ACCOUNTS
of chemical research

A Non-Heme Iron-Mediated Chemical Demethylation in DNA and RNA

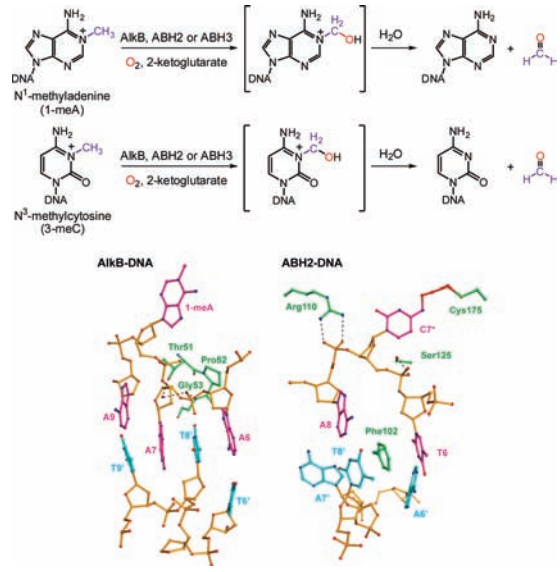
CHENGQI YI, CAI-GUANG YANG, AND CHUAN HE*

Department of Chemistry, The University of Chicago, 929 East 57th Street, Chicago, Illinois 60637

RECEIVED ON AUGUST 13, 2008

CONSPECTUS

DNA methylation is arguably one of the most important chemical signals in biology. However, aberrant DNA methylation can lead to cytotoxic or mutagenic consequences. A DNA repair protein in *Escherichia coli*, AlkB, corrects some of the unwanted methylations of DNA bases by a unique oxidative demethylation in which the methyl carbon is liberated as formaldehyde. The enzyme also repairs exocyclic DNA lesions—that is, derivatives in which the base is augmented with an additional heterocyclic subunit—by a similar mechanism. Two proteins in humans that are homologous to AlkB, ABH2 and ABH3, repair the same spectrum of lesions; another human homologue of AlkB, FTO, is linked to obesity. In this Account, we describe our studies of AlkB, ABH2, and ABH3, including our development of a general strategy to trap homogeneous protein−DNA complexes through active-site disulfide cross-linking.



AlkB uses a non-heme mononuclear iron(II) and the cofactors 2-ketoglutarate (2KG) and dioxygen to effect oxidative demethylation of the DNA base lesions 1-methyladenine (1-meA), 3-methylcytosine (3-meC), 1-methylguanine (1-meG), and 3-methylthymine (3-meT). ABH3, like AlkB, works better on single-stranded DNA (ssDNA) and is capable of repairing damaged bases in RNA. Conversely, ABH2 primarily repairs lesions in double-stranded DNA (dsDNA); it is the main housekeeping enzyme that protects the mammalian genome from 1-meA base damage.

The AlkB-family proteins have moderate affinities for their substrates and bind DNA in a non-sequence-specific manner. Knowing that these proteins flip the damaged base out from the duplex DNA and insert it into the active site for further processing, we first engineered a disulfide cross-link in the active site to stabilize the Michaelis complex. Based on the detailed structural information afforded by the active-site cross-linked structures, we can readily install a cross-link away from the active site to obtain the native-like structures of these complexes. The crystal structures show a distinct base-flipping feature in AlkB and establish ABH2 as a dsDNA repair protein. They also provide a molecular framework for understanding the demethylation reaction catalyzed by these proteins and help to explain their substrate preferences. The chemical cross-linking method demonstrated here can be applied to trap other labile protein−DNA interactions and can serve as a general strategy for exploring the structural and functional aspects of base-flipping proteins.

Introduction

Cellular DNA is constantly challenged by endogenous and exogenous chemicals, often resulting in cytotoxic or mutagenic covalent adducts.[1−4] Alkylating agents are one group of such chemicals widely used in clinical settings and prevalent in the environment that create cytotoxic DNA lesions.[5−8]Once damaged, cellular DNA must be promptly repaired. Organisms have evolved a variety of mechanisms to repair these cytotoxic or

Published on the Web 01/27/2009 www.pubs.acs.org
10.1021/ar800178j CCC: $71.50

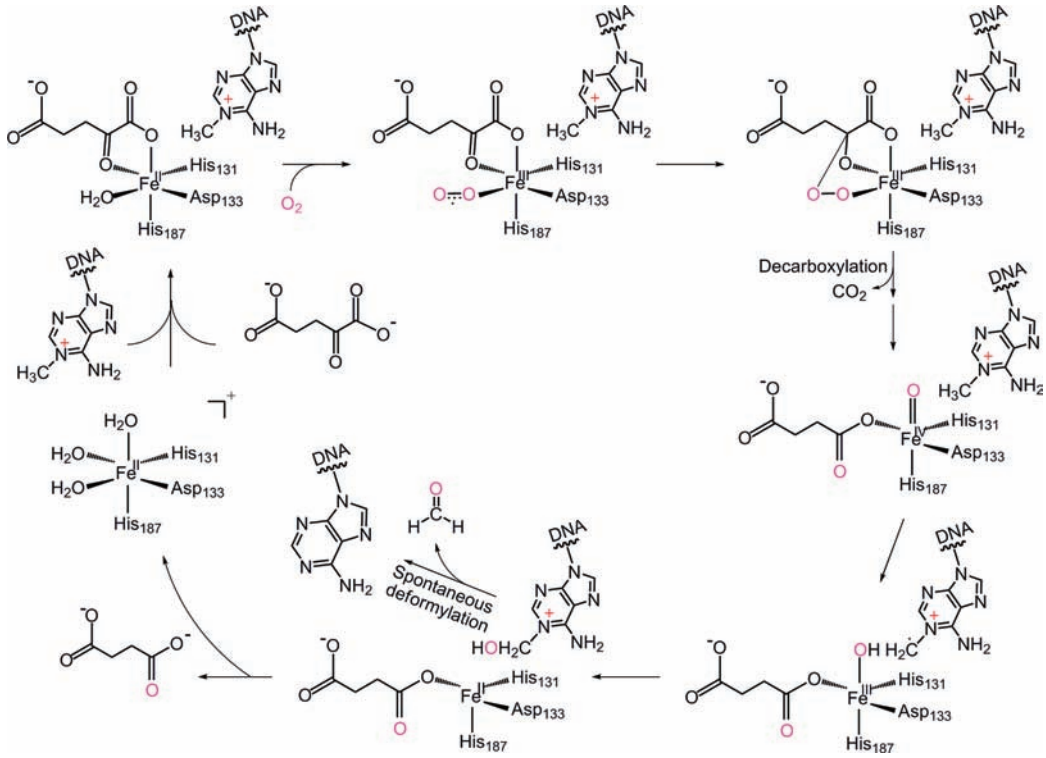


FIGURE 1. Proposed repair mechanism for the AlkB-family proteins. Several repair intermediates are shown.

mutagenic damages.[1,3] In *Escherichia coli*, AlkB protein is one of the four proteins that are activated when *E. coli* is challenged with high doses of methylation agents.[3,9,10] It belongs to a dioxygenase family that uses a non-heme mononuclear iron(II) and the cofactors 2KG and dioxygen to perform an oxidative demethylation of DNA base lesions, 1-methyladenine (1-meA) and 3-methylcytosine (3-meC) (Figure 1).[11−13]

In the event of oxidation reactions, AlkB utilizes an iron(II) site, which is coordinated by the conserved His2/Asp motif, to activate the dioxygen molecule for oxidation of the inert C−H bond. A possible mechanism involves formation of a superoxo radical anion (O_2^-) bound to iron(III), a subsequent bridged peroxo intermediate produced by the attack of the superoxide to the α-keto carbon of an iron-bound 2-KG, and a speculated high-valent iron(IV)−oxo intermediate that hydroxylates the C−H bond of the methyl lesion.[14,15] One of the oxygen atoms from the dioxygen cofactor is incorporated into succinic acid and the other becomes the oxygen of the hydroxyl product (Figure 1). This mechanism is shared by a variety of enzymes within the non-heme iron protein family. In the case of AlkB, the initial hydroxylation at the methyl group on the N^1-position of adenine or N^3-position of cytosine leads to the heterocleavage of C−N bonds, which gives the unmodified base and formaldehyde. The overall effect of the oxidation reaction is the direct repair of the damaged base.

Substrates

The AlkB-family proteins remove methyl adducts that block Watson−Crick base pairs in DNA and RNA. Base lesions 1-meA and 3-meC are the best substrates for AlkB.[11,12] Studies have shown that AlkB prefers 1-meA and 3-meC in a polynucleotide, but it is capable of demethylating shorter nucleotides as well.[16] A polynucleotide is not required for efficient repair, but the presence of a 5′-phosphate seems to be essential.[16] Survival assays for the alkylated bacteriophage have extended AlkB's substrates to include methyl, ethyl, propyl, hydroxyethyl, and hydroxypropyl DNA adducts.[17−19] *In vitro* studies have confirmed that 1-ethyladenine can be repaired by AlkB to give adenine with acetaldehyde produced.[17]

Minor alkylated DNA bases 3-methylthymine (3-meT) and 1-methylguanine (1-meG) are also repaired by AlkB, but less efficiently (Figure 2a).[19−21] Exocyclic DNA adducts such as 1,N^6-ethenoadenine (εA), 3,N^4-ethenocytosine (εC), and 1,N^2-ethenoguanine (εG) are another group of lesions that completely block Watson−Crick base pairing. They can be generated endogenously when cells are under oxidative stress, particularly when DNA is exposed to side products of lipid peroxidation.[22,23] Recent reports from Essigmann's group and our laboratory have revealed that εA can be efficiently repaired by *E. coli* AlkB and its human homologues both *in*

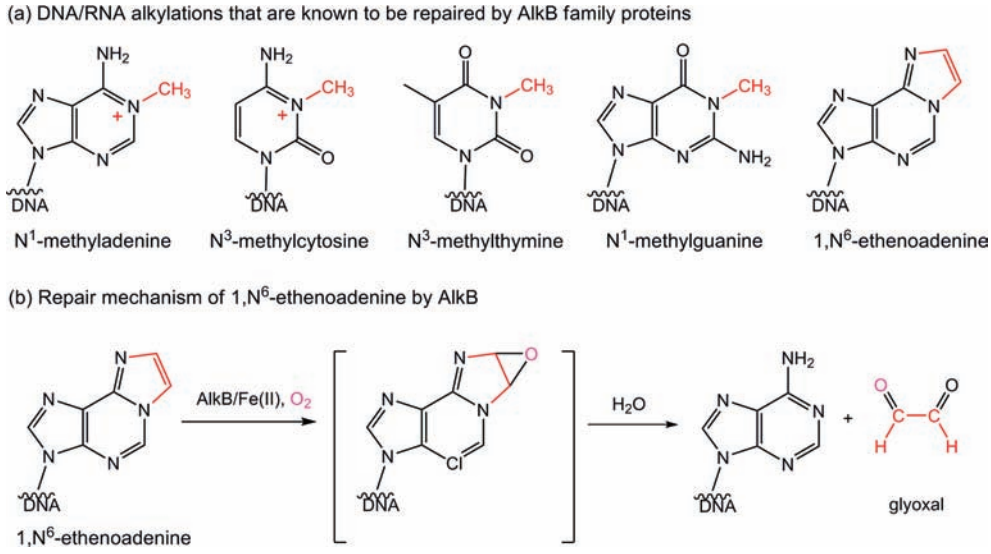


FIGURE 2. (a) Methylated and exocyclic DNA adducts that are repaired by AlkB proteins and (b) proposed repair mechanism for the AlkB-mediated oxidative reversal of εA.

vitro and *in vivo*, and the effect of εC lesions is also mitigated.[24−26] Chemically, εA is oxidized at the etheno bond to give an epoxide intermediate.[25] This result supports the proposed mechanism shown in Figure 2b, in which the putative iron(IV)−oxo intermediate epoxidizes the exocyclic double bond of the base lesion. Subsequent hydrolytic decomposition of the epoxide intermediate liberates the repaired base and dialdehyde glyoxal.[24,25]

The N^1-adenine and N^3-cytosine positions are involved in hydrogen bonding and are inaccessible in dsDNA. If these positions are in a single-stranded oligonucleotide, they are much more susceptible to methylation. Failure to repair these lesions in DNA could lead to cell death.[27] AlkB actually prefers to fix lesions in ssDNA rather than dsDNA (~5−10 fold),[11,28,29] although it is also efficient in repairing lesions in dsDNA. Moreover, studies have revealed its ability to repair 1-meA and 3-meC lesions in RNA as a defense mechanism against alkylation damage.[30,31]

Characterization of the Non-Heme Iron(II) Site of AlkB

Our laboratory reported the overexpression and isolation of the native protein AlkB with bound cofactors directly from *E. coli*.[32] The purified protein sample exhibited a UV−vis band at 560 nm, which is characteristic of a bidentate 2-KG bound to an iron(II) ion (Figure 3). Adding excess ssDNA caused an immediate 9 nm shift to higher energy, suggesting a DNA-binding-induced geometric change at the active site of the protein. This geometric change is further investigated by X-ray absorption spectroscopy, which has indicated that AlkB has a five-coordinate iron(II) geometry in the absence of DNA. Upon

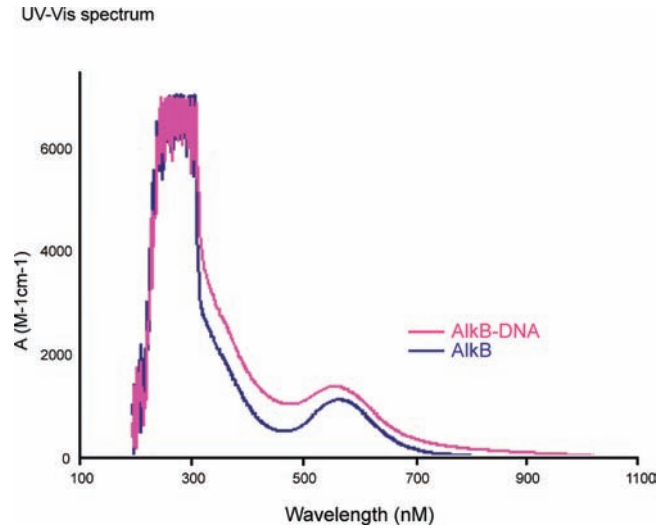


FIGURE 3. UV spectra showing a shift in the absorption peak when AlkB is exposed to ssDNA.

addition of the excess ssDNA, iron(II) becomes six-coordinated.[32] The dioxygen activation for this group of dioxygenases is likely to be regulated by substrate binding. Upon binding of ssDNA, a conformational change may be stimulated, which facilitates the binding of dioxygen to iron(II) and the subsequent oxidation of iron(II).

Human Homologues of *E. coli* AlkB

There are nine human homologues of *E. coli* AlkB; two of them, ABH2 and ABH3, have been shown to repair 1-meA and 3-meC.[17,30,33] A third, FTO, which affects obesity in mammals, recently has been shown to repair 3-meT and 3-meU.[34,35] The fourth one, ABH1, seems to exclusively repair 3-meC in ssDNA with low activity.[36] ABH2 prefers dsDNA while ABH3, like AlkB, works better on single-stranded substrates.[28] AlkB acts

equally well toward 1-meA and 3-meC, while ABH2 is more effective in repairing 1-meA. The situation for ABH3 is just the reverse: it repairs 3-meC more efficiently than 1-meA.[17,33] Both enzymes are ubiquitous in mammalian cells; however, ABH2 is confined within the nucleus, while ABH3 is found in the nucleus and cytoplasm.[17,33,37] Mouse knockout experiments have shown that ABH2-null mice are sensitive to methyl methanesulfonate (MMS)-induced methylation, and they have established ABH2 as the primary enzyme to fix 1-meA lesions in DNA *in vivo*.[37] However, no obvious phenotype has been observed in ABH3-null mice upon treatment with a high but nonlethal dose of MMS. ABH3 may work *in vivo* on targets other than dsDNA. One possibility is RNA because ABH3 has been shown to repair RNA damage *in vitro* and *in vivo*.[30,31] Recently, an exciting discovery indicated that the obesity susceptible gene *FTO* encodes a functional homologue of AlkB.[34] How the DNA/RNA demethylation function of FTO is linked to obesity will be an intensely studied subject in the future.[35]

Structures of AlkB and ABH3

To better understand how these enzymes function at the molecular level, high-resolution crystal structures are strongly desired. The structure of *E. coli* AlkB bound to a short trimer nucleotide d(T-N^1-meA-T) and the structure of ABH3 have been solved recently by two separate groups.[38,39] These structures reveal a common β-strand jelly-roll fold that coordinates a catalytically active iron(II) utilizing a conserved HxD/Ex$_n$H binding motif.[14,15] The AlkB structure consists of three well-defined regions: a catalytic core in the carboxyl-terminal domain, a unique nucleotide-recognition lid, and an N-terminal extension. Several important residues in the active site were also identified. Trp69 and His131 stack against 1-meA to facilitate substrate recognition through $\pi-\pi$ and $\pi-$cation interactions. Tyr76 forms hydrogen bond interactions to the two phosphates 5′ of the methylated base, probably helping to pin the substrate into the active site. Asp135 and Glu136 (through a water molecule) hydrogen bond to the exocyclic amino group of 1-meA, explaining why 1-meA and 3-meC are better substrates than 1-meG and 3-meT.[38] The first reported ABH3 structure also revealed interesting features of the iron(II)/2-KG dependent dioxygenase family. The active site of ABH3 is more polar than that of AlkB and the residue organization seems to be quite different.[39] For detailed comparisons to be made, a bound nucleotide is necessary in the ABH3 structure.

Despite all the efforts mentioned above, there were still significant questions unaddressed. For example, how does AlkB bind longer, more physiologically relevant substrates? AlkB and ABH3 prefer single-stranded substrates, but how do they discriminate dsDNA? Also, the structure of ABH2, the primary house-keeping enzyme in mammalian cells against 1-meA, remained to be solved at the time. These enzymes tend to bind nucleic acids in a weak and non-sequence-specific manner.[32] They are likely to interact with partner proteins and thus gain enhanced affinity toward their physiological substrates *in vivo*. To perform structural studies of their DNA interaction, new strategies must be introduced to stabilize homogeneous complexes. We therefore employed a chemical disulfide cross-linking strategy to covalently stabilize the AlkB−DNA complexes.

Disulfide Cross-Linking To Stabilize Protein on DNA

The disulfide cross-linking technique was first developed by Verdine et al. in the characterizations of structural snapshots of ordinarily transient intermediates and complexes having multiple recognition modes.[40−42] Briefly, the disulfide cross-link is the covalent bond formation between the side chain of a cysteine residue and a thio-tethered modified DNA base (Figure 4a). The cysteine residue is engineered into the DNA binding surface of the protein through site-directed mutagenesis and the alkanethiol tether is synthetically incorporated into DNA. The selection of appropriate covalent linkage is crucial and requires detailed investigation of the system of interest. Generally, trapping strategies fall into two categories: mechanism-based trapping and structure-based trapping.[41] The former depends on a detailed picture of the enzyme mechanism and chemically interrupts the normal course of the reaction. This results in trapping and stabilizing an otherwise transient intermediate in damage recognition or processing. The latter converts a known point of a noncovalent interaction of protein and DNA into a covalent linkage, stabilizing and accumulating a certain state of the complex. For the structural characterizations of the AlkB-family proteins, both aspects of trapping are taken into account. Their combination yields the correct selection of disulfide bonds finally producing the high-resolution structures of the protein−DNA complexes.

Active-Site Cross-Linking

The dilemma we faced at the beginning was the absence of protein−DNA interaction information to set up cross-linking sites. During our initial biochemical studies of AlkB and another class of direct DNA repair proteins, AGT (or C-Ada in *E. coli*), we have developed an active-site cross-linking strategy to circumvent this requirement. Both AlkB and Ada are direct DNA damage repair proteins overexpressed in the adap-

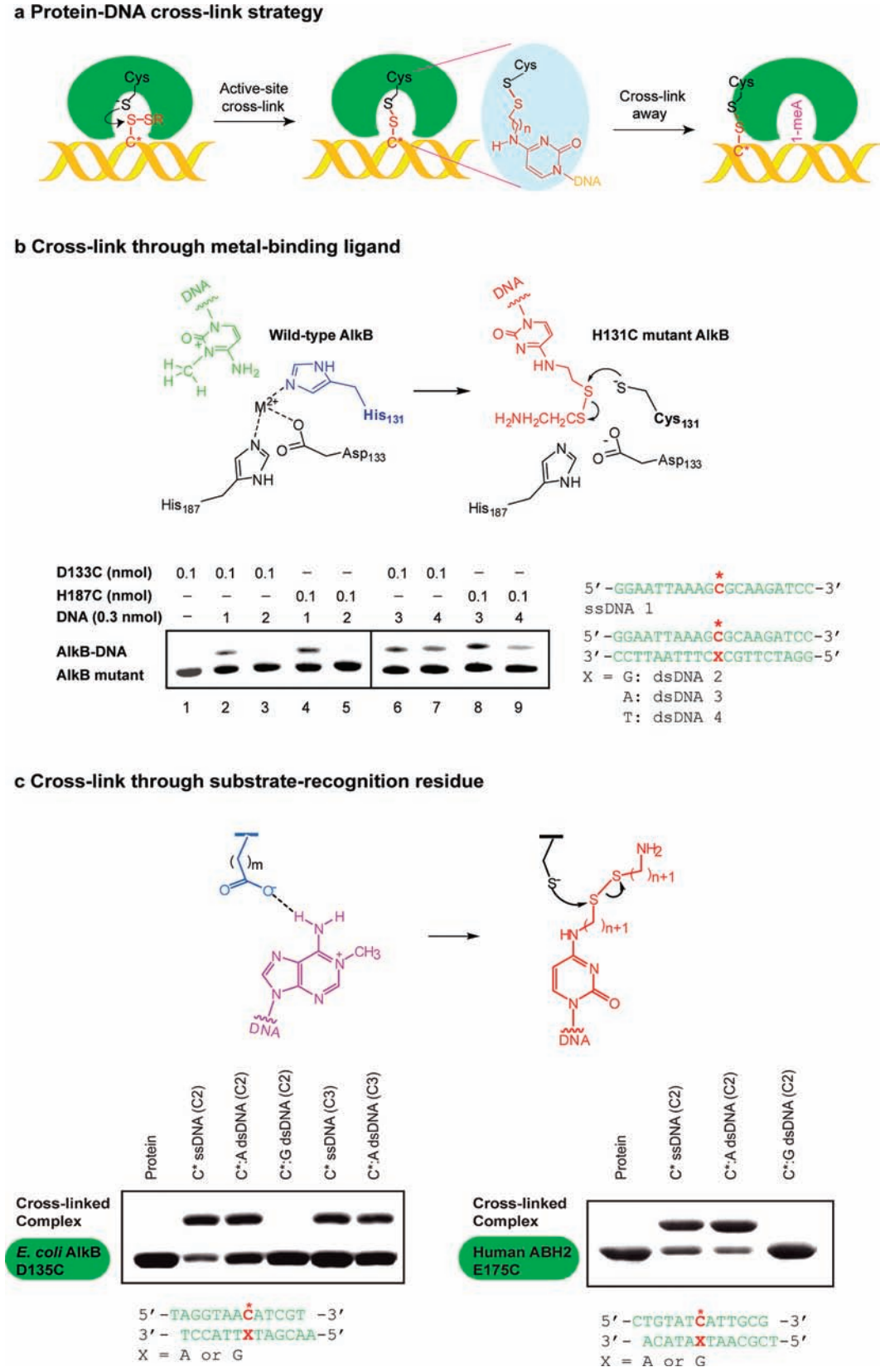


FIGURE 4. (a) The concept of the active-site cross-linking strategy and cross-linking away from the active site. A disulfide-modified cytosine, C* (n = 1 or 2) in a mismatched C*:A base pair can be flipped into the active site of AlkB or ABH2. A covalent disulfide bond can form between C* and an engineered cysteine of the repair protein under equilibrium conditions. (b) Active site cross-linking of AlkB through metal binding ligands. (c) Active site cross-linking of AlkB and ABH2 through substrate recognition residues with various DNA probes.

tive response of *E. coli* to methylating agents.[9,10] The C-terminal domain of Ada (C-Ada) and the human homologue hAGT recognize and repair O^6-meG damage in duplex DNA.[43,44] Both AlkB and the hAGT proteins flip the damaged base into their active sites for lesion recognition and processing. We hypothesized that it might be possible to design a chemical reaction that traps the damage recognition moment in which the lesioned base has been inserted into the active site. The presence and utilization of a reactive cysteine residue in the active site of C-Ada and hAGT for damage removal helped us develop such a cross-linking strategy in our study of these proteins.[45] The oxidation reaction of AlkB itself is difficult to manipulate, but we envisioned that a disulfide-tethered cytosine (C* in Figure 4) may be able to access the substrate binding pocket if there is enough space to accommodate it. This was made possible after the mutation of conserved iron ligand residues to cysteine and the removal of the bound metal ion and cosubstrate 2-KG (Figure 4b).[46] The engineered cysteine residue was expected to attack the disulfide group and form a cross-linked complex if C* can be flipped by AlkB and enter the active site pocket. To our delight, we found that cross-linking did take place when we positioned a mismatched A or T opposite C* in a duplex DNA as shown by the appearance of a new band with retarded mobility on the SDS−PAGE gel (Figure 4b).[46] When we had a matched C*:G base pair in a duplex DNA, we failed to observe reasonable cross-linking. The wild-type AlkB did not form a cross-link with modified dsDNA containing a C*:A or C*:T base pair, which supports the formation of a disulfide bond to the engineered cysteine residue in the active site of AlkB. Sensing an unstable base pair appears to be a key step for AlkB to locate potential base damage in duplex DNA. The same property was also observed for its human homologues.[47]

The cross-linking experiments were performed in the absence of any detailed structural information. The yields were still low, perhaps due to potential constraint introduced through the covalent linkage. In 2006, the first crystal structure of AlkB bound to a short trimer T-N^1-meA-T DNA became available. In the structure, the side chain of Asp135 is within hydrogen-bonding distance of the exocyclic amino group of 1-meA. We decided to convert this noncovalent interaction into a covalent disulfide bond to improve the cross-linking yield. This AlkB D135C mutant was prepared and reactions for cross-linking with various DNA probes bearing C*:A were carried out under equilibrium conditions (Figure 4c). It cross-linked efficiently with the ssDNA probe and a dsDNA containing an unstable central base pair C*:A. In the presence of a stable C*:G base pair, AlkB formed negligible cross-linking product with the DNA probe, indicating that the enzyme is unable to actively break stable base pairs in duplex DNA.

Based on sequence alignment, we also prepared the ABH2 E175C mutant, the counterpart of AlkB D135C. As expected, the ABH2 mutant protein also formed efficient cross-linking with disulfide tethered dsDNA. The active-site cross-linked complexes for both AlkB and ABH2 were readily separated by Mono Q anion-exchange chromatography and the purified complexes were used for crystallization screens. By varying the sequences of DNA, we succeeded in crystallizing and solving the first structures of AlkB−dsDNA and ABH2−dsDNA complexes.[29]

Distal Cross-Linking

An advantage of AlkB D135C and ABH2 E175C cross-linking over the iron ligand mutant is that they allow binding of a metal ion and the cosubstrate 2-KG in the active sites of these proteins.[29] However, in these two structures, mechanisms of lesion recognition were not revealed due to the presence of C* and the disulfide bond in the active sites. This is particularly unsatisfactory for ABH2; the structure and the residue arrangement in the active site were not reported previously. However, detailed interactions between protein and DNA are available from the active-site cross-linked structures. Subsequently, new sites of the disulfide cross-link away from the active site were selected for structure-based trapping. AlkB S129C and ABH2 G169C mutant proteins were chosen to form cross-links distal from the active site pockets. Both residues, AlkB Ser129 and ABH2 Gly169, are two base pairs away from the flipped bases and close to DNA. Both proteins present the major groove of that part of DNA toward themselves. Thus, a tether at the N^4 position of cytosine will point to the protein and is expected to form a covalent bond with an engineered cysteine. The two mutants were prepared, and cross-linking was found to occur as anticipated. Crystals of the complexes with cross-linking away from the active site were obtained using the same DNA sequences that yielded the active-site cross-linking structures under slightly different conditions.[29] Notably, the protein portions of the active-site cross-linked structure and cross-linking away structure overlap perfectly with each other for both AlkB and ABH2, indicating that no noticeable interruption has been introduced by the formation of a covalent disulfide bond.

Overall, disulfide cross-linking reactions were performed under equilibrium conditions, avoiding the kinetically trapped high-energy intermediate states, which are often the concern for photo-cross-linking reactions. The selection of appropriate sites to install a cross-link is crucial in the success of the

structural resolution of the complex. Through active-site cross-linking, we can obtain the first sets of protein−DNA complex structures. Once the detailed protein−DNA interactions are revealed, a cross-link away from the active site can be readily engineered to provide native-like structures of the protein−DNA complexes (Figure 4a).[29]

Structures of AlkB and ABH2 Bound to dsDNA

The structures of AlkB and ABH2 bound to dsDNA showed that both proteins have a common jelly-roll fold, which is shared by other iron(II)/2-KG-dependent dioxygenases, and flip 1-meA into an active site pocket (Figure 5a,b).[14,15] However, the two enzymes interact with dsDNA in quite different manners (Figure 5c).[29] AlkB interacts almost exclusively with the strand bearing the damaged base, while ABH2 makes significant contact with both strands. Moreover, ABH2 has three extra β-sheets compared with AlkB: one is at its N-terminus and the other two are crucial antiparallel sheets between ABH2's β5 and α1. The last two constitute an important part of its base-flipping motifs, a finger residue Phe102 at the tip of the hairpin between β3 and β4 (Figure 5e). ABH2 uses this aromatic finger to intercalate into the duplex stack and fill the DNA gap induced by base-flipping. AlkB does not possess such a motif and must utilize a different strategy to flip the damaged base.

Base Flipping by AlkB

AlkB manipulates the DNA backbone near the flipped 1-meA base so that the two flanking bases, A7 and A9, are "squeezed" and stacked together, ~3.4 Å apart (Figure 6a). The conformations of the sugar rings of the two nucleosides are quite unusual: A9 adopts the C3′-*endo* conformation, standard in A-form DNA, and surprisingly, A7 along with its base is twisted by ~180°. This twist results in an inverted sugar with O4′ pointing toward the 3′ end of the DNA strand. The inversion of the sugar ring is locally enforced by residues 51−53 (Thr-Pro-Gly), which "pinch" the phosphate linking A6 and A7, and through multiple protein−DNA interactions anchoring the two phosphate groups flanking 1-meA.

The complementary stand accommodates the unique conformation of the lesion-containing strand through spontaneous conformational rearrangements. The inversion of A7 would leave a ~6.8 Å (two base pair) gap between A7 and A6. This vacancy is filled by the intercalation of the base opposite 1-meA, which generates a local one-base wide stack that contains bases from both strands (Figure 6a). The base at the 7′ position is forced out of the duplex and is disordered in the structure. To support the observed conformational rearrangements of the complementary strand, the context dependence of AlkB-mediated repair was measured.[29] AlkB can efficiently repair the 1-meA lesion in the sequence used in crystallization, which has two A:T base pairs flanking 1-meA:T. When they were replaced with more stable G:C pairs, repair activity dropped by ~3-fold. The introduction of an abasic site at the 7′ position to the G:C rich sequence, which destabilizes the base pair 7:7′ (the base at the 7′ position is forced out of the duplex in the crystal structure), restored the repair efficiency to the original level. The highest repair activity (at least 6-fold higher compared with the dsDNA sequence used in crystallization) was observed with the ssDNA containing 1-meA (Figure 6b).

Base Flipping and Lesion Recognition by ABH2

ABH2 uses commonly observed strategies to flip the damaged base in dsDNA.[29] First, it has a crucial hairpin motif between β3 and β4 that bears the finger residue Phe102 (Figure 7a). This intercalating residue fills the void created by the flipped 1-meA and maintains the normal length and stacking of duplex DNA. Second, ABH2 has several extra DNA binding motifs that grasp the complementary strand of dsDNA. It uses the positively charged RKK loop (Arg 241−Lys 243) between β14 and β15 and a flexible long loop between β11 and β12 to bind the opposite DNA strand. The presence of these extra binding motifs establishes ABH2 as a dsDNA repair protein (Figure 5b,e).

ABH2 also possesses a more complex residue arrangement in the active site than AlkB to recognize the 1-meA lesion (Figure 7b). Phe124 and His171 stack against 1-meA; an analogous interaction was observed for 1-meA recognition by Trp69 and His131 in AlkB. The N6 and N7 positions of 1-meA are recognized by a network of hydrogen-bonding residues including Tyr122, Glu175, Asp174, and a water molecule. Only the N6 nitrogen atom is recognized by Asp135 and Glu136 in AlkB, while the N7 position of 1-meA is unrecognized. The metal binding site of ABH2 is standard for iron(II)/2-KG-dependent dioxygenases (Figure 7b). Manganese(II), which occupies the iron(II) site without supporting catalysis, is used in these studies. The metal ion is ligated by His171, Asp173, His236, a bidentate 2-KG, and a water molecule in an octahedral geometry. Similar interactions are observed in the AlkB/DNA structure.

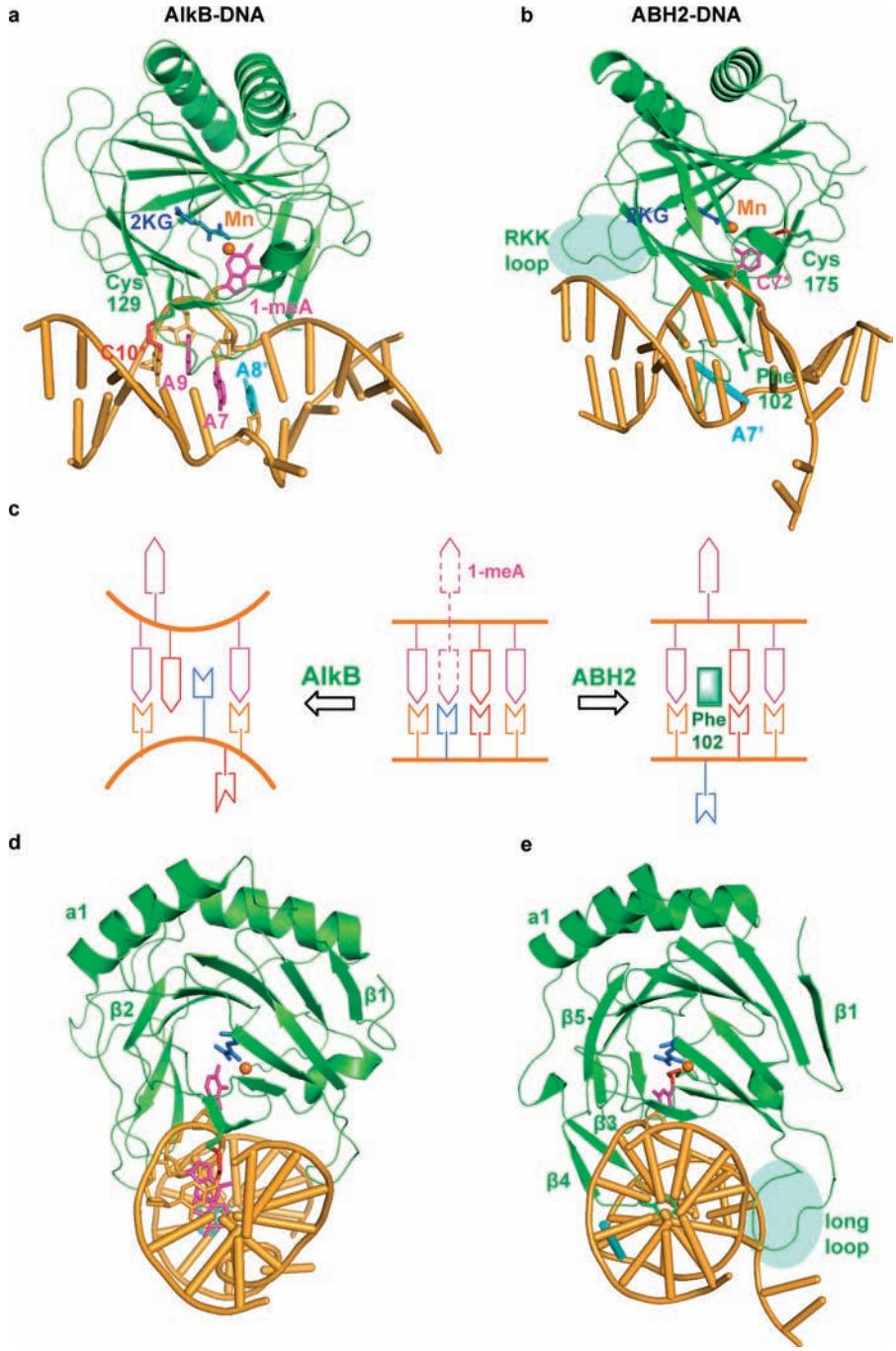


FIGURE 5. Crystal structures of the cross-linking away AlkB−DNA complex and the active site cross-linked ABH2−DNA complex. (a) AlkB−DNA structure with a cross-link away from the active site. The protein is shown in green (the cross-linking residue Cys129 is labeled), manganese(II) in orange, 2-KG in blue, DNA in yellow-orange, the flipped base 1-meA and the two flanking bases A7 and A9 in light magenta, the disulfide bond in red, and the orphaned base A8′ in cyan. (b) Active site cross-linked ABH2−DNA structure. The color coding as in panel a is used. The finger residue Phe102 and the cross-linking residue Cys175 are green, and the orphaned base A7′ is cyan. The location of a DNA binding loop containing an RKK sequence is indicated. (c) Cartoons showing two different strategies used by AlkB and ABH2 to flip out 1-meA from the duplex DNA. (d, e) The same structures as in panels a and b, but rotated 90° to the right. A DNA binding flexible long loop of ABH2 is labeled in panel e.

Preference for ssDNA or dsDNA

Comparisons can be made with the structures of AlkB−dsDNA and ABH2−dsDNA. The AlkB protein, lacking a finger residue, promotes base-flipping by squeezing together the two flanking bases such that they stack onto one another. This base flipping mechanism is unprecedented and has only been observed for AlkB so far. It helps to explain the preference of AlkB for ssDNA over dsDNA. This enzyme interacts almost

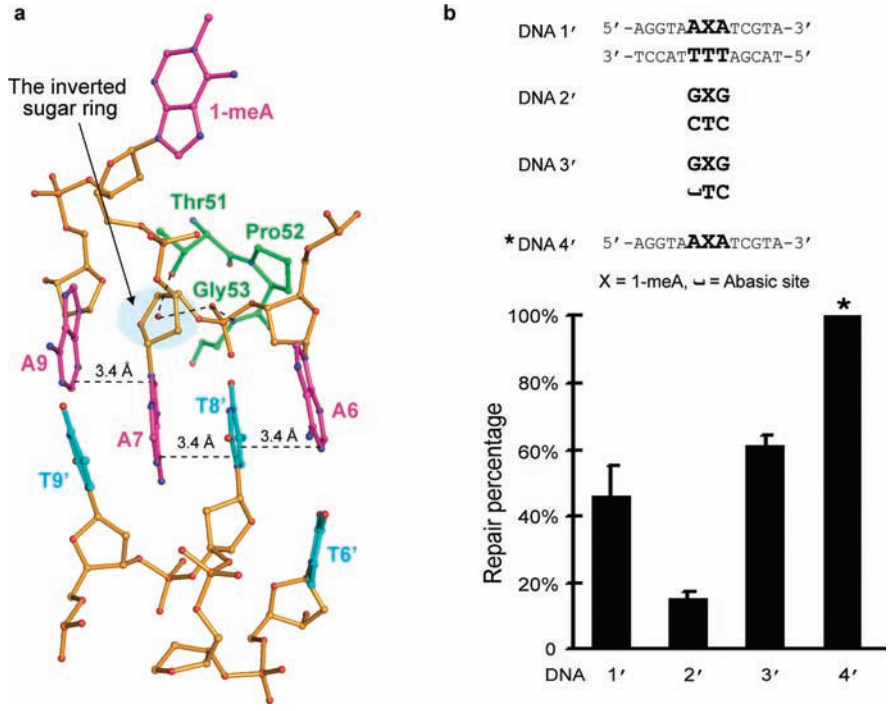


FIGURE 6. (a) Close view of the base-flipping region of AlkB, with the same coloring as in Figure 5a. A7 stacks with the A9:T9′ base pair. Residues Thr51 and Gly53 are hydrogen bonded to the phosphate linking A7 and A6. The sugar ring of A9 adopts the 3′-*endo* conformation, while the sugar ring of A7 is forced to be inverted by ~180° by the protein. T8′ from the complementary strand intercalates between A7 and A6, forming a one-base wide stack containing bases from both strands. (b) Activity assay designed to measure the context dependence of repair efficiency of AlkB (45 min for dsDNA and 15 min for ssDNA).

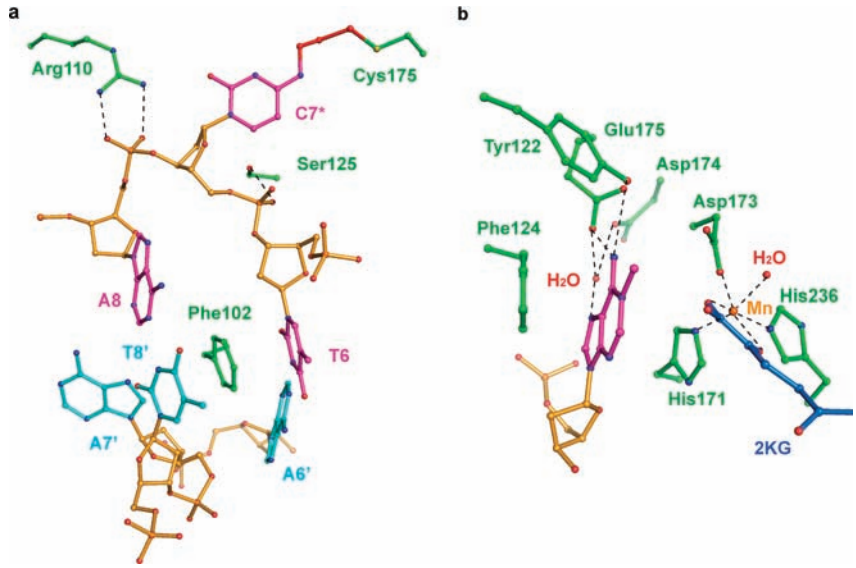


FIGURE 7. (a) Close view of the base-flipping region of ABH2, with the same coloring as in Figure 5a. Phe102 inserts into the DNA duplex with A7′ as the orphaned base. (b) Close view of the ABH2 active site obtained from the cross-linking away structure, with manganese(II) (orange), 2-KG (blue), and 1-meA (light magenta). N6 of 1-meA is within hydrogen bond distance to Tyr122 (3.3 Å) and Glu175 (2.8 Å). N7 forms a hydrogen bond to a water molecule (3.0 Å), which also interacts with Asp174 (2.6 Å) and Glu175 (2.6 Å).

exclusively with the lesion-containing strand, while the complementary strand has to accommodate its unique conformation through spontaneous conformational rearrangements. This distortion of the relatively rigid duplex DNA structure leads to an energetic penalty that must be compensated by the binding energy of the protein to dsDNA. With flexible ssDNA, a much lower energetic penalty is associated with base-flipping, which contributes to AlkB's preference toward

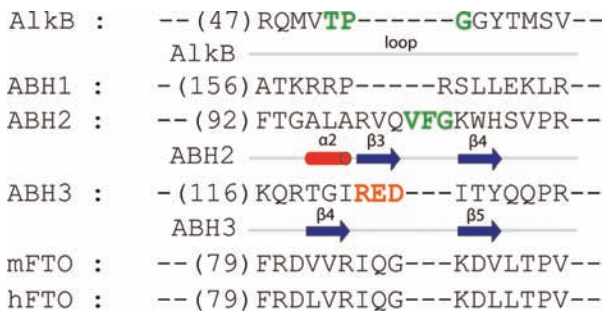


FIGURE 8. Structure−sequence alignment of the AlkB-family proteins in the region containing the base-flipping hairpin motif in ABH2 and ABH3.

ssDNA over dsDNA. On the other hand, ABH2, possessing an intercalating finger residue to maintain the duplex structure after base-flipping, interacts extensively with both strands of duplex DNA. This protein binds dsDNA preferentially over ssDNA, which correlates well with its house-keeping role against 1-meA in human genomes.

Nucleotide-Flipping by the AlkB Family Proteins

A close inspection of the sequence−structure alignment of the AlkB-family proteins reveals several interesting aspects (Figure 8). ABH2 utilizes a hairpin motif bearing a finger residue to intercalate the duplex, while AlkB does not possess such a motif and uses a short "pinch" (residues 51−53 Thr-Pro-Gly) to squeeze together the two flanking bases to facilitate base-flipping.[29] ABH3 also contains a similar β-sheet hairpin motif as presented in ABH2. Within this hairpin region, mutation studies have shown that Arg122 and Glu123 play important roles in dsDNA and ssDNA repair of ABH3.[39] We suggest that this hairpin region may be engaged in a more profound distortion of the DNA duplex when ABH3 recognizes dsDNA. The other two functional homologues, ABH1 and FTO, were shown to work exclusively on single-stranded substrates.[34−36] The conserved region of ABH1 resembles more AlkB. Both human and mouse FTO proteins seem to have a β-sheet hairpin similar to that of ABH3 based on the secondary structure prediction (Figure 8).[34]

Conclusions and Implications

The AlkB-family proteins perform oxidative dealkylation repair of base damages in DNA/RNA.[11,12,30,31] Recent studies of FTO and ABH1, the AlkB-family proteins in humans, suggest that these oxidases may also have broad regulatory functions.[34−36] They likely catalyze oxidative demethylation on key regulatory nucleic acid components that have yet to be revealed. Unlike other DNA base repair family proteins, the majority of the AlkB proteins have a preference for ssDNA over dsDNA.[28,35,36] We have shown that a simple DNA distortion mechanism is used by AlkB to facilitate base-flipping and to discriminate against more rigid dsDNA vs flexible ssDNA.[29] This delicate mechanism could be used by other ssDNA or RNA binding proteins to concentrate on their cellular substrates in the presence of excess dsDNA. The AlkB-family proteins bind their substrate nucleic acids weakly and nonspecifically *in vitro*, often rendering it difficult to obtain a homogeneous protein−DNA complex sample for structural studies.[46] We have developed an active-site cross-linking technique to covalently trap and characterize the base-flipping event.[45−47] After detailed interactions are revealed from the active-site cross-linked structures, the disulfide link can be installed outside of the active site to provide native-like structures of the complexes.[29] This strategy can be widely applied to other base-flipping DNA/RNA repair and modification proteins such as FTO, ABH3, DNA methyltransferases, and DNA/RNA deaminases. Overall, our cross-linking strategy offers a general solution to trap labile protein−DNA and protein−RNA interactions for base-flipping proteins.

This work was supported by the National Institutes of Health (Grant GM071440), the W. M. Keck Foundation Distinguished Young Scholar in Medical Research Program (C.H.), and A Cottrelll Scholar award (C.H.) from the Research Corporation. We thank Ms. Kristel Groth-Garcia for help with editing the manuscript.

BIOGRAPHICAL INFORMATION

Chengqi Yi is a graduate student in the Department of Chemistry at The University of Chicago.

Cai-Guang Yang was a research scientist in the Department of Chemistry at The University of Chicago and now is a professor at Shanghai Institute of Materia Medica, Chinese Academy of Sciences.

Chuan He received his Ph.D. from Massachusetts Institute of Technology, where he worked with Stephen J. Lippard. He did his postdoctoral work with Gregory L. Verdine at Harvard University. He is an associate professor of Chemistry at The University of Chicago.

FOOTNOTES

*To whom correspondence should be addressed. E-mail: chuanhe@uchicago.edu.

REFERENCES

1 Sedgwick, B. Repairing DNA-methylation damage. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 148–157.

2 Rydberg, B.; Lindahl, T. Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J.* **1982**, *1*, 211–216.

3 Lindahl, T.; Sedgwick, B.; Sekiguchi, M.; Nakabeppu, Y. Regulation and expression of the adaptive response to alkylating agents. *Annu. Rev. Biochem.* **1988**, *57*, 133–157.

4 Barrows, L. R.; Magee, P. N. Nonenzymatic methylation of DNA by S-adenosylmethionine in vitro. *Carcinogenesis* **1982**, *3*, 349–351.

5 Taverna, P.; Sedgwick, B. Generation of an endogenous DNA-methylating agent by nitrosation in *Escherichia coli*. *J. Bacteriol.* **1996**, *178*, 5105–5111.

6 Hecht, S. S. DNA adduct formation from tobacco-specific N-nitrosamines. *Mutat. Res.* **1999**, *424*, 127–142.

7 Hurley, L. H. DNA and its associated processes as targets for cancer therapy. *Nat. Rev. Cancer* **2002**, *2*, 188–200.

8 Rajski, S. R.; Williams, R. M. DNA cross-linking agents as antitumor drugs. *Chem. Rev.* **1998**, *98*, 2723–2795.

9 Kataoka, H.; Yamamoto, Y.; Sekiguchi, M. A new gene (*alkB*) of *Escherichia coli* that controls sensitivity to methylmethane sulfonate. *J. Bacteriol.* **1983**, *153*, 1301–1307.

10 Teo, I.; Sedgwick, B.; Kilpatrick, M. W.; McCarthy, T. V.; Lindahl, T. The intracellular signal for induction of resistance to alkylating agents in *E. coli*. *Cell* **1986**, *45*, 315–324.

11 Falnes, P. Ø; Johansen, R. F.; Seeberg, E. AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature* **2002**, *419*, 178–182.

12 Trewick, S. C.; Henshaw, T. F.; Hausinger, R. P.; Lindahl, T.; Sedgwick, B. Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature* **2002**, *419*, 174–178.

13 Aravind, L.; Koonin, E. V. The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. *Genome Biol.* **2001**, *2*, research0007.

14 Schofield, C. J.; Zhang, Z. Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. *Curr. Opin. Struct. Biol.* **1999**, *9*, 722–731.

15 Que, L. J. One motif-many different reactions. *Nat. Struct. Biol.* **2000**, *7*, 182–184.

16 Koivisto, P.; Duncan, T.; Lindahl, T.; Sedgwick, B. Minimal methylated substrate and extended substrate range of *Escherichia coli* AlkB protein, a 1-methyladenine-DNA dioxygenase. *J. Biol. Chem.* **2003**, *278*, 44348–44354.

17 Duncan, T.; Trewick, S. C.; Koivisto, P.; Bates, P. A.; Lindahl, T.; Sedgwick, B. Reversal of DNA alkylation damage by two human dioxygenases. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16660–16665.

18 Dinglay, S.; Gold, B.; Sedgwick, B. Repair in *Escherichia coli* alkB mutants of abasic sites and 3-methyladenine residues in DNA. *Mutat. Res.* **1998**, *407*, 109–116.

19 Delaney, J. C.; Essigmann, J. M. 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.* **2004**, *101*, 14051–14056.

20 Koivisto, P.; Robins, P.; Lindahl, T.; Sedgwick, B. Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases. *J. Biol. Chem.* **2004**, *279*, 40470–40474.

21 Falnes, P. Ø. Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. *Nucleic Acids Res.* **2004**, *32*, 6260–6267.

22 Nair, J.; Barbin, A.; Guichard, Y.; Bartsch, H. 1,N^6-Ethenodeoxyadenosine and 3,N^4-ethenodeoxycytidine in liver DNA from humans and untreated rodents detected by immunoaffinity/^{32}P-postlabelling. *Carcinogenesis* **1995**, *16*, 613–617.

23 Chung, F. L.; Chen, H. J.; Nath, R. G. Lipid peroxidation as a potential endogenous source for the formation of exocyclic DNA adducts. *Carcinogenesis* **1996**, *17*, 2105–2111.

24 Mishina, Y.; Yang, C.-G.; He, C. Direct repair of the exocyclic DNA adduct 1,N^6-ethenoadenine by the DNA repair AlkB proteins. *J. Am. Chem. Soc.* **2005**, *127*, 14594–14595.

25 Delaney, J. C.; Smeester, L.; Wong, C.; Frick, L. E.; Taghizadeh, K.; Wishnok, J. S.; Drennan, C. L.; Samson, L. D.; Essigmann, J. M. AlkB reverses etheno DNA lesions caused by lipid oxidation *in vitro* and *in vivo*. *Nat. Struct. Mol. Biol.* **2005**, *12*, 855–860.

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

27 Chen, B. J.; Carroll, P.; Samson, L. The *Escherichia coli* AlkB protein protects human cells against alkylation-induced toxicity. *J. Bacteriol.* **1994**, *176*, 6255–6261.

28 Falnes, P. O.; Bjoras, M.; Aas, P. A.; Sundheim, O.; Seeberg, E. Substrate specificities of bacterial and human AlkB proteins. *Nucleic Acids Res.* **2004**, *32*, 3456–3461.

29 Yang, C.-G.; Yi, C.; Duguid, E. M.; Sullivan, C. T.; Jian, X.; Rice, P. A.; He, C. Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA. *Nature* **2008**, *452*, 961–965.

30 Aas, P. A.; Otterlei, M.; Falnes, P. Ø.; Vågbo, C. B.; Skorpen, F.; Akari, M.; Sundheim, O.; Bjørås, M.; Slupphaug, G.; Seeberg, E.; Krokan, H. Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature* **2003**, *421*, 859–863.

31 Ougland, R.; Zhang, C. M.; Liiv, A.; Johansen, R. F.; Seeberg, E.; Hou, Y. M.; Remme, J.; Falnes, P. Ø. AlkB restores the biological function of mRNA and tRNA inactivated by chemical methylation. *Mol. Cell* **2004**, *16*, 107–116.

32 Mishina, Y.; Chen, L. X.; He, C. Preparation and characterization of the native iron(II)-containing DNA repair AlkB protein directly from *Escherichia coli*. *J. Am. Chem. Soc.* **2004**, *126*, 16930–16936.

33 Lee, D.-H.; Jin, S.-G.; Cai, S.; Chen, Y.; Pfeifer, G. P.; O'Connor, T. R. Repair of methylation damage in DNA and RNA by mammalian AlkB homologues. *J. Biol. Chem.* **2005**, *280*, 39448–39459.

34 Gerken, T.; Girard, C. A.; Tung, Y.-C. L.; Webby, C. J.; Saudek, V.; Hewitson, K. S.; Yeo, G. S. H.; McDonough, M. A.; Cunliffe, S.; McNeill, L. A.; Galvanovskis, J.; Rorsman, P.; Robins, P.; Prieur, X.; Coll, A. P.; Ma, M.; Jovanovic, Z.; Farooqi, I. S.; Sedgwick, B.; Barroso, I.; Lindahl, T.; Ponting, C. P.; Ashcroft, F. M.; O'Rahilly, S.; Schofield, C. J. The obesity-associated *FTO* gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science* **2007**, *318*, 1469–1472.

35 Jia, G.; Yang, C.-G.; Yang, S.; Jian, X.; Yi, C.; Zhou, Z.; He, C. Oxidative demethylation of 3-methylthymine and 3-methyluracil in single-stranded DNA and RNA by mouse and human FTO. *FEBS Lett.* **2008**, *582*, 3313–3319.

36 Westbye, M. P,; Feyzi, E.; Aas, P. A.; Vågbø, C. B.; Talstad, V. A.; Kavli, B.; Hagen, L.; Sundheim, O.; Akbari, M.; Liabakk, N. B.; Slupphaug, G.; Otterlei, M.; Krokan, H. E. Human AlkB homolog 1 is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA. *J. Biol. Chem.* **2008**, *283*, 25046–25056.

37 Ringvoll, J.; Nordstrand, L. M.; Vagbo, C. B.; Talstad, V.; Reite, K.; Aas, P. A.; Lauritzen, K. H.; Liabakk, N. B.; Bjork, A.; Doughty, R. W.; Falnes, P. Ø.; Krokan, H. E.; Klungland, A. Repair deficient mice reveal mABH2 as the primary oxidative demethylase for repairing 1meA and 3meC lesions in DNA. *EMBO J.* **2006**, *25*, 2189–2198.

38 Yu, B.; Edstrom, W. C.; Benach, J.; Hamuro, Y.; Weber, P. C.; Gibney, B. R.; Hunt, J. F. Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. *Nature* **2006**, *439*, 879–884.

39 Sundheim, O.; Vagbo, C. B.; Bjoras, M.; de Sousa, M. M. L.; Talstad, V.; Aas, P. A.; Drablos, F.; Krokan, H. E.; Tainer, J. A.; Slupphaug, G. Human ABH3 structure and key residues for oxidative demethylation to reverse DNA/RNA damage. *EMBO J.* **2006**, *25*, 3389–3397.

40 Huang, H.; Chopra, R.; Verdine, G. L.; Harrison, S. C. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase at 2.7 Å resolution: Implications of conformational changes for polymerization and inhibition mechanism. *Science* **1998**, *282*, 1669–1675.

41 Verdine, G. L.; Norman, D. P. G. Covalent trapping of protein—DNA complexes. *Annu. Rev. Biochem.* **2003**, *72*, 337–366.

42 Banerjee, A.; Santos, W. L.; Verdine, G. L. Structure of a DNA glycosylase searching for lesions. *Science* **2006**, *311*, 1153–1157.

43 Daniels, D. S.; Woo, T. T.; Luu, K. X.; Noll, D. M.; Clarke, N. D.; Pegg, A. E.; Tainer, J. A. DNA binding and nucleotide flipping by the human DNA repair protein AGT. *Nat. Struct. Mol. Biol* **2004**, *11*, 714–720.

44 Duguid, E. M.; Rice, P. A.; He, C. The structure of the human AGT protein bound to DNA and its implications for damage detection. *J. Mol. Biol.* **2005**, *350*, 657–666.

45 Duguid, E. M.; Mishina, Y.; He, C. How do DNA repair proteins locate potential base lesions? A chemical cross-linking method to investigate the damage-searching mechanism of O6-alkylguanine-DNA alkyltransferases. *Chem. Biol.* **2003**, *10*, 827–835.

46 Mishina, Y.; He, C. Probing the structure and function of the *Escherichia coli* DNA alkylation repair AlkB protein through chemical cross-linking. *J. Am. Chem. Soc.* **2003**, *125*, 8730–8731.

47 Mishina, Y.; Lee, C.-H. J.; He, C. Potential DNA preferences of human and bacterial AlkB proteins as revealed from chemical crosslink studies. *Nucleic Acids Res.* **2004**, *32*, 1548–1554.