

Damage Detection and Base Flipping in Direct DNA Alkylation Repair

Cai-Guang Yang,^[a] Kristel Garcia,^[b] and Chuan He^{*[b]}

To remove a few damaged bases efficiently from the context of the entire genome, the DNA base repair proteins rely on remarkably specific detection mechanisms to locate base lesions. This efficient molecular recognition event inside cells has been extensively studied with various structural and biochemical tools. These studies suggest that DNA base damage can be located by repair proteins by using two mechanisms: a repair protein can probe and detect a weakened base pair that results from mutagenic or cytotoxic base damage; alternatively, a protein can passively capture and stabilize an extrahelical base lesion. Our chemical and struc-

tural studies on the direct DNA repair proteins hAGT, C-Ada and ABH2 suggest that these proteins search for weakened base pairs in their first step of damage searching. We have also discovered a very unique base-flipping mechanism used by the DNA repair protein AlkB. This protein distorts DNA and favors single stranded DNA (ssDNA) substrates over double-stranded (dsDNA) ones. Potentially, it locates base lesions in dsDNA by imposing a constraint that targets less rigid regions of the duplex DNA. The exact mechanism of how AlkB and related proteins search for damage in ssDNA and dsDNA still awaits further studies.

Introduction

DNA base damage, the most common type of DNA damage, occurs at the rate of "several thousand base pairs per cell per day in humans".^[1] Base damage repair is initiated by proteins that recognize the damaged base and then carry out either base excision repair (BER), or fix the damage through direct damage reversal. These repair processes occur extrahelically with the damaged base flipped out of the duplex DNA.^[2–6] To minimize the mutagenic and/or cytotoxic consequences of alkylating agents from both environmental and endogenous sources, evolution has provided human cells with two direct DNA repair systems. These are O⁶-alkylguanine alkyltransferase and the AlkB family oxidative demethylases; both directly eliminate alkylated DNA damage and restore their normal function.^[7] For base-specific DNA repair proteins to act promptly, it is very important that they can locate the very few damaged bases in the entire genome. Questions still remain about the mechanism behind the search for damaged bases and the strategy used to perform base flipping. This article focuses on the structural biochemistry of proteins that enable direct reversal of DNA base damage. This focus is on the initial DNA repair step that is taken by these proteins and the two themes for discussion are: 1) intrahelically or extrahelically damaged base searching and 2) active or passive base flipping.

flipping". Because flipped-out bases are more accessible to solvents or other molecules and are more prone to form complexes with proteins, it is logical that base flipping is involved in many enzymatic DNA modification and repair reactions. Structures of protein–DNA complexes for several base repair and modification proteins have shown that the proteins gain access to their substrates by flipping out and inserting the damaged bases into their active-site pockets; for example, human alkyladenine glycosylase (hAAG), human O⁶-alkylguanine alkyltransferase (hAGT), 8-oxoguanine DNA glycosylase (hOGG), human uracil-DNA glycosylase (UDG or UNG), and the recently solved oxidative DNA/RNA dealkylases, *Escherichia coli* AlkB and its human homologue ABH2, which is bound to double-stranded DNA (dsDNA), all use this repair mechanism (Figure 1).^[8–12]

Structural studies of human AAG^[8,13] and other DNA glycosylases^[11] have established the base-flipping mechanism for repair proteins as a means of recognizing and processing damaged bases in the DNA helix. It is known that once the base is flipped, it is tightly bound into an enzyme active site, but how does the repair protein locate the lesioned base in the first place? A base-repair protein could partially or completely unstack nucleotides from the DNA helix to find the damaged bases. Once a nucleotide has been rotated into an extrahelical

Overall View in Damaged Base Detection

Base-specific DNA repair proteins

As in many other cellular activities, the momentary interaction that occurs between protein and nucleic acid is universal and essential for DNA repair processes. The movement of a DNA base from the base-stacked, hydrogen-bound, intrahelical position to a solvent-exposed, extrahelical position is termed "base

[a] Prof. C.-G. Yang
Shanghai Institute of Materia Medica
Chinese Academy of Sciences
555 Zuchongzhi Road, Shanghai 201203 (China)

[b] K. Garcia, Prof. C. He
Department of Chemistry, The University of Chicago
929 East 57th Street, Chicago, IL 60637 (USA)
Fax: (+1) 773-702-0805
E-mail: chuanhe@uchicago.edu

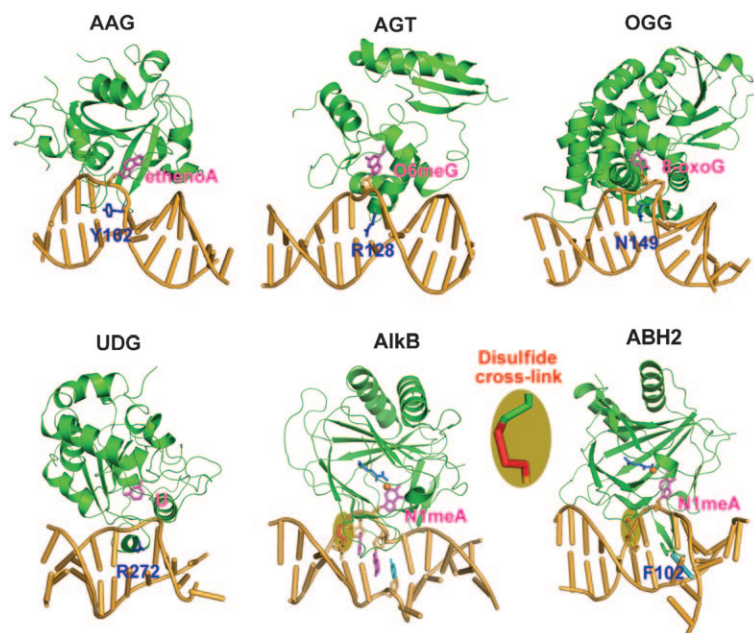


Figure 1. Cartoon representation of selected base-specific repair proteins bound to dsDNA. Proteins are colored green, flipped-out bases are magenta, finger residues are blue, and DNA is bright orange. *E. coli* AlkB (PDB ID: 3BIE) is particularly unique in the way that it flips out a 1-meA damaged base into the active pocket, which lacks a finger residue. Human AAG (PDB ID: 1EWN), AGT (PDB ID: 1T38), OGG (PDB ID: 1EBM), UDG (PDB ID: 4SKN), and ABH2 (PDB ID: 3BUC) proteins flip ϵ A, O^6 -methylG, 8-oxoguanine, uracil, and 1-methyladenine out of the DNA helix into specific recognition pockets by intercalating a finger residue into the double helix, respectively.

conformation, the local melting of the DNA might facilitate the subsequent flipping of neighboring nucleotides as the protein slides along the DNA while searching for damage.^[13] In these nucleotide-flipping protein–DNA complexes, an enzyme amino acid side chain, termed “finger”, is inserted into the site that is left empty by the flipped base. It is thought that this might aid in pushing the damaged base and flipping the nucleotide.^[14] In contrast, an opposing mechanism for enzymatic detection of damaged bases has been provided through the recent dynamic study of human UDG. According to this mechanism, the discrimination of the unwanted bases by UDG is started by thermally induced base-pair openings and not through active participation of the enzyme.^[15] The detection pathway by which the base-transient state is achieved is of fundamental importance and great interest.

DNA damage-detection mechanisms

It is well accepted that most base-specific DNA repair proteins, including AGT and perhaps AlkB proteins, repair the lesioned bases extrahelically. The theoretical study of the energy required for undamaged DNA bases to go through base flipping indicates that the chances of undamaged bases flipping in free solution is quite low and almost unlikely.^[16] Thus, the base flipping must be assisted by repair proteins.^[17] We have proposed three mechanisms for how base-specific proteins locate damaged DNA bases.^[18] First, as shown in Figure 2A, every base could be flipped and checked in an active searching mechanism. To locate the damaged base, this mechanism requires that the protein actually travels along the DNA flipping and checking every base in the active site, which seems to be unlikely considering recent experimental evidence.^[19] Alternatively, the repair proteins do not actively flip out every base. The protein could slide through the duplex DNA and physically detect unstable or non-Watson–Crick base pairing in the intrahelical conformation (Figure 2B). Base modifications incurred from damage could alter the structure of the base-pairing interface and prevent the damaged base from forming a stable Watson–Crick base pair with the opposite base on the complementary strand. Thus, the repair protein might test the stabi-

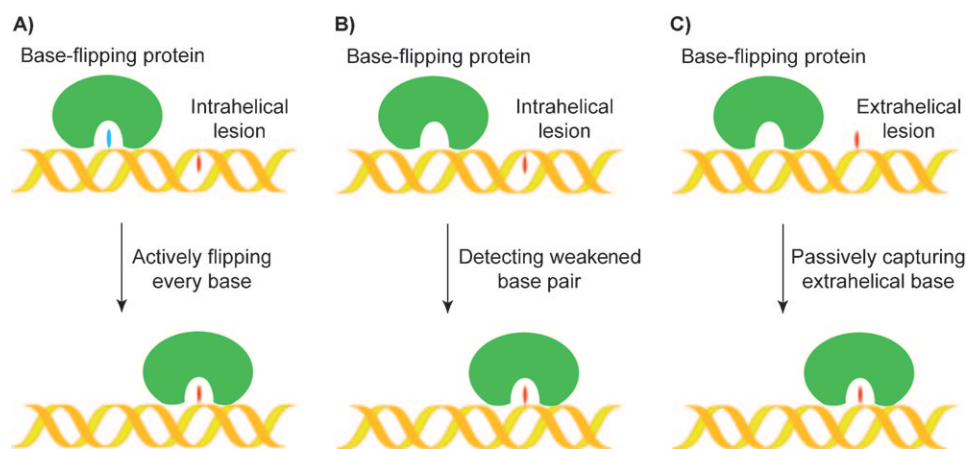


Figure 2. Proposed damage-searching and base-flipping mechanisms for DNA repair proteins. A) In this model, the protein actively flips every base out and checks it in its active-site pocket until the lesion is located. B) Proposed damage-searching model by detecting weakened or non-Watson–Crick base pairs. The repair protein would distort the double-helical structure of DNA and selectively flip out an unstable base. C) DNA repair protein passively captures a transiently extrahelical lesion.

ty of base pairs and assist in flipping the lesioned base in a weak base pairing. Lastly, base-specific repair proteins can passively capture the damaged base that is already “flipped-out” and perform the repair function, as indicated in Figure 2C. In this mechanism, the protein does not need to actively search for a damaged base or sample the conformation because the

base has already spontaneously rotated out of the DNA double helix. It should be noted that the last mode could be assisted by a base repair protein that destabilizes the double-helical structure of DNA and stabilizes the flipped-out base. Because the majority of cytotoxic/mutagenic DNA base damage causes weakened base pairing or non-Watson–Crick geometry in duplex DNA, we tend to favor the mechanism in Figure 2B as a more general pathway for repair proteins to locate the corresponding damage, although other mechanisms could be used by different proteins.

Valuable insight into the potential mechanisms of locating damage has been provided by structural studies of glycosylases that participate in the repair process. Verdine and co-workers have solved a series of structures that capture protein–DNA interactions at different stages of base-flipping on hOGG1 and its bacterial homologue.^[20–21] These proteins repair 8-oxo-G that occurs in low numbers relative to the amount of undamaged bases. They use a fast sliding mechanism to sample millions of base pairs before locating one damaged base.^[19] The active base-flipping mechanism shown in Figure 2A is unlikely to be compatible with the low activation energy that is used by the protein sliding along duplex DNA. Instead it was found that these proteins probe the stability of intrahelical base pairs to locate potentially weakened ones for flipping. The flipped base is further interrogated during the flip inside the active site for the final decision of base excision. The data so far seems to agree with the mechanism shown in Figure 2B.

A key mechanistic question in damaged base recognition is whether the initiating event is exposure of the base from dynamic breathing motions of the base pair, or whether the enzyme actively quickens the base's expulsion by direct interaction with the damaged site. The third hypothesis—passive capture of the base—is supported by a kinetic NMR spectroscopy and X-ray crystallography study, which revealed that UDG uses a passive trapping mechanism to catch thymine and uracil bases that emerge due to the spontaneous breathing motion.^[15] Recognition of a spontaneously flipped base requires a dynamic response from the enzyme. This response allows trapping of the out state during its extremely short extrahelical lifetime. In other words, effective trapping requires dynamic motions of the enzyme that exceed 10^7 per second.^[22] More detailed experimental approaches addressing both the dynamic and diffusion aspects of the initial recognition step are anticipated to support this observation. It will be very interesting to further investigate this mechanism and observe other repair proteins that use similar strategies. It is also interesting that UDG and hOGG1 appear to show distinct differences in their mechanisms to access the corresponding damaged bases. Although, considering that various base lesions possess diverse properties in base-pair stability and geometry, it is actually not surprisingly

to see repair proteins evolve different tricks to best capture particular types of damage from the genome.

Effect of base-pair geometry and stability on base flipping

DNA base-pair geometry and duplex kinetic and thermodynamic stability play very important roles in facilitating base-pair breathing to sample the extrahelical conformation spontaneously.^[23–24] Deviation of such properties from the duplex DNA background provides the most attractive signal for the repair proteins to locate the corresponding lesioned base pairs. It is easy to understand that damaged bases that induce significant perturbations to the DNA structure, such as the thymine dimer, can be readily identified extrahelically. The severe distortion of the DNA duplex structure also lowers the energy needed to undergo base flipping. Other base damage, such as 8-oxo-G, might not induce a structural disturbance, but can still be identified within the helix, perhaps due to the weakened base pair formed by the lesioned base. The energy of the base pairing and the correct geometry of the alignment can be the key detection criterion. Structural disturbance of target bases might cause the differences in the recognition schemes.

The X-ray structural analysis of the duplex DNA that contains *O*⁶-ethylG complexed with minor-groove-binding drugs has provided important information on how carcinogen-modified *O*⁶-alkylG pairs to cytosine (Figure 3).^[25] Depending on the local

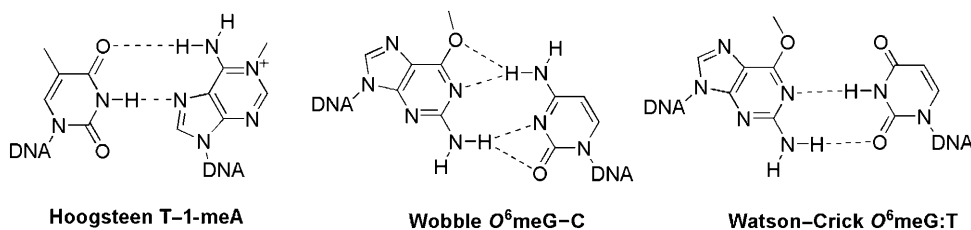


Figure 3. DNA base pairing of Hoogsteen type T–1-meA, wobble-type *O*⁶-meG–C, and the normal Watson–Crick *O*⁶-meG–T.

environment, either wobble pairing or bifurcated hydrogen-bond pairing can occur. The bifurcated configuration is similar to a normal Watson–Crick G:C base pair; however, there might be a dynamic equilibrium between the two configurations of the alkylated G:C base pair that presents an ambiguous signal to the polymerase and subsequently causes it to be edited out. The weakened base-pair stability of the *O*⁶-alkylG:C, combined with the geometrical change from the normal Watson–Crick base pair presents the damaged base for detection by the AGT proteins. In contrast, thymine can pair with *O*⁶-alkylG in only one manner with a configuration similar to a regular Watson–Crick G:C base pair, albeit imperfectly. This is a plausible, though not a definite explanation as to why thymine is found preferentially incorporated across the *O*⁶-alkylG lesion site during replication.

Because the N1-methylation on adenine causes steric clashes in the Watson–Crick base-pairing interface, there is a considerable decrease in the duplex thermal stability. This led to the

proposal that AlkB proteins might passively capture the lesions already in an extrahelical conformation. However, a recent high-resolution NMR spectroscopic study has shown that the N1-methylation of adenine causes a change in the T:A Watson–Crick base pair of DNA double helices to a T:1-meA Hoogsteen base pair (Figure 3).^[26] The methylation does not disrupt base pairing; instead it switches the base-pairing mode. The formation of a Hoogsteen base pair retains the T:1-meA base pairing and stacking within the double helix. This might make the recognition, and thus the repair of 1-meA lesions in dsDNA by AlkB less efficient than in single stranded DNA (ssDNA). Furthermore, the local structural distortion induced by the Hoogsteen base pair could be targeted by the AlkB proteins for damage searching in duplex DNA.

Learning from Direct Reversal of Alkylated DNA

For base-repair proteins, we believe that detecting and capturing an unstable base pair is a simple, efficient, and thus general way to locate potential lesions. In our laboratory we have been studying DNA repair proteins that perform direct reversal of DNA alkylation damage. We have recently proposed that AGT detects DNA lesions by first searching for weakened and/or distorted base pairing. We used an active-site-based disulfide cross-linking strategy with a chemically modified cytosine to study base-flipping by the AGT and AlkB proteins.^[18] The active-site-based disulfide cross-linking method also allowed us to stabilize protein–DNA complexes of the AlkB family proteins. Structural characterization of these complexes provided new insight into damage searching by these proteins.

Disulfide cross-linking of C-Ada and dsDNA

E. coli C-Ada and human AGT use a reactive Cys residue (Cys139 in truncated C-Ada and Cys145 in hAGT) to remove the alkyl groups on either the O⁶-position of guanine or the O⁴-position of thymine. These AGT proteins are quite flexible and can accommodate both purine and pyrimidine substrates. We proposed that this flexibility might allow other modified bases, for example disulfide-tethered cytosine (C*), to access the same substrate-binding pocket and react with the Cys residue to form a disulfide cross-link between the protein and the modified DNA (Figure 4A, B). It should be noted that the disul-

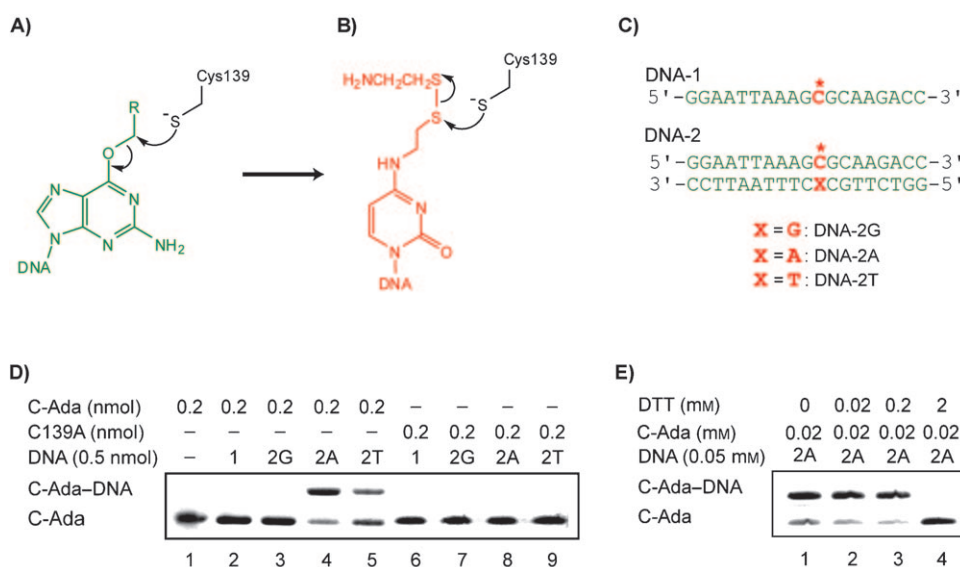


Figure 4. A disulfide cross-linking strategy of C-Ada protein with different dsDNA molecules to investigate damage detection and base flipping. A) C-Ada uses a reactive Cys139 residue to selectively transfer the alkyl modification from O⁶-alkylatedguanine to itself, which represents a suicidal direct repair. B) A thiol-tethered cytosine was introduced in a cross-linking reaction between protein and DNA. C) Various DNA probes, including ssDNA and dsDNA, were used in this study. D) The cross-linking results were shown by nonreduced SDS-PAGE. E) Effect of external thiol (DTT) on the cross-linking reactions.

fide exchange chemistry used for cross-linking is a mechanism-based equilibrium process, and can reach very high yield.^[27] We found by surprise that only when mismatched bases were introduced opposite to the disulfide-tethered C* (C*:A or C*:T base pair) can C* be flipped into the active pocket of C-Ada for the covalent cross-linking reaction to occur (Figure 4).^[18] No cross-linking was observed when a perfectly matched C*:G was used in the same DNA sequence. Therefore, C-Ada appears to locate potential base damage by recognizing a weakened base pair. The protein cannot flip out normal bases that are stabilized by Watson–Crick base pairing in duplex DNA. The exact mechanism still awaits further investigation, but this result indicates that capturing an unstable base pair might be a common damage-searching mechanism for repair proteins.

Disulfide cross-linking of hAGT and dsDNA

Subtle differences were observed between hAGT and C-Ada in their damage-searching modes and offer opportunities to use these two proteins as models to understand the very important mechanism of base-damage detection and flipping. Although it is structurally homologous to the *E. coli* C-Ada protein, human AGT shows some differences. To our knowledge, human AGT can remove alkyl adducts on the O⁶-position of guanine when the alkylated guanine is paired with thymine; such an activity has not been reported for C-Ada.^[28] The O⁶-alkylG:T base pair is fairly stable and can adopt a geometry similar to Watson–Crick base pairs (Figure 3). To detect and remove the alkyl group in this base pair, hAGT might need to use a damage-searching mechanism different from that of C-Ada. The results from our cross-linking experiments show that, like C-Ada, hAGT can efficiently detect damaged bases that form

unstable base pairs; however, this protein also extrudes base lesions that are intrahelically stabilized in duplex DNA, albeit in a less-efficient process.^[18] It is inconceivable that hAGT actively flips every base for damage detection. It might possess additional properties compared with C-Ada to find subtle differences between *O*⁶-alkylG:T and undamaged base pairs. It should be noted that hAGT is very efficient at identifying unstable base pairs, which allows it to quickly find the major target *O*⁶-alkylG:C. The result is in agreement with the previously reported fluorescence and kinetic experiments of repair by human AGT.^[29]

AGT–DNA complex structure analysis

Recently, X-ray structures of human AGT bound to dsDNA were solved by Tainer's and our groups independently.^[9,30] An *O*⁶-methylguanine base and a cross-linkage to a mechanistic inhibitor were used in Tainer's work, and a *p*-xylylenediamine modification of a cytosine was used in our study to stabilize the hAGT–DNA complex. The two structures showed the same base-flipping feature of hAGT to recognize the modified base. Minor-groove binding of dsDNA by AGT was observed. Interestingly, a second hAGT present in our structure binds at the ends of the two neighboring duplex DNA strands and partially inserts an overhang thymine into its active-site pocket (Figure 5). This unique feature suggests that the protein binds

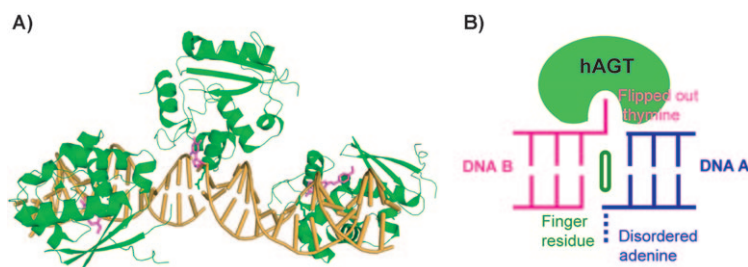


Figure 5. A structure of the human AGT–DNA complex. A) A monomer AGT protein binds at the junction of the two DNA fragments, and recognizes a terminal overhanging thymine. This base is inserted into the active-site pocket only partially rather than forming a base pair with the adenine overhang from the adjacent DNA. B) Cartoon representation of the AGT protein binding the junction of duplex DNA. The finger residue, Arg128, that intercalates inside the duplex and fills the gap left by base flipping, is shown in green.

preferentially to less-rigid regions on DNA. Most likely, the protein exerts a small tension on duplex DNA and finds regions that can best relax this tension due to the presence of damaged, unstable base pairs. It will be very interesting to capture and characterize both hAGT and C-Ada on undamaged duplex DNA. Detailed structural analysis should lead to further insight into the damage-locating mechanism of these proteins.

E. coli AlkB and human ABH2

E. coli AlkB was known to protect against the cytotoxic effects of *S*_N2-methylating agents, however, the explicit function of

AlkB continued to remain unidentified for nearly two decades.^[31] An important lead was provided by a bioinformatic study in 2001. The AlkB protein was predicted to be a mononuclear iron-containing monooxygenase that might perform an oxidative dealkylation function.^[32] The hypothesis was approved experimentally by two independent groups in 2002.^[33–35] The *E. coli* AlkB protein was shown to be a direct dealkylation DNA repair protein by using an unprecedented oxidative dealkylation mechanism.^[33–38] In addition to the main substrates (1-meA and 3-meC), 1-meG, 3-meT, and different etheno-adducts can be repaired by members of the AlkB family by using a similar oxidation mechanism. There are nine potential human homologues of AlkB. Two of these, ABH2 and ABH3, can repair a similar spectrum of base lesions as AlkB, and a third one, FTO, is an important factor involved in obesity.^[39] FTO has been recently shown to repair 3-meT in ssDNA and 3-meU in ssRNA,^[39–40] but how it affects human obesity is still unclear. Interestingly, AlkB and ABH3 display a preference for ssDNA and also quite efficiently demethylates related RNA substrates. In contrast, ABH2 acts as the primary house-keeping enzyme in mammals and repairs endogenously formed 1-meA and ϵ A lesions in duplex DNA.^[41–42]

The structures of AlkB (with dT-(1-medA)-dT)^[43] and ABH3 (in the absence of DNA or RNA)^[44] were solved recently, and provided important insight into the overall architecture of this family of protein. We solved the first structures of AlkB–dsDNA and ABH2–dsDNA complexes stabilized by chemical cross-linking (Figure 1). The mechanisms of base-flipping and damage recognition in dsDNA were revealed from these structures.^[12] The AlkB protein, which lacks a finger residue, uses a unique strategy to recognize dsDNA in an unusual manner. It squeezes together the two bases that flank the flipped out one so that they stack with each other (Figure 6). This distortion is induced by the protein kinking the DNA backbone to invert one of the sugar rings, along with its base, by $\sim 180^\circ$. The distortion of the DNA duplex induced by AlkB eliminates the need for a finger residue to fill the space left by base flipping (Figure 6). Thus, no finger is present in the AlkB–dsDNA complex structure. The passive capture mechanism shown in Figure 2C cannot work for AlkB because the protein has to kink the DNA backbone and open the double helix. The abnormal base stacking and extensive AlkB–DNA backbone interactions suggest that the base flipping is promoted by backbone compression and 180° sugar pucker rotation. Biochemical results support such an enzyme-assisted nucleotide-flipping mode.^[12] The extensive distortion of the dsDNA by AlkB provides an explanation of its preference for ssDNA: these distortions exact an energetic penalty on relatively rigid dsDNA, which should lead to AlkB's preference for more flexible ssDNA. The exact mechanism of damage location by this very unique base-flipping protein needs further studies. In addition, because AlkB has a low affinity to dsDNA, it might be associated with other DNA-processing factors during the damage search. The 1-meA and ϵ A damage that are repaired by AlkB could block replication and

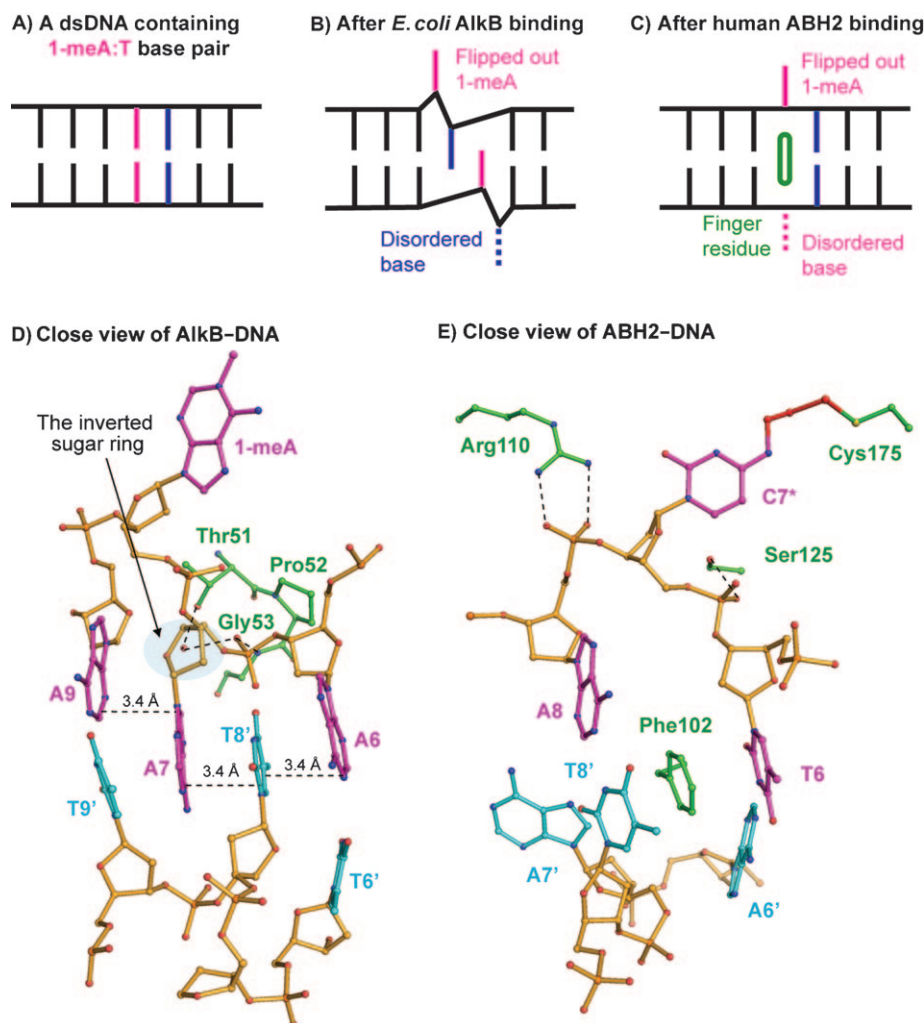


Figure 6. Cartoons for dsDNA conformations in the presence and absence of bound AlkB or ABH2; A) 1-meA can adopt an intrahelical configuration in the double helix by forming a Hoogsteen-type base pair with thymine from the complementary strand. B) *E. coli* AlkB squeezes dsDNA and induces severe distortion of the DNA duplex to facilitate base flipping. C) ABH2 is a standard dsDNA damage repair protein that uses a finger residue (Phe102) to stabilize the duplex structure after base flipping. D) A close view of the structure of how AlkB binds dsDNA. E) A close view of how ABH2 binds dsDNA.

transcription. Perhaps the AlkB protein is recruited when these happen inside cells.

ABH2 works as the main house-keeping enzyme for repairing internally formed 1-meA lesions in mammalian genomes.^[41] The ABH2-DNA structure (Figure 1)^[12] clearly showed ABH2 as a dsDNA repair protein. It interacts extensively with both strands of dsDNA and uses an HTH motif that bears a finger residue to fill the gap left by the flipped-out base. This motif is widespread for base-flipping by base repair proteins. The capacity to interact with the other strand of DNA allows ABH2 to preferentially repair dsDNA lesions, which correlates well with its primary role as a 1-meA damage repair enzyme in mammalian genomes. Based on the active-site cross-linked ABH2-DNA structure we further designed and solved a structure of 1-meA in complex with ABH2. This was aided by a chemical cross-link installed away from the protein active site. The structure provided additional information on lesion recognition by this

human repair protein. ABH2 possesses a more complex residue arrangement than AlkB to specifically recognize a 1-meA lesion in the active site.^[12] The tight lesion recognition might help ABH2 to discriminate 1-meA against other alkylated bases. It is still interesting to see how εA fits in the active-site pocket of ABH2. The active site of AlkB is more flexible, which perhaps helps *E. coli* to recognize and repair a range of different base lesions. The crystal structures of ABH2 complexed with dsDNA are currently unable to offer insight into how the alkylated bases are detected by the protein. Further studies on how ABH2 and normal dsDNA sequences make contact are of fundamental interest and biological importance.

Conclusions

Nucleotide flipping is a common feature in DNA/RNA base repair as well as base-modification proteins. Mechanisms of damage recognition have been established through numerous structural characterizations of the base-specific DNA repair proteins. However, how these repair proteins locate the very few damaged bases among the vast amount of undamaged bases in a genome presents an intriguing

biochemical question. Through recent structural and biochemical studies, several themes start to emerge on the damage-locating mechanisms of these proteins. Many of these proteins seem to detect unstable base pairs in the first step of damage searching. Either by using a probing residue or through small tensions generated on the duplex DNA structure these proteins probe and preferentially recognize "soft" spots on DNA. Then, the located base can be flipped and further checked by the protein. Alternatively, the repair proteins could capture and stabilize a flipped-out base by a more passive mechanism. This mechanism proposed for UDG starts by thermally inducing a base pair opening as opposed to the more active participation of the protein. The protein does stabilize the flipped-out base to facilitate capture of the damaged base. Even within the same repair protein family, there can be subtle differences in mechanisms used for locating damage as we showed with the AGT proteins. Detailed characterization of in-

teractions of these proteins with damaged and undamaged DNA will offer more insight into this intriguing question.

Notably, AlkB recognizes alkylated base damage by using a unique strategy that is distinct from other known base repair proteins. This enables *E. coli* AlkB to recognize and repair alkylated DNA base damage in both flexible ssDNA and relatively rigid dsDNA. How this protein actively searches for base lesions in dsDNA is still unknown. More intriguingly, how this enzyme selectively locates the alkylated DNA bases in ssDNA remains an open question. Are cofactor proteins involved or not in the searching process? Future studies, of both the crystal structures of AlkB and its homologues bound to long ssDNA and additional NMR spectroscopic dynamic experiments, are required to answer these questions. These further efforts might enhance our understanding of how these enzymes initially recognize damaged bases in both ssDNA and dsDNA sequences during DNA repair.

Acknowledgements

The National Institutes of Health (GM071440), the W. M. Keck Foundation Distinguished Young Scholar in Medical Research Program, and the Arnold and Mabel Beckman Foundation Young Investigator Program are acknowledged for their financial support.

Keywords: AlkB • DNA damage • DNA repair • nucleic acids • protein–DNA interactions

- [1] T. Lindahl, *Nature* **1993**, 362, 709–715.
- [2] R. J. Roberts, *Cell* **1995**, 82, 9–12.
- [3] R. J. Roberts, X. Cheng, *Annu. Rev. Biochem.* **1998**, 67, 181–198.
- [4] S. Kumar, X. Cheng, S. Klimasauskas, S. Mi, J. Posfai, R. J. Roberts, G. G. Wilson, *Nucleic Acids Res.* **1994**, 22, 1–10.
- [5] X. Cheng, R. J. Roberts, *Nucleic Acids Res.* **2001**, 29, 3784–3795.
- [6] X. Cheng, R. M. Blumenthal, *Structure* **1996**, 4, 639–645.
- [7] T. Lindahl, B. Sedgwick, M. Sekiguchi, Y. Nakabeppu, *Annu. Rev. Biochem.* **1988**, 57, 133–157.
- [8] A. Y. Lau, M. D. Wyatt, B. J. Glassner, L. D. Samson, T. Ellenberger, *Proc. Natl. Acad. Sci. USA* **2000**, 97, 13573–13578.
- [9] D. S. Daniels, T. T. Woo, K. X. Lu, D. M. Noll, N. D. Clarke, A. E. Pegg, J. A. Tainer, *Nat. Struct. Mol. Biol.* **2004**, 11, 714–720.
- [10] S. D. Bruner, D. P. G. Norman, G. L. Verdine, *Nature* **2000**, 403, 859–866.
- [11] G. Slupphaug, C. D. Mol, B. Kavli, A. S. Arvai, H. E. Krokan, J. A. Tainer, *Nature* **1996**, 384, 87–92.
- [12] C.-G. Yang, C. Yi, E. M. Duguid, C. T. Sullivan, X. Jian, P. A. Rice, C. He, *Nature* **2008**, 452, 961–965.
- [13] A. Y. Lau, O. D. Scharer, L. Samson, G. L. Verdine, T. Ellenberger, *Cell* **1998**, 95, 249–258.
- [14] A. C. Vallur, J. A. Feller, C. W. Abner, R. K. Tran, L. B. Bloom, *J. Biol. Chem.* **2002**, 277, 31673–31678.
- [15] J. B. Parker, M. A. Bianchet, D. J. Krosky, J. I. Friedman, L. M. Amzel, J. T. Stivers, *Nature* **2007**, 449, 433–438.
- [16] U. D. Priyakumar, A. D. MacKerell Jr., *Chem. Rev.* **2006**, 106, 489–505.
- [17] G. L. Verdine, S. D. Bruner, *Chem. Biol.* **1997**, 4, 329–334.
- [18] E. M. Duguid, Y. Mishina, C. He, *Chem. Biol.* **2003**, 10, 827–835.
- [19] P. C. Blainey, A. M. van Oijen, A. Banerjee, G. L. Verdine, X. S. Xie, *Proc. Natl. Acad. Sci. USA* **2006**, 103, 5752–5757.
- [20] A. Banerjee, W. Yang, M. Karplus, G. L. Verdine, *Nature* **2005**, 434, 612–618.
- [21] A. Banerjee, W. L. Santos, G. L. Verdine, *Science* **2006**, 311, 1153–1157.
- [22] J. T. Stivers, *Chem. Eur. J.* **2008**, 14, 786–793.
- [23] H. T. Allawi, J. SantaLucia Jr., *Biochemistry* **1998**, 37, 9435–9444.
- [24] H. T. Allawi, J. SantaLucia Jr., *Nucleic Acids Res.* **1998**, 26, 2694–2701.
- [25] M. Sriram, G. A. van der Marel, H. L. P. F. Roelen, J. H. van Boom, A. H.-J. Wang, *EMBO J.* **1992**, 11, 225–232.
- [26] H. Yang, Y. Zhan, D. Fenn, L. M. Chi, S. L. Lam, *FEBS Lett.* **2008**, 582, 1629–1633.
- [27] C. He, G. L. Verdine, *Chem. Biol.* **2002**, 9, 1297–1303.
- [28] J. Lips, B. Kaina, *Mutat. Res.* **2001**, 487, 59–66.
- [29] H. Zang, Q. Fang, A. E. Pegg, F. P. Guengerich, *J. Biol. Chem.* **2005**, 280, 30873–30881.
- [30] E. M. Duguid, P. A. Rice, C. He, *J. Mol. Biol.* **2005**, 350, 657–666.
- [31] H. Kataoka, Y. Yamamoto, M. Sekiguchi, *J. Bacteriol.* **1983**, 153, 1301–1307.
- [32] L. Aravind, E. V. Koonin, *Genome Biol.* **2001**, 2, research0007.1–0007.8.
- [33] S. Treweek, T. F. Henshaw, R. P. Hausinger, T. Lindahl, B. Sedgwick, *Nature* **2002**, 419, 174–178.
- [34] P. O. Falnes, R. F. Johansen, E. Seeberg, *Nature* **2002**, 419, 178–182.
- [35] T. Duncan, S. C. Treweek, P. Koivisto, P. A. Bates, T. Lindahl, B. Sedgwick, *Proc. Natl. Acad. Sci. USA* **2002**, 99, 16660–16665.
- [36] B. Sedgwick, *Nat. Rev. Mol. Cell Biol.* **2004**, 5, 148–157.
- [37] Y. Mishina, C. He, *J. Inorg. Biochem.* **2006**, 100, 670–678.
- [38] E. Feyzi, O. Sundheim, M. P. Westbye, P. A. Aas, C. B. Vagbo, M. Otterlei, G. Slupphaug, H. E. Krokan, *Curr. Pharm. Biotechnol.* **2007**, 8, 326–331.
- [39] T. Gerken, C. A. Girard, Y. C. Tung, C. J. Webby, V. Saudek, K. S. Hewitson, G. S. Yeo, M. A. McDonough, S. Cunliffe, L. A. McNeill, J. Galvanovskis, P. Rorsman, P. Robins, X. Prieur, A. P. Coll, M. Ma, Z. Jovanovic, I. S. Farooqi, B. Sedgwick, I. Barroso, T. Lindahl, C. P. Ponting, F. M. Ashcroft, S. O'Rahilly, C. J. Schofield, *Science* **2007**, 318, 1469–1472.
- [40] G. Jia, C.-G. Yang, S. Yang, X. Jian, C. Yi, Z. Zhou, C. He, *FEBS Lett.* **2008**, 582, 3313–3319.
- [41] J. Ringvoll, L. M. Nordstrand, C. B. Vagbo, V. Talstad, K. Reite, P. A. Aas, K. H. Lauritzen, N. B. Liabakk, A. Bjork, R. W. Doughty, P. O. Falnes, H. F. Krokan, A. Klungland, *EMBO J.* **2006**, 25, 2189–2198.
- [42] J. Ringvoll, N. M. Moen, L. M. Nordstrand, L. B. Meira, B. Pang, A. Bekkelund, P. C. Dedon, S. Bjelland, L. D. Samson, P. O. Falnes, A. Klungland, *Cancer Res.* **2008**, 68, 4142–4149.
- [43] B. Yu, W. C. Edstrom, J. Benach, Y. Hamuro, P. C. Weber, B. R. Gibney, J. F. Hunt, *Nature* **2006**, 439, 879–884.
- [44] O. Sundheim, C. B. Vagbo, M. Bjoras, M. M. Sousa, V. Talstad, P. A. Aas, F. Drablos, H. E. Krokan, J. A. Tainer, G. Slupphaug, *EMBO J.* **2006**, 25, 3389–3397.

Received: August 27, 2008

Published online on January 14, 2009