Mechanistic Studies of Alkylation Repair DNA Glycosylases

by

Erin Taylor

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biological Chemistry)
in the University of Michigan
2016

Doctoral Committee:

Associate Professor Patrick J. O'Brien, Chair
Professor Carol A. Fierke
Professor Robert S. Fuller
Associate Professor Aaron C. Goldstrohm
Associate Professor Lyle A. Simmons

Dedication

This work is dedicated to those who advocate for basic science research and understand the monumental accomplishments such research can attain.

"…basic research…is a crucial component of the innovation that improves life for everyone."

-Bill Gates

Acknowledgements

Many people have contributed to this work and supported me during my time at the University of Michigan. Thank you to my mentor Pat, and all the members of the O'Brien lab for their advice and being my home for the past five years. I would especially like to thank Abigail Wolfe and Jenna Hendershot for their work detailing the mechanism of AAG, on which this thesis stands. Thank you to Preethi Kesavan, a wonderful undergraduate who took the anisotropy project by the reins and produced very lovely data. To Lyle and all the members of the Simmons Lab, thank you for all the guidance during my sabbatical and helping me start the *Bacillus* project.

Thank you to my family and friends for their constant support and understanding.
And lastly, thank you to Mark, who has made my life in Ann Arbor very special indeed.

Table of Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

1meA	N^1-methyladenine
2meC	O^2-methylcytosine
2meT	O^2-methylthymine
3meA	N^3-methyladenine
3meC	N^3-methylcytosine
3meG	N^3-methylguanine
7meG	N^7-methylguanine
AAG	alkyladenine DNA glycosylase (human) (also called MPG, N-methylpurine DNA glycosylase; ANPG, alkyl-N-purine DNA glycosylase
ABH2	human AlkB homolog 2
Ada	E. coli methyltransferase; induces adaptive response
Ada response	adaptive response
AlkA	3-methyladenine DNA glycosylase II (E. coli)
AlkB	E. coli direct repair methyltransferase
AlkC	3-methyladenine DNA glycosylase with HEAT repeats (Bacillus cereus)
APE	AP (apurinic/apyrimidic) endonuclease
B. subtilis	*Bacillus subtilis*
bAAG	B. subtilis AAG homolog
bAlkA	B. subtilis AlkA homolog

BER	base excision repair pathway
CAA	chloracetaldehyde
CTT	C-terminally tagged AlkA
dRP	deoxyribose phosphate
DY647	DyLight 647 (cyanine 5-like adduct)
E. coli	*Escherichia coli*
εA	1,N^6-ethenoadenine
εC	3,N^4-methylcytosine
εG	1,N^2-ethenoguanine
FAM	6-fluorescein
HEX	6-hexachlorofluorescein
HhH	helix-hairpin-helix
HR	homologous recombination
Hx	hypoxanthine
Kan	kanamycin
MMR	mismatch repair pathway
MMS	methyl methanesulfonate
MNNG	methylnitronitrosoguanidine
NER	nucleotide excision repair pathwy
NHEJ	non-homologous end-joining
O6meG	O^6-methylguanine
SAM	S-adenosylmethionine
Tag	3-methyladenine DNA glycosylase I

WT	wild type
yfjP	2nd *B. subtilis* AlkA homolog
yhaZ	*B. subtilis* AlkC homolog
yxlJ	B. subtilis AAG homolog gene (referred to as *aag* in Chapter 5 results)

Abstract

Alkylating agents are pervasive in the environment and are endogenously produced. DNA nucleobases are highly susceptible to alkylation damage, and resulting lesions have the propensity to be mutagenic and cytotoxic. The base excision repair pathway is primarily responsible for excising alkylated lesions and is initiated by a DNA glycosylase. Glycosylases search the genome for damage, flip the damaged nucleotide into an active site pocket, and hydrolyze the *N*-glycosidic bond between the base and sugar. The *E. coli* AlkA and human AAG have evolved independently to protect cells from a wide range of alkylated lesions such as N^3-methyladenine, N^7-methylguanine and the etheno adduct 1,N^6-ethenoadenine (εA). AlkA and AAG also act on oxidatively damaged bases such as hypoxanthine (Hx). This work characterizes and compares AlkA and AAG to better understand the similarities and differences between these two structurally distinct enzymes that have evolved to repair a similar set of substrates.

The kinetic and thermodynamic framework for the activity of *E. coli* AlkA was determined for εA-DNA. In vitro kinetic assays showed that AlkA binds rapidly and reversibly to εA prior to a slow chemical excision step. The kinetic mechanism of AlkA could be directly compared to the previously determined mechanism of AAG with εA-DNA. Overall, AlkA is a less efficient enzyme than AAG because while it exhibits identical rates of εA excision, AlkA does not bind εA as tightly as AAG. As the expression of AlkA is induced in *E. coli* as part of a response to alkylation damage, the lower efficiency may not be a burden for cellular protection from damaging agents.

Previous work with AAG has shown the protein employs facilitated diffusion to find sites of damage, and therefore specific and nonspecific DNA is important for efficient lesion repair. We used fluorescence anisotropy and electrophoretic mobility shift assays to characterize binding stoichiometry and affinity of AAG for specific and nonspecific DNA. Complex biphasic anisotropy signals were observed for titration of a specific 5′ 6-fluorescein (FAM)-labeled εA substrate with AAG that could be explained by cooperative binding of an AAG dimer. However, 1:1 complexes were observed for specific εA binding when native gel shifts and εA quenching were measured. The large changes in anisotropy and FAM fluorescence indicate that AAG binds DNA preferentially near the FAM lesion site. We compared fluorescein and cyanine-like DyLight 647 dyes to characterize the affinity of AAG for the FAM site (low nanomolar) and nonspecific sites (low micromolar), showing that although AAG binds as a 1:1 complex to a specific site, nonspecific proteins can bind at high density. This work demonstrates the importance of considering protein-fluorophore interactions when analyzing anisotropy data of nucleic acid binding proteins.

Homologs of AlkA and AAG present in *Bacillus subtilis* were also explored because of the rarity of having both glycosylases in the same organism. To determine the contribution of bAlkA and bAAG for methylation repair, in vivo MMS sensitivity assays were performed with single and double deletion strains bAlkA and bAAG. The results indicated that bAlkA was vital for repair of methylated cytotoxic lesions, with bAAG providing some backup activity. To ascertain the substrate specificities of bAlkA and bAAG, single turnover kinetics assays were done with the purified proteins and a variety of base lesions. bAlkA and bAAG exhibited glycosylase activity towards εA and Hx as well as N^1-methyladenine (1meA) and 1,N^2-ethenoguanine (εG) which are known to be substrates of the direct reversal enzyme AlkB in *E.*

coli. It is interesting that *B. subtilis* does not contain an AlkB homolog and thus bAlkA and bAAG have expanded their substrate ranges in order to protect cells from damage. This work lays the groundwork for understanding alkylation repair in *Bacillus subtilis*.

Chapter I

Introduction to DNA Alkylation and DNA Repair Pathways

DNA damaging agents are ubiquitous and cells require multiple pathways to maintain the stability of the genome. This dissertation explores alkylation damage repair by two DNA glycosylases responsible for protection from alkylation stress. The prokaryotic AlkA and eukaryotic AAG glycosylases have independently evolved to remove a similar set of alkylated bases that can be cytotoxic and mutagenic. This work compares and contrasts the kinetic mechanisms of AlkA and AAG which differ in structure and substrate recognition. Substrate binding by AAG is further investigated using fluorescence anisotropy methods which show interesting phenomena dependent on the fluorophore identity. The mechanisms of the two proteins are also explored in *Bacillus subtilis* in which both AlkA and AAG homologs are present. The substrate specificities of the *Bacillus* proteins overlap and are expanded to include additional substrates. This thesis thoroughly characterizes two structurally distinct, yet mechanistically similar glycosylases; thus expanding our knowledge of DNA repair enzymes and pathways.

Genome Instability and Alkylation Damage

Throughout their lifetimes, cells are exposed to agents that have the propensity to damage nucleic acids. Sources of such agents vary widely with some agents being introduced exogenously to cells from the environment and others being produced in cells endogenously.[1-4]

The relentless barrage of damaging agents causes instability in the genome, leading to extremely detrimental effects such as cancer and cell death.[4] In order to avoid such drastic results, it is vital for cells to combat and repair genomic damage.

DNA is known to be very stable molecule, but damaging agents can react with multiple sites including breaking the deoxyribose-phosphate backbone, directly damaging the nucleobases and producing crosslinks between bases. As nucleobases contain a large amount of highly reactive sites, direct damage to bases is most common and occurs on a biologically relevant timescale. It has been estimated that normal human cells experience 10^4 base lesions per cell per day, meaning one out of 300,000 bases is damaged in a healthy human cell.[4] While seemingly insignificant, even a slight modification to a base may affect hydrogen bonding and specificity of DNA and RNA polymerases. Thus, base lesions can halt replication and transcription machinery or cause mutations, resulting in injurious effects even at low amounts.

One common class of base modification is alkylation. Alkylating agents can react with almost all the oxygen and nitrogen atoms in the canonical bases via Sn1 or Sn2 reactions, producing a variety of lesions that range in toxicity. Alkylating agents are found naturally in the environment and are produced in cells. For the simple case of methylation for instance, agents found exogenously include methyl halides that are produced by decaying vegetation and biomass burning, as well as chemicals from cigarette smoke, chemotherapeutic treatments, food and industry.[3] Endogenous methylating agents include metabolites produced by the nitrosation of amines in bacteria, and the abundant methyl donor *S*-adenosylmethionine (SAM) which can spontaneously react with nucleobases.[3,4] Some of the most commonly produced methylated lesions are N^7-methylguanine (7meG), N^3-methyladenine (3meA), O^6-methylguanine (O^6meG) and the backbone methylphosphotriesters (Figure 1-1). Additionally, the N^1-methyladenine

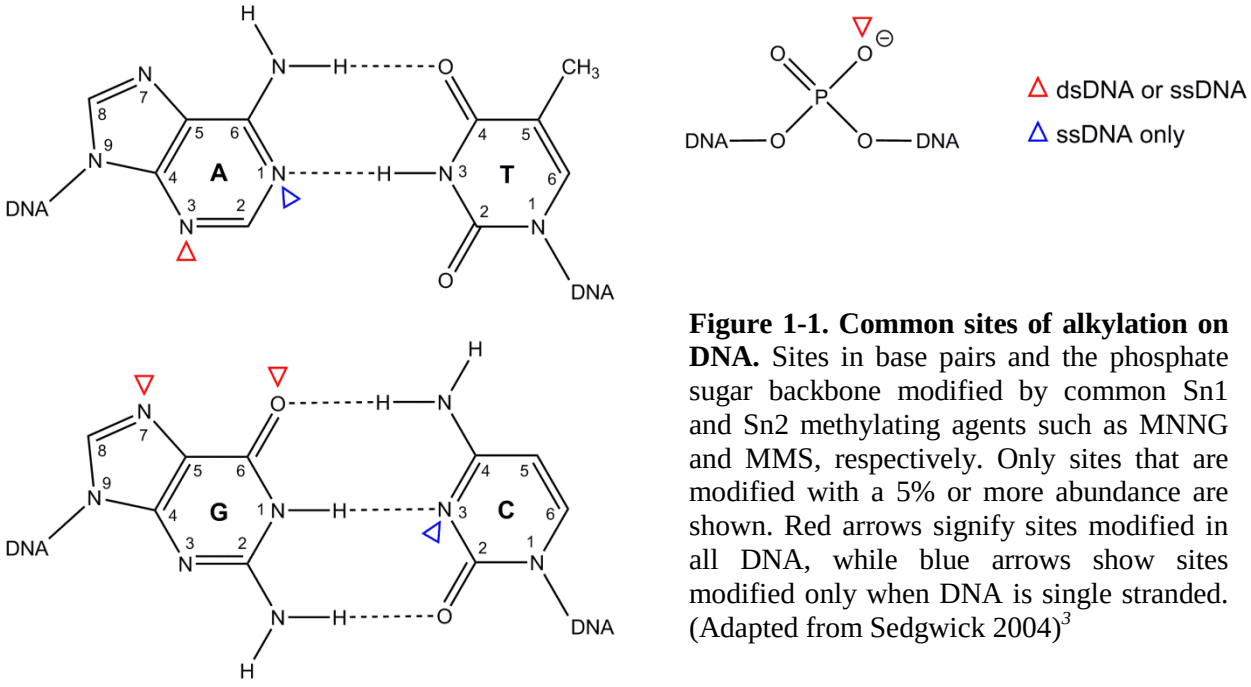


Figure 1-1. Common sites of alkylation on DNA. Sites in base pairs and the phosphate sugar backbone modified by common Sn1 and Sn2 methylating agents such as MNNG and MMS, respectively. Only sites that are modified with a 5% or more abundance are shown. Red arrows signify sites modified in all DNA, while blue arrows show sites modified only when DNA is single stranded. (Adapted from Sedgwick 2004)[3]

(1meA) and N^3-methylcytosine (3meC) are produced, although this is assumed to occur only in single stranded DNA. These lesions range in toxicity with 3meA, 1meA and 3meC being cytotoxic, and O^6meG being mutagenic causing GC→AT transitions.[3] 7meG and backbone methylphosphotriesters are relatively innocuous, although 7meG can spontaneously depurinate and is prone to ring opening to make the secondary lesion Fapy-7meG,[6] and methylphosphotriesters induce the adaptive response as described later. While methyl adducts are the simplest and most commonly studied of alkylated adducts, larger alkyl groups also modify bases causing mutations and halting nucleic acid templated activities.[3]

Multiple repair pathways exist and are highly regulated to complement one another in the protection of the genome. To repair single base lesions, pathways include base excision repair (BER) which removes lesions with small modifications, nucleotide excision repair (NER) which removes bulky lesions by replacing segments of DNA, direct reversal which chemically reverses the modification, and mismatch repair which follows the replication fork and corrects point mutations caused by polymerases.[5,7] Single strand breaks can form in the genome and are

repaired by ligases, but can lead to double strand breaks. For the very deleterious double strand breaks, homologous repair (HR)[8] and non-homologous end joining (NHEJ)[9] pathways as well as error prone alternates rejoin DNA ends, with NHEJ not requiring a homologous template.[10] All of these pathways are vital for cellular health and stability. The base excision repair and direct reversal repair pathways are responsible for the bulk of alkylation repair, although certain modifications, such as bulkier lesions, may require the help of the other repair systems like the NER pathway.[3]

Base Excision Repair of Alkylation Damage

The BER pathway is found across all three domains of life, utilizing a set of enzymes to remove a damaged base and replace it with a fully repaired sequence. BER is initiated by a glycosylase that recognizes a specific base lesion substrate. In general, a monofunctional glycosylase is thought to perform a dissociative Sn1-like hydrolysis reaction in which an active site basic residue positions a water molecule to attack the C1′ position of the deoxyribose ring, thus hydrolyzing the *N*-glycosidic bond between the base and sugar (Figure 1-2). The resulting

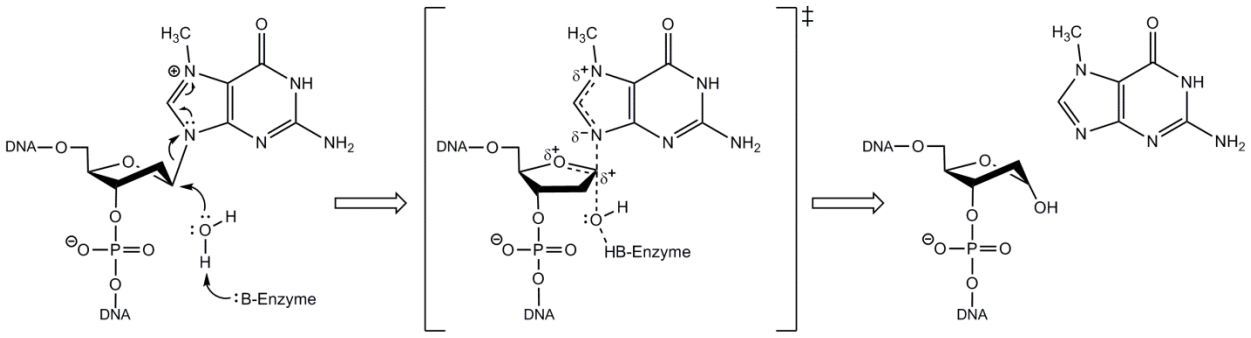


Figure 1-2. Dissociative S_N1-like mechanism for 7meG excision by monofunctional glycosylases. A general base in the glycosylase active site activates a water molecule to attack the C1′ of the deoxyribose sugar. The proposed transition state is stabilized by the protonation of the purine due to alkylation. For neutral lesions, some glycosylases utilize a general acid in the active site to stabilize the transition state. The final products are the free nucleotide and an abasic site. (Adapted from O'Brien 2006)[5]

abasic (apurinic/apyrimidic) site product is acted on by an AP endonuclease that hydrolyzes the phosphodiester bond 3′ to the abasic site, resulting in a single strand nick with a 3′-hydroxyl and 5′-deoxyribosephosphate (5′-dRP). The 5′-dRP is removed by a dRPase (prokaryotes) or the dRP lyase activity of Polymerase β (eukaryotes). The resulting gap can then be filled by a DNA polymerase and the backbone nick is sealed by a DNA ligase. The BER pathway described is defined as short-patch BER with a monofunctional glycosylase that only performs the bond hydrolysis reaction (Figure 1-3). Variations of BER are common and can include the use of bifunctional glycosylases that perform a lyase reaction in addition to base excision. Different sources of the lyase reaction result in other enzymes needed to process the nicked abasic site to make polymerase compatible ends.[5] Additionally, long-patch BER can occur in which the

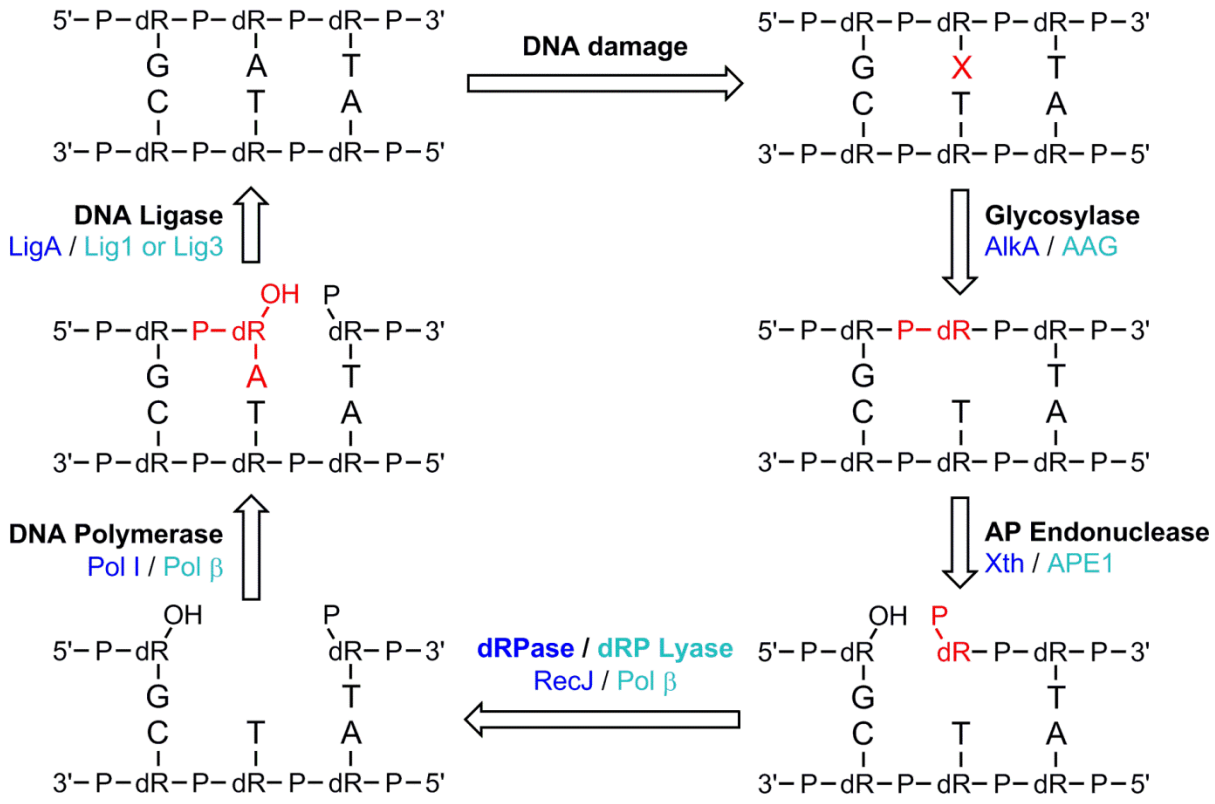


Figure 1-3. Base excision repair pathway for *Escherichia coli* and humans. Short-patch BER with a monofunctional glycosylase is shown with example enzymes from *E. coli* and humans shown in dark and light blue, respectively. (Adapted from O'Brien 2006)[5]

polymerase replaces a longer section of the DNA, making a flap that is removed by a flap endonuclease.[11] This dissertation will focus on the mechanisms of monofunctional glycosylases.

The specificity of BER lies with the glycosylase that has to search the vast amount of undamaged bases in a genome to find a damaged substrate. Most glycosylases gain access to the *N*-glycosidic bond via a process known as base flipping or nucleotide flipping, in which the lesion is flipped 180° out of the DNA duplex and into the enzyme active site.[12-17] The rate of nucleotide flipping partially depends on the strength of the base pair. If base pairing in the DNA duplex is weakened, such as in mismatches or with lesions containing modifications blocking Watson-Crick base pairing, then nucleotide flipping is expected to be faster.[18]

In addition to the dependence on base pair strength, nucleotide flipping provides another specificity filter for glycosylases as flipped base lesions must fit into the active site and fulfill interaction requirements for recognition. Some glycosylases have a very narrow substrate range and only remove one type of lesion. For example, Tag in *E. coli* is specific for N^3-methylated purines. Other glycosylases exhibit a broad substrate range and have the ability to excise multiple types of bases. Alkylation repair glycosylases that exhibit this expansive activity are the *E. coli* AlkA and mammalian AAG.[5]

AlkA and AAG recognize the bulk of alkylation damage, and thus are central for the protection of cells from alkylating agents, and their homologs are present in all domains of life. Other glycosylases that repair alkylation damage include Tag and the lesser known AlkC and AlkD that are found only in limited prokaryotes.[19-21] However, AlkA and AAG have broad substrate ranges and have evolved to be the major players in BER alkylation repair. They have independently evolved to perform this task and differ in their regulation and structure. This dissertation focuses on the mechanisms of these two BER glycosylases.

AlkA Function and Structure

AlkA, otherwise known as 3-methyladenine DNA glycosylase II, was the second glycosylase discovered in *E. coli* to protect cells from alkylation damage.[22,23] Tag (3-methyladenine DNA glycosylase I) was discovered first and is specific for 3meA and 3meG lesions.[19] In contrast, AlkA has been shown to have a wide substrate range with substrates that include 3meA, 7meG, and other minor alkylated purines and pyrimidines (e.g., 2meT; 2meC; 7meA).[23-26] Additionally, AlkA exhibits substantial in vitro activity towards the cyclic adduct 1,N^6-ethenoadenine (εA), created by lipid peroxidation byproducts, and the deaminated adenine adduct, hypoxanthine (Hx), caused by reaction with oxidative species (Figure 1-4).[27,28] AlkA also shows activity on larger alkylation adducts (e.g., ethylated, propylated, etc.), and other oxidative lesions (e.g., xanthosine, oxanine, and 5-formyluracil), although these are less studied.[29-31] AlkA acts on all its substrates with a similar rate enhancement, defining it as a truly nonspecific enzyme. In fact, AlkA also catalyzes the removal of undamaged purines, thus performing gratuitous repair.[32] Despite the same rate enhancement for all substrates, 3meA and 7meG are removed by AlkA at rates 100-fold faster than those of εA and Hx, due to the lability of the *N*-glycosidic bond promoted by the positive charge on N^3 and N^7 alkyl adducts (Figure 1-1).[18] Based on these results and extensive genetic studies, the protonated purines are considered the physiological substrates of AlkA.[18,22,23] Other enzymes appear to have primary responsibility for εA and Hx excision in *E. coli*, notably the direct reversal iron(II) α-ketoglutarate dependent deoxygenase, AlkB, and endonuclease V, respectively.[33-35]

AlkA is a member of the large family of helix-hairpin-helix (HhH) proteins that includes other glycosylases such as Tag, Ogg1 (8-oxoguanine glycosylase, acts on 8-oxoG:C), MutY (mutator, acts on 8-oxoG:A mismatches), and Nth (endonuclease III, acts broadly on modified

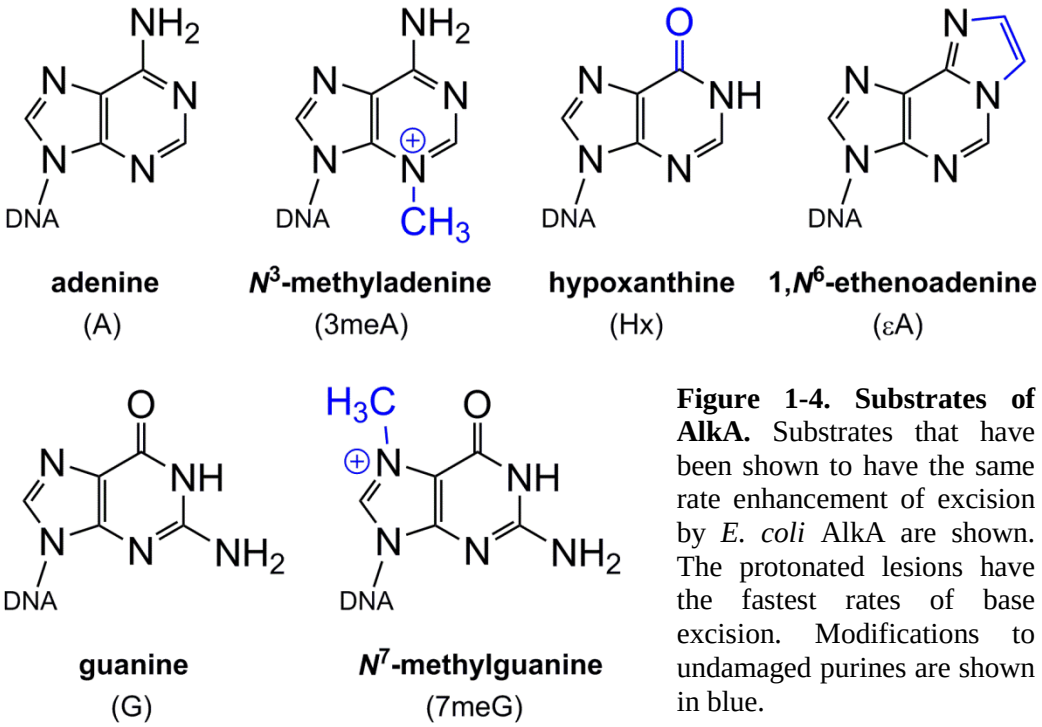


Figure 1-4. Substrates of AlkA. Substrates that have been shown to have the same rate enhancement of excision by *E. coli* AlkA are shown. The protonated lesions have the fastest rates of base excision. Modifications to undamaged purines are shown in blue.

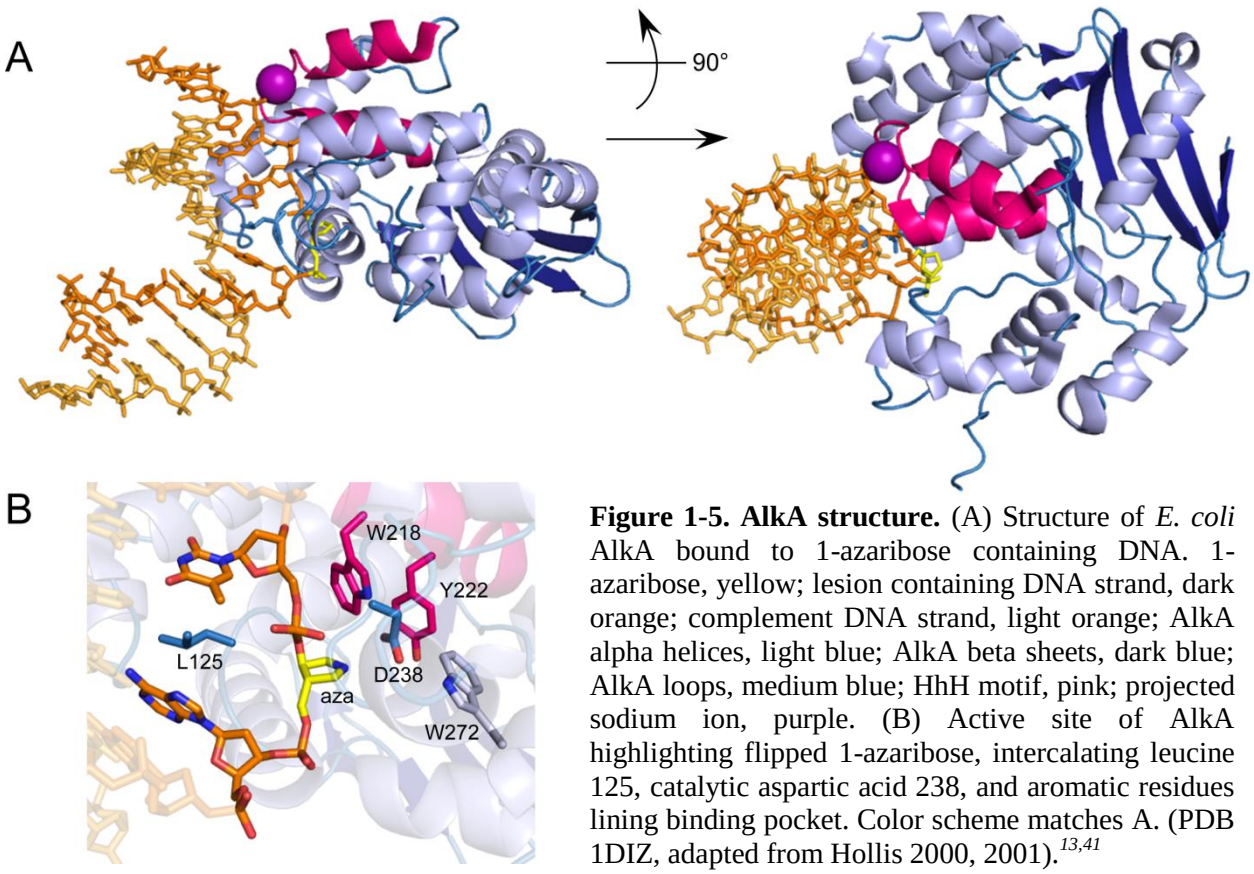


Figure 1-5. AlkA structure. (A) Structure of *E. coli* AlkA bound to 1-azaribose containing DNA. 1-azaribose, yellow; lesion containing DNA strand, dark orange; complement DNA strand, light orange; AlkA alpha helices, light blue; AlkA beta sheets, dark blue; AlkA loops, medium blue; HhH motif, pink; projected sodium ion, purple. (B) Active site of AlkA highlighting flipped 1-azaribose, intercalating leucine 125, catalytic aspartic acid 238, and aromatic residues lining binding pocket. Color scheme matches A. (PDB 1DIZ, adapted from Hollis 2000, 2001).[13,41]

pyrimidines).[36,37] Similar to helix-turn-helix (HtH) motifs, these amino acid structures are used to bind to DNA. However, in contrast to related HtH motifs in which the second helix inserts into the major groove to contact specific sequence contexts, the second helix in HhH points back into the protein and interacts with the minor groove and DNA backbone, aiding in nonspecific DNA binding. Other DNA binding proteins reliant on nonspecific binding also have HhH motifs, including some DNA polymerases.[38] In AlkA, the HhH motif provides most of the few hydrogen bonding contacts between AlkA and the DNA substrate, downstream from the lesion. There may also be a metal mediated interaction between the DNA and HhH motif, which was modeled as a sodium ion in the crystal structure, similar to the HhH motif in DNA polymerase β.[39-41] Compared to other glycosylases, AlkA has a relatively uncharged nonpolar DNA binding cleft, with most interactions being van der Waals contacts between the enzyme and lesion containing DNA strand. Although no crystal structure has been obtained with AlkA bound to a damaged base, the structure of AlkA bound to the transition state mimic 1-azaribose shows the sugar moiety being flipped into the enzyme active site (Figure 1-5A).[41] This flipping of the abasic nucleotide analog is accompanied by the intercalation of leucine 125 into the minor groove to fill the space vacated by the 1-azaribose. The intercalation causes a 66° bend in the DNA backbone and a widening of the minor groove. In the enzyme active site, the conserved catalytic aspartic acid residue 238 is proposed to coordinate the water nucleophile, while various aromatic residues are theoretically positioned for base stacking interactions with a base lesion (Figure 1-5B). The active site is relatively large, consistent with the broad substrate range of AlkA.[13,41]

Prokaryotic Adaptive Response for Alkylation Damage Repair

Repair of alkylation damage in bacteria, notably *E. coli*, is dominated by the adaptive (ada) response, in which alkylation repair proteins are upregulated in response to alkylation

stress. In particular, AlkA expression is regulated by the ada response.[42]

The ada response was originally described in *E. coli* by Samson and Cairns in 1977.[43] Cells exposed to low doses of the Sn1 methylating agent *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine (MNNG) over long periods of time showed genomic mutations during only the first hour of exposure. Furthermore, the *E. coli* could then survive high doses of mutagen that would typically be lethal. *E. coli* was therefore becoming resistant to MNNG when exposed to nonlethal doses. Subsequent work has shown that the cornerstone of the ada response is the direct reversal methyltransferase Ada protein that acts both as a methyltransferase and as a transcription factor, inducing the ada regulon of genes.[3,44-46]

Ada contains two domains that act independently on methylated DNA sites (Figure 1-6). First, the C-terminal domain acts on O^6-methylguanine (O^6meG) by transferring the methyl group to the active site cysteine 321 residue.[47] Secondly, the N-terminal domain of Ada removes the methyl group from *S*-diasteroisomers of DNA backbone methylphophotriesters and transfers it on the active site cysteine 38 residue.[48,49] In both cases, the methylation of cysteine is

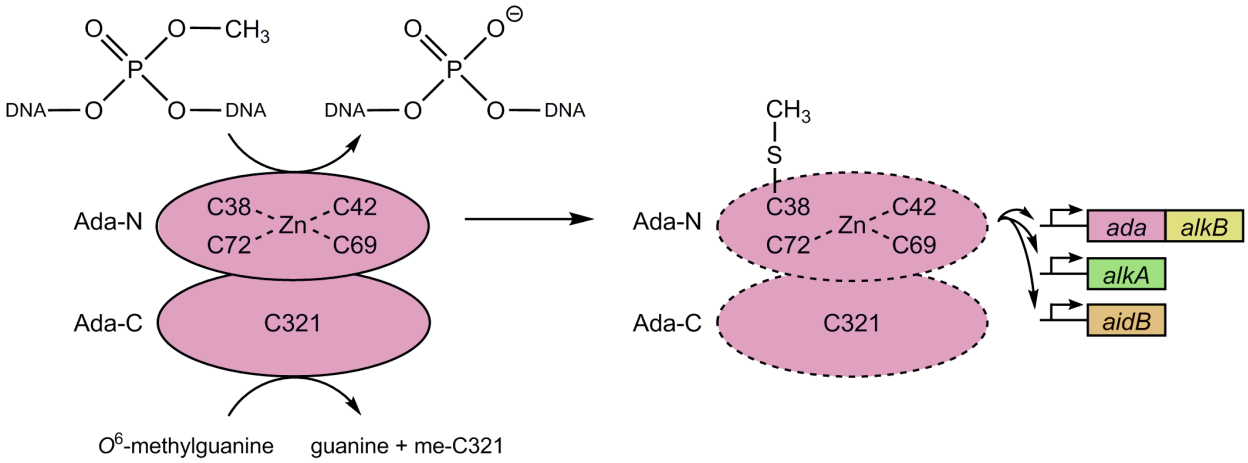


Figure 1-6. Adaptive response in *E. coli*. The two domains of Ada (purple) act as methyltransferases. Ada-N catalyzes the transfer of the methyl group from the DNA backbone to cysteine 38 preventing it from coordinating with a zinc ion. After a projected conformational change, Ada becomes a transcriptional activator for *ada*, *alkB*, *alkA*, and *aidB*. (Adapted from Sedgwick 2004)[3]

irreversible, defining Ada as a suicide enzyme, incapable of multiple turnover. The modification on Cys38 in the N-terminal domain prevents it from coordinating a zinc ion, most likely causing a conformational change in the protein.[50] After this proposed change, Ada becomes a transcriptional activator by binding to and promoting the transcription of the *ada-alkB* operon and *alkA* and *aidB* genes via interactions with RNA polymerase.[51] All four protein products (Ada, AlkA, AlkB, and AidB) are involved in alkylation damage protection. The overexpressed concentrations of each ada response protein vary but generally can range from 4 to 100 fold over the pre-adapted levels.[52-59]

Ada, AlkA, and AlkB have the ability to remove all *N*- and *O*- methylation events on nucleobases and the phosphate backbone and thus protect cells from mutagenesis and cell death.[3] When any of these three proteins is absent, *E. coli* becomes more sensitive to alkylating agents.[60] As previously discussed, Ada directly removes methyl groups from O^6meG and AlkA has a broad range including N^3 and N^7 methylated purines. AlkB is a direct reversal enzyme that oxidatively demethylates 1meA and 3meC via an iron(II), α-ketoglutarate dependent mechanism.[61,62] AlkB also seems to be more important than AlkA for εA repair in *E. coli*.[34] The function of AidB is still unknown but it has been shown to bind DNA nonspecifically and may shield DNA from alkylation events by deactivating alkylating agents with its flavin moiety.[63,64]

The ada response is conserved in many bacterial species highlighting the prevalence of alkylation agents in nature. *Escherichia*, *Bacillus*, *Micrococcus*, *Streptomyces*, *Pseudomonas*, *Enterobacter*, and *Rhodobacter* all express ada response proteins, although the number of homologs and the extent of inducible expression differs across species.[20,54] Ada response protein homologs have also been found in archaea and eukaryotes, including Mag1, an AlkA homolog present in budding and fission yeast, and mammalian AlkB homologs.[65-68] These homologs also

protect cells from alkylation damage but their substrate specificities vary widely, and there is no detectable adaptive response. The ada response proteins are therefore a diverse set of enzymes related in their function as alkylation damage repair proteins.

AAG Function, Structure and the Eukaryotic Repair of Alkylation Damage

The mammalian alkyladenine DNA glycosylase (AAG; also known as N-methylpurine DNA glycosylase, MPG; alkyl-N-purine DNA glycosylase, ANPG) has evolved to carry out a function similar to AlkA in eukaryotes.[69] Surprisingly *Bacillus subtilis* is one of the very few bacterial species to contain an AAG homolog, although there is always a possibility of others unknown.[69,70] Other than these rare instances, AlkA and AAG are mutually exclusive in their expression across species. Mammalian AAG has a similar, though less broad, substrate range to that of AlkA, including 3meA, 7meG, εA and Hx.[27,28,71,72] AAG is the sole alkylation repair glycosylase in all eukaryotes but yeast, and has been shown to be important for in vivo removal of εA although the AlkB homolog, ABH2, overlaps in this function.[73] AAG deficient mice do not show a phenotype under normal conditions but are sensitive to methylating and oxidative agents and AAG knockout mice have also been shown to accumulate εA-like lesions over time.[74-77] Thus, AAG is necessary for genome stability under times of alkylation stress.

AlkA and AAG have evolved to repair a similar set of substrates but they differ greatly in structure (Figure 1-7A). AAG has a unique fold unseen in any other protein. Basic residues line the DNA binding cleft, providing charge-charge interactions between the positively charged residues and negatively charged DNA backbone phosphates. This interaction aids in facilitated diffusion of AAG along DNA during the lesion searching process. Nucleotide flipping is also observed with AAG bound to an εA lesion, and is accompanied by the intercalation of tyrosine

162 and a shallower bend (22°) of the DNA duplex.[78] It is not known if the intercalation-dependent bend angle is a real difference between AAG and AlkA or is just due to the effects of crystal packing in both cases. Ongoing work has shown that the intercalating Tyr162 residue plays a major role in the specificity of AAG, and that large aromatic intercalating residues slow down the rate of nucleotide flipping and provide a selectivity filter for AAG.[79] In the active site of AAG there are base stacking interactions between the flipped εA and the active site residue tyrosine 127 (Figure 1-7B). Additional histidine and tyrosine residues in the active site stabilize the flipped lesion. Glutamic acid residue 125 is proposed to act as the catalytic general base that positions and deprotonates a water molecule.[13,78]

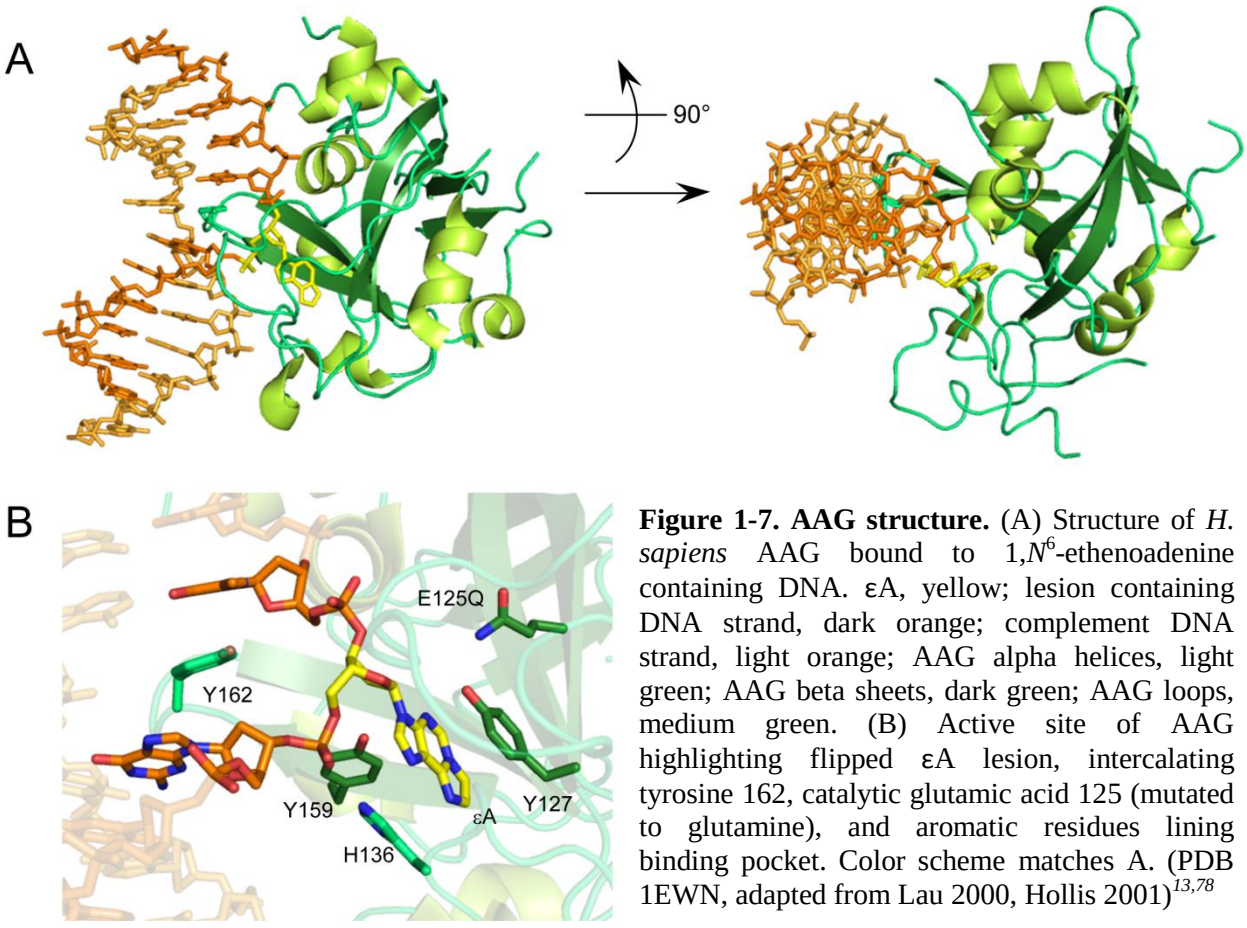


Figure 1-7. AAG structure. (A) Structure of *H. sapiens* AAG bound to 1,N^6-ethenoadenine containing DNA. εA, yellow; lesion containing DNA strand, dark orange; complement DNA strand, light orange; AAG alpha helices, light green; AAG beta sheets, dark green; AAG loops, medium green. (B) Active site of AAG highlighting flipped εA lesion, intercalating tyrosine 162, catalytic glutamic acid 125 (mutated to glutamine), and aromatic residues lining binding pocket. Color scheme matches A. (PDB 1EWN, adapted from Lau 2000, Hollis 2001)[13,78]

Regulation of Alkylation Repair Enzymes in Eukaryotes

No evidence of an ada-like response has been observed in eukaryotes. AAG is expressed constitutively across mouse cell types and does not appear to be inducible in any condition.[72] Even in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* yeast strains that have the AlkA homolog, Mag1, no homolog for Ada is present. Eukaryotes do have inducible responses for other types of DNA damage, notably the DNA damage response for double strand breaks, and other inducible responses to oxidative damage.[80,81] However, none of these responses encompass adaption during alkylation stress. AAG is a more selective enzyme with higher rate enhancements for specific substrates like εA, unlike AlkA that acts as a general nonspecific enzyme.[18,82] The levels of AAG in mammalian tissue seem to be effective and thus overexpression is not necessary for alkylation repair.

Importance of DNA Repair Pathway Regulation

The field of DNA repair is ever expanding due to the large amount of pathways and proteins present and their role in genome stability and disease. In many cancers and diseases, DNA repair pathways are altered, causing higher rates of mutagenesis. Conversely, DNA damaging agents in chemotherapies are sometimes used in conjunction with inhibitors of DNA repair proteins, in order to increase genome instability and thus harm the cancer cells and induce apoptosis.[83-85] Understanding the mechanisms and functions of DNA repair proteins is central to the ability to control such processes.

Glycosylases are fascinating in that they provide the specificity for the base excision repair pathway. *E. coli* have seven different glycosylases and humans have eleven, many being homologs of the bacterial proteins.[5,86] The large number of glycosylase enzymes is observed

across domains of life, emphasizing the pervasive amounts of environmental and endogenous DNA damaging agents that cause genome instability and the importance of BER to repair such damage.

Tight regulation of glycosylase expression and BER is necessary for the health of a cell. For example, if AlkA was constantly expressed at its post adaptive response levels in the absence of DNA damage, it would remove undamaged bases gratuitously.[32] The resulting abasic sites are equally, if not more, damaging than modified bases, due to their penchant for causing single and subsequently double stranded DNA breaks.[87,88] Similarly, studies with overexpressed Mag1, the AlkA homolog in *S. cerevisiae*, and AAG in yeast and mammalian cell lines has led to the increase in mutation frequency and alkylation sensitivity, most likely due to an imbalance between BER and MMR pathways and an accumulation of AP sites.[89-92] Additionally, overexpression of Pol β leads to increased mutagenesis and overexpression of APE1 does not increase the resistance to DNA damaging agents.[93,94] So while BER is an important pathway for genome stability, higher levels of the BER proteins are not helpful and sometimes are harmful for a cell.[95] Researching the specificities and mechanisms of these proteins allows a clearer understanding of their regulation and purpose. Further manipulation of the pathways may play a role in controlling human disease and genome instability.

Current Work and Scope of Thesis

This dissertation focuses on the mechanisms and functions of AlkA and AAG glycosylases, to better define their similarities and differences. The scope of this work is to clearly characterize these two independently evolved enzymes and compare their recognition and binding of DNA, their chemistry of bond hydrolysis, and their in vivo substrate specificities.

While much work has been done on the mechanism of human AAG,[79,96,97] little was known about the mechanism of *E. coli* AlkA. I have thoroughly defined a kinetic and thermodynamic framework for the excision of εA by AlkA. This endogenously produced lesion is excised by AlkA with the same rate enhancement as 3meA, and unlike the protonated lesion, is stable and naturally fluorescent. We can therefore observe nucleotide flipping by AlkA and determine the contribution of flipping to the specificity of the protein. Additionally, as AlkA and AAG share εA as a substrate, direct comparisons were made between the two enzymes. Surprisingly, the rate of bond hydrolysis is identical with AlkA and AAG, and the main differences are in substrate recognition, including the nucleotide flipping step.

Additional mechanistic studies were performed with AlkA to investigate the dependence on salt concentration for DNA binding and base excision. Furthermore, another version of AlkA was studied that contained a short eight amino acid C-terminal tag. While the structure of AlkA suggests that the C-terminus is distant from the active site, differences were observed between the WT and tagged proteins during in vitro biochemical studies. The greatest disparity between the two AlkA proteins was a relief of auto-inhibition when the tagged AlkA was used. This activity occurred when AlkA was in excess over substrate, as the tagged protein binds more weakly to nonspecific DNA and shows reduced interference between multiply bound protein molecules.

We also sought to characterize DNA binding by these glycosylases. We used a variety of methods, notably fluorescence anisotropy, to characterize nonspecific and specific DNA binding by AAG. Our results confirmed that the active enzyme is a monomer and can bind at high density to double stranded and single stranded DNA. We also investigated the affinity of AAG for undamaged DNA and the effect of salt on DNA binding. A surprising feature of this work is

that fluorescein-type dyes on flexible linkers can give very large anisotropy signals which may affect analysis.

Finally, we sought to characterize the *B. subtilis* homologs bAlkA and bAAG, in terms of their substrate specificity and role in DNA damage repair in vivo. The unique overlap of AlkA and AAG homologs in *Bacillus subtilis* has not been thoroughly researched with the limited work focusing on bAlkA and bAAG separately.[69,70] Numerous studies investigate the substrate specificity of *E. coli* AlkA and human AAG but these examples do not have a system in which the two enzymes are in competition for substrates. As *B. subtilis* does not contain AlkB homologs, we hypothesized that the BER enzymes would have an expanded substrate range. Therefore, we tested the canonical AlkB substrates, 1meA and 3meC, and found that both bAlkA and bAAG are able to recognize these lesions.

This work expands our understanding of the BER glycosylases, AlkA and AAG, focusing on the kinetic and thermodynamic mechanisms that define each enzyme and determine their distinct roles in alkylation repair.

References

1. Marnett, L. J., Riggins, J. N., and West, J. D. (2003) Endogenous generation of reactive oxidants and electrophiles and their reactions with DNA and protein, *J Clin Invest 111*, 583-593.
2. Beranek, D. T. (1990) Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents, *Mutation research 231*, 11-30.
3. Sedgwick, B. (2004) Repairing DNA-methylation damage, *Nat Rev Mol Cell Biol 5*, 148-157.
4. Lindahl, T. (1993) Instability and decay of the primary structure of DNA, *Nature 362*, 709-715.
5. O'Brien, P. J. (2006) Catalytic promiscuity and the divergent evolution of DNA repair enzymes, *Chem Rev 106*, 720-752.
6. Tudek, B. (2003) Imidazole ring-opened DNA purines and their biological significance, *Journal of biochemistry and molecular biology 36*, 12-19.
7. Wood, R. D. (1996) DNA repair in eukaryotes, *Annual review of biochemistry 65*, 135-167.
8. Jasin, M., and Rothstein, R. (2013) Repair of strand breaks by homologous recombination, *Cold Spring Harbor perspectives in biology 5*, a012740.
9. Waters, C. A., Strande, N. T., Wyatt, D. W., Pryor, J. M., and Ramsden, D. A. (2014) Nonhomologous end joining: a good solution for bad ends, *DNA repair 17*, 39-51.
10. Ceccaldi, R., Rondinelli, B., and D'Andrea, A. D. (2015) Repair Pathway Choices and Consequences at the Double-Strand Break, *Trends in cell biology*.
11. 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, *Cellular and molecular life sciences : CMLS 66*, 981-993.
12. Cheng, X., and Blumenthal, R. M. (1996) Finding a basis for flipping bases, *Structure 4*, 639-645.
13. Hollis, T., Lau, A., and Ellenberger, T. (2001) Crystallizing thoughts about DNA base excision repair, *Prog Nucleic Acid Res Mol Biol 68*, 305-314.
14. 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.
15. Maiti, A., Morgan, M. T., and Drohat, A. C. (2009) Role of two strictly conserved residues in nucleotide flipping and N-glycosylic bond cleavage by human thymine DNA glycosylase, *The Journal of biological chemistry 284*, 36680-36688.
16. Roberts, R. J., and Cheng, X. (1998) Base flipping, *Annual review of biochemistry 67*, 181-198.
17. Klimasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. (1994) HhaI methyltransferase flips its target base out of the DNA helix, *Cell 76*, 357-369.

18. O'Brien, P. J., and Ellenberger, T. (2004) The Escherichia coli 3-methyladenine DNA glycosylase AlkA has a remarkably versatile active site, *The Journal of biological chemistry 279*, 26876-26884.
19. Riazuddin, S., and Lindahl, T. (1978) Properties of 3-methyladenine-DNA glycosylase from Escherichia coli, *Biochemistry 17*, 2110-2118.
20. Mielecki, D., Wrzesinski, M., and Grzesiuk, E. (2015) Inducible repair of alkylated DNA in microorganisms, *Mutation research. Reviews in mutation research 763*, 294-305.
21. Alseth, I., Rognes, T., Lindback, T., Solberg, I., Robertsen, K., Kristiansen, K. I., Mainieri, D., Lillehagen, L., Kolsto, A. B., and Bjoras, M. (2006) A new protein superfamily includes two novel 3-methyladenine DNA glycosylases from Bacillus cereus, AlkC and AlkD, *Molecular microbiology 59*, 1602-1609.
22. Evensen, G., and Seeberg, E. (1982) Adaptation to alkylation resistance involves the induction of a DNA glycosylase, *Nature 296*, 773-775.
23. Thomas, L., Yang, C. H., and Goldthwait, D. A. (1982) Two DNA glycosylases in Escherichia coli which release primarily 3-methyladenine, *Biochemistry 21*, 1162-1169.
24. Bjelland, S., Bjoras, M., and Seeberg, E. (1993) Excision of 3-methylguanine from alkylated DNA by 3-methyladenine DNA glycosylase I of Escherichia coli, *Nucleic Acids Res 21*, 2045-2049.
25. Bjelland, S., Birkeland, N. K., Benneche, T., Volden, G., and Seeberg, E. (1994) DNA glycosylase activities for thymine residues oxidized in the methyl group are functions of the AlkA enzyme in Escherichia coli, *The Journal of biological chemistry 269*, 30489-30495.
26. McCarthy, T. V., Karran, P., and Lindahl, T. (1984) Inducible repair of O-alkylated DNA pyrimidines in Escherichia coli, *The EMBO journal 3*, 545-550.
27. Saparbaev, M., Kleibl, K., and Laval, J. (1995) Escherichia coli, Saccharomyces cerevisiae, rat and human 3-methyladenine DNA glycosylases repair 1,N6-ethenoadenine when present in DNA, *Nucleic Acids Res 23*, 3750-3755.
28. Saparbaev, M., and Laval, J. (1994) Excision of hypoxanthine from DNA containing dIMP residues by the Escherichia coli, yeast, rat, and human alkylpurine DNA glycosylases, *Proceedings of the National Academy of Sciences of the United States of America 91*, 5873-5877.
29. Masaoka, A., Terato, H., Kobayashi, M., Honsho, A., Ohyama, Y., and Ide, H. (1999) Enzymatic repair of 5-formyluracil. I. Excision of 5-formyluracil site-specifically incorporated into oligonucleotide substrates by alka protein (Escherichia coli 3-methyladenine DNA glycosylase II), *The Journal of biological chemistry 274*, 25136-25143.
30. Terato, H., Masaoka, A., Kobayashi, M., Fukushima, S., Ohyama, Y., Yoshida, M., and Ide, H. (1999) Enzymatic repair of 5-formyluracil. II. Mismatch formation between 5-formyluracil and guanine during dna replication and its recognition by two proteins involved in base excision repair (AlkA) and mismatch repair (MutS), *The Journal of biological chemistry 274*, 25144-25150.
31. Tudek, B., Van Zeeland, A. A., Kusmierek, J. T., and Laval, J. (1998) Activity of Escherichia coli DNA-glycosylases on DNA damaged by methylating and ethylating agents and influence of 3-substituted adenine derivatives, *Mutation research 407*, 169-176.

32. Berdal, K. G., Johansen, R. F., and Seeberg, E. (1998) Release of normal bases from intact DNA by a native DNA repair enzyme, *The EMBO journal 17*, 363-367.
33. Frick, L. E., Delaney, J. C., Wong, C., Drennan, C. L., and Essigmann, J. M. (2007) Alleviation of 1,N6-ethanoadenine genotoxicity by the Escherichia coli adaptive response protein AlkB, *Proceedings of the National Academy of Sciences of the United States of America 104*, 755-760.
34. Maciejewska, A. M., Sokolowska, B., Nowicki, A., and Kusmierek, J. T. (2011) The role of AlkB protein in repair of 1,N(6)-ethenoadenine in Escherichia coli cells, *Mutagenesis 26*, 401-406.
35. Yao, M., Hatahet, Z., Melamede, R. J., and Kow, Y. W. (1994) Purification and characterization of a novel deoxyinosine-specific enzyme, deoxyinosine 3' endonuclease, from Escherichia coli, *The Journal of biological chemistry 269*, 16260-16268.
36. Denver, D. R., Swenson, S. L., and Lynch, M. (2003) An evolutionary analysis of the helix-hairpin-helix superfamily of DNA repair glycosylases, *Mol Biol Evol 20*, 1603-1611.
37. Eichman, B. F., O'Rourke, E. J., Radicella, J. P., and Ellenberger, T. (2003) Crystal structures of 3-methyladenine DNA glycosylase MagIII and the recognition of alkylated bases, *The EMBO journal 22*, 4898-4909.
38. Doherty, A. J., Serpell, L. C., and Ponting, C. P. (1996) The helix-hairpin-helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA, *Nucleic Acids Res 24*, 2488-2497.
39. Pelletier, H., and Sawaya, M. R. (1996) Characterization of the metal ion binding helix-hairpin-helix motifs in human DNA polymerase beta by X-ray structural analysis, *Biochemistry 35*, 12778-12787.
40. Pelletier, H., Sawaya, M. R., Wolfle, W., Wilson, S. H., and Kraut, J. (1996) Crystal structures of human DNA polymerase beta complexed with DNA: implications for catalytic mechanism, processivity, and fidelity, *Biochemistry 35*, 12742-12761.
41. Hollis, T., Ichikawa, Y., and Ellenberger, T. (2000) DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, Escherichia coli AlkA, *The EMBO journal 19*, 758-766.
42. Sedgwick, B., and Lindahl, T. (2002) Recent progress on the Ada response for inducible repair of DNA alkylation damage, *Oncogene 21*, 8886-8894.
43. Samson, L., and Cairns, J. (1977) A new pathway for DNA repair in Escherichia coli, *Nature 267*, 281-283.
44. Schendel, P. F., and Robins, P. E. (1978) Repair of O6-methylguanine in adapted Escherichia coli, *Proceedings of the National Academy of Sciences of the United States of America 75*, 6017-6020.
45. Jeggo, P. (1979) Isolation and characterization of Escherichia coli K-12 mutants unable to induce the adaptive response to simple alkylating agents, *Journal of bacteriology 139*, 783-791.
46. Lindahl, T., Sedgwick, B., Sekiguchi, M., and Nakabeppu, Y. (1988) Regulation and expression of the adaptive response to alkylating agents, *Annual review of biochemistry 57*, 133-157.
47. Demple, B., Sedgwick, B., Robins, P., Totty, N., Waterfield, M. D., and Lindahl, T. (1985) Active site and complete sequence of the suicidal methyltransferase that counters

alkylation mutagenesis, *Proceedings of the National Academy of Sciences of the United States of America 82*, 2688-2692.
48. Sedgwick, B., Robins, P., Totty, N., and Lindahl, T. (1988) Functional domains and methyl acceptor sites of the Escherichia coli ada protein, *The Journal of biological chemistry 263*, 4430-4433.
49. He, C., Wei, H., and Verdine, G. L. (2003) Converting the sacrificial DNA repair protein N-ada into a catalytic methyl phosphotriester repair enzyme, *Journal of the American Chemical Society 125*, 1450-1451.
50. Lin, Y., Dotsch, V., Wintner, T., Peariso, K., Myers, L. C., Penner-Hahn, J. E., Verdine, G. L., and Wagner, G. (2001) Structural basis for the functional switch of the E. coli Ada protein, *Biochemistry 40*, 4261-4271.
51. Landini, P., and Volkert, M. R. (2000) Regulatory responses of the adaptive response to alkylation damage: a simple regulon with complex regulatory features, *Journal of bacteriology 182*, 6543-6549.
52. Nakabeppu, Y., Miyata, T., Kondo, H., Iwanaga, S., and Sekiguchi, M. (1984) Structure and expression of the alkA gene of Escherichia coli involved in adaptive response to alkylating agents, *The Journal of biological chemistry 259*, 13730-13736.
53. Kondo, H., Nakabeppu, Y., Kataoka, H., Kuhara, S., Kawabata, S., and Sekiguchi, M. (1986) Structure and expression of the alkB gene of Escherichia coli related to the repair of alkylated DNA, *The Journal of biological chemistry 261*, 15772-15777.
54. Fernandez de Henestrosa, A. R., and Barbe, J. (1991) Induction of the alkA gene of Escherichia coli in gram-negative bacteria, *Journal of bacteriology 173*, 7736-7740.
55. Landini, P., and Busby, S. J. (1999) Expression of the Escherichia coli ada regulon in stationary phase: evidence for rpoS-dependent negative regulation of alkA transcription, *Journal of bacteriology 181*, 6836-6839.
56. Booth, J. A., Thomassen, G. O., Rowe, A. D., Weel-Sneve, R., Lagesen, K., Kristiansen, K. I., Bjoras, M., Rognes, T., and Lindvall, J. M. (2013) Tiling array study of MNNG treated Escherichia coli reveals a widespread transcriptional response, *Scientific reports 3*, 3053.
57. Saget, B. M., and Walker, G. C. (1994) The Ada protein acts as both a positive and a negative modulator of Escherichia coli's response to methylating agents, *Proceedings of the National Academy of Sciences of the United States of America 91*, 9730-9734.
58. Volkert, M. R., Gately, F. H., and Hajec, L. I. (1989) Expression of DNA damage-inducible genes of Escherichia coli upon treatment with methylating, ethylating and propylating agents, *Mutation research 217*, 109-115.
59. Otsuka, M., Nakabeppu, Y., and Sekiguchi, M. (1985) Ability of various alkylating agents to induce adaptive and SOS responses: a study with lacZ fusion, *Mutation research 146*, 149-154.
60. Volkert, M. R., and Hajec, L. I. (1991) Molecular analysis of the aidD6::Mu d1 (bla lac) fusion mutation of Escherichia coli K12, *Molecular & general genetics : MGG 229*, 319-323.
61. Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., and Sedgwick, B. (2002) Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage, *Nature 419*, 174-178.
62. Falnes, P. O., Johansen, R. F., and Seeberg, E. (2002) AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli, *Nature 419*, 178-182.

63. Rohankhedkar, M. S., Mulrooney, S. B., Wedemeyer, W. J., and Hausinger, R. P. (2006) The AidB component of the Escherichia coli adaptive response to alkylating agents is a flavin-containing, DNA-binding protein, *Journal of bacteriology 188*, 223-230.
64. Bowles, T., Metz, A. H., O'Quin, J., Wawrzak, Z., and Eichman, B. F. (2008) Structure and DNA binding of alkylation response protein AidB, *Proceedings of the National Academy of Sciences of the United States of America 105*, 15299-15304.
65. Grasso, S., and Tell, G. (2014) Base excision repair in Archaea: back to the future in DNA repair, *DNA repair 21*, 148-157.
66. Bjoras, M., Klungland, A., Johansen, R. F., and Seeberg, E. (1995) Purification and properties of the alkylation repair DNA glycosylase encoded the MAG gene from Saccharomyces cerevisiae, *Biochemistry 34*, 4577-4582.
67. Memisoglu, A., and Samson, L. (1996) Cloning and characterization of a cDNA encoding a 3-methyladenine DNA glycosylase from the fission yeast Schizosaccharomyces pombe, *Gene 177*, 229-235.
68. Sedgwick, B., Bates, P. A., Paik, J., Jacobs, S. C., and Lindahl, T. (2007) Repair of alkylated DNA: recent advances, *DNA repair 6*, 429-442.
69. Aamodt, R. M., Falnes, P. O., Johansen, R. F., Seeberg, E., and Bjoras, M. (2004) The Bacillus subtilis counterpart of the mammalian 3-methyladenine DNA glycosylase has hypoxanthine and 1,N6-ethenoadenine as preferred substrates, *The Journal of biological chemistry 279*, 13601-13606.
70. Morohoshi, F., Hayashi, K., and Munkata, N. (1993) Bacillus subtilis alkA gene encoding inducible 3-methyladenine DNA glycosylase is adjacent to the ada operon, *Journal of bacteriology 175*, 6010-6017.
71. Engelward, B. P., Boosalis, M. S., Chen, B. J., Deng, Z., Siciliano, M. J., and Samson, L. D. (1993) Cloning and characterization of a mouse 3-methyladenine/7-methyl-guanine/3-methylguanine DNA glycosylase cDNA whose gene maps to chromosome 11, *Carcinogenesis 14*, 175-181.
72. Smith, S. A., and Engelward, B. P. (2000) In vivo repair of methylation damage in Aag 3-methyladenine DNA glycosylase null mouse cells, *Nucleic Acids Res 28*, 3294-3300.
73. 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 research 68*, 4142-4149.
74. 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, *Proceedings of the National Academy of Sciences of the United States of America 94*, 13087-13092.
75. Meira, L. B., Bugni, J. M., Green, S. L., Lee, C. W., Pang, B., Borenshtein, D., Rickman, B. H., Rogers, A. B., Moroski-Erkul, C. A., McFaline, J. L., Schauer, D. B., Dedon, P. C., Fox, J. G., and Samson, L. D. (2008) DNA damage induced by chronic inflammation contributes to colon carcinogenesis in mice, *J Clin Invest 118*, 2516-2525.
76. Ham, A. J., Engelward, B. P., Koc, H., Sangaiah, R., Meira, L. B., Samson, L. D., and Swenberg, J. A. (2004) New immunoaffinity-LC-MS/MS methodology reveals that Aag null mice are deficient in their ability to clear 1,N6-etheno-deoxyadenosine DNA lesions from lung and liver in vivo, *DNA repair 3*, 257-265.

77. 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, *Molecular and cellular biology 18*, 5828-5837.
78. 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, *Proceedings of the National Academy of Sciences of the United States of America 97*, 13573-13578.
79. Hendershot, J. M., and O'Brien, P. J. (2014) Critical role of DNA intercalation in enzyme-catalyzed nucleotide flipping, *Nucleic Acids Res 42*, 12681-12690.
80. Jackson, S. P., and Bartek, J. (2009) The DNA-damage response in human biology and disease, *Nature 461*, 1071-1078.
81. Davies, K. J. (2000) Oxidative stress, antioxidant defenses, and damage removal, repair, and replacement systems, *IUBMB life 50*, 279-289.
82. O'Brien, P. J., and Ellenberger, T. (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase, *The Journal of biological chemistry 279*, 9750-9757.
83. Wallace, S. S., Murphy, D. L., and Sweasy, J. B. (2012) Base excision repair and cancer, *Cancer letters 327*, 73-89.
84. Guillotin, D., and Martin, S. A. (2014) Exploiting DNA mismatch repair deficiency as a therapeutic strategy, *Experimental cell research 329*, 110-115.
85. Srivastava, M., and Raghavan, S. C. (2015) DNA double-strand break repair inhibitors as cancer therapeutics, *Chemistry & biology 22*, 17-29.
86. Krokan, H. E., and Bjoras, M. (2013) Base excision repair, *Cold Spring Harbor perspectives in biology 5*, a012583.
87. Clauson, C. L., Oestreich, K. J., Austin, J. W., and Doetsch, P. W. (2010) Abasic sites and strand breaks in DNA cause transcriptional mutagenesis in Escherichia coli, *Proceedings of the National Academy of Sciences of the United States of America 107*, 3657-3662.
88. Kathe, S. D., Shen, G. P., and Wallace, S. S. (2004) Single-stranded breaks in DNA but not oxidative DNA base damages block transcriptional elongation by RNA polymerase II in HeLa cell nuclear extracts, *The Journal of biological chemistry 279*, 18511-18520.
89. Posnick, L. M., and Samson, L. D. (1999) Imbalanced base excision repair increases spontaneous mutation and alkylation sensitivity in Escherichia coli, *Journal of bacteriology 181*, 6763-6771.
90. Glassner, B. J., Rasmussen, L. J., Najarian, M. T., Posnick, L. M., and Samson, L. D. (1998) Generation of a strong mutator phenotype in yeast by imbalanced base excision repair, *Proceedings of the National Academy of Sciences of the United States of America 95*, 9997-10002.
91. Hofseth, L. J., Khan, M. A., Ambrose, M., Nikolayeva, O., Xu-Welliver, M., Kartalou, M., Hussain, S. P., Roth, R. B., Zhou, X., Mechanic, L. E., Zurer, I., Rotter, V., Samson, L. D., and Harris, C. C. (2003) The adaptive imbalance in base excision-repair enzymes generates microsatellite instability in chronic inflammation, *J Clin Invest 112*, 1887-1894.

92. Rinne, M., Caldwell, D., and Kelley, M. R. (2004) Transient adenoviral N-methylpurine DNA glycosylase overexpression imparts chemotherapeutic sensitivity to human breast cancer cells, *Molecular cancer therapeutics 3*, 955-967.
93. Chan, K., Houlbrook, S., Zhang, Q. M., Harrison, M., Hickson, I. D., and Dianov, G. L. (2007) Overexpression of DNA polymerase beta results in an increased rate of frameshift mutations during base excision repair, *Mutagenesis 22*, 183-188.
94. Herring, C. J., Deans, B., Elder, R. H., Rafferty, J. A., MacKinnon, J., Barzilay, G., Hickson, I. D., Hendry, J. H., and Margison, G. P. (1999) Expression levels of the DNA repair enzyme HAP1 do not correlate with the radiosensitivities of human or HAP1-transfected rat cell lines, *British journal of cancer 80*, 940-945.
95. Frosina, G. (2000) Overexpression of enzymes that repair endogenous damage to DNA, *Eur J Biochem 267*, 2135-2149.
96. 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.
97. Hendershot, J. M., Wolfe, A. E., and O'Brien, P. J. (2011) Substitution of active site tyrosines with tryptophan alters the free energy for nucleotide flipping by human alkyladenine DNA glycosylase, *Biochemistry 50*, 1864-1874.

Chapter II

Kinetic Mechanism for the Flipping and Excision of 1,N^6-ethenoadenine by AlkA

Abstract

The *Escherichia coli* 3-methyladenine DNA glycosylase II (AlkA), an adaptive response glycosylase with a broad substrate range, initiates base excision repair by flipping a lesion out of the DNA duplex and hydrolyzing the *N*-glycosidic bond. We used transient and steady state kinetics to determine the minimal mechanism for recognition and excision of 1,N^6-ethenoadenine (εA) by AlkA. The natural fluorescence of this endogenously produced lesion allowed us to directly monitor the nucleotide flipping step. We found that AlkA rapidly and reversibly binds and flips out εA prior to *N*-glycosidic bond hydrolysis, which is the rate-limiting step of the reaction. The binding affinity of AlkA for the εA-DNA lesion is only 40-fold tighter than for a nonspecific site and 20-fold weaker than for the abasic-DNA site. The mechanism of AlkA-catalyzed excision of εA was compared to that of the human alkyladenine DNA glycosylase (AAG), an independently evolved glycosylase that recognizes many of the same substrates. AlkA and AAG both catalyze *N*-glycosidic bond hydrolysis to release εA and their overall rates of reaction are within 2-fold of each other. Nevertheless, we find dramatic differences in the kinetics and thermodynamics for binding to εA-DNA. AlkA catalyzes nucleotide flipping an order of magnitude faster than AAG, however the equilibrium for flipping is almost three orders

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

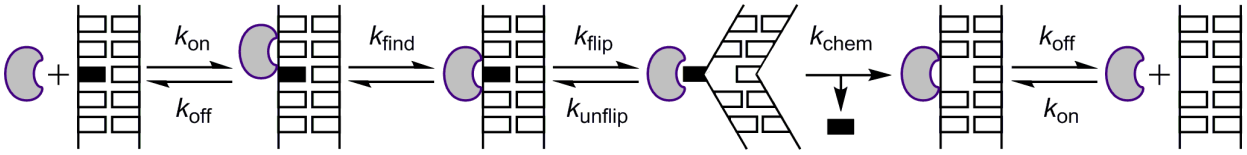
of magnitude more favorable for AAG than for AlkA. These results illustrate how enzymes that perform the same chemistry can use different substrate recognition strategies to effectively repair DNA damage.

Introduction

DNA nucleobases have multiple sites susceptible to alkylation damage, with alkylating agents coming from endogenous[1, 2] and exogenous[3] sources such as methyl methanesulfonate (MMS) and methyl halides.[3, 4] These lesions can exhibit cytotoxicity by blocking DNA replication (e.g., N^3-methyladenine, 3meA), cause mutations (e.g., 1-N^6-ethenoadenine, εA), or be relatively innocuous (e.g., N^7-methylguanine, 7meG).[4, 5] Repair of cytotoxic and mutagenic lesions is vital for cell survival, and organisms have evolved many pathways to deal with DNA damage. In the case of small alkyl modifications, both direct reversal and base excision repair (BER) pathways are used. The BER pathway repairs damage caused by oxidation and deamination in addition to alkylation damage. This pathway is initiated by a variety of different DNA glycosylases that hydrolyze the *N*-glycosidic bond between the base lesion and sugar, generating an abasic site product. Subsequent enzymes nick the backbone, remove the residual deoxyribose phosphate, fill the gap, and seal the nick.[4, 5] Two different superfamilies of glycosylases have evolved to recognize alkylative DNA damage. The first is the helix-hairpin helix superfamily, exemplified by *E. coli* 3-methyladenine DNA glycosylase II (AlkA), and the second is known as the alkyladenine DNA glycosylase (AAG) superfamily. Both AlkA and AAG exhibit a broad and overlapping substrate range.[5, 6]

AlkA homologs are found mostly in prokaryotes and a few eukaryotes, whereas AAG homologs are found predominantly in plants and animals.[7, 8] These two enzymes have been shown to have catalytic activity towards many of the same types of damaged bases, including methylated purines (e.g., 3meA and 7meG),[9, 10] etheno adducts (e.g., εA),[11] and oxidatively

Scheme 2-1. Minimal kinetic mechanism for base excision by AlkA.



damaged purines (e.g., hypoxanthine, Hx).[12] Both enzymes also exhibit detectable activity toward undamaged bases.[13, 14] While AAG and AlkA appear to be similar, they have distinct cellular roles in DNA repair. AlkA is upregulated as part of the adaptive (ada) response when *E. coli* is exposed to low levels of DNA alkylation. The overexpression of alkylation repair proteins allows *E. coli* to survive much higher levels of alkylating agents, which would have been toxic prior to adaptation.[15-17] In contrast, an adaptive response has not been observed in the regulation of AAG expression and it appears to be constitutively expressed.

Previous studies of AAG and AlkA have led to similar proposed mechanisms for the recognition and excision of base lesions as summarized in Scheme 1. Given that damaged bases are rare, initial binding most often occurs at undamaged sites. After searching and locating a site of damage, both glycosylases gain access to the *N*-glycosidic bond by flipping the nucleotide 180° out of the DNA duplex. Crystal structures of AAG and AlkA demonstrate nucleotide flipping with the glycosylases being bound to a flipped εA[18] and a flipped 1-azaribose sugar moiety,[19] respectively. Both enzymes use an active site carboxylate to position a water molecule for *N*-glycosidic bond hydrolysis (E125 in AAG[20] and D238 in AlkA[21, 22]). For AAG-catalyzed excision of εA, the hydrolysis step (k_{chem}) is usually rate-limiting, but product release can be limiting for some substrates (e.g., Hx).[23, 24] For AlkA-catalyzed excision of Hx, k_{chem} appears to be much slower than release of the abasic site.[25] It is likely that k_{chem} is also rate limiting for AlkA-catalyzed excision of εA, but this model has not been tested. A complete kinetic and thermodynamic framework for the AAG-catalyzed reaction has been reported,[26] but much less is

known about AlkA. Without direct observation of the binding and nucleotide flipping steps, it is not clear what step is rate-limiting for AlkA-catalyzed glycosylase activity.

In this study, the kinetic mechanism of AlkA catalyzed base excision is determined for εA-DNA, a mutagenic adduct that is endogenously produced from reactions with lipid oxidation byproducts.[27-29] Although it has been established that both BER and direct repair play a role in the physiological repair of etheno adducts, their relative contributions differ in bacterial and mammalian cells. AAG-initiated BER and direct repair pathways appear to share responsibilities for repair of εA in the mouse.[30] In contrast, genetic experiments in *E. coli* suggest that both AlkA-initiated BER and direct repair are important for repair of etheno adducts,[31] but the direct repair pathway plays the predominant role for repair of εA.[32] However, AlkA exhibits similar catalytic efficiency for the excision of εA as for 3meA and other damaged and undamaged bases.[22] Therefore, we expect that the insights gained regarding the mechanism of εA excision will advance our understanding of AlkA and its broad substrate range. By taking advantage of the intrinsic fluorescence of εA, we can obtain a full kinetic and thermodynamic framework that can be directly compared to previous studies of AAG.[23]

We report the minimal kinetic and thermodynamic framework for AlkA-catalyzed flipping and excision of εA. We found that the steps leading up to the rate-limiting hydrolysis step are under rapid equilibrium. In particular, the εA flipping and unflipping steps are both very rapid, and the complex of AlkA with the flipped out εA is only marginally stable. In comparison, AAG forms a very stable complex with the flipped out εA.[23, 26] This work provides a deeper understanding of how mechanistic differences affect the substrate recognition and catalytic efficiencies of two independently evolved DNA glycosylases and highlights the common features in the repair of damaged nucleotides.

Materials and Methods

Preparation of Proteins

Full-length *E. coli* AlkA was expressed in BL21 (DE3) *E. coli* cells. [21] Cultures were grown in LB to mid-log phase and induced with 0.1 mM IPTG for 4 hours at 37 ^{o}C. Cells were lysed in lysis buffer (50 mM Tris•Cl pH 8.6, 1 mM EDTA, 5% (v/v) glycerol, 50 mM β-ME, and 0.1 mg/mL PMSF) and centrifuged. The supernatant was passed through a DE52 column to remove DNA contamination and loaded onto an S-sepharose column. AlkA was eluted with a NaCl gradient (80 mM to 350 mM) and loaded onto a heparin-sepharose column, equilibrated with 50 mM Tris•Cl pH 8.6, 10 mM NaCl, 1 mM EDTA, 10% (v/v) glycerol, and 10 mM β-ME. AlkA was eluted with a NaCl gradient (100 mM to 700 mM NaCl) and concentrated with an Amicon Ultra concentrator (10 kDa cutoff) to the desired concentration of around 10 mg/mL. Aliquots were snap frozen in liquid nitrogen and stored at -80 °C in 50 mM Tris•Cl pH 8.6, 100 mM NaCl, 1 mM EDTA, 10 mM β-ME, 10% (v/v) glycerol. Prior to use, aliquots were thawed and diluted 1:1 in 1x reaction buffer without BSA (see gel-based general glycosylase assay protocol) and re-frozen in liquid nitrogen for storage. For experiments, the diluted aliquots were thawed and stored at 4 °C for up to one month. UV absorbance was used to estimate the concentration of AlkA, and the active concentration was determined by titration with a tight binding inhibitor (see below).

Preparation of Oligonucleotides

For this study 19- and 25-mer oligonucleotides were used (Figure 2-1), the 19-mers having

an asymmetrically placed lesion.[25] Oligonucleotides were not labeled unless noted in Materials and Methods, in which case a 5′ 6-fluorescein (FAM) or a 6-hexachlorofluorescein (HEX) label was present on the lesion-containing strand. Desalted DNA substrates were purchased from Integrated DNA Technologies (IDT) or the Keck Center at Yale University. Oligonucleotides were purified via denaturing PAGE, extracted, and desalted on reverse phase C18 columns (Sep-Pak, Waters). Concentrations were determined from the A_{260} and the calculated extinction coefficient. For εA substrates, 9400 $M^{-1}cm^{-1}$ was subtracted from the extinction coefficient of an undamaged substrate (adenine in place of εA) to account for the weaker absorbance of εA. The extinction coefficient of pyrrolidine substrates was calculated by replacing the pyrrolidine with dSpacer.

Abasic DNA was produced by incubating 25ThxC DNA with excess Δ80 AAG[20] at 37 °C for at least 50 turnovers. The DNA was phenol/chloroform extracted and desalted with an Illustra MicroSpin G-25 column (GE Healthcare) that had been equilibrated with annealing buffer (10 mM NaMES pH 6.5 and 50 mM NaCl). The concentration of the DNA was determined by the absorbance of the FAM label, and corrected for the purity as determined by denaturing PAGE (>95% pure).

Oligonucleotides were annealed with 1.2-fold excess of the complementary strand by heating to 95 °C for 3 min and cooling to 4 °C at a rate of 0.2 °C/s. Previous work shows that the excess complementary strand does not affect the observed rate constants.[25]

Gel-Based General Glycosylase Activity Assay

Unless otherwise noted, all discontinuous glycosylase activity assays were performed by incubating AlkA and a DNA substrate at 37 °C in reaction buffer (50 mM NaMES (pH 6.1), 100

mM ionic strength (controlled with NaCl), 1 mM EDTA, 1 mM DTT, 0.1 mg/mL BSA). If glycerol was present, 10% (v/v) was used. At varying time points the aliquots of the reaction were quenched in 0.2 M NaOH and placed on ice to prevent the base-catalyzed ring opening of εA. The samples were then heated for 12 min at 70 °C and varying volumes of loading buffer (98% formamide, 1 mM EDTA, bromophenol blue and xylene cyanol) was added to ensure 5-200 fmol of DNA were analyzed by denaturing PAGE (20% (w/v) acrylamide, 1× TBE, 6.6 M urea). Gels were imaged using a Typhoon Trio Fluorescence imager (GE Healthcare) with a 488 nm excitation and a 520 nm band-pass filter to detect fluorescein. The fraction of product for each time point was calculated by dividing the intensity of the product band by the sum of intensities of both product and substrate bands in each lane.

Determination of the Concentration of Active AlkA

A tight binding inhibitor (pyrrolidine-containing DNA) was used to titrate the amount of active AlkA.[25, 33] AlkA (100 or 200 nM) was incubated with excess 5′FAM-19ThxC substrate (500 nM) and varying concentrations of 5'HEX-25TpyC inhibitor (0-800 nM) under the glycosylase assay conditions described above. 19ThxC was used in place of 19TεC as AlkA binds hypoxanthine with a weaker affinity (data not shown), making the pyrrolidine DNA a better inhibitor. The fraction product versus time was fit by linear regression as described in the multiple turnover glycosylase activity protocol. The relative activity was calculated by normalizing to the reaction rate without inhibitor. The concentration of active AlkA was determined by using the quadratic binding equation (Eq 1), where K_d is the dissociation constant for pyr binding, and E and I represent AlkA and the pyr inhibitor, respectively. AlkA was found to be 85% active compared to the estimation that was based on the UV-absorbance (Figure A-1).

The active concentration was used for all subsequent experiments.

$$\frac{\mathrm{V_{obs}}}{\mathrm{V_{max}}} = 1 - \frac{K_\mathrm{d} + [\mathrm{E}] + [\mathrm{I}] - \sqrt{(K_\mathrm{d} + [\mathrm{E}] + [\mathrm{I}])^2 - 4[\mathrm{E}][\mathrm{I}]}}{2[\mathrm{E}]} \quad (1)$$

Multiple Turnover Glycosylase Activity

Excess 5′FAM-19TεC DNA (10-200 nM) was incubated with 1 nM AlkA under the glycosylase assay conditions. The first 8% of reactions (5% for 19ThxC reactions) were used to obtain the rate of reaction (V_{obs}), ensuring linearity as product inhibition occurs after these limits (data not shown).[25] Values of V_{obs}/[E] were plotted against the DNA concentration and fit by the Michaelis-Menten equation (Eq 2), where k_{cat} is the maximal turnover rate constant, S is the substrate, E is AlkA, and the K_m is the concentration at which the observed rate constant is half the maximal rate constant.

$$\frac{\mathrm{V_{obs}}}{[\mathrm{E}]} = \frac{k_\mathrm{cat}[\mathrm{S}]}{K_\mathrm{m} + [\mathrm{S}]} \quad (2)$$

Single Turnover Glycosylase Activity

Excess AlkA (at least 2-fold) was incubated with 5′FAM-19TεC DNA for single turnover reactions under the glycosylase assay conditions. DNA concentration ranged from 5 to 50 nM, and AlkA concentration ranged from 10 nM to 5 μM. In some cases, reactions were compared with the same concentration of AlkA and two different concentrations of DNA to ensure that single turnover conditions were met. The fraction product was fit by a single exponential (Eq 3), where *F* is the fraction product, *A* is the amplitude, k_{obs} is the observed single turnover rate constant, *t* is the reaction time, and *c* is the amount of pre-existing abasic DNA.

$$F = A[1 - \exp(-k_\mathrm{obs}t)] + c \quad (3)$$

The concentration dependence of the single turnover rate constant was fit by a hyperbolic dependence (Eq 4), where k_{max} is the maximal single turnover rate constant, E is AlkA, and the $K_{1/2}$ is the concentration of AlkA at which k_{obs} is 50% of the k_{max} value.

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

Equilibrium Inhibition by Undamaged DNA and the Abasic DNA Product

AlkA (5 nM) was incubated under normal glycosylase assay conditions with excess DNA. Ratios of inhibitor (25TAC or 5′FAM-25TabC) to substrate (5′FAM-19TεC) were varied, while the total [DNA] was kept constant (1 μM for undamaged DNA reactions, or 500 nM for abasic DNA reactions). Linear fits to the fraction product data were performed as described in the multiple turnover glycosylase activity protocol, and the observed rate was normalized to the reaction rate without inhibitor. A competitive inhibition model (Eq 5) was used to fit the [I]/[S] dependence, yielding the ratio of the K_m for substrate to the K_i for inhibitor (K_d for abasic DNA binding).[24] For the undamaged DNA, a microscopic K_d value was determined by multiplying the K_i value by the number (N) of nonspecific binding sites on the 25TAC inhibitor. N was calculated with Equation 6, in which L is the total length of the oligo, l is the site size or footprint of the enzyme (8 bps),[19] and the number of sites is doubled to account for both DNA strands. This calculation assumes that AlkA would bind with equal affinity to all nonspecific sites. For the 25TAC inhibitor, the number of binding sites was calculated to be 36 [2×(25-8+1)=36].

$$\frac{\mathrm{V_{obs}}}{\mathrm{V_{max}}} = \frac{1}{(\frac{[\mathrm{I}]}{[\mathrm{S}]} * \frac{K_\mathrm{m}}{K_\mathrm{i}} + 1)} \quad (5)$$

$$N = 2(L - l + 1) \quad (6)$$

εA Quenching by AlkA Under Steady State Conditions

To measure the stoichiometry of AlkA binding to εA, the εA fluorescence of 400 nM 19AεA DNA was monitored with varying concentrations of AlkA. A PTI QuantaMaster fluorometer using FeliX software was used to measure the εA fluorescence (excitation of 316 nm and emission of 408 nm; both 6 nm band-pass). Reactions were performed at 25 °C in a HEPES pH 7.5 reaction buffer (50 mM HEPES, 100 mM ionic strength (controlled with NaCl), 1 mM EDTA, 1 mM DTT) to slow excision of εA. The spectra were recorded within a minute of AlkA and DNA mixing. The quenching of the εA was normalized to the fluorescence of free 19AεA and fit to a 1:1 binding equation (Eq 7), where *F* is normalized fluorescence, *A* is the amplitude of fluorescence change, E and S are AlkA and the substrate, respectively, and the K_d is the dissociation constant of εA binding.

$$F = 1 - \frac{A\left\{(K_d + [E] + [S]) - \sqrt{(K_d + [E] + [S])^2 - 4[E][S]}\right\}}{2[S]} \quad (7)$$

Stopped-Flow Kinetics

Pre-steady state kinetics experiments were performed on a Hi-Tech SF-61DX2 Stopped-Flow System using Kinetic Studio software (TgK Scientific). The fluorescence of εA was observed using an excitation wavelength of 313 nm and a WG360 long-pass emission filter. Excess AlkA and 100 nM (final concentration) 19AεA or 19TεC DNA (in glycosylase assay buffer with no BSA) were mixed at 25 °C. The εA fluorescence was measured for 2 s, with the average of three independent measurements fit by a single exponential (Eq 3). The AlkA concentration dependence was fit with a hyperbolic function (Eq 4). The k_{max} is independent of AlkA

concentration and describes the nucleotide flipping step, which is an approach to equilibrium (Eq 8). The value of k_{on} was estimated by assuming that at saturating [AlkA] (850 nM), the $k_{on,obs}$ is at least 10-fold faster than the $k_{flip,obs}$ (Eq 9).

$$k_{max} = k_{flip,obs} = k_{flip} + k_{unflip} \quad (8)$$

$$k_{on,obs} = k_{on} * [AlkA] \geq 10 * k_{flip,obs} \quad (9)$$

To measure the dissociation rate from undamaged DNA, 5 μM 19AεA substrate was mixed with a preformed complex of 1 μM AlkA and 500 nM 25TAC. A single phase was observed up to 0.05 s and fit by a single exponential. The k_{off} value was calculated using Equation 10, where k_{obs} is the rate constant for dissociation of AlkA from undamaged DNA and subsequent binding to εA-DNA and $k_{flip,obs}$ is the rate constant of the association of free AlkA with the same εA-DNA.

$$\frac{1}{k_{obs}} = \frac{1}{k_{flip,obs}} + \frac{1}{k_{off}} \quad (10)$$

Pulse-Chase Glycosylase Assays

Single turnover assays were performed with 1 μM AlkA and 100 nM 5'FAM-19TεC substrate and fit by a single exponential as described for single turnover glycosylase activity. Pulse-chase reactions were aged 20 s prior to addition of 10 μM 25TpyC chase, and fit by a straight line with a slope of zero. Reactions with substrate and chase premixed before addition of AlkA and a no enzyme control were also performed, showing no εA cleavage. The observed rate constant for dissociation ($k_{off,obs}$) is given by Equation 11, where *A* is the amplitude in the pulse-chase reaction, k_{max} is the maximal rate from the single turnover reaction.

$$k_{off,obs} = \frac{k_{max}}{A} - k_{max} \quad (11)$$

Double Mixing Stopped-Flow Kinetics

Stopped flow experiments were performed as described above with the following changes. The AlkA and DNA were mixed and aged for 1 s or 5 s, then mixed with chase. The final concentrations after mixing were 425 nM AlkA, 200 nM 19AεA DNA, and 3 or 10 μM 25TpyC chase. The εA fluorescence was measured for 2 s, with averages of three measurements fit by a single exponential (Eq 3) giving a $k_{off,obs}$ value, defined in Equation 12. The k_{unflip} value was calculated according to Equation 13 (rearranged from Eq 8 and 12).

$$k_{\mathrm{off,obs}} = k_{\mathrm{unflip}}\left(\frac{k_{\mathrm{off}}}{k_{\mathrm{off}} + k_{\mathrm{flip}}}\right) \tag{12}$$

$$k_{\mathrm{unflip}} = \frac{k_{\mathrm{off,obs}} * k_{\mathrm{off}} + k_{\mathrm{off,obs}} * k_{\mathrm{flip,obs}}}{k_{\mathrm{off,obs}} + k_{\mathrm{off}}} \tag{13}$$

Results

AlkA-Catalyzed Excision of εA

To determine the kinetic framework for AlkA-catalyzed εA excision, both multiple and single turnover kinetics were measured using a gel-based discontinuous assay. Whereas the multiple turnover rate monitors the entire reaction, including product release, the single turnover reaction only monitors steps up to and including *N*-glycosidic bond hydrolysis, which is irreversible under these conditions (Scheme 1).

Most previous studies of AlkA have included glycerol as a buffer component, but we excluded it because AlkA catalyzes a side reaction with abasic sites to form a glycerol adduct.[34] Single turnover glycosylase assays in the presence and absence of 10% glycerol showed that AlkA is well behaved and slightly more active in the absence of glycerol (Figure A-2). We used an asymmetric 19-mer oligonucleotide containing a single site of damage (19TεC; Figure 2-1), because the binding of multiple AlkA molecules can be inhibitory when the lesion site is located at different distances from a DNA end.[25] This inhibitory behavior was previously observed for the excision of hypoxanthine (Hx) from a 25-mer DNA in buffers containing glycerol, and we confirmed that similar behavior is observed for the excision of εA from a 25-mer DNA in the absence of glycerol (Figure A-3). Therefore, the 19TεC substrate was employed for the

```
19TxC:          5′-CAT CCT x CCT TCT CTC CAT-3′
                3′-GTA GGA T GGA AGA GAG GTA-5′

19AxA:          5′-CAT CAA x AAT TCT CTC CAT-3′
                3′-GTA GTT T TTA AGA GAG GTA-5′

25TxC: 5′-CGA TAG CAT CCT x CCT TCT CTC CAT-3′
       3′-GCT ATC GTA GGA T GGA AGA GAG GTA-5′
```

Figure 2-1. Oligonucleotide substrates used in this study. The nomenclature #NxN describes the length, flanking bases, and lesion (x being ε = εA, Hx = hypoxanthine, pyr = pyrrolidine, A = adenine (undamaged), and ab = abasic). x-DNA denotes the full nucleotide, while x describes the nucleobase.

glycosylase kinetic assays that are described below.

Multiple turnover kinetics were first determined by incubating AlkA with excess DNA substrate. The initial rates of reaction were calculated by fitting the first 8% of reaction by a straight line (Figure 2-2A), because inhibition by the abasic product can be detected above that percentage (data not shown).[25] The dependence on the concentration of εA-DNA followed Michaelis-Menten behavior with a k_{cat} value of 0.24 ± 0.06 min^{-1} and a K_m value of 15 ± 13 nM (Figure 2-2B). In experiments with greater concentration of AlkA, there was no evidence of a burst phase (data not shown). This is similar to what was previously observed with Hx-DNA[25] and suggests that the rate-limiting step is *N*-glycosidic bond hydrolysis or a preceding step.

We next measured the single turnover glycosylase reaction with AlkA in excess over the 19TεC substrate. In all cases, the reaction progress curve followed a single exponential (Figure

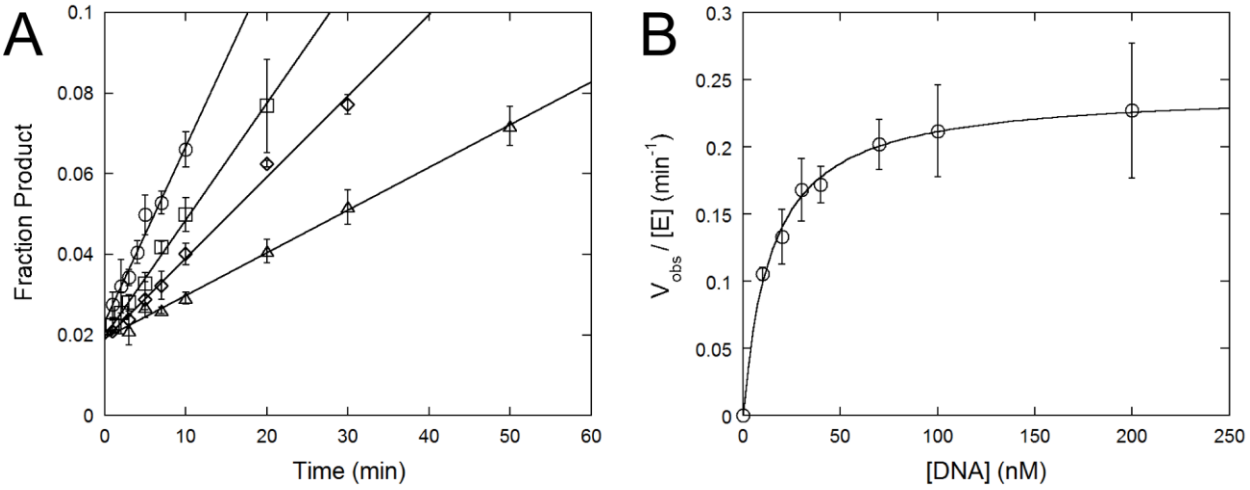


Figure 2-2. Multiple turnover excision of εA by AlkA. (A) Representative time course for excision of 19TεC substrate [40 nM (○), 70 nM (□), 100 nM (◊), and 200 nM (Δ)] by AlkA (1 nM). Linear fits were performed for the first 8% of reaction to ensure linearity and avoid product inhibition. The average of duplicate reactions is shown ± SD. (B) The substrate concentration dependence follows Michaelis-Menten kinetics with a k_{cat} value of 0.24 ± 0.06 min^{-1} and a K_m value of 15 ± 13 nM. The average ± SD is shown (n = 4). This K_M value is poorly defined, but the sensitivity of the assay precludes measurement of steady-state kinetics at lower concentrations of DNA. Note that a wider concentration range is possible for single-turnover kinetics (Figure 2-3B).

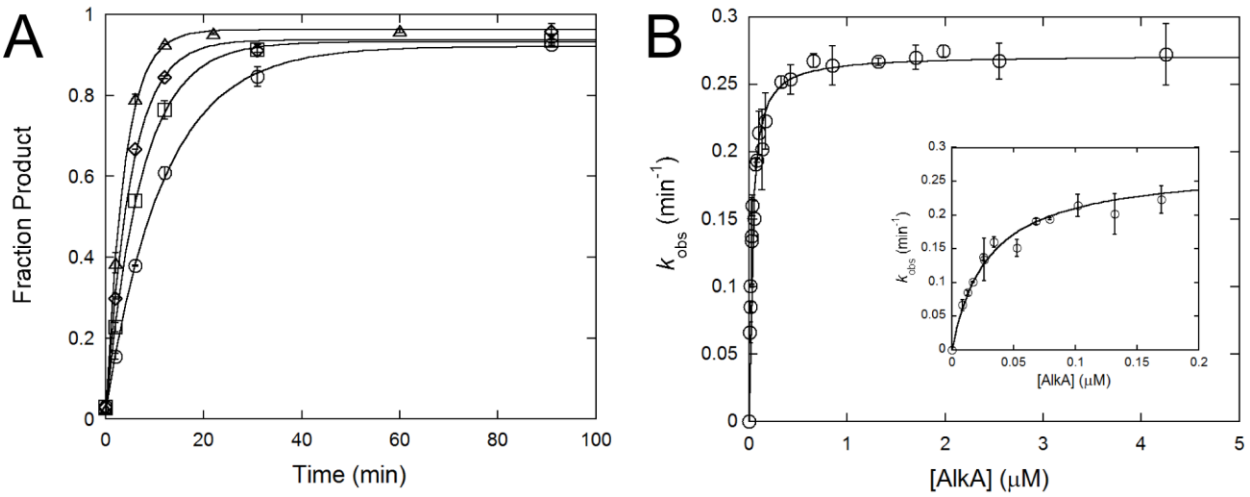


Figure 2-3. **Single turnover excision of εA by AlkA.** (A) Representative time course for single turnover excision of 19TεC substrate (5 nM or 50 nM) by varying concentrations of AlkA (12.75 nM (○), 25.5 nM (□), 68 nM (◊), and 2.55 μM(Δ)). The data were fit by a single exponential. The average of duplicate reactions is shown ± SD. (B) Hyperbolic dependence of the single turnover rate constant on the AlkA concentration with a k_{max} of 0.27 ± 0.01 min^{-1} and a $K_{1/2}$ value of 29 ± 3 nM. Inset: Expanded plot of the lowest AlkA concentrations. The average of duplicate reactions is shown ± SD.

2-3A), and the values of the observed single turnover rate constant (k_{obs}) were obtained for each concentration of enzyme. The AlkA concentration dependence was fit by a hyperbolic dependence (Figure 2-3B), with a maximal single turnover rate constant (k_{max}) of 0.27 ± 0.01 min^{-1} (0.0045 s^{-1}) and a $K_{1/2}$ value of 29 ± 3 nM. These kinetic parameters are almost identical to those observed in the multiple turnover experiments, suggesting that the same step is rate-limiting for single turnover and multiple turnover conditions. Although these data are consistent with either rate-limiting conformational change, such as nucleotide flipping, or rate-limiting *N*-glycosidic bond hydrolysis, direct measurement of nucleotide flipping supports the latter model (see below).

Affinity of AlkA for Abasic Product and Undamaged DNA

In order to measure the affinity of AlkA for binding to the abasic product and undamaged DNA, we performed multiple turnover experiments with mixtures of the 19TεC substrate and 25-

mer DNA competitors that either contained a single abasic site or were undamaged. With a constant amount of total DNA, competitive inhibition can be monitored as previously described.[24] The relative activity is plotted as a function of the ratio of inhibitor to substrate and the curve fit provides the relative K_m/K_i values (Figure 2-4). The K_i values were calculated using the independently determined K_m for εA-DNA (Figure 2-2B). For the abasic-containing DNA this K_i value of 1.2 ± 0.1 nM is equal to the dissociation constant (K_d). This tight binding to the abasic site corroborates the significant product inhibition that occurs under multiple turnover conditions. For the undamaged DNA, the observed inhibition constant is 33 ± 16 nM, which reflects the affinity of AlkA for the total number of nonspecific binding sites along the 25-mer undamaged DNA duplex. Assuming a site size of 8 base pairs,[19] there are 36 overlapping binding sites (N) on the undamaged DNA (Eq 6) with an average K_d value of 1.2 μM for an individual nonspecific binding site ($K_{d,nsDNA}$ = N×K_i = 36×33 nM = 1.2 μM). Additional experiments with different DNA sequences and longer length gave the same dissociation constant for an average

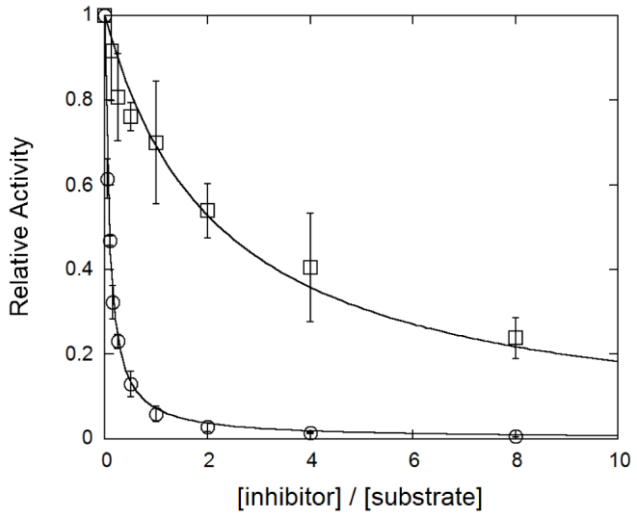


Figure 2-4. Affinity for abasic and undamaged DNA by AlkA determined with competition experiments. Multiple turnover competition between the 19TεC substrate and 25TabC (○) or 25TAC (□) inhibitor DNA. The initial rate of product formation (up to 8%) was measured and the relative activity (V_{obs}/V_{max}) was fit by the model for competitive inhibition (Eq 5). The average ± SD is shown (n ≥ 3). The K_m/K_i values for abasic and undamaged inhibitor were 12.8 ± 0.8 and 0.45 ± 0.15, respectively. Using the K_m value of 15 nM for εA-DNA (Figure 2-2B), $K_{i,obs}$ values of 1.2 ± 0.1 nM for the abasic product, and 33 ± 16 nM for undamaged DNA were calculated. This macroscopic value for undamaged DNA can be used to calculate a microscopic $K_{d,nsDNA}$ value of 1.2 μM for an individual nonspecific site (See Materials and Methods).

nonspecific site (Figure A-4; $K_{d,nsDNA} = 1.2 \pm 0.2$ μM).

Before proceeding to stopped-flow experiments, we repeated some of the glycosylase assays at a lower temperature. This is because we expected that DNA binding and nucleotide flipping might be very fast and would be more easily measured at 25 ^{o}C. The rate of the *N*-glycosidic bond cleavage step, measured with saturating AlkA, was found to be ~3-fold slower at 25 ^{o}C (0.080 ± 0.003 min^{-1} = 0.0013 s^{-1}; Figure A-5). The steady state kinetic parameters and the competition with undamaged DNA were also investigated at 25 ^{o}C and very similar K_m and K_d values were obtained (Figure A-6). These observations suggest that the main difference between AlkA activity at 25 ^{o}C and 37 ^{o}C is that the rate-limiting step (*N*-glycosidic bond cleavage) is slower at the lower temperature.

DNA Binding and Nucleotide Flipping by AlkA

We sought to use the intrinsic fluorescence of εA to monitor the DNA binding and nucleotide flipping steps carried out by AlkA. The difference in fluorescence between free DNA and bound DNA provides the signal to detect binding and nucleotide flipping. Our efforts were guided by previous studies of the binding of AAG to εA-DNA.[23] It is advantageous to use different sequence contexts in which the basal εA-DNA fluorescence is either high or low, because it is difficult to predict how strongly the fluorescence will be quenched upon binding to the protein. As the fluorescence of 19TεC is relatively low in duplex DNA, we also prepared a 19AεA substrate that has much higher basal fluorescence (Figure 2-1).[23]

To establish if AlkA binding causes a detectable change in εA-DNA fluorescence, we performed steady state fluorescence experiments by incubating different concentrations of AlkA with the 19AεA DNA. Fluorescence measurements were conducted after enzyme and substrate

were incubated for one minute to give time for binding to occur, but prior to *N*-glycosidic bond cleavage. The fluorescence of the εA-DNA was strongly quenched by one equivalent of AlkA (Figure 2-5A), suggesting that a monomer of AlkA tightly binds to each εA lesion site. It is likely that the decrease in εA-DNA fluorescence is attributed to interaction between the flipped out εA base and the tryptophan residues that line the AlkA active site.[19]

We next used stopped-flow fluorescence to monitor the kinetics of DNA binding and nucleotide flipping. When the 19AεA substrate was rapidly mixed with excess AlkA, a fast quenching of εA fluorescence was detected that followed a single exponential (Figure 2-5B) and exhibited good signal to noise (Figure 2-5C). The rate constants calculated for each AlkA concentration were roughly fit by a hyperbolic function, but the observed rate constant was almost fully saturated even at the lowest concentration of AlkA (Figure 2-5D). The concentration independence at high concentration of AlkA suggests that the εA quenching step being observed corresponds to the nucleotide flipping step ($k_{flip,obs}$ = 242 ± 11 s^{-1}), as a DNA binding event would be dependent on AlkA concentration. The observed rate constant for formation of the flipped out complex is equal to the sum of the forward (k_{flip}) and reverse (k_{unflip}) rate constants for flipping (Eq 8). Similar results were obtained with the 19TεC substrate, but the fluorescence change was much smaller (Figure A-7). Although we cannot tell from these data whether or not there is a detectable signal for binding of AlkA prior to nucleotide flipping, the rapid concentration-independent nucleotide flipping establishes that DNA binding is very fast. We can estimate a lower limit for the association rate constant (k_{on}) of 3×10^9 $M^{-1}s^{-1}$, because binding must be at least 10-fold faster than the concentration-independent change in fluorescence (Figure 2-5D; Eq 9). There is no evidence for the formation of a nonspecific AlkA•DNA complex, suggesting that if this forms it quickly converts to the specific complex. However, multiple AlkA

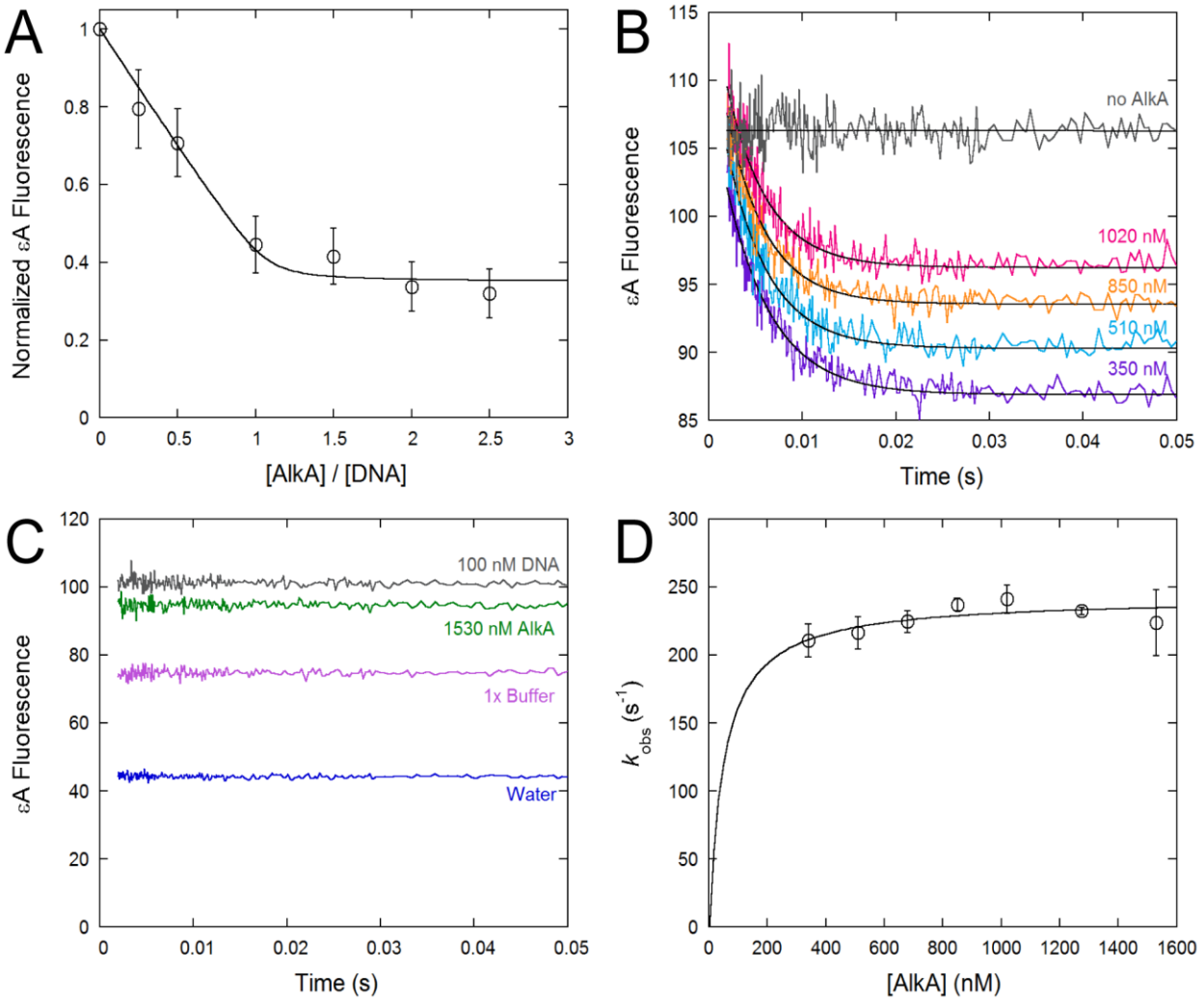


Figure 2-5. Quenching of εA fluorescence by AlkA indicates rapid nucleotide flipping. (A) Steady state quenching of εA-DNA fluorescence was measured using 400 nM 19AεA DNA and increasing concentrations of AlkA. The normalized fluorescence was fit by a 1:1 binding model (Eq 7). The average ± SD is shown (n = 4). (B) Representative time-dependent changes in εA-DNA fluorescence as observed by single mixing stopped-flow with 19AεA substrate (100 nM) and varying excess AlkA. The fluorescence is reported in arbitrary units and each trace is the average of three shots. Data were fit by a single exponential. (C) Controls to evaluate the signal to noise. The fluorescence signal of 100 nM 19AεA substrate was set to 100, and the signal from water, buffer, or the highest [AlkA] alone are shown. (D) The k_{obs} values from single exponentials were plotted versus AlkA concentration and fit with a hyperbolic function. This fit does not yield information regarding the $K_{1/2}$ for AlkA binding, because there is not sufficient signal to test lower concentrations of AlkA. The maximal value of 242 ± 11 s^{-1} is independent of AlkA concentration above 850 nM AlkA. The average ± SD is shown ($n \geq 2$).

proteins can bind under conditions of excess AlkA and this could obscure a slower searching process. Therefore, it can be informative to test conditions of excess DNA (see below).

Nucleotide Unflipping and Dissociation from DNA

To test whether AlkA is committed to catalysis once it binds εA-DNA, we performed a single turnover pulse-chase assay. AlkA was incubated with 19TεC substrate for 20 s before the addition of excess tight-binding pyrrolidine (pyr-DNA) chase. AlkA will thus be bound to the flipped out εA as formation of the specific complex occurs on the millisecond time scale (Figure 2-5B), but excision occurs on the minute time scale (Figure 2-3). No εA excision was detected after the addition of chase (Figure 2-6A), indicating that the specific AlkA complex is not committed to base excision and rapidly dissociates. Assuming that 5% product could have been

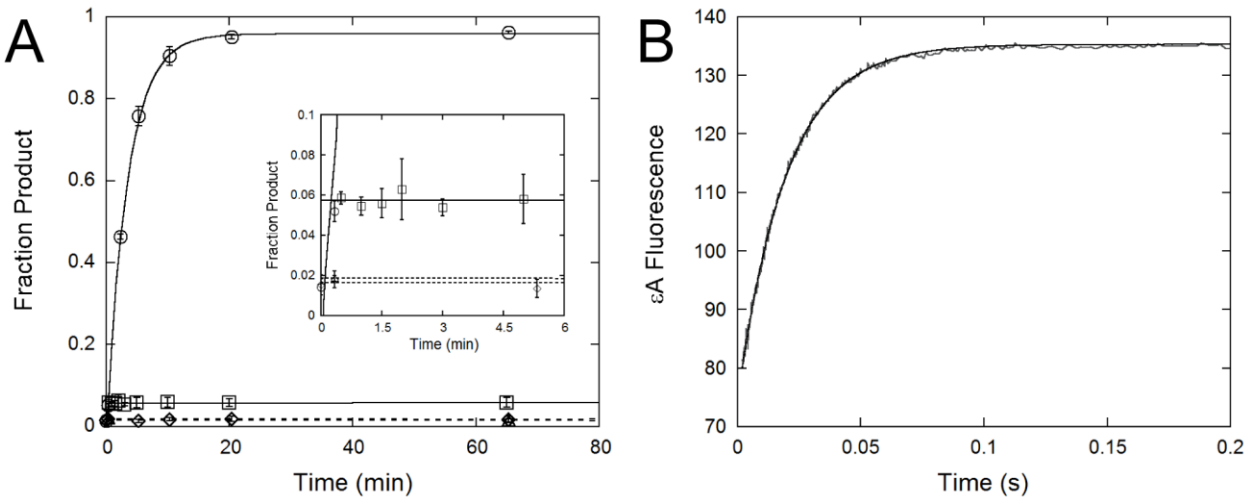


Figure 2-6. AlkA unflips εA and dissociates rapidly. (A) Time course for single turnover pulse-chase assay. 19TεC substrate (100 nM) was mixed with 850 nM AlkA, aged for 20 s, and mixed with 10 μM 25TpyC chase. The reaction without chase (○) was fit by a single exponential matching that in Figure 2-1A. The reaction with chase (□) was fit by a straight line, as zero percent of the reaction progressed after the 20s age time. No enzyme (Δ) and premixed εA and chase controls (◊) show no reaction occurring. Inset: Low fraction product data showing that no εA excision occurred in the presence of chase. The average ± SD is shown (n = 3). (B) Representative εA-DNA fluorescence from double mixing stopped-flow assay in which AlkA (425 nM) and 19AεA DNA (200 nM) was mixed, aged for 5 s, and mixed with 25TpyC chase (3 μM). Data were fit by a single exponential with a $k_{off,obs}$ value of 51 s^{-1}. Fluorescence is reported in arbitrary units and is the average of three shots. Replicates give an average $k_{off,obs}$ value of 52 ± 2 s^{-1} (Figure A-8).

readily observed, the value of the observed rate constant for dissociation of the εA-DNA•AlkA complex ($k_{off,obs}$) must be at least 20-fold faster than the rate constant for *N*-glycosidic bond hydrolysis (Eq 11). A lower limit of 0.09 s^{-1} can therefore be assigned for $k_{off,obs}$.

Given that AlkA is not committed to catalysis, and all of the substrate dissociates in the presence of chase, we turned to double mixing stopped-flow experiments to directly measure the dissociation of substrate. These experiments were performed with ratios of reagents similar to those used in the pulse-chase assay. AlkA and 19AεA were mixed and aged for 5 s to allow for the formation of the specific complex, which was subsequently chased with pyr-DNA competitor. A single exponential increase in εA fluorescence was observed as the εA-DNA dissociates from AlkA (Figure 2-6B; $k_{off,obs} = 52 \pm 2\ s^{-1}$). The length of the age time (1 or 5 s) and the concentration of chase (3 or 10 μM) did not affect the value of $k_{off,obs}$ that was obtained (Figure A-8).

We hypothesized that the slow step in dissociation of εA-DNA would correspond to unflipping of the εA lesion, followed by rapid dissociation from DNA. However, if dissociation from undamaged DNA was on the same time scale as flipping then this could complicate the interpretation. Therefore, we also performed a stopped-flow experiment to assay the dissociation of AlkA from undamaged DNA (Figure 2-7). In this experiment, the nonspecific complex is first formed between AlkA and undamaged DNA and then mixed with a saturating amount of 19AεA substrate. The signal of εA-DNA binding to AlkA is monitored, but the observed rate constant is dependent on both AlkA dissociation and association. The contribution from the dissociation step can be determined by comparing this rate constant to the rate constant obtained for free AlkA binding to εA-DNA under the same conditions. In the stopped-flow experiment with free AlkA and excess DNA the εA fluorescence is quenched in a single phase with an observed rate

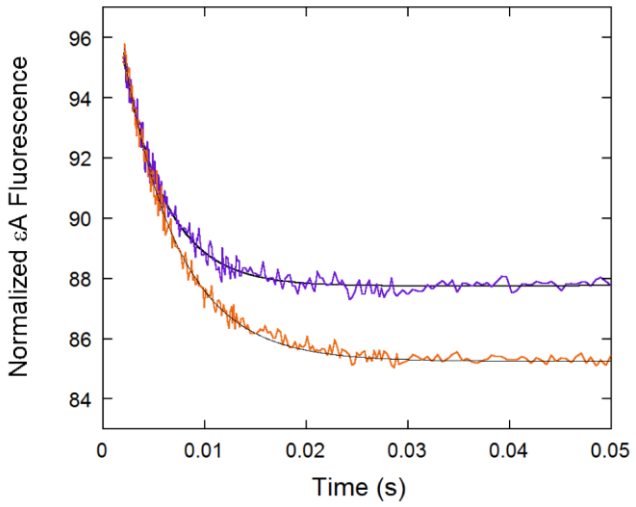


Figure 2-7. Association of AlkA under conditions of excess DNA and measurement of dissociation from nonspecific DNA. Representative εA-DNA fluorescence when AlkA (1 μM) was mixed with excess 19AεA substrate (5 μM) in the stopped-flow (purple). The single phase was fit by a single exponential and gave a k_{obs} value of 233 ± 11 s^{-1}. The nonspecific AlkA•DNA complex was formed by pre-incubating AlkA with 25TAC DNA (500 nM). This complex was mixed with 19AεA DNA (orange), showing a single exponential with a k_{obs} value of 186 ± 7 s^{-1}. Fluorescence traces are the average of triplicate reactions from a single day and the arbitrary fluorescence was normalized by dividing by the expected value at the time of mixing and multiplying by 100. This experiment was repeated and the reported rate constants reflect the average ± SD (n = 3).

constant of 233 ± 11 s^{-1} (Figure 2-7; purple trace). This value is almost identical to the value that was observed with excess AlkA protein ($k_{flip,obs}$ = 242 ± 11 s^{-1}; Figure 2-4). This observation suggests that the searching steps (k_{find}) performed by a single AlkA molecule are very rapid. When the nonspecific AlkA•DNA complex was mixed with excess εA-DNA, the single phase was slowed slightly with a k_{obs} value of 186 ± 7 s^{-1}. The modest reduction in rate indicates that AlkA dissociates very quickly from undamaged DNA with a rate constant of 920 s^{-1} (Eq 10). This confirms that the unflipping step is almost fully rate-limiting for the dissociation of AlkA from the specific complex with εA-DNA.

Calculation of the Kinetic and Thermodynamic Framework

Many of the experiments that were performed do not directly yield microscopic rate constants, due to reversibility or to multiple steps that are partially rate-limiting. However, these independent experiments could be combined to determine the individual kinetic parameters (Figure 2-8). These calculations are briefly described below and the equations and additional

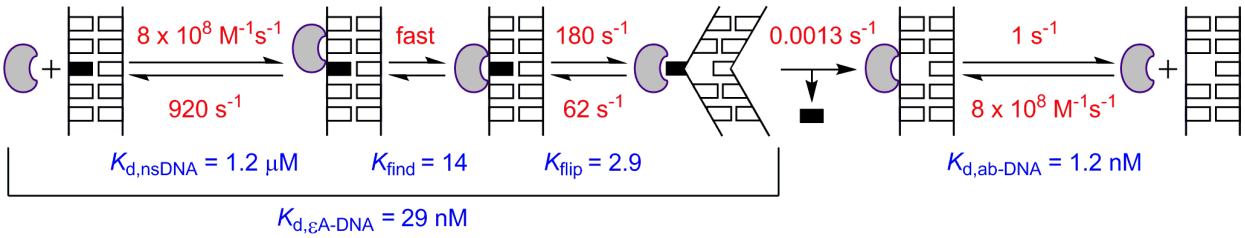


Figure 2-8. Minimal kinetic mechanism for AlkA-catalyzed excision of εA. Rate constants are shown in red and equilibrium constants are shown in blue. The internal equilibria, K_{find} for locating the site of damage, and K_{flip} for flipping out the damaged nucleotide, are expressed in the forward direction. We assume that AlkA has a similar mechanism for binding abasic DNA as for binding damaged nucleotides, with nonspecific binding, fast searching, and flipping of the abasic nucleotide. These steps are omitted for simplicity. The affinity for the abasic site is ~1000-fold tighter than for a nonspecific site, which includes both the equilibrium for flipping and the equilibrium for finding the site. The values describe the reaction at 25 °C (see Appendix for more information).

details are provided in the Materials and Methods.

The association and dissociation steps for AlkA are very fast and the most poorly defined parameters that we have measured. The dissociation of undamaged DNA was calculated to be 920 s^{-1} (Figure 2-7; see text above). Using this value and the K_d value for nonspecific sites (1.2 μM) that was determined by competition (Figure 2-4), an association rate constant of 8×10^8 $M^{-1}s^{-1}$ can be calculated ($k_{on} = k_{off}/K_d$). This value is consistent with the 3×10^9 $M^{-1}s^{-1}$ limit estimated from the rapid binding of AlkA with εA-DNA (Figure 2-5D). Assuming that AlkA binds to its εA substrate, abasic DNA and undamaged DNA at the same rate, we can use the K_d value of the abasic product (1.2 nM; Figure 2-4) to calculate a value of 1 s^{-1} for the product dissociation $k_{off} = k_{on} \times K_d$). This calculated dissociation rate constant is much faster than the steady state k_{cat} value of 0.004 s^{-1}, consistent with the conclusion that *N*-glycosidic bond hydrolysis is rate-limiting for multiple and single turnover reactions. After AlkA binds to a nonspecific site it must find the site of damage. These DNA searching steps (k_{find}) are all much faster than the rate of nucleotide flipping, because identical reaction rates are observed under conditions of excess DNA (where each protein needs to conduct its search independently; Figure

2-7) and conditions of excess protein (where the searching time will be accelerated by having many proteins searching; Figure 2-5D).

The rates and equilibrium constants for nucleotide flipping can be calculated from the forward and reverse binding reactions. The nucleotide unflipping step is the rate-limiting step in the dissociation reaction, as the $k_{off,obs}$ value of 52 s^{-1} (Figure 2-6B) is much slower than the k_{off} value of 920 s^{-1}. Combining these results with the observed rate constant for flipping in the forward direction ($k_{flip,obs}$ = 242 s^{-1}; Figure 2-5D) yields a value of 62 s^{-1} for k_{unflip} (Eq 13). By subtracting the k_{unflip} value from the observed flipping rate constant, the forward rate constant for flipping (k_{flip}) is calculated to be 180 s^{-1} (Eq 8). The equilibrium constant for flipping of 2.9 is defined as the ratio of the nucleotide flipping and unflipping rate constants ($K_{flip} = k_{flip}/k_{unflip}$). This indicates that the extrahelical recognition complex is only slightly favorable for AlkA.

The stopped-flow fluorescence studies demonstrate that DNA binding, searching and nucleotide flipping are all much faster than base excision. Therefore, the single turnover rate constant is approximately equal to the rate constant for *N*-glycosidic bond cleavage (k_{chem}) and the $K_{1/2}$ is equal to the K_d for εA-DNA (Figure 2-3B). For this multi-step binding mechanism, the overall dissociation constant for the AlkA•εA-DNA complex is the product of the individual equilibria along the reaction coordinate ($K_{d,\varepsilon A\text{-}DNA} = K_{d,nsDNA} \times 1/K_{find} \times 1/K_{flip}$). Thus, we were able to obtain an estimate for the equilibrium between non-specific sites and the lesion site (K_{find} = 14; Figure 2-8). This indicates that AlkA is able to bind significantly more tightly to an εA•T site than to a nonspecific site even prior to the nucleotide flipping step.

Discussion

We report the minimal kinetic and thermodynamic framework for the AlkA-catalyzed excision of εA, including measurements for DNA binding and nucleotide flipping. This framework reveals some clear differences in substrate recognition by AlkA and AAG, two glycosylases that independently evolved to have similar functions in protecting the genome against alkylative DNA damage.

Minimal AlkA Mechanism

Previous studies of AlkA have assumed that the slow glycosylase activity reflects rate-limiting *N*-glycosidic bond hydrolysis, however it was not possible to rule out the alternative model that binding or nucleotide flipping is rate-limiting.[22, 25] By directly measuring the rates for binding and nucleotide flipping with εA-DNA we confirmed that hydrolysis is the rate-limiting step for AlkA. All of the other steps in the AlkA mechanism are much faster than this step, which ensures that εA-DNA is sampled and rejected many times prior to being excised (Figure 2-8). This mechanism is consistent with a generalist strategy of removing many different damaged bases with a wide range of shapes and chemical properties.

The overall affinity for εA-DNA is 40-fold tighter than for a nonspecific site, demonstrating specific recognition of this damaged base. However, given the vast excess of undamaged sites this specificity is unlikely to allow for efficient targeting of εA lesions in vivo. This is consistent with the results from deletion experiments suggesting that direct repair is responsible for the majority of εA repair in *E. coli*.[32] Nevertheless, the origin of this 40-fold specificity is interesting

and potentially informative for understanding how AlkA recognizes other lesions. We determined a value of K_{flip} = 3, which indicates that the flipping of εA into the active site of AlkA only provides 3-fold stabilization relative to a bound state in which εA is not flipped out. By independently measuring the affinity of undamaged DNA and the overall equilibrium constant for εA-DNA binding, we can calculate a theoretical value for binding to the damaged site (K_{find} = 14). We do not know the nature of this early recognition complex, but it may take advantage of DNA intercalation and/or DNA bending to achieve tighter binding.

Our kinetic experiments reveal that AlkA searches DNA very rapidly. Many other DNA glycosylases use facilitated diffusion to search DNA for sites of damage, and in most cases strong electrostatic interactions are an important feature.[35-38] However, the AlkA DNA binding interface is not positively charged and few electrostatic DNA contacts are observed in crystal structures of AlkA•DNA complexes.[19] This raises the possibility that the searching mechanism of AlkA may be quite different from those employed by other glycosylases.

Biochemical assays suggest that a major role of AlkA is to remove positively charged alkylated bases bearing a destabilized *N*-glycosidic bond. The substrates have sites of alkylation that include N^3 and N^7 of purines and O^2 of pyrimidines. [9, 39] It appears that AlkA has adopted a strategy of providing similar rate enhancement for excision of these damaged bases and undamaged nucleotides that have more stable *N*-glycosidic bonds.[13, 22] For this mechanism to operate, it is important that AlkA rapidly search DNA and sample nucleotides in its active site. The kinetic parameters that we have described for the excision of εA support this model. Although this mechanism provides very broad protection, it does have a cost associated with it, in the removal of undamaged bases.[13] Indeed, overexpression of AlkA causes increased frequency of mutations attributed to larger numbers of abasic sites.[13, 40] *E. coli* minimizes this

cost by controlling the expression of AlkA, inducing its expression under conditions of chronic exposure to DNA alkylation damage as part of the ada transcriptional response.[15]

We found that AlkA binds more than 20-fold more tightly to the abasic product than to the εA lesion. This suggests that AlkA flips out the abasic site to allow for more favorable contacts with the DNA and it is likely that the crystal structure of AlkA bound to the abasic site analog 1-aza, 2'-deoxyribose is a reasonable model for this complex. In this structure the DNA is sharply bent, the sugar is flipped out and a conserved leucine (L125) intercalates into the DNA duplex.[19] Although it has been proposed that DNA glycosylases protect reactive abasic sites from undesirable reactions, it has been observed that AlkA catalyzes promiscuous reactions of abasic sites with alcohols to generate *O*-glycosidic DNA adducts.[34] Despite this tight binding, the observed dissociation is very fast and unlikely to limit the rate of AlkA catalysis on known substrates.

Comparisons Between AlkA and AAG

AlkA and AAG independently evolved to remove an extensive set of alkylated and oxidatively damaged bases, but they operate under different physiological niches. Both glycosylases show in vitro activity on the cyclic adduct εA,[11] allowing us to directly compare their mechanisms.

Given their independent evolution, it is striking that there are some structural parallels between AlkA and AAG. Both enzymes have conserved carboxylate groups to position a water molecule and directly hydrolyze the *N*-glycosidic bond, and both use DNA intercalation to stabilize a flipped-out lesion.[6] However, the three-dimensional structure and identity of the residues that contact the DNA are completely different for the two enzymes. Remarkably, the

rates of *N*-glycosidic bond cleavage are almost identical when measured under similar reaction conditions (Table A-1).[26] In contrast, the affinities for εA-DNA are dramatically different. AlkA binds εA-DNA relatively weakly with a K_d value of 29 nM compared to 20 pM for AAG.[23] The difference in binding makes AAG 1000-fold more efficient than AlkA at εA removal (Table A-1). Crystal structures of AAG bound to εA-DNA reveal favorable contacts with the εA nucleobase, including a hydrogen bond between a backbone amide and the N^6 of εA.[18] This suggests that AAG has evolved to recognize εA as an important physiological substrate, whereas AlkA exhibits fortuitous activity. Whereas DNA binding is similarly fast for both enzymes, the rate constants for flipping and unflipping are very different. Flipping is 50-fold slower and unflipping is 40,000-fold slower for AAG than for AlkA. This reflects an 800-fold difference in the equilibrium constant for nucleotide flipping (Table A-1). Furthermore, the slow unflipping by AAG causes a strong commitment to catalysis such that most DNA binding events are productive. As discussed above, AlkA exhibits rapid reversible εA-DNA binding without any commitment to catalysis, allowing it to quickly sample a structurally diverse set of damaged bases.

Recently, a mutant of AAG was described in which the intercalating residue (Y162) was mutated to alanine.[26] This residue is critical for stabilizing the extrahelical εA complex. The Y162A AAG mutant is no longer committed to excision of εA and its behavior is remarkably similar to that of WT AlkA. The rate constant for nucleotide flipping by Y162A AAG is virtually identical to AlkA (170 s^{-1} compared to 180 s^{-1}; Table A-1). The Y162A AAG mutant also exhibits a K_{flip} value of 17, which is much closer to that of AlkA than WT AAG. Thus, it appears that much of the kinetic differences between AlkA and AAG can be attributed to the intercalating tyrosine that is found in AAG. In AAG this residue allows for much more stable

binding of the extrahelical lesion, but also slows the nucleotide flipping step. Although it is clearly an over-simplification to attribute the differences between enzymes to a single interaction, it is intriguing to note that kinetic studies of other DNA glycosylases show a similar trend. Uracil DNA glycosylase, like AlkA, uses an intercalating leucine and exhibits very rapid nucleotide flipping (k_{flip} = 1200 s^{-1}).[26,27] Formamidopyrimidine DNA glycosylase employs an intercalating phenylalanine that is more similar to the tyrosine that AAG uses, and this enzyme exhibits slower nucleotide flipping (k_{flip} = 1–12 s^{-1}).[41-43] Additional mutational studies are needed to understand the mechanistic differences between the different families of DNA repair enzymes.

Conclusions

By directly measuring DNA binding and nucleotide flipping we have gained new insights into how AlkA recognizes DNA damage. We expect that the fast flipping and rapid searching that we have observed for an εA-DNA substrate will extend to other substrates of AlkA, and the *N*-glycosidic bond cleavage step is likely to be rate-limiting for most substrates. By rapidly and reversibly sampling sites, AlkA is able to maintain a remarkably promiscuous active site pocket that accommodates damaged bases of very different sizes. This strategy is ideally suited to alkyl adducts that have destabilized *N*-glycosidic bonds, because little catalytic power is needed to perform the reaction on a biologically relevant time scale.

References

1. Lindahl, T. (1993) Instability and decay of the primary structure of DNA, *Nature 362*, 709-715.
2. Marnett, L. J., Riggins, J. N., and West, J. D. (2003) Endogenous generation of reactive oxidants and electrophiles and their reactions with DNA and protein, *J Clin Invest 111*, 583-593.
3. Beranek, D. T. (1990) Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents, *Mutat Res 231*, 11-30.
4. Sedgwick, B. (2004) Repairing DNA-methylation damage, *Nat Rev Mol Cell Biol 5*, 148-157.
5. O'Brien, P. J. (2006) Catalytic promiscuity and the divergent evolution of DNA repair enzymes, *Chem Rev 106*, 720-752.
6. Hollis, T., Lau, A., and Ellenberger, T. (2001) Crystallizing thoughts about DNA base excision repair, *Prog Nucleic Acid Res Mol Biol 68*, 305-314.
7. Denver, D. R., Swenson, S. L., and Lynch, M. (2003) An evolutionary analysis of the helix-hairpin-helix superfamily of DNA repair glycosylases, *Mol Biol Evol 20*, 1603-1611.
8. Eisen, J. A., and Hanawalt, P. C. (1999) A phylogenomic study of DNA repair genes, proteins, and processes, *Mutat Res 435*, 171-213.
9. Thomas, L., Yang, C. H., and Goldthwait, D. A. (1982) Two DNA glycosylases in Escherichia coli which release primarily 3-methyladenine, *Biochemistry 21*, 1162-1169.
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. Saparbaev, M., Kleibl, K., and Laval, J. (1995) Escherichia coli, Saccharomyces cerevisiae, rat and human 3-methyladenine DNA glycosylases repair 1,N6-ethenoadenine when present in DNA, *Nucleic Acids Res 23*, 3750-3755.
12. Saparbaev, M., and Laval, J. (1994) Excision of hypoxanthine from DNA containing dIMP residues by the Escherichia coli, yeast, rat, and human alkylpurine DNA glycosylases, *Proc Natl Acad Sci U S A 91*, 5873-5877.
13. 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.
14. 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.
15. Evensen, G., and Seeberg, E. (1982) Adaptation to alkylation resistance involves the induction of a DNA glycosylase, *Nature 296*, 773-775.
16. Sedgwick, B., and Lindahl, T. (2002) Recent progress on the Ada response for inducible repair of DNA alkylation damage, *Oncogene 21*, 8886-8894.
17. Samson, L., and Cairns, J. (1977) A new pathway for DNA repair in Escherichia coli,

Nature 267, 281-283.
18. 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.
19. Hollis, T., Ichikawa, Y., and Ellenberger, T. (2000) DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, Escherichia coli AlkA, *EMBO J 19*, 758-766.
20. 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.
21. 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.
22. 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.
23. 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.
24. 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.
25. Zhao, B., and O'Brien, P. J. (2011) Kinetic mechanism for the excision of hypoxanthine by Escherichia coli AlkA and evidence for binding to DNA ends, *Biochemistry 50*, 4350-4359.
26. Hendershot, J. M., and O'Brien, P. J. (2014) Critical role of DNA intercalation in enzyme-catalyzed nucleotide flipping, *Nucleic Acids Res*. 42, 12681-90.
27. Pandya, G. A., and Moriya, M. (1996) 1,N6-ethenodeoxyadenosine, a DNA adduct highly mutagenic in mammalian cells, *Biochemistry 35*, 11487-11492.
28. Nair, J., Barbin, A., Velic, I., and Bartsch, H. (1999) Etheno DNA-base adducts from endogenous reactive species, *Mutat Res 424*, 59-69.
29. Gros, L., Ishchenko, A. A., and Saparbaev, M. (2003) Enzymology of repair of etheno-adducts, *Mutat Res 531*, 219-229.
30. Calvo, J. A., Meira, L. B., Lee, C. Y., Moroski-Erkul, C. A., Abolhassani, N., Taghizadeh, K., Eichinger, L. W., Muthupalani, S., Nordstrand, L. M., Klungland, A., and Samson, L. D. (2012) DNA repair is indispensable for survival after acute inflammation, *J Clin Invest 122*, 2680-2689.
31. Delaney, J. C., Smeester, L., Wong, C., Frick, L. E., Taghizadeh, K., Wishnok, J. S., Drennan, C. L., Samson, L. D., and Essigmann, J. M. (2005) AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo, *Nature structural & molecular biology 12*, 855-860.
32. Maciejewska, A. M., Sokolowska, B., Nowicki, A., and Kusmierek, J. T. (2011) The role of AlkB protein in repair of 1,N(6)-ethenoadenine in Escherichia coli cells, *Mutagenesis 26*, 401-406.
33. Scharer, O. D., Nash, H. M., Jiricny, J., Laval, J., and Verdine, G. L. (1998) Specific binding of a designed pyrrolidine abasic site analog to multiple DNA glycosylases, *J Biol Chem 273*, 8592-8597.
34. Admiraal, S. J., and O'Brien, P. J. (2013) DNA-N-glycosylases process novel O-

glycosidic sites in DNA, *Biochemistry 52*, 4066-4074.
35. Ganesan, A. K., Seawell, P. C., Lewis, R. J., and Hanawalt, P. C. (1986) Processivity of T4 endonuclease V is sensitive to NaCl concentration, *Biochemistry 25*, 5751-5755.
36. Hedglin, M., Zhang, Y., and O'Brien, P. J. (2013) Isolating contributions from intersegmental transfer to DNA searching by alkyladenine DNA glycosylase, *J Biol Chem 288*, 24550-24559.
37. Schonhoft, J. D., Kosowicz, J. G., and Stivers, J. T. (2013) DNA translocation by human uracil DNA glycosylase: role of DNA phosphate charge, *Biochemistry 52*, 2526-2535.
38. Francis, A. W., and David, S. S. (2003) Escherichia coli MutY and Fpg utilize a processive mechanism for target location, *Biochemistry 42*, 801-810.
39. Bjelland, S., Birkeland, N. K., Benneche, T., Volden, G., and Seeberg, E. (1994) DNA glycosylase activities for thymine residues oxidized in the methyl group are functions of the AlkA enzyme in Escherichia coli, *J Biol Chem 269*, 30489-30495.
40. Frosina, G. (2000) Overexpression of enzymes that repair endogenous damage to DNA, *Eur J Biochem 267*, 2135-2149.
41. Fedorova, O. S., Nevinsky, G. A., Koval, V. V., Ishchenko, A. A., Vasilenko, N. L., and Douglas, K. T. (2002) Stopped-flow kinetic studies of the interaction between Escherichia coli Fpg protein and DNA substrates, *Biochemistry 41*, 1520-1528.
42. Koval, V. V., Kuznetsov, N. A., Zharkov, D. O., Ishchenko, A. A., Douglas, K. T., Nevinsky, G. A., and Fedorova, O. S. (2004) Pre-steady-state kinetics shows differences in processing of various DNA lesions by Escherichia coli formamidopyrimidine-DNA glycosylase, *Nucleic Acids Res 32*, 926-935.
43. 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.

Appendix A

Figures and Table to Accompany Chapter II

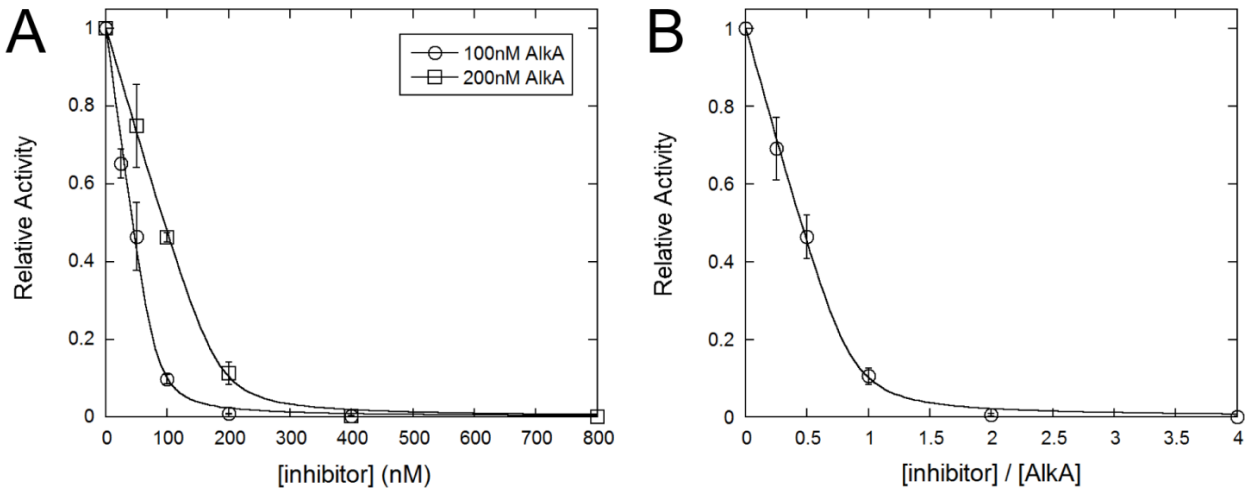


Figure A-1. Determination of the concentration of active AlkA. (A) Titrations using varying concentrations of the 25-mer pyrrolidine inhibitor were performed with either 100 or 200 nM AlkA and excess 19ThxC (500 nM) substrate under multiple turnover conditions. The initial rate of product formation was measured and the relative activity (V_{obs}/V_{max}) was plotted versus the concentration of inhibitor. The average ± SD is shown (n = 3). (B) The relative activity of the two AlkA concentrations were averaged and plotted versus the ratio of inhibitor to AlkA. The quadratic binding equation gave an 85% activity of AlkA (Eq 1). The average ± SD is shown (n = 6).

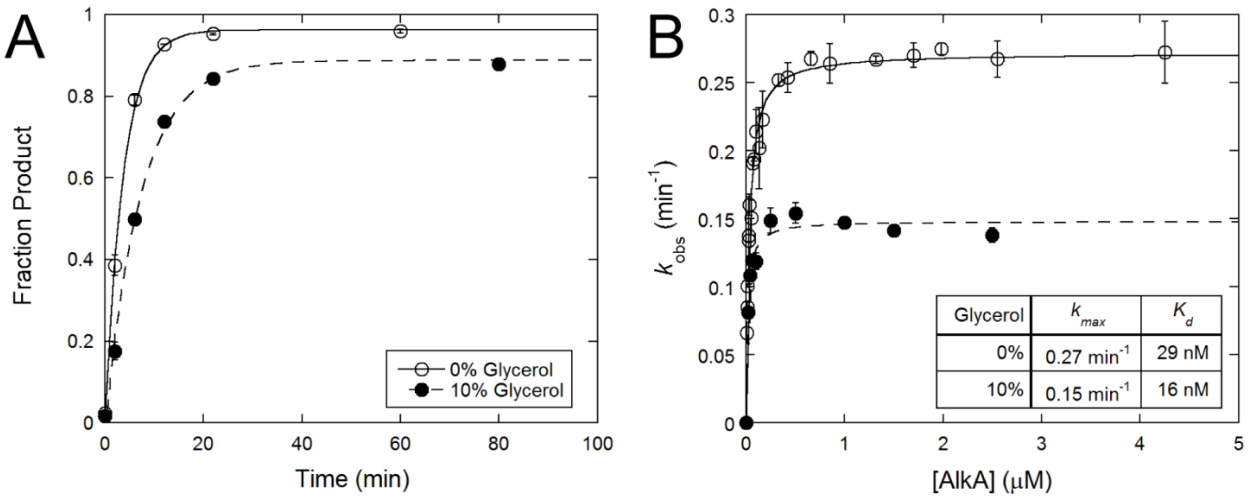


Figure A-2. Glycerol has an inhibitory effect on AlkA-catalyzed excision of εA. (A) Representative time course of single turnover excision of 19TεC substrate (50 nM) by excess AlkA (2.5 μM), with or without 10% (v/v) glycerol present in the reaction buffer. The data were fit by a single exponential (Eq 3). The average of duplicate experiments ± SD is shown. Note that the apparent endpoint of the reaction is decreased in the presence of glycerol due to the AlkA-catalyzed formation of a glycerol-DNA adduct that is stable under the conditions of the glycosylase assay.[1] (B) Hyperbolic dependence of the single-turnover rate constant on the AlkA concentration. The average ± SD is shown ($n \geq 2$). The data obtained in the absence of glycerol are taken from Figure 2-3B for comparison.

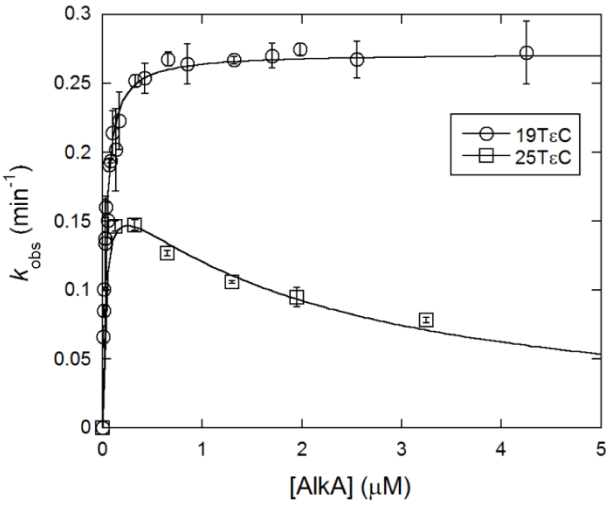


Figure A-3. Single turnover excision with 25TεC substrate. The single turnover rate constant for excision of εA from the 25TεC substrate (50 nM) by AlkA is shown as a function of the concentration of AlkA. The data were fit with an inhibitory model[2] using the K_d for εA-DNA of 29 nM from Figure 2-1 (This gives an inhibitory K_i value of 2 μM). The average of duplicate experiments ± SD is shown. The data for the asymmetric 19TεC substrate are taken from Figure 2-3B for comparison.

A

```
25TAC:              5′-CGATAGCATCCTACCTTCTCTCCAT-3′
                    3′-GCTATCGTAGGATGGAAGAGAGGTA-5′

49TAC: 5′-GACATGATTGCCCGATAGCATCCTACCTTCTCTCCATGCGTCAATTGTC-3′
       3′-CTGTACTAACGGGCTATCGTAGGATGGAAGAGAGGTACGCAGTTAACAG-5′

25(alt):            5′-GATCATACTAATACTCGGCCCAAAG-3′
                    3′-CTAGTATGATTATGAGCCGCCTTTC-5′
```

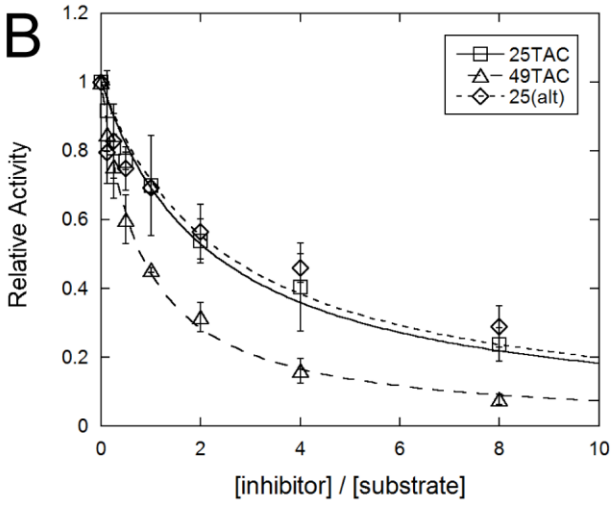


Figure A-4. AlkA affinity for undamaged DNA. (A) Additional undamaged DNA inhibitor oligonucleotides used in this figure. 25TAC is also shown in Figure 2-1. 49TAC contains the central 25TAC sequence and the alternative 25(alt) varies in sequence compared to 25TAC. (B) Multiple turnover competition between 19TεC substrate and 49TAC or 25(alt) inhibitor DNA with 25TAC replotted from Figure 2-4 for comparison. The average ± SD ($n \geq 2$). Using the K_m value of 15 nM for εA-DNA (Figure 2-2B), K_m/K_i values of 1.25 ± 0.20 and 0.40 ± 0.04 were observed for the 49TAC and 25(alt), respectively. A $K_{i,obs}$ value of 12 ± 2 nM and a final K_d of 1.0 ± 0.2 μM was calculated for the 49TAC inhibitor (84 binding sites; Eq 6). The 25(alt) inhibitor has a $K_{i,obs}$ of 38 ± 5 nM and a final K_d of 1.4 ± 0.2 μM (36 binding sites). All three undamaged DNA inhibitors give an average K_d of 1.2 μM for undamaged DNA as shown in Figure 2-8. These data suggest that multiple AlkA molecules bind undamaged DNA nonspecifically without strongly favoring end binding or specific DNA sequences.

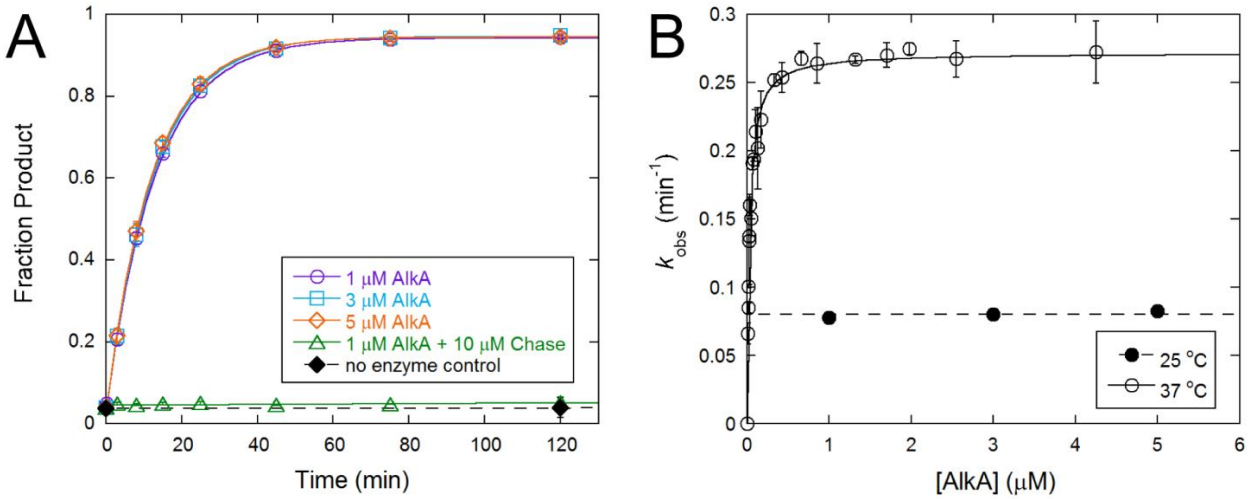


Figure A-5. Effect of temperature on the single turnover glycosylase reaction catalyzed by AlkA. (A) Time courses for single turnover excision of 19TεC at 25 °C with the indicated concentration of AlkA. Both 50 nM and 100 nM substrate were reacted in duplicate with saturating AlkA concentrations. Chase experiment was performed with pyr-DNA chase as described for Figure 2-6 in the main text. The average ± SD is shown (n ≥ 2). (B) The single-turnover rate constant for excision of εA from the 19TεC substrate at 25 °C. The observed rate constant was independent of the concentration of AlkA from 1–5 μM, giving a k_{max} value of 0.080 ± 0.003 min^{-1} (0.0013 ± 0.0001 s^{-1}). The average ± SD is shown (n = 4). The data at 37 °C are replotted from Figure 2-3B for comparison.

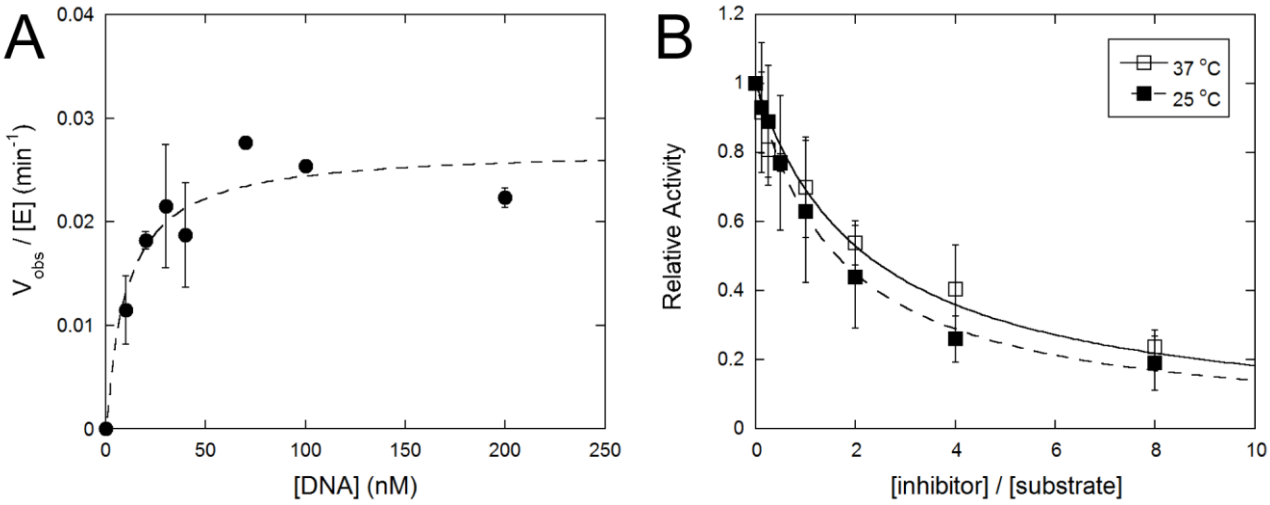


Figure A-6. Effect of temperature on the multiple turnover glycosylase reaction catalyzed by AlkA. (A) Multiple turnover excision of εA from 19TεC substrate (10-200 nM) by AlkA (1 nM) at 25 °C. The substrate concentration dependence follows Michaelis-Menten kinetics with a k_{cat} value of 0.027 ± 0.01 min^{-1} and a K_m value of 11 ± 1 nM. The average of duplicate experiments ± SD is shown. (B) Multiple turnover competition between the 19TεC substrate and 25TAC undamaged DNA inhibitor at 25 °C. The relative activity (V_{obs}/V_{max}) was fit by the model for competitive inhibition (Eq 5). The average of duplicate experiments ± SD is shown. The K_m/K_i value was 0.62, giving a K_i value for the undamaged inhibitor of 18 ± 14 nM (using the K_m value of 11 nM for εA-DNA). The data from 37 °C are replotted from Figure 2-4 for comparison. The K_m and K_i values calculated for AlkA binding to the different DNA sites is identical within error at 25 and 37 °C. Although the dissociation constant for abasic DNA was not obtained at 25 °C, the experiments with undamaged DNA suggest that AlkA binding affinity is not significantly affected by temperature.

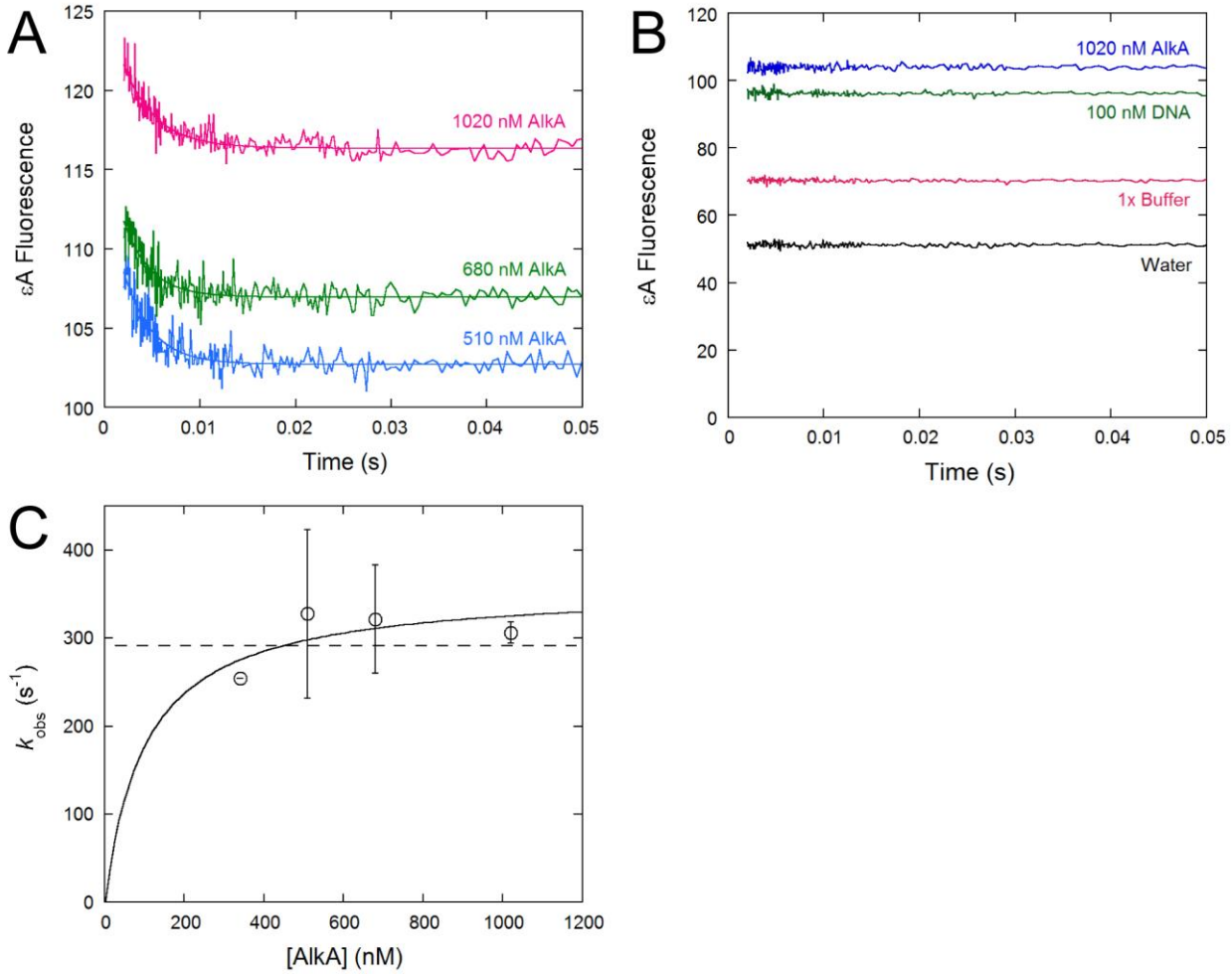


Figure A-7. Single mixing stopped-flow for association of AlkA with the 19TεC substrate. (A) Stopped-flow was performed with 100 nM 19TεC substrate and varying AlkA concentrations. The representative εA-DNA fluorescence is reported in arbitrary units and is the average of three shots. The single phase was fit by a single exponential. (B) Controls to evaluate the signal to noise. The fluorescence signal for 100 nM 19TεC substrate was set to 100, and the signal from water, buffer, or the highest [AlkA] alone are shown. (C) The k_{obs} values were plotted versus the AlkA concentration and fit by a hyperbolic function to demonstrate that the reaction is mostly saturated at the concentrations that were tested (solid line; k_{max} value of 360 ± 70 s^{-1}). The low signal to noise precluded testing lower concentrations of AlkA and therefore it is not possible to determine the $K_{1/2}$ value. Averaging all of the data gives a very similar value of 290 ± 60 s^{-1} (dotted line). This fit represents $k_{flip,obs}$ which is independent of AlkA concentration under saturating conditions. The average ± SD is shown ($n \geq 2$).

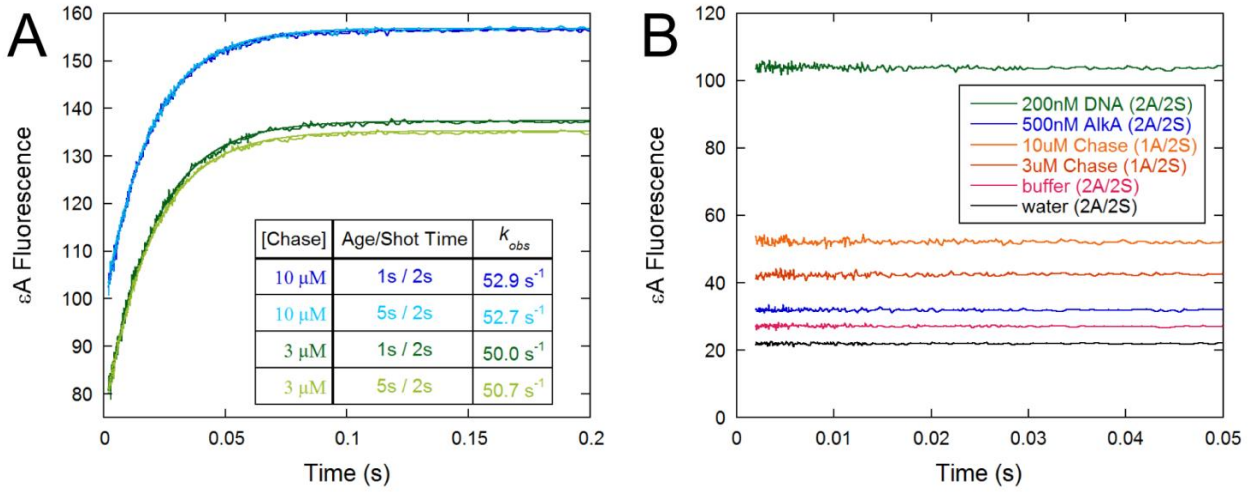


Figure A-8. Controls for double mixing experiments to measure substrate dissociation. (A) εA-DNA fluorescence from double mixing stopped-flow assay in which AlkA (425 nM) and 19AεA DNA (200 nM) was mixed, aged for 1 or 5 s, and mixed with 25-mer pyrrolidine chase (3 μM or 10 μM). The change in fluorescence (average of 3 shots) was fit by single exponential with an average k_{obs} value of 52 ± 2 s^{-1}. (B) Controls to evaluate the signal to noise. The fluorescence for 200 nM 19AεA substrate was set to 100, and the signal from water, buffer, chase DNA, or AlkA alone are shown.

Table A-1. Kinetic parameters for recognition and excision of εA by AlkA and AAG.

	AlkA[a]	**Y162A AAG**[b]	**WT AAG**[b]
k_{flip} (s^{-1})	180 ± 25	170 ± 16	3.6 ± 0.7
k_{unflip} (s^{-1})	62 ± 22	10 ± 1	(1.6 ± 0.3) x 10^{-3}
K_{flip}	2.9	17	2300
k_{chem} (s^{-1})	(1.4 ± 0.1) x 10^{-3}	(3.8 ± 0.1) x 10^{-4}	(8.0 ± 0.6) x 10^{-4}

[a]AlkA parameters are all obtained using the 19AεA substrate (19TεC for k_{chem}) at 25 °C . [b]AAG WT and Y162A parameters were obtained using a 25TεC and 25AεA substrates, respectively, at 25 °C.[3]

References

1. Admiraal, S. J., and O'Brien, P. J. (2013) DNA-N-glycosylases process novel O-glycosidic sites in DNA, *Biochemistry 52*, 4066-4074.
2. Zhao, B., and O'Brien, P. J. (2011) Kinetic mechanism for the excision of hypoxanthine by Escherichia coli AlkA and evidence for binding to DNA ends, *Biochemistry 50*, 4350-4359.
3. Hendershot, J. M., and O'Brien, P. J. (2014) Critical role of DNA intercalation in enzyme-catalyzed nucleotide flipping, *Nucleic Acids Res 42*, 12681-12690.

Chapter III

Effect of Salt on Base Excision by AlkA Glycosylase

Abstract

The bacterial AlkA glycosylase initiates base excision repair by hydrolyzing the bond between a damaged base and the deoxyribose backbone of DNA. AlkA exhibits activity towards a variety of substrates including the cyclic 1,N^6-ethenoadenine (εA) and deaminated hypoxanthine (Hx) adducts. While the mechanisms of AlkA-catalyzed excision of εA and Hx have been described, accurate comparisons cannot be made between the substrates. Here we have characterized the activity of AlkA on Hx and εA substrates under identical conditions and show that while AlkA excises Hx at a faster rate, it has a tighter affinity for εA, thus catalytic efficiencies for the two substrates are very similar. These reactions were performed with varying salt concentration showing that the binding of AlkA to substrates is salt dependent. As was previously observed, high concentrations of AlkA are auto-inhibitory, presumably due to interference from nonspecifically bound proteins. Surprisingly, the addition of a short C-terminal tag relieved this auto-inhibition. This work emphasizes the complexities of the AlkA mechanism and sets up future experimentation with other physiological substrates of AlkA.

Introduction

The *Escherichia coli* AlkA is a DNA glycosylase that initiates the base excision repair pathway by hydrolyzing the *N*-glycosidic bond between a damaged nucleotide and the sugar-phosphate backbone of DNA. As part of the bacterial adaptive response, the expression of AlkA is upregulated in response to DNA damaging agents.[1,2] The substrate specificity of AlkA encompasses a broad range including *N*-methylated purines (e.g., 3-methyladenine, 7-methylguanine) and *O*-methylated pyrimidines (2-methylthymine, 2-methylcytosine).[2-5] AlkA also shows significant activity on the deaminated adenine adduct, hypoxanthine (Hx), and the cyclic adduct, 1,N^6-ethenoadenine (εA), a byproduct of lipid peroxidation (Figure 3-1).[6,7] Remarkably, AlkA has similar rate enhancement for all of its substrates, suggesting that it has a promiscuous base lesion binding pocket.[8,9] The broad substrate range even includes undamaged bases, notably adenine, guanine and bases in mismatches.[9,10]

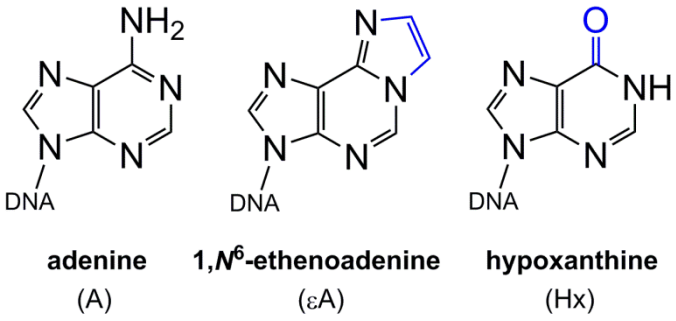
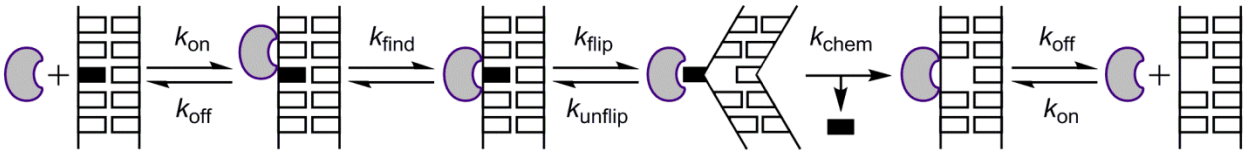


Figure 3-1. Structures of εA and Hx lesions in comparison to their source undamaged adenine.

A kinetic mechanism has been proposed for the AlkA-catalyzed excision of εA as documented in Chapter 2 and shown in Scheme 3-1.[12] AlkA binds nonspecifically to DNA prior to finding a damaged lesion. Nonspecific binding by other glycosylases is usually dependent on electrostatic interactions between the positively charged binding cleft of the enzymes and the negatively charged DNA.[13-18] The human AAG glycosylase, for example, uses these electrostatic

Scheme 3-1. Minimal kinetic mechanism for base excision by AlkA. (Reprinted from Taylor 2015)[11]



interactions to aid in facilitated diffusion during the lesion searching step.[19-21] In contrast, AlkA has a relatively neutral binding cleft. However, the structure of AlkA shows a putative sodium ion binding site between the helix-hairpin-helix domain of AlkA and the DNA.[8] Occupancy of this site could affect DNA binding, leading to a dependence on the monovalent salt concentration.

Following nonspecific binding, AlkA recognizes a damaged base. The recognition of a substrate includes flipping the nucleotide from the DNA duplex into the active site.[8] The equilibrium of flipping is negatively correlated to the stability of the base pair and presumptively positively correlated to the stability of the interactions between the flipped base and enzyme active site and the intercalating residue and duplex DNA. The flipping equilibrium for εA in duplex DNA, which is not capable of Watson-Crick base pairing, is only marginally stable (K_{flip} = 3).[12] For comparison, the human AAG glycosylase stably binds the flipped out εA lesion with a K_{flip} value of 2300. In addition, for excision of Hx by AlkA, significant inhibition is observed when AlkA is at high concentrations. These results are explained by the model that AlkA prefers to bind to undamaged sites, blocking access to the Hx lesion.[11] This model is expounded further to explain why the multiple turnover rate of Hx excision by AlkA is slower than its single turnover rate. These observations demonstrate that AlkA is not efficient at recognition of Hx and εA.

Based on the reported Michaelis constants, previous work suggests a modest preference of AlkA for εA over Hx lesions.[11,12] However, the conditions used in these studies are not

identical and thus the rates cannot be directly compared. The discrimination between two substrates is given by the relative k_{cat}/K_m values, which describes all of the steps up to and including irreversible bond cleavage. By placing limited amounts of AlkA in competition for Hx and εA excision, we can evaluate the ratio of k_{cat}/K_m values for both substrates, thereby distinguishing the preferred substrate. These competition methods can easily be expanded to include many other substrates in the study of AlkA specificity.

The contribution of active site residues to catalysis has not been probed beyond the active site catalytic base (D238). To potentially facilitate mutant studies, we purified and characterized a tagged form of AlkA. We expressed AlkA with a C-terminal 6xHis tag containing a TEV protease cleavage site. After cleavage by TEV, an eight amino acid remnant remained on the C-terminus of AlkA. The crystal structure of AlkA bound to DNA shows the C-terminus as a flexible loop region. While not near the active site of the enzyme, the flexibility of the region and the expanded amino acid sequence have the possibility of affecting the activity of the enzyme.

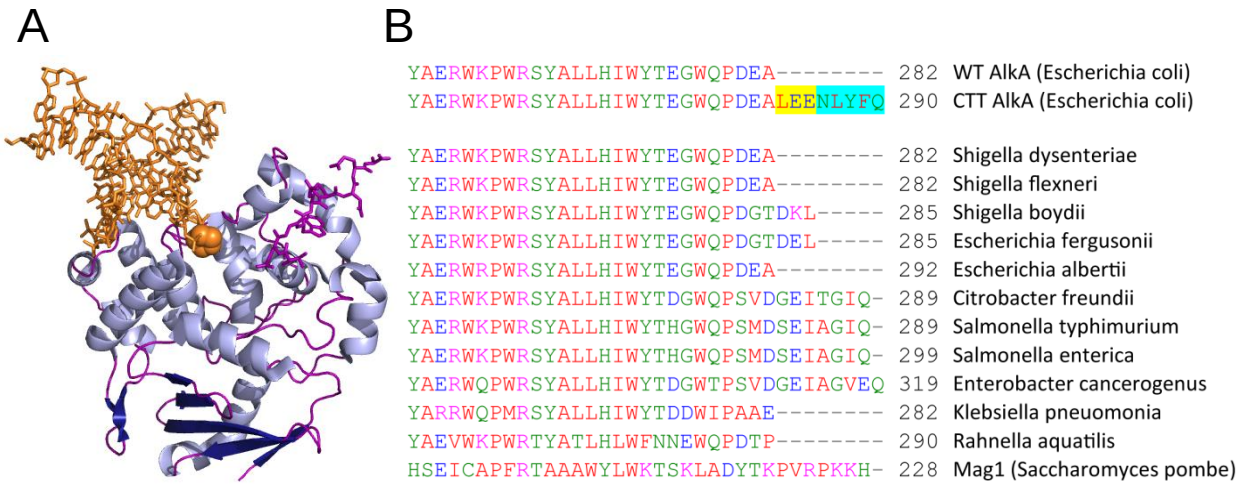


Figure 3-2. AlkA C-terminus. (A) Crystal structure of WT AlkA (blues) binding specifically to a 1-azaribose lesion (orange, spheres) (PDB 1DIZ).[10] The C-terminus forms an extended loop and is shown as purple sticks. (B) ClustalW2 alignment of WT *E. coli* AlkA with various bacterial homologs and the eukaryotic Mag1. For comparison, CTT AlkA is shown with its eight amino acid remnant of a C-terminal His6 affinity tag. Highlighted in blue and yellow are the upstream TEV protease cleavage site and flexibility linker, respectively.

Comparison of the kinetic properties of WT and tagged (CTT) AlkA revealed a significant difference between the two proteins under single turnover conditions. Interestingly, there is a loose conservation of the C-terminal sequences observed across some bacterial species (Figure 3-2). Our mechanistic studies give evidence that the C-terminal region may affect the activity of these enzymes, and caution should be taken when modifying wild type enzymes.

This study focuses on the excision of εA and Hx by the wild type and modified *E. coli* AlkA proteins. We performed experiments under a variety of conditions to more thoroughly understand the differences in substrates and proteins. We especially concentrated on the effect of salt on AlkA binding and base excision. This report opens the way for future AlkA experiments with additional substrates and ion identities.

Materials and Methods

Preparation of Proteins

Full length wild type *E. coli* AlkA was expressed and purified as previously described in Chapter 2 of this thesis.[12] The CTT AlkA pET24 plasmid (with C-terminal 6xHis tag) was expressed in BL21 (DE3) *E. coli* cells. Cell cultures were grown in LB to mid-log phase and induced with IPTG. Cells were homogenized in lysis buffer (100 mM Tris-Cl pH 7.4, 300 mM NaCl, 5 mM β-mercaptoethanol (BME), 1 mM phenylmethylsulfonyl (PMSF) and 5% glycerol (v/v)) and centrifuged. The supernatant was mixed with 0.2% polyethyleneimine (PEI) at 4 °C and centrifuged again. The supernatant was loaded on a HisTrap HP 5 mL NTA-Ni^{2+} sepharose column (GE Healthcare) equilibrated with 100 mM Tris-Cl pH 7.4 and 5% glycerol (v/v) buffer. An imidazole gradient (15 mM to 500 mM) was run over the column and the 6xHis-AlkA eluted around 210 mM imidazole. 5 mM BME and 1 mM EDTA was added along with 1 mg TEV protease for every 100 mg of protein and the reaction was incubated at 16 °C overnight. The cleaved protein was loaded on a HiTrap SP HP 5 mL sepharose column (GE Healthcare) equilibrated with a 50 mM Tris-Cl pH 7.4, 25 mM NaCl and 5% glycerol (v/v) buffer. A NaCl gradient (25 mM to 500 mM) was applied to the column and the CTT AlkA eluted around 200 mM NaCl. The protein was dialyzed in storage buffer (25 mM HEPES pH 7.5, 100 mM NaCl, 0.1 mM EDTA, 5 mM BME, and 10% glycerol (v/v)) overnight. The protein was concentrated with an Amicon Ultra concentrator (15 kDa cutoff) to the desired concentration of around 10 mg/mL. Aliquots were snap frozen in liquid nitrogen and stored at -80 °C for future use. Prior to use aliquots were thawed and diluted 1:1 in 100% glycerol (v/v) and stored at -20 °C. WT AlkA

was also stored in 50% glycerol for these assays in contrast to the zero glycerol buffer described in Chapter 2.

Preparation of Oligonucleotides

For these assays, symmetric 25-mer (5′-CGATAGCATCCTXCCTTCTCTCCAT-3′) and asymmetric 19-mer (5′-CATCCTXCCTTCTCTCCAT-3′) 5′ 6-fluorescein (FAM) labeled Hx or εA oligonucleotides were used. A 5′ 6-hexachlorofluorescein (HEX)-labeled 25-mer pyrrolidine (pyr) inhibitor was also used in the determination of active AlkA concentration assays. Desalted oligonucleotides were purchased from Integrated DNA Technologies (IDT) or the Keck Center at Yale University. The DNA was purified with denaturing PAGE, extracted, and desalted with reverse phase C18 columns (Sep-Pak, Waters). The concentrations of the forward and complement strands were determined with A_{260} absorbance. The extinction coefficients of εA and pyr substrates were estimated by subtracting 9400 M^{-1} cm^{-1} from an undamaged A substrate or using a dSpacer, respectively. The DNA substrates were annealed by heating the reaction to 95 °C for 3 min and cooling to 4 °C at a rate of 0.2 °C/s. The complementary strands were in a 1.2-fold excess of the lesion containing strands.

General Glycosylase Activity Assay

AlkA and DNA were mixed in reaction buffer (50 mM NaMES pH 6.1, 1 mM EDTA, 1 mM DTT, 0.1 mg/mL BSA, 10% glycerol (v/v) and appropriate concentration of NaCl). Ionic strength was corrected to 50 mM, 100 mM, or 150 mM with NaCl (26 mM, 76 mM, and 126 mM Cl^-, respectively). Reaction mixtures were incubated at 37 °C and 3 μL aliquots were quenched in 0.2 M NaOH at various time points. The samples were heated at 70 °C for 15 min,

centrifuged, and neutralized with loading buffer (98% formamide, 1 mM EDTA). 5 to 200 fmol of DNA were loaded into each lane of denaturing PAGE (20% (w/v) acrylamide, 1× TBE, 6.6 M urea). Gels were analyzed using a Typhoon Trio Fluorescence imager (GE Healthcare) with 488 nm excitation and a 520 nm band-pass filter to detect fluorescein. The fraction product was calculated by dividing the intensity of the product band by the sum of product and substrate band intensities.

Determination of Active AlkA Concentration

The active concentration of both WT and CTT AlkA was determined with an active site titration using a tight binding inhibitor pyrrolidine-containing (pyr) DNA. Varying concentrations (0-800 nM) of 25-mer 5′HEX-pyr DNA was mixed with 1 μM 19-mer 5′FAM-Hx DNA and 100 or 200 nM AlkA under the reaction conditions described above (100 mM ionic strength condition). The fraction product versus time was fit by linear regression up to 8% reaction. The reaction rates were normalized to the reaction without inhibitor and fit with a quadratic binding equation (Eq 1), where K_d is the dissociation constant for pyr-DNA binding, and E and I represent AlkA and the pyr inhibitor, respectively. The WT AlkA was found to be 63% active and the CTT AlkA was 50% active compared to the A_{280} absorbance measurements (Figure 3-3). All subsequent AlkA concentrations were corrected to the amount of active enzyme.

$$\frac{V_{obs}}{V_{max}} = 1 - \frac{K_d + [E] + [I] - \sqrt{(K_d + [E] + [I])^2 - 4[E][I]}}{2[E]} \quad (1)$$

Single Turnover Glycosylase Assays

Single turnover reactions were performed by mixing 50 nM DNA and increasing

concentrations of AlkA (at least a 2-fold excess). Time dependence data were fit to a single exponential and the AlkA concentration dependence was fit to a hyperbolic dependence if no inhibition was present (Eq 2 without (1+[E]/K_i); Eq 4 in Chapter 2). In the case of inhibition at high enzyme concentrations, a simple noncompetitive inhibition model was used in which k_{max} is the maximal single turnover rate of reaction, E is the concentration of AlkA, and K_d and K_i are the dissociation constants for the stimulatory and inhibitory phases, respectively (Eq 2).[11]

$$k_{obs} = \frac{k_{\max} * [\mathrm{E}]}{K_\mathrm{d} + [\mathrm{E}] * \left(1 + \frac{[\mathrm{E}]}{K_\mathrm{i}}\right)} \tag{2}$$

Multiple Turnover Glycosylase Assays

Multiple turnover reactions were performed by mixing 10 nM AlkA with increasing concentrations of DNA. The fraction product versus time was fit by linear regression up to 8% reaction, and the DNA concentration dependence was fit with the Michaelis-Menten equation where k_{cat} is the maximal turnover rate constant, S is the substrate, E is AlkA, and the K_m is the concentration at which the observed rate constant is half the maximal rate constant (Eq 3). The lowest concentration of DNA tested was 50 nM, in which case only 0.5 turnovers were measured in the initial rate period. However, due to the lack of a burst phase, the measurement was considered accurate.

$$\frac{\mathrm{V_{obs}}}{[\mathrm{E}]} = \frac{k_\mathrm{cat}[\mathrm{S}]}{K_\mathrm{m} + [\mathrm{S}]} \tag{3}$$

Competition Assays to Measure Relative k_{cat}/K_m Values

Direct competition reactions were performed by mixing 100 nM each of 25-mer and 19-mer DNA (200 nM total DNA) and 20 nM AlkA. Time dependence plots were fit with linear regression up to 8% reaction. The ratio of initial velocities (V_A/V_B) is proportional to the ratio of

k_{cat}/K_m values for each substrate when equal concentration of substrates is used (Eq 4).[22]

$$\frac{V_A}{V_B} = \frac{[A] * \left(\frac{k_{cat}}{K_m}\right)_A}{[B] * \left(\frac{k_{cat}}{K_m}\right)_B} \quad (4)$$

Results and Discussion

Purification of AlkA with a C-terminal Extension

To aid in purification of AlkA, it was expressed with a hexahistidine (His) tag on its C-terminal end. After purification with a nickel affinity column, the protein was treated with TEV protease to remove the His tag from the C-terminus. A non-native remnant of 8 amino acids remained after cleavage (Figure 3-2). It was expected that this remnant would not affect glycosylase activity, because the C-terminus is not near the DNA binding site.[8] Preliminary experiments showed that the C-terminally tagged (CTT) AlkA exhibited glycosylase activity. However, a marked difference between the two proteins was apparent. An alignment of bacterial homologs reveals some patterns of conservation at the C-terminus (Figure 3-2). This raised the possibility that C-terminal modification could impact the function of AlkA. We therefore performed a variety of in vitro kinetics assays with both WT and the C-terminally tagged (CTT) AlkA proteins on εA and Hx substrates to thoroughly compare the proteins. We also used different ionic strength conditions to test the effect of salt on AlkA binding. For all assays, the

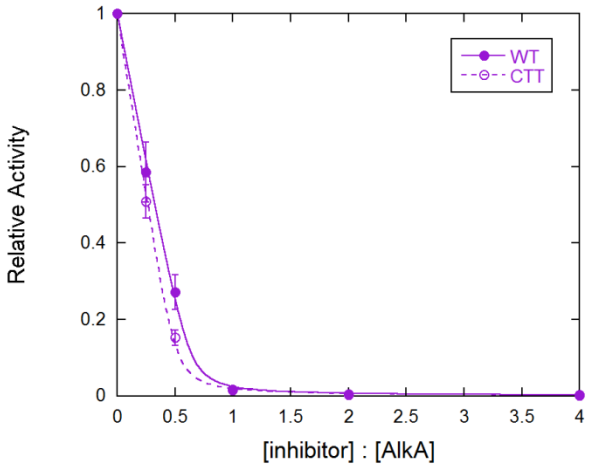


Figure 3-3. Determination of the concentrations of WT and CTT AlkA. The average of 100 and 200 nM AlkA reactions with 1 μM 5′FAM 19-mer Hx DNA and varying 5′HEX 25-mer pyr inhibitor concentrations. The relative activity of reactions versus the ratio of inhibitor to AlkA concentration was fit with a quadratic binding equation (Eq 1). WT AlkA was found to be 63% active and CTT AlkA was 50% active. The average ± SD is shown ($n \geq 4$).

active concentration of WT and CTT proteins was first determined by titrating the proteins with a tight binding transition state mimic (Figure 3-3).

Ionic Strength Dependence of Glycosylase Activity

To understand the effects of salt on AlkA-catalyzed bond hydrolysis, WT or CTT AlkA was incubated with hypoxanthine DNA and various ionic strengths under single turnover conditions. As previously reported, glycosylase activity of WT AlkA is inhibited at high enzyme concentrations due to interference between multiple bound enzymes, and this inhibition is dependent on the position of the lesion (Figure 3-4A).[11] Interestingly, the addition of the C-terminal tag relieved much of this inhibition, allowing the CTT AlkA to display a faster k_{max} value for bond hydrolysis at a saturating concentration of enzyme (Figure 3-4B).

Ionic strength also affected the single turnover activity of AlkA on Hx-containing DNA. An increase in ionic strength (controlled by NaCl) caused an increase in the K_d values for both WT and CTT AlkA on Hx (Figure 3-4; Table 4-1). Additionally, the inhibition observed at high concentrations of WT and CTT AlkA was absent at the ionic strength of 150 mM. These results are indicative of the influence ionic strength has on nonspecific binding by AlkA. At higher salt concentrations, the electrostatic interactions between AlkA and DNA are decreased and thus the nonspecific binding by AlkA is weaker. At high AlkA concentrations, this weaker nonspecific binding does not interfere with the specific binding of Hx by AlkA and the inhibition is alleviated. The effects of salt are also seemingly independent of cation identity, because the enzyme dependence was the same when sodium was replaced with potassium (Figure 3-4A). It is worthwhile to note the complicated nature of these results as the WT AlkA k_{max} value is decreased for the 150 mM ionic strength condition while the CTT AlkA k_{max} value is unchanged

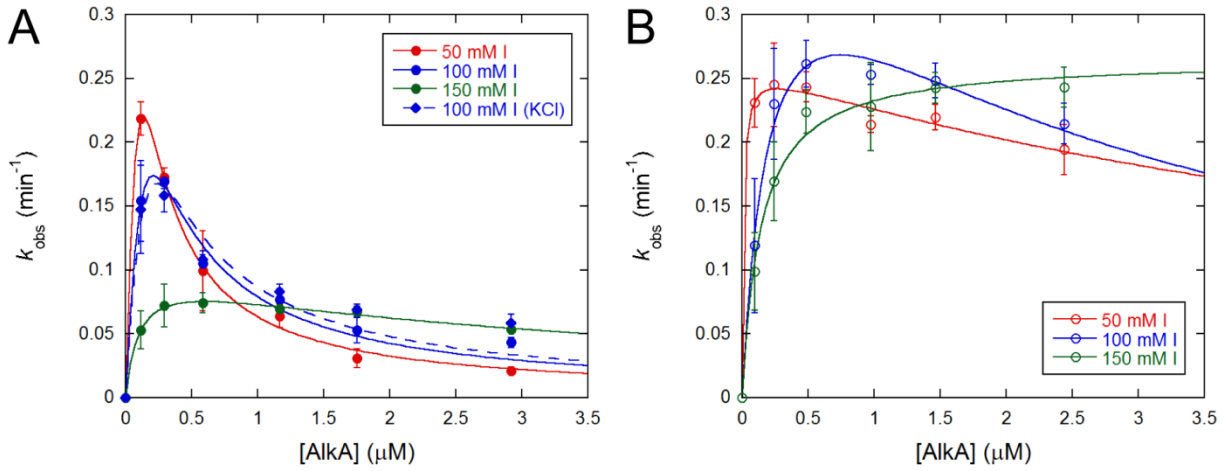


Figure 3-4. Ionic strength dependence of single turnover excision of hypoxanthine. Single turnover activity of WT (A) and CTT (B) AlkA on 5′FAM 25-mer Hx-DNA substrate under varying ionic strength conditions. All data were fit using the inhibitory site model (Eq 2), excepting the CTT AlkA 150 mM ionic strength data which was fit for a hyperbolic dependence. The average ± SD is shown (n ≥ 4; KCl n = 2). Kinetic parameters are shown in Table 3-1.

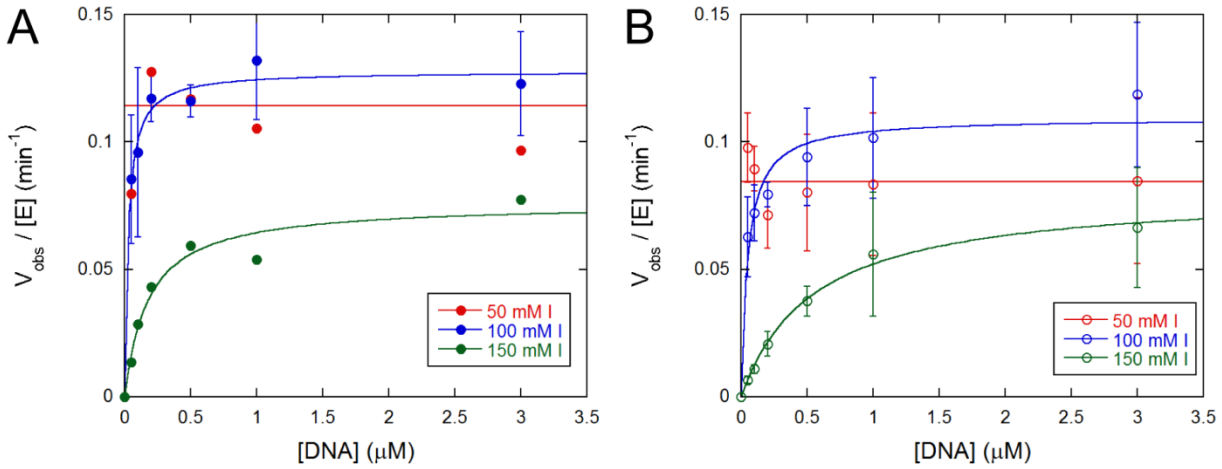


Figure 3-5. Ionic strength dependence of multiple turnover excision of hypoxanthine. Multiple turnover activity of WT (A) and CTT (B) AlkA on 5′FAM 25-mer Hx substrate under varying ionic strength conditions. 50 mM ionic strength data were fit with a straight line indicating the average of all data points. 100 mM and 150 mM ionic strength data were fit with the Michaelis-Menten equation (Eq 3). The average ± SD is shown for all CTT AlkA data and WT AlkA 100 mM I data (n ≥ 3). WT AlkA 50 mM and 150 mM I data was performed only once. Kinetic parameters are shown in Table 3-1.

compared to the other salt conditions. Our hypothesis is that there is a protein specific salt effect (cation or anion) that is specifically slowing the WT AlkA protein rate of excision.

We next explored the effects of salt on the multiple turnover kinetics for Hx excision. No burst phase has ever been observed for AlkA-catalyzed release of Hx or εA,[11,12] suggesting that multiple turnover also reports on *N*-glycosidic bond hydrolysis. First examining WT AlkA, we observe slower k_{cat} values compared to single turnover k_{max} values for the 50 mM and 100 mM ionic strength conditions (Figure 3-5A). We attribute these differences to competition for AlkA binding between undamaged sites and the Hx site. The 150 mM ionic strength condition exhibits identical maximal rates for both single and multiple turnover experiments with WT AlkA, indicating the weakening of nonspecific binding by AlkA at high salt and the preference of AlkA for the Hx lesion. However, as seen with the WT AlkA under single turnover conditions, this model does not explain why we see a faster k_{cat} value at lower salt conditions. This discrepancy again seems to be specific to WT AlkA and may be some effect of the anion identity. Further experimentation will be needed to resolve this issue.

Curiously, the CTT AlkA behaves very similarly to WT AlkA under multiple turnover conditions. The k_{cat} values for the two proteins are identical within error for all salt conditions although the K_m values for CTT AlkA-catalyzed excision of Hx appear to be weaker compared to WT AlkA•Hx binding (Figure 3-5B; Table 3-1). This weaker lesion binding by CTT AlkA could explain why the k_{cat} values are further removed from the related single turnover k_{max} values. The difference between k_{max} and k_{cat} values is especially apparent at the 150 mM ionic strength condition. Unlike WT AlkA in which the maximal rates of excision are identical whether enzyme or substrate is in excess, the CTT AlkA has a much reduced maximal rate of base excision for the highest salt concentration. One explanation is that nonspecific sites and the

lesion site could be in competition for CTT AlkA binding more than the WT AlkA. However, this model overlooks the single turnover inhibition relief of CTT AlkA at the high salt concentration. These data demonstrate the slight nuances of complicated models and further work would help ease confusion. As previously discussed, additional cations and anions could be tested for their effect on AlkA activity and ionic strength dependencies.

Comparing Activity of AlkA on εA and Hx Lesions

We next measured single and multiple turnover kinetics of εA excision to compare to Hx excision in the same conditions. For both AlkA proteins, εA was excised at a slower k_{max} than Hx (and slower than the value previously documented in the absence of glycerol) (Figure 3-6A).[12] There was also less inhibition observed for εA excision, mostly likely explained by the tighter K_d values observed with εA, thus reducing the influence of other AlkA binding events. We also tested the effect of the fluorophore label by moving the label or changing its identity. This was done because of the model that the position of the lesion from the 5′-end of the DNA affects the inhibition of AlkA on Hx lesions.[11] Neither alteration affected the single turnover activity of AlkA on εA.

We also compared the excision of εA and Hx by WT and CTT AlkA under multiple turnover conditions and again saw that εA was excised more slowly than Hx and that WT and CTT AlkA have similar k_{cat} values for εA excision (Figure 3-6B). The WT AlkA activity clearly shows εA binding more tightly than Hx under these conditions although the CTT AlkA experiments are a bit more ambiguous due to the limited concentrations tested.

To more fully compare these two substrates we directly mixed εA and Hx DNA in competition with limited amounts of AlkA. 25-mer εA and 19-mer Hx substrates were in equal

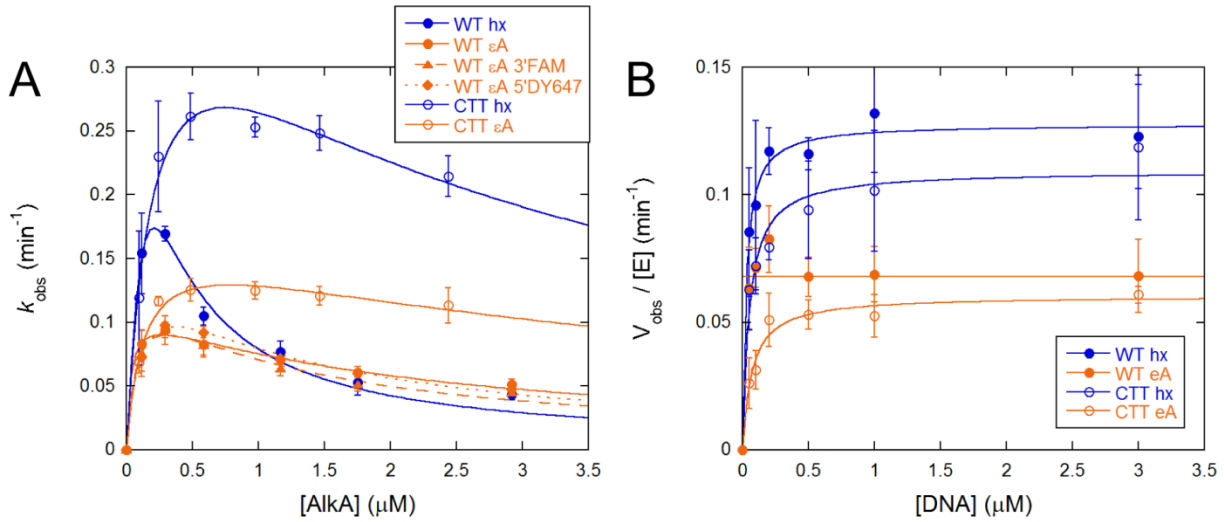


Figure 3-6. Glycosylase activity of AlkA on εA-DNA and comparison to Hx-DNA. Single (A) and multiple (B) turnover AlkA-catalyzed activity on 25-mer Hx and εA substrates under 100 mM ionic strength conditions. All substrates have a 5′FAM unless otherwise noted. All single turnover data were fit by the inhibitory model as described in materials and methods. Multiple turnover data was fit with the Michaelis-Menten equation excluding the WT AlkA εA data that was fit with a straight line defining the average V_{obs}/[E] values. All Hx data is repeated from Figures 3-4 and 3-5 for comparison. The average ± SD is shown (n ≥ 2). Kinetic parameters are shown in Table 3-1.

Table 3-1. Kinetic parameters for Hx and εA excision by WT and CTT AlkA.

	Hx						εA	
	50 mM I		100 mM I		150 mM I		100 mM I	
	WT	CTT	WT	CTT	WT	CTT	WT	CTT
k_{max} (min^{-1})[a]	ND	0.26 ± 0.02	*0.28 ± 0.01*[c]	0.42 ± 13	0.10 ± 0.02	0.26 ± 0.01	0.11 ± 0.03	0.17 ± 0.3
K_d (nM)[a]	ND	10 ± 7	*42*[c]	208 ± 186	99 ± 150	139 ± 59	33 ± 59	128 ± 54
K_i (nM)[a]	34 ± 110	2700 ± 4500	*220 ± 30*[c]	2700 ± 2000	3600 ± 1300	NA	2200 ± 1300	4800 ± 1900
k_{cat} (min^{-1})[b]	0.11	0.08 ± 0.02	0.13 ± 0.02	0.11 ± 0.03	0.076	0.08 ± 0.03	0.07 ± 0.01	0.06 ± 0.01
K_m (nM)[b]	≤5	≤5	27 ± 52	49 ± 33	177	547 ± 167	≤5	69 ± 11

[a]Single turnover kinetic parameters
[b]Multiple turnover kinetic parameters
[c]Kinetic parameters reported in Zhao 2010[11]

excess to AlkA, thus any difference in initial rates of reaction is directly proportional to the ratios of k_{cat}/K_m values for each substrate (Figures 3-7A and 3-7C). To control for length dependence, competition of 25-mer and 19-mer substrates both containing Hx lesions was also performed (Figures 3-7B and 3-7D). For a simple example, at 150 mM ionic strength the WT AlkA favored the 25-mer εA substrate 1.55-fold over the 19-mer Hx substrate. (Figure 3-7A; green traces).

Under the same condition the two Hx substrates showed very similar initial rates of base excision, with the 25-mer being favored 1.09-fold over the 19-mer DNA (Figure 3-7B; green traces). Therefore the overall preference for εA over Hx substrates by WT AlkA is 1.42-fold at 150 mM ionic strength (Table 3-2). An identical preference was also observed with the CTT AlkA (Figures 3-7C and 3-7D; green traces; Table 3-2). These data indicate that AlkA exhibits very similar catalytic efficiencies for the excision of Hx and εA with only a slight preference for εA sites.

The direct competition between εA and Hx was also performed at 50 mM (red traces) and 100 mM (blue traces) ionic strength (Figure 3-7). The results for the 100 mM ionic strength show a very slight preference for the εA substrate by both AlkA proteins (Table 3-2) with the 100 mM ionic strength condition having a similar explanation to the higher salt condition. At the 50 mM ionic strength condition, εA is not preferred to Hx and their k_{cat}/K_m ratios are equal to (CTT AlkA) or close to 1 (WT AlkA). These results can be explained by the model that AlkA binds nonspecific DNA more tightly at low salt. Since the binding of AlkA at the lesion sites is in competition with the other undamaged sites at 50 mM ionic strength (Figures 3-5) the lesions are undistinguishable from undamaged DNA and thus the slight preference of εA to Hx is lost.

These results, in combination with the higher salt concentrations and previously described single and multiple turnover kinetics assays show a clear dependence of AlkA binding on salt concentration. It is now unclear whether this dependence is due to the ionic strength and thus electrostatic interactions between the protein and DNA, or the identity of the salt ions.

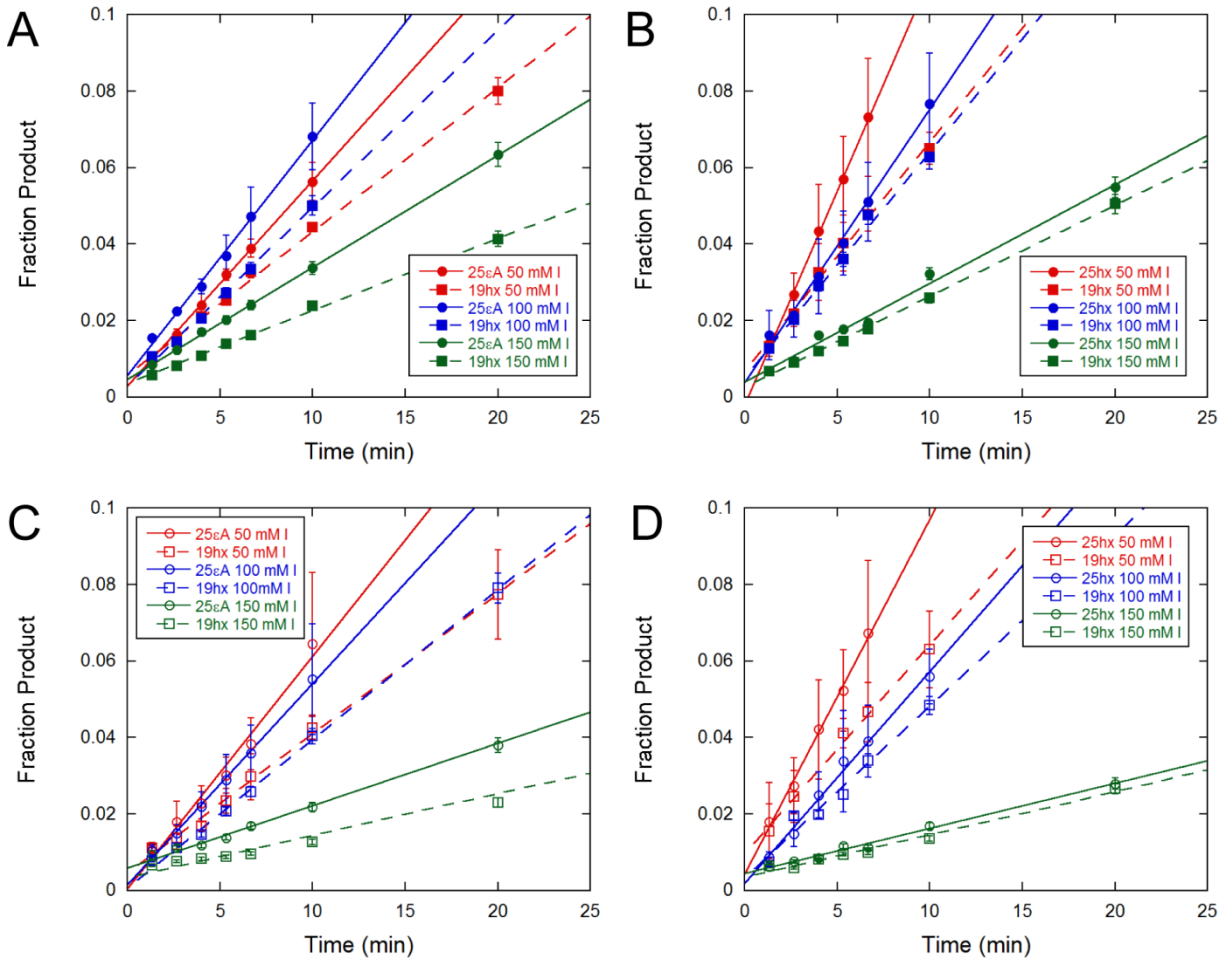


Figure 3-7. Direct competition between 25-mer and 19-mer substrates with εA and Hx lesions. Initial rates of reaction for WT AlkA-catalyzed excision of two competing substrates. Competition was between 5′FAM 25-mer εA (A) or Hx (B) and 19-mer Hx substrates under various ionic strength conditions. The first 8% of reactions were fit by a straight line for all data. CTT AlkA reactions on the same substrates are shown in C and D. The average ± SD is shown (n = 2). The ratios of initial rates are shown in Table 3-2.

Table 3-2. Ratios of k_{cat}/K_m values[a] for competition experiments.

Ionic Strength	*AlkA*	*25hx / 19hx*	*25 ε A / 19hx*	***ε A / hx***
50 mM I	WT	1.87 ± 0.39	1.43 ± 0.15	**0.76 ± 0.18**
	CTT	1.70 ± 0.15	1.65 ± 0.02	**0.97 ± 0.07**
100 mM I	WT	1.21 ± 0.02	1.32 ± 0.07	**1.09 ± 0.06**
	CTT	1.23 ± 0.11	1.34 ± 0.11	**1.09 ± 0.13**
150 mM I	WT	1.09 ± 0.12	1.55 ± 0.05	**1.42 ± 0.16**
	CTT	1.05 ± 0.01	1.49 ± 0.03	**1.42 ± 0.03**

[a]Ratios of initial rates are proportional to the ratios of k_{cat}/K_m values (Eq 4).

Dependence of AlkA Activity on pH

Single turnover CTT AlkA activity on Hx was also investigated with varying pH (Figure 3-8). We observe faster k_{max} values with decreasing pH value, and pH 5.2 showcasing the fastest rate of reaction. These preliminary results suggest that AlkA prefers a different pH to that optimal for εA excision and other substrates. Previous reported data has shown optimal activity of WT AlkA on εA:T and G:T mismatches at a pH value of 6.1. Higher pH values above 7 are optimal for 7meG excision by WT AlkA.[9] What all these results indicate is that AlkA has poorer activity toward neutral bases at physiological pH values. The protonated purines (3meA and 7meG) have positive charge at physiological pH values and this stabilizes the accumulation of negative charge on the nucleobase in the transition state (Figure 1-2). AlkA may require an active site general acid to stabilize the transition state with neutral lesions, thus lower pH values are optimal. It is interesting that the human AAG glycosylase exhibits the same optimal pH values as AlkA for Hx, εA and 7meG and uses a general acid-base mechanism to excise neutral lesions.[23] The inefficient general acid catalysis at physiological pH is expected to impair the ability of AlkA to excise neutral lesions such as εA and Hx in vivo. Therefore the transition state stabilization may act as another selectivity filter for physiological AlkA activity. Reports suggest that *E. coli* relies on AlkB and endonuclease V to repair εA and Hx, as opposed to AlkA.[24,25]

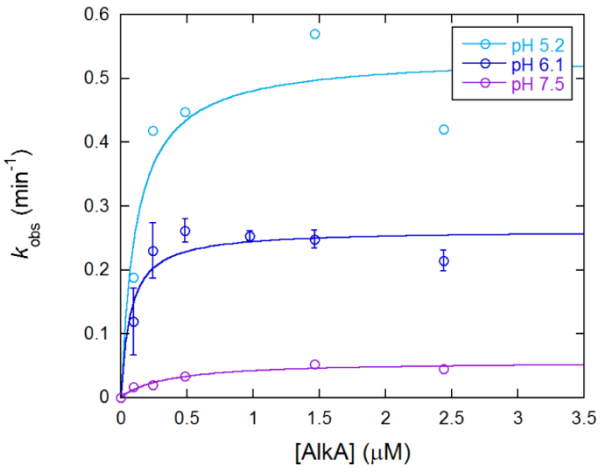


Figure 3-8. pH dependence of single turnover CTT AlkA activity on Hx substrates. Hyperbolic dependence of Hx excision from 5′FAM 25-mer by CTT AlkA under varying pH conditions. The pH 6.1 data were repeated from Figure 3-4 but fit with a hyperbolic curve to show comparison with other pH conditions. The pH 5.2 and 7.5 were performed once and the pH 6.1 data averages are shown ± SD (n = 9).

Conclusions

We have directly compared the activity of two forms of the *E. coli* AlkA enzyme on εA and Hx substrates under varying salt conditions. Our results show that the 8-amino acid remnant on the C-terminus of AlkA does not affect its activity during multiple turnover conditions, but there is a significant distinction between the WT and CTT proteins when the enzymes are in excess. Specifically, the 8-amino acid tag relieved auto-inhibition that is observed when WT AlkA acts on 25-mer Hx and εA substrates. The observation sheds some light on the auto-inhibition at high AlkA concentrations, and raises the possibility that this biochemical property serves some function in the subset of AlkA proteins with conserved C-termini. Affinity tags are commonplace and can greatly aid in protein purification. However, the surprising result that the CTT AlkA has improved activity compared to WT demonstrates the importance of comparing tagged and untagged forms of proteins. Due to its catalytic promiscuity, AlkA could be a useful tool for removing a wide variety of damaged bases from DNA in analyses such as the comet assay or for the quantification of damaged bases. The recombinant CTT form of AlkA may be superior for such applications as it can be used in excess without worry of auto-inhibition.

We have also shown that AlkA acts with similar catalytic efficiency on εA and Hx lesions. While AlkA binds εA slightly more tightly than Hx, the rate of Hx excision is slightly faster. The combination of these two kinetic parameters shows only a slight 1.4-fold preference of AlkA for εA over Hx lesions at the highest ionic strength condition. These results demonstrate that AlkA has a permissive active site that can accommodate a broad range of substrates. The promiscuity of AlkA is a good strategy for removing lesions of various shapes or charge, and

predicts a base line level of protection against new forms of DNA damage that have not been previously encountered.

Finally, the initial look at salt dependence of AlkA shows that salt has significant effects on the catalytic activity of AlkA. At higher salt concentrations, the affinity of AlkA for DNA appears to be weakened, similar to other glycosylases such as AAG. The structure of AlkA suggests there are few electrostatic interactions with the DNA, but a sodium ion may interface between the helix-hairpin-helix domain and the DNA backbone.[8] Our results suggest an ionic strength dependence, a salt ion effect, or both dependencies occurring when AlkA binds DNA.

It will be important to extend these biochemical studies to AlkA-catalyzed excision of *N*-methylated purines as these positively charged lesions are considered the physiological substrates of AlkA. For example, 7meG, while having a relatively labile *N*-glycosidic bond in comparison to εA and Hx, can be produced and used in such studies.[26] In addition, 3-methyl-3-deaza adenine (3d3mA) is a stable analog of 3meA that could potentially be protonated at low pH to mimic the 3meA positive charge.[27] This work lays the foundation for studying the substrate specificity of AlkA.

References

1. Evensen, G., and Seeberg, E. (1982) Adaptation to alkylation resistance involves the induction of a DNA glycosylase, *Nature 296*, 773-775.
2. Thomas, L., Yang, C. H., and Goldthwait, D. A. (1982) Two DNA glycosylases in Escherichia coli which release primarily 3-methyladenine, *Biochemistry 21*, 1162-1169.
3. Bjelland, S., Bjoras, M., and Seeberg, E. (1993) Excision of 3-methylguanine from alkylated DNA by 3-methyladenine DNA glycosylase I of Escherichia coli, *Nucleic Acids Res 21*, 2045-2049.
4. Bjelland, S., Birkeland, N. K., Benneche, T., Volden, G., and Seeberg, E. (1994) DNA glycosylase activities for thymine residues oxidized in the methyl group are functions of the AlkA enzyme in Escherichia coli, *The Journal of biological chemistry 269*, 30489-30495.
5. McCarthy, T. V., Karran, P., and Lindahl, T. (1984) Inducible repair of O-alkylated DNA pyrimidines in Escherichia coli, *The EMBO journal 3*, 545-550.
6. Saparbaev, M., and Laval, J. (1994) Excision of hypoxanthine from DNA containing dIMP residues by the Escherichia coli, yeast, rat, and human alkylpurine DNA glycosylases, *Proceedings of the National Academy of Sciences of the United States of America 91*, 5873-5877.
7. Saparbaev, M., Kleibl, K., and Laval, J. (1995) Escherichia coli, Saccharomyces cerevisiae, rat and human 3-methyladenine DNA glycosylases repair 1,N6-ethenoadenine when present in DNA, *Nucleic Acids Res 23*, 3750-3755.
8. Hollis, T., Ichikawa, Y., and Ellenberger, T. (2000) DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, Escherichia coli AlkA, *The EMBO journal 19*, 758-766.
9. O'Brien, P. J., and Ellenberger, T. (2004) The Escherichia coli 3-methyladenine DNA glycosylase AlkA has a remarkably versatile active site, *The Journal of biological chemistry 279*, 26876-26884.
10. Berdal, K. G., Johansen, R. F., and Seeberg, E. (1998) Release of normal bases from intact DNA by a native DNA repair enzyme, *The EMBO journal 17*, 363-367.
11. Zhao, B., and O'Brien, P. J. (2011) Kinetic mechanism for the excision of hypoxanthine by Escherichia coli AlkA and evidence for binding to DNA ends, *Biochemistry 50*, 4350-4359.
12. Taylor, E. L., and O'Brien, P. J. (2015) Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by AlkA, *Biochemistry 54*, 898-908.
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. Bennett, S. E., Sanderson, R. J., and Mosbaugh, D. W. (1995) Processivity of Escherichia coli and rat liver mitochondrial uracil-DNA glycosylase is affected by NaCl concentration, *Biochemistry 34*, 6109-6119.
15. Francis, A. W., and David, S. S. (2003) Escherichia coli MutY and Fpg utilize a

processive mechanism for target location, *Biochemistry 42*, 801-810.
16. Blainey, P. C., van Oijen, A. M., Banerjee, A., Verdine, G. L., and Xie, X. S. (2006) A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA, *Proceedings of the National Academy of Sciences of the United States of America 103*, 5752-5757.
17. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0, *Bioinformatics 23*, 2947-2948.
18. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool, *Journal of molecular biology 215*, 403-410.
19. Hedglin, M., and O'Brien, P. J. (2010) Hopping enables a DNA repair glycosylase to search both strands and bypass a bound protein, *ACS chemical biology 5*, 427-436.
20. Hedglin, M., Zhang, Y., and O'Brien, P. J. (2013) Isolating contributions from intersegmental transfer to DNA searching by alkyladenine DNA glycosylase, *The Journal of biological chemistry 288*, 24550-24559.
21. Zhang, Y., and O'Brien, P. J. (2015) Repair of Alkylation Damage in Eukaryotic Chromatin Depends on Searching Ability of Alkyladenine DNA Glycosylase, *ACS chemical biology*.
22. Baldwin, M. R., and O'Brien, P. J. (2010) Nonspecific DNA binding and coordination of the first two steps of base excision repair, *Biochemistry 49*, 7879-7891.
23. 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.
24. Maciejewska, A. M., Sokolowska, B., Nowicki, A., and Kusmierek, J. T. (2011) The role of AlkB protein in repair of 1,N(6)-ethenoadenine in Escherichia coli cells, *Mutagenesis 26*, 401-406.
25. Yao, M., Hatahet, Z., Melamede, R. J., and Kow, Y. W. (1994) Purification and characterization of a novel deoxyinosine-specific enzyme, deoxyinosine 3' endonuclease, from Escherichia coli, *The Journal of biological chemistry 269*, 16260-16268.
26. Asaeda, A., Ide, H., Asagoshi, K., Matsuyama, S., Tano, K., Murakami, A., Takamori, Y., and Kubo, K. (2000) Substrate specificity of human methylpurine DNA N-glycosylase, *Biochemistry 39*, 1959-1965.
27. Plosky, B. S., Frank, E. G., Berry, D. A., Vennall, G. P., McDonald, J. P., and Woodgate, R. (2008) Eukaryotic Y-family polymerases bypass a 3-methyl-2'-deoxyadenosine analog in vitro and methyl methanesulfonate-induced DNA damage in vivo, *Nucleic Acids Res 36*, 2152-2162.

Chapter IV

Distinguishing Specific and Nonspecific Complexes of Alkyladenine DNA Glycosylase

Abstract

Human alkyladenine DNA glycosylase (AAG) recognizes a broad range of alkylated and deaminated purine lesions and initiates the base excision DNA repair pathway by excising the damaged base. Previous work has shown that AAG is capable of facilitated diffusion, rapidly scanning nonspecific sites to locate sites of damage. Crystallographic data and biochemical data is consistent with AAG acting as a monomer to search DNA, but the ability of AAG to conduct intersegmental transfer raises the possibility that AAG may act as a transient dimer. Therefore, we have employed biochemical experiments to investigate the stoichiometry of AAG. We observe complex spectroscopic signals for AAG binding to nonspecific DNA, and demonstrate that these signals arise from relatively tight affinity for nonspecific DNA and gratuitous interaction between AAG and fluorescein. The resulting immobilization of the dye gives very large changes in fluorescence anisotropy, providing a sensitive assay for binding of AAG near to the dye. We take advantage of the natural fluorescence of the alkylated base, 1,N^6-ethenoadenine (εA) to independently measure interaction of AAG with the site of DNA damage. Quenching of εA-fluorescence by a 1:1 complex of AAG strongly suggests that AAG is active as a monomer, and the fluorescence anisotropy measurements are consistent with this model. Nevertheless, two

This work was inspired by preliminary data from Abby Wolfe and data was collected by Preethi Kesavan and Erin Taylor.

AAG molecules can bind to an oligonucleotide hairpin with 11 bp of duplex. This demonstrates that AAG is capable of binding at high density to nonspecific DNA. At 200 mM NaCl, equilibrium binding measurements reveal micromolar binding affinity to nonspecific sites, which results in nanomolar macroscopic dissociation constants for binding to short, undamaged oligonucleotides. Complex equilibrium binding curves and large anisotropy changes have been observed for other DNA repair glycosylases, suggesting that this behavior is not unique to AAG. Fluorescein-derivatives are stable to solid phase synthesis and deprotection chemistry and have strong fluorescence, making them popular choices for fluorescence anisotropy binding assays. However, many protein-DNA complexes exhibit complex behavior that has been interpreted in a variety of ways. Our results provide a unifying conceptual framework for understanding the behavior of dye-labeled DNA and the binding of proteins that exhibit nonspecific DNA binding.

Introduction

Humans have eleven genes encoding DNA repair glycosylases and these belong to four distinct superfamilies. Glycosylases initiate the base excision repair pathway by finding and excising damaged bases from genomic DNA.[1,2] These enzymes differ in their substrate specificity and catalytic mechanisms, but are thought to share some common features in how they search DNA to locate rare sites of damage. For example, most of these enzymes exhibit facilitated diffusion whereby nonspecific DNA binding interactions are used by the enzymes to diffuse along DNA in an efficient search for sites of damage.[3-7] The ability to bind tightly to nonspecific DNA is therefore critical to their biological function; however, it can complicate biochemical experiments monitoring DNA binding. Most structures of DNA glycosylases support monomeric models for damage recognition,[8] but biochemical evidence has been used in several cases to argue in favor of dimer models.[9-11] In the current work we have examined the stoichiometry and nonspecific DNA binding affinity of human alkyladenine DNA glycosylase (AAG).

AAG excises a wide range of *N*-alkylated purines, including 3-methyladenine, 7-methylguanine, and the lipid peroxidation byproduct, 1,N^6-ethenoadenine (εA).[12-14] In addition to the alkylated bases, AAG is responsible for removing hypoxanthine and xanthosine formed from the deamination of adenine and guanine, respectively.[15,16] The crystal structure of AAG in complex with εA-containing DNA reveals a positively charged DNA binding surface.[17] Mutation of these conserved residues leads to weakened DNA binding in vitro and reduced DNA repair activity in vivo.[17,18] Furthermore, the rate of dissociation of AAG from DNA is strongly

dependent on the monovalent salt concentration.[19,20]

Previous work has investigated the searching mechanism of AAG, demonstrating that facilitated diffusion is dominated by a rapid hopping mechanism.[21] Biochemical experiments also demonstrated that AAG is capable of intersegmental transfer, the direct transfer from one DNA segment or molecule to another without fully dissociating from DNA.[22] Theoretical treatment of intersegmental transfer has argued that this would be most easily achieved by a multimer with two or more DNA binding sites.[23] Attempts to crystallize AAG in complex with nonspecific DNA have not been successful, and therefore it is important to determine whether nonspecific complexes of AAG are monomers or higher order multimers. DNA binding studies for several different DNA glycoyslases have found evidence for multimers of the glycosylase binding to DNA.[9-11,24,25] It is expected that dimers would function differently than monomers in terms of DNA searching efficiency and therefore it is important to characterize the oligomerization state.

We used electrophoretic mobility shift assays (EMSA) and fluorescence anisotropy methods to study the binding of AAG to undamaged DNA and compared it to the binding of lesion-containing DNA. Both of these techniques have been widely used to characterize nucleic acid binding proteins. EMSA has the advantage of physically separating different complexes via native gels, but the equilibrium can be perturbed by electrophoresis and weak complexes may not be adequately trapped in the gel matrix.[26] Fluorescence anisotropy measures the tumbling of a fluorescently labeled molecule in solution and can be a versatile method for determining equilibrium binding in solution. In our studies, a DNA substrate labeled with a fluorophore tumbles faster and thus has a lower anisotropy value compared to a slower tumbling DNA:enzyme complex. Fluorescence anisotropy has the advantage of being measured in solution

and it is sensitive to the fluorescence lifetime and the frequency of tumbling in solution. However, the identity of the fluorophore and the linker that is used to attach it to DNA can have a major impact on the anisotropy measurements and there is always the possibility that a given protein will make gratuitous interactions with a particular chemical entity.[27,28] Therefore, it is important to compare binding measurements made with chemically distinct fluorophores.

We describe a comprehensive study of the affinity and stoichiometry of AAG binding to specific and nonspecific DNA. Fluorescence anisotropy titrations with fluorescein-labeled DNA could be interpreted with simple dimerization models, but by employing multiple approaches and systematically varying the length and labeling of the DNA, we provide evidence that interactions between the protein and the fluorescein are responsible for our observations. These experiments demonstrate that AAG is capable of binding with high affinity to nonspecific DNA and it can bind at high density (2 AAG molecules binding to an 11mer duplex). Direct measurement of εA-DNA quenching demonstrates that AAG functions as a monomer and the monomer model is consistent with all of the experimental data. These observations with AAG provide new insight into the interpretation of many different DNA binding studies that have been performed, in particular those that have employed fluorescein labels.

Materials and Methods

Preparation of Proteins and DNA

The catalytic domain of AAG lacking the first 79 amino acids was expressed in *Escherichia coli* BL21 (DE3) cells and purified following previous methods.[29] The concentration of AAG was estimated with UV absorbance, and the active concentration was determined by titration of εA-DNA as described below.[30]

Oligonucleotides were purchased from Integrated DNA Technologies or the Keck Center at Yale University. Unless otherwise noted, oligonucleotides were labeled on the 5′ end of the lesion-containing strand with a 6-fluorescein (FAM), 6-hexacholorofluorescein (HEX) or DyLight 647 (DY647). Oligonucleotides were purified with denaturing PAGE, extracted and desalted with a C18 reverse phase column. Concentrations of oligonucleotides were determined from A_{260} and calculated extinction coefficient. The extinction coefficients for DY647 substrates were estimated using the absorbance of cyanine 5. The extinction coefficients for εA-containing oligos were estimated by subtracting 9400 $M^{-1}cm^{-1}$ from the calculated extinction coefficients for the analogous A-containing oligos and the extinction coefficient of the pyrrolidine-containing substrates assumed no absorbance for the pyrrolidine. Oligonucleotides were mixed in annealing buffer (10 mM NaMES pH 6.5, 50 mM NaCl) with a 1.2-fold excess of the complementary strand and annealed by heating to 95 °C for 3 min, and cooling to 4 °C at a rate of 0.2 °C/s. All sequences used in this study are shown in Figure B-1.

Measurement of Fluorescence Anisotropy and Total Fluorescence

DNA was mixed with varying amounts of AAG in binding buffer (50 mM NaHEPES, pH 7.5, 100 or 200 mM ionic strength controlled with NaCl, 1 mM EDTA, 1 mM DTT, ± 0.1 mg/mL BSA) at 25 °C. BSA was not necessary and did not affect the anisotropy measurements (data not shown) but was included in some experiments to stabilize protein for multiple additions and mixings of AAG. All AAG additions were equilibrated for at least 45 sec prior to fluorescence detection and all titrations were completed in less than 10 min prior to εA excision. There was no significant difference of εA quenching between 10 min and 1 min reactions indicating that εA was not excised on this time scale. (data not shown). Fluorescence was detected using a T-format PTI QuantaMaster (FeliX software) or L-format Horiba Fluoromax-3 fluorometer (DataMax software). Excitation/emission wavelengths were 493/518 nm for FAM, 531/547 nm for HEX, 643/660 nm for DY647 (all 6 nm band-pass). Excitation and emission scans were taken using a vertical excitation and magic angle (54.7°) emission. The two emission filters were corrected for differences with the G-factor (Eq 1), and the anisotropy (r) and total fluorescence (F_T) of the end label were calculated using equations 2 and 3, where V and H denote the position of the excitation/emission polarization filters (V= 0°; H= 90°) and I is intensity of fluorescence.

$$G = \frac{I_{HV}}{I_{HH}} \tag{1}$$

$$r = \frac{I_{VV} - G * I_{VH}}{I_{VV} + 2 * G * I_{VH}} \tag{2}$$

$$F_T = I_{VV} + 2 * G * I_{VH} \tag{3}$$

For 1:1 binding curves observed under equilibrium binding conditions, the anisotropy was corrected to account for the quenching of the FAM fluorophore upon AAG binding.

Equation 4 describes the ratio of bound to free FAM-DNA where F is the molar concentration of the FAM-DNA, r is the anisotropy, Q is the molar fluorescence of the FAM-DNA, f and b represent the free and bound FAM-DNA states, and r and M represent the measured and corrected anisotropy values, respectively.[31] Equation 4 can be rearranged to solve for M shown in Equation 5.[32]

$$\frac{F_b}{F_f} = \left[\frac{r - r_f}{r_b - r}\right] * \frac{Q_f}{Q_b} = \frac{M - r_f}{r_b - M} \tag{4}$$

$$M = \frac{\left[\frac{r - r_f}{r_b - r}\right] * \frac{Q_f}{Q_b} * r_b + r_f}{1 + \left[\frac{r - r_f}{r_b - r}\right] * \frac{Q_f}{Q_b}} \tag{5}$$

For titration conditions, 100 or 300 nM DNA was mixed with AAG in 250 μL reaction volume (3 mm path length quartz cuvette). For equilibrium binding conditions, 5 or 15 nM DNA was mixed with AAG in 3 mL reaction volumes for greater sensitivity (10 mm path length quartz cuvette). For competition anisotropy, 150 nM 19AAA (undamaged) 5′FAM DNA and 150 nM AAG was premixed then titrated with various inhibitors under the 200 mM ionic strength condition. The change in anisotropy was normalized (r_N) from 0 (DNA alone) to 1 (AAG:DNA, no inhibitor) and fit using a one-site competitive binding model where IC_{50} is the concentration of inhibitor that reduces binding of AAG to labeled DNA by half (Eq 6).

$$r_N = \frac{1}{1 + \frac{[I]}{IC_{50}}} \tag{6}$$

The resulting IC_{50} values were converted to a macroscopic K_i value for each inhibitor using Equation 7 where S is the labeled DNA substrate, I is the unlabeled DNA inhibitor , and the K_d is the dissociation constant for the substrate (assumed to be 68 nM from the total FAM

fluorescence at 200 mM ionic strength (Figure 4-8)).

$$K_{\mathrm{i}} = \frac{\mathrm{IC}_{50}}{\left(1 + \frac{[\mathrm{S}]}{K_{\mathrm{d}}}\right)} \tag{7}$$

To find the microscopic K_{i} value for a single undamaged binding site, the macroscopic K_{i} values for each competitor were multiplied by the number of binding sites (N) on each competitor (Eq 8) in which L is the total length of the oligo, l is the site size or footprint of the enzyme (6 bp for duplex DNA; 5 bp for ssDNA),[17] and the number of sites is doubled to account for both DNA strands. This calculation assumes that AAG binds with equal affinity to all nonspecific sites.

$$N = 2(L - l + 1) \tag{8}$$

Anisotropy Data Fitting

All anisotropy and total fluorescence data, unless otherwise noted, was fit with DynaFit (BioKin).[33] One, two, and 3-site models were used to assess the various data sets. Example DynaFit models and fit parameters are detailed in the Supplemental Data.

Determination of AAG Concentration by εA-Quenching

The active concentration of AAG was determined by mixing varying amounts of AAG with 200 or 400 nM 19AEA (εA) 5′FAM DNA in binding buffer (100 mM ionic strength) without BSA (BSA interferes with εA fluorescence). Binding reactions were mixed for 45 sec at 25 °C before recording the εA fluorescence at excitation/emission wavelengths of 316/408 nm (both 6 nm band-pass). The εA fluorescence was normalized and fit to a 1:1 binding equation (Eq 9) where F is normalized fluorescence, A is the amplitude of fluorescence change, E and S are AAG and the substrate, respectively, and the K_{d} is the dissociation constant of εA binding.

Two protein preparations were used in this study and were 81% and 62% active, respectively (Figure B-2). All assays used the corrected active concentrations of AAG.

$$F = 1 - \frac{A\left\{(K_{\mathrm{d}} + [\mathrm{E}] + [\mathrm{S}]) - \sqrt{(K_{\mathrm{d}} + [\mathrm{E}] + [\mathrm{S}])^2 - 4[\mathrm{E}][\mathrm{S}]}\right\}}{2[\mathrm{S}]} \quad (9)$$

Electrophoretic Mobility Shift Assays (EMSA)

Horizontal gels were cast with 1% agarose and 2% acrylamide and run with TGE buffer (25 mM Tris Base, 190 mM Glycine, 1 mM EDTA). 1% agarose was melted in 1x TGE and cooled until warm to touch prior to addition of 2% acrylamide (37.5 acrylamide:bis-acrylamide ratio), 10% ammonium persulfate and TEMED. DNA and AAG were incubated for 15 min at 4 °C in binding buffer + 10% glycerol (v/v). Longer reaction times did not affect the fraction bound. We observed that bromophenol blue led to altered migration of the AAG•DNA complexes and therefore samples did not include loading dye (data not shown). Lanes were loaded with 50-500 fmol of DNA samples and electrophoresed at 120 V for 30-60 min at 4 °C. Gels were imaged using a Typhoon Trio Fluorescence imager (GE Healthcare) with 488 nm excitation and 520 nm band pass emission filters to detect FAM-labeled DNA species.

Results and Discussion

Considerable progress has been made in understanding the kinetic mechanism by which AAG searches nonspecific DNA and flips out an alkylated lesion,[6,17,19,21,22,34] but thermodynamic characterization of nonspecific DNA binding has remained elusive. In the current study, we employed EMSA and a variety of fluorescence methods to define the stoichiometry and thermodynamics of AAG binding to nonspecific DNA. We investigated two commonly employed fluorophores, fluorescein (FAM) and Dylight647 (DY647), and compared their effects on the DNA binding properties of AAG. FAM is a negatively charged dye that is commonly attached to DNA via a flexible linker. Previous studies suggest that FAM does not interact strongly with DNA, leading to weak coupling between the fluorophore and the DNA.[35] In contrast, cyanine dyes such as DY647 are positively charged and stack onto the ends of DNA duplexes.[35-37] This leads to a tighter coupling between the fluorophore and the DNA molecule. By comparing anisotropy and total fluorescence with two independent labels we are able to interpret the spectroscopic signals in terms of DNA binding and stoichiometry, whereas subsets of the data would be consistent with multiples models for AAG binding.

For this study, we employed a well characterized truncated form of AAG (Δ80 AAG) that has an identical rate of *N*-glycosidic bond cleavage on εA lesions, lacks the poorly ordered N-terminus, and has been used in crystallographic studies.[17,29,38]

Fluorescence Measurements of AAG Binding to εA-DNA with a 5′FAM Label.

Previous work suggests that a 19-mer oligonucleotide duplex is sufficiently large to

```
5′(FAM)-TAG CAT CAA X AAT TCT CTC-3′
     3′-ATC GTA GTT T TTA AGA GAG-5′
```

Figure 4-1. Representative DNA sequence used in this study. 19-mer double stranded DNA (AEA) was labeled with a 5′ 6-FAM and contains a centrally located εA lesion (X). All sequences used in this study are found in Figure B-1.

satisfy optimal binding of AAG.[39] Therefore, we examined binding of AAG to a 5′-6-carboxyfluorescein (FAM) labeled 19mer with a centrally located εA•T lesion site (Figure 4-1). AAG binds rapidly and tightly to εA-DNA,[19] and hydrolysis of the *N*-glycosidic bond is relatively slow under the conditions employed. This provides the opportunity to use steady-state fluorescence measurements to monitor titration of AAG to its preferred substrate and binding can be measured prior to catalysis. The intrinsic fluorescence of the εA lesion provides an independent method for determining the interaction of AAG with DNA, because the εA fluorescence is quenched upon binding in the active site pocket.[19,34]

Titrations were performed with a fixed concentration of DNA (100 nM or 300 nM 5′FAM-labeled 19AEA DNA duplexes) and increasing amounts of AAG (Figure 4-2; DynaFit models described in Figure B-3). We observed that the fluorescence anisotropy exhibited a biphasic behavior (Figure 4-2A), with the first equivalent of AAG causing a modest change in anisotropy and the second equivalent causing a much larger change in anisotropy. Replotting these data as a function of the ratio of AAG to DNA confirms that this experiment was conducted under titration conditions ([DNA]>>K_d; Figure 4-2B). The changes in total FAM fluorescence exhibit a similar biphasic process (Figure 4-2 C and D). The first equivalent of AAG gave little or no quenching of the FAM fluorescence and a second equivalent decreased the fluorescence to approximately 70% of the level of the free DNA. This pattern could be interpreted in terms of a cooperative binding transition in which two molecules of AAG bound to a single εA lesion site. However, we also measured the direct quenching of εA fluorescence

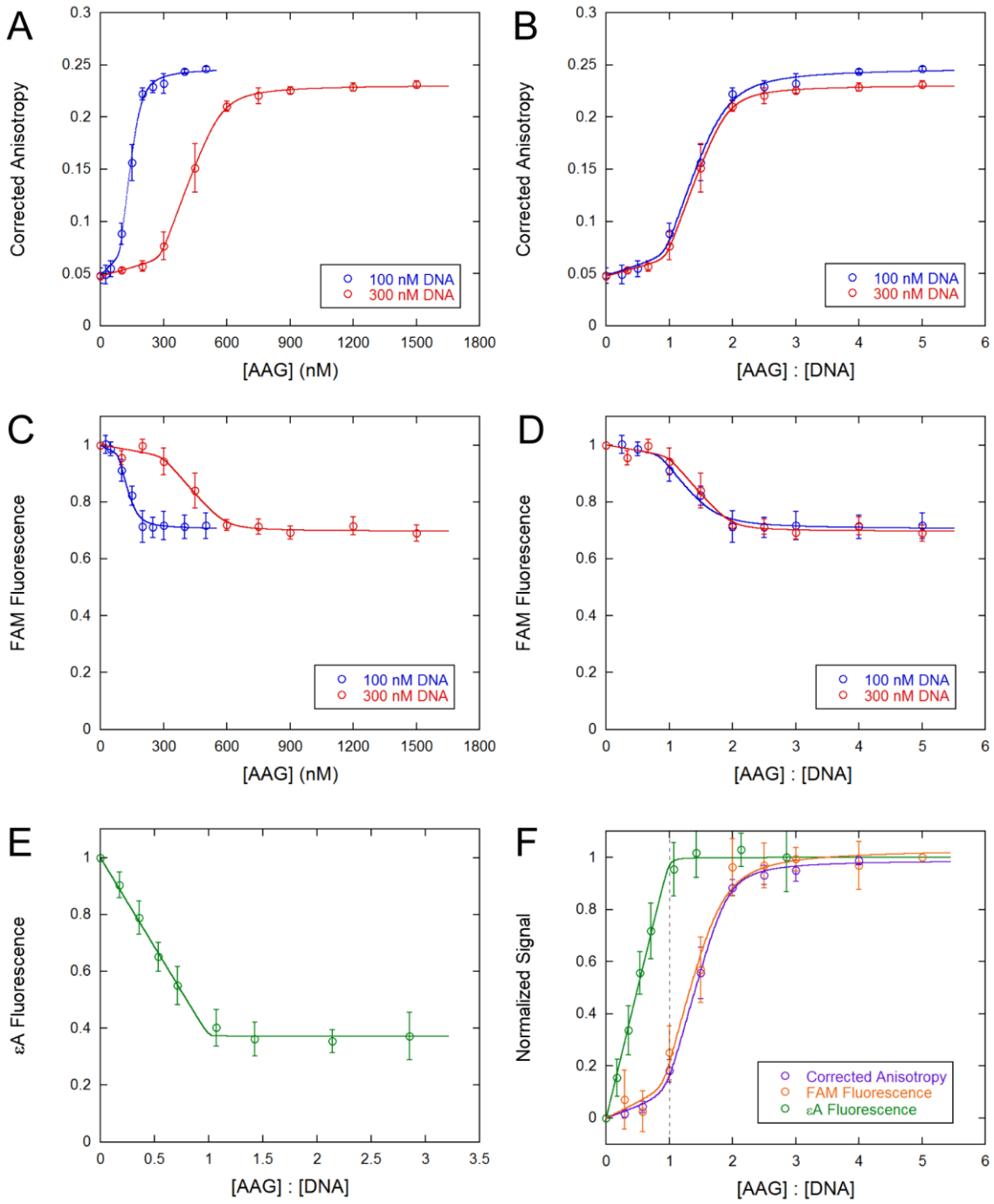


Figure 4-2. Fluorescent measurements for binding of AAG to 5′FAM-labeled εA-DNA under titration conditions. Fluorescence anisotropy (A, B), normalized total FAM fluorescence (C, D), and normalized εA fluorescence (E) of the FAM-labeled 19AEA DNA measured with increasing amounts of AAG. Two DNA concentrations (100 nM and 300 nM) were used and signals were plotted versus AAG concentration (A, C) or the ratio of AAG to DNA concentration (B, D, E). To facilitate comparison, the data at 100 and 300 nM DNA were averaged together and plotted in panel F. The εA quenching was fit by a 1:1 binding model (Eq 9) and DynaFit was used to fit anisotropy and total fluorescence data with a 2-site binding model (Figure B-3). The anisotropy was corrected to account for FAM quenching (Eq 5). The average ± SD is shown ($n \geq 4$).

using the same DNA under the same conditions and this shows a single phase, consistent with a 1:1 specific complex (Figure 4-2E; corrected from Figure B-2). The comparison of these three spectroscopic signals, εA quenching, FAM quenching, and FAM anisotropy indicates that the FAM label monitors binding of at least two molecules of AAG (Figure 4-2F). To account for the 30% quenching of FAM fluorescence, we corrected the anisotropy signal for binding of the second equivalent of AAG (Figures 4-2 and B-4). However, since there is not a significant difference between corrected and uncorrected anisotropy signals, and there is added complexity in the corrections of subsequent substrates, we will be reporting raw anisotropy values for all following assays unless noted.

Measurement of AAG Binding to Undamaged DNA with a 5′FAM Label

We next performed titrations of undamaged 19-mer DNA duplex with a 5′FAM label in order to compare the anisotropy and total 5′FAM fluorescence with that of the εA-DNA. Titrations performed at 100 nM and 300 nM DNA gave identical dependence on AAG concentration once they are plotted as a function of the number of equivalents (Figure B-5). AAG induced a very different pattern of changes in anisotropy and FAM fluorescence for the undamaged DNA as compared to the εA-DNA (Figure 4-3). The overall anisotropy change for the undamaged DNA is the same as for the εA-DNA, but the first equivalent caused a large change in anisotropy (>0.1) followed by a small change, which is the opposite of the behavior on the εA-DNA (Figure 4-3A). In other words, the first binding event on the undamaged DNA matched the second binding event on the εA-DNA, but two molecules of AAG bind to each DNA molecule. The quenching of total FAM fluorescence occurs upon the binding of the first equivalent of AAG for the undamaged DNA and for the second equivalent of AAG for the εA-

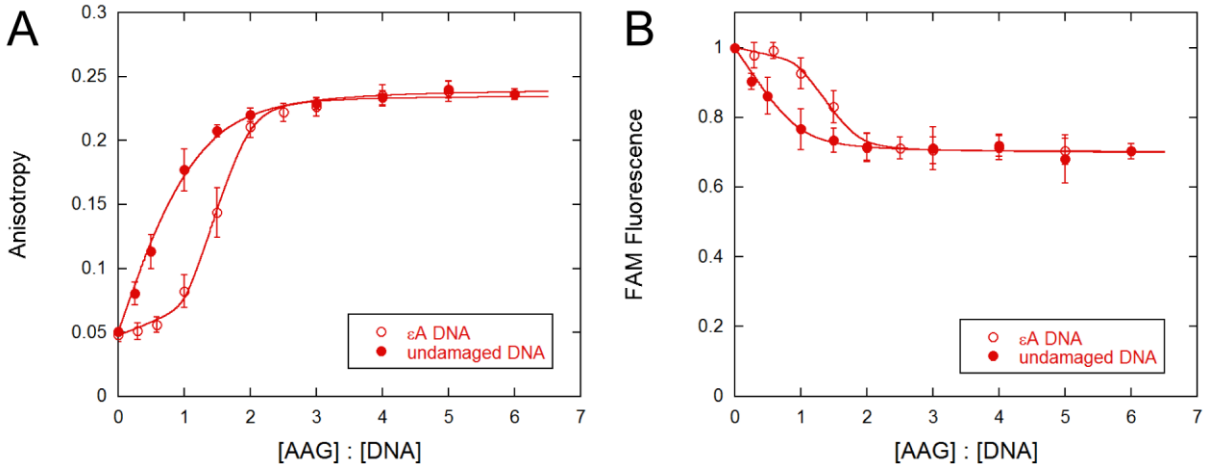


Figure 4-3. Comparison of AAG binding to 5′FAM-labeled εA-containing and undamaged DNA under titration conditions. Anisotropy (A) and normalized FAM fluorescence (B) for AAG binding to 19AAA undamaged (solid circles) or 19AEA εA (open circles) 5′FAM DNA. All data is the average of 100 nM and 300 nM DNA concentrations (Figures 4-2 and B-5). The undamaged DNA data were fit with DynaFit using a 2-site and 1-site model for anisotropy and total fluorescence, respectively (Figure B-6). The εA data were fit with 2-site models as described previously. The average ± SD is shown ($n \geq 4$).

DNA (Figure 4-3B). This suggests that the FAM quenching is attributed to AAG binding near to the 5′FAM-labeled DNA end.

To evaluate if the behavior of the 19mer εA-DNA is representative of tight binding, we performed analogous titration experiments with 25-mer DNA duplexes containing either a 5′FAM or 5′HEX and either a central εA•T lesion site or pyrrolidine•T transition state analog (pyr) which is known to bind tightly to AAG,[40] and we compared this to a 25-mer undamaged DNA. These experiments showed almost identical changes in anisotropy and total fluorescence as were observed for the 19-mer substrates (Figure B-7). Fitting of the anisotropy titrations with DynaFit revealed a slightly better fit of 3 AAG molecules to the 25-mers, as compared to the 2 AAG molecules observed for the 19-mers. Similar to the 19-mer substrates, the quenching of the total fluorescence appeared to follow a single AAG binding event for the undamaged DNA, and the second AAG binding event for the specific εA or pyr-containing DNA.

We sought an independent method for evaluating the model that multiple AAG molecules

were binding to a single DNA, and found EMSA to be a suitable approach. Using mixed agarose/acrylamide native gel conditions we probed the stoichiometry of AAG binding to the same 5′FAM-labeled 19-mer εA-DNA and undamaged DNA as was analyzed in the anisotropy experiments in Figure 4-3. As predicted by the FAM measurements, two sequential binding events could be observed for both DNA samples. For the εA-DNA, there were discrete bands corresponding to 1 and 2 equivalents of AAG (Figure 4-4A). Although the undamaged DNA also showed two distinct binding events, the bands were not as well defined as with the εA-DNA. It is

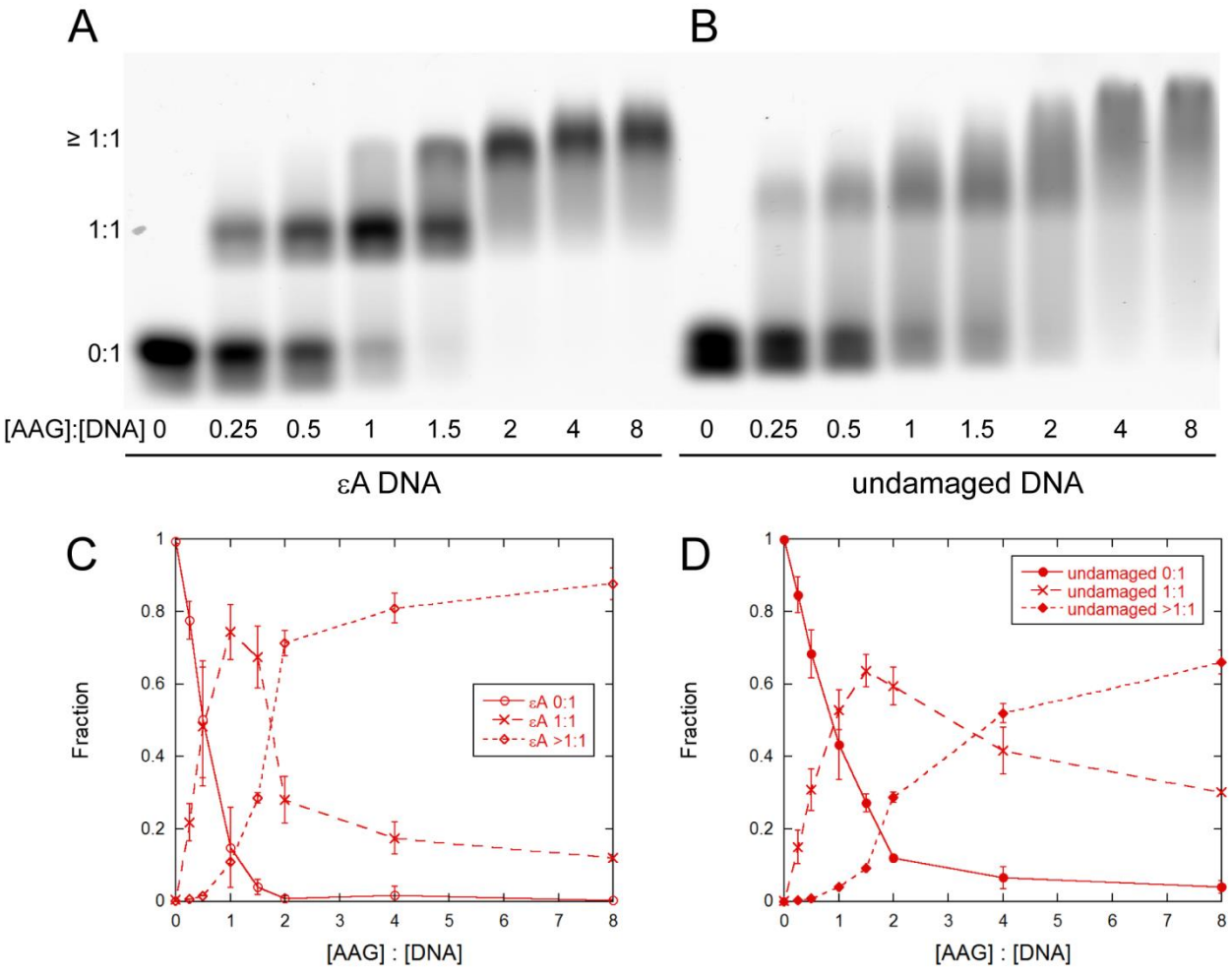


Figure 4-4. Representative EMSA with 5′FAM-labeled εA-containing and undamaged DNA. Equilibrium titration of AAG with 300 nM FAM-labeled 19AEA (A) and 19AAA (B) DNA. Stoichiometry of binding is consistent with the stoichiometry obtained from anisotropy measurements with the same DNA (Figure 4-3). Quantifications of independent experiments are averaged in C and D (n = 3).

expected that the nonspecific binding to the undamaged DNA reflects a larger ensemble of binding states and nonspecific complexes could dissociate and re-associate in the gel (note the smear between bands in the undamaged DNA in Figure 4-4B and in the second binding event for the εA-DNA in Figure 4-4A).

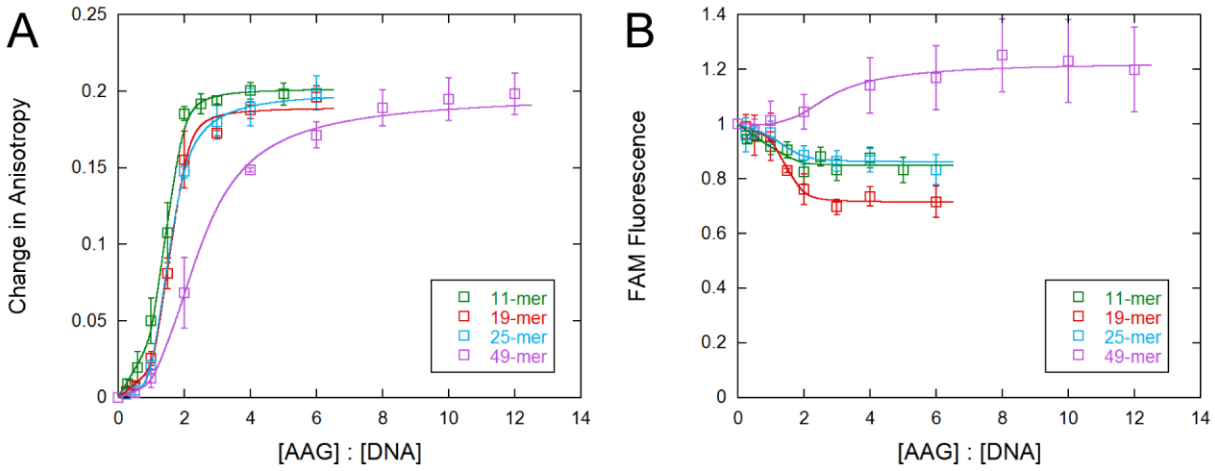
Taken together, both EMSA and FAM fluorescence provide strong evidence that two AAG molecules can bind to a 19-mer DNA duplex. The stoichiometric quenching of FAM fluorescence in the first binding event for undamaged DNA and in the second binding event for εA-DNA suggests that AAG has a propensity to bind near to the 5′FAM label. In turn, interaction with AAG appears to restrict the mobility of the FAM moiety and can account for the large change in anisotropy that correlates with the FAM quenching.

Binding Density of AAG

In order to explore the effect of DNA length on the anisotropy signals, we titrated several different length oligonucleotides that contained a single εA•T lesion site and a 5′FAM label. Binding of AAG near the 5′FAM causes a large anisotropy due to restricted mobility of the fluorophore, providing a robust signal for binding (Figure 4-5; DynaFit models in Figure B-8). As was observed for the 19-mer substrates, the first AAG binding event is to the lesion site and there is little change in anisotropy associated with this binding event. As the length was increased beyond 19 bp, additional molecules of AAG were required to achieve the full anisotropy change. For example, the titration curve for the 25-mer was best fit by a 3-site model. The 49-mer titration curve shows more than 3 binding events, but for simplicity we have employed a 3-site DynaFit model. We attribute this to competition between nonspecific binding sites away from the 5′FAM and the sites near to it, because the total fluorescence tracks closely

with this same stoichiometry. We verified that binding of AAG to the 49-mer was measured under titration conditions as it showed a much different binding density than the 19 and 25-mer oligonucleotides (Figure B-9). It should be noted that the 49-mer has a different sequence context at the 5′FAM site, which accounts for the enhancement of fluorescence upon binding to AAG.[41] Although it was not possible to definitively resolve complexes containing more than 3 bound AAG molecules using native gels, it is clear that at least 3 molecules can bind to the 25-mer and more than 3 to the 49-mer (Figure B-10). It should be noted that there was a slight change to the DNA sequence context for the comparison of DNA lengths (TEC vs. AEA; Figure B-1), but comparison of FAM-labeled DNA in both sequence contexts confirmed that the central sequence did not alter the stoichiometry or signals associated with AAG binding (Figure B-11).

We also investigated the possibility that a short hairpin oligonucleotide could be used to



4-5. Binding density of AAG on εA DNA duplexes. Anisotropy (A) and normalized total fluorescence (B) of 5′FAM labeled TEC DNA of different lengths. The 11-mer and 19-mer anisotropy and total fluorescence data were both fit to a 2-site binding model in DynaFit. The 25-mer anisotropy data were fit to a 3-site model and the total fluorescence to a 2-site model. The 49-mer anisotropy and total fluorescence data were both fit to a 3-site binding model. In all cases, the first binding event corresponds to the recognition of εA by AAG. The second binding event corresponds to AAG binding to the 5′FAM. Subsequent binding events are averaged together for the 49-mer substrate and are in competition with the FAM site. DynaFit parameters are shown in Figure B-8. The average of two substrate concentrations ± SD is shown (n = 3-6).

limit the nonspecific binding sites. Systematic analysis of optimal binding of abasic and hypoxanthine-containing oligonucleotides guided this design, with an upstream 6 bp segment and a downstream 2 bp segment joined by an internal GAAA hairpin (Figure B-1; 11-mer DNA).[39] Even on this minimal hairpin, a clear signature for two AAG binding events can be observed in the anisotropy and total fluorescence (Figure 4-5; 11-mer DNA), as well as in the EMSA (Figure B-10). This suggests that two AAG molecules bind in close proximity with an estimated footprint of 6 bp per AAG that is consistent with DNA contacts in the crystal structure of AAG bound to DNA.[17] This high binding density is not affected by the hairpin, because an asymmetric 19-mer (with an upstream segment of 6 bp and a downstream segment of 12 bp) gives very similar binding behavior (Figure B-12). To achieve this high density binding we infer that AAG must bind to the opposite sides of the DNA duplex, rather than side-by-side. Crude models suggest that this binding mode could be achieved without obvious steric clash (Figure B-13). It is intriguing that TDG is capable of binding in a similar fashion to adjacent sites on DNA.[9,25] We cannot ignore with a binding footprint of 6 bp, more than two AAG molecules may bind a 19-mer DNA. However, due to the supposed high affinity of AAG near the 5′FAM site, our methods limit our detection of these higher order complexes.

5′-Fluorophore Identity Affects Anisotropy Signal

To investigate the role of the 5′-label, we prepared 19-mer undamaged and εA-DNA with 5′DY647. DY647 is a more chemically stable analog of cyanine 5 (Cy5) that can be incorporated during solid phase synthesis. Whereas FAM does not stack on the end of the DNA,[35] NMR studies with Cy5 labeled DNA shows the fluorophore stably stacking on the end of DNA.[36] Stacking of the DY647 label on the end of the duplex would be expected to give tight coupling

between the tumbling of the fluorophore and of the DNA duplex. Consistent with this expectation, the anisotropy of the DY647-labeled DNA is much greater than for the FAM-labeled DNA (r = 0.25 and 0.05, respectively for the 19-mer DNA duplexes; Figure 4-6A).

Titration of DY647-labeled DNA shows a much smaller increase in anisotropy than for the FAM-labeled DNA, and no significant difference was observed between εA-DNA and undamaged DNA (Figure 4-6A). Due to the initial tight coupling of DY647 anisotropy to the free DNA, the change in anisotropy observed for binding both equivalents of AAG is similar and much less than was observed for binding adjacent to the FAM-labeled DNA (Figure 4-6B). The total fluorescence increases slightly for the DY647-labeled DNA, and while the εA-DNA could be fit to a 2-site model compared to the 1-site model for undamaged DNA, the difference is not dramatic (Figure 4-6C). These results may

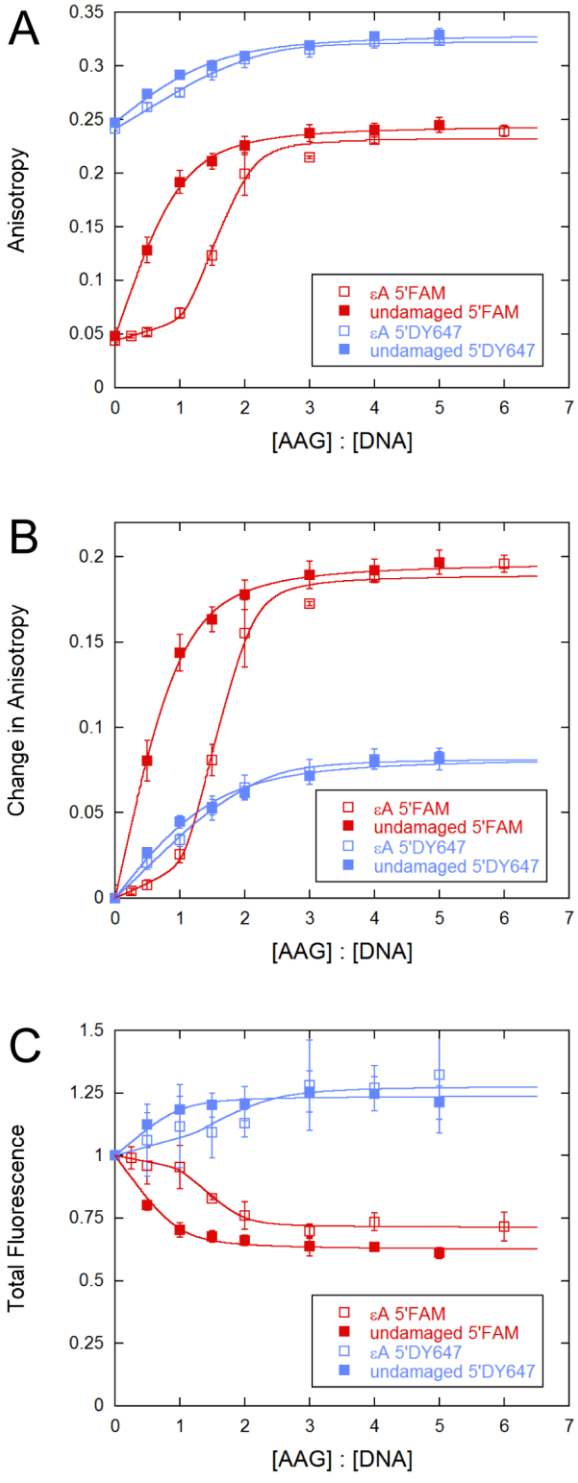


Figure 4-6. Comparison of 5′FAM and 5′DY647 labels in fluorescence measurements. Raw anisotropy (A), change in anisotropy (B) and normalized total fluorescence (C) of 5′-end labeled 19TEC and 19TAC DNA. AAG was incubated with 100 nM or 300 nM DNA under titration conditions for each sample. All the εA data was fit to a 2-site model, while the undamaged anisotropy and total fluorescence was fit to a 2-site and 1-site model, respectively (Figure B-14).

indicate that AAG is still preferentially binding near the 5′-end, but overall the 5′DY647 labeled DNA better reports on tumbling of the molecular AAG:DNA complex.

These data suggest that the large anisotropy changes for binding of AAG to undamaged DNA can be attributed to binding near to the 5′FAM. To address the question of whether AAG interacts directly with the FAM or to a specific DNA sequence context or structure upstream of the εA site, we measured the anisotropy of AAG binding to undamaged DNA with an alternate sequence or with the FAM on the 5′-end of the complement strand (Figure 4-7; Figure B-1). We found that there is no significant difference in the fluorescence anisotropy for these different DNA,

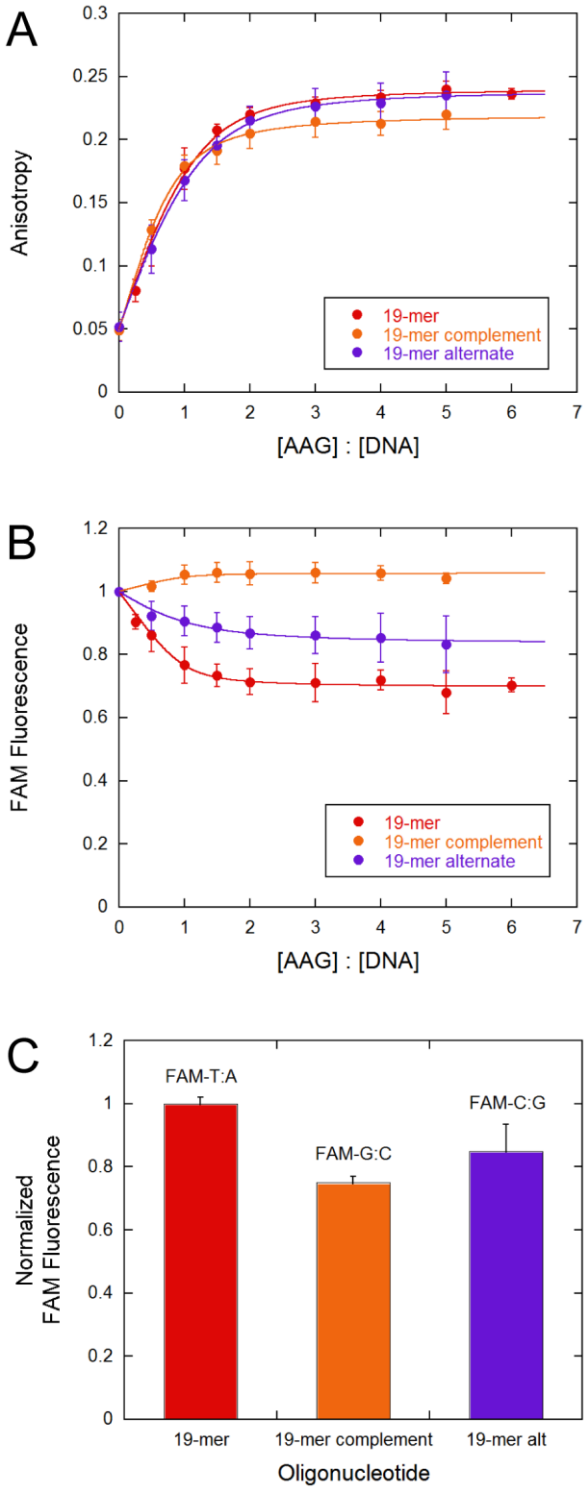


Figure 4-7. Fluorescence measurements for FAM-labeled undamaged DNA substrates. (A) Anisotropy and (B) normalized total FAM fluorescence of various FAM labeled undamaged 19-mer substrates to show that AAG is specifically binding FAM and thus inducing a large change in signal. The FAM was located on the 5′ end of the indicated strand adjacent to a T•A pair (19AAA, red), a C•G pair (19-alt, purple), or a G•C pair (19-complement; orange) (Figure B-1). Anisotropy and FAM fluorescence was fit to a 2-site and 1-site model as described for undamaged DNA substrates in Figure B-6. (C) Total FAM fluorescence intensity of the different DNA species is shown. The decreased fluorescence of FAM is attributed to quenching interaction with guanosine at the terminal base pair. The intrinsic differences in FAM fluorescence in the free DNA contribute to the observed changes in FAM fluorescence upon AAG binding. The average ± SD is shown ($n \geq 3$).

strongly suggesting that AAG is binding specifically to the FAM label regardless of the DNA sequence or location of the label. The total FAM fluorescence for the DNA duplexes show differences in the amount of fluorescent change, with the alternative sequence quenching to a lesser extent and the 5′FAM on the complement strand exhibiting a small increase in FAM fluorescence. The different extent of quenching can be at least partially attributed to the differences in the terminal base pair for the different oligonucleotides that were tested (Figure 4-7C). The relative fluorescence is consistent with quenching by a terminal guanosine nucleotide and the DNA with a terminal T•A base pair gives the largest amount of quenching upon AAG binding to DNA.[41]

Although the fortuitous interaction of AAG with a 5′FAM dye provides a large signal change that makes it easy to detect binding, it is expected that it contributes to DNA binding affinity relative to a DNA that is not labeled. In order to quantify the effect of the fluorescein in DNA binding affinity we first measure the equilibrium binding to undamaged DNA containing a 5′FAM label and then use competition experiments to measure the affinity of unlabeled DNA.

Affinity of AAG for Undamaged DNA Labeled with a 5′FAM

To measure the affinity of AAG to the FAM-labeled DNA we used a much lower concentration of DNA (5 nM) and varied the concentration of AAG (Figure 4-8). Under the standard conditions of 100 mM NaCl, both anisotropy and total fluorescence showed high affinity binding that was best fit by a single binding event, presumably because the first protein bound gives rise to the full amount of quenching and most of the change in anisotropy. As the apparent K_d value obtained from the fit is close to the concentration of the labeled-DNA it is likely that these conditions do not reflect equilibrium binding and this value should be

considered an upper limit to the true K_d value (Figure B-15 A and B). Increasing the DNA concentration to 15 nM shifted the apparent K_d value, confirming that binding affinity at 100 mM NaCl was too tight for us to measure with the FAM fluorescence ($K_d \leq 5$ nM). Although we could not get good signal to noise at lower concentrations of FAM-labeled DNA, we could increase the salt concentration to 200 mM ionic strength with NaCl, as this is predicted to weaken DNA binding affinity and facilitate quantitative analysis. The anisotropy and total FAM fluorescence for AAG binding both showed an increased K_d value at this higher salt concentration (Figure 4-8 A and B). Hyperbolic fits to a single site binding model yielded K_d values of 68 nM for the quenching of the FAM fluorescence and 146 nM for the anisotropy value (Figure B-15 C and D). However, since the equilibrium binding shows a one-site binding model, we corrected the anisotropy to account for the 30% quenching in total FAM fluorescence. This correction provides a K_d value of 96 nM, which is in reasonable agreement with the K_d value obtained from the total fluorescence (Figure 4-8A). This suggests that this measurement reflects equilibrium binding, and increased the concentration of DNA to 15 nM yielded identical curves and a very similar K_d value of 71 nM for FAM quenching (Figure B-12).

We also performed EMSA assays with the same undamaged DNA (5 nM) at both 100 mM and 200 mM ionic strength (Figure 4-8C). Quantification of the EMSA shows K_d values that are in reasonable agreement with the measurement of FAM quenching in solution (Figure 4-8D). We consider the solution-based measurement to be more reliable than the EMSA, because some of the AAG•DNA complexes appear to dissociate during electrophoresis.

Affinity of AAG for Undamaged DNA

To measure the affinity of AAG for undamaged and unlabeled DNA, we could take

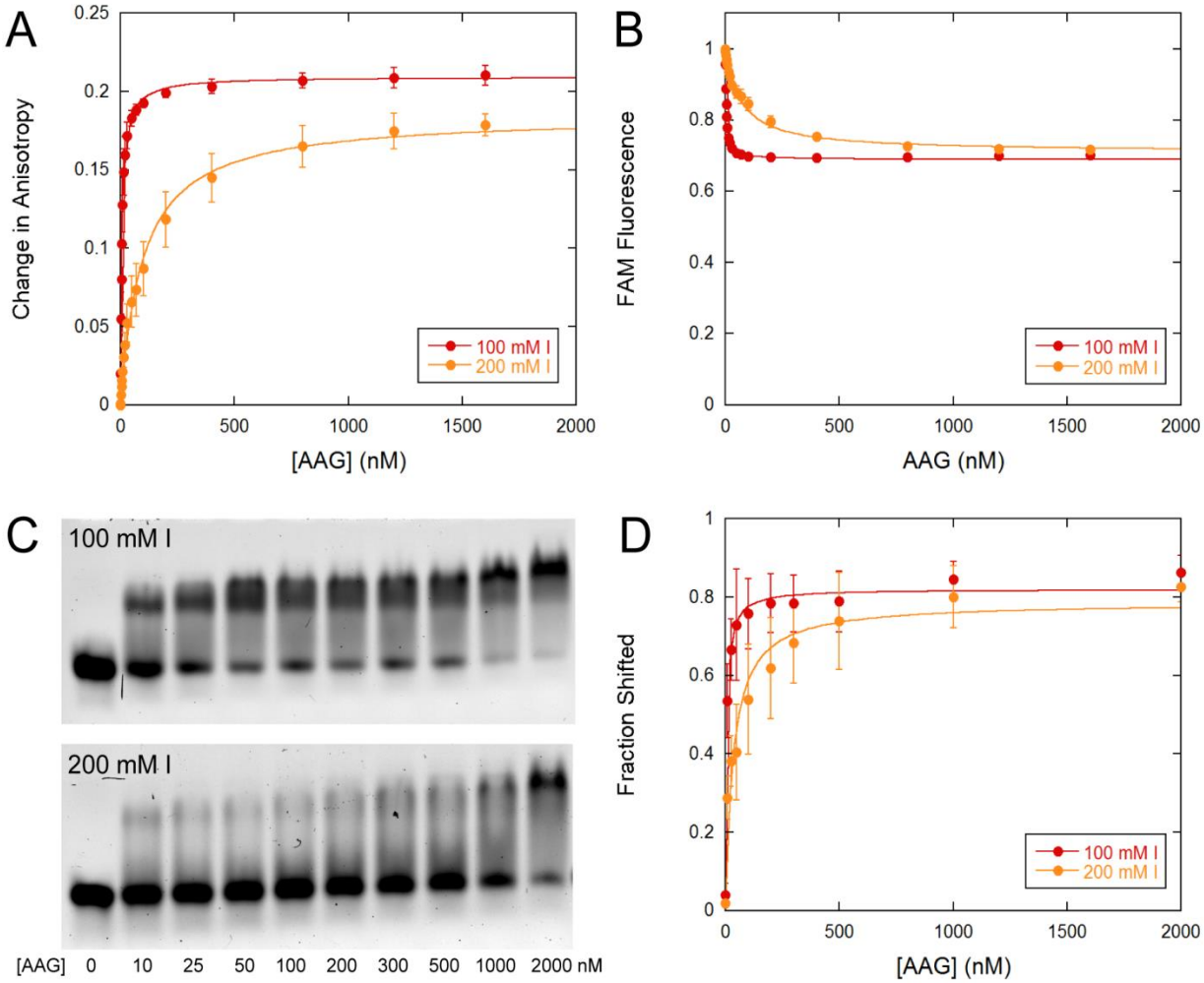


Figure 4-8. Equilibrium binding to 5′FAM-labeled undamaged DNA duplex at two concentrations of salt. Corrected anisotropy (A) and normalized total fluorescence (B) of AAG binding to the 5′FAM of 5 nM 19AAA DNA. The data was fit to a hyperbolic dependence representing one protein binding event. The K_d values for the anisotropy/total fluorescence data are 7/5 nM and 96/68 nM at 100 and 200 mM ionic strength, respectively. The average ± SD is shown (n = 3, 100 mM I; n = 2, 200 mM I). (C) Representative EMSA showing shift of 5 nM undamaged 5′FAM labeled 19-mer DNA upon binding of AAG. (D) Quantification of EMSA fit to a hyperbolic dependence. The K_d values are 6 and 34 nM for the 100 and 200 mM ionic strength conditions, respectively. The average ± SD is shown (n = 3). Both the anisotropy and EMSA data was corrected to account for the 30% quenching of FAM upon AAG binding.

advantage of the large changes in anisotropy that was observed for binding to the FAM-labeled DNA (Figure 4-8A) and compete labeled and unlabeled species. Therefore, we pre-incubated 5′-FAM labeled 19-mer undamaged DNA with equimolar amounts of AAG allowing binding to occur prior to adding various undamaged DNA to compete with binding to the FAM-labeled

DNA (Figure 4-9; Table 4-1). Unlabeled DNA exhibited a macroscopic K_i value of 112 nM (Table 4-1). In comparison, an analogous 5′-DY647-labeled 19-mer undamaged DNA gave a K_i value of 93 nM that is essentially the same, indicating that AAG does not binding preferentially to DY647 (Figure 4-9A). Additionally, we looked at the ability of different sized inhibitors to compete with the FAM-bound AAG. Since longer undamaged DNA contains more nonspecific binding sites, the macroscopic K_i values of such inhibitors are expected be tighter than shorter inhibitors. However, only a modest correlation between length of the inhibitor and the inhibition

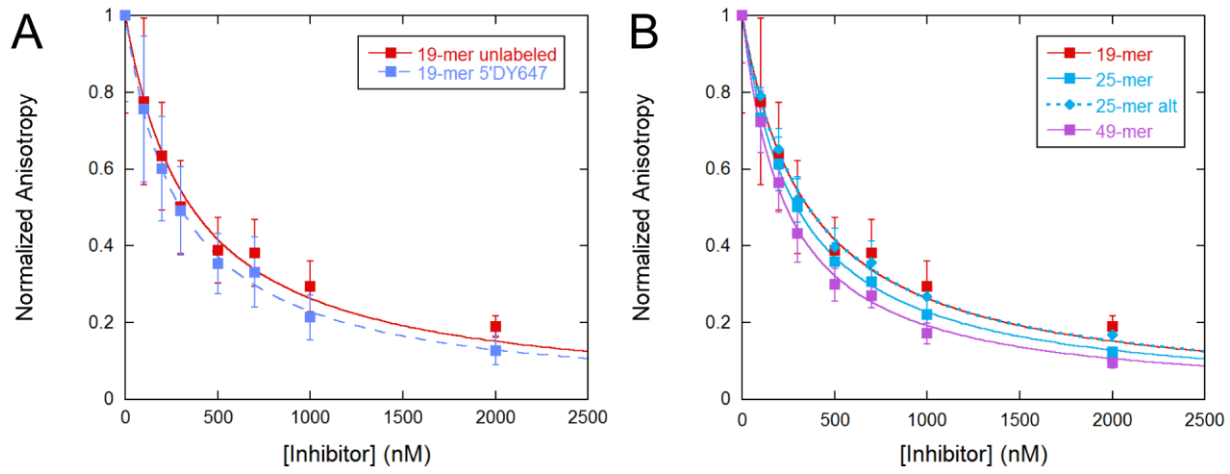


Figure 4-9. Anisotropy competition assay to measure binding of unlabeled undamaged DNA duplexes. Anisotropy of 150 nM AAG and 150 nM 5′FAM-labeled 19AAA DNA incubated with increasing amounts of inhibitor DNA at 200 mM ionic strength conditions. (A) Comparison of unlabeled and 5′-DY647 labeled undamaged 19TAC DNA inhibitors. (B) Comparison of different sized undamaged unlabeled TAC inhibitors. All data were fit with a one-site competitive binding model (Eq 6) and results are shown in Table 4-1. The average ± SD is shown (n ≥ 3).

Table 4-1. Analysis of anisotropy competition data

Inhibitor	*IC_{50} (nM)*	*K_i macroscopic (nM)*	*# binding sites*	*K_i microscopic (nM)*
19-mer	357.8	111.6	28	3125
19-mer DY647	297.3	92.7	28	2597
25-mer	295.2	92.1	40	3683
25-mer alt	358.9	112.0	40	4478
49-mer	238.3	74.3	88	6541
19-mer ss	689.5	215.1	30	6452
19-mer ss alt	607.9	189.6	30	5689

Equations 6-8 were used to convert the normalized anisotropy values into IC_{50}, macroscopic K_i, and microscopic K_i values, respectively. The K_d value used for the 5′FAM-19AAA substrate was assumed to be 68 nM based on the FAM quenching data (Figure 4-8B). The number of binding sites was estimated with a 6 bp footprint for duplex DNA and a 5 bp footprint for single strand DNA. All data was collected at 200 mM ionic strength condition.

constant was observed (Figure 4-9B; Table 4-1). An alternate 25-mer sequence was also tested and showed similar results, indicating that AAG is not binding a specific sequence context. These data may show that AAG molecules are preferentially binding to the ends of unlabeled DNA, thus cancelling the effect of a longer sequence. However, the 49-mer is a better inhibitor than the 19 and 25-mer DNAs, demonstrating some influence of the increased number of nonspecific binding sites. A microscopic K_i value of AAG for undamaged DNA can be calculated by multiplying the number of binding sites on each inhibitor by the macroscopic K_i value for the entire oligonucleotide (Table 4-1). While this calculation assumes AAG binds equally to all sites and the ends of DNA may be unequally preferred, we can conclude that AAG binds undamaged DNA in the low micromolar range.

Affinity of AAG for Single-Stranded DNA

AAG exhibits a preference for excision of lesion bases present in double-stranded (ds) DNA, but it has been shown to catalyze excision of some lesions present on single-stranded (ss) DNA, particularly εA.[42-44] To compare binding of ssDNA and dsDNA, we used fluorescence anisotropy methods to observe the binding of AAG to undamaged and εA-DNA in a single-stranded context (Figures 4-10 and B-16). Single strand sequences were chosen that exhibited no secondary complexes at reaction temperatures. As was observed for dsDNA, AAG binding to ssDNA caused a large change in anisotropy. Under titration conditions, it took 2 molecules of AAG to realize the full extent of fluorescence quenching and more than 4 molecules of AAG to cause the full change in anisotropy (Figure 4-10 A and B). This suggests that a greater number of binding sites exist on ssDNA as compared to dsDNA. As AAG efficiently excises εA from ssDNA under single turnover conditions (the maximal single turnover rate constant is only 5-fold

less for ssDNA than for dsDNA,[44] it was surprising to observe no significant difference between titration of εA-containing and undamaged DNA. This apparent lack of discrimination could be explained by the flexibility of ssDNA, because binding to the lesion and to the 5′FAM are not mutually exclusive. Both a specific complex and a nonspecific complex are able to interact to some extent with the FAM label, causing fluorescence quenching and enhanced fluorescence anisotropy. Subsequent binding of additional proteins causes more complete/stable interaction with the FAM label and also increases the tumbling time for the protein-DNA complex.

The EMSA approach did not prove to be very informative for characterizing the ssDNA complexes with AAG, because the DNA was retarded into lower mobility complexes without

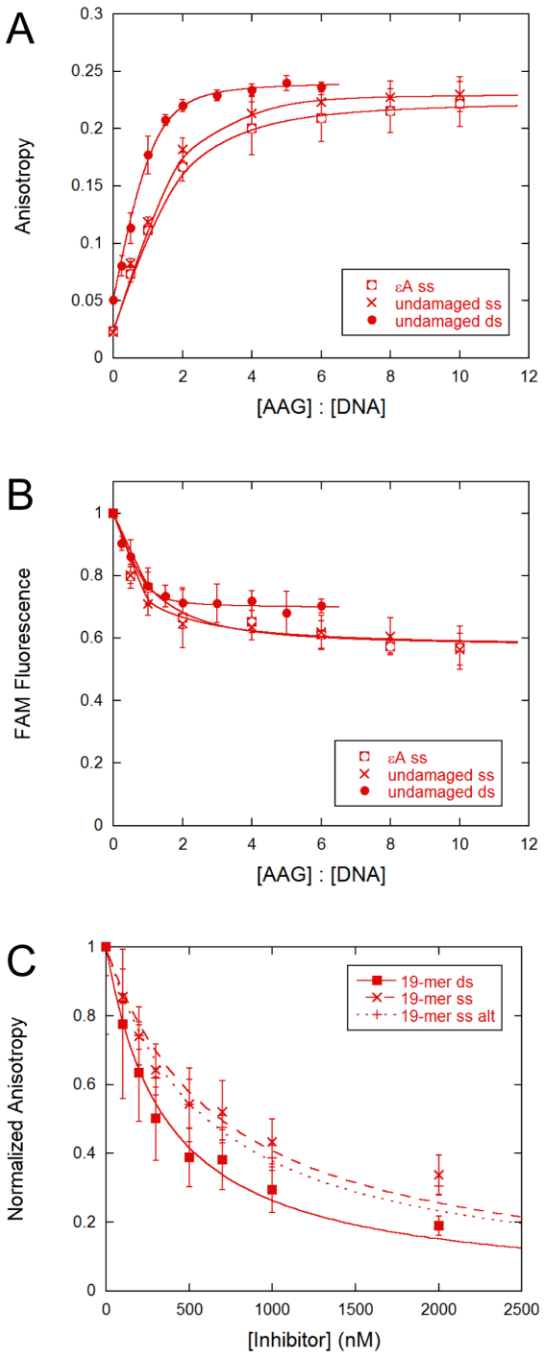


Figure 4-10. Fluorescence measurements of single stranded 19-mer DNA. (A) Anisotropy and (B) normalized total FAM fluorescence of 5′FAM-labeled single stranded 19AEA and 19AAA DNA substrates under titration conditions. Anisotropy and FAM fluorescence data were fit to a 4-site and 2-site model in DynaFit, respectively. The average ± SD is shown ($n \geq 3$). Double stranded 19AAA data is repeated from Figure 4-3. (C) Anisotropy of 150 nM AAG and 150 nM 5′FAM-labeled 19AAA DNA incubated with increasing amounts of inhibitor DNA at 200 mM ionic strength conditions. Inhibitors are double stranded 19TAC (repeated from Figure 4-9), single stranded 19TAC and a single stranded alternate 19-mer. Data were fit to a one site competition binding model and results are shown in Table 4-1. The average ± SD is shown ($n \geq 2$).

formation of discrete bands (Figure B-17). It is likely that an ensemble of states can form in which AAG is bound to many different sites, leading to a wide range of mobility. In addition, ssDNA affinity appears to be weaker than dsDNA affinity and this could make it difficult to trap the AAG•ssDNA complex in the gel. To measure of the affinity for ssDNA, we performed anisotropy competition assays (Figure 4-10C). Interestingly, the macroscopic K_d values for 19mer ssDNAs was only 2-fold larger than the K_d values for 19mer dsDNAs (Figure 4-10C; Table 4-1). The difference in affinity for ds and ss DNA can be attributed to both differences in the number of binding sites and in the average affinity for nonspecific ssDNA and dsDNA sites. If we assume a similar site size of 5 and 6 bp nucleotides for ss and ds substrates, respectively (4 AAG on 19-mer ss; 2 AAG on 11-mer ds) then the microscopic K_d for AAG binding is actually very similar (Table 4-1). This behavior is reminiscent of results obtained with other DNA repair proteins, whereby nonspecific binding interactions have similar energy for duplex and single-strand nucleic acids.[45]

Conclusions

Systematic comparison of anisotropy and fluorescence quenching for nonspecific and specific DNA with FAM and DY647 labeled oligonucleotides has enabled a minimal characterization of AAG binding energy and binding density. Examination of the literature for other DNA binding proteins provides evidence that the behavior that we observe for AAG may be relatively common. Most, if not all DNA repair proteins use nonspecific DNA binding interactions to search DNA for target sites. This affinity for nonspecific DNA can complicate

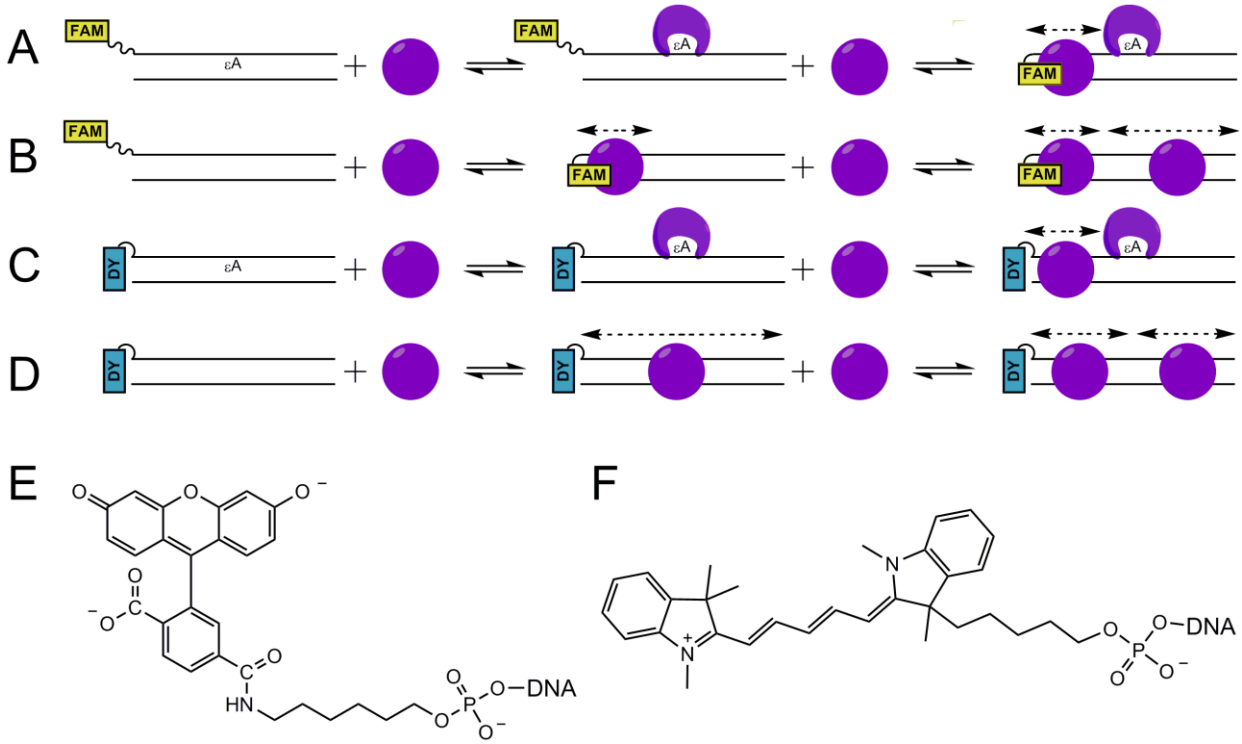


Figure 4-11. Schematic model depicting AAG binding events on εA and undamaged DNA. In the case of 5′FAM-labeled oligonucleotides, AAG binds εA specifically (A). If εA is absent (B) or has already been bound then AAG directly interacts with the FAM, causing a restriction of the fluorophore and coupling it to the protein:DNA complex. With 5′DY647 labeled DNA (C and D), there is no proof of AAG directly interacting with the fluorophore. Structures of 5′FAM and 5′DY647 are shown in E and F, respectively.

equilibrium measurements. We present a unified model to explain the observations with different labels and different assays that could be useful for characterization of other DNA repair enzymes (Figure 4-11). On the one hand it would appear to be advantageous to avoid using fluorescein-based dyes for anisotropy studies, because they are environmentally sensitive and poorly coupled to the tumbling of the nucleic acid to which they are attached. Cyanine and Rhodamine-based dyes appear to have more simple spectroscopic properties.[35,36,46] On the other hand, these properties can be easily interrogated by measurement of total fluorescence and this can give rise to unique signals for the measurement of DNA binding. In the specific case of AAG, the large change in anisotropy associated with binding near to a FAM label provides a robust assay for competitive binding and allows specific and nonspecific binding events to be readily distinguished.

Stoichiometry and Binding Density for AAG

Although the anisotropy and total fluorescence measurements that we observe for AAG binding to short oligonucleotides is consistent with a dimerization model for AAG binding, a more complete analysis supports the previously favored monomer model. This model invokes specific binding by an AAG monomer followed by high affinity, but weaker, binding to nonspecific sites. We found that AAG can bind with high density to both duplex and single strand DNA. For ssDNA, we can estimate a binding density of one protein molecule per 5 nucleotides based upon the observation that 4 proteins can bind to a 19-mer oligonucleotide. This is similar to the binding density of AAG on duplex DNA, with 2 AAG molecules binding to a short hairpin corresponding to a density of approximately one protein molecule per 6 bp. This density is very close to the footprint that is observed in the crystal structure of the AAG•DNA

complex of 5-6 bp (a 6 bp footprint was assumed for the duplex DNA).[17] The absence of strong interference between closely binding proteins suggests that AAG is capable of searching very small sites of accessible DNA in chromatin.

Binding Affinity of AAG for Nonspecific DNA

Using anisotropy competition assays, we were able to determine the binding affinity of AAG for ss and ds DNA in the absence of labels. In the presence of 200 mM NaCl and pH 7.5, macroscopic K_d values of 100-200 nM were measured for short oligonucleotides (Table 4-1). Considering a model of overlapping nonspecific binding sites, we calculate a rough estimate of 3 μM for both ss and ds DNA for an average affinity site. It is surprising that similar binding affinity is observed for ssDNA and dsDNA, because the glycosylase activity of AAG shows a strong preference for duplex DNA. For example, a 6000-fold preference for duplex over single strand DNA has been reported for AAG-catalyzed excision of εA.[43] This may suggest that nonspecific electrostatic interactions are similar with ssDNA and dsDNA, but only in dsDNA is AAG capable of making more extensive specific contacts. This is consistent with the crystal structure of AAG, in which DNA intercalation and contacts with the opposing strand would not be possible in ssDNA. For example, I161 and Y162 intercalate into the DNA and appear to stabilize the extrahelical base in the structure of AAG bound to εA-DNA.[17] In this same structure, there are extensive contacts between the DNA and T143 and M164 and electrostatic interactions between R141 and K229 and the phosphodiester backbone. The similar affinity between duplex and ssDNA is consistent with the previous findings that ssDNA contributes to processive searching of DNA,[22] and the observation that the AAG•ssDNA complex is sufficiently long-lived to allow for DNA annealing.

Consideration of Fluorescent Labeling Strategy for Characterization of Protein-DNA Complexes

Fluorescence anisotropy is an established and versatile technique for characterizing protein-DNA complexes.[27,28] It is generally advised that multiple labeling strategies be employed to vary the identity and position of the label, but in practice it is rare for systematic comparisons to be performed. In the current study we observe dramatic differences between FAM-labeled and DY647-labeled DNA. These data suggest that AAG preferentially interacts with a terminal FAM label and this gives rise to large changes in the fluorescence anisotropy of the protein•DNA complex. Once this is understood, fluorescence anisotropy can provide a robust assay for DNA binding. Fortuitously, this behavior also makes it easy to distinguish specific and nonspecific binding. Results obtained with many different DNA binding proteins suggest that studies of FAM-labeled DNA falls into one of two general behaviors. The most simple involves relatively small anisotropy values for binding, which suggests that there is no interaction between the protein and the FAM-label.[47,48] More commonly, large anisotropy values have been obtained.[10,49] In light of the results that we observe for AAG, we propose that very large observed changes in the anisotropy of FAM-labeled DNA should be treated with caution. Several different DNA glycosylases have been studied with FAM or the closely related HEX labels and binding has been reported to give large changes in anisotropy.[10,49] In these cases it would be useful to measure the total fluorescence to ask whether there is an enhancement or reduction in fluorescence. Even if no change is observed, this could reflect similarities between the collisional quenching that can be observed with guanosine nucleotides and with the protein environment that interacts with the dye (see Fig 4-7).[41] Therefore, it may also be necessary to test a different type of dye. We find that DY647 exhibits different properties and similarly Texas Red has been reported as an excellent choice for a fluorescent label that is tightly coupled to DNA and is

insensitive to local environment.[35,46]

References

1. O'Brien, P. J. (2006) Catalytic promiscuity and the divergent evolution of DNA repair enzymes, *Chem Rev 106*, 720-752.
2. Krokan, H. E., and Bjoras, M. (2013) Base excision repair, *Cold Spring Harbor perspectives in biology 5*, a012583.
3. Schonhoft, J. D., Kosowicz, J. G., and Stivers, J. T. (2013) DNA translocation by human uracil DNA glycosylase: role of DNA phosphate charge, *Biochemistry 52*, 2526-2535.
4. Blainey, P. C., van Oijen, A. M., Banerjee, A., Verdine, G. L., and Xie, X. S. (2006) A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA, *Proceedings of the National Academy of Sciences of the United States of America 103*, 5752-5757.
5. Porecha, R. H., and Stivers, J. T. (2008) Uracil DNA glycosylase uses DNA hopping and short-range sliding to trap extrahelical uracils, *Proceedings of the National Academy of Sciences of the United States of America 105*, 10791-10796.
6. Hedglin, M., and O'Brien, P. J. (2008) Human alkyladenine DNA glycosylase employs a processive search for DNA damage, *Biochemistry 47*, 11434-11445.
7. Lee, A. J., Warshaw, D. M., and Wallace, S. S. (2014) Insights into the glycosylase search for damage from single-molecule fluorescence microscopy, *DNA repair 20*, 23-31.
8. Brooks, S. C., Adhikary, S., Rubinson, E. H., and Eichman, B. F. (2013) Recent advances in the structural mechanisms of DNA glycosylases, *Biochimica et biophysica acta 1834*, 247-271.
9. Morgan, M. T., Maiti, A., Fitzgerald, M. E., and Drohat, A. C. (2011) Stoichiometry and affinity for thymine DNA glycosylase binding to specific and nonspecific DNA, *Nucleic Acids Res 39*, 2319-2329.
10. Grippon, S., Zhao, Q., Robinson, T., Marshall, J. J., O'Neill, R. J., Manning, H., Kennedy, G., Dunsby, C., Neil, M., Halford, S. E., French, P. M., and Baldwin, G. S. (2011) Differential modes of DNA binding by mismatch uracil DNA glycosylase from Escherichia coli: implications for abasic lesion processing and enzyme communication in the base excision repair pathway, *Nucleic Acids Res 39*, 2593-2603.
11. Rubinson, E. H., Gowda, A. S., Spratt, T. E., Gold, B., and Eichman, B. F. (2010) An unprecedented nucleic acid capture mechanism for excision of DNA damage, *Nature 468*, 406-411.
12. Engelward, B. P., Boosalis, M. S., Chen, B. J., Deng, Z., Siciliano, M. J., and Samson, L. D. (1993) Cloning and characterization of a mouse 3-methyladenine/7-methyl-guanine/3-methylguanine DNA glycosylase cDNA whose gene maps to chromosome 11, *Carcinogenesis 14*, 175-181.
13. Saparbaev, M., Kleibl, K., and Laval, J. (1995) Escherichia coli, Saccharomyces cerevisiae, rat and human 3-methyladenine DNA glycosylases repair 1,N6-ethenoadenine when present in DNA, *Nucleic Acids Res 23*, 3750-3755.
14. Smith, S. A., and Engelward, B. P. (2000) In vivo repair of methylation damage in Aag

3-methyladenine DNA glycosylase null mouse cells, *Nucleic Acids Res 28*, 3294-3300.
15. Saparbaev, M., and Laval, J. (1994) Excision of hypoxanthine from DNA containing dIMP residues by the Escherichia coli, yeast, rat, and human alkylpurine DNA glycosylases, *Proceedings of the National Academy of Sciences of the United States of America 91*, 5873-5877.
16. Wuenschell, G. E., O'Connor, T. R., and Termini, J. (2003) Stability, miscoding potential, and repair of 2'-deoxyxanthosine in DNA: implications for nitric oxide-induced mutagenesis, *Biochemistry 42*, 3608-3616.
17. 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, *Proceedings of the National Academy of Sciences of the United States of America 97*, 13573-13578.
18. Zhang, Y., and O'Brien, P. J. (2015) Repair of Alkylation Damage in Eukaryotic Chromatin Depends on Searching Ability of Alkyladenine DNA Glycosylase, *ACS chemical biology*.
19. 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.
20. 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.
21. Hedglin, M., and O'Brien, P. J. (2010) Hopping enables a DNA repair glycosylase to search both strands and bypass a bound protein, *ACS chemical biology 5*, 427-436.
22. Hedglin, M., Zhang, Y., and O'Brien, P. J. (2013) Isolating contributions from intersegmental transfer to DNA searching by alkyladenine DNA glycosylase, *The Journal of biological chemistry 288*, 24550-24559.
23. Berg, O. G., Winter, R. B., and von Hippel, P. H. (1981) Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory, *Biochemistry 20*, 6929-6948.
24. 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, *Journal of molecular biology 381*, 13-23.
25. Maiti, A., Morgan, M. T., Pozharski, E., and Drohat, A. C. (2008) Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition, *Proceedings of the National Academy of Sciences of the United States of America 105*, 8890-8895.
26. Hellman, L. M., and Fried, M. G. (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions, *Nature protocols 2*, 1849-1861.
27. Anderson, B. J., Larkin, C., Guja, K., and Schildbach, J. F. (2008) Using fluorophore-labeled oligonucleotides to measure affinities of protein-DNA interactions, *Methods in enzymology 450*, 253-272.
28. Heyduk, T., Ma, Y., Tang, H., and Ebright, R. H. (1996) Fluorescence anisotropy: rapid, quantitative assay for protein-DNA and protein-protein interaction, *Methods in enzymology 274*, 492-503.
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. Hendershot, J. M., and O'Brien, P. J. (2014) Critical role of DNA intercalation in enzyme-catalyzed nucleotide flipping, *Nucleic Acids Res*.
31. Dandliker, W. B., Hsu, M. L., Levin, J., and Rao, B. R. (1981) Equilibrium and kinetic inhibition assays based upon fluorescence polarization, *Methods in enzymology 74 Pt C*, 3-28.
32. Invitrogen (2006) *Fluorescence Polarization Technical Research Guide; Chapter 3 Immunoassays*, Vol. 4th Ed, 4th ed.
33. Kuzmic, P. (1996) Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase, *Analytical biochemistry 237*, 260-273.
34. Hendershot, J. M., Wolfe, A. E., and O'Brien, P. J. (2011) Substitution of active site tyrosines with tryptophan alters the free energy for nucleotide flipping by human alkyladenine DNA glycosylase, *Biochemistry 50*, 1864-1874.
35. Norman, D. G., Grainger, R. J., Uhrin, D., and Lilley, D. M. (2000) Location of cyanine-3 on double-stranded DNA: importance for fluorescence resonance energy transfer studies, *Biochemistry 39*, 6317-6324.
36. Iqbal, A., Wang, L., Thompson, K. C., Lilley, D. M., and Norman, D. G. (2008) The structure of cyanine 5 terminally attached to double-stranded DNA: implications for FRET studies, *Biochemistry 47*, 7857-7862.
37. Kroutil, O., Romancova, I., Sip, M., and Chval, Z. (2014) Cy3 and Cy5 dyes terminally attached to 5'C end of DNA: structure, dynamics, and energetics, *The journal of physical chemistry. B 118*, 13564-13572.
38. O'Connor, T. R. (1993) Purification and characterization of human 3-methyladenine-DNA glycosylase, *Nucleic Acids Res 21*, 5561-5569.
39. Baldwin, M. R., and O'Brien, P. J. (2012) Defining the functional footprint for recognition and repair of deaminated DNA, *Nucleic Acids Res 40*, 11638-11647.
40. Scharer, O. D., Nash, H. M., Jiricny, J., Laval, J., and Verdine, G. L. (1998) Specific binding of a designed pyrrolidine abasic site analog to multiple DNA glycosylases, *The Journal of biological chemistry 273*, 8592-8597.
41. Nazarenko, I., Pires, R., Lowe, B., Obaidy, M., and Rashtchian, A. (2002) Effect of primary and secondary structure of oligodeoxyribonucleotides on the fluorescent properties of conjugated dyes, *Nucleic Acids Res 30*, 2089-2195.
42. Lee, C. Y., Delaney, J. C., Kartalou, M., Lingaraju, G. M., Maor-Shoshani, A., Essigmann, J. M., and Samson, L. D. (2009) Recognition and processing of a new repertoire of DNA substrates by human 3-methyladenine DNA glycosylase (AAG), *Biochemistry 48*, 1850-1861.
43. Lyons, D. M., and O'Brien, P. J. (2009) Efficient recognition of an unpaired lesion by a DNA repair glycosylase, *Journal of the American Chemical Society 131*, 17742-17743.
44. Hedglin, M., Zhang, Y., and O'Brien, P. J. (2015) Probing the DNA structural requirements for facilitated diffusion, *Biochemistry 54*, 557-566.
45. Nevinsky, G. A. (2011) Structural, thermodynamic, and kinetic basis for the activities of some nucleic acid repair enzymes, *Journal of molecular recognition : JMR 24*, 656-677.
46. Unruh, J. R., Gokulrangan, G., Wilson, G. S., and Johnson, C. K. (2005) Fluorescence properties of fluorescein, tetramethylrhodamine and Texas Red linked to a DNA aptamer, *Photochemistry and photobiology 81*, 682-690.
47. Reid, S. L., Parry, D., Liu, H. H., and Connolly, B. A. (2001) Binding and recognition of GATATC target sequences by the EcoRV restriction endonuclease: a study using

fluorescent oligonucleotides and fluorescence polarization, *Biochemistry 40*, 2484-2494.
48. Ozers, M. S., Hill, J. J., Ervin, K., Wood, J. R., Nardulli, A. M., Royer, C. A., and Gorski, J. (1997) Equilibrium binding of estrogen receptor with DNA using fluorescence anisotropy, *The Journal of biological chemistry 272*, 30405-30411.
49. Bharti, S. K., and Varshney, U. (2010) Analysis of the impact of a uracil DNA glycosylase attenuated in AP-DNA binding in maintenance of the genomic integrity in Escherichia coli, *Nucleic Acids Res 38*, 2291-2301.

Appendix B

Figures and Tables to Accompany Chapter IV

Figure B-1. DNA oligonucleotides used in this study. For all DNA molecules, X may represent εA (1,N^6-ethenoadenine; E), A (undamaged) or pyrrolidine (pyr) as described for each experiment. FAM (6-fluorescein), DY647 (DyLight 647), or HEX (6-hexachlorofluorescein) are attached on the 5′-end of all nucleotides when described, with some special cases being labeled on the 5′-end of the complement strand. Lesion containing strands and complements are shown in upper and lowercase, respectively.

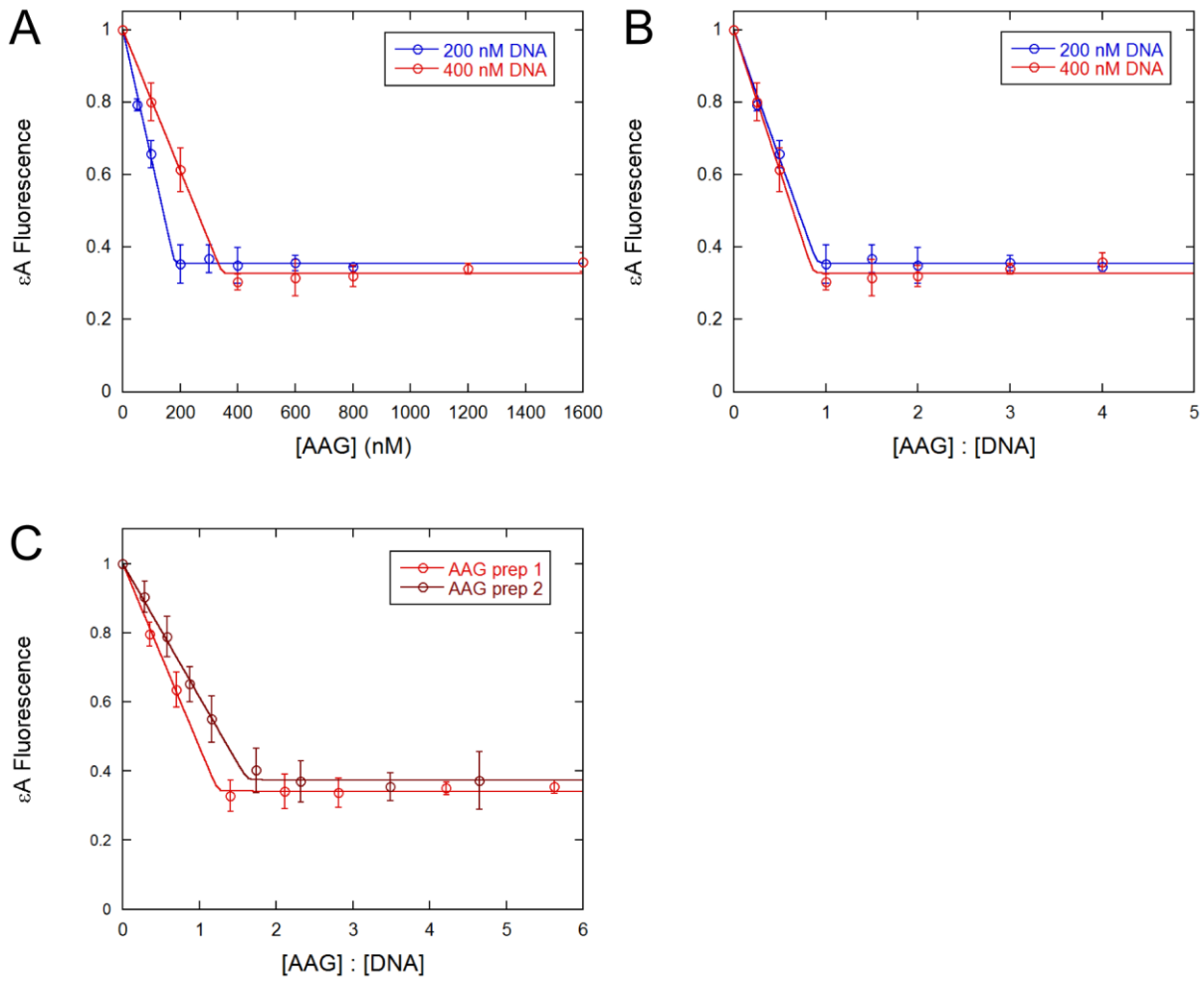


Figure B-2. Quenching of εA fluorescence by AAG. The εA fluorescence of 200 nM and 400 nM 5′FAM labeled 19AEA DNA was normalized and plotted versus the corrected concentration of AAG (A) or the corrected ratio of AAG to DNA concentration (B). The data were fit by a 1:1 binding equation (Eq 9; Chapter 4) and show a 1:1 ratio of AAG to DNA. The average ± SD is shown (n ≥ 3). (C) The εA fluorescence was normalized and averaged for both concentrations of DNA tested with two AAG preparations. The concentration of AAG was estimated from the UV absorbance at 280 nm. The data was fit by a 1:1 binding equation to determine the active concentration of AAG. The first AAG prep is 81% active and the second is 62% active, relative to the UV absorbance. This is within the range of ~50%-100% that has previously been measured for other preparations of AAG.[1] Each preparation was corrected for the active fraction and these values are used in all subsequent experiments.

A

```
[task]
        data = equilibria
        task = fit[model]
        2 site ?
[mechanism]
        AAG + DNA <==> AAG.DNA : Kd1 dissoc
        AAG.DNA + AAG <==> AAG.AAG.DNA : Kd2 dissoc
[constants]
        Kd1 = 0.020
        Kd2 = 1 ?
[concentrations]
        DNA = 300 ?
[responses]
        intensive
        DNA = 0.047353
        AAG.DNA = 0.062 ?
        AAG.AAG.DNA = 0.23170 ?
[data]
        mesh from 0 to 1650 step 1
        variable AAG
        file .\[insert data file location]
[output]
        directory           .\[insert output folder location]
```

Figure B-3. DynaFit model and fits for 5′FAM-labeled 19-mer εA DNA.
(A) Representative DynaFit model. The 300 nM 19AEA corrected anisotropy is shown (Figure 4-2 A).
(B) Parameters for each Dynafit model corresponding to Figure 4-2.

B

Anisotropy				
Parameters		**100nM Corrected**	**300nM Corrected**	**Normalized Corrected**
guesses	*DNA*	100 nM float	300 nM float	100 nM float
	Kd1	20 pM fix	20 pM fix	20 pM fix
	Kd2	1 nM float	1 nM float	1 nM float
	r(DNA)	0.04807 fix	0.047353 fix	1e-20 float (0)
	r(AAG.DNA)	0.062 float	0.062 float	0.071 float
	r(AAG.AAG.DNA)	0.245 float	0.23170 float	1 float
fits	*DNA*	93.3 nM (93.3%)	285.8 (95.3%)	95.5 (95.5%)
	Kd2	6.5 nM	10.2 nM	3.2 nM
	r(AAG.DNA)	0.066	0.63	0.083
	r(AAG.AAG.DNA)	0.248	0.231	0.99
Total Fluorescence				
Parameters		**100nM Corrected**	**300nM Corrected**	**Normalized Corrected**
guesses	*DNA*	100 nM float	300 nM float	100 nM float
	Kd1	20 pM fix	20 pM fix	20 pM fix
	Kd2	6.6 nM fix	6.6 nM fix	6.6 nM fix
	r(DNA)	1 fix	1 fix	1e-20 float (0)
	r(AAG.DNA)	1 float	1 float	0.08 float
	r(AAG.AAG.DNA)	0.7151 float	0.705878 float	1 float
fits	*DNA*	85.0 (85.0%)	296.0 (98.7%)	90.5 (90.5%)
	r(AAG.DNA)	0.969	0.963	0.108
	r(AAG.AAG.DNA)	0.703	0.696	1.03

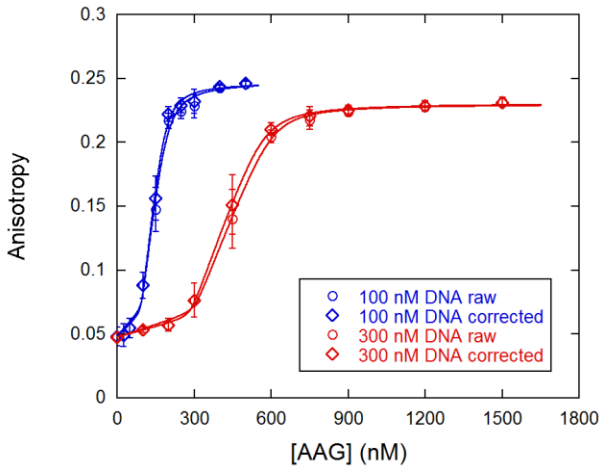


Figure B-4. Correction of 5′FAM 19-mer εA anisotropy. The second AAG binding event of the 19AEA anisotropy was corrected for the 30% quenching of FAM fluorescence using Equation 5 (Chapter 4). The first binding event was not corrected due to the insignificant change in FAM fluorescence up to the 1:1 AAG to DNA ratio (Figure 4-2). The corrected anisotropy values are shown in Figure 4-2 but due to the complexity of binding events for undamaged DNA and longer DNA substrates, as well as the insignificant difference between uncorrected and corrected anisotropy signals, all other anisotropy data collected under titration conditions will not be corrected.

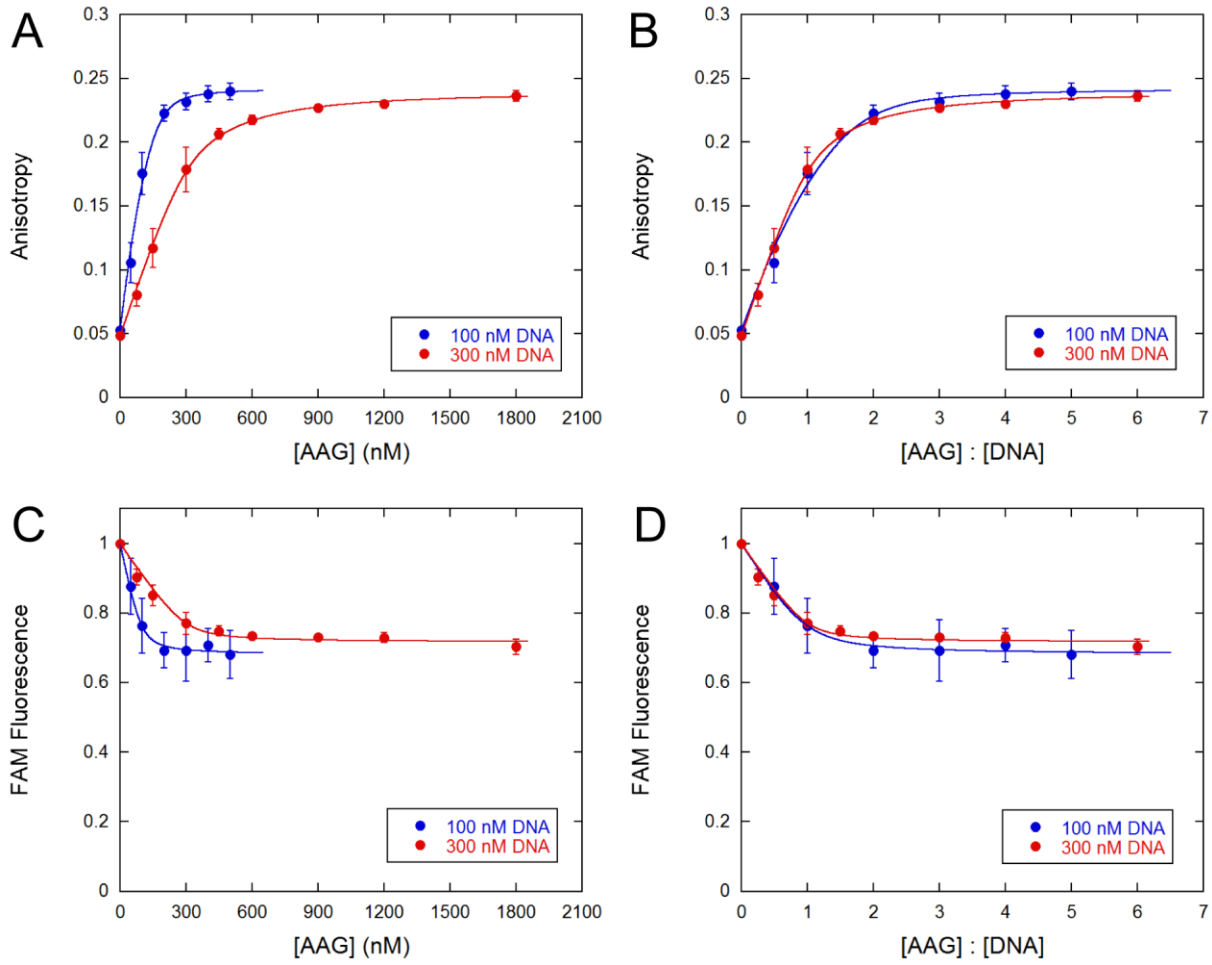


Figure B-5. Fluorescence measurements for binding of AAG to 5′FAM-labeled undamaged DNA under titration conditions. Fluorescence anisotropy (A, B) and normalized total FAM fluorescence (C, D) of the FAM-labeled 19AAA DNA when mixed with increasing amounts of AAG. Two DNA concentrations (100 nM and 300 nM FAM) were used and signals were either plotted versus AAG concentration (A, C), or the ratio of AAG to DNA concentration (B, D) to showcase titration conditions of all binding events. Anisotropy data were fit to a 2-site model and total FAM fluorescence was fit to a 1-site model using Dynafit (Figure B-6). The first binding event describes AAG binding near the 5′FAM site, while the second anisotropy binding event is subsequent AAG molecules binding the substrate, and thus changing the tumbling of the molecular complex. The average ± SD is shown ($n \geq 3$).

A

```
[task]
        data = equilibria
        task = fit
[model]
        1 site ?
[mechanism]
        AAG + DNA <==> AAG.DNA : Kd1 dissoc
[constants]
        Kd1 = 1 ?
[concentrations]
        DNA = 100
[responses]
        intensive
        DNA = 1
        AAG.DNA = 0.70 ?
[data]
        mesh from 0 to 650 step 1
        variable AAG
        file .\[insert data file location]
[output]
        directory           .\[insert output folder location]
 [end]
```

Figure B-6. DynaFit model and fits for 5′FAM-labeled 19-mer undamaged DNA.
(A) Representative DynaFit model. The averaged 19AAA FAM fluorescence is shown (Figure 4-3 B).
(B) Parameters for each undamaged DNA Dynafit model corresponding to Figures B-5 and 4-3.

B

Anisotropy (2-site model)				
Parameters		**100 nM DNA**	**300 nM DNA**	**Average**
guesses	*DNA*	100 nM fix	300 nM fix	100 nM fix
	Kd1	6.5 nM fix	6.5 nM fix	6.5 nM fix
	Kd2	100 nM float	100 nM float	100 nM float
	r(DNA)	0.0526 fix	0.0483 fix	0.0504 fix
	r(AAG.DNA)	0.175 float	0.179 float	0.177 float
	r(AAG.AAG.DNA)	0.24 float	0.237 float	0.24 float
fits	*Kd2*	24.8 nM	177.2 nM	50.5 nM
	r(AAG.DNA)	0.208	0.201	0.215
	r(AAG.AAG.DNA)	0.242	0.241	0.241
Total Fluorescence (1-site model)				
Parameters		**100 nM DNA**	**300 nM DNA**	**Average**
guesses	*DNA*	100 nM fix	300 nM fix	100 nM fix
	Kd1	1 nM float	1 nM float	1 nM float
	r(DNA)	1 fix	1 fix	1 fix
	r(AAG.DNA)	0.70 float	0.71 float	0.70 float
fits	*Kd1*	9.1 nM	13.7 nM	7.0 nM
	r(AAG.DNA)	0.68	0.72	0.70

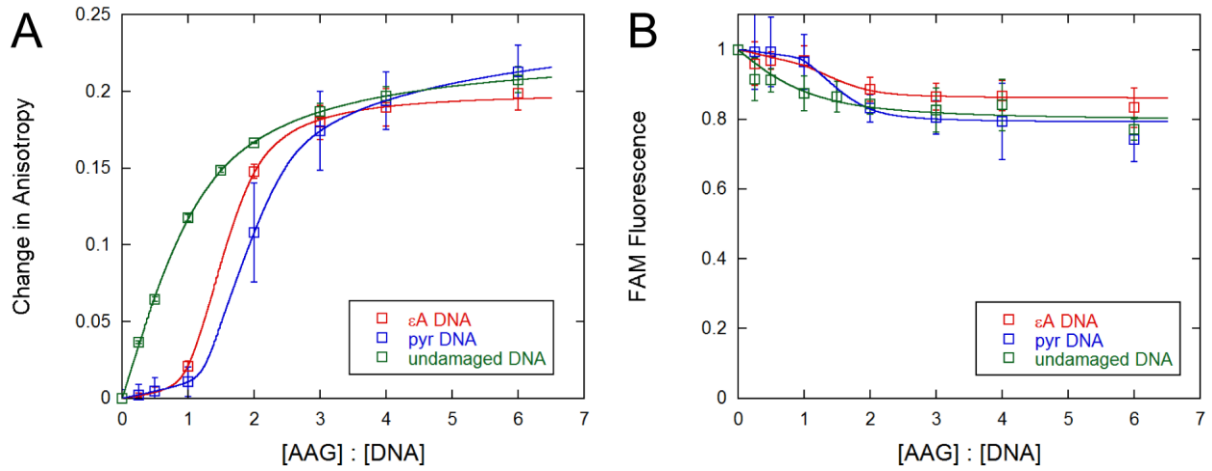


Figure B-7. Fluorescence measurements for binding of AAG to 25-mer specific and nonspecific DNA. (A) Anisotropy and (B) normalized total FAM fluorescence of 25-mer TXC substrates containing εA, pyrrolidine, or undamaged adenine sites. The pyr is 5′-labeled with a hexachlorofluorescein (HEX) while the εA and undamaged substrates are labeled with a 5′FAM.

	11-mer anisotropy		11-mer total fluor		19-mer anisotropy		19-mer total fluor	
Parameters	*Guess*	*Fit*	*Guess*	*Fit*	*Guess*	*Fit*	*Guess*	*Fit*
Kd1	0.02	-	0.02	-	0.02	-	0.02	-
Kd2	3.5	-	3.5	-	1 ?	3.5	3.5	-
DNA	100 ?	99.3	100 ?	97.2	100 ?	106.4	100 ?	99.5
r(DNA)	1.00E-20	-	1	-	1.00E-20	-	1	-
r(AAG.DNA)	0.050 ?	0.036	0.916 ?	0.917	0.025 ?	0.017	0.953 ?	0.953
r(AAG.AAG.DNA)	0.200 ?	0.202	0.85 ?	0.848	0.196 ?	0.190	0.71 ?	0.712

	25-mer anisotropy		25-mer total fluor		49-mer anisotropy		49-mer total fluor	
Parameters	*Guess*	*Fit*	*Guess*	*Fit*	*Guess*	*Fit*	*Guess*	*Fit*
Kd1	0.02	-	0.02	-	0.02	-	0.02	-
Kd2	1 ?	1.634	1.6	7.2	2	-	2	-
Kd3	100 ?	55.318	-	-	100 ?	68.7	100 ?	79.4
DNA	100 ?	97.87	100	-	100	-	100	-
r(DNA)	1.00E-20	-	1	-	1.00E-20	-	1	-
r(AAG.DNA)	0.0209 ?	0.0056	0.96 ?	0.956	0.0123 ?	0.00783	1 ?	1.00
r(AAG.AAG.DNA)	0.148 ?	0.168	0.86	-	0.0682 ?	0.065	1.04 ?	1.03
r(AAG.AAG.AAG.DNA)	0.2	-	-	-	0.21	-	1.23	-

Figure B-8. DynaFit guesses and fits for different sized 5′FAM-labeled εA substrates. 11-mer, 19-mer, 25-mer and 49-mer data as shown in Figure 4-5 (all fit for 100 nM DNA and normalized change in anisotropy) were fit by DynaFit. All K_d and DNA concentrations are in nanomolar and parameters were either fixed or floated (?).

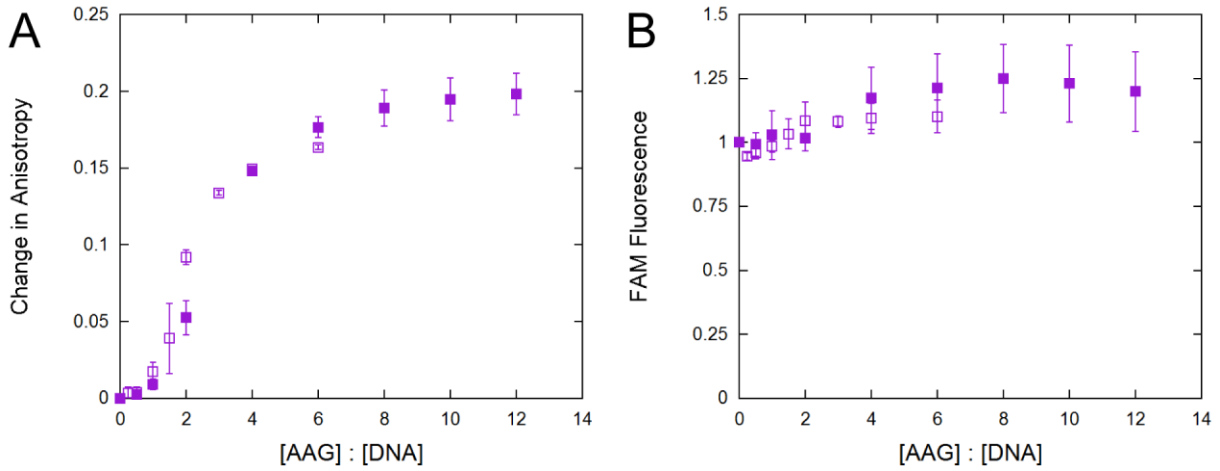


Figure B-9. Fluorescence measurements for 5′FAM-labeled 49-mer TEC DNA. (A) Anisotropy was measured with 100 and 300 nM DNA under the standard conditions (25 ^{o}C, 100 mM NaCl). (B) Total fluorescence was measured in the same samples. Data were plotted as the ratio of AAG to DNA to demonstrate that titration conditions were achieved (K_d < 100 nM). Mean ± SD (n=2).

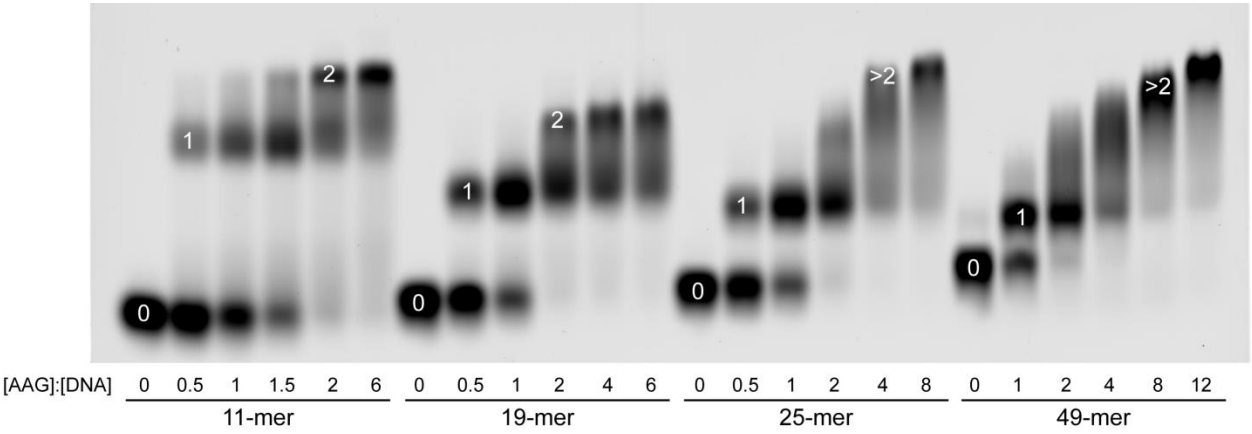


Figure B-10. Native gel-shift assays with 5′FAM-labeled εA-DNA of various sizes. AAG shifts 300 nM 11-mer, 19-mer, 25-mer, and 49-mer TEC DNA. All DNAs have an εA lesion and are 5′FAM labeled. Shown in white is the number of AAG molecules bound to each DNA. Only a small distinction is made between substrates without AAG due to the low concentration of acrylamide used. A larger change is observed between 0 and 1 AAG bound species for the smaller substrates due to the larger effect of a single AAG on the mass/charge ratio. Specifically, the 11-mer has a higher percentage of backbone phosphates "neutralized" by a single AAG molecule compared to the 49-mer.

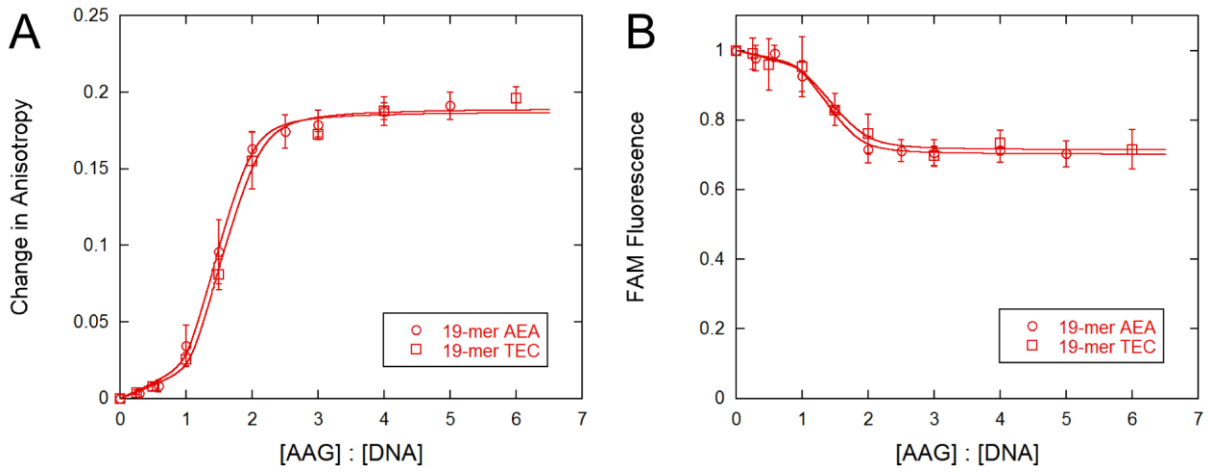


Figure B-11. Comparing the AEA and TEC sequences of 5′FAM-labeled εA-DNA under titration conditions. Anisotropy (A) and normalized total FAM fluorescence (B) of 19-mer εA DNA. Anisotropy and total fluorescence data were fit with a 2-site binding model in DynaFit (K_{d1} = 20 pM, K_{d2} ≈ 3 nM). The average of two concentrations ± SD are shown (n ≥ 4). These data show that a change in sequence context around the εA lesion does not alter the anisotropy and FAM fluorescence signals.

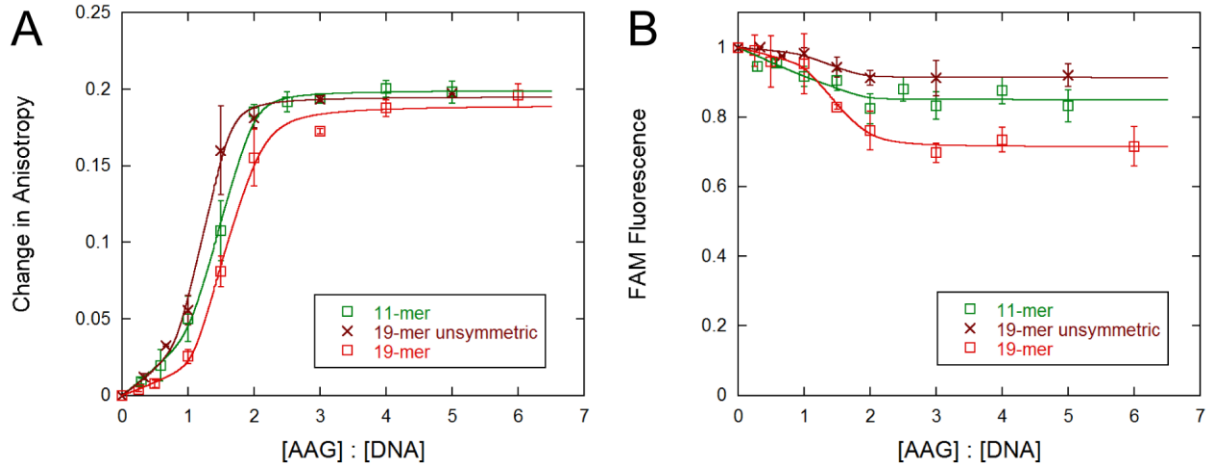


Figure B-12. Comparing substrates with symmetric and unsymmetric εA lesions. Anisotropy (A) and normalized total FAM fluorescence (B) of 11-mer, 19-mer and unsymmetric 19-mer 5′FAM labeled TEC εA DNA. The unsymmetric 19-mer and 11-mer are more similar than the two 19-mers because the εA is located the same distance away from the 5′FAM end (Figure B-1). The anisotropy and total fluorescence data of these substrates were fit with a 2-site binding model using DynaFit (K_{d1} = 20 pM, K_{d2} ≈ 3 nM). The FAM is quenched to a greater extent in the symmetric 19-mer substrate due to the sequence of the 5′end as described in Figure 4-7. The average ± SD is shown (n ≥ 2).

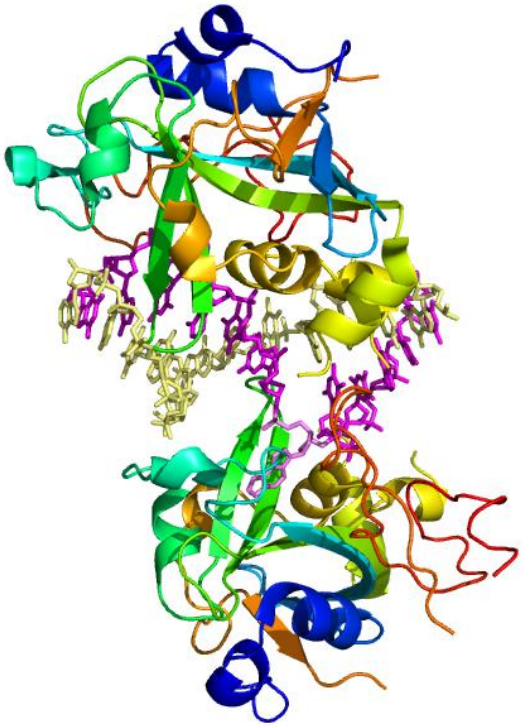


Figure B-13. Model suggesting that two AAG molecules can bind to an 11-mer εA-DNA. The crystal structure of AAG bound to εA-DNA (εA in light pink; bottom AAG; PDB 1EWN)[3] shows AAG flipping the εA into its active site for a specific binding event. The AAG molecule was duplicated and manually modeled on a nonspecific binding site upstream from the εA site (top AAG). The DNA substrate is a 13-mer with 1 bp overhangs which is close in upstream structure to the 11-mer hairpin DNA used in anisotropy assays (Figure 4-5). This crude model suggests that two AAG molecules can easily bind a short DNA duplex; one at the εA site and one upstream (near the 5′FAM in our assays).

	eA anisotropy		undamaged anisotropy		eA total fluor		undamaged total fluor	
Parameter	*Guess*	*Fit*	*Guess*	*Fit*	*Guess*	*Fit*	*Guess*	*Fit*
DNA	100 ?	121.4	100 ?	111.2	121.4	-	100	-
Kd1	0.02	-	6.5	-	0.02	-	6.5 ?	7.35
Kd2	10 ?	13.0	100 ?	58.2	10 ?	12.98	-	-
r(DNA)	0.24128	-	0.24723	-	1	-	1	-
r(AAG.DNA)	0.275 ?	0.285	0.29180 ?	0.305	1.117 ?	1.086	1.24 ?	1.24
r(AAG.AAG.DNA)	0.324	-	0.33	-	1.28	-	-	-

Figure B-14. DynaFit guesses and fits for different sized 5′DY647-labeled substrates. 19-mer TEC and TAC DNA data as shown in Figure 4-6 (all fit for 100 nM DNA and raw anisotropy) were fit by DynaFit. All K_d and DNA concentrations are in nanomolar and parameters were either fixed or floated (?).

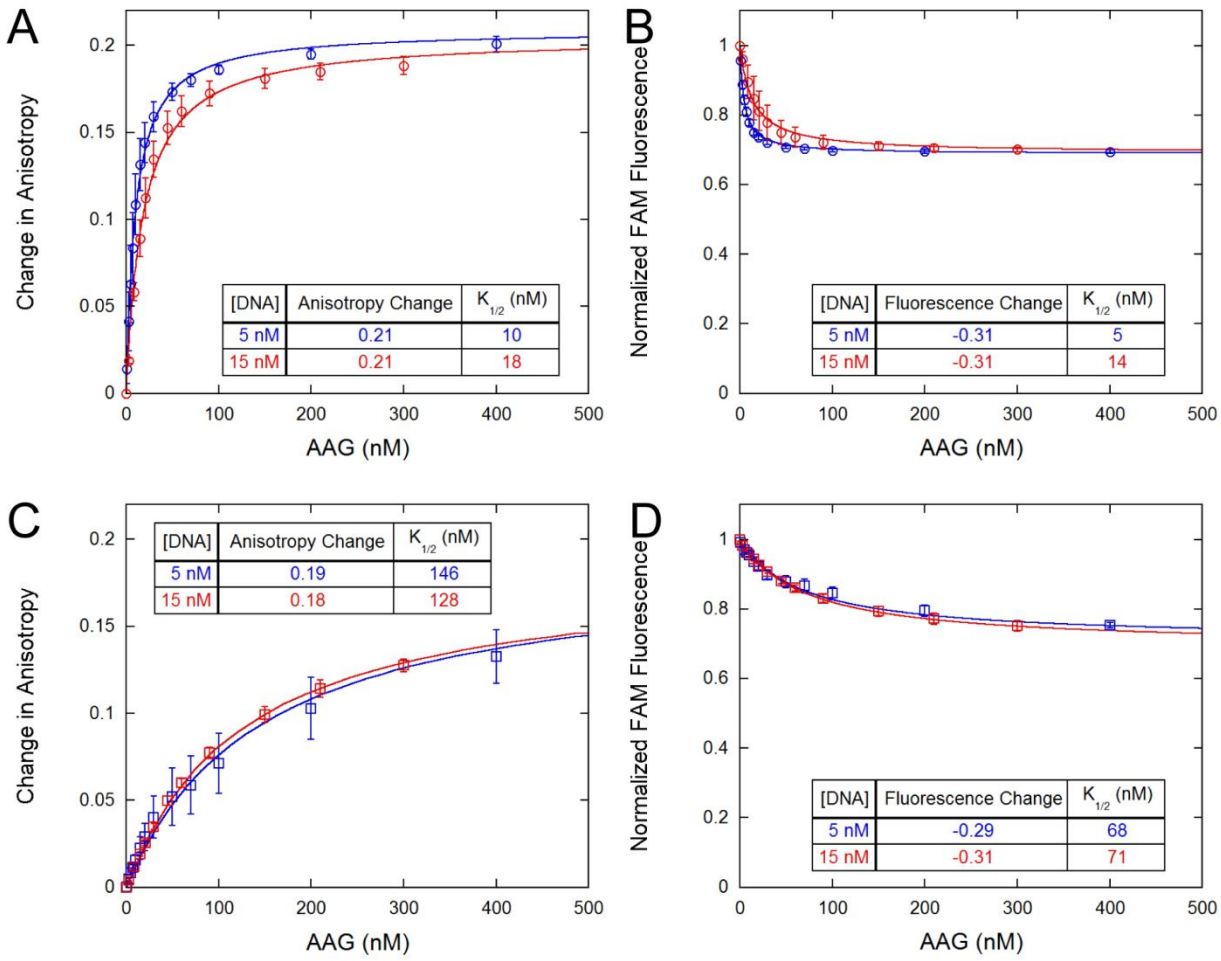


Figure B-15. Equilibrium binding of AAG to 5′FAM-labeled undamaged DNA. Anisotropy (A,C) and normalized total FAM fluorescence (B,D) for AAG binding to 5 nM and 15 nM 5′FAM 19 AAA undamaged DNA. 100 mM ionic strength (A,B) and 200 mM ionic strength (C,D) conditions were tested. Only the 200 mM ionic strength condition shows true equilibrium binding. All data were fit to a single site Michaelis-Menten binding equation. The raw anisotropy values are shown. The average ± SD is shown ($n \geq 2$).

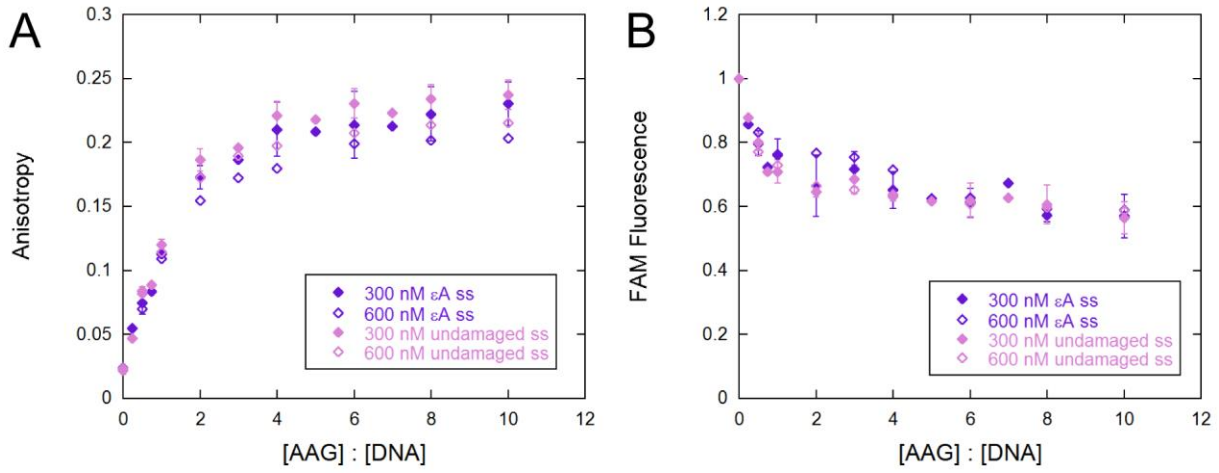


Figure B-16. Titration of ssDNA with AAG. 300 and 600 nM 19-mer 5′FAM-labeled AEA or AAA DNA was incubated with increasing amounts of AAG. Titrations were measured at 25 ^{o}C with the standard binding buffer and 100 mM NaCl. The data is plotted versus the ratio of AAG to DNA to demonstrate that titration conditions were achieved (K_d << 300 nM). The mean ± SD is plotted (n ≥ 2).

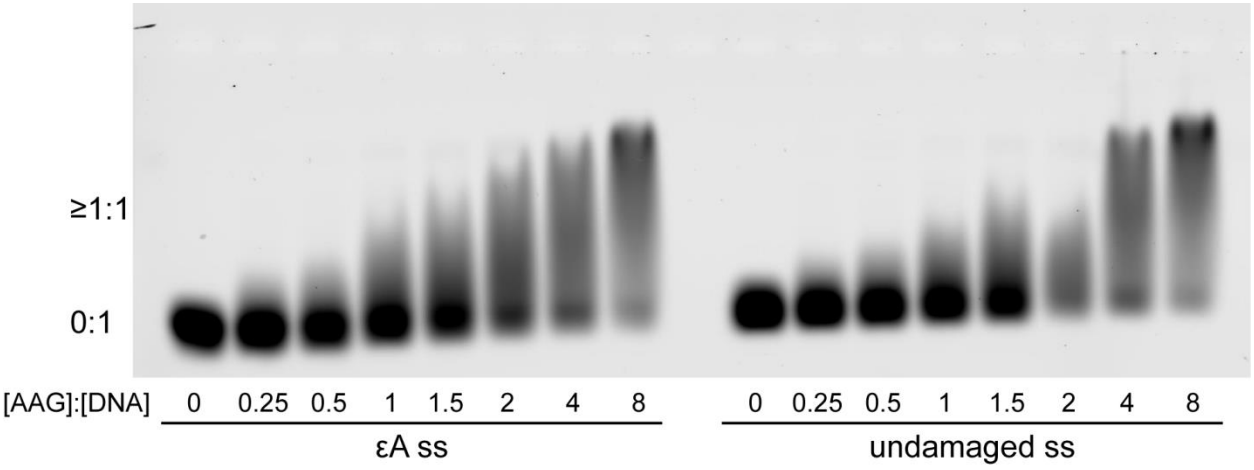


Figure B-17. Native gel-shifts with single stranded DNA under titration conditions. 300 nM 19-mer 5′FAM-labeled AEA or AAA DNA was incubated with increasing amounts of AAG.

References

1. Hendershot, J. M., and O'Brien, P. J. (2014) Critical role of DNA intercalation in enzyme-catalyzed nucleotide flipping, *Nucleic Acids Res*.
2. Kuzmic, P. (1996) Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase, *Analytical biochemistry 237*, 260-273.
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, *Proceedings of the National Academy of Sciences of the United States of America 97*, 13573-13578.

Chapter V

Investigation of the Roles of *Bacillus subtilis* AlkA and AAG Glycosylases

Abstract

DNA glycosylases excise damaged nucleobases and initiate base excision repair to protect genomic DNA in all forms of life. The *E. coli* AlkA and human AAG glycosylases have independently evolved to remove an extensive set of alkylated base lesions and are almost always mutually exclusive in their expression. *Bacillus subtilis* is unique in that it contains both AlkA and AAG homologs. Here we have investigated the roles of bAlkA and bAAG in alkylation repair by assessing MMS sensitivities of bAlkA and bAAG deletion strains, and performing in vitro kinetics studies with a variety of substrates. Results suggest that bAlkA is necessary for MMS resistance and may have specificity for cytotoxic lesions. In contrast, bAAG is more efficient at the removal of mutagenic etheno lesions. bAlkA and bAAG also show activity on canonical substrates of AlkB, a direct repair enzyme that is lacking in *B. subtilis*. This is the first work to directly compare bAlkA and bAAG and lays the groundwork for future experimentation of *B. subtilis* glycosylases.

This work was done in collaboration with Lyle Simmons in the Molecular, Cellular and Developmental Biology Department at the University of Michigan.

Introduction

Alkylating agents are produced endogenously and multiple sources are found in the environment. These sources damage DNA, commonly via direct modification of the nucleobases. Alkylated bases are commonly mutagenic or cytotoxic by interfering with DNA templated activities such as replication and transcription. The constant assault on cells by alkylating agents requires quick and effective pathways designed to reverse and repair the damage. The direct reversal and base excision repair pathways repair the bulk of alkylation damage in cells.[1]

In *Escherichia coli*, an adaptive (ada) response has evolved to respond to elevated levels of alkylating agents in the environment. When *E. coli* are exposed to low doses of alkylating agents such as N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) and methyl methanesulfonate (MMS), they are able to subsequently survive at high doses of the drugs that would otherwise be fatal.[2] The source of this adaptation is the direct reversal protein, Ada, which demethylates O^6-methylguanines (O^6meG) and methylphosphotriesters on the DNA backbone by transferring the methyl groups to active site cysteine residues.[3-6] When Ada demethylates methylphosphotriesters, the ensuing methylation of Cys38 causes a proposed conformational change in the enzyme and activates Ada as a transcription factor. Via interactions with RNA polymerase, Ada induces expression of the ada regulon which includes the combined ada and alkB operon as well as the alkA and aidB genes.[7,8] The protein products Ada, AlkB, AlkA and AidB are all upregulated and contribute to alkylation repair.[1,9]

Ada, AlkB and AlkA repair all the possible *N*- and *O*-methylated sites that occur on DNA. As described above, Ada acts on O^6meG and methylphosphotriesters.[6] AlkB is another

direct reversal enzyme that primarily acts on N^1-methyladenine (1meA) and N^3-methylcytosine (3meC) lesions that are formed in single stranded DNA.[10,11] AlkB also acts as the major repair protein of 1,N^6-ethenoadenine (εA), a cyclic adduct produced endogenously from reactions with lipid peroxidation byproducts.[12] AlkA is a glycosylase that initiates the base excision repair pathway by hydrolyzing the *N*-glycosidic bond between the deoxyribose and damaged base. AlkA has a broad substrate range that includes *N*-methylated purines (3meA, 7meG),[13,14] *O*-methylated pyrimidines (2meT, 2meC),[15] εA,[16] and oxidatively damaged lesions (hypoxanthine (Hx), oxanine, xanthosine, and 5-formyluracil).[17-19] AidB does not have a known function but does bind DNA and may deactivate alkylating agents with its flavin moiety, thus protecting cells from damage.[20,21]

While the inducible ada response is only observed in bacteria, homologs of the repair proteins are also found in archaea and eukaryotes.[22-26] AlkA homologs are found in both budding and fission yeast and multiple AlkB homologs have been studied in mammalian systems.[25-28] While there is no AlkA homolog present in higher eukaryotes, alkyladenine DNA glycosylase (AAG) has independently evolved to repair a similar broad range of alkylated and oxidative lesions in plants and animals.[16,17,29] AAG and AlkA are structurally distinct but perform the same chemistry.[30] AAG and AlkA are almost always mutually exclusive in their expression, with most organisms containing only AlkA or AAG homologs.

The Gram-positive soil bacterium, *Bacillus subtilis*, showcases an ada response similar to *E. coli*. In *E. coli*, a single Ada protein has two distinct domains to demethylate methylphosphotriesters and O^6meG lesions.[1] In *B. subtilis* these functions are performed by the two proteins, AdaA and AdaB, respectively.[31] AdaA has been shown to be a transcriptional activator and promotes the ada response during alkylation stress. An AlkA homolog (bAlkA) is

regulated by AdaA in *B. subtilis* and has been shown to protect cells from propylated adducts when induced.[31,32]

Unlike *E. coli*, *B. subtilis* contains another putative AlkA homolog (YfjP), an AAG homolog (YxlJ; hereafter referred to as bAAG) and a proposed homolog of the HEAT-repeat alkyl repair glycosylase AlkC from *Bacillus cereus* (YhaZ).[33-35] The numerous alkylation repair proteins present in *B. subtilis* may be evolutionary products of the high levels of alkylating agents naturally found in a soil environment. Alignments of the *B. subtilis* AlkA (bAlkA) and YfjP with the *E. coli* AlkA show some similarity and conservation of the catalytic and active site residues (Figure 5-1). The *E. coli* AlkA is 27% identical to YfjP and the two *B. subtilis* proteins are 44% identical. The bAlkA is also 53% identical to the *Bacillus cereus* AlkA and around 25% identical to AlkA homologs in *Bacillus halodurans*, *Archaeglobus flugidas*, and *Saccharomyces cerevisiae*. While the YfjP protein seems to be closest in homology to the *E. coli* AlkA, we are focusing this research on the bAlkA protein as it has been shown to protect *B. subtilis* from alkylation damage and is part of the conserved ada response. YfjP may play a role as a constitutively expressed protein that protects cells from alkylation damage, similar to Tag, an *E. coli* protein that has narrow specificity for 3meA and 3meG lesions only.

As mentioned above, *B. subtilis* also surprisingly contains an AAG homolog, YxlJ (bAAG). The alignment of bAAG and the human AAG proteins show 33% identity (Figure 5-2). Previous work has also shown that bAAG partially rescues MMS resistance in *E. coli* lacking Tag and AlkA proteins. bAAG acts on Hx, εA, 3meA, and 3meG, but in contrast to the human enzyme, has not been shown to excise 7meG lesions.[34]

We want to understand the roles of these alkylation repair glycosylases in *B. subtilis.* Uniquely, *B. subtilis* contains both AlkA and AAG homologs. The canonical *E. coli* AlkA and

Figure 5-1. Alignment of *E. coli* and *B. subtilis* AlkA proteins. Clustal Omega alignment of full length AlkA from *E. coli* and two of the proposed AlkA homologs in *B. subtilis* (bAlkA and YfjP). The catalytic aspartic acid 238 is highlighted in purple, the intercalating residue in blue and active site aromatic residues in yellow. The helix-hairpin-helix domain in the *E. coli* AlkA is highlighted in gray. The *E. coli* AlkA is 27% identical to the *B. subtilis* YfjP and the two *B. subtilis* proteins are 44% identical. The *B. subtilis* AlkA protein is also around 25% identical to AlkA homologs in *Bacillus halodurans*, *Archaeglobus flugidas*, and *Saccharomyces cerevisiae*, and 53% identical to the *Bacillus cereus* AlkA (not shown).

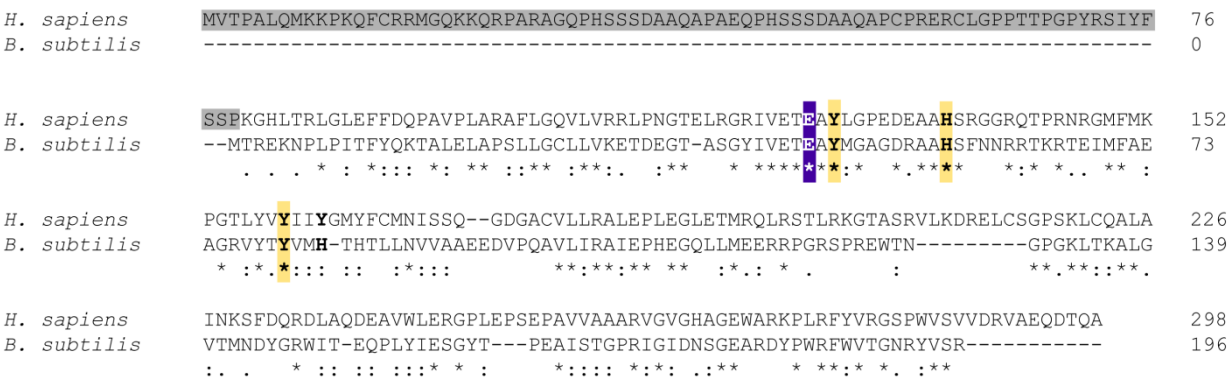


Figure 5-2. Alignment of human and *B. subtilis* AAG proteins. Clustal Omega alignment of full length AAG from humans and *B. subtilis* AAG homolog (YxlJ; bAAG) showing 33% identity. The catalytic glutamic acid 125 is highlighted in purple, the intercalating residue in blue and active site aromatic residues in yellow. The human AAG N-terminus domain involved in DNA binding is highlighted in gray and is not necessary for catalysis.

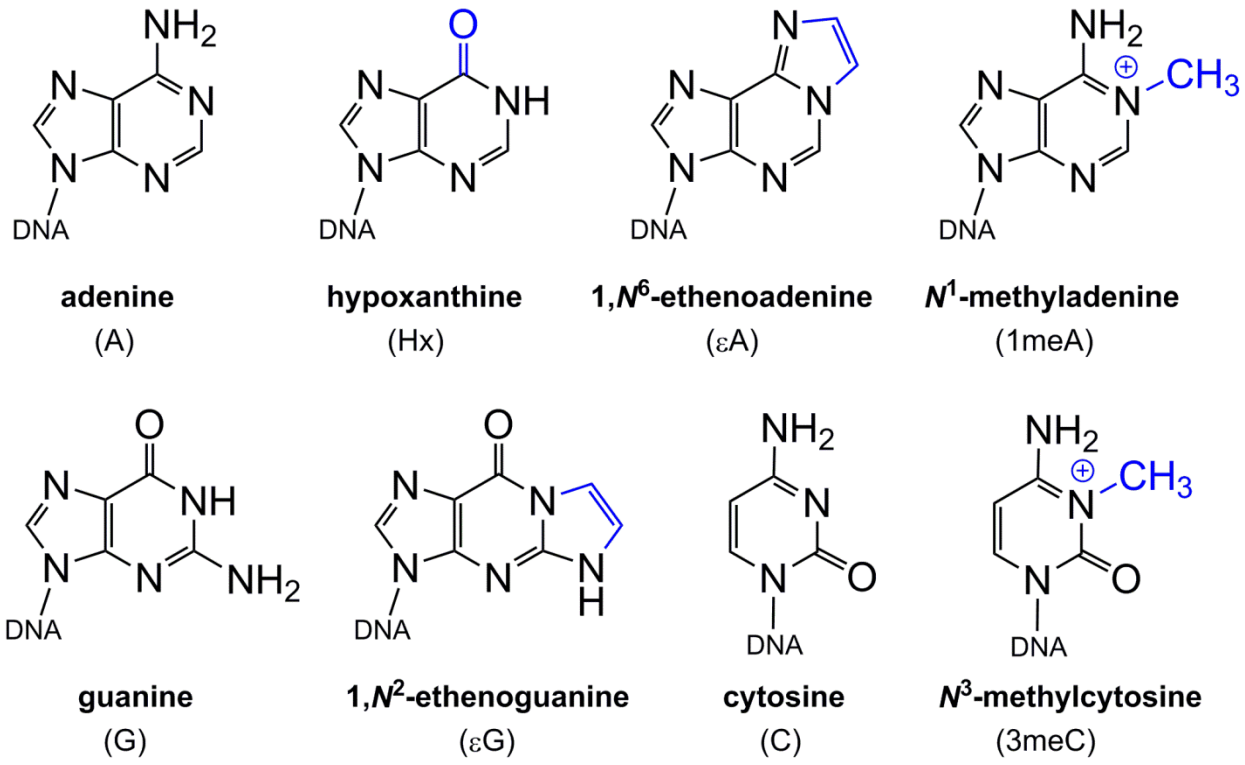


Figure 5-3. Normal and damaged nucleotides tested for bAlkA and bAAG activity. Modifications to the undamaged bases are shown in blue.

mammalian AAG proteins have independently evolved to repair similar alkylation and oxidative damage. Why *B. subtilis* has evolved to contain both glycosylases is not understood. One model is that the two proteins are redundant and both may not have to be present for cells to be protected against alkylating agents prevalent in a soil environment. Another model is that bAlkA and bAAG have evolved to act on different substrate ranges than their *E. coli* and human counterparts. We wish to test the importance of these genes with cell survival assays and in vitro kinetic studies using a variety of substrates. As no AlkB homologs have been identified in the *B. subtilis* genome, we hypothesize that the substrate ranges of either bAlkA or bAAG have expanded to include the AlkB substrates, 1meA and 3meC. We also tested the etheno substrates 1,N^2-ethenoguanine (εG) and 1,N^6-ethenoadenine (εA) as well as the deaminated hypoxanthine (Hx), because these are common forms of damage that are repaired by base excision repair or

direct repair pathways (Figure 5-3).[16,17,36-41]

This work is preliminary and looks to compare these two alkylation repair glycosylases in vivo and in vitro. Initial experiments suggest that bAlkA and bAAG are necessary for alkylation damage repair in *B. subtilis* as deletion of the genes renders cells sensitive to MMS. Their substrate ranges do overlap, although their catalytic efficiencies for the substrates vary. Both bAlkA and bAAG act on Hx, εA, 1meA, and εG, and bAlkA shows some activity on 3meC. It should be noted that the substrate specificities of YfjP and YhaZ have not been studied and could be unique or overlapping with bAlkA and bAAG. These studies imply that bAlkA and bAAG have evolved to remove similar substrates but the enzymes may play different predominant roles in the different lesion repair in cells.

Materials and Methods

Design and Assembly of pMiniMAD2 Plasmid for bAlkA and bAAG Deletions

Primers were designed for the 500 bp regions upstream and downstream of *alkA* and *yxlJ* (AAG) genes to be inserted into the pMiniMAD2 plasmid vector (Table 5-1) as described previously.[42] It should be noted that both genes are in the reverse orientation to the *Bacillus subtilis* genome map. However, for the design of the plasmid, the upstream and downstream regions are denoted as if the genes were in the forward orientation. Primers 1 and 2 overlap the pMiniMAD2 vector and the 5′-end of the 500 bp downstream section (reverse) and upstream section (forward) respectively. Primers 3 and 4 overlap the upstream and downstream sections in the reverse and forward directions, respectively (Figure 5-4).

The upstream and downstream 500 bp regions for *alkA* and *yxlJ* were PCR amplified from the PY79 *B. subtilis* strain using Q5 DNA polymerase (20 ng genomic DNA, 0.4 µM primers, 0.24 mM dNTPs, 1x Q5 Pol Reaction Buffer (NEB) and 0.04 units/µL Q5 DNA polymerase (NEB)). The pMiniMAD2 vector was also amplified using vector DNA and primers described in Table 5-1 and Figure 5-4. The size of the upstream and downstream regions (inserts)

Table 5-1. Primers designed for *alkA* and *yxlJ* deletions.

Gene	Primer	Sequence	direction	length	melting T (°C)	length (whole)	melting T (whole) (°C)
alkA	1A	5' CACAGATGCGTAAGGAGAAAATACCGCTTGTTCTGTGTTTGGAAAGATACATACATTTTC 3'	reverse	35	61	60	70
	2A	5' CCCAGGCTTTACACTTTATGCTTCCCGTAACCTTGCGAAGAGTGTTACTGTG 3'	forward	27	61	52	72
	3A	5' GGAGAATGAAAAAATCATTTCAACTTTAACCGGAATGTACTCAATAAAAAAAG 3'	reverse	40	61	53	64
	4A	5' GTTAAAGTTGAAATGATTTTTTCATTCTCCTTATCTTAAATAATTATAAGAAAACTCACG 3'	forward	43	59	60	63
yxlJ	1Y	5' CACAGATGCGTAAGGAGAAAATACCGACAAATGCCGGTGTTTTCTTGCC 3'	reverse	24	61	49	72
	2Y	5' CCCAGGCTTTACACTTTATGCTTCCCGTGGCCGCTTAGCACACC 3'	forward	19	62	44	75
	3Y	5' GAGAGGTCGATTCCATGTTTAAGGAAGAGGCAATCAGGGG 3'	reverse	27	61	40	69
	4Y	5' CTTCCTTAAACATGGAATCGACCTCTCCTTTTACCGATTTTTG 3'	forward	29	60	43	66

Red sequences denote change from one section (e.g., pMiniMAD2) to another (e.g., downstream region). Primers were designed so that the red sequence portions would have melting temperatures of 60 °C and the entire length would be a maximum of 60 bases.

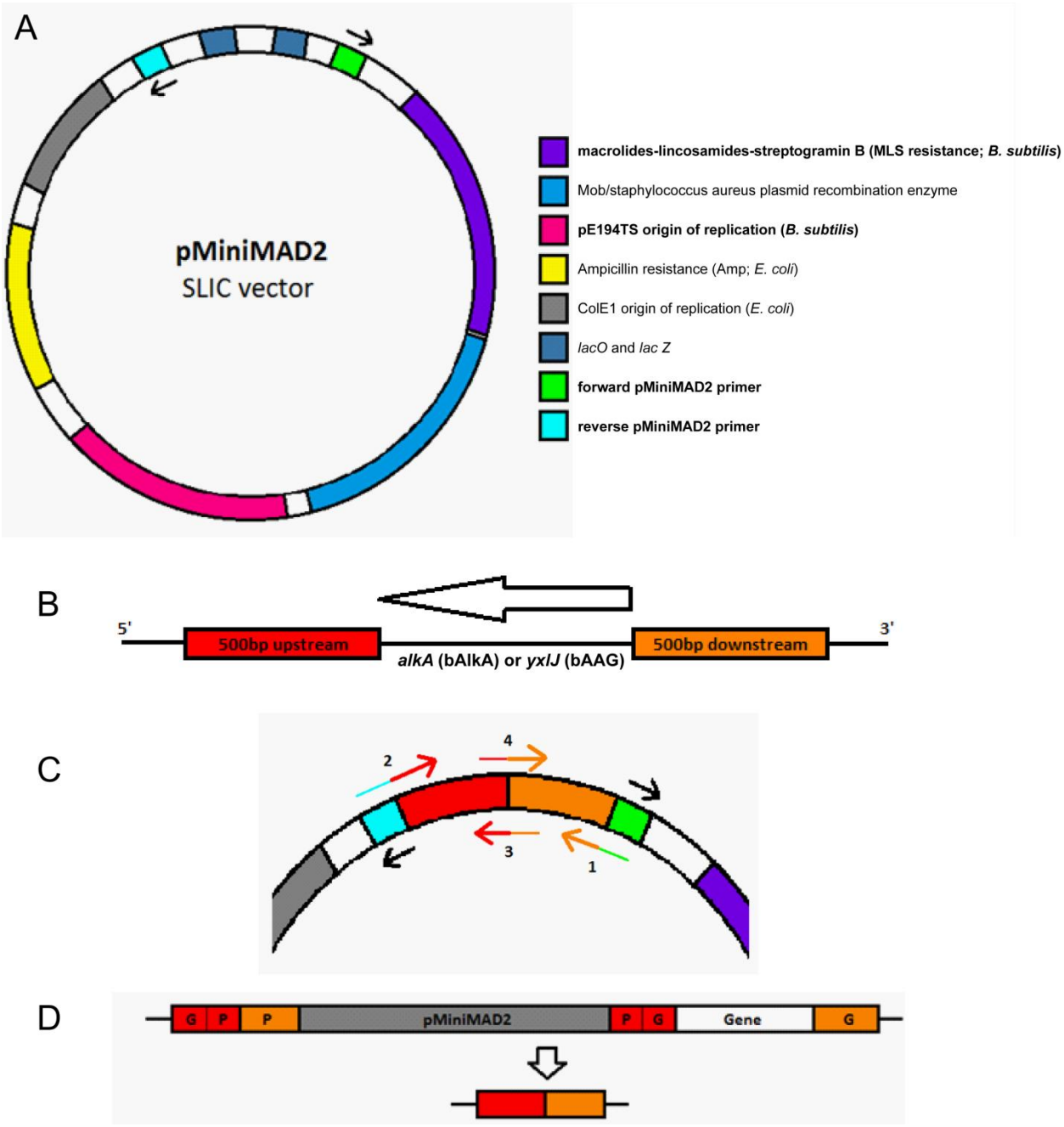


Figure 5-4. pMiniMAD2 plasmid design for *alkA* and *yxlJ* deletions. (A) The pMiniMAD2 vector is designed for transformation in both *E. coli* and *B. subtilis*. (B) General map of *alkA* and *yxlJ* gene region in *B. subtilis*. The genes are oriented in the reverse direction while the upstream and downstream 500 bp regions are in the forward direction. (C) Primer design for incorporation of upstream and downstream regions (inserts) into pMiniMAD2 vector. (D) Incorporation of pMiniMAD2 plasmid into genome and ejection of plasmid from genome that occur during back dilutions of *B. subtilis* as described in Methods.

and vector was verified via agarose gel electrophoresis, and the PCR products were purified with the QIAquick PCR Purification Kit (Qiagen). In short, 5x volume of Buffer PB was added to the PCR reactions and centrifuged in QIAquick spin columns. 750 µL of Buffer PE was added to the column and centrifuged. 50 µL ddH_2O was used to elute the DNA.

The pMiniMAD2 plasmid was assembled with the inserts of either *alkA* or *yxlJ* using the Gibson Assembly Protocol (NEB). In short, 40 ng of vector was added to 3-fold amounts of each insert and 1x Gibson Master mix (2x Master mix includes 2x isothermal reaction buffer (5x contains 500 mM Tris-Cl pH 7.5, 50 mM $MgCl_2$, 1 mM dNTPs, 3.1 mM PEG-8000, 5 mM NAD, 50 mM DTT), 0.008 units/µL T_5 exonuclease, 0.05 units/µL Phusion DNA polymerase, 8 units/µL Taq DNA ligase) and incubated at 50 °C for 1hr.

The Gibson reactions described above were immediately used to transform chemically competent MC1061 *E. coli* cells (plated on LB + ampicillin). 5 mL cultures of triplicate transformants (LB + amp) were incubated overnight at 37 °C. Cells were centrifuged and plasmid was extracted with the QIAprep Miniprep kit (Qiagen). In short, the pellet was resuspended in 250 µL Buffer P1 and then vigorously mixed with 250 µL Buffer P2. After incubation at RT for 5 min, 350 µL Buffer N3 was added and mixed vigorously. The reaction was incubated on ice for 10 min and centrifuged. The supernatant was cleaned with the QIAquick PCR Purification kit (Qiagen). The cleaned pMiniMAD2 plasmids were sequenced to verify incorporation of both upstream and downstream regions of *alkA* and *yxlJ* genes.

Deletion of bAlkA and bAAG

PY79 *B. subtilis* cells were made chemically competent by inoculating 3 mL LM media (LB, 3 mM $MgSO_4$) and growing at 37 °C until OD reaches 1. 150 µL of the culture was used to

inoculate 3 mL MD media (1x phosphate-citrate buffer (10x contains 107 g K_2HPO_4, 60 g KH_2PO_4, 10 g sodium citrate•2 H_2O), 2% glucose, 50 μg/mL L-tryptophan, 50 μg/mL phenylalanine, 11 μg/mL ferric ammonium citrate, 2.5 mg/mL potassium aspartate, 3 mM $MgSO_4$) and incubated at 37 °C for 5 hrs.

200 μL of the chemically competent PY79 culture was transformed with 1 μg of pMiniMAD2 plasmids described above, plated onto LB + MLS (macrolides-lincosamides-streptogramin B), and incubated at 37 °C overnight. The *B. subtilis* transformants were back diluted to allow for incorporation of the gene into the plasmid followed by ejection of the plasmid from the cells. In short, triplicate colonies from the transformations were used to inoculate 3 mL LB media. Cultures were incubated for 2 hrs at 37 °C then 3 hrs at RT. 10-fold dilutions of the cultures in LB media were made and incubated at RT for 12 hrs. The 10-fold dilutions were repeated twice at RT (12 hrs each), followed by a 10-fold dilution incubated at 37 °C until an OD of 1.1 was reached. A $1:10^5$ dilution of the cultures were made in 0.85% saline, plated onto LB and incubated at 37 °C overnight. Twelve colonies from each LB plate were streaked onto LB and LB + MLS plates which were incubated at 37 °C overnight. Colonies that are MLS sensitive were screened with PCR to verify target gene deletions. In short, overnight cultures of each colony were grown in LB at RT. A small volume (20 μL) of each culture was incubated at 4 °C for 3 hrs to lyse the cells. 10 μL PCR reactions (1x GoTaq Green Mastermix (Promega), 1 μM primers (1 and 2), 0.5 μL cell culture) were performed and 1 kB PCR products analyzed with agarose gel electrophoresis.

Single *alkA* and *yxlJ* deletions were produced using the protocol described above. A double *alkA*/*yxlJ* deletion was also produced by transforming on *alkA* deletion PY79 strain with the pMiniMAD2 plasmid containing upstream and downstream *yxlJ* inserts. All following steps

are identical to the single dilution protocol. All single and double deletions were verified for a second time with PCR (Q5 DNA polymerase reaction (NEB), Primers 1 and 2), cleaned and sequenced. After successful deletions, cells were maintained by restreaking everyday onto LB and incubating at RT.

MMS Sensitivity of alkA and yxlJ Deletion Strains

LB agar plates (50 mL volume) were poured and with 0, 50, 100, or 200 μg/mL methyl methanesulfonate (MMS). Two colonies each of WT PY79, *alkA* and *yxlJ* single deletion strains and the *alkA*/*yxlJ* double deletion strain were grown to exponential phase (OD between 0.5 and 0.8) at 37 °C. Cultures were normalized to lowest OD sample via dilutions with 0.85% saline. 20 μL of normalized cell samples were placed in first and seventh (duplicate) column of 96-well plate. Using a Precision X2 Spot Plating Robot and Precision Power V2 Software (Biokin), 10x serial dilutions (200 μL volumes) were made for columns 2-6 and 8-12. 5 μL of each dilution was spotted onto LB + MMS plates. Plates were incubated at 37 °C and imaged at 16, 18, 20, and 24 hrs. Both colonies of each strain were plated in duplicate.

Expression and Purification of bAlkA and bAAG Proteins

Primers (5 and 6) were designed to amplify and clone the *alkA* or *yxlJ* genes into a 6xHis-SUMO vector with kanamycin (kan) resistance (Table 5-2). The 6xHis-SUMO vector and inserts (*alkA* and *yxlJ* genes) were PCR amplified with Q5 DNA polymerase, verified with gel electrophoresis, cleaned, and assembled with the Gibson Assembly Protocol as described above. The 6xHis-SUMO region is directly connected to the 5' end of the gene region (N-terminus). The Gibson reactions were used to transform *E. coli* MC1061 cells (plated onto LB + kan), and

Table 5-2. Primers designed for cloning of *alkA* and *yxlJ*.

Gene	Primer	Sequence	direction	length	melting T (°C)	length (whole)	melting T (whole) (°C)
alkA	5A	5' GGCTCACCGCGAACAGATTGGAGGTGTGACATGGCATGAAGTCAATGATGTTATTG 3'	forward	31	61	56	74
	6A	5' CAGTGGTGGTGGTGGTGGTGCTCGATCAGTAAAGGACACGCCATAAATAAAAGGTAG 3'	reverse	32	61	57	74
yxlJ	5Y	5' CTCACCGCGAACAGATTGGAGGTGTGACGAGAGAAAAAAATCCGCTGCC 3'	forward	26	62	49	74
	6Y	5' GGTGGTGGTGGTGCTCGACTAACGCGACACGTACCTGTTTCC 3'	reverse	24	61	42	75

Red sequences denote change from one section (e.g., 6xHis-SUMO vector) to another (gene). Primers were designed so that the red sequence portions would have melting temperatures of 60 °C and the entire length would be a maximum of 60 bases.

the plasmids were purified with the QIAprep Miniprep kit as described above. To verify successful gene insertion the clean PCR products were sequenced and colony PCR was performed on the successful transformants using the Green Mastermix Protocol (with primers 5 and 6) as described above excluding the cell lysing step.

BL21 (DE3) *E. coli* cells were transformed with the PCR-clean 6xHis-SUMO-*alkA* or –*yxlJ* plasmids, plated on LB + kan (25 μg/mL) and incubated at 37 °C overnight. Cell cultures were grown in LB + kan at 37 °C until an OD of 0.7 was reached. 200 μM IPTG was added to each culture to induce protein expression and cultures were incubated for an additional 3 hrs (OD values around 1.5). Cells were centrifuged for 30 min at 4 °C and pellets were frozen at -80 °C for future purification.

Pellets were resuspended in lysis buffer (50 mM Tris-Cl pH 8, 10% D-sucrose (w/v), 300 mM NaCl, 20 mM spermidine hydrochloride) and cells were lysed with French press (3 runs). Cell lysates were run over HisTrap FF Crude 5 mL column (GE Healthcare) equilibrated with 20 mM Tris-Cl pH 8, 2 M NaCl, 5% glycerol, 25 nM imidazole). An imidazole gradient was used to elute 6xHix-SUMO-bAlkA and 6xHis-SUMO-bAAG. SUMO protease was added to protein solutions, and reaction was dialyzed with protease buffer (50 mM Tris-Cl pH 8, 150 mM NaCl, 5% glycerol, 1 mM DTT) at 4 °C overnight. Dialyzed solutions were run through 2.5 mL Ni^{2+} NTA beads in a gravity drip column equilibrated with protease buffer. Fractions were collected from the protein solution flow through and concentrated down to 1 mL using Millipore

Centrifugal concentration tubes (30 kDa for bAlkA, 10 kDa for bAAG). The bAlkA protein was run through a Superdex Gel Filtration 120 mL column (GE Healthcare) equilibrated with protease buffer to remove higher mass protein contaminants. The bAAG protein was applied to a HiPrep 26/10 Desalting 50 mL column (GE Healthcare) equilibrated with Q binding buffer (20 mM Tris-Cl pH 8, 50 mM NaCl, 5% glycerol, 1 mM DTT). The bAAG protein was subsequently loaded on a HiTrapQ HP 5 mL anion exchange column (GE Healthcare) equilibrated with Q binding buffer. A NaCl gradient was used to elute the protein (50 mM to 1 M). Both bAlkA (36 kDa) and bAAG (22 kDa) proteins were concentrated down to 2 mL and flash frozen in liquid nitrogen for storage. Aliquots of the proteins were thawed and concentrated to approximately 50 μM, resuspended in glycosylase assay buffer (50 mM NaHEPES pH 7.5, 76 mM NaCl (100 mM ionic strength), 1 mM EDTA, 1 mM DTT) and reconcentrated to approximately 50 μM prior to experimentation using Amicon Ultra 0.5 mL 3K centrifugal filters (Millipore).

Preparation of Oligonucleotides

25-mer oligonucleotides with a central lesion and 5′ 6-fluorescein (FAM) (5′-CGATAGCATCCTXCCTTCTCTCCAT-3′) were purchased from the Keck Center at Yale University. The lesions studied were hypoxanthine (Hx), 1,N^6-ethenoadenine (εA), 1,N^2-ethenoguanine (εG), 3,N^4-ethenocytosine (εC), N^1-methyadenine (1meA), and N^3-methylcytosine (3meC). Unlabeled complementary oligonucleotides were purchased from Integrated DNA Technologies (IDT). Oligonucleotides were purified with denaturing PAGE, extracted, and desalted with reverse phase C18 columns (Sep-Pak, Waters). The concentrations of the Hx, εA, and unlabeled complement strands were determined from the A_{260} and the calculated extinction coefficient (subtracting 9400 $M^{-1}s^{-1}$ from an undamaged adenine containing oligonucleotide for

the εA extinction coefficient). For the εG, εC, 1meA, and 3meC oligonucleotides, the difference in extinction coefficient between the lesion and an undamaged base is unknown, therefore the UV absorbance of the 5′FAM (A_{493}) was used to determine the concentration. Upon investigation, no significant difference was observed between the A_{493} and A_{260} measurements for these substrates so extinction coefficients of the undamaged oligonucleotides could be used in the future.

Oligonucleotides were annealed with a 1.2-fold excess of complementary strand by heating to 95 °C for 3 min and cooling to 4 °C at a rate of 0.2 °C/s. Complementary strands contained the opposing base that would naturally occur across from each lesion. Specifically, εA, Hx and 1meA were paired with thymine, εC and 3meC were paired with guanine, and εG was paired with cytosine.

Single Turnover Glycosylase Activity Assays

bAlkA or bAAG was mixed in excess concentrations (50 nM to 5 μM) with 20 nM DNA substrate in reaction buffer (50 mM NaHEPES pH 7.5, 76 mM NaCl (100 mM controlled ionic strength), 1 mM EDTA, 1 mM DTT, 0.1 mg/mL BSA). A single 5 μM reaction was performed at pH 6.1 (50 mM NaMES). Mixtures were incubated at 25 °C and aliquots were quenched with 0.2 M NaOH at various time points. Samples were placed on ice immediately after quenching for up to 90 min prior to the next step. Samples were then heated to 70 °C for 12 min, centrifuged and mixed with loading buffer (98% formamide, 1 mM EDTA, bromophenol blue and xylene cyanol). 25 fmol of DNA was loaded for each sample onto denaturing PAGE (20% (w/v) acrylamide, 1× TBE, 6.6 M urea) and analyzed with a Typhoon Trio Fluorescence imager (GE Healthcare) using a 488 nm excitation and 520 nm band-pass filter to detect the FAM.

The fraction of product for each time point was calculated by dividing the intensity of the product band by the sum of intensities of both product and substrate bands in each lane. The fraction of product versus time plots were fit by a single exponential (Eq 1) where F is the fraction product, A is the amplitude, k_{obs} is the observed single turnover rate constant, t is the reaction time, and c is the amount of pre-existing abasic DNA.

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

The enzyme concentration dependence of the observed rate constants was fit by a hyperbolic dependence (Eq 2), where k_{max} is the maximal single turnover rate constant, E is bAlkA, and the $K_{1/2}$ is the concentration of bAlkA at which k_{obs} is 50% of the k_{max} value.

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

All single turnover glycosylase assays were performed once and the concentrations of bAlkA and bAAG were assumed to be 100% active. The εC substrate appeared to be incorrectly produced or damaged during purification as 65% was found to be abasic in the absence of enzyme and independent of incubation time (data not shown). Therefore the εC substrate was left out of our analysis. All other substrates were stable up to the 5 hrs tested. The fraction product end points did vary with each substrate (96% for Hx, 95% for εA, 86% for εG, 76% for 1meA, 63% for 3meC) indicating that other form (e.g., ring opening of etheno products) or contaminant is present in the mixture and does not react with bAlkA or bAAG.

Results and Discussion

Deletion of bAlkA and bAAG Genes and MMS Resistance

In order to study the roles of bAlkA and bAAG in alkylation repair, we deleted the *alkA* and *yxlJ* (hereafter named *aag*; bAAG) genes from the *B. subtilis* genome utilizing the pMiniMAD2 vector system first described by Patrick and Kearns (Figure 5-4).[42] Clean single deletion strains of *alkA* and *aag* were produced as well as a double deletion strain. The deletions did not affect the growth of *B. subtilis* under normal conditions (Figure 5-5; 0 μg/mL MMS plates).

The deletion strains were plated on LB media containing increasing concentrations of the general S_N2 methylating agent methyl methanesulfonate (MMS), which produces methylated lesions such as 7meG and 3meA. The sensitivities of the deletion strains were compared to WT *B. subtilis* (Figure 5-5). No MMS sensitivity was detected at the lowest concentration of 50 μg/mL MMS. At 100 μg/mL MMS, the single *alkA* and double deletion strains showed increased MMS sensitivity in comparison to the WT strain, while the *aag* deletion behaved similarly to WT. At the highest MMS concentration tested (200 μg/mL), all the deletion strains were highly sensitized to MMS. Again, the single *alkA* deletion appeared to be more sensitized than the *aag* deletion, leading to one model that only bAlkA is the necessary enzyme for methylation repair in *B. subtilis.* However, as shown at the longer growth time of 24 hrs, the double deletion strain is the most sensitized to MMS, more so than the single *alkA* deletion strain. Therefore, both bAlkA and bAAG repair methylation damage in vivo. Although their niches may be redundant, these data suggest that bAlkA is the more important of the two enzymes for repair of MMS-induced

damage in vivo. Previous work has shown that the bAAG protein has relatively low activity towards 3meA and 3meG lesions which are the cytotoxic lesions most commonly produced by MMS.[34] We hypothesize that bAlkA exhibits greater activity than bAAG for these lesions and therefore is more necessary for MMS resistance in vivo. We did not seek to adapt the cells to

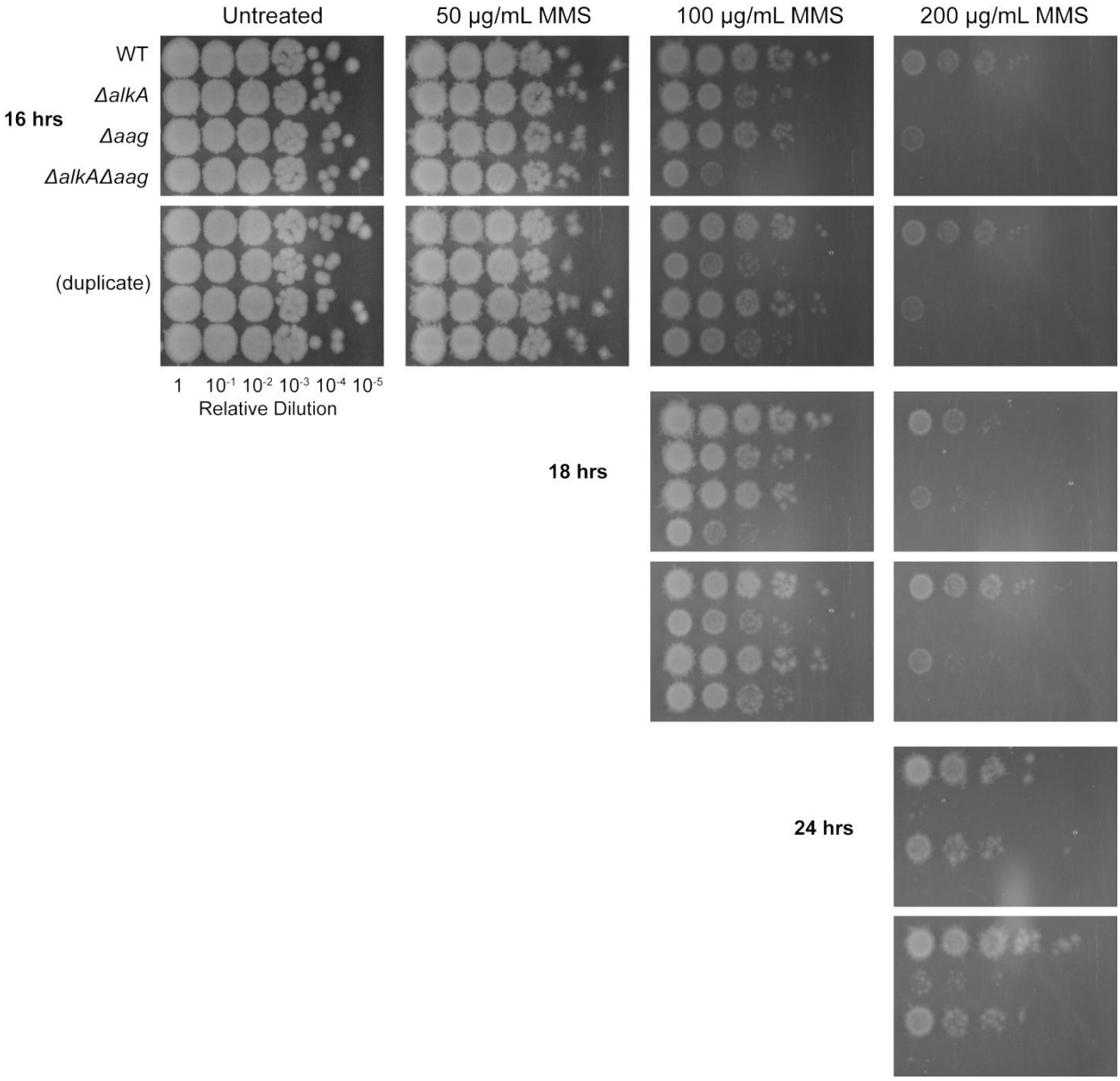


Figure 5-5. MMS resistance of *B. subtilis* strains lacking bAlkA and bAAG proteins. *alkA* and *aag* genes were deleted from the *B. subtilis* genome. Cells were plated in ten-fold dilutions on LB with increasing concentrations of MMS and grown at 37 °C overnight. The deletion strains show varying sensitivities to MMS over time.

induce bAlkA overexpression, although the ada response may have occurred. Liquid culture cell survival assays may be more appropriate for testing the induction of the ada response in relation to MMS sensitivities. Future experimentation is required to better understand in vivo repair by bAlkA and bAAG.

In vitro Activity of B. subtilis Glycosylases

WT bAlkA and bAAG were purified for the purpose of studying their glycosylase activity on various substrates. Single turnover kinetics assays were performed with bAlkA and bAAG using Hx, εA, εG, 1meA and 3meC substrates to evaluate the ability of the *B. subtilis* glycosylases to excise substrates of the *E. coli* AlkA, human AAG, and *E. coli* AlkB proteins. Concentration dependencies were examined at pH 7.5, and we compared the activity at pH 6.1 for the highest concentration (5 μM) of each protein.. A representative time course and enzyme concentration dependence is shown in Figure 5-6, specifically showing the activity of bAlkA on

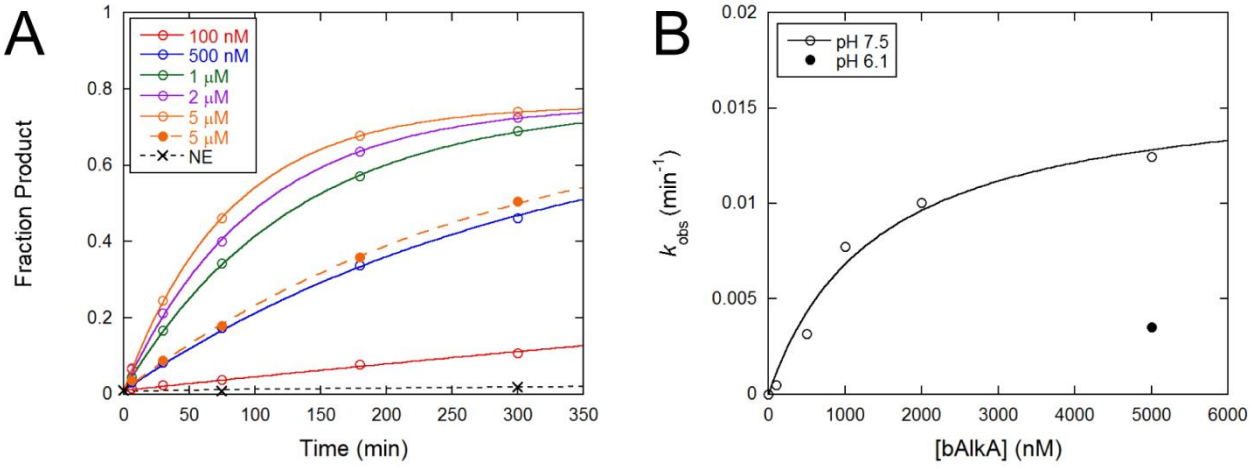


Figure 5-6. Single turnover excision of 1meA by bAlkA. (A) Representation time course for single turnover glycosylase activity assays by the *B. subtilis* enzymes. 20 nM 5'FAM-labeled DNA containing a 1meA lesion was mixed with varying concentrations of bAlkA at 25 °C. Data was fit by a single exponential and was performed once. (B) Representative enzyme concentration dependence showing 1meA excision by bAlkA. Data was fit to a hyperbolic dependence and was performed once. All data in A and B are performed at pH 7.5 (open circles) with an additional bAlkA concentration (5 μM) performed at pH 6.1 (closed circle). Kinetic parameters are reported in Table 5-3.

the 1meA substrate.

The single turnover kinetic parameters for each substrate are detailed in Table 5-3. bAlkA was shown to be active on all substrates tested and bAAG exhibited glycosylase activity on all the substrates tested but 3meC. Catalytic efficiencies calculated from the single turnover kinetic parameters show that bAAG more efficiently removes all the substrates tested except for 3meC

Table 5-3. Kinetic parameters for single turnover excision by bAlkA and bAAG.

Substrate	*Parameter*	*bAlkA*	*bAAG*
Hx	k_{max} (min^{-1})	0.53	8.4
	$K_{1/2}$ (nM)	3500	760
	$k_{max}/K_{1/2}$ $(M^{-1}s^{-1})$	**$2.5 * 10^3$**	**$1.9 * 10^5$**
	optimal pH	6.1 (2.8-fold)	6.1 (1.4-fold)
εA	k_{max} (min^{-1})	4.9	3.5
	$K_{1/2}$ (nM)	950	≤ 5
	$k_{max}/K_{1/2}$ $(M^{-1}s^{-1})$	**$8.6 * 10^4$**	**$\geq 1.2 * 10^7$**
	optimal pH	6.1=7.5	6.1=7.5
εG	k_{max} (min^{-1})	0.28	5.1
	$K_{1/2}$ (nM)	570	65
	$k_{max}/K_{1/2}$ $(M^{-1}s^{-1})$	**$8.2 * 10^3$**	**$1.3 * 10^6$**
	optimal pH	6.1 (≥ 10-fold)	6.1 (1.7-fold)
1meA	k_{max} (min^{-1})	0.016	0.018
	$K_{1/2}$ (nM)	1400	≤ 10
	$k_{max}/K_{1/2}$ $(M^{-1}s^{-1})$	**$1.9 * 10^2$**	**$\geq 3.0 * 10^4$**
	optimal pH	7.5 (3.6-fold)	7.5 (5-fold)
3meC	k_{max} (min^{-1})	0.038	NA
	$K_{1/2}$ (nM)	440	NA
	$k_{max}/K_{1/2}$ $(M^{-1}s^{-1})$	**$1.4 * 10^3$**	**NA**
	optimal pH	6.1 (1.2-fold)	NA

All parameters were determined at pH 7.5 and 25 °C.

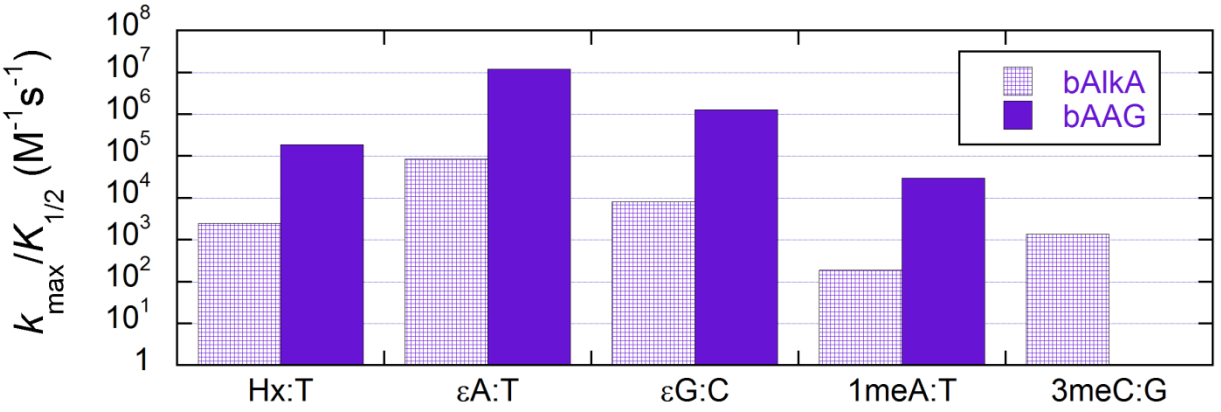


Figure 5-7. Catalytic efficiencies of bAlkA and bAAG. The ratio of single turnover k_{max} and $K_{1/2}$ parameters estimate the catalytic efficiencies of bAlkA and bAAG for all lesions tested. Values plotted are shown in Table 5-3 and represent data collected at pH 7.5 and 25 °C.

on which it has no activity (Table 5-3; Figure 5-7). Similar to the human AAG protein, bAAG binds εA with sub-nanomolar affinity. Interestingly, bAAG also binds 1meA very tightly, nominating it as the potential in vivo repair enzyme for the methylated lesion. The human AAG has also been shown to bind 1meA and 3meC although it does not exhibit glycosylase activity toward the lesions.[39] Furthermore, bAAG shows robust activity on the εG substrate, that is much greater than the activity of the human enzyme on the same lesion.[41] In humans, the AlkB homolog, ABH2, also repairs εG and thus the bAAG may have evolved to replace this missing function, or human AAG lost this ability because ABH2 has activity towards εG.[40]

In contrast, bAlkA shows similar activity but relatively low efficiency for excision of the substrates tested. This could indicate that other substrates are preferred by bAlkA (e.g., 3meA and 3meG). As shown with the MMS resistance assays, bAlkA seems to play a more major role in methylation repair in vivo than bAAG and the *N*-methylated purines are the most commonly produced cytotoxic lesions by MMS. In addition, bAlkA acts on 3meC which is cytotoxic while bAAG does not exhibit activity towards the lesion. These data suggest bAlkA is necessary for cytotoxic lesion removal while bAAG is more efficient at protection against mutagenic lesions.

However, we cannot eliminate the model that bAlkA is just a less efficient repair enzyme than bAAG. Since bAlkA is part of the ada response in *B. subtilis*, this lack of efficiency may be counter balanced by its overexpression during times of alkylation stress. This model would be reminiscent of the *E. coli* AlkA protein homolog.

These single turnover kinetics assays are preliminary and set the stage for future experimentation. Although only pH 7.5 parameters are shown in Table 5-3, a pH 6.1 condition was additionally tested at the highest enzyme concentration. Similarly to *E. coli* AlkA and human AAG, the *B. subtilis* proteins exhibit faster rates of Hx excision at a lower pH condition (fold change shown in Table 5-3). In contrast, 1meA excision is faster at the higher pH 7.5 value. We hypothesize that bAlkA and bAAG may use a similar general acid-base mechanism to their respective homologs, in which protonated substrates are stabilized during the transition state due to their delocalized positive charge, while neutral substrates require an active site acid to stabilize the transition state.[43,44] More experimentation is needed to expand on these models.

In a final comparison of the *B. subtilis* proteins with their *E. coli* and human counterparts, we assessed the maximal single turnover rate of excision of Hx and εA under the pH 7.5 conditions for all the enzymes (Figure 5-8). bAlkA and bAAG are dramatically faster than the *E. coli* AlkA and human AAG enzymes at Hx and εA excision. It should be noted that the *E. coli* and human glycosylase activity at pH 7.5 was estimated from previous reports and is at the 37 °C condition, thus an even greater disparity between the homologs is expected at 25 °C.[43,45,46] Due to the large differences in doubling times of *Bacillus* and human cells, we would hypothesize that bAAG would exhibit faster activity than human AAG in order to protect cells during times of alkylation stress. bAAG also shows tight binding to many of its substrates ($K_{1/2}$ values in the low nanomolar range; Table 5-3), suggesting that bAAG is a much more efficient glycosylase on

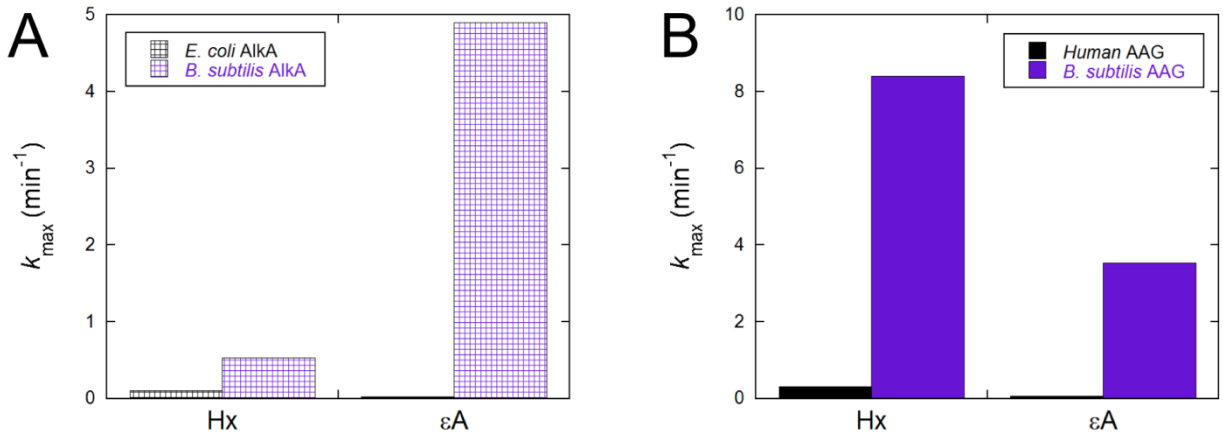


Figure 5-8. Comparisons between *B. subtilis*, *E. coli*, and human AlkA and AAG. Single turnover k_{max} values for Hx and εA excision are shown. *B. subtilis* values are repeated from Table 5-3 and are under pH 7.5 and 25 °C conditions. *E. coli* and human data are shown for 37 °C conditions and estimated at pH 7.5. Specifically, AlkA-Hx data was estimated from the excision of Hx by CTT AlkA at pH 7.5 (Chapter 3) correcting for the effect of glycerol (Chapter 2),[45] AlkA-εA data was estimated from the excision of εA by WT AlkA (Chapter 2)[45] and corrected for pH effects,[46] AAG-Hx and AAG-εA data was estimated from k_{max} values previously described and corrected for pH effects.[43]

these substrates compared to human AAG.

In contrast, the faster single turnover rate constants for bAlkA compared to *E. coli* AlkA cannot be explained by the growth rates of the bacteria, which are similar. As shown in Table 5-3, bAlkA binds Hx and εA in the low micromolar range compared to the low nanomolar range previously described for *E. coli*.[45,47] Therefore, the overall catalytic efficiencies (k_{cat}/K_m) for *E. coli* AlkA and bAlkA may be comparable. Again, bAAG seems to be the more efficient enzyme compared to bAlkA, but like AlkA in *E. coli*, bAlkA is overexpressed as part of an adaptive response and may not be required to be as efficient. Overall, these data do show that *Bacillus subtilis* has evolved to contain multiple alkylation repair glycosylases that are appropriately effective in their protection cells from alkylation damage.

Conclusions

The in vivo roles of *B. subtilis* alkylation repair glycosylases are relatively unknown. Here we have started an investigation of bAlkA and bAAG enzymes. Both glycosylases protect cells from methylation damage as shown with the increased MMS sensitivities of bAlkA and bAAG deletion strains. In vitro kinetics assays also showed that bAlkA and bAAG exhibit broad substrate ranges, including canonical AlkB substrates. These data lay the groundwork for future experimentation with the *B. subtilis* enzymes. In depth kinetics studies and expanded cell survival assays in various DNA damaging agents will expand our knowledge of these glycosylases. In addition, to truly understand alkylation repair in *B. subtilis*, investigation into the other putative alkylation repair glycosylases should also be explored. Such studies will help us understand why the *Bacillus* genome has evolved to contain so many alkylation repair glycosylases in contrast to the *E. coli* and eukaryotic systems.

References

1. Sedgwick, B. (2004) Repairing DNA-methylation damage, *Nat Rev Mol Cell Biol 5*, 148-157.
2. Samson, L., and Cairns, J. (1977) A new pathway for DNA repair in Escherichia coli, *Nature 267*, 281-283.
3. Demple, B., Sedgwick, B., Robins, P., Totty, N., Waterfield, M. D., and Lindahl, T. (1985) Active site and complete sequence of the suicidal methyltransferase that counters alkylation mutagenesis, *Proceedings of the National Academy of Sciences of the United States of America 82*, 2688-2692.
4. Sedgwick, B., Robins, P., Totty, N., and Lindahl, T. (1988) Functional domains and methyl acceptor sites of the Escherichia coli ada protein, *The Journal of biological chemistry 263*, 4430-4433.
5. He, C., Wei, H., and Verdine, G. L. (2003) Converting the sacrificial DNA repair protein N-ada into a catalytic methyl phosphotriester repair enzyme, *Journal of the American Chemical Society 125*, 1450-1451.
6. McCarthy, T. V., and Lindahl, T. (1985) Methyl phosphotriesters in alkylated DNA are repaired by the Ada regulatory protein of E. coli, *Nucleic Acids Res 13*, 2683-2698.
7. Lindahl, T., Sedgwick, B., Sekiguchi, M., and Nakabeppu, Y. (1988) Regulation and expression of the adaptive response to alkylating agents, *Annual review of biochemistry 57*, 133-157.
8. Landini, P., and Volkert, M. R. (2000) Regulatory responses of the adaptive response to alkylation damage: a simple regulon with complex regulatory features, *Journal of bacteriology 182*, 6543-6549.
9. Sedgwick, B., and Lindahl, T. (2002) Recent progress on the Ada response for inducible repair of DNA alkylation damage, *Oncogene 21*, 8886-8894.
10. Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., and Sedgwick, B. (2002) Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage, *Nature 419*, 174-178.
11. Falnes, P. O., Johansen, R. F., and Seeberg, E. (2002) AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli, *Nature 419*, 178-182.
12. Maciejewska, A. M., Sokolowska, B., Nowicki, A., and Kusmierek, J. T. (2011) The role of AlkB protein in repair of 1,N(6)-ethenoadenine in Escherichia coli cells, *Mutagenesis 26*, 401-406.
13. Thomas, L., Yang, C. H., and Goldthwait, D. A. (1982) Two DNA glycosylases in Escherichia coli which release primarily 3-methyladenine, *Biochemistry 21*, 1162-1169.
14. Bjelland, S., Bjoras, M., and Seeberg, E. (1993) Excision of 3-methylguanine from alkylated DNA by 3-methyladenine DNA glycosylase I of Escherichia coli, *Nucleic Acids Res 21*, 2045-2049.
15. Bjelland, S., Birkeland, N. K., Benneche, T., Volden, G., and Seeberg, E. (1994) DNA glycosylase activities for thymine residues oxidized in the methyl group are functions of

the AlkA enzyme in Escherichia coli, *The Journal of biological chemistry 269*, 30489-30495.
16. Saparbaev, M., Kleibl, K., and Laval, J. (1995) Escherichia coli, Saccharomyces cerevisiae, rat and human 3-methyladenine DNA glycosylases repair 1,N6-ethenoadenine when present in DNA, *Nucleic Acids Res 23*, 3750-3755.
17. Saparbaev, M., and Laval, J. (1994) Excision of hypoxanthine from DNA containing dIMP residues by the Escherichia coli, yeast, rat, and human alkylpurine DNA glycosylases, *Proceedings of the National Academy of Sciences of the United States of America 91*, 5873-5877.
18. Masaoka, A., Terato, H., Kobayashi, M., Honsho, A., Ohyama, Y., and Ide, H. (1999) Enzymatic repair of 5-formyluracil. I. Excision of 5-formyluracil site-specifically incorporated into oligonucleotide substrates by alka protein (Escherichia coli 3-methyladenine DNA glycosylase II), *The Journal of biological chemistry 274*, 25136-25143.
19. Terato, H., Masaoka, A., Kobayashi, M., Fukushima, S., Ohyama, Y., Yoshida, M., and Ide, H. (1999) Enzymatic repair of 5-formyluracil. II. Mismatch formation between 5-formyluracil and guanine during dna replication and its recognition by two proteins involved in base excision repair (AlkA) and mismatch repair (MutS), *The Journal of biological chemistry 274*, 25144-25150.
20. Rohankhedkar, M. S., Mulrooney, S. B., Wedemeyer, W. J., and Hausinger, R. P. (2006) The AidB component of the Escherichia coli adaptive response to alkylating agents is a flavin-containing, DNA-binding protein, *Journal of bacteriology 188*, 223-230.
21. Bowles, T., Metz, A. H., O'Quin, J., Wawrzak, Z., and Eichman, B. F. (2008) Structure and DNA binding of alkylation response protein AidB, *Proceedings of the National Academy of Sciences of the United States of America 105*, 15299-15304.
22. Fernandez de Henestrosa, A. R., and Barbe, J. (1991) Induction of the alkA gene of Escherichia coli in gram-negative bacteria, *Journal of bacteriology 173*, 7736-7740.
23. Mielecki, D., Wrzesinski, M., and Grzesiuk, E. (2015) Inducible repair of alkylated DNA in microorganisms, *Mutation research. Reviews in mutation research 763*, 294-305.
24. Grasso, S., and Tell, G. (2014) Base excision repair in Archaea: back to the future in DNA repair, *DNA repair 21*, 148-157.
25. Bjoras, M., Klungland, A., Johansen, R. F., and Seeberg, E. (1995) Purification and properties of the alkylation repair DNA glycosylase encoded the MAG gene from Saccharomyces cerevisiae, *Biochemistry 34*, 4577-4582.
26. Memisoglu, A., and Samson, L. (1996) Cloning and characterization of a cDNA encoding a 3-methyladenine DNA glycosylase from the fission yeast Schizosaccharomyces pombe, *Gene 177*, 229-235.
27. Duncan, T., Trewick, S. C., Koivisto, P., Bates, P. A., Lindahl, T., and Sedgwick, B. (2002) Reversal of DNA alkylation damage by two human dioxygenases, *Proceedings of the National Academy of Sciences of the United States of America 99*, 16660-16665.
28. Ougland, R., Rognes, T., Klungland, A., and Larsen, E. (2015) Non-homologous functions of the AlkB homologs, *Journal of molecular cell biology*.
29. Engelward, B. P., Boosalis, M. S., Chen, B. J., Deng, Z., Siciliano, M. J., and Samson, L. D. (1993) Cloning and characterization of a mouse 3-methyladenine/7-methyl-guanine/3-methylguanine DNA glycosylase cDNA whose gene maps to chromosome 11, *Carcinogenesis 14*, 175-181.

30. Hollis, T., Lau, A., and Ellenberger, T. (2001) Crystallizing thoughts about DNA base excision repair, *Prog Nucleic Acid Res Mol Biol 68*, 305-314.
31. Morohoshi, F., Hayashi, K., and Munakata, N. (1990) Bacillus subtilis ada operon encodes two DNA alkyltransferases, *Nucleic Acids Res 18*, 5473-5480.
32. Morohoshi, F., Hayashi, K., and Munkata, N. (1993) Bacillus subtilis alkA gene encoding inducible 3-methyladenine DNA glycosylase is adjacent to the ada operon, *Journal of bacteriology 175*, 6010-6017.
33. Lenhart, J. S., Schroeder, J. W., Walsh, B. W., and Simmons, L. A. (2012) DNA repair and genome maintenance in Bacillus subtilis, *Microbiology and molecular biology reviews : MMBR 76*, 530-564.
34. Aamodt, R. M., Falnes, P. O., Johansen, R. F., Seeberg, E., and Bjoras, M. (2004) The Bacillus subtilis counterpart of the mammalian 3-methyladenine DNA glycosylase has hypoxanthine and 1,N6-ethenoadenine as preferred substrates, *The Journal of biological chemistry 279*, 13601-13606.
35. Alseth, I., Rognes, T., Lindback, T., Solberg, I., Robertsen, K., Kristiansen, K. I., Mainieri, D., Lillehagen, L., Kolsto, A. B., and Bjoras, M. (2006) A new protein superfamily includes two novel 3-methyladenine DNA glycosylases from Bacillus cereus, AlkC and AlkD, *Molecular microbiology 59*, 1602-1609.
36. Mroczkowska, M. M., Kolasa, I. K., and Kusmierek, J. T. (1993) Chloroacetaldehyde-induced mutagenesis in Escherichia coli: specificity of mutations and modulation by induction of the adaptive response to alkylating agents, *Mutagenesis 8*, 341-348.
37. Gros, L., Maksimenko, A. V., Privezentzev, C. V., Laval, J., and Saparbaev, M. K. (2004) Hijacking of the human alkyl-N-purine-DNA glycosylase by 3,N4-ethenocytosine, a lipid peroxidation-induced DNA adduct, *The Journal of biological chemistry 279*, 17723-17730.
38. Gros, L., Ishchenko, A. A., and Saparbaev, M. (2003) Enzymology of repair of etheno-adducts, *Mutation research 531*, 219-229.
39. Lee, C. Y., Delaney, J. C., Kartalou, M., Lingaraju, G. M., Maor-Shoshani, A., Essigmann, J. M., and Samson, L. D. (2009) Recognition and processing of a new repertoire of DNA substrates by human 3-methyladenine DNA glycosylase (AAG), *Biochemistry 48*, 1850-1861.
40. Zdzalik, D., Domanska, A., Prorok, P., Kosicki, K., van den Born, E., Falnes, P. O., Rizzo, C. J., Guengerich, F. P., and Tudek, B. (2015) Differential repair of etheno-DNA adducts by bacterial and human AlkB proteins, *DNA repair 30*, 1-10.
41. Saparbaev, M., Langouet, S., Privezentzev, C. V., Guengerich, F. P., Cai, H., Elder, R. H., and Laval, J. (2002) 1,N(2)-ethenoguanine, a mutagenic DNA adduct, is a primary substrate of Escherichia coli mismatch-specific uracil-DNA glycosylase and human alkylpurine-DNA-N-glycosylase, *The Journal of biological chemistry 277*, 26987-26993.
42. Patrick, J. E., and Kearns, D. B. (2008) MinJ (YvjD) is a topological determinant of cell division in Bacillus subtilis, *Molecular microbiology 70*, 1166-1179.
43. 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.
44. O'Brien, P. J. (2006) Catalytic promiscuity and the divergent evolution of DNA repair enzymes, *Chem Rev 106*, 720-752.
45. Taylor, E. L., and O'Brien, P. J. (2015) Kinetic mechanism for the flipping and excision

of 1,N(6)-ethenoadenine by AlkA, *Biochemistry 54*, 898-908.
46. O'Brien, P. J., and Ellenberger, T. (2004) The Escherichia coli 3-methyladenine DNA glycosylase AlkA has a remarkably versatile active site, *The Journal of biological chemistry 279*, 26876-26884.
47. Zhao, B., and O'Brien, P. J. (2011) Kinetic mechanism for the excision of hypoxanthine by Escherichia coli AlkA and evidence for binding to DNA ends, *Biochemistry 50*, 4350-4359.

Chapter VI

Conclusions and Future Directions

Conclusions

This work has compared and contrasted the mechanisms of two alkylation repair glycosylases, AlkA and AAG. While structurally different, AlkA and AAG have independently evolved to fill a vital niche of alkylation damage repair in bacteria, eukaryotes and archaea. These enzymes initiate the vital base excision repair pathway and protect cells from damaging agents found naturally in the environment and produced endogenously. Here we have described thorough biochemical experimentation detailing the kinetic and thermodynamic framework of AlkA, the binding of AAG to damaged and undamaged DNA, and the substrate specificities of two *Bacillus subtilis* homologs of the glycosylases. This work expands our understanding of these two independently evolved glycosylases and their roles in alkylation damage repair.

Kinetic Mechanism of Escherichia coli AlkA and Comparison to AAG

We have characterized the kinetic mechanism of AlkA-catalyzed excision of the endogenously produced 1,N^6-ethenoadenine (εA) (Chapter 2).[1] Our work shows that AlkA rapidly and reversibly associates with the lesion, flipping it between the active site and the DNA duplex at almost equally fast rates. The flipping equilibrium is defined by K_{flip} (k_{flip}/k_{unflip}) has a value of 3 for AlkA binding to εA–DNA. When compared to the human AAG binding to εA–DNA which has a K_{flip} value of 2300,[2] AlkA is relatively unstable in the binding of εA-DNA and

this is reflected in the 1000-fold difference in K_d values for εA binding. However, it is surprising that AlkA and AAG exhibit identical rates of εA excision. The differences in binding imply different roles of these two enzymes in vivo. AlkA is part of the adaptive response in *E. coli* and is overexpressed during times of alkylation stress.[3] It can be hypothesized that AlkA is less efficient on specific substrates in order to be generalist. Its permissive active site allows for a variety of bases with structural differences and the enzyme even performs gratuitous repair on undamaged bases.[4,5] The ada response and the induction of AlkA expression optimizes this trade-off in efficiency and promiscuity, as AlkA is only overexpressed during times of alkylation stress.[3]

Additional studies with AlkA show that the excision of εA and the deaminated adenine lesion, hypoxanthine (Hx), are affected by salt concentration (Chapter 3). While other glycosylases including AAG have positively charged binding clefts that aid in electrostatic interactions between protein and DNA, AlkA has relatively few charged residues.[6-10] However, in vitro single turnover and multiple turnover kinetics assays show that AlkA binds undamaged DNA weaker at higher salt concentrations and these effects may rely on the identity of the salt ions. We have also shown that while AlkA binds εA tighter than Hx, it excises Hx at a faster rate, thus having almost identical catalytic efficiencies for the excision of the lesions. This work again shows that AlkA is a promiscuous enzyme that can act on many substrates. Finally, we characterized a modified version of AlkA that has an 8 amino acid remnant on its C-terminus, leftover from affinity purification. Interestingly, the tagged AlkA behaves better at high concentrations and does not promote auto-inhibition as observed with WT AlkA. Examination of the conservation of the C-termini of AlkA homologs raises the possibility that some, but not all AlkA homologs, would be auto-inhibitory at high concentrations of protein. Possible functional

roles for such high density binding would be shielding of undamaged sites from reactive species or attenuation of gratuitous AlkA-catalyzed excision of undamaged bases. Additional work would be needed to understand how extension of the C-terminus alters DNA binding and catalysis of AlkA, because it is not near the active site or DNA binding cleft. Regardless of the mechanism, these results suggest that the C-terminally modified form of AlkA would be a useful reagent for removing damaged bases from genomic DNA samples in the quantification of lesion sites or the recovery and amplification of environmental DNA.

Analysis of Binding by AAG using Fluorescence Anisotropy Methods

We employed fluorescence anisotropy methods to analyze the binding of AAG to specific εA-DNA and undamaged DNA substrates to determine if AAG binds as a dimer or monomer (Chapter 4). Anisotropy and the total fluorescence of 5′ 6-fluorescein (FAM) εA–DNA substrates yielded complex biphasic signals that could be explained by a dimer model. However, by using the natural fluorescence of εA to our advantage we show that AAG binds εA specifically in a 1:1 binding complex and this complex is defined by a small anisotropy and FAM quenching signal change. The large signal changes observed correspond to AAG binding near the 5′FAM site and interacting with the fluorophore, causing a loss of flexibility in the FAM linker. Compared to a cyanine-like dye, DyLight 647 (DY647), which stacks on the end of duplex DNA, the tumbling of the FAM is much more sensitive to AAG binding events. We used this model to investigate the affinity of AAG for the FAM site and for undamaged unlabeled DNAs. AAG binds DNA near to the 5′FAM in the low nanomolar range, while a single nonspecific binding site has an affinity in the low micromolar range. As these methods are commonly used for the analysis of DNA binding proteins, including other glycosylases,[11-13] we show that caution should be used in

the analysis of complex binding models in which multiple proteins bind with high density to DNA.

Roles of Bacillus subtilis AlkA and AAG Homologs

The Gram-positive *B. subtilis* is unique in that it contains both AlkA and AAG homologs in their genome.[14,15] As these two glycosylases are almost always mutually exclusive in their expression, we chose to study the *B. subtilis* homologs for differences in their substrate specificities (Chapter 5). The genes were deleted from the genome which increased the sensitivity of *B. subtilis* to MMS, a common methylating agent. While the bAlkA deletion strain was sensitized to a greater extent than the bAAG deletion strain, a double deletion strain was additive in the MMS sensitivity, suggesting that both bAlkA and bAAG protect cells from alkylation damage. In addition, bAlkA and bAAG were purified and their glycosylase activities were analyzed under single turnover conditions. Both enzymes displayed broad substrate ranges and excised Hx, εA, 1,N^2-ethenoguanine (εG), and N^1-methyladenine (1meA). bAlkA also excised N^3-methylcytosine (3meC). 1meA and 3meC are canonical substrates of the direct reversal AlkB protein which does not have a homolog in *B. subtilis*. Our results suggest that bAAG is the more catalytically efficient enzyme for these lesion excisions, however the in vivo data and previous work with bAAG[15] suggests bAlkA plays the more major role in the repair of cytotoxic lesions caused by MMS treatment such as 3meA and 3meC. This work lays the groundwork for future experimentation with the *Bacillus* enzymes.

Future Directions

AlkA Kinetic Mechanism and Substrate Specificity

We have clearly defined the kinetic and thermodynamic framework for AlkA-catalyzed excision of εA and previous work has characterized the excision of Hx by AlkA.[16] However, many questions remain regarding the substrate specificity of AlkA. It is interesting that while AlkA acts with similar rate enhancement on all substrates tested,[4] and exhibits robust activity on εA and Hx in vitro, these substrates are not deemed the physiological substrates of AlkA, but instead are acted on by AlkB and endonuclease V.[17,18] Future experimentation could include directly competing different substrates for AlkA excision as described in Chapter 3. I hypothesize that inefficient binding could be responsible for low k_{cat}/K_m values which would be expected to limit the effectiveness of AlkA in the physiological context where there is a vast excess of undamaged sites.

There is abundant evidence that deletion of AlkA results in a strong sensitivity to DNA alkylating agents, and it seems likely that the ability of AlkA to excise 3meA and 3meG provides the basis for the protective effect of AlkA. However, 3me purine lesions are chemically unstable and attempts to incorporate site-specific lesions through solid phase synthesis have been unsuccessful. The stable analog, 3-deaza-3-methyladenine has been synthesized and can be commercially obtained. This analog is isosteric with 3meA, but it is not charged at physiological pH. We hypothesized that at low pH it may be possible to protonate N^7 which could result in a positively charged purine that might be a substrate for AlkA. Preliminary experiments with the

analog showed limited excision by AlkA that could be caused by contamination of the lesion. Mass spectrometry analysis may shed some light on the purity of the synthetic DNA and identify the chemical structure of the small amount of reactive material.

In addition to studying other substrates, the ionic strength dependence of AlkA can be further explored. Our work shows that there may be some effects of ion identity. The structure of AlkA bound to DNA shows a putative sodium ion binding site in the interface of the helix-hairpin-helix binding motif and the DNA backbone.[10] Other studies with AAG have also shown that the chloride ion commonly used in our assays inhibits AAG activity. More work is needed to understand the complex nature of the ionic strength dependence data and determine the most optimal buffer for AlkA study.

Finally, all of this work has been performed with WT AlkA (or the tagged version thereof). Only the catalytically dead mutant in which the catalytic base is mutated (D238N) has been described. The active site of AlkA contains many aromatic residues proposed to interact with the flipped nucleotide. As the only structure of AlkA bound specifically to DNA is with a 1-azaribose moiety, it is unknown how a flipped nucleotide would bind in the active site.[10] Being that AlkA is promiscuous, different configurations may occur with different lesions. Fortuitously, two of the enzyme active site residues are tryptophans (W218; W272) which are inherently fluorescent.[10] Previous work on AAG has shown that Trp residues can act as reporters for nucleotide flipping and protein binding.[19,20] In contrast to AAG in which the intercalating tyrosine and active site residues were mutated to Trp, AlkA naturally contains these lesions in the active site and preliminary work has shown Trp fluorescence changes in AlkA when rapidly mixed with DNA (data not shown). The other Trp residues in the protein are not near the active site and can potentially be mutated to reduce the amount of fluorescence background. The

definition of nucleotide flipping rates by AlkA have been described for εA-DNA.[1] However, using the active site tryptophan reporters, additional substrates that are not fluorescent could also be easily studied without modifying the nucleotide or placing an unnatural base next to the lesion as employed by previous groups.[21-23]

Furthermore, mutating different active site residues or the intercalating leucine residue would help determine their function contributions to nucleotide flipping and base excision. It is interesting that when the intercalating tyrosine residue of AAG is mutated to an alanine, AAG loses its tight selective binding for εA.[2] If the intercalating leucine of AlkA was thus mutated to a large aromatic residue, this may stabilize the flipped out lesion:AlkA complex.

Fluorescence Anisotropy Methods for DNA Binding Analysis

We have thoroughly examined DNA binding by AAG using fluorescence anisotropy and native gel shift methods. However, this is one example of glycosylase binding used to explain a model involving fluorophore complexities. This work could easily be expanded to include other glycosylases and DNA binding enzymes like AP endonuclease I, a downstream base excision repair protein. Preliminary work with the tagged WT AlkA showed similar biphasic anisotropy patterns using the tight binding pyrrolidine inhibitor (data not shown). Other glycosylases such as mismatch uracil DNA glycosylase (Mug) have also shown this activity and could be explained by our fluorophore model.[12] For a more expansive comparison, different proteins and a wider variety of fluorophores and linkers should be examined. As fluorescence anisotropy with fluorescein derivatives is widespread, this work could lead to improved DNA binding assays in many different systems.

Bacillus subtilis Alkylation Repair Glycosylases

Perhaps the greatest amount of future directions lies with the *B. subtilis* glycosylases. As described, only preliminary work has been done on these relatively uncharacterized proteins to define their substrate specificities and physiological niches for alkylation repair. Liquid culture cell survival assays are key to show the redundancy or necessity of bAlkA and bAAG in vivo as the cells are exposed to a better controlled MMS concentration and cell survival can be accurately quantified. While only MMS was used in our preliminary studies, other methylating agents, such as MNNG which produces S_N1 lesions, or compounds such as chloracetaldehyde (CAA) that produce etheno adducts could be tested. Additionally, the adaption of bAlkA overexpression can be studied to determine the increased protection to these agents.

The in vitro kinetics assays with bAlkA and bAAG have only just begun and lay the framework for future experiments. Optimization of the protein purification protocols and buffer components should be explored in order to produce larger quantities of stable protein. In addition, assays determining the active concentrations of the enzymes are necessary in order to accurately compare single turnover and multiple turnover kinetic parameters. While the bAAG protein seems to bind εA with high affinity and the steady state εA-DNA fluorescence could potentially report on the bAAG concentration, the affinity of bAlkA for εA is in the low micromolar range. Perhaps unlike its *E. coli* counterpart, multiple turnover burst analysis may be possible with bAlkA to determine its active concentration. Subsequently, the substrate specificities could easily be better defined for these proteins by putting different substrates in competition for glycosylase activity. It is interesting that bAlkA and bAAG work on AlkB substrates. While the human AAG shows binding towards these lesions it has not been reported to have activity.[24] Expanding the kinetics assays to include other substrates such as 7meG and

oxidatively damaged substrates (e.g., xanthosine) could also prove enlightening.

Finally, we must not forget that *B. subtilis* contains two other putative alkylation repair glycosylases, *yfjP* and *yhaZ*. The *yfjP* protein seems to actually be a closer homolog to the *E. coli* AlkA, although it is not part of the adaptive response. Sequence similarity between *yhaZ* and the *Bacillus cereus* AlkC protein suggests that the fourth glycosylase is part of a newly defined superfamily of glycosylases.[25] These proteins are not well understood and may have expanded unique substrate specificities that include bulkier adducts produced by competing microorganisms in addition to *N*-methyl purines.[25,26] Deletion and cloning of these two proteins is necessary for the overall understanding of alkylation repair in *B. subtilis*. Without the study of all four glycosylases in concert, the true physiological roles of each substrate cannot be fully determined.

References

1. Taylor, E. L., and O'Brien, P. J. (2015) Kinetic mechanism for the flipping and excision of 1,N(6)-ethenoadenine by AlkA, *Biochemistry 54*, 898-908.
2. Hendershot, J. M., and O'Brien, P. J. (2014) Critical role of DNA intercalation in enzyme-catalyzed nucleotide flipping, *Nucleic Acids Res*.
3. Thomas, L., Yang, C. H., and Goldthwait, D. A. (1982) Two DNA glycosylases in Escherichia coli which release primarily 3-methyladenine, *Biochemistry 21*, 1162-1169.
4. O'Brien, P. J., and Ellenberger, T. (2004) The Escherichia coli 3-methyladenine DNA glycosylase AlkA has a remarkably versatile active site, *The Journal of biological chemistry 279*, 26876-26884.
5. Berdal, K. G., Johansen, R. F., and Seeberg, E. (1998) Release of normal bases from intact DNA by a native DNA repair enzyme, *The EMBO journal 17*, 363-367.
6. Hedglin, M., and O'Brien, P. J. (2008) Human alkyladenine DNA glycosylase employs a processive search for DNA damage, *Biochemistry 47*, 11434-11445.
7. Zhang, Y., and O'Brien, P. J. (2015) Repair of Alkylation Damage in Eukaryotic Chromatin Depends on Searching Ability of Alkyladenine DNA Glycosylase, *ACS chemical biology*.
8. Eichman, B. F., O'Rourke, E. J., Radicella, J. P., and Ellenberger, T. (2003) Crystal structures of 3-methyladenine DNA glycosylase MagIII and the recognition of alkylated bases, *The EMBO journal 22*, 4898-4909.
9. Hollis, T., Lau, A., and Ellenberger, T. (2001) Crystallizing thoughts about DNA base excision repair, *Prog Nucleic Acid Res Mol Biol 68*, 305-314.
10. Hollis, T., Ichikawa, Y., and Ellenberger, T. (2000) DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, Escherichia coli AlkA, *The EMBO journal 19*, 758-766.
11. Bharti, S. K., and Varshney, U. (2010) Analysis of the impact of a uracil DNA glycosylase attenuated in AP-DNA binding in maintenance of the genomic integrity in Escherichia coli, *Nucleic Acids Res 38*, 2291-2301.
12. Grippon, S., Zhao, Q., Robinson, T., Marshall, J. J., O'Neill, R. J., Manning, H., Kennedy, G., Dunsby, C., Neil, M., Halford, S. E., French, P. M., and Baldwin, G. S. (2011) Differential modes of DNA binding by mismatch uracil DNA glycosylase from Escherichia coli: implications for abasic lesion processing and enzyme communication in the base excision repair pathway, *Nucleic Acids Res 39*, 2593-2603.
13. Morgan, M. T., Maiti, A., Fitzgerald, M. E., and Drohat, A. C. (2011) Stoichiometry and affinity for thymine DNA glycosylase binding to specific and nonspecific DNA, *Nucleic Acids Res 39*, 2319-2329.
14. Morohoshi, F., Hayashi, K., and Munkata, N. (1993) Bacillus subtilis alkA gene encoding inducible 3-methyladenine DNA glycosylase is adjacent to the ada operon, *Journal of bacteriology 175*, 6010-6017.

15. Aamodt, R. M., Falnes, P. O., Johansen, R. F., Seeberg, E., and Bjoras, M. (2004) The Bacillus subtilis counterpart of the mammalian 3-methyladenine DNA glycosylase has hypoxanthine and 1,N6-ethenoadenine as preferred substrates, *The Journal of biological chemistry 279*, 13601-13606.
16. Zhao, B., and O'Brien, P. J. (2011) Kinetic mechanism for the excision of hypoxanthine by Escherichia coli AlkA and evidence for binding to DNA ends, *Biochemistry 50*, 4350-4359.
17. Maciejewska, A. M., Sokolowska, B., Nowicki, A., and Kusmierek, J. T. (2011) The role of AlkB protein in repair of 1,N(6)-ethenoadenine in Escherichia coli cells, *Mutagenesis 26*, 401-406.
18. Yao, M., Hatahet, Z., Melamede, R. J., and Kow, Y. W. (1994) Purification and characterization of a novel deoxyinosine-specific enzyme, deoxyinosine 3' endonuclease, from Escherichia coli, *The Journal of biological chemistry 269*, 16260-16268.
19. Hendershot, J. M., Wolfe, A. E., and O'Brien, P. J. (2011) Substitution of active site tyrosines with tryptophan alters the free energy for nucleotide flipping by human alkyladenine DNA glycosylase, *Biochemistry 50*, 1864-1874.
20. Hendershot, J. M., and O'Brien, P. J. (2014) Critical role of DNA intercalation in enzyme-catalyzed nucleotide flipping, *Nucleic Acids Res 42*, 12681-12690.
21. 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.
22. Bernards, A. S., Miller, J. K., Bao, K. K., and Wong, I. (2002) Flipping duplex DNA inside out: a double base-flipping reaction mechanism by Escherichia coli MutY adenine glycosylase, *The Journal of biological chemistry 277*, 20960-20964.
23. 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.
24. Lee, C. Y., Delaney, J. C., Kartalou, M., Lingaraju, G. M., Maor-Shoshani, A., Essigmann, J. M., and Samson, L. D. (2009) Recognition and processing of a new repertoire of DNA substrates by human 3-methyladenine DNA glycosylase (AAG), *Biochemistry 48*, 1850-1861.
25. Alseth, I., Rognes, T., Lindback, T., Solberg, I., Robertsen, K., Kristiansen, K. I., Mainieri, D., Lillehagen, L., Kolsto, A. B., and Bjoras, M. (2006) A new protein superfamily includes two novel 3-methyladenine DNA glycosylases from Bacillus cereus, AlkC and AlkD, *Molecular microbiology 59*, 1602-1609.
26. Mullins, E. A., Shi, R., Parsons, Z. D., Yuen, P. K., David, S. S., Igarashi, Y., and Eichman, B. F. (2015) The DNA glycosylase AlkD uses a non-base-flipping mechanism to excise bulky lesions, *Nature*.