HHS Public Access

Author manuscript

DNA Repair (Amst). Author manuscript; available in PMC 2021 December 01.

Published in final edited form as:

DNA Repair (Amst). 2020 December ; 96: 102944. doi:10.1016/j.dnarep.2020.102944.

Insights into the Direct Oxidative Repair of Etheno Lesions: MD and QM/MM Study on the Substrate Scope of ALKBH2 and AlkB

Stefan A. P. Lenz[a], **Deyu Li**[b], **Stacey D. Wetmore**[*,a]

[a]Department of Chemistry and Biochemistry, University of Lethbridge, 4401 University Drive West, Lethbridge, Alberta Canada, T1K 3M4

[b]Department of Biomedical and Pharmaceutical Sciences, College of Pharmacy, University of Rhode Island, Kingston, RI 02881, USA.

Abstract

E. coli AlkB and human ALKBH2 belong to the AlkB family enzymes, which contain several α-ketoglutarate (α-KG)/Fe(II)-dependent dioxygenases that repair alkylated DNA. Specifically, the AlkB enzymes catalyze decarboxylation of α-KG to generate a high-valent Fe(IV)-oxo species that oxidizes alkyl groups on DNA adducts. AlkB and ALKBH2 have been reported to differentially repair select etheno adducts, with preferences for 1,N^6-ethenoadenine (1,N^6-εA) and 3,N^4-ethenocytosine (3,N^4-εC) over 1,N^2-ethenoguanine (1,N^2-εG). However, N^2,3-ethenoguanine (N^2,3-εG), the most common etheno adduct, is not repaired by the AlkB enzymes. Unfortunately, a structural understanding of the differential activity of *E. coli* AlkB and human ALKBH2 is lacking due to challenges acquiring atomistic details for a range of substrates using experiments. This study uses both molecular dynamics (MD) simulations and ONIOM(QM:MM) calculations to determine how the active site changes upon binding each etheno adduct and

[*]**Corresponding Author:** Stacey.wetmore@uleth.ca. Telephone: (403) 329-2323.

CRediT author statement

Stefan Lenz: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – Original Draft, Writing – Review & Editing, Visualization **Deyu Li:** Conceptualization, Resources, Writing – Review & Editing, Project administration, Funding acquisition **Stacey Wetmore:** Conceptualization, Methodology, Validation, Resources, Data Curation, Writing – Original Draft, Writing – Review & Editing, Supervision, Project administration, Funding acquisition

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website. rmsd between replicate trajectories relative to representative reference structures (Tables S1–S2). Key distances and hydrogen-bonding data (Tables S3–S8). Overlays of AlkB MD representative and X-ray crystal structures (Figure S1). Active site conformation with water distribution highlighted (Figure S2 and S10). Overlays of AlkB and ALKBH2 complexes (Figures S3, S5, S6, S9, and S13). Key interactions between AlkB/ALKBH2 and substrate interactions (Figures S4, S12, and S13). Histograms of catalytic DNA-enzyme distances (Figures S7 and S8). Radius of gyration for water density (Figure S11). Representative structures for each system and parameter sets for the iron-binding sets and each substrate.

Notes

The authors declare no competing financial interests.

characterizes the corresponding catalytic impacts. Our data reveal that the preferred etheno substrates (1,N^6-εA and 3,N^4-εC) form favorable interactions with catalytic residues that situate the lesion near the Fe(IV)-oxo species and permit efficient oxidation. In contrast, although the damage remains correctly aligned with respect to the Fe(IV)-oxo moiety, repair of 1,N^2-εG is mitigated by increased solvation of the active site and a larger distance between Fe(IV)-oxo and the aberrant carbons. Binding of non-substrate N^2,3-εG in the active site disrupts key DNA–enzyme interactions, and positions the aberrant carbon atoms even further from the Fe(IV)-oxo species, leading to prohibitively high barriers for oxidative catalysis. Overall, our calculations provide the first structural insight required to rationalize the experimentally-reported substrate specificities of AlkB and ALKBH2 and thereby highlight the roles of several active site residues in the repair of etheno adducts that directly correlates with available experimental data. These proposed catalytic strategies can likely be generalized to other α-KG/Fe(II)-dependent dioxygenases that play similar critical biological roles, including epigenetic and post-translational regulation.

Keywords

direct repair; alkylation damage; molecular dynamics; AlkB enzymes; QM/MM methods

INTRODUCTION

DNA is often modified by endogenous and exogenous agents that alkylate nucleobases and threaten the fidelity of the genome.[1–3] For example, etheno adducts can form when DNA reacts with metabolites of vinyl chloride or the unsaturated products of lipid peroxidation.[4, 5] To date, four etheno adducts have been characterized: 1,N^6-ethenoadenine (1,N^6-εA), 3,N^4-ethenocytosine (3,N^4-εC), 1,N^2-ethenoguanine (1,N^2-εG), and N^2,3-ethenoguanine (N^2,3-εG; Figure 1). These lesions are replication blocks for standard DNA polymerases,[6, 7] and successful bypass by translesion synthesis polymerases sometimes results in mutagenic consequences *in vivo*.[6, 8] Specifically, A→T and A→G substitution mutations occur upon replication of DNA-containing 1,N^6-εA,[6, 8] while C→A and C→T mutations arise when 3,N^4-εC-containing DNA is copied.[6] Both frameshift and substitution (G→A and G→T) mutations result when 1,N^2-εG or 2′-F-N^2,3-εG is replicated, with the 2′-F-N^2,3-εG analogue studied due to the highly labile glycosidic bond of the free N^2,3-εG nucleotide and the anticipated minimal effect of the 2′ fluoro substituent.[7, 9]

Since the accumulation of mutations destabilizes the genome, leading to carcinogenesis,[10] successful repair of mutagenic adducts such as the etheno lesions is paramount.[1] Indeed, deficiencies in the repair of 1,N^6-εA and 3,N^4-εC have been associated with adenocarcinoma lung tumours[11, 12] and colorectal carcinoma.[13] Furthermore, ineffective repair of N^2,3-εG has been implicated as a cause of vinyl chloride-induced carcinogenesis.[7] Several DNA repair pathways act on etheno adducts,[14] including base excision repair. Specifically, AAG initiates removal of 1,N^6-εA and 1,N^2-εG from DNA as part of the base excision repair (BER) pathway,[15, 16] while repair of 3,N^4-εC can be initiated by DNA glycosylases TDG,[17] SMUG1,[18] and MBD4,[19] and N^2,3-εG is a substrate of 3-methyladenine DNA glycosylase II.[20] Transription-coupled nucleotide excision repair has

also been reported to act on 3,N^4-εC-containing DNA.[21] In *E. coli* and humans, the AlkB family enzymes catalyzes oxidative dealkylation of 1,N^6-εA[6, 22] and 3,N^4-εC,[6] which expands the substrate scope of AlkB beyond methylated DNA.[23–29] In fact, despite numerous pathways that process alkylated DNA,[14] Delaney and coworkers revealed that only AlkB proficient cells mitigate the mutagenicity and replication block of 1,N^6-εA and 3,N^4-εC.[6]

E. coli AlkB and other members of the AlkB family enzymes are part of the nonheme α-ketoglutarate (α-KG)/Fe(II)-dependent dioxygenase superfamily that utilizes molecular oxygen and decarboxylation of α-KG to drive the oxidation of organic molecules (Figure 2).[30] Members of this superfamily catalyze oxidative transformations and thereby play critical roles within biological pathways.[31–36] Indeed, the AlkB enzymes are present in prokaryotes and eukaryotes, including 9 human homologues (ALKBH1–8 and FTO), although only ALKBH2 and ALKBH3 have been reported to repair alkylated DNA.[30, 37–39] Among the four etheno lesions, 1,N^6-εA and 3,N^4-εC are generally efficiently dealkylated by the AlkB proteins,[6, 22, 40] while N^2,3-εG is not repaired by *E. coli* AlkB[7] or other AlkB enzymes including ALKBH2 and ALKBH3.[7, 41] Furthermore, 1,N^2-εG damage is reversed slower than 1,N^6-εA or 3,N^4-εC by AlkB, ALKBH2, and ALKBH3.[40, 41]

In addition to adduct specificity, the AlkB enzymes exhibit different strand preferences when targeting etheno lesions. Specifically, both *E. coli* AlkB and ALKBH3 have higher repair efficiencies for etheno adducts in single-stranded DNA (ssDNA) compared to double-stranded DNA (dsDNA),[40] which correlates with a ssDNA preference for the repair of methyl lesions.[23–28, 39] In contrast, ALKBH2 preferentially repairs both methyl and etheno adducts in dsDNA over ssDNA.[2, 27, 40, 42] Crystal structures of *E. coli* AlkB and ALKBH2 bound to DNA explain the strand specificity by suggesting unique base flipping strategies for each enzyme.[43] While ALKBH2 uses an aromatic finger residue (Phe102) to facilitate base flipping,[43] *E. coli* AlkB bends and compresses the DNA backbone to flip the lesion into the active site and stabilizes the extrahelical conformation by forcing the flanking nucleobases into a stacked arrangement.[43] Despite conservation of several DNA-binding motifs, the Phe102 finger residue of ALKBH2 is not conserved in ALKBH3,[39, 43] and an DNA–ALKBH3 complex has yet to be crystallized to provide structural clues for the base-flipping mechanism. Despite differential strand preferences for the AlkB enzymes, crystal structures of *E. coli* AlkB in complex with ssDNA and dsDNA reveal a remarkably similar active site conformation (rmsd = 0.912 Å),[44] indicating that the base flipping step is likely responsible for the strand specificity. Furthermore, the AlkB enzymes prefer 1,N^6-εA and 3,N^4-εC over 1,N^2-εG regardless of the ssDNA versus dsDNA context,[27, 40, 42] providing evidence that the substrate specificity of the AlkB enzymes is affected by the steps following base flipping.

Crystal structures of *E. coli* AlkB[43–47] and ALKBH2[43, 46–48] bound to cofactors and lesion-containing DNA reveal that these enzymes utilize similar active site architectures to bind substrates (Figure 3). Specifically, a 2-His-1-carboxylate (Asp) facial triad binds Fe(II) and α-KG, which facilitates generation of a high-valent iron-oxo intermediate (Figures 2 and 3).[39, 43–48] Several crystal structures have been solved in the presence of the cofactors and Mn(II), which is anticipated to bind to the facial triad in a similar geometry as Fe(II). In *E.*

coli AlkB (ALKBH2 numbering in brackets), a hydrogen bond exists between the 5′-phosphate of the lesion and Tyr76 (Ser125), and π−π interactions are present between the nucleobase and His131 (His171) and Trp69 (Phe124). For 1-methyladenine (1MeA) or 3-methylcytosine (3MeC), Asp135 (Glu175) hydrogen bonds with the exocyclic amino group, which likely contributes to the substrate preference for A and C lesions over G and T lesions.[25, 27, 43] However, direct hydrogen bonds are not present between the nucleobase and active site residues when AlkB (ALKBH2) is in complex with 1,N^6-εA-containing DNA, although Asp135 (Glu175), Glu136, and Tyr78 (Tyr122) are positioned to interact with the nucleobase through water-mediated hydrogen-bonding networks.[43–48] Unfortunately, crystal structures of AlkB or ALKBH2 bound to 3,N^4-εC, 1,N^2-εG or N^2,3-εG have yet to be solved and therefore the active site conformation upon binding these lesions is currently unknown.

Computational studies have provided valuable insight into the function of AlkB. Specifically, DFT and QM/MM studies have been used to examine *E.* coli AlkB-catalyzed decarboxylation of α-KG to yield succinate and the Fe(IV)–oxo intermediate (Figure 2), which confirm a favored quintet spin state for the Fe(IV)-oxo species.[49–52] Furthermore, molecular dynamics (MD) simulations suggest that changes in discrete interactions between methyl lesions and the active sites of AlkB, ALKBH2, and FTO have a significant effect on relative repair efficiencies.[53–55] QM/MM calculations of the *E. coli* AlkB-mediated repair of 1,N^6-εA predicted a catalytic mechanism involving oxidation of the ethene bridge by the oxo ligand via oxy zwitterion and glyoxide intermediates.[56] This contrasts the proposed epoxide and glycol intermediates (Figure 2),[6, 7, 22, 45, 57, 58] based on mass spectrometry detection of intermediate species with corresponding masses,[6, 22, 57, 58] and *in crystallo* trapping of the glycol intermediate.[45] Nevertheless, the QM/MM predicted glyoxide intermediate[56] is indistinguishable from the glycol intermediate *in crystallo* and the aldehyde byproduct has an equivalent molecular mass to the glycol intermediate.

While previous mechanistic and structural studies have provided important insights regarding AlkB- and ALKBH2-mediated repair of 1,N^6-εA, the structural basis for the relative activity of the AlkB enzymes toward other etheno lesions is poorly understood. Indeed, to the best of our knowledge, no theoretical study has examined the activity of AlkB on an etheno lesion other than 1,N^6-εA or the activity of ALKBH2 on any etheno adduct. Nevertheless, when 3,N^4-εC, 1,N^2-εG, and N^2,3-εG are superimposed onto a crystal structure of 1,N^6-εA-containing DNA bound to AlkB,[45] N^2,3-εG falls the furthest from the metal center (by ~0.5 Å compared to 1,N^6-εA), which was proposed to mitigate repair.[7] However, structural studies of nonheme Fe(II) and α-KG-dependent enzymes reveal that the distance between the substrate and metal atoms can vary significantly.[59] It is therefore unclear whether a longer distance between N^2,3-εG and Fe(II) entirely mitigates catalysis or whether other factors contribute to the inability of AlkB to oxidize N^2,3-εG. Moreover, NMR studies suggest that AlkB is highly dynamic and that decarboxylation of α-KG to succinate leads to an enhanced dynamic state,[60, 61] suggesting significant changes could occur in the position of the lesion relative to the Fe(IV)-oxo moiety and other catalytic active site residues. Thus, it is critical to obtain information about the dynamics of the active site conformation when the AlkB enzymes are bound to different etheno lesions to understand the reported differential rates of repair.

The current study combines MD simulations and QM/MM calculations to probe how AlkB or ALKBH2 bind the four etheno lesions (1,N^6-εA, 3,N^4-εC, 1,N^2-εG, and N^2,3-εG). Our work provides the first structural evidence of an AlkB enzyme bound to the 3,N^4-εC and 1,N^2-εG substrates and the N^2,3-εG non-substrate, which reveals how each lesion affects the position of catalytic residues and active site water molecules. As a result, new proposals can be made regarding the catalytic roles of active site residues that directly correlate with and thereby rationalize available experimental data. To understand the impact of these structural changes on the oxidation mechanism, we characterize the ONIOM(QM:MM) reaction pathway for the first step of AlkB-mediated repair of representative etheno lesions. Overall, our data provide a structural explanation for the experimentally-determined preference of the AlkB enzymes for 1,N^6-εA and 3,N^4-εC over 1,N^2-εG, and explain why N^2,3-εG is not an AlkB or ALKBH2 substrate. More broadly, our data sheds light on the catalytic strategies used by the AlkB enzymes to repair alkylated DNA, which can be extended to other α-KG/Fe(II)-dependent dioxygenases that play pivotal roles in numerous pathways,[31] including transcription,[32, 33] epigenetic regulation,[34] biosynthesis of base J in parasitic trypanosomes,[35] and tRNA modification pathways.[36]

COMPUTATIONAL METHODS

In total, 8 unique complexes were considered that contain AlkB or ALKBH2 bound to DNA containing one of four etheno lesions (namely 1,N^6-εA, 3,N^4-εC, 1,N^2-εG, and N^2,3-εG, Figure 1). X-ray crystal structures of AlkB (PDB IDs: 3O1U and 3O1S)[45] and ALKBH2 (PDB IDs: 3RZK and 3RZJ)[48] bound to lesion-containing DNA were used to build all models. In these crystal structures, AlkB is in complex with succinate, Fe(II), and oxidized 1,N^6-εA or 3MeC, while ALKBH2 is in complex with α-KG, Mn(II), and 1,N^6-εA or 3MeC. Starting from the structure of AlkB bound to oxidized 1,N^6-εA (PDB ID: 3O1U), the coordination distance between Fe(II) and the oxygen attached to the nucleobase was shortened to generate the Fe(IV)–oxo AlkB complex. The iron-binding site was then optimized with B3LYP/6–31G(d,p) (H, C, N, and O) and B3LYP/LANL2DZ (Fe(IV)) using Gaussian 16 (revision B.01).[62] Subsequently, the iron-binding site was superimposed onto the metal-binding site of the three other X-ray crystal structures. For the ALKBH2 models, α-KG was replaced with succinate. Starting structures for 3,N^4-εC bound by AlkB or ALKBH2 were generated from the 3MeC-containing crystal structures since these pyrimidine lesions differ by only a single heavy atom. Models of AlkB or ALKBH2 bound to the εG lesions were generated from the X-ray crystal structures of 1,N^6-εA bound by the corresponding enzyme. The nucleotide chi (χ) angle was adjusted to minimize steric clashes within the active site and increase the distance between the nucleobase and the oxo ligand for the 3O1U-derived complexes (±30° for AlkB bound to 1,N^6-εA, 1,N^2-εG, and N^2,3-εG; ±5° for other complexes). Protonation states were initially assigned using H++,[63] and adjusted to preserve local hydrogen-bonding networks. All crosslinking mutations necessary for crystallization were reverted to the wild-type phenotype, while the overhanging 5′ and 3′ termini of the bound DNA were truncated.

The Amber ff14SB forcefield[64] was used to model each complex, with the parameter set assigned to non-standard etheno lesions being supplemented with GAFF parameters.[65] Each etheno adduct was assigned restrained electrostatic potential (RESP) charges calculated

based on B3LYP/6–31G(d) optimized geometries (Gaussian 16, revision B.01).[62] The Metal Center Parameter Builder (MCPB.py)[66] was utilized to assign bonding, angle, dihedral, and non-bonding parameters based on B3LYP (6–31G(d,p) for H, C, N, and O, and LANL2DZ for Fe(IV)) optimized structures of the iron-binding site according to the Seminaro method.[67] Each complex was neutralized with Na^+ counterions and placed in an octahedral TIP3P water box, ensuring that at least 10 Å exists between the DNA–protein complex and the water box boundary.

The Amber 16 GPU-accelerated and the MPI version of PMEMD[68, 69] was used for all equilibration and production steps, while the CPU version of PMEMD was utilized for minimization steps.[70] A 10 Å electrostatic cutoff was implemented using the particle mesh Ewald method. Prior to unrestrained simulation, the system was subjected to several rounds of minimization, heating, and equilibration. Representative structures (clustering methodology described below) were chosen from these simulations as starting structures for three 500 ns MD replica simulations for each system, which differ in the initial velocity. Due to small differences between the replicas (active site rmsds range between 0.7 and 1.4 Å; Tables S1–S2), one replica for each system was extended to 1 μs MD, which yielded data in good agreement with that from the shorter simulations. Therefore, the main text focuses on the results from the 1 μs MD production simulations. The CPPTRAJ module[71] implemented in Amber 16 was used to analyze the geometric data obtained from trajectories.

To determine the structural impact of differing active site configurations isolated from MD simulations on the first step in the catalytic pathway, representative MD structures for select etheno lesions were used as ONIOM(QM:MM) starting points. The ONIOM(QM:MM) models were minimized using the MD force field described above and a 100.0 kcal mol^{-1} $Å^{-2}$ restraint to fix excluded atoms. Subsequently, all water molecules greater than 5 Å from the DNA or protein were removed. The system was separated into two layers according to the ONIOM formulism: 1) the QM region was treated with B3LYP/6–31G(d,p) for H, C, N, and O atoms and B3LYP/LANL2DZ for Fe, and 2) the MM region was treated with Amber ff14SB supplemented with GAFF, and MCPB.py-derived parameters for the bound lesion and iron-ligating site. The QM layer for AlkB contains the sidechains (truncated at the Cα −Cβ bond) of Trp69, His131, Asp133, His187, and Asp135, succinate, Fe(IV), the oxo ligand, the iron-ligated water, the bound nucleotide (truncated at the N9–C1′ bond), and water molecules (typically 4–5) within 5 Å of the aberrant carbon atoms of the bound lesion. Water molecules more than 9 Å from the bound nucleotide were held fixed. Each reactant complex (RC) was ONIOM(QM:MM) optimized using electrostatic embedding (Gaussian 16, revision B.01).[62] Subsequently, a relaxed potential energy surface scan was performed from the optimized RC by fixing and reducing the reaction coordinate (e.g., distance between the oxo ligand and C10) to ~1.3 Å, initially in 0.100 Å increments and subsequently in 0.05 Å increments in the TS region. Stationary points were identified from the scans, optimized without constraints, and characterized using frequency calculations. Approximate Gibbs energies with scaled (0.9813) zero-point corrections and unscaled thermal corrections are reported in the figures and tables. Reported Gibbs energies neglect entropic and enthalpic contributions from the MM environment due to the harmonic approximation applied to QM atoms.

Full computational methods are provided in the Supporting Information.

RESULTS AND DISCUSSION

Average Dynamic Distance Between the Bound Lesion and Fe(IV)-oxo Group Correlates with Relative AlkB and ALKBH2 Repair Efficiencies.

As discussed in the Introduction, the 3,N^4-εC, 1,N^2-εG, and N^2,3-εG nucleotides have been superimposed onto DNA containing 1,N^6-εA in complex with AlkB, which revealed that the etheno substrates (i.e., 1,N^6-εA, 3,N^4-εC, and 1,N^2-εG) are in close proximity to the Fe(IV)-oxo group, while N^2,3-εG is far from this group.[7] Although this static prediction correlates with kinetic and mass spectrometry data reflecting the relative repair rates,[6, 7, 40] this approach does not account for the reported highly dynamic AlkB complex.[60, 61] To more accurately assess how the series of etheno lesions is positioned with respect to the Fe(IV)-oxo moiety, MD simulations were performed on the corresponding DNA-AlkB and ALKBH2 complexes.

In all DNA-protein structures, the preferred octahedral coordination geometry of iron is maintained, with average distances to Fe(IV) of ~2.2 Å for His residues, ~1.9 Å for succinate, Asp, and water, and ~1.6 Å for the oxo ligand (Tables S3–S4). Furthermore, the positions of His131, His187, and Asp133 deviate by only ~0.1 Å for the AlkB complex with 1,N^6-εA compared to the corresponding X-ray crystal structure (Figure S1 and Table S3),[45] while the oxo coordination distance (1.6 Å) is equivalent to that in a previously reported QM/MM structure.[56] These features provide confidence that the modelled Fe(IV)-oxo species permits accurate geometries of the active site.

For AlkB bound to 1,N^6-εA, 3,N^4-εC, or 1,N^2-εG, the distance between the Fe(IV)-oxo group and closest carbon atom (C10 or C11) in the etheno bridge is 3.3–3.4 Å (Figure 1 and Figure 4a–c). In contrast, the distance between N^2,3-εG and the Fe(IV)-oxo moiety is longer (4.0 Å; Figure 4d and Table S3). A similar trend occurs for the ALKBH2 complexes, where substrates are positioned near the oxo ligand (3.4–3.8 Å; Figure 4e–h), while N^2,3-εG is significantly further away (5.2 Å, Figures 4e–h and Table S4). Although the substrates are slightly further from the Fe(IV)-oxo moiety in ALKBH2 compared to AlkB, the standard deviation in this distance is also greater for ALKBH2 (±0.4–0.7 Å) compared to AlkB (±0.3 Å). The greater flexibility of the bound substrate in AlkBH2 arises since there are fewer π-interactions between the nucleobase and active site residues (Figure 3).

The distance between the Fe(IV)-oxo group and closest nucleobase heavy atom predicted in the present work by MD simulations for 1,N^6-εA, 3,N^4-εC, and 1,N^2-εG match the equivalent distance reported for a QM/MM reactant of AlkB bound to 1,N^6-εA (3.2 Å).[56] Furthermore, for both the AlkB and ALKBH2 complexes, the distance between the oxo ligand and 1,N^6-εA, 3,N^4-εC, or 1,N^2-εG matches the predicted and equivalent distance for AlkB or ALKBH2 bound to 3-methylcytosine (3MeC), which is among the best substrates for each enzyme.[54] Since Fe(IV)-oxo intermediates are highly reactive and not possible to trap *in crystallo*, the distance between the metal ion and aberrant carbon atoms must be compared. Our calculations predict that 1,N^6-εA and 3,N^4-εC are positioned ~4.5 Å from the metal center of AlkB or ALKBH2, which matches the corresponding distances (4.3–4.5

Å) in X-ray crystal structures of either enzyme in complex with 1,N^6-εA.[45, 48] In contrast, 1,N^2-εG falls 5.0 Å from the metal center (Tables S3–S4), while the corresponding distance for N^2,3-εG is even longer (5.3–6.1 Å; Tables S3–S4). Moreover, these distances are consistent across all replicates of the same DNA-protein complex (distances within 0.5 Å; Tables S5–S6). These comparisons provide confidence in the accuracy of our MD predicted structural information for the DNA-AlkB and ALKBH2 complexes.

Our predicted structural data correlates with the relative AlkB and ALKBH2 repair efficiencies of the etheno lesions,[7, 27, 40, 42] with the shortest distance between Fe(IV)-oxo and the aberrant carbon atoms arising for the strong substrates 1,N^6-εA and 3,N^4-εC, a longer distance for the weaker substrate 1,N^2-εG, and the longest distance for the non-substrate N^2,3-εG. Nevertheless, the proximity of N^2,3-εG to Fe(IV) is comparable to X-ray crystal structures of other α-KG/Fe(II)-dependent dioxygenases, including JmjC-methyl lysyl demethylases that positions substrates up to 5.8 Å away from the metal ion.[32, 59] While our data suggests that the proximity of the lesion to the Fe(IV)-oxo group may be one factor that influences catalysis, other factors likely also play significant roles in dictating whether the DNA-enzyme complex is biologically active.

π-interactions Stabilize Etheno Adducts within the Active Sites of AlkB and ALKBH2.

Crystal structures of AlkB or ALKBH2 bound to 1,N^6-εA suggest no direct hydrogen bonds occur between the nucleobase and active site residues (Figure 3).[43–48] However, π−π stacking interactions are formed between the nucleobase and His131/Trp69 in AlkB, and His171/Phe124 in ALKBH2. Tyr76 in AlkB is also in proximity to form π-stacking interactions with the bound 1,N^6-εA substrate nucleobase. Such π-interactions are ubiquitous throughout biology,[72] and have been shown to play a critical role in positioning alkylated DNA adducts in AAG-mediated repair.[73–75] Therefore, such interactions may also be critical components in substrate binding and catalysis for the repair of etheno lesions by the AlkB family enzymes.

Across MD simulations on DNA-AlkB complexes, all alkylated nucleobases maintain a π-stacking interaction with the sidechain of His131 in AlkB (His171 in ALKBH2), with a distance between the rings of ~3.5–4.5 Å (Tables S3–S4 and Figure 5). In AlkB, a T-shaped π-interaction also occurs between Tyr76 and each nucleobase (distance = 4.5–5.4 Å, Table S3), which matches the crystallographic orientation of Tyr76 with respect to AlkB for 1,N^6-εA (Figure 3).[45, 47] Interestingly, Tyr76 also forms a direct or water-bridged hydrogen bond with the 5′-phosphate (Table S7 and Figures 5 and S2), while Ser125 forms a direct hydrogen-bonding interaction for all ALKBH2 substrates (Table S8 and Figure 5). X-ray crystal structures of AlkB and ALKBH2 bound to 1,N^6-εA have led to the proposal that these interactions are important for stabilizing a closed and catalytically-conducive active site configuration.[43, 44]

Although the average stacking distance between the nucleobase and Trp69 is ~5.1–5.8 Å when AlkB is bound to the etheno substrates (1,N^6-εA, 3,N^4-εC, or 1,N^2-εG), the distance increases to 6.6 Å when AlkB is bound to a non-substrate (N^2,3-εG; Table S3). Indeed, the nucleobases of the substrates significantly overlap with Trp69, which contrasts the poor alignment of N^2,3-εG and Trp69 (Figures 5 and S3). Similarly, a π−π interaction exists

between Phe124 and the nucleobase for ALKBH2 bound to 1,N^6-εA, 3,N^4-εC, or 1,N^2-εG (distance = ~4.0–4.9 Å), but is generally not present for N^2,3-εG (distance = 8.3 Å; Table S4 and Figures 5 and S3–S4). In fact, the N^2,3-εG nucleotide π-contact with Phe124 is replaced with hydrogen-bonding interactions with Asn169 (occupancy = 70.4%, distance = 3.0 Å , angle = 145°) and Arg254 (occupancy = 74.0%, distance = 2.9 Å, angle = 144°, Table S8). Furthermore, there is disruption of the hydrogen bond between the 5′-phosphate and Ser125 for bound N^2,3-εG (occupancy = 16.2%, distance = 2.7 Å, angle = 164°; Figures S3 and S4 and Table S8). When compared to the consistent location of the three substrates in the active site (Figures 5 and S3), this altered positioning of N^2,3-εG may contribute to the lack of AlkB and ALKBH2 repair activity. The altered positioning of N^2,3-εG relative to the substrates is also consistent in all replicate simulations (distance = ~6.5 Å for AlkB and distance = ~7.6–9.4 Å for ALKBH2; Tables S5 and S6). Furthermore, when the AlkB and ALKBH2 active sites are compared when bound to the same etheno substrate, similar orientations of the nucleobase with respect to Trp69 in AlkB and Phe124 in ALKBH2 suggest these residues fulfill similar roles in alkylated DNA repair (Figure S5).

Overall, we propose that π-interactions play key roles in the AlkB and ALKBH2-mediated repair of alkylated DNA, including positioning substrates in catalytically-conducive conformations. This suggestion correlates with the critical binding and catalytic roles that π-interactions have been shown to play in AAG-mediated repair of damaged DNA (including 1,N^6-εA),[73–75] with disruption of these interactions mitigating AAG activity.[74, 75] Interestingly, a π-contact is present between an active site tyrosine (Tyr1902) and a bound substrate (5-methylcytosine) in ten-eleven translocation enzyme 2 (TET2), another α-KG/Fe(II)-dependent dioxygenase that oxidizes DNA (distance = 4.1 Å).[34] Our data suggest this π-contact may stabilize the position of the 5-methyl group in the vicinity of the Fe(II) center in TET2 and thereby facilitate oxidation. Thus, π-interactions are likely not only important for the repair of alkylated DNA, but rather may also play a catalytic role in the active sites of other α-KG/Fe(II)-dependent dioxygenases.

Other Catalytic Residues and Water are Aligned for AlkB- or ALKBH2-mediated Oxidation of 1,N^6-εA and 3,N^4-εC, but Disrupted for 1,N^2-εG and N^2,3-εG.

In the literature, Asp135 or Glu136 in AlkB (Glu175 in ALKBH2) have been implicated in the substrate preference of A and C lesions over T and G lesions (Figure 3).[25, 27, 43] Moreover, Asp135, Glu136, and Tyr78 were proposed to play an active role in AlkB-catalyzed repair of 1,N^6-εA, including activation of water to assist the formation of the glycol[57] or glyoxide[56] repair intermediate. Indeed, crystal structures of AlkB or ALKBH2 in complex with 1,N^6-εA suggest these residues may interact with the nucleobase through water-mediated hydrogen-bonding networks. Moreover, the placement of Tyr78 (Tyr122 in ALKB2) suggests possible (direct or water-mediated) interactions with the bound substrate. Therefore, these residues may also be important for binding and repair.

Our simulations of AlkB and ALKBH2 bound to DNA containing 1,N^6-εA reveal that the position of Asp135/Glu136 or Glu175, and Tyr78 or Tyr122 with respect to the lesion is similar to the matches the corresponding AlkB or ALKBH2 X-ray crystal structure (Figures 6, S6–S9).[45, 48] Specifically, the distance between the C10–C11 bond midpoint and Asp135

or Glu136 in AlkB falls between 6.1 and 7.1 Å (Table S3 and Figure S7, while Glu175 is generally < 6 Å from the lesion for ALKBH2 (Figure S8). Although the nucleobase-active site residue relative orientations are similar when either enzyme is bound to 3,N^4-εC (Figure 6a and d), there are slight differences, likely due to the smaller nucleobase. These slight changes permit increased solvation of the oxo ligand for AlkB or ALKBH2 and may not be favorable for the first step of repair (occupancy = 76–83%; Table S7–S8 and Figure S10). Nevertheless, a persistent solvent bridge exists between N6 of 1,N^6-εA or N4 of 3,N^4-εC and the carboxylate groups of Asp135 and/or Glu136 in AlkB (Glu175 in ALKBH2) (occupancy = 100%; Tables S7–S8 and Figure S11). Furthermore, a solvent-mediated hydrogen bond stabilizes the position of Tyr78 (Tyr122 in ALKBH2) and Glu136 (Glu175 in ALKBH2). Thus, the solvent-bridged interactions between the active site residues and N6 of 1,N^6-εA or N4 of 3,N^4-εC likely plays an important role in substrate binding. Thus, our simulations suggest that Asp135, Glu136, Tyr78 and active site water are in position to facilitate AlkB repair of 1,N^6-εA and 3,N^4-εC, while Glu175 and Tyr122 may play equivalent roles for ALKBH2. Nevertheless, the position of the residues could change during repair due to the highly dynamic nature of the active site of both enzymes.[60, 61]

For 1,N^2-εG or N^2,3-εG bound by AlkB or ALKBH2, the distances between the lesion and Asp135, Glu136, and Tyr78 in AlkB (Glu175 and Tyr122 in ALKBH2) increase significantly (by ~2–8 Å) compared to either enzyme bound to 1,N^6-εA or 3,N^4-εC (Tables S3–S4, Figures 6 and S6–S8). These changes occur in part due to electrostatic repulsion between O6 of the guanine nucleobase and the carboxylate group of Asp135 or Glu175 (Figure 6). For 1,N^2-εG, the oxo ligand is highly solvated, such that a water-bridged interaction exists between the oxo ligand and N2H of the nucleobase, which may mitigate lesion oxidation (Figure S10c and g). In contrast, water hydrogen bonds with N1H of N^2,3-εG and the oxo ligand is minimally solvated (occupancy < 32%, Figure S10d and h and Tables S7–S8). Nevertheless, solvent bridges occur between O6 of the εG lesions and the carboxylate groups of Asp135/Glu136 in AlkB or Glu175 in ALKBH2 (occupancy 83–100%; Tables S7–S8 and Figure S11). Furthermore, water is generally > 3.5 Å away from the aberrant carbon atoms (Figure S11), which indicates that water is not readily available for the formation of the glycol or glyoxide intermediate, although the position of the residues could change during repair.[54,55] Interestingly, the aberrant atoms of N^2,3-εG are situated in a hydrophobic region of the AlkB active site , while N^2,3-εG forms a cation-π interaction with Arg100 when bound by ALKBH2 (distance = 6.5 Å, Table S4 and Figure S12). These interactions are likely anti-catalytic due to the anticipated poor stabilization of polar and charged repair intermediates, including the positive charge forming in the oxy zwitterion intermediate previously predicted by QM/MM studies for AlkB-catalyzed repair of 1,N^6-εA.[56] Thus, the nonoptimal position of active site residues and water for oxidative repair may be additional factors for the weak activity of the AlkB enzymes toward 1,N^2-εG and the lack of activity toward N^2,3-εG.[7, 40]

The Lesion-Dependent Active Site Conformation Enhances AlkB-catalyzed Oxidation of 1,N^6-εA, but Significantly Impedes Oxidation of N^2,3-εG.

QM/MM mechanistic studies have proven to be a valuable tool for assessing the impact of changes in the active site conformations on the catalytic activity.[73, 75] Since MD simulations

reveal similar active site conformations for AlkB and ALKBH2 bound to the same lesion (Figures S5–S6) and there is more experimental data for AlkB than ALKBH2 to validate our predictions, we focus on AlkB-catalyzed DNA repair. Furthermore, due to the previous computational work on 1,N^6-εA repair and the differential activity of the AlkB repair enzymes toward N^2,3-εG, these adducts were chosen for further examination as representative examples of the extremes in activity. Specifically, ONIOM(QM:MM) models were built from representative MD structures to model RC of the first reaction step for 1,N^6-εA and N^2,3-εG repair, namely the oxidation of the ethylene bridge of the lesion by the oxo ligand coordinated to Fe(IV). The ONIOM(QM:MM) reactant complex for this first reaction step closely match the MD representative structures for both lesions (rmsd ~0.3–0.4 Å, Figure S13), and further highlight the unique lesion-specific active site conformation discussed based on the MD simulations.

For 1,N^6-εA, the lesion is positioned close to the Fe(IV)-oxo moiety (~3.0 Å), which is likely critical for oxidative catalysis (Figure 7a). A solvent bridge is present between N6 of 1,N^6-εA and Asp135, which stabilizes the position of the substrate in the active site. As the reaction proceeds from the RC to the transition state complexes (TS), the 1,N^6-εA nucleobase approaches the oxo ligand (oxo···C10 distance = 2.4 Å), while the oxo coordination distance increases slightly (by ~0.06 Å). The TS is stabilized by tightening of the π-interactions between the nucleobase and His131 (~0.15 Å decrease from the RC). Despite the nucleobase approaching the Fe(IV)-oxo group, the stacking distance between the nucleobase and Trp69 is essentially unchanged compared to the RC, indicating that Trp69 remains closely associated with the nucleobase (Figure 7a). This supports our previous suggestion that His131 and Trp69 play catalytic roles in addition to facilitating substrate binding. In the resulting intermediate complexes (IC), a bond is formed between the oxo ligand and the substrate (oxo···C10 distance = ~1.3 Å), while the oxo ligand remains coordinated to iron (distance = ~2.1 Å). The overall Gibbs reaction barrier (~1.4 kcal mol^{-1}) is low due to the high reactivity of the oxyferryl intermediate and the optimal position of the substrate with respect to the oxo ligand, His131, and Trp69. Although our QM/MM predicted barrier for 1,N^6-εA oxidation is smaller than that previously reported (~5.3 kcal mol^{-1}),[56] our model differs by the inclusion of Trp69 in the high-level layer, which may lower the barrier by stabilizing the charge developing on the nucleobase throughout the reaction. Indeed, the interactions between the nucleobase and His131/Trp69 are present in all (MD or QM/MM) DNA-enzyme complexes considered in the present work regardless of the bound substrate, which emphasizes the potential key mechanistic roles of these residues.

Unlike for 1,N^6-εA, the bound N^2,3-εG nucleotide is located far from the oxo ligand in the reactant complex corresponding to AlkB repair (3.506 Å; Figure 7b). Furthermore, although a π-contact is formed between the nucleobase and His131 (4.3 Å), the distance between the nucleobase and Trp69 is much longer for N^2,3-εG than 1,N^6-εA (by approximately 1.1 Å; Figure 7). Along with these differences in the RC, the N^2,3-εG TS is later compared to that of the strong etheno substrate 1,N^6-εA (oxo···C11 distance = 1.754 Å compared to 1.683 Å), which likely arises due to the lack of stabilization afforded to the nucleobase by the elongated π-contacts and the positioning of the aberrant atoms in a hydrophobic region. Indeed, although the His131···nucleobase distance decreases by ~0.2 Å, the Trp69···nucleobase distance increases by 0.4 Å in the TS compared to the RC, which

suggests that Trp69 is unable to stabilize the nucleobase in the TS. These structural features result in a high barrier for N^2,3-εG oxidation (~14.7 kcal mol^{-1}). However, the first step of AlkB-mediated repair is expected to occur rapidly since Fe(IV)-oxo intermediates are highly reactive. Indeed, kinetic data indicates that the overall barrier for AlkB-catalyzed repair of 1,N^6-εA is ~22 kcal mol^{-1},[6, 22] while QM/MM calculations predict that the rate-determining step is one of the final steps of the reaction,[56] which has an associated barrier (~18.6 kcal mol^{-1}) much larger than the first oxidation (~5.3 kcal mol^{-1}) for 1,N^6-εA repair. Therefore, the high barrier for the first reaction step for N^2,3-εG oxidation predicted by our ONIOM(QM:MM) models supports the proposal from our MD simulations that the DNA-enzyme complex for this etheno lesion does not adopt a catalytically-conducive conformation.

CONCLUSIONS

By providing the first structural insight into key active site features of repair enzymes bound to an etheno lesion other than 1,N^6-εA, the present combined MD and ONIOM(QM:MM) study sheds light on previously reported experimental data that highlights the differential activity of AlkB and ALKBH2 toward four etheno adducts.[7, 40] MD simulations reveal that π-interactions position the good substrates (1,N^6-εA and 3,N^4-εC) in close proximity to the Fe(IV)-oxo moiety. Furthermore, solvent bridges exist between catalytic Asp or Glu residues and N6 of 1,N^6-εA or N4 of 3,N^4-εC, which facilitate repair. In contrast, electrostatic repulsion occurs between the O6 carbonyl of the εG adducts and these active site residues. Indeed, the spatial separation between 1,N^2-εG and Asp/Glu residues increases solvation of the oxo ligand and increases the distance between the aberrant carbons and the Fe(IV)-oxo species. This likely mitigates oxidative catalysis and provides a structural rationale for the observed reduced activity of the AlkB repair enzymes toward this substrate. However, the biggest difference in binding for the G etheno lesions arises in the case of N^2,3-εG for which key active site π-interactions with the substrate are disrupted and the lesion is positioned very far from the Fe(IV)-oxo group. ONIOM(QM:MM) calculations reveal that changes in the position of the etheno adduct in the AlkB active site can have significant consequences on the first chemical step of etheno lesion repair. Indeed, ONIOM(QM:MM) predict a low barrier for 1,N^6-εA oxidation due to optimal binding of the substrate, while the binding configuration of N^2,3-εG significantly increases the barrier of an otherwise very low energy process. The large predicted oxidation barrier explains why N^2,3-εG lesion escapes AlkB-mediated repair. Overall, our data provide structural and energetic explanations for the experimentally-observed relative activity of the AlkB enzymes toward different etheno lesions, and suggest a complex interplay of several factors is critical for successful oxidation, including the placement of the substrate near the Fe(IV)-oxo moiety, maintenance of π-interactions between the nucleobase and aromatic residues, and limited solvation of the oxo ligand. As a result, our findings contribute to the growing body of literature clarifying the catalytic strategies used by α-KG/Fe(II)-dependent dioxygenases, which have diverse biological roles, including epigenetic regulation and post-translational modifications.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

Computational resources from the New Upscale Cluster for Lethbridge to Enable Innovative Chemistry (NUCLEIC) and those provided by Westgrid and Compute/Calcul Canada are greatly appreciated.

Funding

Support for this research was provided to S.D.W. by the Natural Sciences and Engineering Research Council of Canada (NSERC, 2016-04568), the Canada Foundation for Innovation (22770) and the Board of Governors Research Chair Program at the University of Lethbridge, and to D.L. by the National Institutes of Health (R15 CA213042 and R01 ES028865). S.A.P.L. acknowledges NSERC (CGS-D), Alberta Innovates-Technology Futures (AI-TF) and the University of Lethbridge for student scholarships.

ABBREVIATIONS

1MeA	1-methyladenine
3MeC	3-methylcytosine
AAG	alkyladenine DNA glycosylase
AlkA	3-methyladenine DNA glycosylase II
BER	base excision repair
CG	conjugate gradient
COM	center-of-mass
dsDNA	double-stranded DNA
IC	intermediate complex
MBD4	methyl-CpG binding domain protein 4
MCPB	Metal Center Parameter Builder
MD	molecular dynamics
MM	molecular mechanics
endonuclease	VIII-like 1
ONIOM	Our own N-layered Integrated Molecular Orbital and Molecular Mechanics
QM	quantum mechanics
RC	reactant complex
RESP	restrained electrostatic potential
rmsd	root-mean-square-deviation
SD	steepest descent

SMUG1	single-strand selective monofunctional uracil DNA glycosylase
ssDNA	single-stranded DNA
TDG	thymine-DNA glycosylase
TS	transition-state complex
1,N^2-ethenoguanine	1,N^2-εG
1,N^6-ethenoadenine	1,N^6-εA
3,N^4-ethenocytosine	3,N^4-εC
N^2,3-ethenoguanine	N^2,3-εG
α-KG	α-ketoglutarate

REFERENCES

[1]. Knijnenburg TA, Wang L, Zimmermann MT, Chambwe N, Gao GF, Cherniack AD, Fan H, Shen H, Way GP, Greene CS, Liu Y, et al. The Cancer Genome Atlas Research, N., Monnat RJ, Xiao Y, and Wang C (2018) Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas, Cell reports 23, 239–254.e236. [PubMed: 29617664]

[2]. Fu D, and Samson LD (2012) Direct Repair of 3,N(4)-Ethenocytosine by the Human ALKBH2 Dioxygenase is blocked by the AAG/MPG Glycosylase, DNA Repair 11, 46–52. [PubMed: 22079122]

[3]. Srinivas US, Tan BWQ, Vellayappan BA, and Jeyasekharan AD (2019) ROS and the DNA damage response in cancer, Redox Biology 25, 101084. [PubMed: 30612957]

[4]. Nair U, Bartsch H, and Nair J (2007) Lipid peroxidation-induced DNA damage in cancer-prone inflammatory diseases: A review of published adduct types and levels in humans, Free Radical Biol. Med 43, 1109–1120. [PubMed: 17854706]

[5]. Shrivastav N, Li D, and Essigmann JM (2010) Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation, Carcinogenesis 31, 59–70. [PubMed: 19875697]

[6]. Delaney JC, Smeester L, Wong C, Frick LE, Taghizadeh K, Wishnok JS, Drennan CL, Samson LD, and Essigmann JM (2005) AlkB reverses etheno DNA lesions caused by lipid oxidation *in vitro* and *in vivo*, Nat. Struct. Mol. Biol 12, 855–860. [PubMed: 16200073]

[7]. Chang S. c., Fedeles BI, Wu J, Delaney JC, Li D, Zhao L, Christov PP, Yau E, Singh V. Jost, et al. (2015) Next-generation sequencing reveals the biological significance of the N^2,3-ethenoguanine lesion *in vivo*, Nucleic Acids Res. 43, 5489–5500. [PubMed: 25837992]

[8]. Pandya GA, and Moriya M (1996) 1,N^6-ethenodeoxyadenosine, a DNA adduct highly mutagenic in mammalian cells, Biochemistry 35, 11487–11492. [PubMed: 8784204]

[9]. Zhao L, Christov PP, Kozekov ID, Pence MG, Pallan PS, Rizzo CJ, Egli M, and Guengerich FP (2012) Replication of N^2,3-Ethenoguanine by DNA Polymerases, Angew. Chem. Int. Ed 51, 5466–5469.

[10]. Hanahan D, and Weinberg Robert A. (2011) Hallmarks of Cancer: The Next Generation, Cell 144, 646–674. [PubMed: 21376230]

[11]. Tudek B, Zdżalik-Bielecka D, Tudek A, Kosicki K, Fabisiewicz A, and Speina E (2017) Lipid peroxidation in face of DNA damage, DNA repair and other cellular processes, Free Radical Biol. Med 107, 77–89. [PubMed: 27908783]

[12]. Speina E, Zielińska M, Barbin A, Gackowski D, Kowalewski J, Graziewicz MA, Siedlecki JA, Oliński R, and Tudek B (2003) Decreased Repair Activities of 1,N^6-Ethenoadenine and 3,N^4-

Ethenocytosine in Lung Adenocarcinoma Patients, Cancer Res. 63, 4351–4357. [PubMed: 12907604]

[13]. Obtułowicz T, Winczura A, Speina E, Swoboda M, Janik J, Janowska B, Cieśla JM, Kowalczyk P, Jawien A, Gackowski D, et al. (2010) Aberrant repair of etheno-DNA adducts in leukocytes and colon tissue of colon cancer patients, Free Radical Biol. Med 49, 1064–1071. [PubMed: 20600828]

[14]. Soll JM, Sobol RW, and Mosammaparast N (2017) Regulation of DNA Akylation Damage Repair: Lessons and Therapeutic Opportunities, Trends Biochem. Sci 42, 206–218. [PubMed: 27816326]

[15]. Lee C-YI, Delaney JC, Kartalou M, Lingaraju GM, Maor-Shoshani A, Essigmann JM, and Samson LD (2009) Recognition and Processing of a New Repertoire of DNA Substrates by Human 3-Methyladenine DNA Glycosylase (AAG), Biochemistry 48, 1850–1861. [PubMed: 19219989]

[16]. O'Brien PJ, and Ellenberger T (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase, J. Biol. Chem 279, 9750–9757. [PubMed: 14688248]

[17]. Saparbaev M, and Laval J (1998) 3,N^4-ethenocytosine, a highly mutagenic adduct, is a primary substrate for *Escherichia coli* double-stranded uracil-DNA glycosylase and human mismatch-specific thymine-DNA glycosylase, Proc. Natl. Acad. Sci. U.S.A 95, 8508–8513. [PubMed: 9671708]

[18]. Kavli B, Sundheim O, Akbari M, Otterlei M, Nilsen H, Skorpen F, Aas PA, Hagen L, Krokan HE, and Slupphaug G (2002) HUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup, J. Biol. Chem 277, 39926–39936. [PubMed: 12161446]

[19]. Petronzelli F, Riccio A, Markham GD, Seeholzer SH, Genuardi M, Karbowski M, Yeung AT, Matsumoto Y, and Bellacosa A (2000) Investigation of the substrate spectrum of the human mismatch-specific DNA N-glycosylase MED1 (MBD4): fundamental role of the catalytic domain, J. Cell Physiol 185, 473–480. [PubMed: 11056019]

[20]. Matijasevic Z, Sekiguchi M, and Ludlum DB (1992) Release of N^2,3-ethenoguanine from chloroacetaldehyde-treated DNA by *Escherichia coli* 3-methyladenine DNA glycosylase II, Proc. Natl. Acad. Sci. U.S.A 89, 9331–9334. [PubMed: 1409640]

[21]. Chaim IA, Gardner A, Wu J, Iyama T, Wilson DM III, and Samson LD (2017) A novel role for transcription-coupled nucleotide excision repair for the in vivo repair of 3,N^4-ethenocytosine, Nucleic Acids Res. 45, 3242–3252. [PubMed: 28115629]

[22]. Mishina Y, Yang C-G, and He C (2005) Direct Repair of the Exocyclic DNA Adduct 1,N^6-Ethenoadenine by the DNA Repair AlkB Proteins, J. Am. Chem. Soc 127, 14594–14595. [PubMed: 16231911]

[23]. Trewick SC, Henshaw TF, Hausinger RP, Lindahl T, and Sedgwick B (2002) Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage, Nature 419, 174–178. [PubMed: 12226667]

[24]. Aas PA, Otterlei M, Falnes PØ, Vågbø CB, Skorpen F, Akbari M, Sundheim O, Bjørås M, Slupphaug G, Seeberg E, and Krokan HE (2003) Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA, Nature 421, 859–863. [PubMed: 12594517]

[25]. Koivisto P, Duncan T, Lindahl T, and Sedgwick B (2003) Minimal Methylated Substrate and Extended Substrate Range of Escherichia coli AlkB Protein, a 1-Methyladenine-DNA Dioxygenase, J. Biol. Chem 278, 44348–44354. [PubMed: 12944387]

[26]. Delaney JC, and Essigmann JM (2004) Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli, Proc. Natl. Acad. Sci. U.S.A 101, 14051–14056. [PubMed: 15381779]

[27]. Falnes PØ, and Rognes T (2003) DNA repair by bacterial AlkB proteins, Res. Microbiol 154, 531–538. [PubMed: 14527653]

[28]. Duncan T, Trewick SC, Koivisto P, Bates PA, Lindahl T, and Sedgwick B (2002) Reversal of DNA alkylation damage by two human dioxygenases, Proc. Natl. Acad. Sci. U.S.A 99, 16660–16665. [PubMed: 12486230]

[29]. Falnes PØ, Johansen RF, and Seeberg E (2002) AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli, Nature 419, 178–182. [PubMed: 12226668]

[30]. Muller TA, and Hausinger RP (2015) CHAPTER 8 AlkB and Its Homologues - DNA Repair and Beyond, In 2-Oxoglutarate-Dependent Oxygenases, pp 246–262, The Royal Society of Chemistry.

[31]. Hausinger RP (2015) CHAPTER 1 Biochemical Diversity of 2-Oxoglutarate-Dependent Oxygenases, In 2-Oxoglutarate-Dependent Oxygenases, pp 1–58, The Royal Society of Chemistry.

[32]. Aik W, McDonough MA, Thalhammer A, Chowdhury R, and Schofield CJ (2012) Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases, Curr. Opin. Struct. Biol 22, 691–700. [PubMed: 23142576]

[33]. Ougland R, Lando D, Jonson I, Dahl JA, Moen MN, Nordstrand LM, Rognes T, Lee JT, Klungland A, Kouzarides T, and Larsen E (2012) ALKBH1 is a Histone H2A Dioxygenase Involved in Neural Differentiation, Stem Cells 30, 2672–2682. [PubMed: 22961808]

[34]. Hu L, Li Z, Cheng J, Rao Q, Gong W, Liu M, Shi YG, Zhu J, Wang P, and Xu Y (2013) Crystal Structure of TET2-DNA Complex: Insight into TET-Mediated 5mC Oxidation, Cell 155, 1545–1555. [PubMed: 24315485]

[35]. Gommers-Ampt JH, Van Leeuwen F, de Beer ALJ, Vliegenthart JFG, Dizdaroglu M, Kowalak JA, Crain PF, and Borst P (1993) β-d-glucosyl-hydroxymethyluracil: A novel modified base present in the DNA of the parasitic protozoan T. brucei, Cell 75, 1129–1136. [PubMed: 8261512]

[36]. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, and Rao A (2009) Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1, Science 324, 930–935. [PubMed: 19372391]

[37]. Aravind L, and Koonin EV (2001) The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases, Genome Biology 2, research0007.0001–research0007.0008. [PubMed: 11276424]

[38]. Kurowski MA, Bhagwat AS, Papaj G, and Bujnicki JM (2003) Phylogenomic identification of five new human homologs of the DNA repair enzyme AlkB, BMC Genomics 4, 48. [PubMed: 14667252]

[39]. Sundheim O, Vågbø CB, Bjørås M, Sousa MML, Talstad V, Aas PA, Drabløs F, Krokan HE, Tainer JA, and Slupphaug G (2006) Human ABH3 structure and key residues for oxidative demethylation to reverse DNA/RNA damage, EMBO J. 25, 3389–3397. [PubMed: 16858410]

[40]. Zdżalik D, Domańska A, Prorok P, Kosicki K, van den Born E, Falnes PØ, Rizzo CJ, Guengerich FP, and Tudek B (2015) Differential repair of etheno-DNA adducts by bacterial and human AlkB proteins, DNA Repair 30, 1–10. [PubMed: 25797601]

[41]. Fedeles BI, Singh V, Delaney JC, Li D, and Essigmann JM (2015) The AlkB Family of Fe(II)/α-Ketoglutarate Dependent Dioxygenases: Repairing Nucleic Acid Alkylation Damage and Beyond, J. Biol. Chem 20734–20742. [PubMed: 26152727]

[42]. Ringvoll J, Moen MN, Nordstrand LM, Meira LB, Pang B, Bekkelund A, Dedon PC, Bjelland S, Samson LD, Falnes PØ, and Klungland A (2008) AlkB homologue 2-mediated repair of ethenoadenine lesions in mammalian DNA, Cancer Res. 68, 4142–4149. [PubMed: 18519673]

[43]. Yang C-G, Yi C, Duguid EM, Sullivan CT, Jian X, Rice PA, and He C (2008) Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA, Nature 452, 961–965. [PubMed: 18432238]

[44]. Holland PJ, and Hollis T (2010) Structural and Mutational Analysis of Escherichia coli AlkB Provides Insight into Substrate Specificity and DNA Damage Searching, PLOS ONE 5, e8680. [PubMed: 20084272]

[45]. Yi C, Jia G, Hou G, Dai Q, Zhang W, Zheng G, Jian X, Yang C-G, Cui Q, and He C (2010) Iron-catalysed oxidation intermediates captured in a DNA repair dioxygenase, Nature 468, 330–333. [PubMed: 21068844]

[46]. Yu B, and Hunt JF (2009) Enzymological and structural studies of the mechanism of promiscuous substrate recognition by the oxidative DNA repair enzyme AlkB, Proc. Natl. Acad. Sci. U.S.A 106, 14315–14320. [PubMed: 19706517]

[47]. Yu B, Edstrom WC, Benach J, Hamuro Y, Weber PC, Gibney BR, and Hunt JF (2006) Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB, Nature 439, 879–884. [PubMed: 16482161]

[48]. Yi C, Chen B, Qi B, Zhang W, Jia G, Zhang L, Li CJ, Dinner AR, Yang C-G, and He C (2012) Duplex interrogation by a direct DNA repair protein in search of base damage, Nat. Struct. Mol. Biol 19, 671–676. [PubMed: 22659876]

[49]. Liu H, Llano J, and Gauld JW (2009) A DFT Study of Nucleobase Dealkylation by the DNA Repair Enzyme AlkB, J. Phys. Chem. B 113, 4887–4898. [PubMed: 19338370]

[50]. Fang D, Lord RL, and Cisneros GA (2013) *Ab Initio* QM/MM Calculations Show an Intersystem Crossing in the Hydrogen Abstraction Step in Dealkylation Catalyzed by AlkB, J. Phys. Chem. B 117, 6410–6420. [PubMed: 23642148]

[51]. Cisneros GA (2010) DFT study of a model system for the dealkylation step catalyzed by AlkB, Interdiscip. Sci 2, 70–77. [PubMed: 20640798]

[52]. Latifi R, Minnick JL, Quesne MG, de Visser SP, and Tahsini L (2020) Computational studies of DNA base repair mechanisms by nonheme iron dioxygenases: selective epoxidation and hydroxylation pathways, Dalton Transactions 49, 4266–4276. [PubMed: 32141456]

[53]. Waheed SO, Ramanan R, Chaturvedi SS, Ainsley J, Evison M, Ames JM, Schofield CJ, Christov CZ, and Karabencheva-Christova TG (2019) Conformational flexibility influences structure–function relationships in nucleic acid N-methyl demethylases, Org. Biomol. Chem 17, 2223–2231. [PubMed: 30720838]

[54]. Bian K, Lenz SAP, Tang Q, Chen F, Qi R, Jost M, Drennan CL, Essigmann JM, Wetmore SD, and Li D (2019) DNA repair enzymes ALKBH2, ALKBH3, and AlkB oxidize 5-methylcytosine to 5-hydroxymethylcytosine, 5-formylcytosine and 5-carboxylcytosine *in vitro*, Nucleic Acids Res. 47, 5522–5529. [PubMed: 31114894]

[55]. Zhu C, and Yi C (2014) Switching Demethylation Activities between AlkB Family RNA/DNA Demethylases through Exchange of Active-Site Residues, Angew. Chem. Int. Ed 53, 3659–3662.

[56]. Wang B, Usharani D, Li C, and Shaik S (2014) Theory Uncovers an Unusual Mechanism of DNA Repair of a Lesioned Adenine by AlkB Enzymes, J. Am. Chem. Soc 136, 13895–13901. [PubMed: 25203306]

[57]. Li D, Delaney JC, Page CM, Yang X, Chen AS, Wong C, Drennan CL, and Essigmann JM (2012) Exocyclic Carbons Adjacent to the N6 of Adenine are Targets for Oxidation by the Escherichia coli Adaptive Response Protein AlkB, J. Am. Chem. Soc 134, 8896–8901. [PubMed: 22512456]

[58]. Li D, Delaney JC, Page CM, Chen AS, Wong C, Drennan CL, and Essigmann JM (2010) Repair of DNA Alkylation Damage by the Escherichia coli Adaptive Response Protein AlkB as Studied by ESI-TOF Mass Spectrometry, J. Nucleic Acids 2010, 369434. [PubMed: 21048928]

[59]. Aik WS, Chowdhury R, Clifton IJ, Hopkinson RJ, Leissing T, McDonough MA, Nowak R, Schofield CJ, and Walport LJ (2015) CHAPTER 2 Introduction to Structural Studies on 2-Oxoglutarate-Dependent Oxygenases and Related Enzymes, In 2-Oxoglutarate-Dependent Oxygenases, pp 59–94, The Royal Society of Chemistry.

[60]. Bleijlevens B, Shivarattan T, Flashman E, Yang Y, Simpson PJ, Koivisto P, Sedgwick B, Schofield CJ, and Matthews SJ (2008) Dynamic states of the DNA repair enzyme AlkB regulate product release, EMBO Rep. 9, 872–877. [PubMed: 18617893]

[61]. Bleijlevens B, Shivarattan T, van den Boom KS, de Haan A, van der Zwan G, Simpson PJ, and Matthews SJ (2012) Changes in Protein Dynamics of the DNA Repair Dioxygenase AlkB upon Binding of Fe2+ and 2-Oxoglutarate, Biochemistry 51, 3334–3341. [PubMed: 22443471]

[62]. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, et al. (2018) Gaussian 16, Revision B.01 ed, Gaussian, Inc, Wallingford CT.

[63]. Anandakrishnan R, Aguilar B, and Onufriev AV (2012) H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations, Nucleic Acids Res. 40, W537–W541. [PubMed: 22570416]

[64]. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, and Simmerling C (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, J. Chem. Theory Comput 11, 3696–3713. [PubMed: 26574453]

[65]. Wang JM, Wolf RM, Caldwell JW, Kollman PA, and Case DA (2004) Development and testing of a general amber force field, J. Comput. Chem 25, 1157–1174. [PubMed: 15116359]

[66]. Li P, and Merz KM (2016) MCPB.py: A Python Based Metal Center Parameter Builder, J. Chem. Inf. Model 56, 599–604. [PubMed: 26913476]

[67]. Seminario JM (1996) Calculation of intramolecular force fields from second-derivative tensors, Int. J. Quantum Chem 60, 1271–1277.

[68]. Götz AW, Williamson MJ, Xu D, Poole D, Le Grand S, and Walker RC (2012) Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born, J. Chem. Theory Comput 8, 1542–1555. [PubMed: 22582031]

[69]. Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, and Walker RC (2013) Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald, J. Chem. Theory Comput 9, 3878–3888. [PubMed: 26592383]

[70]. Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Crowley M, Walker RC, Zhang W, et al. (2008) AMBER Tools, Version 1.0 ed, University of California, San Francisco.

[71]. Roe DR, and Cheatham TE (2013) PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data, J. Chem. Theory Comput 9, 3084–3095.– [PubMed: 26583988]

[72]. Wilson KA, Kellie JL, and Wetmore SD (2014) DNA-protein pi-interactions in nature: abundance, structure, composition and strength of contacts between aromatic amino acids and DNA nucleobases or deoxyribose sugar, Nucleic Acids Res. 42, 6726–6741. [PubMed: 24744240]

[73]. Rutledge LR, and Wetmore SD (2011) Modeling the chemical step utilized by human alkyladenine DNA glycosylase: a concerted mechanism aids in selectively excising damaged purines, J. Am. Chem. Soc 133, 16258–16269. [PubMed: 21877721]

[74]. Lenz SAP, and Wetmore SD (2016) Evaluating the Substrate Selectivity of Alkyladenine DNA Glycosylase: The Synergistic Interplay of Active Site Flexibility and Water Reorganization, Biochemistry 55, 798–808. [PubMed: 26765542]

[75]. Lenz SAP, and Wetmore SD (2017) QM/MM Study of the Reaction Catalyzed by Alkyladenine DNA Glycosylase: Examination of the Substrate Specificity of a DNA Repair Enzyme, J. Phys. Chem. B 121, 11096–11108. [PubMed: 29148771]

Highlights

- π−π interactions play key roles in positioning AlkB and ALKBH2 substrates for repair
- π−π interactions are disrupted when AlkB or ALKBH2 bind N2,3-εG
- Calculations predict small oxidation barriers for optimally placed substrates
- A large oxidation barrier is predicted for N_2,3-εG repair
- The large barrier explains why N_2,3-εG escapes AlkB-mediated repair

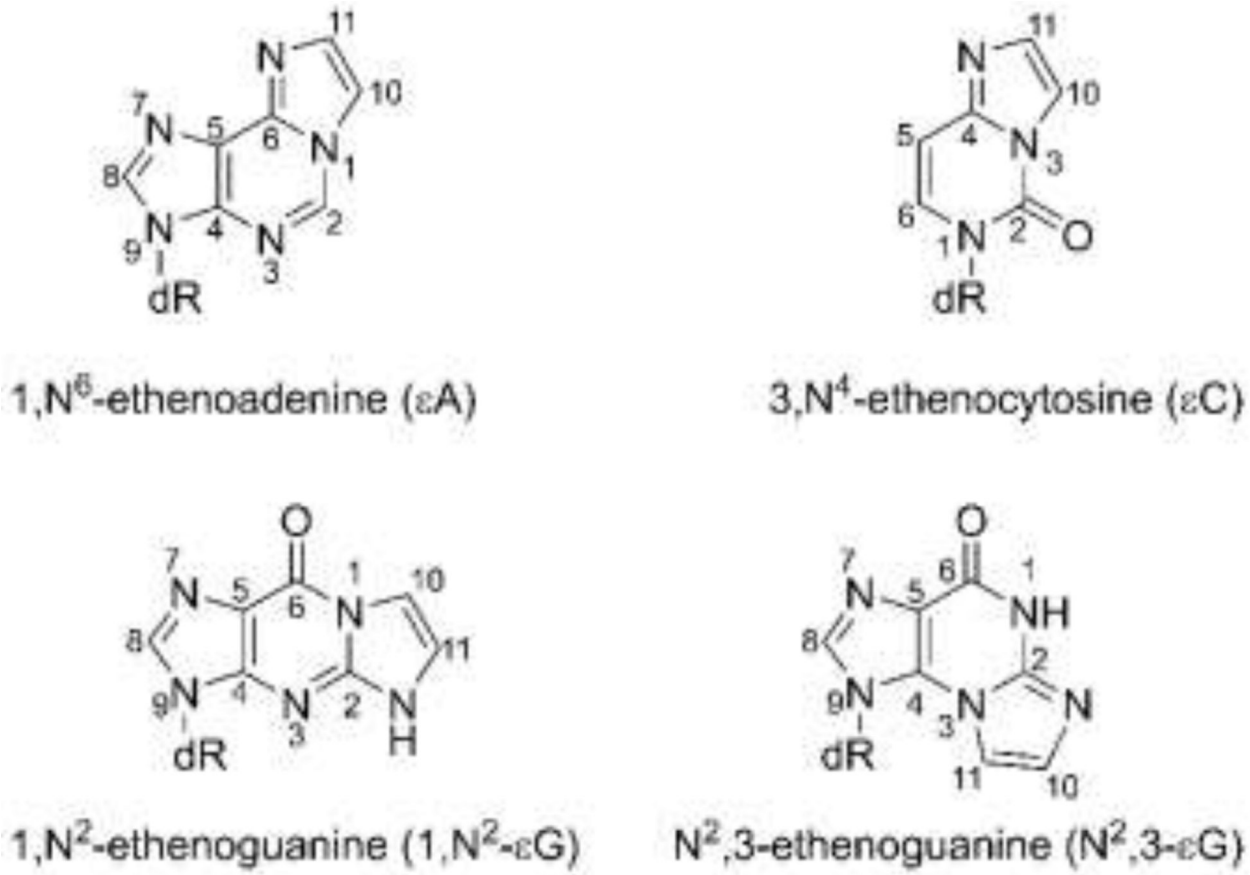

Figure 1.
Structures and chemical numbering of etheno adducts.

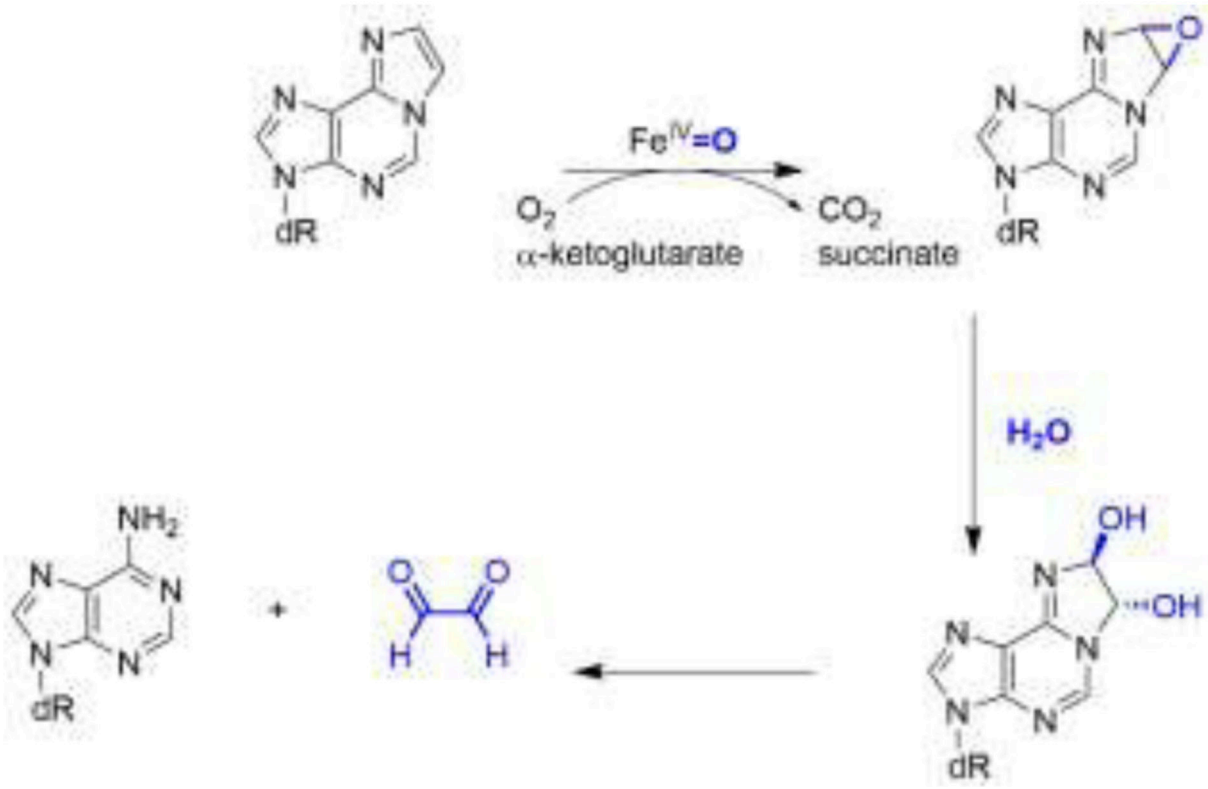

Figure 2.
Proposed mechanism for the repair of 1,N^6-εA catalyzed by the AlkB enzymes.[6]

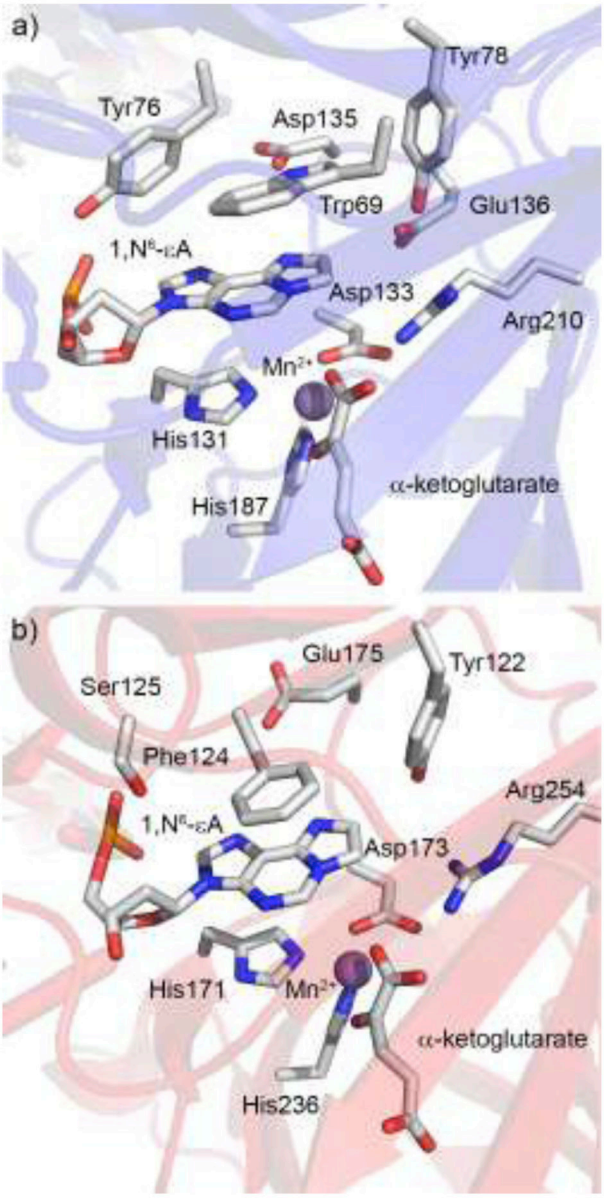

Figure 3.

X-ray crystal structures of the a) AlkB complex bound to 1,N^6-ɛA (PDB ID: 3O1P)[45] and b) ALKBH2 complex bound to 1,N^6-ɛA (PDB ID: 3RZK).[48]

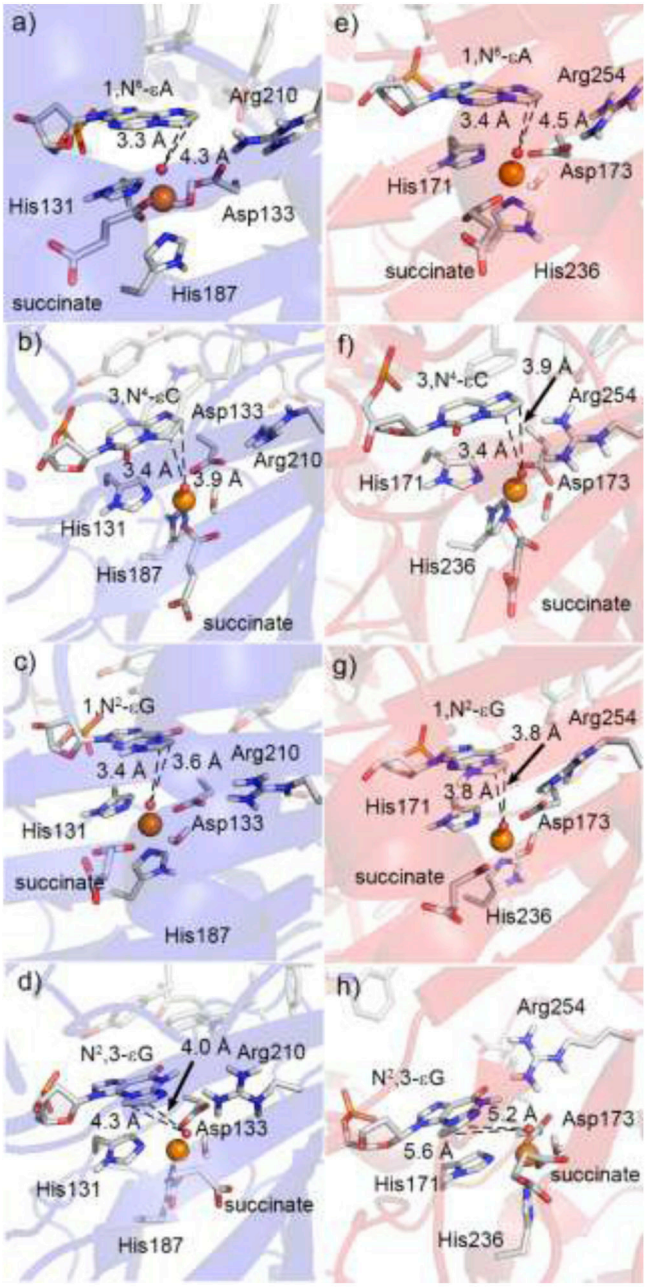

Figure 4.

Representative MD structures of the AlkB (a-d, blue) or ALKBH2 (e-h, red) complex bound to 1,N^6-εA (a,e), 3,N^4-εC (b,f), 1,N^2-εG (c,g), or N^2,3-εG (d,h) containing DNA, highlighting the average distances between the Fe(IV)-oxo group and the aberrant carbon atoms of the nucleobase.

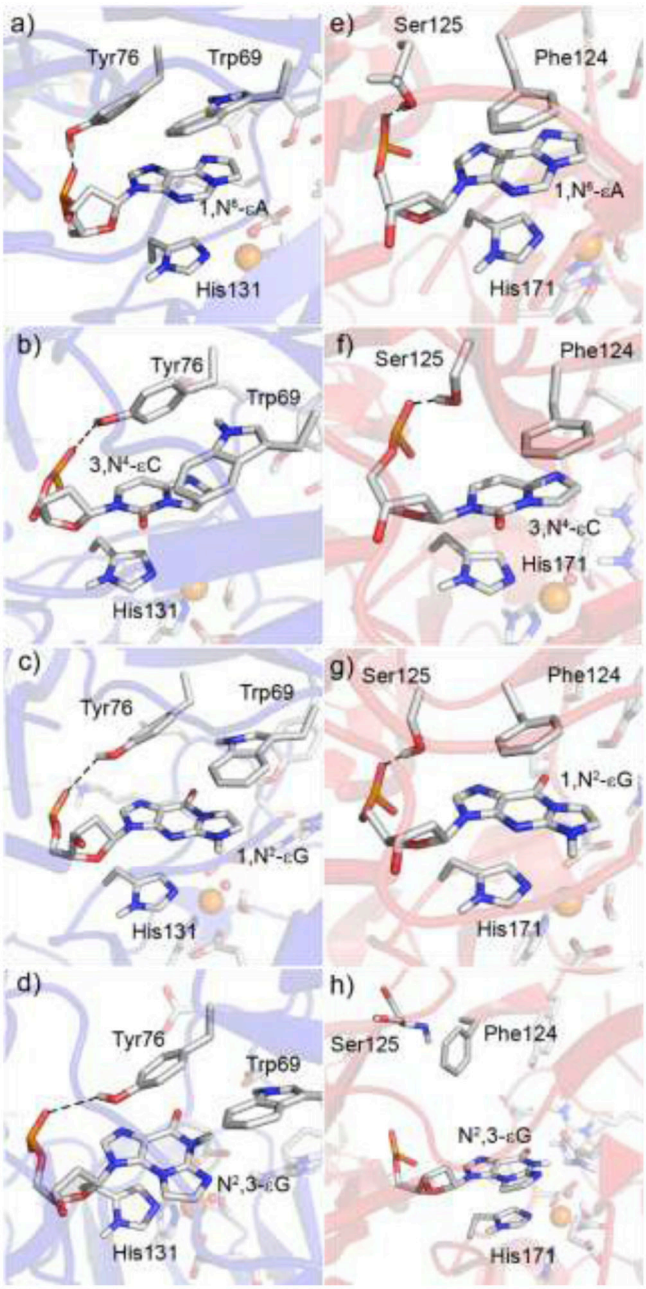

Figure 5.

Representative MD structures of the AlkB (a-d, blue) or ALKBH2 (e-h, red) complex bound to 1,N^6-εA (a,e), 3,N^4-εC (b,f), 1,N^2-εG (c,g), or N^2,3-εG (d,h) containing DNA, highlighting the hydrogen-bonding interactions between the 5′-phosphate and active site residues, and the π-interactions between the nucleobase and aromatic residues.

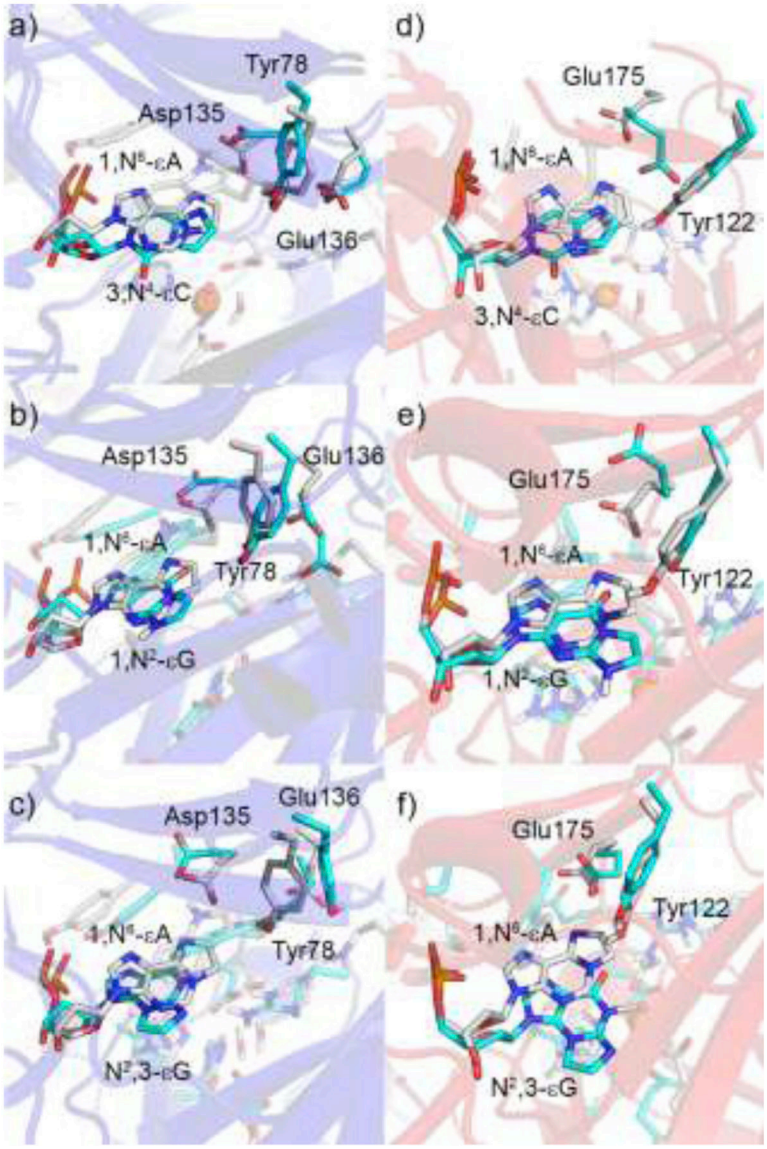

Figure 6.

Overlays of representative MD structures of the AlkB (a-c blue) or ALKBH2 (d-f red) complex bound to 1,N^6-εA and 3,N^4-εC (a,b), 1,N^2-εG (c,d), or N^2,3-εG (e,f) containing DNA, highlighting the lesion-dependent positions of Asp135, Glu136, and Tyr78 for AlkB, and Glu175 and Tyr122 for ALKBH2.

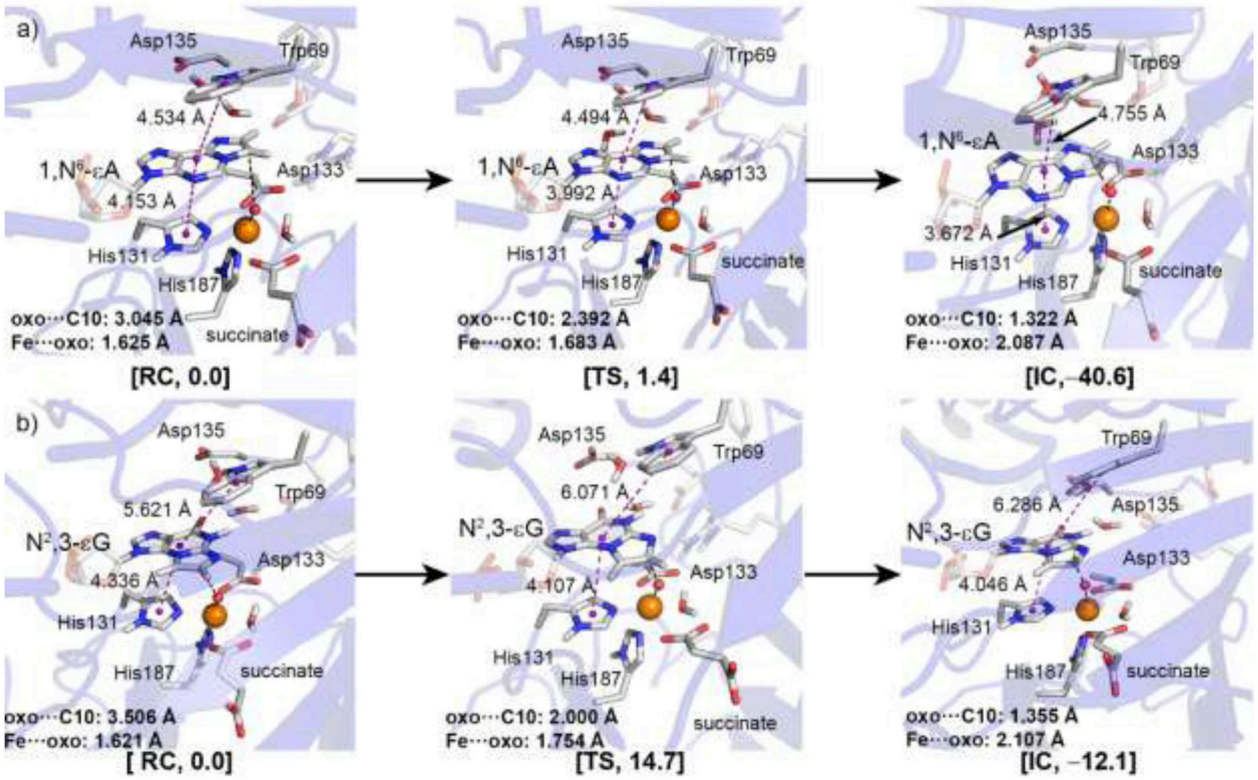

Figure 7.

ONIOM(QM:MM) stationary points (QM layer in tubes, MM layer in transparent sticks or cartoon) corresponding to the first step of the AlkB-catalyzed oxidation of a) 1,N^6-εA or b) N^2,3-εG. Relative approximate Gibbs energies (kcal mol^{-1}) are displayed in square brackets. Center-of-mass of the nucleobase, Trp69, and His131 are represented as purple spheres.