Structural basis for the inhibition of human alkyladenine DNA by 3,N4-ethenocytosine containing DNA

Citation: Lingaraju, Gondichatnahalli M. et al. "Structural Basis for the Inhibition of Human Alkyladenine DNA Glycosylase (AAG) by 3,N4-Ethenocytosine-containing DNA." Journal of Biological Chemistry 286.15 (2011) : 13205 -13213. © 2011 by American Society for Biochemistry and Molecular Biology.

As Published: http://dx.doi.org/10.1074/jbc.M110.192435

Publisher: American Society for Biochemistry and Molecular Biology, Inc.

Persistent URL: http://hdl.handle.net/1721.1/64727

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

⌘ *Author's Choice*

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 15, pp. 13205–13213, April 15, 2011

Structural Basis for the Inhibition of Human Alkyladenine DNA Glycosylase (AAG) by 3,N^4-Ethenocytosine-containing DNA*[S]

Received for publication, October 7, 2010, and in revised form, February 14, 2011 Published, JBC Papers in Press, February 24, 2011, DOI 10.1074/jbc.M110.192435

Gondichatnahalli M. Lingaraju[‡§], C. Ainsley Davis[¶1], Jeremy W. Setser[¶2], Leona D. Samson[‡§‖3], and Catherine L. Drennan[‡¶‖‡‡4]**

From the [‡]Center for Environmental Health Sciences, the Departments of [§]Biological Engineering, [¶]Chemistry, and [‖]Biology, the [**]Koch Institute for Integrative Cancer Research, and the [‡‡]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Reactive oxygen and nitrogen species, generated by neutrophils and macrophages in chronically inflamed tissues, readily damage DNA, producing a variety of potentially genotoxic etheno base lesions; such inflammation-related DNA damage is now known to contribute to carcinogenesis. Although the human alkyladenine DNA glycosylase (AAG) can specifically bind DNA containing either 1,N^6-ethenoadenine (ϵA) lesions or 3,N^4-ethenocytosine (ϵC) lesions, it can only excise ϵA lesions. AAG binds very tightly to DNA containing ϵC lesions, forming an abortive protein-DNA complex; such binding not only shields ϵC from repair by other enzymes but also inhibits AAG from acting on other DNA lesions. To understand the structural basis for inhibition, we have characterized the binding of AAG to DNA containing ϵC lesions and have solved a crystal structure of AAG bound to a DNA duplex containing the ϵC lesion. This study provides the first structure of a DNA glycosylase in complex with an inhibitory base lesion that is induced endogenously and that is also induced upon exposure to environmental agents such as vinyl chloride. We identify the primary cause of inhibition as a failure to activate the nucleotide base as an efficient leaving group and demonstrate that the higher binding affinity of AAG for ϵC *versus* ϵA is achieved through formation of an additional hydrogen bond between Asn-169 in the active site pocket and the O^2 of ϵC. This structure provides the basis for the design of AAG inhibitors currently being sought as an adjuvant for cancer chemotherapy.

Genotoxic etheno (ϵ)-lesions such as 3,N^4-ethenocytosine (ϵC)[5] and 1,N^6-ethenoadenine (ϵA) are endogenously generated when DNA is attacked by reactive aldehydes. These reactive compounds are generated as byproducts of lipid peroxidation that is induced by reactive oxygen and nitrogen species. Neutrophils and macrophages generate large quantities of reactive oxygen and nitrogen species in tissues undergoing chronic inflammation (1, 2), and it is widely accepted that such inflammation increases the risk of colon cancer in ulcerative colitis and Crohn disease patients and increases the risk of liver cancer in Wilson disease and hemochromatosis patients (1, 3). In fact, increased levels of ϵ-lesions in the DNA of tissues undergoing chronic inflammation have been reported for each of these diseases (4). Depending on the type of DNA polymerase, ϵC mispairs with A, T, or C during DNA replication, resulting in both transition and transversion mutations (5). In contrast, ϵA primarily gives rise to A:T to T:A transversion mutations (6). These mutagenic ϵ-lesions are generally removed via the base excision repair pathway, initiated by lesion-specific DNA glycosylases that cleave the *N*-glycosidic bond between the damaged base and the deoxyribose sugar (5, 7). In humans, several DNA glycosylases can excise ϵC, namely thymine DNA glycosylase, methyl-CpG binding domain protein, and single strand monofunctional uracil DNA glycosylase (5). In contrast, there is only one DNA glycosylase known to excise ϵA lesions in humans, alkyladenine DNA glycosylase (AAG) (5, 8).

AAG (also known as MPG and ANPG) has been previously characterized by crystallography. The crystal structures of an N-terminally truncated, but catalytically active, construct of AAG (Δ79AAG), both bound to a pyrrolidine abasic site mimic (Pyr) and bound to an ϵA-containing piece of DNA, suggest a mode by which AAG recognizes a wide range of lesions while still discriminating against undamaged bases (9, 10). When substrate is bound, AAG can excise the damaged base through acid/base catalysis (11). A putative catalytic water molecule, revealed by crystallographic studies, is proposed to act as a nucleophile as it is ideally positioned to attack the *N*-glycosidic bond present between the ϵA base and the deoxyribose sugar (10). This water molecule is also in contact with the side chain

* This work was supported, in whole or in part, by National Institutes of Health Grants P30-ES002109 (to C. L. D. and L. D. S.), GM65337 (to C. L. D.), GM65337-03S2 (to C. A. D.), and CA055042 and CA092584 (to L. D. S.).

⌘ *Author's Choice*—Final version full access.

The atomic coordinates and structure factors (code 3QI5) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental experimental procedures, Table S1, and Figs. S1–S6.

[1] Present address: Bethune-Cookman University, 640 Dr. Mary McLeod Bethune Blvd., Daytona Beach, FL 32114.

[2] Supported by a Repligen Koch Institute for Integrated Cancer Research Graduate Fellowship.

[3] An American Cancer Society Research Professor. To whom correspondence may be addressed: 77 Massachusetts Ave., Cambridge, MA 02193. Tel.: 617-258-7813; Fax: 617-253-8099; E-mail: lsamson@mit.edu.

[4] A Howard Hughes Medical Institute Investigator. To whom correspondence may be addressed: 77 Massachusetts Ave., Cambridge, MA 02193. Tel.: 617-253-5622; Fax: 617-258-7847; E-mail: cdrennan@mit.edu.

[5] The abbreviations used are: ϵC, 3,N^4-ethenocytosine; ϵA, 1,N^6-ethenoadenine; AAG, alkyladenine DNA glycosylase; Hx, hypoxanthine; 7-meG, 7-methylguanine; ATL, alkyltransferase-like; AP, apurinic/apyrimidinic; r.m.s.d., root mean square deviation.

of Glu-125, which is proposed to be the catalytic base responsible for activating the water for nucleophilic attack. Consistent with this proposal, introduction of an E125Q mutation completely abolishes AAG activity (9–11). The identity of the general acid is unknown. Interestingly, in this structure of Δ79AAG bound to ϵA:T-containing DNA, the ϵA lesion remained intact (10). Later experiments showed that the catalytic activity of Δ79AAG is significantly inhibited in the presence of a variety of divalent metal ions including Mn^{2+}, Zn^{2+}, Ca^{2+}, Cd^{2+}, Ni^{2+}, and most importantly, Mg^{2+}, which was contained in the crystallization buffer (9, 10, 12, 13). However, the structures did not show any Mg^{2+} ions bound (9, 10).

In addition to repairing ϵA, AAG can repair other reactive oxygen and nitrogen species and alkylation-induced DNA damage, including lesions hypoxanthine (Hx), 1,N^2-ethenoguanine, 8-oxoguanine, 3-methyladenine, 7-methylguanine (7-meG), and 3-methylguanine (2, 8). However, despite this broad substrate specificity, there is a growing list of lesions to which AAG can bind while failing to excise the lesion. In addition to ϵC-containing DNA (14), this list now includes 3-methyluracil, 3-ethyluracil, 3-methylthymine, and 3-methylcytosine (15). This binding without cleavage can result in the formation of stable abortive complexes between AAG and damaged DNA. The ϵC-AAG abortive complex has been shown to inhibit AAG glycosylase activity *in vivo* in human cells and to result in replication blockage (14), increasing the genotoxicity of ϵC lesions *in vivo*. In tissues undergoing chronic inflammation, which have higher ϵC content (4), the formation of AAG-ϵC abortive complexes may significantly diminish the repair of other AAG substrates, ultimately resulting in the accumulation of various DNA lesions in addition to ϵC. Interestingly, in ulcerative colitis patients, the colon epithelium undergoing chronic inflammation was found to have increased AAG expression, perhaps indicating an adaptive response triggered by increased levels of DNA damage (16). This adaptation might compensate for the AAG hijacked by ϵC lesions.

To understand the structural basis for the inhibition of AAG by ϵC-containing DNA, we have carried out biochemical and crystallographic studies on AAG, using a truncated form of the protein (Δ79) that was previously shown to have full catalytic activity (11, 15, 17). Our crystal structure of Δ79AAG bound to a DNA duplex containing ϵC:G (G paired opposite ϵC), in combination with previous (11) and current (this work) biochemical analysis, suggests that the failure of AAG to activate the leaving group (ϵC) by protonation is likely the primary reason for its inability to remove ϵC from the DNA. This structure also shows that a divalent metal ion, Mn^{2+}, can bind to the base opposite the ϵC lesion, changing its sugar pucker and providing the first structural framework for considering the molecular basis for metal ion inhibition of AAG.

EXPERIMENTAL PROCEDURES

Δ79AAG Plasmid Construction, Creation of Mutants, and Protein Preparation—Constructs of full-length AAG and a truncated form of AAG (with 84 residues at the N terminus deleted) were cloned into pET19b-PPS vectors for protein expression (UniprotID: P29372). These constructs encoded the wild-type AAG sequence with an N-terminal 10× histidine tag followed by the PreScission Protease cleavage site (PPS) (GE Healthcare). The precision protease treatment to remove the histidine tag leaves behind four extra amino acids (GPHM) from the expression vector at the N terminus of AAG. For full-length AAG, this cleavage site results in a GPHM sequence prior to residue 1, and for the truncated form of AAG, this cleavage site results in the addition of residues Gly-80, Pro-81, His-82, and Met-83 prior to residue 84. We refer to this truncated protein construct as Δ79AAG, although the actual AAG protein sequence starts at Thr-84. It should be noted that the AAG protein from previous structural studies was also referred to as Δ79AAG (9, 10). However, in those former studies, all residues contained in the construct are of the wild-type sequence. For the creation of Δ79AAG-N169L and Δ79AAG-N169A mutants, PCR-based site-directed mutagenesis was performed with the primers shown in supplemental Table S1, and successful incorporation was confirmed by DNA sequencing (MWG Biotech). Δ79AAG protein expression and purification were done similarly to previously described protocols (11, 18). For a detailed procedure, refer to the supplemental Experimental Procedures.

Gel Mobility Shift Assays—The binding affinity of Δ79AAG to different DNA oligonucleotides was measured using gel mobility shift assays. DNA oligonucleotides were ^{32}P-labeled and purified using Sephadex G-25 quick spin columns (detailed protocol in the supplemental Experimental Procedures). Binding reactions were set up as 10-μl solutions containing 1× binding buffer (50 mM Hepes-NaOH, pH 7.5, 100 mM NaCl, 1 mM EDTA, 9.5% v/v glycerol, 50 μg/ml BSA, and 5 mM DTT), 2 nM ^{32}P-labeled oligonucleotide, and 0–1000 nM of the purified Δ79AAG protein. The reaction samples were incubated on ice for 30 min, and the products were resolved using 6% native-PAGE in 0.5× Tris-borate-EDTA buffer at 110 V for 3 h at 4 °C. The extent of complex formation was quantified and analyzed by phosphorimaging. Dissociation constant (K_d) values were calculated by fitting the binding data to the binding equation (Equation 1) using GraphPad Prism (GraphPad software, Inc.)

$$Y = B_{max}{*}X/(K_d + X) \quad \text{(Eq. 1)}$$

where Y is the specific binding, B_{max} is the maximal binding, X is the concentration of protein, and K_d is the dissociation constant. Each experiment was repeated at least three times, and the data represent the average of at least three independent experiments.

DNA Glycosylase Assays—DNA glycosylase assays were set up as 10-μl solutions containing 1× glycosylase assay buffer (50 mM Tris-HCl, pH 7.8, 100 mM NaCl, 1 mM EDTA, 50 μg/ml BSA, and 5 mM DTT), 2 nM ^{32}P-labeled oligonucleotide, and 25 nM of either the purified full-length enzyme or the Δ79AAG enzyme. The reactions were carried out at 37 °C. Aliquots (10 μl) from particular time points were mixed with piperidine to the final concentration of 0.2 M and heated at 75 °C for 15 min. The piperidine treatment cleaves all abasic (AP) sites resulting in single strand breaks at the region of AP sites. This procedure was followed by the addition of one sample volume of 90% formamide buffer with dye markers. The samples were heated at 75 °C for 15 min, and the products were resolved using 20%

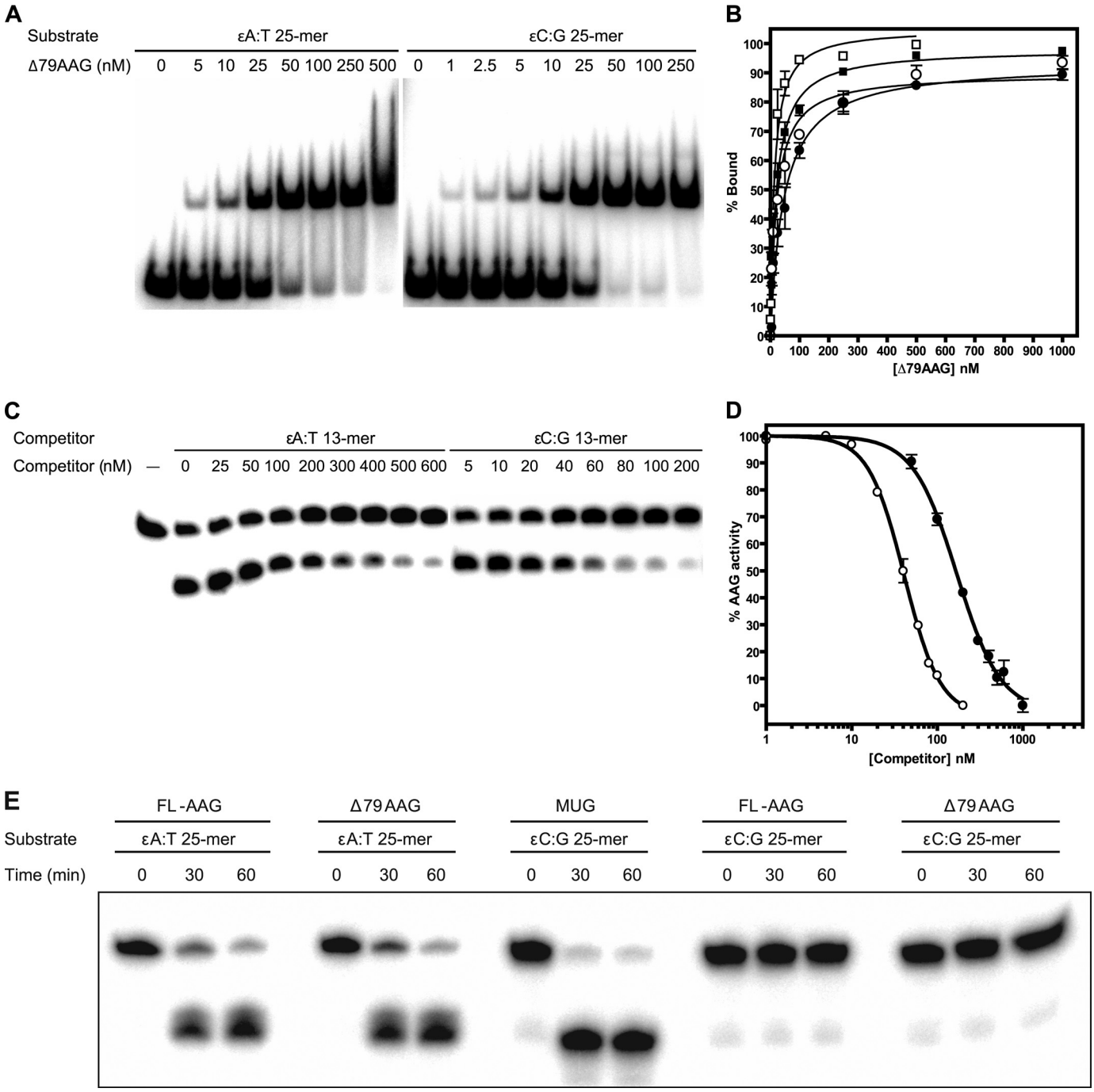

FIGURE 1. **Biochemical characterization of AAG variants with oligomers containing etheno lesions.** *A*, representative gels showing the results of gel mobility shift assays for Δ79AAG binding to ϵA:T (ϵA paired opposite T) and ϵC:G (ϵC paired opposite G) 25-mer DNA duplexes. 2 nM ^{32}P-labeled oligonucleotide was incubated with an indicated concentration of Δ79AAG in the 1× binding buffer, and the resulting protein-DNA complexes were resolved using 6% native-PAGE. The *top band* corresponds to the protein-DNA complex, and the *bottom band* corresponds to the free DNA. *B*, graphical representation of Δ79AAG binding to ϵA:T 25-mer (■), ϵC:G 25-mer (□), ϵA:T 13-mer (●), and ϵC:G 13-mer (○) oligonucleotide duplexes as measured by gel shift mobility assays. *Error bars* indicate the S.D. of at least three independent trials. *C*, representative gels showing the results from competition DNA glycosylase assays. The activity of Δ79AAG on labeled ϵA:T 25-mer duplex is inhibited in the presence of indicated concentrations of ϵA:T 13-mer and ϵC:G 13-mer unlabeled competitors. 2 nM ^{32}P-labeled ϵA:T 25-mer duplex oligonucleotide was incubated in 1× glycosylase assay buffer with 5 nM Δ79AAG enzyme and increasing concentrations of cold competitors at 37 °C for 30 min. Following hot alkali treatment, the products were resolved using 20% denaturing urea-PAGE. The *top band* represents the uncleaved DNA substrate, and the *bottom band* represents the cleaved product. *D*, graphical representation of the data from competition DNA glycosylase assays, looking at the inhibition of Δ79AAG activity on labeled ϵA:T 25-mer duplex substrate by ϵA:T 13-mer (●) and ϵC:G 13-mer (○) unlabeled competitors. *Error bars* indicate the S.D. of at least three independent trials. *E*, gel results of DNA glycosylase assays for full-length AAG (*FL-AAG*) and truncated Δ79AAG on ϵA:T and ϵC:G 25-mer DNA duplexes with *E. coli* mismatch uracil DNA glycosylase (*MUG*) as a positive control. Single strand breaks are observed in this urea-denaturing gel when piperidine treatment cleaves AP sites (*bottom band*). When AAG fails to cleave the lesion base, no AP site is formed, and the oligomer remains intact (*top band*).

denaturing urea-PAGE in 1× Tris-borate-EDTA buffer at 450 V for 2 h. The extent of substrate cleavage was quantified and analyzed by phosphorimaging.

Competition DNA Glycosylase Assays—Competition DNA glycosylase assays were performed to measure the inhibition of Δ79AAG activity on ϵA-containing duplex DNA substrate by ϵA and ϵC duplexes. The reactions were set up as 20-μl solutions containing 1× glycosylase assay buffer, 1 nM ^{32}P-labeled ϵA:T (T paired opposite ϵA) 25-mer oligonucleotide duplex DNA (5′-GCA ATC TAG CCA ϵAGT CGA TGT ATG C-3′), 5 nM of the purified Δ79AAG enzyme, and increasing concentrations of competitor DNA (0–3000 nM). The reactions were carried out at 37 °C for 30 min. After incubation, NaOH was added to a final concentration of 0.2 M followed by heating at 75 °C for 15 min. Similar to piperidine treatment, hot alkali treatment with NaOH cleaves all AP sites and creates DNA single strand breaks at the AP sites. Upon cooling, one sample volume of 90% formamide buffer with dye markers was added into the reaction mixture. The samples were heated at 75 °C for 15 min before loading, and the products were resolved using 20% denaturing urea-PAGE in 1× Tris-borate-EDTA buffer at 450 V for 2 h. The extent of substrate cleavage was quantified and analyzed by phosphorimaging. The experiment with each competitor was repeated at least three times. To calculate the IC_{50} (50% inhibitory concentration), the competition data were fitted to the sigmoidal dose-response curve (Equation 2) using GraphPad Prism

$$Y = Y_{min} + (Y_{max} - Y_{min})/(1 + 10^{LogIC_{50} - X}) \quad \text{(Eq. 2)}$$

where X is the logarithm of competitor concentration, Y_{max} and Y_{min} are the maximum and minimum values of the percentage of AAG activity (Y), and $LogIC_{50}$ is the logarithm of IC_{50}.

Crystallographic Procedures—Δ79AAG was crystallized in the presence of pre-annealed ϵC-containing 13-mer DNA. Crystals were grown with a precipitant of 100 mM sodium cacodylate, pH 6.0, 200 mM manganese chloride, and 20% polyethylene glycol-3350. The x-ray diffraction data were collected at the Advanced Light Source on beamline 12.3.1 at 100 K to 2.2 Å resolution. The structure of the Δ79AAG-ϵC:G (ϵC paired opposite G) complex, with two molecules in the asymmetric unit, was determined by molecular replacement using PHASER (19) and the Δ79AAG protein coordinates from the Δ79AAG-Pyr:T complex structure (Protein Data Bank (PDB) ID 1BNK) as a search model (9). The models were further refined in CNS (20) and Refmac 5.4 (21, 22). For a more detailed description, see the supplemental Experimental Procedures.

The final model of the Δ79AAG-ϵC:G complex converged at an R factor of 23.9 (R_{free} = 28.4). For residues 80–298, the following have no electron density and are therefore not included in the model: the residues 203–207, 265–268, and 295–298 in chain A and the residues 205–206, 265–266, and 294–298 in chain B. Due to a lack of interpretable electron density for the side chains of some residues in the structure (Arg-201, Leu-249, and Glu-253 in chain A), these residues were modeled as alanines. Each protein molecule in the asymmetric unit has a 13-mer DNA duplex associated with it. One nucleotide of each of these duplexes is disordered (A26).

TABLE 1

Dissociation constant (K_d) values measured using gel shift assays and 50% inhibitory concentration (IC_{50}) for the inhibition of Δ79AAG activity on ϵA:T 25-mer, measured using competition DNA glycosylase assay at 37 °C, in the presence of increasing concentration of cold competitor 13-mer duplexes

Δ79AAG	Oligonucleotide	K_d[a] ± S.D.	IC_{50}[a] (95% confidence intervals)
		nM	nM
Wild type	ϵA:T 25-mer	20 ± 2	
Wild type	ϵC:G 25-mer	13 ± 2	
Wild type	ϵA:T 13-mer	46 ± 6	163 (152–174)
Wild type	ϵC:G 13-mer	21 ± 3	39 (38–41)
N169L	ϵC:G 25-mer	31 ± 4	
N169A	ϵC:G 25-mer	47 ± 6	

[a] Average of at least three independent experiments.

TABLE 2

Data collection and refinement statistics of the Δ79AAG-DNA complex

	Δ79AAG-ϵC:G
Data collection	
Space group	P1
Cell constants	
a, b, c	a = 41.23 Å, b = 41.22 Å, c = 82.14 Å
α, β, γ	α = 81.23°, β = 88.4°, γ = 89.15°
Beamline	12.3.1
Wavelength (Å)	1.116
Resolution (Å)[a]	41-2.20 (2.28-2.20)
No. of total observations	71943
No. of unique observations	26278
Completeness (%)[a]	96.6 (95.2)
$\langle I/\sigma(I)\rangle$[a]	17.1 (4.9)
R_{sym} (%)[a,b]	5.8 (19.2)
Model refinement	
R_{work} (%)[c]	23.9
R_{free} (%)[c]	28.4
B-factors (Å^2)	
Protein	15.2
DNA	17.9
Water	14.1
Mn^{2+} ion	Chain A: 27.9 chain B: 44.0
r.m.s.d. bonds (Å)	0.007
r.m.s.d. angles (°)	1.1
Number of atoms	
Protein	3257 (2 molecules/asu)[d]
DNA	1024
Water	232
Mn^{2+} ion	2
Ramachandran plot (%)	
Most favored	89.7
Additionally allowed	10.0
Generously allowed	0.3

[a] Values in parentheses correspond to the highest resolution shell.
[b] $R_{sym} = \sum|I_{hkl} - \langle I_{hkl}\rangle|/\sum I_{hkl}$, where I is the intensity of a reflection hkl and ⟨I⟩ is the average over symmetry-related reflections of hkl.
[c] $R_{work} = \sum||F_o| - |F_c||/\sum|F_o|$ in which F_o and F_c are the observed and calculated structure factor amplitudes, respectively. R_{free} is calculated from 5% of the reflections not used in the model refinement.
[d] asu, asymmetric unit.

RESULTS

AAG Binding Studies—The binding affinity of Δ79AAG to the ϵA:T (ϵA paired opposite T) lesion containing 25-mer and 13-mer duplexes was compared with the binding affinity to ϵC:G (ϵC paired opposite G) duplexes, using gel mobility shift assays (Fig. 1, *A* and *B*; and Table 1). As shown in Table 1, when either ϵA or ϵC lesions are present in a given sequence context, Δ79AAG consistently binds the ϵC:G duplex with ~2-fold higher affinity as compared with that of the ϵA:T duplex. In addition, Δ79AAG binds the ϵA:T 25-mer duplex (K_d = 20 ± 2 nM) with ~2-fold higher affinity as compared with the ϵA:T 13-mer duplex used for crystallization (K_d = 46 ± 6 nM). Cor-

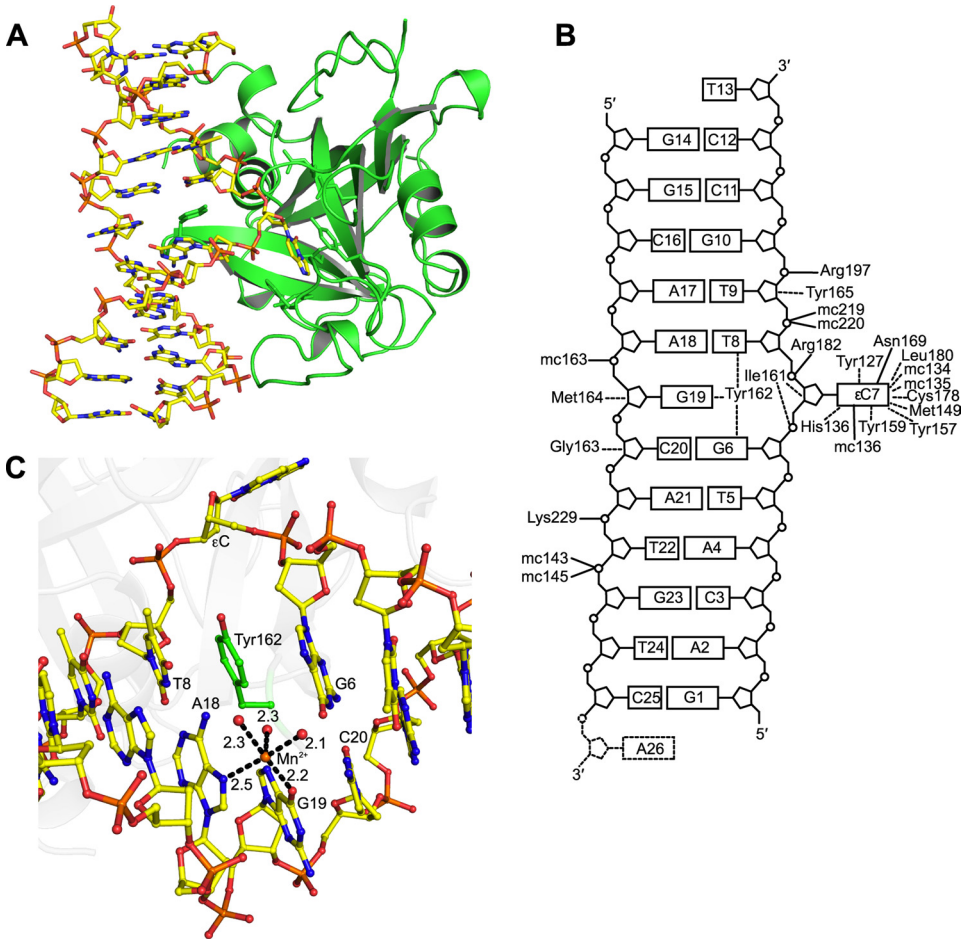

FIGURE 2. **Structure of Δ79AAG with ϵC inhibitor DNA.** *A*, overall structure of the Δ79AAG-ϵC:G complex. The protein is in *green ribbons* with Tyr-162 in *stick form*. The DNA is colored as follows: *yellow*, carbon; *red*, oxygen; *blue*, nitrogen; and *orange*, phosphorus. *B*, schematic illustration of the interactions between the amino acid side chains (three-letter code) and main chains (*mc*) of the AAG protein and ϵC-containing DNA in the structure of the Δ79AAG-ϵC:G complex. Hydrogen bonds are indicated by a *solid line*, and van der Waals interactions are indicated by a *dashed line*. The disordered A26 nucleotide is in *dashed lines*. *C*, intercalation of Tyr-162 (carbons in *green*) in the Δ79AAG-ϵC:G structure and coordination of metal ion Mn^{2+} (*orange sphere*) to the A18 and G19 of the DNA strand opposite the ϵC lesion (carbons in *yellow*) and to three water molecules (*red spheres*). Distances (*dashed lines*) are measured in angstroms. The electron density for the metal ion site is shown in supplemental Fig. S4*A*. Other non-carbon atoms are colored as in *A*.

respondingly, Δ79AAG also binds the ϵC:G 25-mer duplex ($K_d = 13 \pm 2$ nM) with ~2-fold higher affinity as compared with the ϵC:G 13-mer duplex ($K_d = 21 \pm 3$ nM). These results indicate that the binding affinity of Δ79AAG to the DNA containing the same lesion varies depending on the length of the DNA duplex. The binding studies also show that in a given sequence context, Δ79AAG binds ϵC:G duplex with higher affinity as compared with that of ϵA:T duplex.

Catalytic Ability of AAG for ϵC-containing DNA—Following our binding studies, we tested the DNA glycosylase activity of both full-length AAG and Δ79AAG on ϵA and ϵC residues present in the 25-mer oligonucleotide duplexes (Fig. 1*E*). As shown in the representative gel, both full-length AAG and Δ79AAG robustly removed ϵA from the ϵA:T 25-mer duplex. However, the activity of both full-length AAG and Δ79AAG on ϵC residues from the same duplex was completely absent. In contrast, the positive control, *Escherichia coli* mismatch uracil DNA glycosylase (MUG) (Trevigen, Inc.), shows robust catalytic activity on ϵC contained in an ϵC:G 25-mer duplex (Fig. 1*E*). Further, we tested the activity of Δ79AAG on ϵC-containing 13-mer oligonucleotide duplexes used for crystallization, with ϵC paired opposite different bases. The results show that Δ79AAG does not have activity on ϵC:G, ϵC:A, ϵC:C, or ϵC:T 13-mer duplexes (supplemental Fig. S1).

Inhibition of AAG by ϵC-containing DNA—We employed competition assays to measure the inhibition of Δ79AAG activity on ϵA:T 25-mers using ϵA:T and ϵC:G 13-mer competitor DNA oligonucleotides (Fig. 1, *C* and *D*; and Table 1). The catalytic activity of Δ79AAG on labeled ϵA:T 25-mer duplex was measured at 37 °C in the presence of an increasing concentration of the aforementioned cold competitors (Fig. 1*C*). The cleavage data were fitted to Equation 2 (see "Experimental Procedures") to calculate the 50% inhibitory concentration (IC_{50}) for each cold competitor used in the experiment (Fig. 1*D*). The results obtained correlate with the binding measurements described above. As shown in Table 1, the ϵA:T 13-mer DNA duplex binds ~4-fold weaker (IC_{50} = 163 nM) than the ϵC:G 13-mer (IC_{50} = 39 nM).

Overall Structure of the Δ79AAG-ϵC DNA Inhibitor Complex—The crystal structure of Δ79AAG-ϵC:G at 2.2 Å resolution was determined by molecular replacement using the structure of Δ79AAG-Pyr:T complex as a search model (9). Dif-

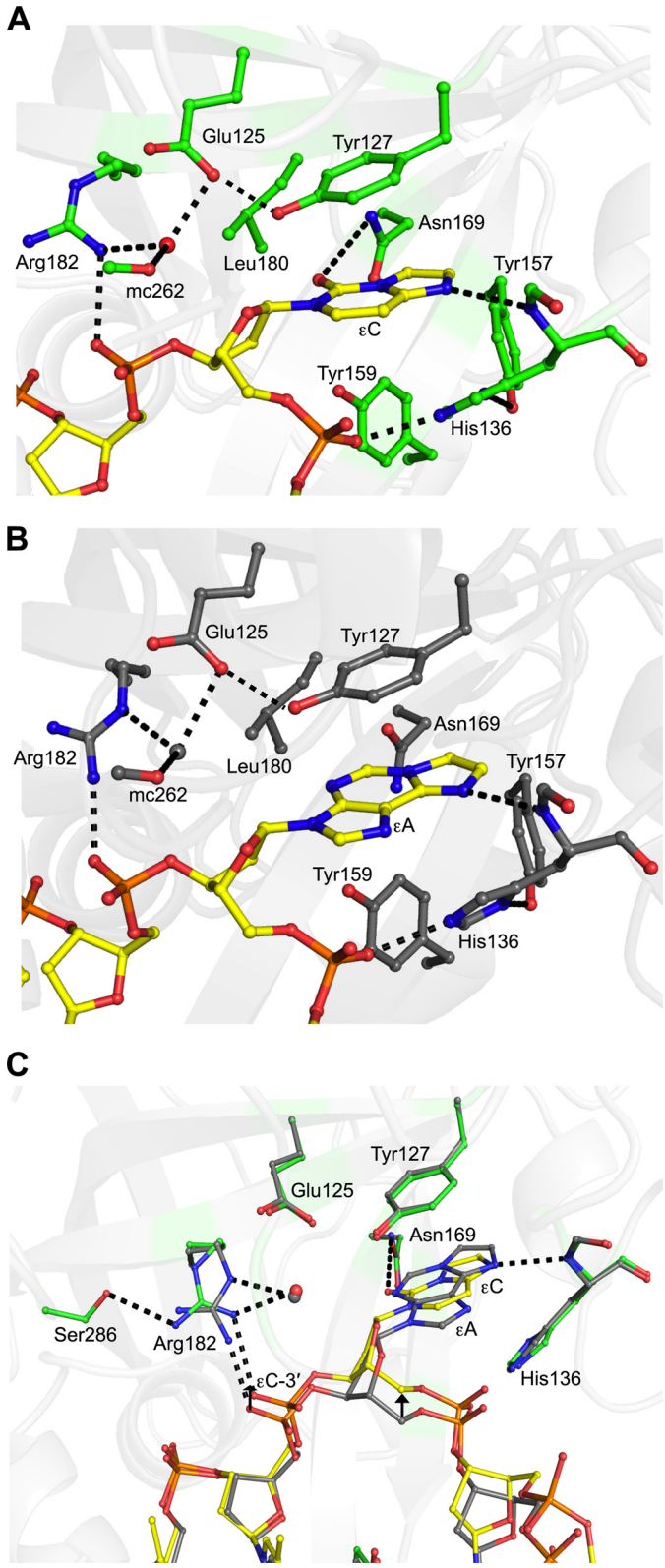

FIGURE 3. **Active site architecture of AAG.** *A*, active site of the Δ79AAG-ϵC:G inhibitor complex with DNA and amino acid carbons in *yellow* and *green*, respectively, and a putative catalytic water molecule as a *red sphere*. Hydrogen bonds are indicated as *dashed lines*. A stereo view of the electron density of this interaction is shown in supplemental Fig. S6. *mc*, main chain. *B*, active site architecture of the Δ79AAG-ϵA:T substrate complex (carbons and catalytic water colored *gray*) showing the amino acids interacting with a flipped ϵA base and the DNA backbone (carbons in *yellow*) (PDB ID 1F4R (10)). *C*, comparison of the active site architecture of Δ79AAG-ϵC:G (protein carbons colored *green*, DNA carbons colored *yellow*, and catalytic water colored *red*) and Δ79AAG-ϵA:T (all carbon atoms and catalytic water colored *gray*). The movement of the DNA backbone in the Δ79AAG-ϵC:G complex due to the binding of ϵC is indicated by *black arrows*. All non-carbon atoms are colored as in Fig. 2.

ference electron density maps calculated in the absence of DNA show interpretable electron density for DNA backbone and for the ϵC base in the active site pocket of AAG. In comparison with the previously published Δ79AAG-DNA complexes (9, 10), the Δ79AAG-ϵC inhibitor complex crystallized in a different space group (P1). The final model of Δ79AAG-ϵC:G has been refined to the *R* factor of 23.9 (R_{free} = 28.4) (Table 2).

The overall structure of Δ79AAG bound to an ϵC DNA lesion includes the insertion of Tyr-162 into the DNA and the flip of the ϵC nucleotide into the enzyme active site (Fig. 2). The two copies of this inhibitor-protein complex in the asymmetric unit are quite similar to each other, with a root mean square deviation (r.m.s.d.) for all α-carbon atoms of ~0.6 Å, and to that of the Δ79AAG-ϵA:T substrate complex (10), also having an r.m.s.d. of ~0.6 Å. With the exception of some loops and two disordered regions (supplemental Fig. S2), the only notable difference in our inhibitor structure is that the octahedrally coordinated Na^+ metal ion modeled in the structure of the Δ79AAG-ϵA:T complex was found to be absent. Instead, this site is occupied by the N-terminal amino group of residue 80 of a symmetrically equivalent Δ79AAG molecule (supplemental Fig. S3). Similar to the Na^+ ion, the N-terminal amine in this position also interacts with the main chain carbonyls of Met-149, Gly-174, Ala-177, and Ser-171 and the side chain of Ser-172.

Protein-DNA Interactions—In both molecules in the asymmetric unit of the Δ79AAG-ϵC:G structure, the DNA duplex is largely B-form and is bent away from the protein by about 22°, with the bend primarily centered on the flipped ϵC nucleotide (Fig. 2*A*). For the most part, all protein-DNA interactions are similar between the Δ79AAG-ϵC:G inhibitor complex and the Δ79AAG-ϵA:T substrate complex (Fig. 2 and supplemental Fig. S2). Tyr-162 makes the most important protein contact, given that it is inserted into the DNA duplex, replacing the ϵC lesion, and forming van der Waals contacts with the opposite base, G19 (Fig. 2). A potential steric clash of Tyr-162 with G19 is prevented by shifting this opposite base out of the minor groove, leaving it without a base-pairing partner (Fig. 2, *B* and *C*). This sliding rearrangement, in which G19 is orphaned in terms of hydrogen bonding, is disruptive and presumably has an energetic penalty associated with it. The deoxyribose of the orphaned base does make favorable van der Waals interactions with Met-164 (Fig. 2*B* and supplemental Fig. S4*C*), counteracting some of these energetic costs. As was the case in the Δ79AAG-ϵA:T substrate complex, no direct hydrogen-bonding interactions are present between the protein and the base opposite the DNA lesion, indicating that specific recognition of this base is not required for the binding of AAG to DNA.

Metal Ion Mn^{2+} in the Δ79AAG-ϵC:G Structure—The Δ79AAG-ϵC:G complex was crystallized in the presence of 200 mM $MnCl_2$, a condition under which the ability of AAG to excise ϵA from a DNA duplex is impaired (12, 13). We find electron density consistent with a metal ion near the base oppo-

site the ϵC lesion (Fig. 2*C* and supplemental Fig. S4), at a distance of ~16 Å to the AAG active site (C1′ of ϵC). Mn^{2+} refines well in this electron density with no positive or negative difference electron density. In contrast, refinement of a water molecule or a sodium ion (also present in the crystallization buffer) leads to positive difference electron density, suggesting that the correct atom in this site is heavier than water and sodium, consistent with Mn^{2+}. Anomalous difference density is also present at both sites in the asymmetric unit at approximate σ levels of 8 and 5 for chain A and chain B, respectively, consistent with the presence of Mn^{2+} ions (supplemental Fig. S4*A*). At the wavelength of data collection ($\lambda = 1.116$ Å), sodium would not give rise to an anomalous signal, ruling out the other metal ion contained in the crystallization buffer as being present in this site. The refined Mn^{2+} is coordinated to the O^6 of G19 (the base opposite ϵC), to the N7 of A18, and to three water molecules (Fig. 2*C*). The coordination of Mn^{2+} appears to induce a significant change in the phosphodiester backbone in the region around G19, resulting in a change in the sugar pucker to a C2′-exo conformation (supplemental Fig. S4*B*).

Active Site Architecture of Δ79AAG-ϵC DNA Complex—The ϵC base lesion is recognized and stabilized by hydrogen bonds, along with van der Waals interactions (Fig. 3 and supplemental Fig. S5). Similar to the recognition of ϵA (Fig. 3*B*), ϵC is stacked between Tyr-127 on one side and His-136 and Tyr-159 on the other side (Fig. 3*A*). Tyr-159 makes edge-to-face contact with the lesion base. The specificity to discriminate between the ϵC lesion and undamaged cytosine appears to be achieved through a hydrogen bond donated from the main chain amide of His-136 to the N^4 of the ϵC base. This interaction is similar to the damage-specific recognition of ϵA (*versus* undamaged adenine) by AAG, which is made through a hydrogen bond donated by the main chain amide of His-136 to the N^6 of ϵA. An additional hydrogen bond is also observed between the carboxamide nitrogen of the side chain of Asn-169 and the O^2 of ϵC. Mutants of Asn-169, namely N169L and N169A, show respective ~2- and ~4-fold reduced affinity for the ϵC:G 25-mer oligonucleotides as compared with that of wild-type Δ79AAG (Table 1), indicating that the additional hydrogen bond donated from Asn-169 to ϵC contributes to the increased affinity of Δ79AAG for ϵC inhibitor DNA. Interestingly, in the Δ79AAG-ϵC:G inhibitor complex, the putative catalytic water molecule (proposed to act as a nucleophile) occupies the same position as in the Δ79AAG-ϵA:T substrate complex (Fig. 3). Also in agreement with the structure of the Δ79AAG-ϵA:T substrate complex, this water molecule interacts with Glu-125, Arg-182, and main chain carbonyl oxygen of Val-262. Glu-125, which is proposed to activate the water molecule for nucleophilic attack (10, 11), is held in its position through a hydrogen bond donated from Tyr-127 to its carboxyl group. The side chain of Arg-182 further contacts the 3′-phosphate of the ϵC nucleotide.

In comparison with the recognition of the ϵA substrate, the ϵC inhibitor recognition induces slight changes in the active site, mostly with respect to DNA backbone positions (Fig. 3*C*). Because ϵC is smaller than ϵA, the DNA backbone must be pulled farther into the active site such that ϵC is recognized with optimal molecular interactions. In a modeling exercise in which the DNA backbone is held rigid, AAG fails to form optimal hydrogen bonding interactions between both the main chain amide of His-136 and the N^4 of ϵC and the side chain of Asn-169 and the O^2 of ϵC. Perhaps as a result of this slight repositioning of the DNA backbone, the side chain of Arg-182 adopts a different conformation (Fig. 3*C*). The side chain of Arg-182 still interacts with the 3′-phosphate of the ϵC nucleotide and still hydrogen-bonds to the putative catalytic water molecule. Overall, however, comparison of the active site of the Δ79AAG-ϵC inhibitor complex with that of the Δ79AAG-ϵA:T substrate complex shows that with the exception of Arg-182, all the residues involved in lesion recognition and catalysis maintain similar orientations.

DISCUSSION

AAG plays an important role in the maintenance of genomic integrity, presumably through its ability to recognize, bind, and excise a wide range of DNA base lesions. It was therefore surprising that AAG also has the ability to recognize and bind a number of DNA base lesions that it is incapable of excising, in particular the ϵC lesion. Moreover, the tight binding of AAG to ϵC leads to the inhibition of its catalytic activity and, in addition, is known to shield ϵC from ABH2-mediated direct reversal repair.[6] To understand the structural basis for the inhibition of AAG by ϵC-containing DNA, we solved the crystal structure of Δ79AAG bound to a 13-mer ϵC:G (ϵC paired opposite G) duplex.

Given that AAG can bind ϵC-containing DNA, we anticipated that the lack of activity may be a result of one or more of the following factors. (i) AAG might fail to flip ϵC into its active site, or it might flip ϵC into an alternative binding pocket that lacks the appropriate catalytic residues; (ii) the binding mode of ϵC in the active site might not favor accommodation of the water molecule thought to act as a nucleophile in the reaction; (iii) the side chain of the putative catalytic base (Glu-125) might adopt a non-productive conformation that fails to activate the putative catalytic water molecule; (iv) AAG might be unable to protonate ϵC, failing to activate it for departure.

The crystal structure shows that AAG successfully flips the ϵC inhibitor into the same active site pocket that binds the ϵA substrate, ruling out the first possibility. We also find that the putative catalytic water molecule is present in the inhibitor complex, ruling out the second possibility. Furthermore, as was observed in the structure of the Δ79AAG-ϵA:T substrate complex, this water molecule is in contact with Glu-125, as would be required for its activation. The water molecules are ideally positioned to attack the *N*-glycosidic bond. Thus, we infer that the inability of AAG to remove ϵC is unlikely to be due to a problem with nucleophilic activation or attack, ruling out the third possibility from our list.

We next examined the fourth possibility, that the failure to excise ϵC is due to a problem with leaving group activation. In a previous biochemical study, O'Brien and Ellenberger (11) measured the pH rate profiles for the excision of neutral Hx and ϵA lesions by AAG and for its excision of the positively charged 7-meG lesion under single turnover conditions. They found that the pH rate profiles for ϵA and Hx excision follow a bell-

[6] D. Fu and L. D. Samson, submitted for publication.

shaped curve, indicating that for the excision of neutral lesions, AAG uses the action of both a general acid and a general base (Glu-125). The general base can activate a catalytic water molecule, whereas the general acid is expected to facilitate the protonation of neutral lesions, making the lesion base a better leaving group (11). In contrast, the pH rate profile for the excision of 7-meG shows only a single ionization corresponding to a general base, suggesting that leaving group activation of 7-meG is not necessary because the base is already positively charged. To help pinpoint the site of protonation, the activity of AAG on Hx was compared with its activity on 7-deaza-Hx, and although AAG greatly enhances the rate of Hx excision ($\sim 10^8$), the same lesion with N7 changed to C7 is not cleaved by AAG, directly implicating the involvement of the N7 position in catalysis (see Fig. 4*B* for numbering) (11). Although this study was unable to identify a specific residue as the general acid, the crystal structure of a Δ79AAG(E125Q)-ϵA:T substrate complex shows a water molecule in contact with the equivalent position to N7 of Hx, that is, N7 of ϵA (Fig. 4*A*) (10), raising the possibility that a protein-bound water molecule could be responsible for protonation. Once protonated, the AAG active site is designed to stabilize the protonated form of the base through a hydrogen bond between N7H of ϵA and the backbone carbonyl oxygen of Ala-134 (Fig. 4*A*). Given these findings on the catalytic significance of protonation of the N7 of Hx and ϵA, it is important to consider the equivalent position in the ϵC base. A superposition shows that unlike ϵA, ϵC has a carbon (C5) in the position equivalent to N7 and thus cannot be protonated at that site (Fig. 4*A*). Therefore, as opposed to our findings with respect to possibilities one through three, it appears that the failure of AAG to cleave ϵC could be due to an inability to activate the ϵC leaving group by protonation. Because AAG is reported to bind and not cleave a number of different pyrimidine lesions, including 3-methyluracil, 3-ethyluracil, and 3-methylthymine (15), this mechanism of inhibition may be broadly applicable.

Given that AAG cannot repair ϵC lesions, it is interesting that AAG binds this lesion so tightly. The molecular basis for the approximate 2-fold higher affinity of AAG for the ϵC:G duplex (as compared with the substrate ϵA:T duplex) can be attributed to an additional hydrogen bond formed between the carboxamide side chain of Asn-169 and the O^2 of ϵC. Mutation of Asn-169 to residues that cannot maintain this hydrogen bond (Leu and Ala) completely abolished this 2-fold binding effect (Table 1), suggesting that this one hydrogen bond is chiefly responsible for the higher affinity of ϵC-containing DNA. Thus, in addition to its previously proposed role of serving to help discriminate between damaged and undamaged guanine (24), Asn-169 appears to play a role in the recognition and binding of pyrimidine DNA lesions. It is important to note that Asn-169 is strictly conserved among AAG-related glycosylases (9).

Although inhibition of AAG by divalent metal ions (Mg^{2+}, Mn^{2+}, Zn^{2+}, Ca^{2+}, Cd^{2+}, and Ni^{2+}) has been well documented (12, 13), no previous crystal structure displayed electron density consistent with such an ion, although these cations were used in the crystallization buffers (9, 10). Here we find density consistent with the presence of Mn^{2+} in position to coordinate the base opposite the ϵC lesion (G19) (Fig. 2*C* and supplemental Fig. S4). Binding of Mn^{2+} to this site

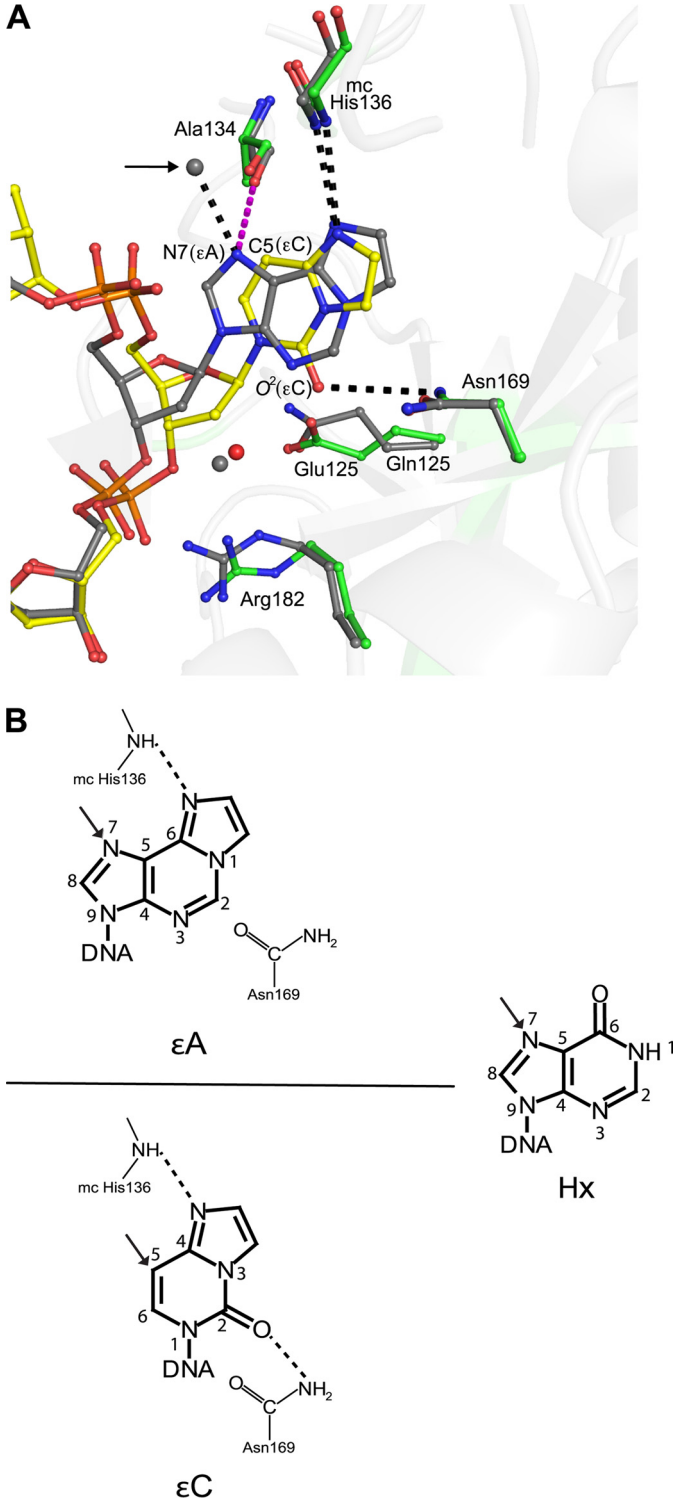

FIGURE 4. **Activation of the leaving group by protonation.** *A*, superposition of the structure of the Δ79AAG-ϵC:G inhibitor complex (amino acid carbons in *green*; DNA carbons in *yellow*; and water as a *red sphere*) with that of the Δ79AAG-E125Q-ϵA:T substrate complex (amino acid carbons, DNA carbons, and waters are in *gray*) (PDB ID 1EWN (10)). All other non-carbon atoms are colored as in Fig. 2. The hydrogen bonding contacts are indicated as *black dashed lines*. A *magenta dashed line* represents a putative hydrogen bonding contact that could stabilize a protonated state of ϵA. An *arrow* indicates a water molecule that could protonate N7 of ϵA. *mc*, main chain. *B*, schematic illustration of the recognition of ϵA (*top*) and ϵC (*bottom*) by active site residues and of the structure of Hx (*right*) with each base numbered to correlate with the text. *Arrows* indicate the proposed protonation site on ϵA (N7) (*top*) and the corresponding position in ϵC (C5) and Hx (N7).

appears to influence the pucker of the sugar, yielding a C2′-exo configuration (supplemental Fig. S4*B*). This occurrence is the first time that this sugar pucker has been observed in an AAG structure. In our Mn^{2+}-containing structure, the binding of ϵC to the active site is nearly identical to that observed in the Δ79AAG-ϵA:T structure. Thus, the inhibitory effect of Mn^{2+} does not appear to be due to a large conformational change in the active site, although the electrostatics of the active site could be affected by the presence of the positively charged ion ~16 Å away. Other dynamic movements of protein or DNA that are important in catalysis might also be affected by the Mn^{2+} coordination. Although binding of divalent metal cations to protein and DNA is common, it is intriguing that we find this divalent metal bound to such an important site in this protein-DNA complex. We now have a physical model for the influence of divalent cations on AAG activity that can be tested.

With a molecular view of the AAG-ϵC abortive complex in hand, it is interesting to consider what the physiological benefits in forming this complex might be. Abortive complexes between alkyltransferase-like (ATL) proteins and a lesion that it cannot excise, O^6-alkylguanine, have recently been observed (25). ATLs are known to interact with proteins in another DNA repair pathway, nucleotide excision repair, suggesting that ATLs may function to present alkylated DNA to nucleotide excision repair proteins for repair (25, 26). Given the difficulty of finding a single damaged DNA base in the midst of the genome, it makes sense that once found, some DNA repair proteins may be designed to "hand over" lesion bases that they cannot themselves repair to an alternative repair pathway. Although it is too early to infer directly from ATLs, preliminary data do show that AAG can interact with human nucleotide excision repair proteins hHR23A and hHR23B (27). Obviously, more studies are necessary to establish whether AAG-ϵC abortive complexes interact with nucleotide excision repair proteins, resulting in ϵC repair, but this idea is intriguing and, given the recent studies on ATL, not without precedence. Originally identified by its ability to excise alkylated 3-methyladenine and 7-meG lesions (8, 23), the role of AAG in DNA repair is far more complex than once thought. Given the importance of repairing reactive oxygen and nitrogen species-generated DNA damage for tissues undergoing chronic inflammation, a complete understanding of the physiological function of AAG is essential.

Acknowledgments—The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the United States Department of Energy under Contract Number DE-AC02-05CH11231.

REFERENCES

1. Coussens, L. M., and Werb, Z. (2002) *Nature* **420,** 860–867
2. Wiseman, H., and Halliwell, B. (1996) *Biochem. J.* **313,** 17–29
3. Nair, U., Bartsch, H., and Nair, J. (2007) *Free Radic. Biol. Med.* **43,** 1109–1120
4. Bartsch, H., and Nair, J. (2002) *Cancer Detect. Prev.* **26,** 308–312
5. Gros, L., Ishchenko, A. A., and Saparbaev, M. (2003) *Mutat. Res.* **531,** 219–229
6. Levine, R. L., Yang, I. Y., Hossain, M., Pandya, G. A., Grollman, A. P., and Moriya, M. (2000) *Cancer Res.* **60,** 4098–4104
7. Zharkov, D. O. (2008) *Cell Mol. Life Sci.* **65,** 1544–1565
8. Wyatt, M. D., Allan, J. M., Lau, A. Y., Ellenberger, T. E., and Samson, L. D. (1999) *BioEssays* **21,** 668–676
9. Lau, A. Y., Schärer, O. D., Samson, L., Verdine, G. L., and Ellenberger, T. (1998) *Cell* **95,** 249–258
10. Lau, A. Y., Wyatt, M. D., Glassner, B. J., Samson, L. D., and Ellenberger, T. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97,** 13573–13578
11. O'Brien, P. J., and Ellenberger, T. (2003) *Biochemistry* **42,** 12418–12429
12. Adhikari, S., Toretsky, J. A., Yuan, L., and Roy, R. (2006) *J. Biol. Chem.* **281,** 29525–29532
13. Wang, P., Guliaev, A. B., and Hang, B. (2006) *Toxicol. Lett.* **166,** 237–247
14. Gros, L., Maksimenko, A. V., Privezentzev, C. V., Laval, J., and Saparbaev, M. K. (2004) *J. Biol. Chem.* **279,** 17723–17730
15. Lee, C. Y., Delaney, J. C., Kartalou, M., Lingaraju, G. M., Maor-Shoshani, A., Essigmann, J. M., and Samson, L. D. (2009) *Biochemistry* **48,** 1850–1861
16. 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) *J. Clin. Invest.* **112,** 1887–1894
17. O'Brien, P. J., and Ellenberger, T. (2004) *J. Biol. Chem.* **279,** 9750–9757
18. Klapacz, J., Lingaraju, G. M., Guo, H. H., Shah, D., Moar-Shoshani, A., Loeb ,L. A., and Samson, L. D. (2010) *Mol. Cell* **37,** 843–853
19. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) *J. Appl. Crystallogr.* **40,** 658–674
20. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54,** 905–921
21. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763
22. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255
23. Samson, L., Derfler, B., Boosalis, M., and Call, K. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88,** 9127–9131
24. Connor, E. E., and Wyatt, M. D. (2002) *Chem. Biol.* **9,** 1033–1041
25. Tubbs, J. L., Latypov, V., Kanugula, S., Butt, A., Melikishvili, M., Kraehenbuehl, R., Fleck, O., Marriott, A., Watson, A. J., Verbeek, B., McGown, G., Thorncroft, M., Santibanez-Koref, M. F., Millington, C., Arvai, A. S., Kroeger, M. D., Peterson, L. A., Williams, D. M., Fried, M. G., Margison, G. P., Pegg, A. E., and Tainer, J. A. (2009) *Nature* **459,** 808–813
26. Margison, G. P., Butt, A., Pearson, S. J., Wharton, S., Watson, A. J., Marriott, A., Caetano, C. M., Hollins, J. J., Rukazenkova, N., Begum, G., and Santibáñez-Koref, M. F. (2007) *DNA Repair* **6,** 1222–1228
27. Miao, F., Bouziane, M., Dammann, R., Masutani, C., Hanaoka, F., Pfeifer, G., and O'Connor, T. R. (2000) *J. Biol. Chem.* **275,** 28433–28438