Published in final edited form as:
Nature. 2008 April 24; 452(7190): 961–965. doi:10.1038/nature06889.

Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA

Cai-Guang Yang[1,*], **Chengqi Yi**[1,*], **Erica M. Duguid**[1], **Christopher T. Sullivan**[1], **Xing Jian**[1], **Phoebe A. Rice**[2], and **Chuan He**[1]

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

2*Department of Biochemistry and Molecular Biology, The University of Chicago, 929 East 57th Street, Chicago, Illinois 60637, USA*

Abstract

Escherichia coli AlkB and its human homologues ABH2 and ABH3 repair DNA/RNA base lesions by using a direct oxidative dealkylation mechanism. ABH2 has the primary role of guarding mammalian genomes against 1-meA damage by repairing this lesion in double-stranded DNA (dsDNA), whereas AlkB and ABH3 preferentially repair single-stranded DNA (ssDNA) lesions and can repair damaged bases in RNA. Here we show the first crystal structures of AlkB–dsDNA and ABH2–dsDNA complexes, stabilized by a chemical cross-linking strategy. This study reveals that AlkB uses an unprecedented base-flipping mechanism to access the damaged base: it squeezes together the two bases flanking the flipped-out one to maintain the base stack, explaining the preference of AlkB for repairing ssDNA lesions over dsDNA ones. In addition, the first crystal structure of ABH2, presented here, provides a structural basis for designing inhibitors of this human DNA repair protein.

Cellular DNA is constantly subjected to modifications by environmental and endogenous chemicals, which can result in covalent changes[1,2]. Methylating (or alkylating) agents are a common group of DNA modifiers that introduce damage primarily to the heterocyclic bases of DNA, with mutagenic and/or cytotoxic consequences. Alkylating agents are also widely used in cancer therapy and exert anticancer effects by creating cytotoxic DNA lesions in tumour cells. Many of these alkylation DNA damages are detected and repaired by proteins that are conserved across kingdoms.

The *E. coli* AlkB protein is a direct dealkylation DNA repair protein[3–5]. It uses a mononuclear iron(II) site and cofactors 2-ketoglutarate (2KG) and dioxygen to perform an unprecedented oxidative demethylation of DNA base lesions 1-meA, 3-meC, 1-meG and 3-meT (Supplementary Fig. 2)[6–11]. AlkB also removes etheno DNA lesions by using a similar oxidation mechanism[12,13]. There are nine potential human homologues of AlkB. Two of

Correspondence and requests for materials should be addressed to C.H. (chuanhe@uchicago.edu).
*These authors contributed equally to this work.

Author Contributions C.-G.Y. and C.Y. solved all AlkB–dsDNA and ABH2–dsDNA structures with help from E.M.D. (crystallography), C.T.S. (initial construct of ABH2 and crystallography) and X.J. (biochemistry). P.A.R. contributed to protein crystallography. C.H. designed the overall project and wrote the manuscript with C.-G.Y. and C.Y. All authors discussed results and commented on the manuscript.

Author Information Atomic coordinates are deposited in Protein Data Bank under accession numbers 3BKZ ((Mn/2KG) AlkB–DNA1), 3BI3 ((Mn/2KG) AlkB–DNA5), 3BIE ((Mn/2KG) AlkB–DNA4), 3BTX (ABH2–DNA2), 3BTZ (ABH2–DNA3), 3BU0 ((Mn/2KG) ABH2–DNA2), 3BTY (ABH2–DNA6) and 3BUC ((Mn/2KG)ABH2–DNA6).

these, ABH2 and ABH3, can repair the same spectrum of base lesions as AlkB[14–16]. A third one, FTO, has been recently shown to repair 3-meT in DNA, and it affects obesity in mammals through an unknown mechanism[17]. ABH2 acts as the primary housekeeping enzyme in mammals for repairing endogenously formed 1-meA lesions in duplex DNA[18]. The *in vivo* function of ABH3 is still unclear but it has been shown to repair base lesions in RNA as well as DNA[15]. Although the structures of AlkB (with dT-(1-medA)-dT)[19] and ABH3 (in the absence of DNA or RNA)[20] were solved recently, the molecular mechanism underlying the different substrate preferences of ABH2 and AlkB remains unknown. Furthermore, the structure of ABH2 has, to our knowledge, never been reported.

Cross-linking to stabilize protein–DNA complexes

We report here the first crystal structures of AlkB–dsDNA and ABH2–dsDNA complexes. The AlkB family proteins bind DNA weakly[21] and form labile complexes with damage-containing DNA[22], which makes crystallization of their protein–DNA complexes challenging. To overcome this difficulty we used chemical cross-linking methods[23,24]; initially using an active site disulphide cross-linking strategy that we developed previously (Fig. 1a)[25,26]. Base-repair proteins flip damaged bases and insert them into the active site for processing. Therefore, we reasoned, a cysteine residue engineered into the active site of AlkB may form a disulphide cross-link, at equilibrium, with a disulphide-modified cytosine (C* in a C*:A base pair) flipped into the active site of the repair protein (Fig. 1a)[27]. Of several mutations tested, AlkB D135C and ABH2 E175C gave the best cross-linking yields with disulphide-modified DNA under equilibrium conditions (Supplementary Fig. 3).

Structure determination

For crystallization, truncated but catalytically active AlkB-ΔN11[19] and ABH2-ΔN55 were used (Supplementary Fig. 4). Of multiple DNA sequences screened, a 13-base-pair duplex DNA1, with a central C*:A base pair (Fig. 1a), gave the best crystals of the active site cross-linked AlkB–DNA complex. These crystals diffracted to 1.65 Å and the structure was solved by molecular replacement using the previously reported AlkB structure as the search model[19]. Two crystal forms of active site cross-linked ABH2–DNA complexes were grown using different 13-base-pair dsDNAs containing central C*:A pairs. The first, with DNA2, diffracted to 2.0 Å, and was solved by multi-wavelength anomalous dispersion (MAD) phasing using bromine-labelled (5-Br dU) DNA. The second form, with DNA3, diffracted to 3.0 Å and was solved by molecular replacement using the first as a search model. Both the AlkB and the ABH2 truncations used in this study still efficiently repair 1-meA in dsDNA with the sequences shown in the crystal structures (Supplementary Figs 5 and 6).

General features of the protein–DNA interactions

AlkB and ABH2 have similar folds (a common jelly-roll fold that is shared by all iron(II)/2-KG-dependent oxygenases[28–30]), and both flip the damaged base into a deep active site pocket. However, their interactions with dsDNA are very different (Fig. 1b–e). First, AlkB interacts almost exclusively with the damaged strand, whereas ABH2 makes significant contacts with both DNA strands (Supplementary Fig. 7). Second, the two enzymes use very different mechanisms to flip the damaged base. ABH2 uses an aromatic finger residue, Phe 102, which resembles the intercalating hairpin motif observed with DNA glycosylases[31,32], to intercalate into the duplex stack and fill the DNA gap resulting from the base flipping. AlkB lacks this finger residue; instead it manipulates the DNA backbone near the flipped base such that the bases flanking it stack with one another. Two slightly different DNA conformations were observed in the AlkB–DNA1 structure, but in both cases a similar distortion of the lesion-containing strand was observed (Fig. 1b, d, and Supplementary Fig. 8).

Cross-linking AlkB to dsDNA away from the active site

The protein portion of the AlkB–DNA1 structure overlaps almost perfectly with the previously published AlkB structure (Supplementary Fig. 9a). However, to confirm that the DNA distortions observed are not due to cross-linking to the active site of the protein, we also engineered a disulphide cross-link away from the active site, to AlkB S129C (Fig. 2a)[32,33]. Two such AlkB–dsDNA complexes were crystallized: one with DNA4, a 13-base-pair dsDNA containing C10*, 1-meA and an abasic site (Fig. 2b); and one with DNA5, containing C10* two base pairs away from the 1-meA (Fig. 2c). These were solved at 1.68 Å and 1.90 Å, respectively, by molecular replacement.

Unique binding of duplex DNA by AlkB

All three structures of the AlkB–dsDNA complexes show a base-flipping feature distinct from all known structures of base-flipping proteins bound to DNA. The two bases that flank the flipped 1-meA, A7 and A9, are 'squeezed' such that they stack on one another, about 3.4 Å apart (Fig. 3a and Supplementary Fig. 10c). The sugar rings of these two nucleosides adopt unusual conformations. One (A9) is C3′-endo, standard in A-form DNA (Supplementary Fig. 10a). The other (A7) shows the most surprising feature: with its base, it is twisted by about 180°, resulting in an inverted sugar (O4′ points towards the 3′ end of the DNA strand) while the base stacks with A9:T9′ (Fig. 3a). The inversion of this sugar ring is locally enforced by residues 51–53 (Thr-Pro-Gly), which form a rigid hydrogen bonding network with the phosphate group between A7 and A6 (Fig. 3a and Supplementary Fig. 10b), and through multiple protein–DNA interactions anchoring the two phosphate groups flanking the 1-meA (Supplementary Fig. 7a). These features were not observed in the AlkB structure containing a short T-(1-meA)-T[19], most likely because the ssDNA trimer is too short to reach much of the DNA-binding surface of AlkB. As a result, the positions of the two bases that flank the 1-meA are different in the previous AlkB structure and the current AlkB–dsDNA structures (Supplementary Fig. 9c). It remains to be seen how a long ssDNA interacts with AlkB.

In all three AlkB–dsDNA structures, the protein stabilizes only the distorted conformation of the DNA immediately flanking the 1-meA; the complementary strand appears to accommodate this through spontaneous conformational rearrangements. The base opposite the 1-meA (T8′ in DNA4 or A8′ in DNA1 and DNA5) intercalates between A7 and A6 to generate a one-base-wide stack containing bases from both strands, which maintains the length of the normal duplex DNA (Fig. 3a and Supplementary Fig. 10c). The base at the 7′ position is forced out of the helix and is disordered in our structures (in DNA4 it is replaced with an abasic site) (Figs 1b and 2b, c). To support the observation that the complementary strand distorts to accommodate base flipping in AlkB–dsDNA complexes, we measured the context dependence of AlkB-mediated 1-meA repair (Supplementary Fig. 11). AlkB can catalytically repair this lesion in the dsDNA sequence used in our crystals (replacing C8*:A8′ with 1-meA:T in DNA1), which has two A:T base pairs flanking 1-meA:T. When the two flanking base pairs (7:7′ and 9:9′) were replaced with G:C, repair activity dropped approximately threefold. The activity of this substrate was restored to the original level by incorporating an abasic site at position 7′, which destabilizes the base pair containing the base that is forced out of the double helix in our structures. The highest catalytic repair activity was observed with the ssDNA strand containing 1-meA, as expected.

ABH2 binding to duplex DNA

The structure of ABH2 differs from AlkB in having extra DNA-binding motifs that grasp the complementary strand of the duplex DNA (Fig. 1). ABH2 uses a short loop with a positively charged RKK sequence (Arg 241–Lys 243) and an additional long, flexible loop carrying DNA-binding residues Arg 198, Gly 204 and Lys 205 to bind the opposite DNA strand (Fig. 1c, e

and Supplementary Fig. 7b). The two bases that flank the flipped out C7*, A8 and T6, are approximately the same distance from each other as they would be in canonical duplex DNA (Fig. 3b). The orphaned base A7′ is flexible and poorly ordered. Because the normal duplex DNA length is maintained by the intercalating finger Phe 102, no further distortion of the DNA duplex is induced by ABH2.

To capture the structure of ABH2 with an intact active site, we engineered a disulphide cross-link away from the active site, to ABH2 G169C[32,33]. The cross-linked complex with a new 13-base-pair duplex probe containing C9* (DNA6; Fig. 4a and Supplementary Fig. 12) was crystallized and the structure was solved at 2.35 Å by molecular replacement. A cofactor-containing structure was also determined (at 2.6 Å) by soaking these crystals in 2KG and Mn^{2+} (which occupies Fe^{2+}-binding site but does not support catalysis). The cofactors are bound by ABH2 in a standard conformation resembling that seen in AlkB and other iron(II)/2KG-based oxygenases[19,28]. The metal ion is ligated by His 171, Asp 173 and His 236, a bidentate 2KG and a water molecule in an octahedral geometry (Fig. 4b, and Supplementary Figs 13 and 14).

Recognition of 1-meA by ABH2

The ABH2–dsDNA6 structure provides an undistorted view of the active site (Fig. 4b). A total of five active site residues are engaged in the recognition of 1-meA. Phe 124 and His 171 stack against the 1-meA; an analogous stacking interaction is also observed for 1-meA recognition by Trp 69 and His 131 in AlkB (Supplementary Fig. 9c). Three residues, Tyr 122, Glu 175 and Asp 174, together with a water molecule, form an extensive hydrogen-bonding network that contacts the N6 and N7 nitrogen atoms of 1-meA in ABH2. This tight recognition of 1-meA is not observed in the AlkB structure, as only two residues, Asp 135 and Glu 136 (mediated through a water molecule), were involved in the recognition of the N6 nitrogen of 1-meA by AlkB[19]. The positively charged 1-meA may be favourably recognized through electrostatic interaction with the nearby carboxylate of Glu 175 (Asp 135 of AlkB) and stacking with the π system of Phe 124 (Trp 69 of AlkB). These features may help ABH2 and AlkB discriminate against the neutral base A.

Discussion

The AlkB family oxygenases represent a novel class of DNA base-repair enzymes that use an oxidative dealkylation mechanism. Certain members of this family, AlkB and ABH3, exhibit a distinct preference to repair ssDNA and RNA. We present the first structures of AlkB–dsDNA and ABH2–dsDNA complexes stabilized by chemical cross-linking, showing how AlkB and ABH2 bind and repair damage in dsDNA. The AlkB protein, lacking a finger residue, facilitates base flipping by squeezing together the two bases flanking the flipped-out one such that they stack on one another. This AlkB-mediated DNA distortion, together with the fact that AlkB mostly binds to the lesion-containing strand of DNA, explains the preference of AlkB for ssDNA substrates. The presence of the complementary strand in dsDNA rigidifies the duplex structure and serves as a ‘non-competitive’ inhibitor. The binding energy of complex formation must be used to pay an energetic penalty to flip 1-meA from a duplex DNA. A much lower energetic penalty would be associated with AlkB binding to the same base lesion in flexible ssDNA.

The ABH2–DNA structures clearly show ABH2 as a dsDNA repair protein, as the protein interacts extensively with both strands of dsDNA and uses a finger residue to intercalate in the DNA duplex to maintain continuous stacking. It has been commonly observed that single-base DNA glycosylases tend to contact only the lesion-containing strand of duplex DNA[31,34,35]. The extra capacity to interact with the other strand of DNA allows ABH2 to repair dsDNA

lesions preferentially, which correlates well with its primary role as an enzyme that repairs 1-meA damage in mammalian genomes. Our structures also reveal a more complex arrangement of protein-(1-meA) interactions in the active site of ABH2 than AlkB, which provides a structural basis for designing small molecules that may shut down the ABH2-based repair activity potentially to improve efficacy of alkylation anticancer treatments.

Now with structures of AlkB–dsDNA, ABH2–dsDNA and ABH3 available, comparisons can be made (Fig. 5 and Supplementary Fig. 15). Surprisingly, ABH3, which shows substrate preferences similar to AlkB, shares some of the DNA-binding motifs of ABH2 in an overlay of these three structures. The long loop that binds the second DNA strand in ABH2 (between β11 and β12) is present in the sequence though disordered in the structure of ABH3 (between β10 and β11), but is much shorter in AlkB. Both ABH2 and ABH3 also contain a b-sheet hairpin motif (β3–β4 and β4–β5, respectively) that is absent in AlkB. In ABH2 this hairpin contains the aromatic finger residue, Phe 102, which is not conserved in ABH3. ABH3 also lacks the RKK motif used by ABH2 to contact the complementary strand of double-stranded DNA. How ABH3 may use these motifs to bind ssDNA or RNA preferentially remains to be seen in future structures of ABH3–DNA complexes.

Both AlkB and ABH2 are weak DNA-binding proteins. They likely interact with other DNA-binding proteins to stay close to DNA. ABH2, with a finger residue, may search for DNA damage by similar mechanisms proposed for DNA repair glycosylases[33,36]. AlkB, using a very different base-flipping mechanism, may impose a strain to the backbone of duplex DNA to facilitate identification of weakened base pairs, as suggested by the structures presented here. However, how does it locate damaged bases in flexible ssDNA? Further studies are required to answer these intriguing questions.

Lastly, the active site cross-linking strategy to trap labile protein–DNA complexes could be powerful for structurally characterizing other base-flipping proteins such as FTO and ABH3. As demonstrated here, once 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.

Methods Summary

The disulphide cross-linked protein–DNA complexes were purified using Mono-Q anion exchange chromatography, and crystals were grown by hanging-drop vapour diffusion. Diffraction data were collected from cryo-preserved crystals at beamlines 19BM and 14BM at the Advanced Photon Source, Argonne National Laboratory, Argonne, Illinois. The structures of AlkB–dsDNA complexes were solved by molecular replacement, and the ABH2–dsDNA structure was solved by MAD phasing with a bromine-labelled DNA. Data collection and refinement parameters for all structures are given in Supplementary Tables 1 and 2. Detailed procedures are presented in Supplementary Information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

R. Zhang and other beamline staff for assistance with data collection; Y. Luo and the Proteomics and Informatics Services Facility at Research Resources Center (University of Illinois at Chicago) for liquid chromatography–mass spectrometry analysis; and X. Yang, H. Chen and P. R. Chen for discussions. We also thank T. Lindahl and B. Sedgwick for the gift of the *abh2* gene. This work was supported by National Institutes of Health (GM071440 to C.H. and a PCBio fellowship for C.T.S.), the W. M. Keck Foundation Distinguished Young Scholar in Medical Research Program (C.H.), and the Arnold and Mabel Beckman Foundation Young Investigator Program (C.H.). Data collection was

performed at beamlines 19BM (Structure Biology Center) and 14BM (BioCARS) at the Advanced Photon Source at Argonne National Laboratory; financial support for these beamlines comes from the National Institutes of Health and the United States Department of Energy.

References

1. Friedberg, EC.; Walker, GC.; Siede, W. DNA Repair and Mutagenesis. Ch 1. ASM Press; Washington DC: 1995.
2. Wood RD, Mitchell M, Sgouros J, Lindahl T. Human DNA repair genes. Science 2001;291:1284–1289. [PubMed: 11181991]
3. Sedgwick B. Repairing DNA-methylation damage. Nature Rev Mol Cell Biol 2004;5:148–157. [PubMed: 15040447]
4. Mishina Y, He C. Oxidative dealkylation DNA repair mediated by the monomuclear non-heme iron AlkB proteins. J Inorg Biochem 2006;100:670–678. [PubMed: 16469386]
5. Lindahl T, Sedgewick B, Sekiguchi M, Nakabeppu Y. Regulation and expression of the adaptive response to alkylating agents. Annu Rev Biochem 1988;57:133–157. [PubMed: 3052269]
6. Falnes PO, Johansen RF, Seeberg E. AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. Nature 2002;419:178–182. [PubMed: 12226668]
7. Trewick SC, Henshaw TF, Hausinger RP, Lindahl T, Sedgwick B. Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. Nature 2002;419:174–178. [PubMed: 12226667]
8. Aravind L, Koonin EV. The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. Genome Biol 2001;2:research0007. [PubMed: 11276424]
9. 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. [PubMed: 15269201]
10. Falnes PO. Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. Nucleic Acids Res 2004;32:6260–6267. [PubMed: 15576352]
11. Delaney JC, Essigmann JM. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in *alkB Escherichia coli*. Proc Natl Acad Sci USA 2004;101:14051–14056. [PubMed: 15381779]
12. Mishina Y, Yang CG, 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. [PubMed: 16231911]
13. Delaney JC, et al. AlkB reverses etheno DNA lesions caused by lipid oxidation *in vitro* and *in vivo*. Nature Struct Mol Biol 2005;12:855–860. [PubMed: 16200073]
14. Duncan T, et al. Reversal of DNA alkylation damage by two human dioxygenases. Proc Natl Acad Sci USA 2002;99:16660–16665. [PubMed: 12486230]
15. Aas PA, et al. Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. Nature 2003;421:859–863. [PubMed: 12594517]
16. Lee DH, et al. Repair of methylation damage in DNA and RNA by mammalian AlkB homologues. J Biol Chem 2005;280:39448–39459. [PubMed: 16174769]
17. Gerken T, et al. The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. Science 2007;318:1469–1472. [PubMed: 17991826]
18. Ringvoll J, et al. Repair deficient mice reveal mABH2 as the primary oxidative demethylase for repairing 1meA and 3meC lesions in DNA. EMBO J 2006;25:2189–2198. [PubMed: 16642038]
19. Yu B, et al. Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. Nature 2006;439:879–884. [PubMed: 16482161]
20. Sundheim O, et al. Human ABH3 structure and key residues for oxidative demethylation to reverse DNA/RNA damage. EMBO J 2006;25:3389–3397. [PubMed: 16858410]
21. Mishina Y, Chen LX, 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. [PubMed: 15612731]
22. Dinglay S, Trewick SC, Lindahl T, Sedgwick B. Defective processing of methylated single-stranded DNA by *E. coli* AlkB mutants. Genes Dev 2000;14:2097–2105. [PubMed: 10950872]

23. Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase at 2.7 Å resolution: implications of comformational changes for polymerization and inhibition mechanism. Science 1998;282:1669–1675. [PubMed: 9831551]
24. Verdine GL, Norman DPG. Covalent trapping of protein–DNA complexes. Annu Rev Biochem 2003;72:337–366. [PubMed: 14527324]
25. 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. [PubMed: 12862460]
26. Duguid EM, Mishina Y, He C. How do DNA repair proteins locate potential base lesions? A chemical crosslinking method to investigate the damage-searching mechanism of O^6-alkylguanine-DNA alkyltransferases. Chem Biol 2003;10:827–835. [PubMed: 14522053]
27. Mishina Y, Lee CH, He C. Interaction of human and bacterial AlkB proteins with DNA as probed through chemical cross-linking studies. Nucleic Acids Res 2004;32:1548–1554. [PubMed: 15004242]
28. Schofield CJ, Zhang Z. Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. Curr Opin Struct Biol 1999;9:722–731. [PubMed: 10607676]
29. Lange SJ, Que L Jr. Oxygen activating nonheme iron enzymes. Curr Opin Chem Biol 1998;2:159–172. [PubMed: 9667935]
30. Elkins JM, et al. X-ray crystal structure of *Escherichia coli* taurine/α-ketoglutarate dioxygenase complexed to ferrous iron and substrates. Biochemistry 2002;41:5185–5192. [PubMed: 11955067]
31. Lau AY, Scharer OD, Samson L, Verdine GL, Ellenberger T. Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanismf for nucleotide flipping and base excision. Cell 1998;95:249–258. [PubMed: 9790531]
32. Fromme JC, Banerjee A, Huang SJ, Verdine GL. Structural basis for removal of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase. Nature 2004;427:652–656. [PubMed: 14961129]
33. Banerjee A, Santos WL, Verdine GL. Structure of a DNA glycocylase searching for lesions. Science 2006;311:1153–1157. [PubMed: 16497933]
34. Slupphaug G, et al. A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. Nature 1996;384:87–92. [PubMed: 8900285]
35. Bruner SD, Norman DPG, Verdine GL. Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. Nature 2000;403:859–866. [PubMed: 10706276]
36. Parker JB, et al. Enzymatic capture of an extrahelical thymine in the search for uracil in DNA. Nature 2007;449:433–437. [PubMed: 17704764]

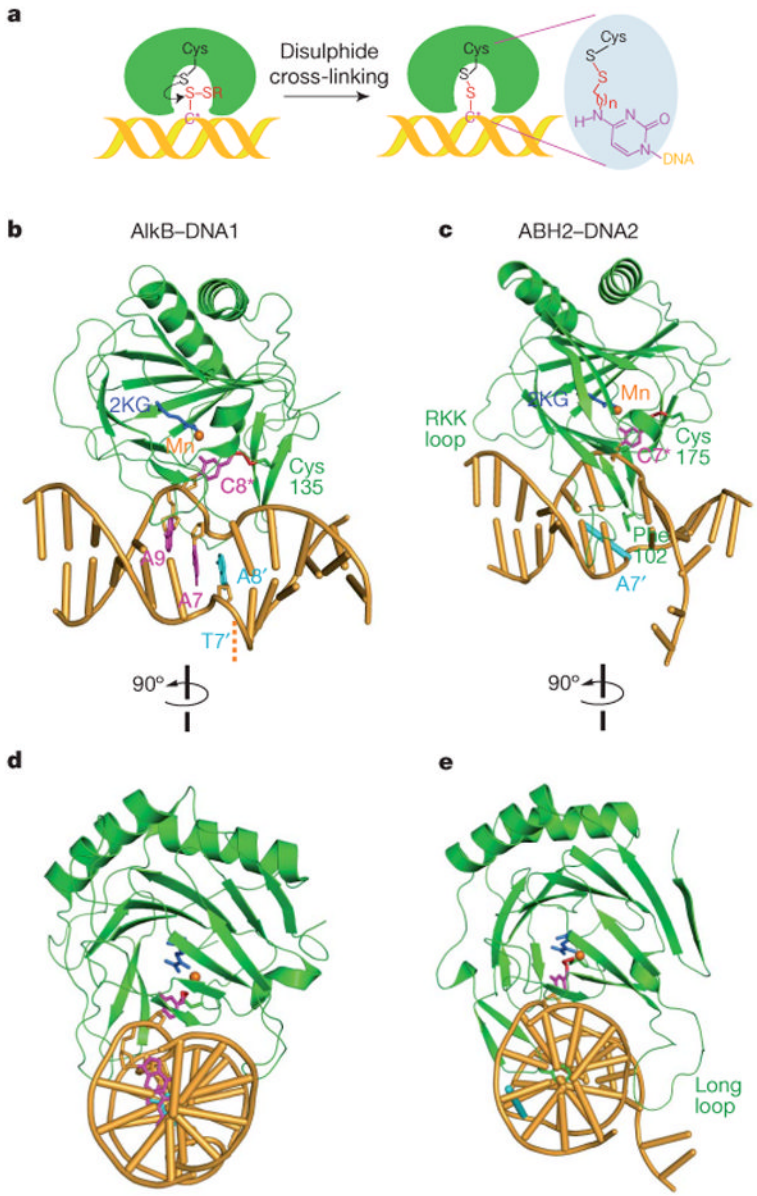

Figure 1. Crystal structures of the active site cross-linked AlkB–DNA1 and ABH2–DNA2 complexes

a, The active site cross-linking strategy. A disulphide-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 disulphide bond can form between C* and an engineered Cys in the active site of the repair protein under equilibrium conditions. **b**, Cartoon of the AlkB–DNA1 complex. Two slightly different DNA conformations are observed in this structure; one is shown here, the other in Supplementary Fig. 8. The protein is shown in green (the cross-linking residue Cys 135 is labelled), manganese(II) in orange, 2KG in blue, DNA in yellow–orange, the flipped base C8* and the two bases flanking C8* in light magenta, the disulphide bond in red and the orphaned base A8′ in cyan. The approximate location of the disordered base T7′ is shown as a dotted line. **c**, Cartoon trace of the ABH2–DNA2 complex. The same colour coding as in **b** is used. The finger residue Phe 102 and the cross-linking residue Cys 175 are green, and the orphaned base A7′ is cyan. The location of a DNA-binding loop containing a RKK sequence is indicated.

d, e, Same structure as in **b** and **c**, rotated 90° to the right. A DNA-binding flexible long loop is labelled. See Supplementary Fig. 15 for a topological diagram of ABH2.

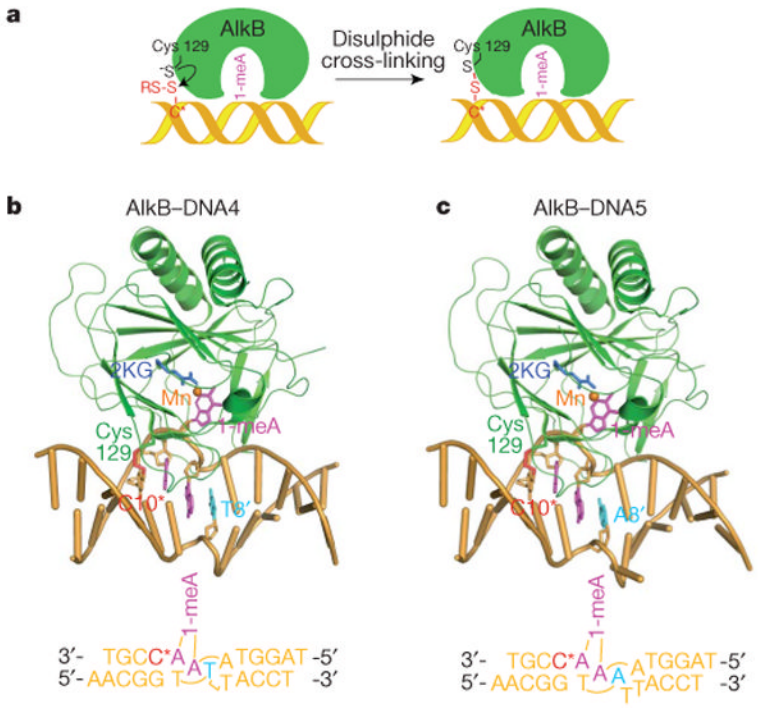

Figure 2. Crystal structures of AlkB–DNA4 and AlkB–DNA5 complexes with 1-meA recognized by an intact active site

a, A strategy to set up a disulphide cross-link away from the active site of AlkB. **b**, Cartoon of the AlkB–DNA4 complex with 1-meA (light magenta) recognized in the active site of AlkB. The complex was stabilized by a disulphide link between an engineered Cys 129 (green) and C10* that is two base pairs away from 1-meA. The colour coding is the same as in Fig. 1b. DNA4 has an abasic site 'ب' (position 7′) 3′ to T8′ of the 1-meA:T base pair. **c**, Cartoon trace of the AlkB–DNA5 complex; DNA5 has a central 1-meA:A base pair replacing a C*:A in DNA1 and a C10* cross-linking site two base pairs away from 1-meA.

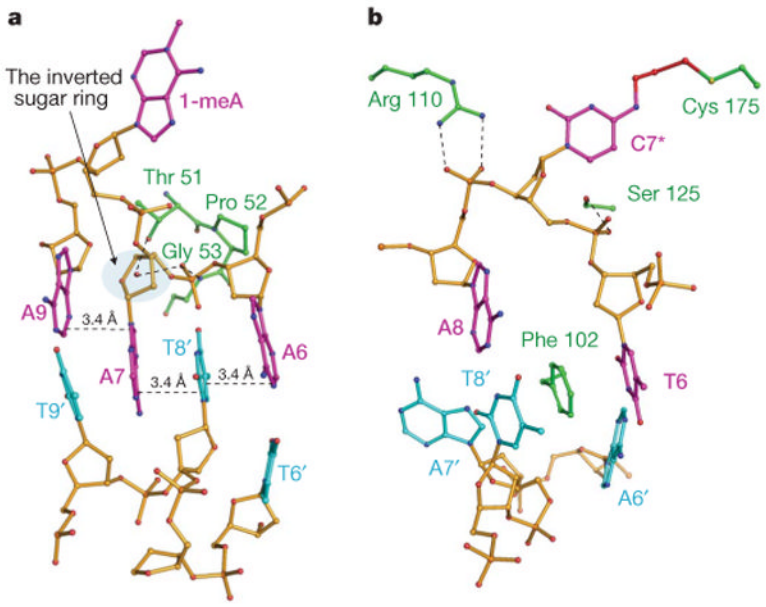

Figure 3. Close views of the base-flipping regions

a, AlkB–DNA4 complex, with the same colouring as in Fig. 1b. A7 stacks with the A9:T9′ base pair. Thr 51 and Gly 53 are hydrogen bonded to the phosphate linking A7 and A6 (Supplementary Fig. 10b). The sugar ring of A9 adopts the 3′-endo conformation whereas the sugar ring of A7 is forced to be inverted by about 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**, ABH2–DNA2 complex, with the same colour coding as in Fig. 1b. Phe 102 inserts into the DNA duplex with A7′ as the orphaned base.

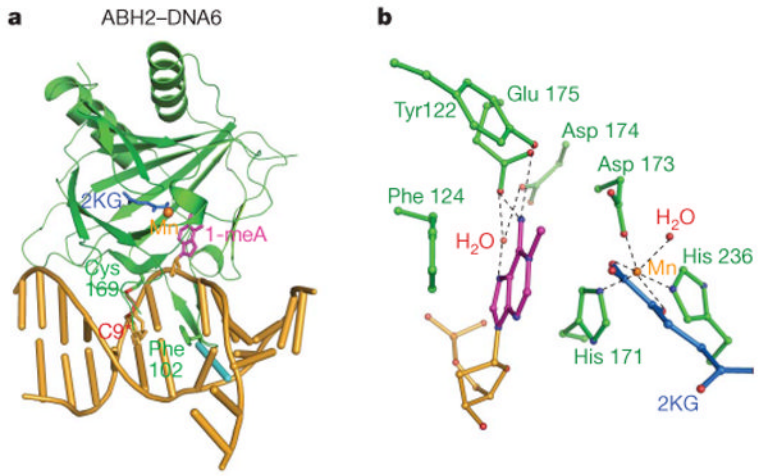

Figure 4. Crystal structure of the ABH2–DNA6 complex with 1-meA recognized by an intact active site

a, Cartoon trace of the ABH2–DNA6 complex with 1-meA recognized in the active site of ABH2 and a disulphide cross-link (Cys 169 to C9*) installed two base pairs away from the flipped base. The same colour scheme is used as in Fig. 1b. **b**, View of the active site of ABH2 with manganese(II) (orange), 2KG (blue) and 1-meA (light magenta). N6 of 1-meA is within hydrogen-bond distance to Tyr 122 (3.3 Å) and Glu 175 (2.8 Å). N7 forms a hydrogen bond to a water (3.0 Å) which also interacts with Asp 174 (2.6 Å) and Glu 175 (2.6 Å).

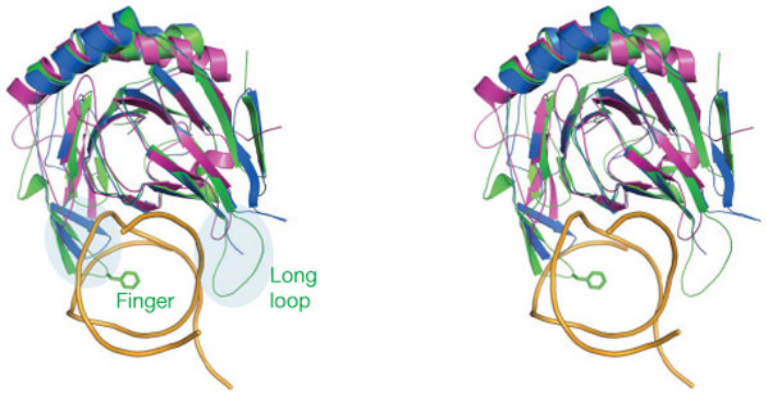

Figure 5. Structural comparison of AlkB, ABH2 and ABH3 (stereo view)

Least-squares superposition of structures of AlkB (in magenta, from AlkB–DNA1 without showing DNA1), ABH2–DNA2 (protein in green and DNA backbone in orange) and ABH3 (in blue, from Protein Data Bank 2IUW). The hairpin and flexible long-loop motifs are shaded, and the finger residue Phe 102 from ABH2 is labelled in green.