J|A|C|S
COMMUNICATIONS

Published on Web 06/28/2003

Probing the Structure and Function of the *Escherichia coli* DNA Alkylation Repair AlkB Protein through Chemical Cross-Linking

Yukiko Mishina and Chuan He*

Department of Chemistry, The University of Chicago, 5735 South Ellis Avenue, Chicago, Illinois 60637

Received February 12, 2003; E-mail: chuanhe@uchicago.edu

The *Escherichia coli* AlkB protein has been known to play an important role in alkylated DNA damage repair;[1,2] however, the exact function of AlkB was only recently discovered. Using sequence profile analysis, AlkB was predicted to be a member of the 2-ketoglutarate- and iron-dependent dioxygenase superfamily that contains a metal ion binding site.[3] Subsequent studies revealed that the repair of the cytotoxic 1-methyladenine and 3-methylcytosine (Figure 1A) in single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA) by *E. coli* AlkB was dependent on the addition of dioxygen, 2-ketoglutarate, and iron(II). The DNA repair function of AlkB represents a new type of activity that directly removes DNA-alkylation damage through a metal-mediated mechanism. A novel oxidative demethylation mechanism was proposed as is shown in Figure 1B.[4,5] Human homologues of AlkB were found to exhibit similar repair reactivities.[6]

Despite breakthroughs in identifying possible cellular targets for AlkB, the exact search and repair mechanism of AlkB is still unknown. Moreover, the putative active site composed of His131, Asp133, and His187 predicted by the sequence alignment has yet to be directly confirmed. Previous results provided contradictory evidence about whether AlkB preferentially repairs DNA lesions in ssDNA or dsDNA.[4,5] To understand AlkB−DNA interactions and the damage repair mechanism, a structure of an AlkB/DNA complex is desirable. However, homogeneous AlkB/DNA complexes are almost impossible to obtain for structural studies using traditional methods for several reasons. First, AlkB binds DNA nonsequence-specifically and shows a weak affinity to undamaged ssDNA and dsDNA (unpublished result, Mishina and He).[2] Second, the affinity of AlkB to methylated ssDNA is only 2-fold higher than the binding to unmethylated ssDNA.[2] Third, because both 1-methyladenine and 3-methylcytosine lesions are unstable, it is almost impossible to stably incorporate these modified bases into a synthetic oligonucleotide through solid-state DNA synthesis. These modified bases, if generated, would decompose gradually in an aqueous environment.[7] Here, we report the use of a chemical disulfide cross-linking technique[8,9] to probe the structure and function of *E. coli* AlkB. This method, an equilibrium process that relies on protein/DNA interactions, enables us to trap and isolate homogeneous AlkB/DNA complexes for structural characterization.

As a DNA repair protein, AlkB can potentially use its substrate binding site to check every base in ssDNA or dsDNA for DNA lesions (Figure 2A). We synthesized an oligonucleotide having a modified cytosine base with a thiol tether introduced at the N^4-position. We envision that the modified cytosine could access the substrate-binding site if there is enough space to accommodate it. This space can be created by removing the bound metal ion and the cosubstrate, 2-ketoglutarate (Figure 2B). Mutation of one of the three putative metal-binding residues, H131, D133, and H187, to a Cys residue would allow a disulfide cross-link between the engineered Cys residue and thiol-tethered DNA (H131C is shown in Figure 2B). Thus, we prepared and purified H131C, D133C,

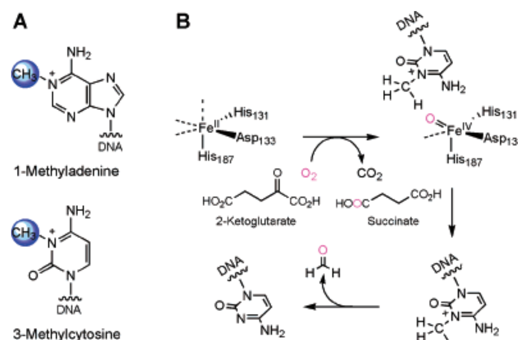


Figure 1. (A) 1-Methyladenine and 3-methylcytosine lesions in DNA. (B) Proposed oxidative demethylation of 1-methyladenine and 3-methylcytosine lesions in DNA by AlkB in a putative iron(II)-containing active site.

and H187C mutant AlkB proteins and treated them with EDTA to remove metal ions and 2-ketoglutarate. A random single-stranded DNA (ssDNA-1) with a thiol-tethered cytosine (C*) was synthesized and purified (Figure 2C). A two-carbon thiol tether, protected with 2-aminoethanethiol mixed disulfide, was introduced into C* by a convertible nucleoside methodology.[10,11] Annealing ssDNA-1 with its complementary strand afforded dsDNA-2. dsDNA-3 and dsDNA-4 in which mismatched bases A and T were introduced opposite C* in the complementary strand, respectively, were also prepared (Figure 2C).

To assay for cross-link formation between the modified DNA and mutant proteins, we incubated 3 equiv of DNA (DNA-1−DNA-4) with 1 equiv of mutant AlkB at 4 °C for 24 h. The reaction was then quenched and examined using SDS-polyacrylamide gel electrophoresis as was described previously.[8] The formation of a cross-linked complex between ssDNA-1 and H131C AlkB was observed by the appearance of a new band with retarded mobility (Figure 2D, lane 2). Approximately 28% of the H131C AlkB cross-linked to ssDNA-1, as was estimated from the intensities of the two bands after 24 h. Very small amounts of cross-linked complexes, estimated to be less than 1%, were observed between dsDNA-2 and H131C AlkB (Table S1, lane 3), whereas both dsDNA-3 and dsDNA-4, with mismatched bases opposite C*, cross-linked with H131C AlkB (lanes 4, 5). Less than 1% of cross-linked products was observed in control cross-linking reactions between all DNA and wild-type AlkB under the same conditions (lanes 6−9). Thus, the cross-linking is selective for the engineered Cys in H131C AlkB. The D133C and H187C mutant proteins, containing mutations to the two other putative metal-binding ligands, also effectively cross-linked with ssDNA-1, dsDNA-3, and dsDNA-4 under the same conditions (Figure 2F). These AlkB mutants only formed trace amounts of complexes with dsDNA-2 (Table S1 summarizes estimated cross-linking yields for all reactions).

If the cross-linking site is indeed a metal-binding site as predicted, the binding of a metal ion would inhibit the cross-linking reaction. The influence of metal ions on the cross-linking reaction was studied

10.1021/ja034636c CCC: $25.00 © 2003 American Chemical Society

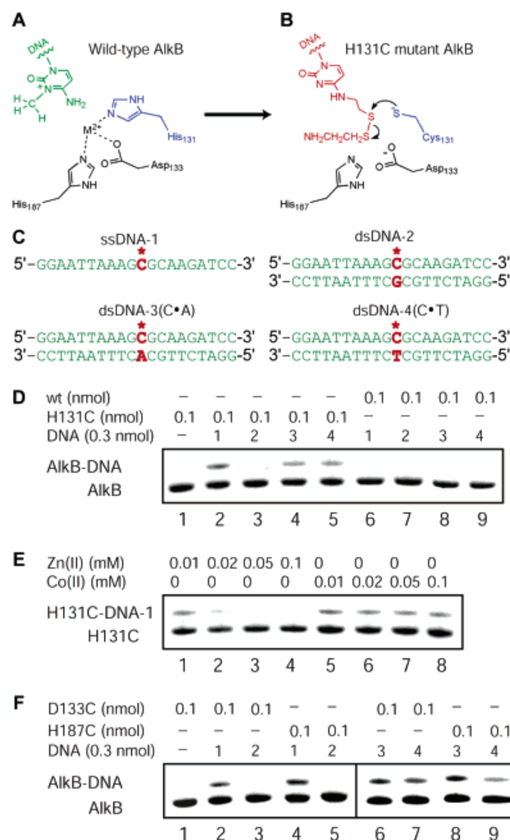
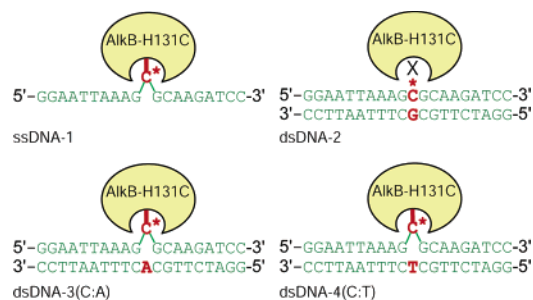


Figure 2. Disulfide cross-linking. (A) Proposed substrate binding in the putative AlkB active site. (B) Replacement of one of the three putative metal-binding ligand His131 to Cys131 and introduction of a thiol-tethered C* (star denotes the modified cytosine site) in DNA provide the partners for the formation of a disulfide cross-link. (C) Oligonucleotides used in cross-linking. (D) SDS gel analysis of the cross-linking reactions between H131C AlkB and DNA-(1−4). Lane 1 is a size standard for AlkB. Disulfide cross-linking between the protein and DNA results in the appearance of a new band having retarded mobility. Lanes 2−5: cross-linking results between H131C AlkB and DNA-(1−4) after 24 h of incubation at 4 °C. Lanes 5−9: controls with wild-type AlkB. (E) SDS gel analysis of the cross-linking reaction between H131C (10 μM) and DNA-1 (30 μM) in the presence of $ZnCl_2$ and $CoCl_2$ at 4 °C. The protein was preincubated with metal for 1 h. (F) SDS gel analysis of the cross-linking reactions between D133C AlkB and DNA-(1−4) and H187C AlkB and DNA-(1−4). The reactions were analyzed after 24 h of incubation at 4 °C.

using H131C AlkB. The presence of $ZnCl_2$ (2−5 equiv) inhibited the formation of cross-linked complexes between the mutant protein and DNA-1 (Figure 2E and Figure S1B). $CoCl_2$ had similar but less dramatic effects on the cross-linking reaction (Figures 2E and S1B). This result is expected because zinc(II) binds Cys residue more tightly than does cobalt(II). To confirm the existence of a disulfide bond between mutant AlkB and DNA, we treated H131C AlkB with a strong reducing reagent, dithiothreitol (DTT). The cross-linking disappeared in the presence of 100 equiv of DTT but was stable with 10 equiv of DTT (Figure S1A). Incubating isolated H131C/DNA-1 complex with DTT resulted in the cleavage of the disulfide bond, and 100 equiv of DTT can reduce the cross-linked disulfide bond (Figure S1A).

The mutant AlkB cross-links with ssDNA-1 with relatively high efficiency, suggesting that AlkB checks every base in ssDNA for potential lesions by inserting each base into a substrate binding pocket (Scheme 1). The mutant AlkB proteins formed less than 1% of cross-linked product with dsDNA-2, where C* is held in

Scheme 1. Schematics of H131C AlkB/DNA Cross-Linking



the helical structure by hydrogen bonding to its complementary base G. It appears from this result that AlkB cannot break base pairs and flip out every base for damage checking in dsDNA. Instead, it may require existing perturbations in the duplex DNA structure. In the case of 1-methyladenine and 3-methylcytosine, their inability to base pair with their complementary bases causes them to rotate out of their position in duplex DNA, allowing AlkB to capture them for repair. To test this hypothesis, we introduced mismatched bases opposite C* in dsDNA-3 and dsDNA-4. This allows the thiol-tethered C* to readily rotate out of the duplex DNA structure and be captured by mutant AlkB for cross-linking (Scheme 1). Our results clearly supported this damage searching mechanism as is shown by the effective cross-linking observed between all mutant proteins and dsDNA-3 and dsDNA-4 (Figure 2).

In summary, our results suggest that the pocket formed by His131, Asp133, and His187 is the substrate binding active site. Efficient chemical cross-linking was observed between modified ssDNA-1 and the engineered Cys residue at all of these positions. The effect of metal ions, especially zinc(II), on the cross-linking reaction indicates that the cross-linking site is a metal-binding site. Our results also suggest that AlkB checks every base in ssDNA for potential damage. However, it does not flip out every base from dsDNA when searching for damage; instead, it captures damaged bases that can easily rotate out of the duplex structure because they do not stably base pair with bases in the opposite strand. The cross-linking reaction can be run on a large scale, allowing homogeneous AlkB/DNA complexes, which are otherwise difficult to prepare, to be isolated for further structural characterization.

Acknowledgment. This research was supported by The University of Chicago.

Supporting Information Available: Experimental details. Figure S1 and Table S1 (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Dinglay, S.; Gold, B.; Sedgwick, B. *Mutat. Res.* **1998**, *407*, 109−116.
(2) Dinglay, S.; Trewick, S. C.; Lindahl, T.; Sedgwick, B. *Genes Dev.* **2000**, *14*, 2097−2105.
(3) Aravind, L.; Koonin, E. V. *Genome Biol.* **2001**, *2*, 007.1−007.8.
(4) Trewick, S. C.; Henshaw, T. F.; Hausinger, R. P.; Lindahl, T.; Sedgwick, B. *Nature* **2002**, *419*, 174−178.
(5) Falnes, P. O.; Johansen, R. F.; Seeberg, E. *Nature* **2002**, *419*, 178−182.
(6) Duncan, T.; Trewick, S. C.; Koivisto, P.; Bates, P. A.; Lindahl, T.; Sedgwick, B. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16660−16665.
(7) Chang, C.-j.; Gomes, J. D.; Byrn, S. R. *J. Org. Chem.* **1983**, *48*, 5151−5160.
(8) He, C.; Verdine, G. L. *Chem. Biol.* **2002**, *9*, 1297−1303.
(9) Huang, H.; Harrison, S. C.; Verdine, G. L. *Chem. Biol.* **2000**, *7*, 355−364.
(10) Xu, Y. Z.; Zheng, Q.; Swann, P. F. *J. Org. Chem.* **1992**, *57*, 3839−3845.
(11) MacMillan, A. M.; Verdine, G. L. *J. Org. Chem.* **1990**, *55*, 5931−5933.

JA034636C