan residues on the opposite surface of AAG from the DNA binding interface. When the Y127W and Y159W mutants were mixed with –TEC- DNA, then two phases of decreasing fluorescence were observed (Figure 4.4). These data were best

fit by the same equation that was used to monitor εA fluorescence, and yielded the same rate constants as were observed for εA fluorescence changes. These data reveal that tryptophan at position 127 and 159 in the active site pocket are sensitive to the initial AAG•DNA complex. This generally suggests 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, these tryptophan residues are quenched in the flipped out complex with the same observed rate constant as the εA is quenched. The crystal structure of wildtype AAG bound to εA-containing DNA shows that Y127 and Y159 both interact with the εA base, so it is expected that W residues at these sites would also interact.

**Pulse Chase Experiments to Determine the Rate of Substrate Dissociation**

We next performed pulse-chase experiments in which wildtype or mutant AAG proteins were initially mixed with fluorescently labeled –TEC- DNA and then chased with excess unlabeled –TEC- DNA (Figure 4.5). The wildtype protein gave the same results as previously reported, with 40% of the substrate partitioning forward to product and ~60% dissociating. Under these conditions, the observed rate constant is simply the reverse rate constant for nucleotide flipping [$k_{obs} = k_{-flip} = 0.002\ s^{-1}$; (*5*)]. In contrast, the addition of chase completely blocked the reaction by the Y127W and Y159W mutants (Figure 4.5B and C). 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 (Table 4.1).

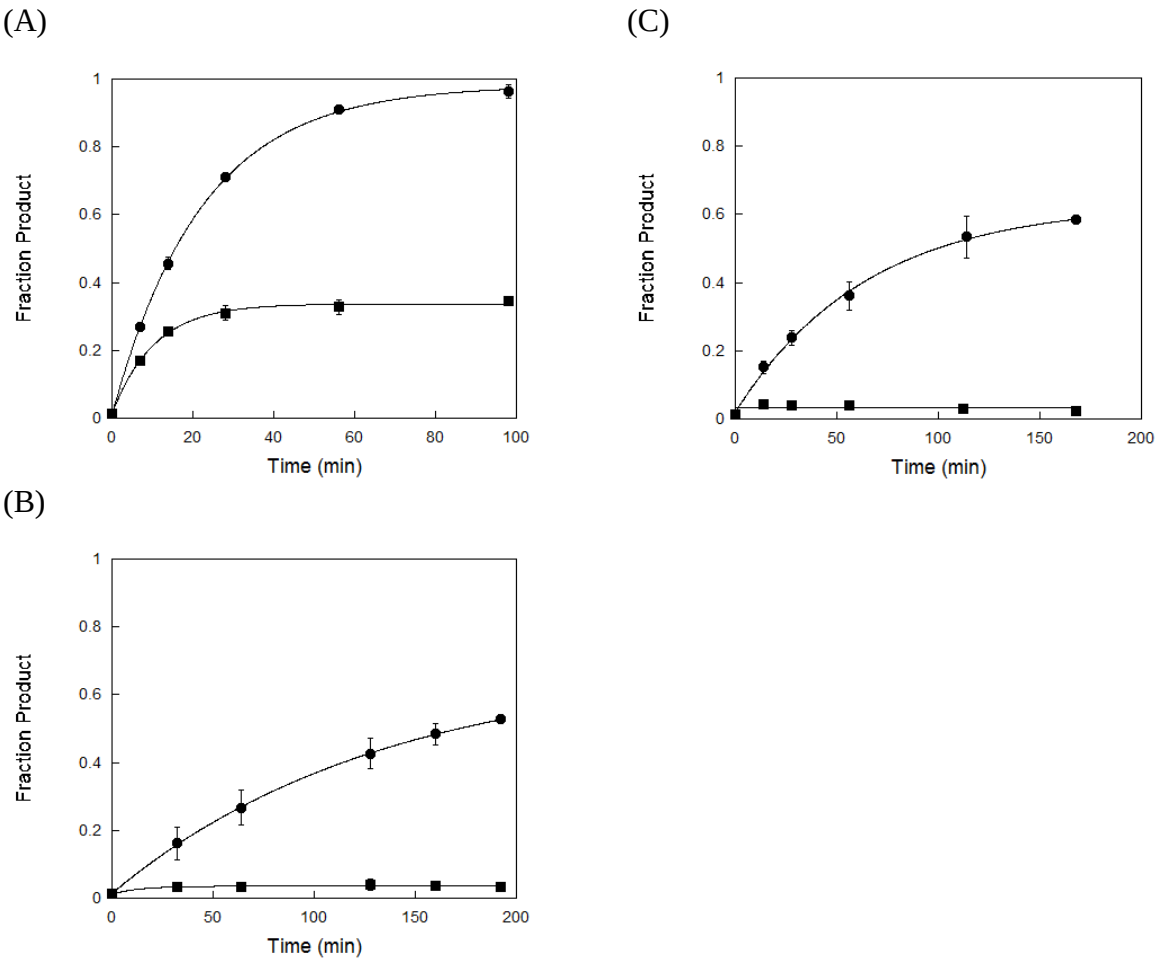



Figure 4.5. Pulse-chase experiment to measure dissociation of –TEC- DNA substrate. The partitioning between the forward reaction (base excision) and substrate dissociation were measured for wildtype (A), Y127W (B), and Y159W (C) AAG. Either wildtype of mutant AAG was mixed with FAM-labeled –TEC- DNA in the presence (■) or absence (●) of excess unlabeled –TEC- DNA, as described in the Materials and Methods. Data are the average of two independent experiments.

# Discussion

## Comparison of Wildtype and W Mutant AAG

We examined the functional consequences of mutating each of the tyrosine residues in the AAG active site pocket. Tryptophan was chosen as a conservative replacement with the intent to be able to directly monitor conformational changes and binding of base lesions in the active site. Although initial experiments showed robust

activity towards an εA lesion, a more complete analysis demonstrated that these mutants severely impair the equilibrium for flipping by AAG.

Tryptophan mutations at Y159 and Y127 each have small overall effects on excision of εA as compared with wildtype AAG (2.7- and 4-fold, respectively). However, large effects from these mutations are seen when other kinetic parameters are investigated. For Y127W AAG, slower flipping kinetics are observed. The rate constant for the flipping step is decreased at least 3-fold. From the pulse-chase data, we know that the unflipping rate constant is increased by at least 10-fold. This leads to the equilibrium constant for flipping being destabilized as compared to wildtype AAG ($\leq 50$ compared to 1300). Y159W AAG, on the other hand, exhibits a faster observed rate constant for the flipping step than wildtype AAG. This could be explained by a faster flipping rate constant or by a faster unflipping rate constant or some combination of the two. The pulse-chase data shows at least a 10-fold increase in the unflipping rate constant for this mutant as well. The result of the changes in these rate constants leads to destabilization of the equilibrium constant for flipping ($\leq 470$ compared to 1300 for wildtype AAG). It is likely that the flipping equilibrium constant is still favorable for these two mutants and the εA lesion since there is not a great effect on the chemical step. This is consistent with the idea that εA is easier to flip than many other lesions that AAG recognizes since it does not hydrogen bond with the opposing base.

These experiments show that there are significant effects on the rate constants and equilibrium constants for flipping of εA when either active site tyrosine is mutated to tryptophan. These effects are seen even though εA is expected to be easy to flip due to its lack of base pairing with the opposing base. It is expected that experiments with the

mutant proteins and stably base paired lesions will show larger effects. Indeed, in preliminary experiments, these mutations have large effects on excision of stably paired lesions, such as inosine and 7-methylguanine. It is likely that the destabilization of the equilibrium constant for nucleotide flipping in the AAG mutants can explain why the excision of other lesions are more strongly influenced by these mutations. This suggests that these mutations will have limited utility in studying other AAG substrates in the future.

**Comparison of Fluorescent Signals**

Utilizing the tryptophan reporters allows the occupancy of AAG active site pocket to be monitored at the same time as εA fluorescence. In these experiments, the identical rate constants for each fluorescent signal supports the previous assignment that the quenching of εA fluorescence corresponds to εA flipping fully into the AAG active site pocket (*5*). The initial recognition complex also shows detectable changes in fluorescence of these active site tryptophan residues. This suggests that there may be conformational changes in AAG that accompany either a nonspecific or initial recognition complex which occurs before the fully flipped complex. Nevertheless, the observation that tryptophan and εA fluorescence are tightly correlated suggests that the intermediate seen in the stopped-flow experiment changes the environment of both fluorophores. This is likely to be due to a partially flipped complex, where the εA fluorescence is increased when it is no longer stacked in the DNA but not yet stacked in the active site of AAG and the tryptophan would be partially quenched due to either interaction with the εA or a change in the protein conformation that changes the environment of the tryptophan. The strongly correlated fluorescence signals also strengthen the previous assignment that the

most highly quenched intermediate is the flipped out complex. This is likely due to the stacking of the two fluorophores against each other in the active site, each one quenching the fluorescence of the other.

**Functional Insights into Active Site Tyrosines**

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 invariantly 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 have been identified (*22-23*). 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.

From the data above, the mutation at Y127W slows the rate of flipping. This could be due to the larger bulk of the tryptophan hindering flipping of εA. The unflipping is also increased by at least 10-fold compared to wildtype AAG, which suggests that the bulk of tryptophan also causes εA to be less stabile in the active site pocket.

The mutation at Y159W appears to accelerate the rate constant for nucleotide flipping. This could be due to a 5-fold effect on flipping, a 6000-fold effect on unflipping, or a combination of increase in the two. It is not clear how replacing the tyrosine with tryptophan would increase the flipping rate constant, but it is likely that the unflipping rate constant is increased due to the extra bulk of the tryptophan in the active site that

would perturb the εA lesion that is bound there. It is not possible to draw further conclusions about which rate constant is increased unless more experiments are done.

In order to draw more specific conclusions about the effects of the tryptophan mutations on the flipping and unflipping rate constants, dissociation experiments would need to be done to measure the rate constant for unflipping directly. In these experiments, the dissociation would be measured in the stopped-flow instrument utilizing a pulse-chase type scheme to capture the protein after it dissociates from the labeled DNA. The experiments would be informative if two phases were seen for dissociation of AAG from the labeled DNA. A faster phase would be due to the protein falling off from a nonspecifically bound complex and a slower phase would be due to the protein-DNA complex first unflipping then AAG dissociating from the DNA. Measuring two phases would allow for the determination of the unflipping rate constant. Taking this value along with the observed flipping rate constant obtained in the stopped-flow binding experiments described above, the flipping rate constant could be calculated. If both the forward and reverse rate constants for flipping were determined, the functional effects of the tyrosine to tryptophan mutations could be better understood.

# References


1. Wyatt, M. D., Allan, J. M., Lau, A. Y., Ellenberger, T. E., and Samson, L. D. (1999) 3-methyladenine DNA glycosylases: structure, function, and biological importance, *Bioessays 21*, 668-676.

2. McCullough, A. K., Dodson, M. L., and Lloyd, R. S. (1999) Initiation of base excision repair: glycosylase mechanisms and structures, *Annu Rev Biochem 68*, 255-285.

3. Lau, A. Y., Scharer, O. D., Samson, L., Verdine, G. L., and Ellenberger, T. (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision, *Cell 95*, 249-258.

4. Lau, A. Y., Wyatt, M. D., Glassner, B. J., Samson, L. D., and Ellenberger, T. (2000) Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG, *Proc Natl Acad Sci U S A 97*, 13573-13578.

5. Wolfe, A. E., and O'Brien, P. J. (2009) Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by human alkyladenine DNA glycosylase, *Biochemistry 48*, 11357-11369.

6. Kuznetsov, N. A., Koval, V. V., Zharkov, D. O., Nevinsky, G. A., Douglas, K. T., and Fedorova, O. S. (2005) Kinetics of substrate recognition and cleavage by human 8-oxoguanine-DNA glycosylase, *Nucleic Acids Res 33*, 3919-3931.

7. Kuznetsov, N. A., Koval, V. V., Zharkov, D. O., Vorobjev, Y. N., Nevinsky, G. A., Douglas, K. T., and Fedorova, O. S. (2007) Pre-steady-state kinetic study of substrate specificity of Escherichia coli formamidopyrimidine--DNA glycosylase, *Biochemistry 46*, 424-435.

8. Stivers, J. T., Pankiewicz, K. W., and Watanabe, K. A. (1999) Kinetic mechanism of damage site recognition and uracil flipping by Escherichia coli uracil DNA glycosylase, *Biochemistry 38*, 952-963.

9. Wong, I., Lundquist, A. J., Bernards, A. S., and Mosbaugh, D. W. (2002) Presteady-state analysis of a single catalytic turnover by Escherichia coli uracil-DNA glycosylase reveals a "pinch-pull-push" mechanism, *J Biol Chem 277*, 19424-19432.

10. Chelli, R., Gervasio, F. L., Procacci, P., and Schettino, V. (2002) Stacking and T-shape competition in aromatic-aromatic amino acid interactions, *J Am Chem Soc 124*, 6133-6143.

11. Bordo, D., and Argos, P. (1991) Suggestions for "safe" residue substitutions in site-directed mutagenesis, *J Mol Biol 217*, 721-729.

12. Chen, C. Y., Guo, H. H., Shah, D., Blank, A., Samson, L. D., and Loeb, L. A. (2008) Substrate binding pocket residues of human alkyladenine-DNA glycosylase critical for methylating agent survival, *DNA Repair (Amst) 7*, 1731-1745.

13. O'Brien, P. J., and Ellenberger, T. (2003) Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines, *Biochemistry 42*, 12418-12429.

14. Hedglin, M., and O'Brien, P. J. (2008) Human alkyladenine DNA glycosylase employs a processive search for DNA damage, *Biochemistry 47*, 11434-11445.

15. Secrist, J. A., 3rd, Barrio, J. R., Leonard, N. J., and Weber, G. (1972) Fluorescent modification of adenosine-containing coenzymes. Biological activities and spectroscopic properties, *Biochemistry 11*, 3499-3506.

16. Biswas, T., Clos Ii, L. J., SantaLucia Jr, J., Mitra, S., and Roy, R. (2002) Binding of Specific DNA Base-pair Mismatches by N-Methylpurine-DNA Glycosylase and Its Implication in Initial Damage Recognition, *Journal of Molecular Biology 320*, 503-513.

17. Abner, C. W., Lau, A. Y., Ellenberger, T., and Bloom, L. B. (2001) Base Excision and DNA Binding Activities of Human Alkyladenine DNA Glycosylase Are Sensitive to the Base Paired with a Lesion, *Journal of Biological Chemistry 276*, 13379-13387.

18. Fersht, A. (1999) *Structure and mechanism in protein science : a guide to enzyme catalysis and protein folding*, W.H. Freeman, New York.

19. Hsieh, J., Walker, S. C., Fierke, C. A., and Engelke, D. R. (2009) Pre-tRNA turnover catalyzed by the yeast nuclear RNase P holoenzyme is limited by product release, *RNA 15*, 224-234.

20. Baldwin, M. R., and O'Brien, P. J. (2009) Human AP Endonuclease I Stimulates Multiple-Turnover Base Excision by Alkyladenine DNA Glycosylase, *Biochemistry 48*, 6022-6033.

21. O'Connor, T. R. (1993) Purification and characterization of human 3-methyladenine-DNA glycosylase, *Nucleic Acids Res 21*, 5561-5569.

22. Guo, H. H., Choe, J., and Loeb, L. A. (2004) Protein tolerance to random amino acid change, *Proc Natl Acad Sci U S A 101*, 9205-9210.

23. Chen, C.-Y., Guo, H., Shah, D., Blank, A., Samson, L., and Loeb, L. (2008) Substrate binding pocket residues of human alkyladenine-DNA glycosylase critical for methylating agent survival, *DNA repair 7*, 1731-1745.

# Chapter 5

# Preliminary NMR Studies to Characterize Changes in Conformation and Dynamics of AAG[4]

The repair of single base damage is an important process within the cell because failure to repair it will lead to mutations that cause cancers and apoptosis. This type of repair is achieved through the base excision repair pathway (*1-2*). The first step of this pathway is for a DNA glycosylase to find the site of damage within the genome and excise the damaged base. One of these enzymes, alkyladenine DNA glycosylase (AAG) has been well studied. The crystal structure was determined (*3-4*), which is very informative but static and it doesn't provide any information about the motions of the protein. The kinetics have also been studied (*5-7*) which provides a deeper understanding of the mechanism by which AAG excises damaged bases. This information is limited in its ability to describe the movements of the individual amino acids in the protein and the time scales during which they occur. In order to get this kind of information, studies of the dynamics of AAG must be performed.

Understanding the dynamics of AAG will help to build on the information that is already known about the enzyme. It will allow a deeper understanding of the molecular motions that occur as AAG binds to a DNA oligonucleotide. Specifically, it is possible to measure the motions of individual AAG residues and see what timescales these motions

[4] Special thanks to Evgenia Nikolova and Hashim Al-Hashimi, who we collaborated with on this project. Abigail Wolfe prepared the protein and DNA samples for the NMR experiments. Evgenia ran the NMR experiments and analyzed the data with some help from Abigail Wolfe.

occur on. The initial step of the base excision pathway is for AAG to recognize a damaged base. This is achieved by utilizing a base flipping mechanism. From the crystal structure, several key residues can be identified in the AAG active site that likely play a role in this base flipping mechanism. It would be very interesting to get dynamics information about these residues to see if and how the movements change when AAG goes from free in solution to bound to either undamaged or damaged DNA oligonucleotides.

An excellent technique to address this question of dynamics is NMR. Many dynamics experiments have been done with NMR to address the molecular motions of proteins (*8*-*9*). A wide range of timescales (picoseconds to milliseconds) can be examined using different NMR experiments. We were interested specifically in doing residual dipolar coupling (RDC) experiments to probe the motions of AAG, focusing on the nanosecond to microsecond timescale (*10*-*12*).

Since NMR experiments have not been performed for AAG before, a protocol for expression of AAG in minimal media was developed to allow for $^{15}N$ labeling. A purification protocol had to be optimized as well to increase the yield of the protein. Preliminary HSQC NMR experiments were done to assess the spectrum of AAG in solution as well as determine if there were changes to the spectrum when AAG is bound to DNA.

# Materials and Methods

## Preparation of Uniformly $^{15}N$ Labeled E12Q Δ80 AAG

A plasmid containing the E125Q Δ80 AAG gene that can be induced by IPTG (*5*) was transformed into a BL21(DE3) strain of *E. coli* that also contained an expression plasmid for GroEL/GroES genes also inducible by IPTG (*13*). To uniformly $^{15}N$ label the expressed protein, the cells were grown in M9 minimal media containing $^{15}NH_4Cl$ as the sole source of nitrogen (*14*). Once the $OD_{600}$ reached 0.8, expression was induced with 0.1 mM IPTG at 25 °C for ~36 hours. Cells were harvested, resuspended in lysis buffer (20 mM KPi, pH 7.4, 1 M NaCl, 10% glycerol, 0.1 % NP-40 and 5 mM BME) then frozen at -80 °C.

The protein was then purified according to the following protocol. All steps were performed at 4 °C unless otherwise indicated. The cells were thawed and protease inhibitors were added (1 mM PMSF and the recommended amount of a premade mixture of inhibitors from Sigma [containing AEBSF, E-64, bestatin, leupeptin, aprotinin, and sodium EDTA]) to prevent cleavage of AAG. Cells were lysed with 4 passes through a cell homogenizer. The lysate was cleared by centrifugation for 30 minutes at 38,000 x g. The DNA was removed from the supernatant by the addition of PEI, where 0.2% PEI was added then mixed for 30 minutes. The supernatant was cleared by centrifugation as above then loaded onto a 5 mL HisTrap Nickel column, which was run at 2 mL/minute. Once AAG was loaded, the column was washed with 10 column volumes of lysis buffer(see above) followed by 10 column volumes of Buffer A (20 mM KPi, pH 7.4, 150 mM NaCl, 10% glycerol, 5 mM BME). Another wash of 6 column volumes of 3% Buffer B (20 mM KPi, pH 7.4, 150 mM NaCl, 10% glycerol, 500 mM imidizole, 5 mM BME) diluted with

Buffer A was done. AAG was eluted by running a linear gradient from 3-100% Buffer B (diluted with Buffer A) over 3 column volumes followed by washing with 8 column volumes of Buffer B. The fractions containing AAG (fractions 48-83; 1mL each) were pooled and 5 mM BME and 1 mM EDTA were added. 0.2 mg/mL of rhinovirus 3C protease was added and incubated at 16 °C overnight to cleave the His tag from the AAG protein. The solution containing AAG was diluted 4-fold with Buffer C (20 mM KPi, pH 7.4, 100 mM NaCl, 10% glycerol, 5 mM BME, 1 mM EDTA) to decrease the ionic strength. Some protein precipitated during the dilution step, so the supernatant was cleared by centrifugation as above. In order to remove residual nucleic acids and AAG-DNA complexes that were not precipitated in the PEI precipitation, AAG was passed through a 5 mL Q column. Some of the protein remained bound to the Q column, presumably through interactions with the DNA. From there, the protein solution was loaded onto a 5 mL S column. The column was run at 2 mL/minute and was washed with 8 column volumes of Buffer C. AAG was eluted with a linear gradient of 0-50% Buffer D (20 mM KPi, pH 7.4, 2 M NaCl, 10% glycerol, 5 mM BME, 1 mM EDTA) diluted with Buffer C over 8 column volumes and a wash of 8 column volumes of Buffer D. The fractions containing AAG (fractions 17-22; 3mL each) were pooled and dialyzed into storage buffer (50 mM NaHEPES, pH 7.5, 100 mM NaCl, 10% glycerol, 1 mM EDTA, 1 mM DTT). The concentration was determined by UV absorbance, using the calculated extinction coefficient of $2.5 \times 10^4$ $M^{-1}$ $cm^{-1}$. AAG was then concentrated to 276 μM and stored at -80 °C. The yield from this preparation of $^{15}N$ AAG was 11.5 mg/L of original culture.

**Preparation of DNA Oligonucleotides**

5'phospho 2'deoxyethenoAMP (εAMP) was purchased from Sigma. The 13mer DNA oligonucleotides used in these experiments were obtained from commercial sources. The sequences are in Scheme 5.1. The A·T oligonucleotides, containing only normal deoxynucleotides, were synthesized and deprotected with standard phosphoramidite chemistry. These oligonucleotides were not purified prior to the annealing step.

| | |
|---|---|
| εA·T | 5′- GACATG**E**TTGCCT |
| | 3′-ACTGTACTAACGG |
| A·T | 5′- GACATG**A**TTGCCT |
| | 3′-ACTGTACTAACGG |

Scheme 5.1

Those oligonucleotides that contained a single εA lesion were synthesized and deprotected using Ultra-Mild chemistry according to the supplier's recommendations. The εA-containing oligonucleotide was synthesized by the W. M. Keck Facility at Yale University (New Haven, CT). The oligonucleotides were desalted using SephadexG-25 and purified using denaturing polyacrylamide gel electrophoresis, as previously described (*15*). The concentrations of the single-stranded oligonucleotides were determined from the absorbance at 260 nm, using the calculated extinction coefficients for all oligonucleotides except the one containing εA, where 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 versus that of A

(*16*). The two oligonucleotide strands were annealed at a 1:1 ratio by being heated to 90 °C and cooled slowly to 4 °C.

**Preparation of Samples for NMR Experiments**

$^{15}$N AAG was quickly thawed then dialyzed into NMR buffer (10 mM NaMES, pH 6.1, 100 mM NaCl, 1 mM DTT, 0.1 mM EDTA). 300 μL samples were made for each NMR experiment. The sample of AAG free in solution was prepared by mixing 270 μL of 255 μM AAG with 30 μL $D_2O$. The sample of AAG bound to εA·T DNA was made by mixing $^{15}$N AAG and εA·T DNA at a 1:1 ratio to a final concentration of 250 μM then $D_2O$ was added to 10%. Similarly, the sample of AAG bound to A·T DNA was made by mixing $^{15}$N AAG and A·T DNA at a 1:1.1 ratio to a final concentration of 196 μM then $D_2O$ was added to 10%. The final sample of AAG bound to εAMP was prepared by mixing 196 μM AAG and 1M εAMP followed by the addition of $D_2O$ to 10%.

**HSQC NMR Experiments**

$^{1}$H, $^{15}$N HSQC correlation maps of uniformly $^{15}$N labeled AAG in the free and bound form (see above for concentrations) were collected using a 30 ppm spectral width and 300 indirect points. Spectra were acquired with either 48 (free AAG, AAG bound to εA·T and AAG bound to εAMP) or 128 (AAG bound to A·T) scans for a total of ~ 1 day. All NMR experiments were performed at 25 °C using a Bruker Avance 600 MHz NMR spectrometer equipped with a 5 mm triple-resonance cryogenic probe. Data were processed and analyzed using nmrPipe (*17*).

# Results and Discussion

## Preparation of NMR Samples

For NMR, we need to use a catalytically inactive mutant, E125Q, and I found that this AAG mutant behaves very different from WT AAG during both expression and purification. Additionally, uniformly $^{15}N$ labeled AAG was needed to perform the NMR experiments. In order to label the protein at each nitrogen, AAG was expressed in minimal media supplemented with $^{15}NH_4Cl$. Optimal expression conditions were worked out by expressing AAG in minimal media without the label so that this valuable resource would not be wasted. The regular purification conditions for wildtype AAG was tried first and conditions were changed from there (*5*).

There were several changes to the expression conditions to optimize them for expression of $^{15}N$-labeled E125Q AAG. The first was that the expression was done in M9 minimal media to allow for incorporation of $^{15}N$ into the protein, whereas typical expression of WT AAG is done in LB media. The expression of E125Q AAG was greatly enhanced by co-expressing the chaperone proteins, GroEL/GroES, but this co-expression is unnecessary for expressing WT AAG. While the best expression for both E125Q AAG and WT AAG is to induce at an OD of 0.8 and 25 °C, a much longer time after induction is needed for optimal expression of E125Q AAG (36 hours) as compared to WT AAG (~15 hours).

In addition to changing the conditions for expression, several changes in the purification protocol were necessary to optimize the yield of E125Q AAG during purification. E125Q AAG was more problematic during the purification because it binds

more tightly to genomic DNA than WT AAG does. One change to the purification protocol for E125Q AAG was to increase the NaCl concentration in the lysis buffer from 0.5 to 1 M. During the PEI precipitation, it was necessary to increase the percent PEI to 0.2% from 0.1% and increasing the mixing time from 10 to 30 minutes. These changes aid in removing more DNA that is bound to the protein. Finally, an additional step was added in order to remove any DNA that is still bound to E125Q AAG during the purification. Before loading onto the S column, the protein solution was passed through a Q column so that any DNA in the solution would bind to the Q column. However, free E125Q AAG would flow through the Q column and bind only to the S column. This step is unnecessary for WT AAG purification because DNA is no longer present in the solution after the PEI precipitation.

While there are not many significant changes to the protocols, these changes do allow for much greater yield and purity of E125Q AAG, which is necessary when using expensive labeling reagents as in NMR experiments.

**Stability Problems with NMR Samples**

When running the NMR experiments on the free AAG sample, it was clear that AAG was not stable under the experimental conditions as it precipitated quickly. Although the HSQC spectrum was able to be measured, it would be desirable to find better conditions for future experiments so that AAG would remain stable throughout the experiment. Similar stability problems were seen with AAG in complex with the A·T DNA and the εAMP. Again, HSQC spectra were obtained with these samples, but better conditions (and more stable samples) would likely lead to better quality spectra in the

future. Stability problems were not seen with AAG in complex with εA·T DNA, presumably due to the fact that AAG binds strongly to the εA site.

**HSQC Spectra of AAG**

From the HSQC NMR experiments, we see that the spectrum of AAG free in solution is well resolved with ~300 peaks (Figure 5.1). There are 210 amide nitrogens in the protein as well as some nitrogens in the side chains. This is a reasonable number of peaks to observe. The fact that the spectrum is well resolved indicates that this molecule is amenable to NMR techniques.

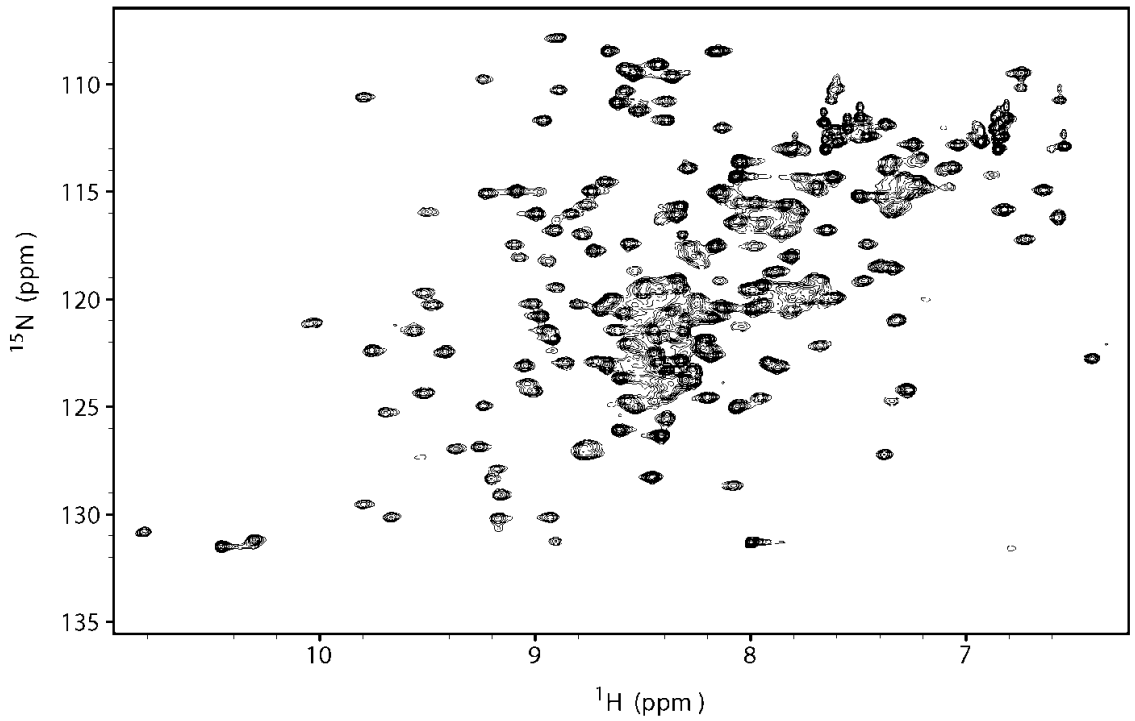



Figure 5.1. HSQC Spectrum of Free AAG. The spectrum of free AAG gives many well resolved peaks.

When AAG is bound to εA·T DNA, there are changes to the chemical shifts for most of the peaks in the spectrum (Figure 5.2 and Figure 5.3). This indicates that there is a change in the environment for many of the residues in AAG between the bound and

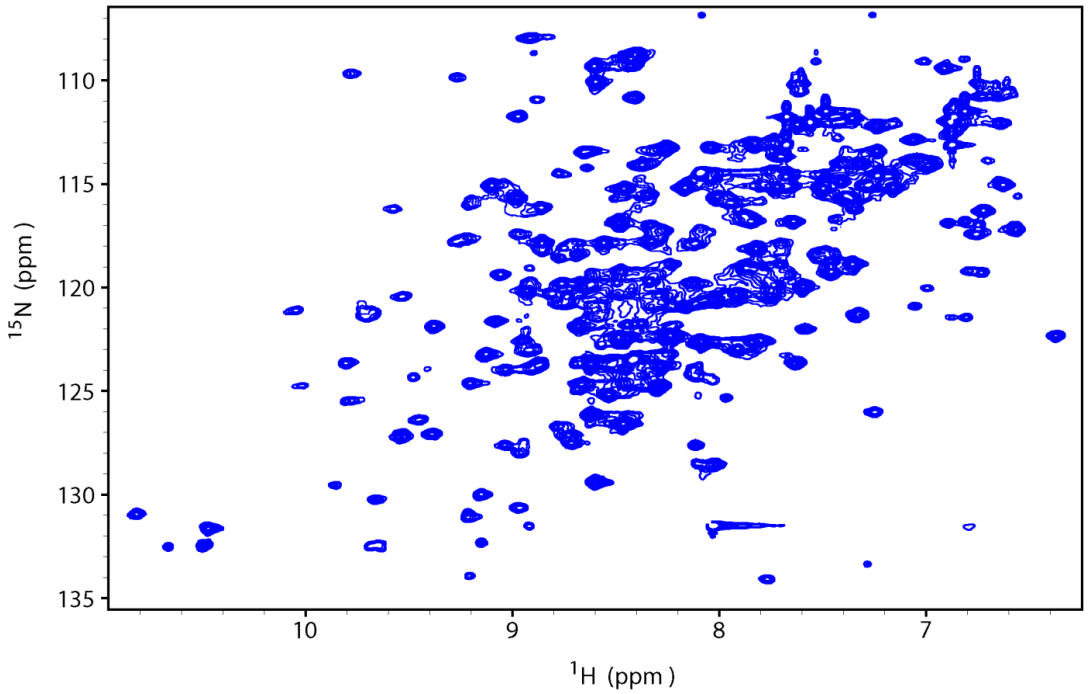



Figure 5.2. HSQC Spectrum of AAG Bound to εA·T DNA. The spectrum of AAG bound to εA·T DNA gives many well resolved peaks, but they are broader than seen in the spectrum of free AAG (Figure 5.1).

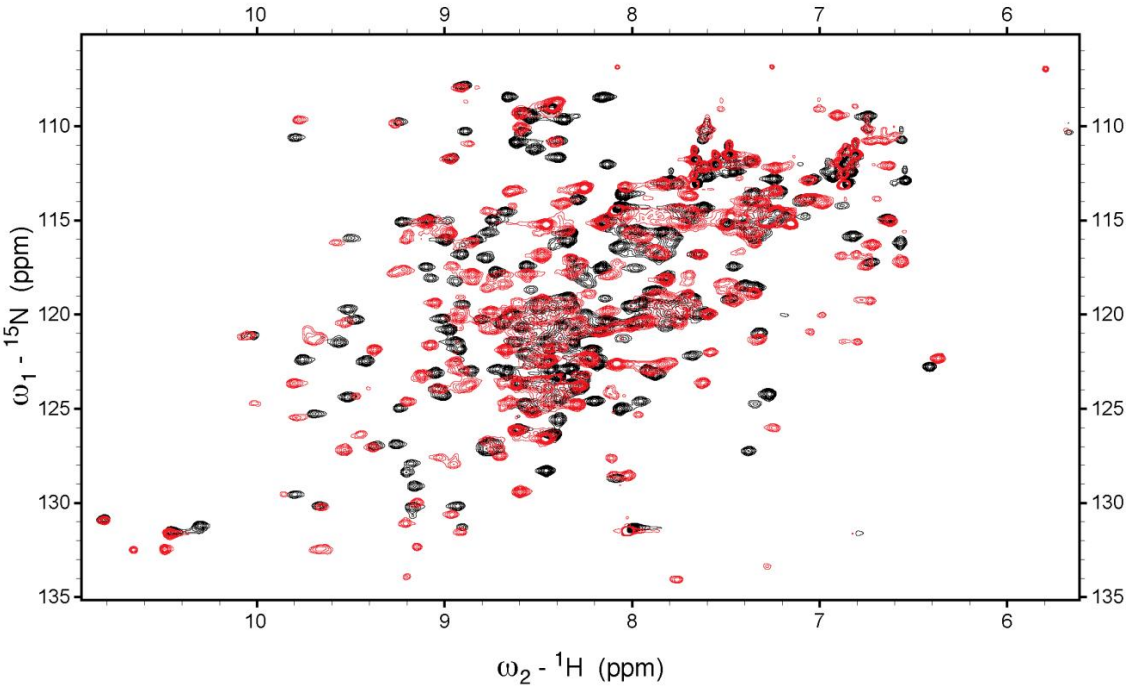



Figure 5.3. Comparison of the Spectra of Free AAG and AAG Bound to εA·T DNA. In black is the spectrum of free AAG (Figure 5.1) and in red is the spectrum of AAG bound to εA·T DNA (Figure 5.2) at a 1:1 ratio. Many of the AAG peaks are shifted between the free and bound states. Also, the peaks in the bound spectrum are visibly broader than those in the free AAG spectrum.

unbound states. This is consistent with the idea that a global shift in conformation is necessary for AAG to bind to DNA in the first place and many of the active site residues and other side chains that contact the DNA are anticipated to move in order to allow base flipping to occur. It is expected that at 1:1 ratio of AAG: εA·T DNA the protein will be specifically bound to the εA site, with the base fully flipped.

In the spectrum of AAG bound to A-T DNA, the peaks are very broad (Figure 5.4 and Figure 5.5). It is much harder to distinguish the individual peaks in this spectrum. Thus, it is difficult to correlate any peaks from this spectrum to that of free AAG. This likely indicates that there is not a single species as seen in the εA·T DNA spectrum. This would occur if AAG was bound at different sites all along the DNA strand, which is

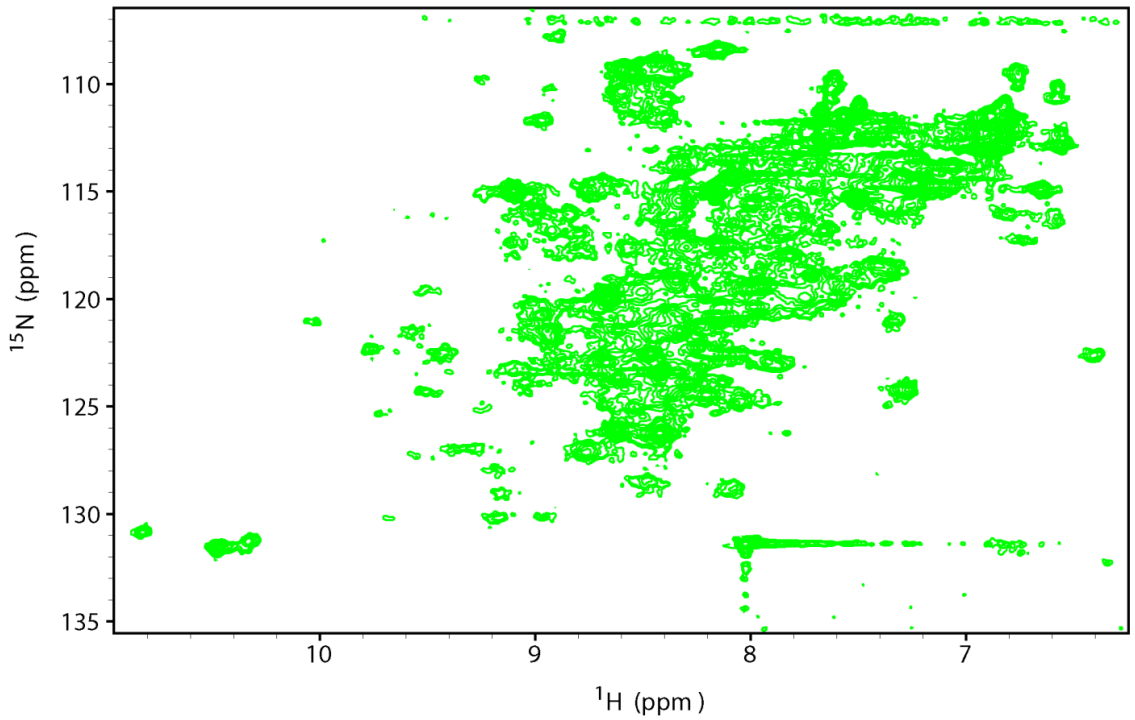



Figure 5.4. HSQC Spectrum of AAG Bound to A·T DNA. In the spectrum of AAG bound to A·T DNA, the peaks are not very well resolved and they are much broader than seen in the spectrum of free AAG (Figure 5.1).

expected when AAG is bound to undamaged DNA. It could also be due to heterogeneity in the unpurified DNA.

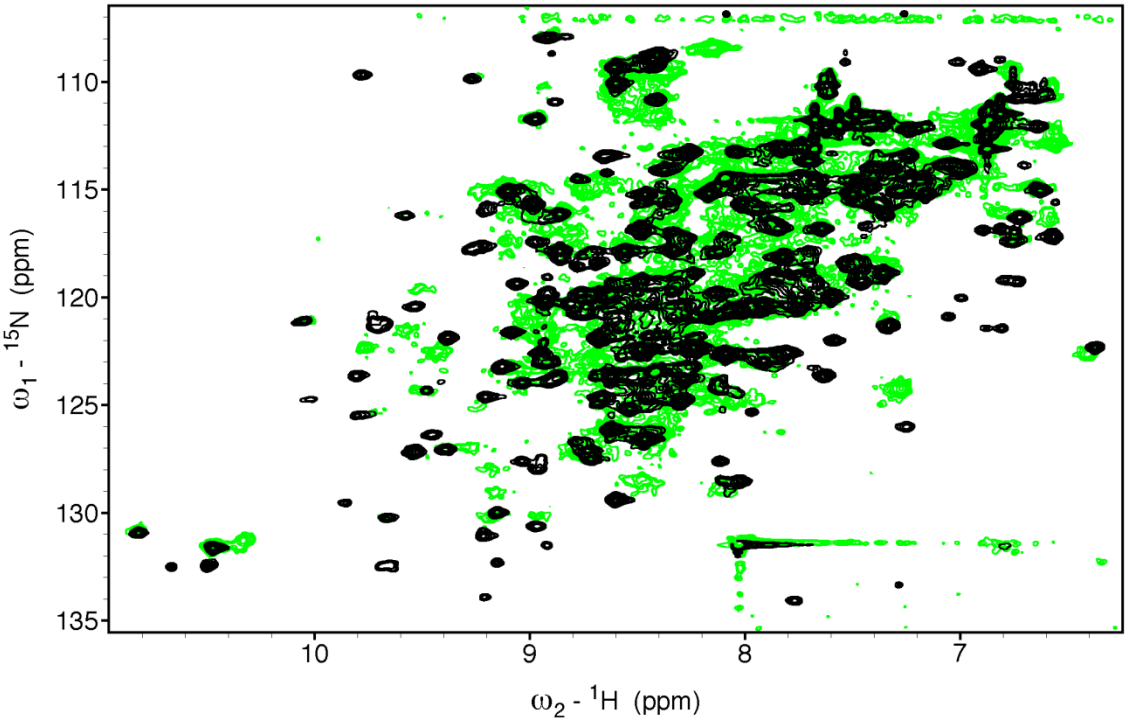



Figure 5.5. Comparison of the Spectra of AAG bound to εA·T and A·T DNA oligonucleotides. In black is the spectrum of AAG bound to εA·T DNA (Figure 5.2) and in green is the spectrum of AAG bound to A·T DNA (Figure 5.4). The spectra are similar but there is much greater broadening of the peaks in the spectrum of AAG bound to A·T DNA.

The spectrum for εAMP bound to AAG looks identical to that of AAG alone (Figure 5.6 and Figure 5.7). Not many peak shifts were expected for this complex in the first place since it is predicted to bind in the active site and only contact a few residues there but not cause a global shift in the protein as seen when DNA is bound. The lack of change in the spectrum could indicate that there is no change in the spectrum upon binding εAMP, a change does occur, but is not detected, or the nucleotide was not bound in the active site under these conditions.

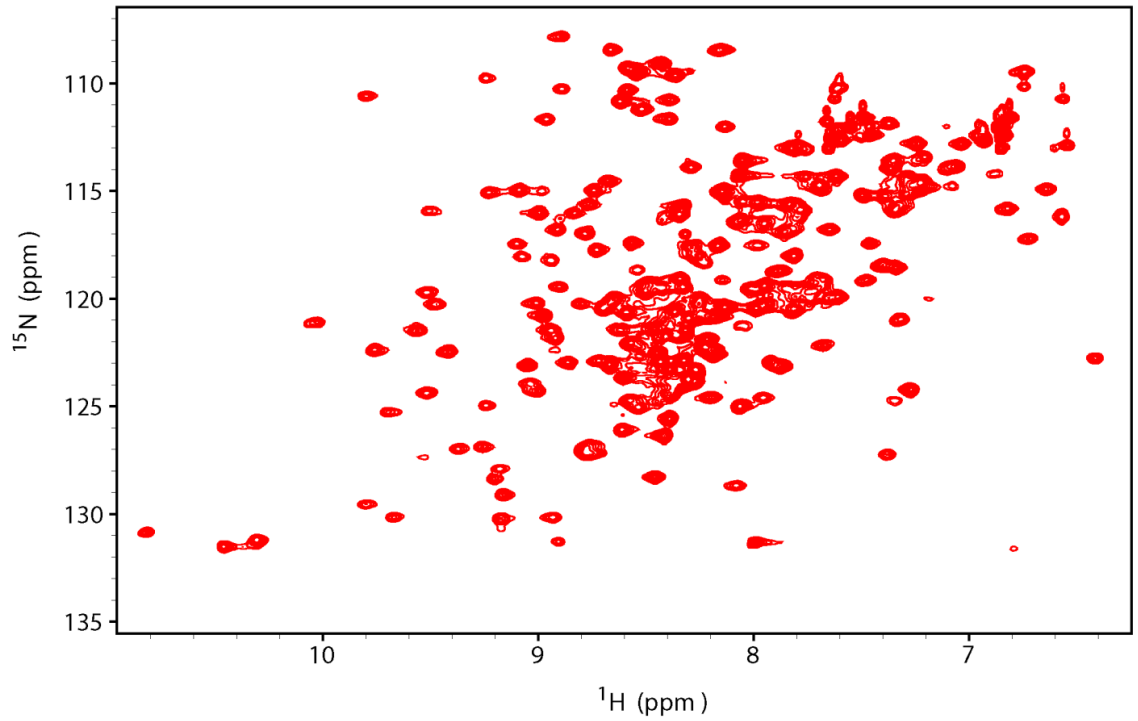



Figure 5.6. HSQC Spectrum of AAG bound to εAMP. The spectrum of AAG bound to εAMP gives many well resolved peaks and is identical to the spectrum of free AAG (Figure 5.1).

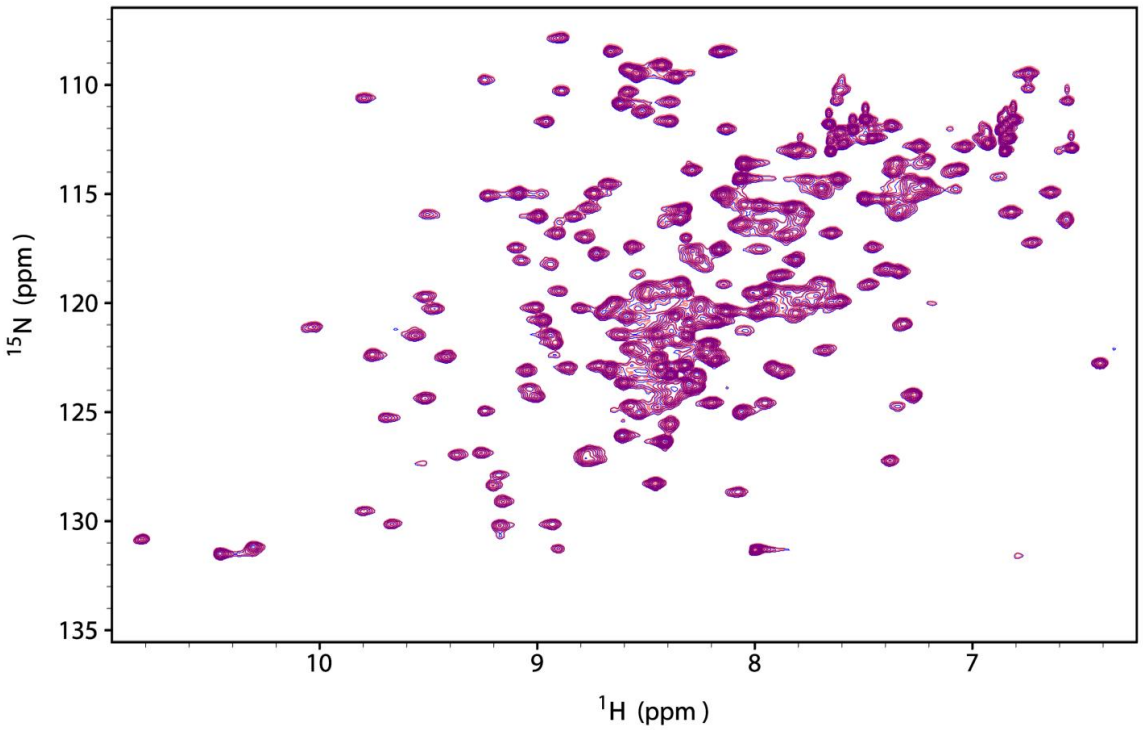



Figure 5.7. Comparison of the Spectra of Free AAG and AAG Bound to εAMP. The spectrum of free AAG (Figure 5.1) is in blue and the spectrum of AAG bound to εAMP (Figure 5.6) is in red. The plot is an overlay of the two spectra, which gives purple when the two spectra match exactly. The entire spectrum here is purple, showing that the two spectra are identical.

Since the peaks for the HSQC spectrum have not be previously assigned for AAG, it is impossible to know which peaks in the spectra correspond to which amino acids in AAG at this time. This information is imperative to determine which residues are moving when bound to different substrates. Nevertheless, the purification protocol described here and the HSQC spectra obtained will allow future NMR experiments to be done to determine the conformation of free AAG and additional DNA complexes and the timescales of motions of AAG.

## Future Directions

There are two possibilities for getting molecular information out of the NMR experiments. The first is to make differentially labeled AAG and assign the complete spectra for both the free enzyme and AAG bound to εA·T DNA. This option would require a more complicated labeling scheme and would take a relatively long time to complete the experiments and assign all of the residues. A second option would be to only label the tyrosines in AAG. This can be done by supplementing the minimal media with $^{15}$N-tyrosine and unlabeled $NH_4Cl$, which would be a simple protein stock to make. AAG only has 6 tyrosine residues. Three of them are in the active site with two directly contacting the damaged base (as seen in the crystal structure). A fourth tyrosine, Y162, is intercalated into the DNA duplex, so it expected to move when AAG forms a flipped conformation with damaged DNA. A fifth tyrosine is also close to the DNA interface, while the final tyrosine is partially buried and not near the DNA or active site. Many of these tyrosine residues would be expected to be in a different environment when AAG binds to DNA and/or when the base is flipped. These would likely be the most interesting residues to look at even if the whole AAG spectrum was assigned. This, in addition to the

relative ease of making tyrosine labeled AAG, makes it an attractive avenue that could be pursued in the future. Once AAG was labeled at the tyrosines, it would be simple to run the dynamics experiments on the 6 peaks to see what types of motions are occurring for these residues as well as their time-scales in the presence and absence of damaged and undamaged DNA. To identify individual peaks in the spectra, single mutations could be made to change each tyrosine residue to some other amino acid, such as tryptophan (see Chapter 4 for data that these AAG mutants are active). By using single mutants, it would be possible to correlate the amino acid that was no longer tyrosine with the peak that disappeared from the spectra. This way, all six peaks in the spectra could be assigned to the tyrosines in AAG. It is possible that making the mutations could change the structure of AAG, which would shift the other 5 peaks that are seen in the spectra. This would potentially make it difficult to correlate the correct peak with the residue in AAG.

# References


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

2. Almeida, K. H., and Sobol, R. W. (2007) A unified view of base excision repair: Lesion-dependent protein complexes regulated by post-translational modification, *DNA Repair 6*, 695-711.

3. Lau, A. Y., Wyatt, M. D., Glassner, B. J., Samson, L. D., and Ellenberger, T. (2000) Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG, *Proc Natl Acad Sci U S A 97*, 13573-13578.

4. Lau, A. Y., Scharer, O. D., Samson, L., Verdine, G. L., and Ellenberger, T. (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision, *Cell 95*, 249-258.

5. O'Brien, P. J., and Ellenberger, T. (2003) Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines, *Biochemistry 42*, 12418-12429.

6. O'Brien, P. J., and Ellenberger, T. (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase, *J Biol Chem 279*, 9750-9757.

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

8. Mittermaier, A. K., and Kay, L. E. (2009) Observing biological dynamics at atomic resolution using NMR, *Trends in Biochemical Sciences 34*, 601-611.

9. Grzesiek, S., and Sass, H. J. (2009) From biomolecular structure to functional understanding: new NMR developments narrow the gap, *Curr Opin Struct Biol 19*, 585-595.

10. Lange, O. F., Lakomek, N. A., Fares, C., Schroder, G. F., Walter, K. F., Becker, S., Meiler, J., Grubmuller, H., Griesinger, C., and de Groot, B. L. (2008) Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution, *Science 320*, 1471-1475.

11. Bouvignies, G., Bernado, P., Meier, S., Cho, K., Grzesiek, S., Bruschweiler, R., and Blackledge, M. (2005) Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings, *Proc Natl Acad Sci U S A 102*, 13885-13890.

12. Lakomek, N. A., Lange, O. F., Walter, K. F., Fares, C., Egger, D., Lunkenheimer, P., Meiler, J., Grubmuller, H., Becker, S., de Groot, B. L., and Griesinger, C. (2008) Residual dipolar couplings as a tool to study molecular recognition of ubiquitin, *Biochem Soc Trans 36*, 1433-1437.

13. Amrein, K. E., Takacs, B., Stieger, M., Molnos, J., Flint, N. A., and Burn, P. (1995) Purification and characterization of recombinant human p50csk protein-tyrosine kinase from an Escherichia coli expression system overproducing the bacterial chaperones GroES and GroEL, *Proc Natl Acad Sci U S A 92*, 1048-1052.

14. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular cloning : a laboratory manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

15. Hedglin, M., and O'Brien, P. J. (2008) Human alkyladenine DNA glycosylase employs a processive search for DNA damage, *Biochemistry 47*, 11434-11445.

16. Secrist, J. A., 3rd, Barrio, J. R., Leonard, N. J., and Weber, G. (1972) Fluorescent modification of adenosine-containing coenzymes. Biological activities and spectroscopic properties, *Biochemistry 11*, 3499-3506.

17. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J Biomol NMR 6*, 277-293.

# Appendix

**Description of Equations for Binding at Equimolar Concentrations of Enzyme and DNA.**

A general scheme for two-step, reversible binding is shown in Equation A.1. If the two steps can be directly distinguished, for example by different spectroscopic signals, then this can be simplified to Equation A.2, in which there is an observed rate constant for the first and second steps. The value of $k_{1,\,obs}$ is given by Equation A.3. However, at micromolar concentrations of AAG the rate of dissociation is significantly slower than the rate of association and therefore the observed rate constant is approximately equal to the rate constant for association (Equation A.4). Similarly, the observed rate constant for the second step is equal to the sum of the rate constants for the forward and reverse flipping steps (Equation A.5). As discussed in the text, the pulse-chase experiment revealed that $k_2$ is much greater than $k_{-2}$, so that this simplifies to Equation A.6.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E\cdot S \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E\cdot S' \qquad \text{(A.1)}$$

$$E + S \xrightarrow{k_{1,obs}} E\cdot S \xrightarrow{k_{2,obs}} E\cdot S' \qquad \text{(A.2)}$$

$$k_{1,\,obs} = k_1\,[E] + k_{-1} \qquad \text{(A.3)}$$

$$k_{1,\,obs} \sim k_1\,[E] \qquad \text{(A.4)}$$

$$k_{2,\,obs} = k_2 + k_{-2} \qquad \text{(A.5)}$$

$$k_{2,\,obs} \sim k_2 \qquad \text{(A.6)}$$

For an irreversible binding step, the rate of formation of the complex is given by Equation A.7 (*1*). Under conditions in which both binding partners are present at the same concentration (i.e., [AAG] = [DNA]), then Equation A.8 is the appropriate expression. As described by Fersht, integration gives a relatively simple expression (Equation A.9). This can be rearranged to solve for the concentration of the bound species as a function of time (Equation A.10). The second step of the binding reaction is a first order reaction that is described by a single exponential (Equation A.11).

$$d[E\bullet S]/dt = k_1\,[E][S] \qquad \text{(A.7)}$$

$$d[E\bullet S]/dt = k_1[E]_0^{\,2} \qquad \text{(A.8)}$$

$$1/([E]_0\text{-}[ES]) - 1/[E]_0 = k_1 t \qquad \text{(A.9)}$$

$$[E \bullet S] = \frac{[E]_0^{\,2} k_1 t}{[E]_0 k_1 t + 1} \qquad \text{(A.10)}$$

$$[E \bullet S'] = [E]^0(1 - e^{-k_{2,obs}t}) \qquad \text{(A.11)}$$

The fluorescence of the sample (F) can be considered to be a linear combination of the fluorescence of each of the species present in solution at a given time after mixing (Equation A.12). In this expression $F_{free}$ is the fluorescence of the unbound εA-DNA, $\Delta F_{E\bullet S}$ is the change in fluorescence between free DNA and the E•S complex, and $\Delta F_{E\bullet S'}$ is the change in fluorescence between the E•S and the E•S' complexes. Substitution of Equation A.12 with the time dependence for the formation of the initial E•S complex (Equation A.10) and the time dependence of the formation of E•S' (Equation A.11) gives Equation A.13 that was used in the text, in which $F_{free} = C$, $\Delta F_{E\bullet S} = Y$, and $\Delta F_{E\bullet S'} = Z$.

$$F = F_{free} + \Delta F_{E\bullet S}[E\bullet S] + \Delta F_{E\bullet S'}[E\bullet S'] \qquad (A.12)$$

$$F = C + Y\left(\frac{E_o^2 k_1 t}{1 + E_o k_1 t}\right) - Z(1 - e^{-k_{2,obs} t}) \qquad (A.13)$$

## Reference

1. Fersht, A. (1999) *Structure and Mechanism in Protein Science*, pp 200-201, W.H. Freeman and Company, New York.