Biological Sciences, Biochemistry

SUPPORTING INFORMATION
for

Enzymological and structural studies of the mechanism of promiscuous substrate recognition by the oxidative DNA repair enzyme AlkB.

Bomina Yu [1] and John F. Hunt [1,2] *

[1] Department of Biological Sciences, 702A Fairchild Center, MC2434,
Columbia University, New York, NY 10027;

and

[2] Northeast Structural Genomics Consortium.

* Corresponding author: (212)-854-5443 voice; (212)-865-8246 FAX;
jfhunt@biology.columbia.edu.

Final version: Wednesday, August 5, 2009 @ 4:17 PM.

Supplemental Discussion

Comparison with previous enzymological results on repair of m3C by E. coli AlkB. Previous studies have suggested that *E. coli* AlkB repairs m1A and m3C lesions with comparable efficiency (1-6). We also observe similar net catalytic efficiency (k_{cat}/K_M ratio) for these two substrates (Table 1). However, our Michaelis-Menten analyses indicate offsetting order-of-magnitude differences in the values of the two contributing kinetic parameters (Table 1). Few of the previously published studies have presented detailed kinetic analyses, and most have employed heterogeneously chemically alkylated polymers of varying length (Table S1). Only one previous study directly evaluated individual kinetic parameters and showed little difference between k_{cat} and K_M values for m1A and m3C lesions in synthetic DNA 19-mers (6). This discrepancy compared to our results could be attributable to substrate depletion in the earlier work, which produces systematic underestimation of reaction velocity especially at high substrate concentrations. Our preliminary repair assays on the TmCT trimer (Fig. S1A), which were conducted under standard conditions in the literature, showed a systematic underestimation of both k_{cat} and K_M due to substrate depletion. Therefore, we used shorter reaction times and explicitly verified that there was no more than 20% substrate turnover in any assay used for Michaelis-Menten analysis (Fig. S1B).

Practical limitations in the amount of the TmCT substrate that can be used in an assay prevent direct measurement of the reaction rate at fully saturating concentrations of this trinucleotide substrate, producing greater uncertainty in the values of the kinetic parameters for this substrate compared to TmAT (for which V_{max} can be measured directly at saturating substrate concentrations). However, Michaelis-Menten analysis of m3C and m1A lesions in the pentanucleotides CAmCAT and CAmAAT show comparable differences in the individual k_{cat} and

K_M parameters. Thanks to the significantly lower K_M values for these longer substrates, saturating substrate concentrations were easily achieved in these assays, and an equivalent pattern of k_{cat}/K_M compensation was observed (Fig. 2 & Table 1 in the main text).

Additional analyses related to repair of ethenoadenine. Consistent with previously published results (7, 8), our enzyme assays demonstrate that 1,N^6-ethenoadenine (εA) is repaired by *E. coli* AlkB, although very inefficiently compared to m1A and m3C bases (with ~1000-fold lower k_{cat}/K_M). Previously, epoxide and glycol derivatives of the adenine base were detected by electrospray ionization time-of-flight mass spectrometry and hypothesized to be intermediates in the repair of 1,N^6-ethenoadenine (8). Formation of such intermediates could potentially interfere with HPLC observation of product generation. While our HPLC assays did not show any additional peaks other than the 1,N^6-ethenoadenine substrate and the adenine product (Fig. S2), it is not known whether epoxide or glycol species would be resolved using this technique. Even if these oxidized species were accumulating in our enzyme assays, AlkB would have poor efficiency in actually repairing the modified base. However, such accumulation seems unlikely. Increasing the enzyme concentration in these assays resulted in a linearly proportional increase in the repair of TεAT (data not shown). These observations indicate that there is no significant substrate depletion even at the highest enzyme concentration tested, which was 10-fold higher than in assays on the TmAT and TmCT substrates, indicating that *E. coli* AlkB has comparatively very low activity even in initial oxidation of the εA base.

To assess whether steric hindrance may interfere with accommodation of the larger εA base in the active site of AlkB, it was modeled into the Fe(II)-2OG-TmAT complex structure (Fig. S12). After substituting the structure of the alternatively modified base, the model was refined (9) against the diffraction data obtained for the Fe(II)-2OG-TmAT-bound structure. The R-factors increased by only a few tenths of a percent compared to the proper structure with TmAT. While negative density appeared for the extra carbon atom, the TεAT stereochemistry was easily accommodated with no rearrangement of surrounding protein sidechains (Fig. S12). The closest contacts to the additional carbon atom were residues W69, D133 and R210 (which were at least 3.4, 3.5 and 3.6 Å away, respectively). The extra bulk of the larger alkyl modification might produce steric hindrance contributing to the reduced turnover rate of the TεAT substrate (Table 1), perhaps due to interference with CO_2 release or oxyferryl ligand migration dynamics (10). However, the previously published observation that AlkB repairs εA more efficiently than the smaller εC substrate (8) indicates that steric hindrance is not the only factor contributing to the inefficient repair of etheno nucleotides. The more complex chemistry involved in repairing etheno lesions could represent an inherent limitation in the utility of AlkB-like oxidative repair even though the εA base can be accommodated in the active site (Fig. S12).

Prior to the discovery of the ability of AlkB to repair εA, repair of this lesion and the equivalent εC lesion was attributed exclusively to DNA glycosidases in the base excision repair (BER) pathway (11). Moreover, the AlkA enzyme in this pathway was found to be more important than AlkB for repair of εA *in vivo* in *E. coli* (8). While redundancy in critical enzyme functions can provide several evolutionary advantages (12), the existence of a more efficient pathway for repairing etheno lesions lessens the evolutionary pressure for AlkB to optimize its efficiency in repairing them.

Conformations of bound trinucleotide substrates containing 1-methyladenosine vs. 3-methylcytosine. The TmCT substrate is bound to AlkB in a similar conformation to that previously observed for TmAT. The flanking thymine bases stack with each other while the central methylated base is flipped out to point in the opposite direction, enabling encapsulation of this single base in the active site (Fig. S8). When the different trinucleotide structures are superimposed based on least-squares alignment of the ribose ring of the central methylated nucleotide, the flanking thymine nucleotides are observed to have very similar conformations (Fig. S8A). The only significant differences are variations in the orientation of the ribose ring and the phosphate dihedral angles on the 5' terminal thymine. While the electron density for this phosphate is either weak or absent, indicating significant disorder within individual structures, the electron density for the ribose ring is mostly well defined, and its orientation in both TmCT structures is turned ~90° compared to that observed in TmAT structures (Fig. S8A). The TmCT structures all have the ribose ring of the 5' terminal thymine oriented essentially parallel to that in the 3' terminal thymine, whereas these rings are nearly perpendicular to one another in all but one of the TmAT structures. The lone exception is the TmAT substrate in the Co(II)-bound structure (yellow in Fig. S8A) in which the electron density for the 5' ribose is weak and poorly defined.

The consistent difference in 5' ribose conformation in the m1A *vs.* m3C structures is probably caused by the deeper penetration of the m3C substrate into the active site coupled to the observed movements of FL1 and FL2 (Fig. 4). However, the exact stereochemical explanation for this conformational difference is not clear. No obvious steric clashes occur when either trinucleotide substrate is modeled into structures obtained with the other one bound, and there are few interatomic contacts closer than 3.9 Å to the 5' ribose in any of the structures. A single weak hydrogen bond (3.5-3.6 Å) between residue Y76 in FL2 is consistently observed to the O3' atom of the ribose in all of the m1A structures showing a clearly parallel orientation of the two ribose rings. This interaction or weaker interactions with the many atoms in FL1 located just beyond nominal van der Waals interaction distance may couple the observed conformation of the 5' ribose to the differences in the conformations of the two flexible loops observed in the m1A *vs.* m3C structures (Figs. S3A, S4A, & S5A).

Recently, three crystal structures have been reported with *E. coli* AlkB-ΔN11 chemically crosslinked to different double-stranded DNA substrates (10), and these structures potentially provide additional information regarding the conformational preferences of bound oligonucleotide substrates. In one of these structures, a sulfur-derivatized cytosine in the substrate is crosslinked to residue 135, which was mutated from aspartate to cysteine. The crosslinked cytosine base is located in the active site binding slot sandwiched between residues H131 and W69 but displaced by ~2 Å from the m3C location observed in our non-crosslinked structures. In the other two crosslinked crystal structures, the double-stranded DNA substrates contain an m1A base bound in the active site and a disulfide-tethered cytosine two nucleotides away towards the 3' terminus of the same strand. These substrates were crosslinked to residue 129, which was mutated from serine to cytosine. The m1A bases in both of these crosslinked structures are bound in a similar conformation to that observed in our non-crosslinked structures but displaced by ~0.3 Å towards the entrance to the binding slot. Except for these displacements within the binding slot, the ribose group of the modified base and both of the covalently bound phosphate groups adopt similar conformations to those observed in all of our crystal structures.

In contrast, the flanking nucleotides on the bound DNA strand show substantially different conformations in the crosslinked structures compared to the conformation of the flanking bases observed in all of our crystal structures. The crosslinked structures all show a very similar oligonucleotide conformation on the 3' side where the crosslink is located in both m1A structures. However, this crosslink was engineered to preserve the binding geometry observed in the structure with cytosine crosslinked in the active site. If that structure was perturbed by the presence of the active-site crosslink, as seems possible based on the ~2 Å displacement of the cytosine base compared to our non-crosslinked structures, the common 3' oligonucleotide conformation observed in all of the crosslinked structures could be biased. Moreover, the 5' oligonucleotide conformation varies considerably between the structures obtained using the different crosslinking strategies. These conformational variations in the DNA substrate are coupled to variations in the conformations of FL1 and FL2 in AlkB. Therefore, the conformational preferences of the DNA segments 3' and 5' to the modified base and the stereochemistry of their interaction with the enzyme remain difficult to interpret confidently in spite of the extensive collection of crystal structures. The conformational plasticity of the loops in the nucleotide-recognition lid in AlkB potentially contributes to accommodating diverse conformations of the DNA strand flanking the damaged base.

Potential allosteric control of O_2 access to the active site of AlkB. During the reaction cycle of AlkB, molecular oxygen binds to iron and oxidizes the 2OG co-substrate to succinate and CO_2 (10). This reaction results in formation of a highly reactive oxygen atom on the Fe cofactor, a complex that is called an oxy-ferryl intermediate (Fe(IV)=O). Following release of CO_2, this reactive oxygen atom migrates towards the site vacated by the CO_2 and hydroxylates the alkyl moiety of the nucleotide substrate, which is later spontaneously hydrolyzed to release an aldehyde product. The first and second phases of this multi-step enzyme-bound redox reaction must be tightly coupled to prevent release of the reactive oxygen atom and generation of oxidative stress in the cytosol (13, 14). Such coupling cannot be ensured unless O_2 binds to AlkB last after the DNA substrate, because binding of O_2 and 2OG in the absence of DNA could result in gratuitous conversion of 2OG to succinate and release the reactive oxygen species. However, the binding of 2OG and DNA blocks solvent access to the active site. Therefore, O_2 was hypothesized to enter the active site via an alternative route through the protein, specifically the putative O_2-diffusion tunnel shown in Figs. 3B & S3B (15). (It is noteworthy that several members of the Fe-2OG dioxygenase superfamily self-hydroxylate active site residues upon prolonged O_2 exposure in the absence of their primary substrate (16-18). This reaction has been observed at residue W178 in AlkB (19), which is the active-site tryptophan residue that undergoes a rotamer flip coupled to opening/closing of the putative O_2-diffusion tunnel (Fig. 3B & S3B).)

The conformational transition opening the O_2-diffusion tunnel in *E. coli* AlkB appears to be controlled by inter-protomer packing interactions in the available crystal structures. The observed conformation does not correlate with either the identity of the bound ligands or the spacegroup symmetry of the crystals (Table S3), which were all grown from very similar mother liquors. However, it does correlate perfectly with the occurrence of inter-protomer crystal packing contacts to both the 158-164 loop (in CC4) and the entrance of the O_2-diffusion tunnel (spanning CC3 and CC5). Different inter-protomer contacts at these two sites are consistently observed in all open structures compared to all closed structures (Fig. S7). The sidechain of R161 has high B-factors in all open structures, where it is remote from crystal packing contacts, while it packs tightly against the sidechain of D39 from a different protomer in all closed structures (Fig. S7A). The global conformational transition occurs cooperatively throughout the crystal lattice and couples formation of these inter-protomer packing contacts (Fig. S7) to conformational changes in protein segments participating in CC3, CC4, and CC5 (Figs. 3 & S4-S5). On this basis, we infer that this transition is an intrinsic property of the protein likely to involve a relatively modest free energy change, as observed for some other functional allosteric protein conformational changes (20, 21).

While it appears to be controlled by lattice-packing interactions in crystals, this allosteric conformational transition could be coupled to the functional state of the protein when it is not constrained in a crystal lattice. The transition couples local stereochemical changes in the active site to remote conformational changes in the 158-164 loop connecting strands β-C4 and β-C5. Residue R161 in this loop interacts weakly with the trinucleotide trimer in our structures in the open conformation (Fig. S8B & C) but more strongly with the larger nucleotide substrate in the DNA-crosslinked structures reported by another group (10) and discussed above. The crosslinked structures also show additional DNA-interactions involving other residues in the 158-164 loop.

The allosteric conformational transition in AlkB couples the conformation of this DNA-interacting loop to the conformation near the catalytic metal ion, which moves 0.3-0.4 Å relative to the conformationally invariant core of the protein during the transition (Fig. S5B and Table S4). In comparison, differences of less than 0.2 Å in metal-ion position are observed between structures in the same conformational state (Fig. S11). The displacement of the metal ion during the conformational transition is mediated mostly by small dihedral angle changes in the sidechains of its liganding residues, two of which are located in the conformationally invariant core (H131 and D133). Therefore, the global allosteric conformational transition in AlkB could potentially modulate the chemical reactivity of the catalytic Fe(II) ion center simultaneously with the accessibility of the O_2-diffusion tunnel and the conformational preferences of the DNA-interacting 158-164 loop.

The analyses presented here suggest a structural pathway by which binding of the alkylated nucleotide substrate might be able to influence the energetics of the observed conformational change and thereby promote opening and closing of the O_2-diffusion pathway for optimal O_2-binding, CO_2 product release, and/or oxy-ferryl migration. However, while such elegant substrate-dependent kinetic control of protein dynamics is possible in solution, the underlying allosteric conformational change in AlkB is clearly influenced by lattice packing interactions in the available crystal structures (Fig. S7). A rapid conformational change in AlkB was recently observed in solution by NMR (22). However, there is little overlap in the residues with broadened or shifted NMR resonances and those involved in the conformational change we observe in the crystal structures (Fig. S9).

Binding of the primary oxidation substrate (*i.e.*, alkylated DNA for AlkB) is believed to trigger O_2 binding in all members of the Fe-2OG-dioxygenase superfamily (13). A putative O_2-diffusion tunnel was first described in clavaminate synthase (23) and allosteric conformational changes of short disordered loops covering the active site have been suggested for other members of the superfamily (13, 24, 25). However, the crystallographic results reported in this paper represent the first time that two distinct global conformational states were clearly documented in an Fe-2OG-dioxygenase.

Structural comparison of E. coli AlkB to human ABH2 and ABH3. There are nine eukaryotic AlkB homologues (ABH1-8 and FTO) (26-29). ABH1-3 and FTO have been confirmed to repair endocyclic alkylation lesions on nucleic acid bases (1, 2, 10, 30), but they exhibit differences in their substrate specificities. For example, ABH3 has been shown to preferentially dealkylate single-stranded substrates and to be active on RNA. In contrast, AHB2 prefers double-stranded substrates and is not active on RNA (1, 2).

Crystal structures of human ABH2 and ABH3 have been published (10, 31) (Fig. S10). The structure of ABH2, in complex with Mn(II) and 2OG and chemically crosslinked to a double-stranded DNA molecule, revealed that the methylated base was flipped out of the DNA helix and inserted into the active site with the aromatic rings of residues H171 and F124 stacked on either side of the DNA base, analogous to the H131-DNA-W69 interaction in AlkB (Fig. S12B). But unlike the *E. coli* AlkB structures, ABH2 made considerable contacts to the complimentary, non-alkylated DNA strand. A phenylalanine at residue 102 of ABH2 inserted into the DNA helix in the space left by the flipped-out base. This phenylalanine finger maintains the distance between the flanking nucleotide bases such that no distortion occurs in the bound DNA helix. These protein-DNA interactions, which are absent in AlkB, explain the preference of ABH2 for double-stranded DNA over single-stranded DNA substrates.

The crystal structure of human ABH3 was determined with Fe(II) and 2OG bound in the active site, but without a bound nucleotide substrate (31). When this structure was compared to that of *E. coli* AlkB, distinct differences were observed in the region corresponding to the "nucleotide-binding lid", suggesting that ABH3 might bind nucleic acids in a different manner (31) (Fig. S10). This proposal was supported by the absence in ABH3 of any obvious equivalent of the aromatic residues that base-stack with the alkylated DNA base in AlkB (W69) and ABH2 (F124). Another difference in active-site stereochemistry is that there is no equivalent in ABH3 (or ABH2) of residue W178 in AlkB, which "flips" its conformation during opening/closing of the O_2-diffusion tunnel (Fig. S3B & S4B). Instead, the crystal structure of ABH3 is hydroxylated on L177 (31) (Fig. S12). In spite of these differences, given the conformational plasticity observed in comparing the DNA-bound *vs.* nucleotide-free structures of *E. coli* AlkB (15), it is conceivable that flexible regions in ABH3 could change their conformation during the binding of a DNA substrate to produce DNA-recognition stereochemistry similar to that of AlkB.

Alternatively, it is possible that the primary function of ABH3 is something distinct from DNA repair. Protein characteristics that allow an enzyme to act promiscuously can also make it more evolutionarily adaptable (12), and ABH3 may have evolved another enzymatic activity while also retaining an AlkB-like activity. Knockout mice lacking ABH3 have phenotypes indistinguishable from wild-type animals, while those lacking ABH2 exhibited accumulation of endogenous m1A lesions as well as impaired DNA repair after methylmethane sulfonate damage (32). These experiments demonstrate DNA-repair activity under physiological conditions for ABH2 but not ABH3. While *in vitro* DNA repair by ABH3 has been demonstrated (1, 2, 30), its primary biological activity may involve something other than repairing m1A and m3c lesions in damaged DNA.

Supplemental Methods

Protein crystallization. AlkB-ΔN11 was crystallized with the Fe(II) cofactor in an anaerobic glovebox (COY Laboratories Products Inc., Grass Lake, MI) by the hanging-drop vapor diffusion method at room temperature. The well solution (18% PEG 3350, 0.2 M sodium formate) and $Fe(NH_4)_2(SO_4)_2$ crystals were brought into the glovebox at least 24 hr before setting up reactions, after which the $Fe(NH_4)_2(SO_4)_2$ crystals were dissolved in degassed water. Protein, 2OG, and TmCT solutions were brought into the glove box immediately before use and mixed to final concentrations of 0.4 mM, 5.4 mM, and 7.0 mM, respectively. $Fe(NH_4)_2(SO_4)_2$ was added to a final concentration of 2.2 mM prior to mixing 1 μl of this protein solution with an equal volume of the well solution. Tetragonal bipyramidal crystals were obtained by streak-seeding with crystals grown previously with TmAT instead of TmCT (15). Co-crystals of AlkB-ΔN11 and Mn(II) were grown in the same manner but in an aerobic environment (*i.e.*, not in the glovebox) by substituting $Fe(NH_4)_2(SO_4)_2$ with $MnCl_24H2O$. Crystals were cryo-protected by brief passage through 22% PEG 3350, 0.2 M sodium formate, and 10% glycerol before freezing in liquid propane.

Structure determination and refinement. Single wavelength diffraction datasets for the Mn(II) and Fe(II) containing crystals were collected at cryogenic temperature on beamlines X12C and X4C, respectively, at the National Synchrotron Light Source at Brookhaven National Laboratory. After data reduction and scaling in DENZO/SCALEPACK (33), the structures were solved by molecular replacement using COMO (34) with the previously published AlkB-ΔN11 structure as the search model (15). Structures were completed via iterative rounds of manual model building and computational refinement using "O" (35), COOT (36), and CNS (9). After initial least-squares superposition of all Cα atoms in "O" (35), invariant regions of the core β-strands (residues 19-22, 115-134, 143-158, 166-178, 186-189, 204-210) in all structures were selected for the structural alignment of the catalytic core.

Supplemental References

1. Aas PA*, et al.* (2003) Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature* 421(6925):859-863.

2. Falnes PO, Bjoras M, Aas PA, Sundheim O, & Seeberg E (2004) Substrate specificities of bacterial and human AlkB proteins. *Nucleic Acids Res* 32(11):3456-3461.

3. Falnes PO, Johansen RF, & Seeberg E (2002) AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli. *Nature* 419(6903):178-182.

4. Trewick SC, Henshaw TF, Hausinger RP, Lindahl T, & Sedgwick B (2002) Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. *Nature* 419(6903):174-178.

5. Delaney JC & Essigmann JM (2004) Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. *Proc Natl Acad Sci U S A* 101(39):14051-14056.

6. Roy TW & Bhagwat AS (2007) Kinetic studies of Escherichia coli AlkB using a new fluorescence-based assay for DNA demethylation. *Nucleic Acids Res*.

7. Mishina Y, Yang CG, & He C (2005) Direct Repair of the Exocyclic DNA Adduct 1,N(6)-Ethenoadenine by the DNA Repair AlkB Proteins. *J Am Chem Soc* 127(42):14594-14595.

8. Delaney JC*, et al.* (2005) AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo. *Nat Struct Mol Biol* 12(10):855-860.

9. Brunger AT*, et al.* (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54 (Pt 5):905-921.

10. Yang CG*, et al.* (2008) Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA. (Translated from eng) *Nature* 452(7190):961-965.

11. Guliaev AB, Singer B, & Hang B (2004) Chloroethylnitrosourea-derived ethano cytosine and adenine adducts are substrates for Escherichia coli glycosylases excising analogous etheno adducts. (Translated from eng) *DNA Repair (Amst)* 3(10):1311-1321.

12. O'Brien PJ (2006) Catalytic promiscuity and the divergent evolution of DNA repair enzymes. (Translated from eng) *Chem Rev* 106(2):720-752.

13. Purpero V & Moran GR (2007) The diverse and pervasive chemistries of the alpha-keto acid dependent enzymes. (Translated from eng) *J Biol Inorg Chem* 12(5):587-601.

14. Valko M, Rhodes CJ, Moncol J, Izakovic M, & Mazur M (2006) Free radicals, metals and antioxidants in oxidative stress-induced cancer. (Translated from eng) *Chem Biol Interact* 160(1):1-40.

15. Yu B*, et al.* (2006) Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. *Nature* 439(7078):879-884.

16. Liu A*, et al.* (2001) Alternative reactivity of an alpha-ketoglutarate-dependent iron(II) oxygenase: enzyme self-hydroxylation. (Translated from eng) *J Am Chem Soc* 123(21):5126-5127.

17. Ryle MJ, Koehntop KD, Liu A, Que L, Jr., & Hausinger RP (2003) Interconversion of two oxidized forms of taurine/alpha-ketoglutarate dioxygenase, a non-heme iron hydroxylase: evidence for bicarbonate binding. (Translated from eng) *Proc Natl Acad Sci U S A* 100(7):3790-3795.

18. Ryle MJ*, et al.* (2003) O2- and alpha-ketoglutarate-dependent tyrosyl radical formation in TauD, an alpha-keto acid-dependent non-heme iron dioxygenase. (Translated from eng) *Biochemistry* 42(7):1854-1862.

19. Henshaw TF, Feig M, & Hausinger RP (2004) Aberrant activity of the DNA repair enzyme AlkB. *J Inorg Biochem* 98(5):856-861.

20. Gunasekaran K, Ma B, & Nussinov R (2004) Is allostery an intrinsic property of all dynamic proteins? (Translated from eng) *Proteins* 57(3):433-443.

21. Tsai CJ, del Sol A, & Nussinov R (2008) Allostery: absence of a change in shape does not imply that allostery is not at play. (Translated from eng) *J Mol Biol* 378(1):1-11.

22. Bleijlevens B*, et al.* (2008) Dynamic states of the DNA repair enzyme AlkB regulate product release. (Translated from Eng) *EMBO Rep* .

23. Zhang Z*, et al.* (2002) Crystal structure of a clavaminate synthase-Fe(II)-2-oxoglutarate-substrate-NO complex: evidence for metal centered rearrangements. *FEBS Lett* 517(1-3):7-12.

24. Muller I*, et al.* (2004) Crystal structure of the alkylsulfatase AtsK: insights into the catalytic mechanism of the Fe(II) alpha-ketoglutarate-dependent dioxygenase superfamily. *Biochemistry* 43(11):3075-3088.

25. O'Brien JR, Schuller DJ, Yang VS, Dillard BD, & Lanzilotta WN (2003) Substrate-induced conformational changes in Escherichia coli taurine/alpha-ketoglutarate dioxygenase and insight into the oligomeric structure. (Translated from eng) *Biochemistry* 42(19):5547-5554.

26. Kurowski MA, Bhagwat AS, Papaj G, & Bujnicki JM (2003) Phylogenomic identification of five new human homologs of the DNA repair enzyme AlkB. *BMC Genomics* 4(1):48.

27. Wei YF, Carter KC, Wang RP, & Shell BK (1996) Molecular cloning and functional analysis of a human cDNA encoding an Escherichia coli AlkB homolog, a protein involved in DNA alkylation damage repair. *Nucleic Acids Res* 24(5):931-937.

28. Duncan T*, et al.* (2002) Reversal of DNA alkylation damage by two human dioxygenases. *Proc Natl Acad Sci U S A* 99(26):16660-16665.
29. Jia G*, et al.* (2008) Oxidative demethylation of 3-methylthymine and 3-methyluracil in single-stranded DNA and RNA by mouse and human FTO. (Translated from Eng) *FEBS Lett* .
30. Lee DH*, et al.* (2005) Repair of methylation damage in DNA and RNA by mammalian AlkB homologues. *J Biol Chem* 280(47):39448-39459.
31. Sundheim O*, et al.* (2006) Human ABH3 structure and key residues for oxidative demethylation to reverse DNA/RNA damage. (Translated from eng) *Embo J* 25(14):3389-3397.
32. Ringvoll J*, et al.* (2006) Repair deficient mice reveal mABH2 as the primary oxidative demethylase for repairing 1meA and 3meC lesions in DNA. (Translated from eng) *Embo J* 25(10):2189-2198.
33. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307-326.
34. Jogl G, Tao X, Xu Y, & Tong L (2001) COMO: a program for combined molecular replacement. *Acta Crystallogr D Biol Crystallogr* 57(Pt 8):1127-1134.
35. Jones TA, Zou JY, Cowan SW, & Kjeldgaard (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47 (Pt 2):110-119.
36. Cowtan EP (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* D60:2126-2132.
37. Welford RW, Schlemminger I, McNeill LA, Hewitson KS, & Schofield CJ (2003) The selectivity and inhibition of AlkB. *J Biol Chem* 278(12):10157-10161.
38. Koivisto P, Duncan T, Lindahl T, & Sedgwick B (2003) Minimal methylated substrate and extended substrate range of Escherichia coli AlkB protein, a 1-methyladenine-DNA dioxygenase. *J Biol Chem* 278(45):44348-44354.
39. Koivisto P, Robins P, Lindahl T, & Sedgwick B (2004) Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases. *J Biol Chem* 279(39):40470-40474.
40. Mishina Y, Chen LX, & He C (2004) Preparation and characterization of the native iron(II)-containing DNA repair AlkB protein directly from Escherichia coli. *J Am Chem Soc* 126(51):16930-16936.
41. Falnes PO (2004) Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. *Nucleic Acids Res* 32(21):6260-6267.

42. Frick LE, Delaney JC, Wong C, Drennan CL, & Essigmann JM (2007) Alleviation of 1,N6-ethanoadenine genotoxicity by the Escherichia coli adaptive response protein AlkB. (Translated from eng) *Proc Natl Acad Sci U S A* 104(3):755-760.

43. Drenth J (1994) *Principles of protein x-ray crystallography* (Springer-Verlag, New York).

Table S1. Summary of *E. coli* AlkB substrates and their kinetic properties.

Substrates		Result			Detection Method	Reference
		Km (μM)	Kcat (min^{-1})	Kcat/Km (min^{-1} μM^{-1})		
1-methyladenine (m1A)						
– poly(dA), oligo(dA rich), ss- and ds-DNA treated with [^{13}C]MeI[a], [^{3}H]MNU[b], DMS[c], MMS[d]		turnover = 0.02 – 0.2 min^{-1}			HPLC/scintillation[e], MALDI-TOF[f], GC-MS[g]	(1-4, 6, 8, 37)
– poly(dA) treated with [^{13}C]MeI		1.4 ± 0.2	11.7 ± 1.3	8.6	HPLC/scintillation	(38)
– monomer (dA) treated with DMS		2.0	4.2	2.1	HPLC	(38)
– dimer (TA) treated with DMS		4.4	3.7	0.8	HPLC	(38)
– trimer (TAT) treated with DMS		2.8 ± 0.9	7.4 ± 0.6	2.6	HPLC	(38)
– trimer (T-m1A-T) synthesized, single m1A		1.2	4.5	3.8	HPLC	(15)
* *– trimer (T-m1A-T) synthesized, single m1A*		*1.4 ± 0.5*	*2.7 ± 0.8*	*1.9*	*HPLC*	*this study*
* *– pentamer (CA-m1A-AT) synthesized, single m1A*		*0.06 ± 0.01*	*5.4 ± 1.3*	*97.0*	*HPLC*	*this study*
– 19-mers, synthesized, single m1A	ss-DNA	5.4 ± 0.9	3.7 ± 0.2	0.7	FDH-coupled[h]	(6)
	ds-DNA	6.2 ± 1.3	3.1 ± 0.2	0.5		
– ss vector from 16-mer, synthesized, single m1A		bypass: to 100 % mutagenesis: not mutagenic			CRAB[i]/REAP[j] (*in vivo*)	(5)
3-methylcytosine (m3C)						
– poly(dC) treated with [^{13}C]MeI or [^{3}H]MNU		turnover = 0.02 – 0.2 min^{-1}			HPLC/scintillation	(1, 4, 39)
– 19-mers, synthesized with single m3C	ss-DNA	3.4 ± 0.6	2.2 ± 0.1	0.7	FDH-coupled	(6)
	ds-DNA	9.3 ± 2.4	3.3 ± 0.2	0.4		
* *– trimer (T-m3C-T) synthesized, single m3C*		*23.4 ± 4.7*	*21.3 ± 3.6*	*0.9*	*HPLC*	*this study*
* *– pentamer (CA-m3C-AT) synthesized, single m3C*		*0.29 ± 0.03*	*22.7 ± 9.5*	*78.3*	*HPLC*	*this study*
– ss vector from 16-mer, synthesized, single m3C		bypass: to 100 % mutagenesis: 30 % → 0 %			CRAB/REAP (*in vivo*)	(5)
– unmethylated 11-mer (T_5CT_5)		AlkB : K_d = 30 ± 7 μM AlkB-Fe(II) : K_d = 43 ± 14 μM			ITC[k]	(40)
random sequence (m1A or m3C)						
– 27-, 33-mer treated with MMS		??			1-[^{14}C]2OG[l] turnover	(37)

– M13 phage treated with MMS or [^{3}H]MNU, ss-DNA	turnover = 0.02 – 0.2 min^{-1} +++	HPLC, scintillation, survival in *E. coli* ΔAlkB	(1, 3, 28)
– λ Cl 857 phage treated with MMS, ds-DNA	+++	survival in *E. coli* ΔAlkB	(1)
– 24-mer treated with DMS or [^{3}H]MNU	??	scintillation, primer extension	(2, 41)
1-methylguanosine (m1G)			
* *– trimer (T-m1G-T) synthesized, single m1G*	*no repair*	*HPLC*	*this study*
– tRNA (random sequence) treated with [^{13}C]MeI	~100 % repaired	HPLC/scintillation	(41)
– ss vector from 16-mer, synthesized, single m1G	bypass: 2 % → 16 % mutagenesis: 80% → 4 %	CRAB/REAP (*in vivo*)	(5, 8)
3-methylthymidine (m3T)			
* *– trimer (T-m3T-T) synthesized*	*no repair*	*HPLC*	*this study*
– poly(dT) treated with [^{13}C]MeI or DMS	turnover = ~ 0.005 min^{-1}	HPLC/scintillation	(39)
– 24-mer, synthesized with single m3T	turnover = ~ 0.005 min^{-1}	primer extension	(41)
– ss vector from 16-mer, synthesized, single m3T	bypass: hardly removed mutagenesis: 60 % → 40 %	CRAB/REAP (*in vivo*)	(5)
1,*N*^{6}-ethenoadenine (εA); 3,*N*^{4}-ethenocytosine (εC)			
* *– trimer (T-εA-T) synthesized, single εA*	*59.7 ± 14.2 0.13 ± 0.05 0.002*	*HPLC*	*this study*
– trimer (T-εA-T) synthesized, single εA	turnover = ~ 0.5 min^{-1}	HPLC	(7)
– 11-mer (T$_5$-εA-T$_5$) synthesized, single εA	turnover = ~ 0.7 min^{-1}	HPLC	(7)
– 16-mer, synthesized, single εA	turnover = ~ 0.02 min^{-1}	MALDI-TOF	(8)
– ss vector from 16-mer, synthesized, single εA	bypass: 9 % → 100 % mutagenesis: 37 % → 3 %	CRAB/REAP (*in vivo*)	(8, 42)
– ss vector from 16-mer, synthesized, single εC	bypass: 15 % → 25 % mutagenesis: 82 % → 37 %	CRAB/REAP (*in vivo*)	(8)
1,*N*^{6}-ethanoadenine (EA)			
– 16-mer, synthesized with single EA	??	MALDI-TOF	(42)
– ss vector from 16-mer, synthesized, single EA	bypass: 14 % → 100 % mutagenesis: 4 % → 2 %	CRAB/REAP (*in vivo*)	(42)
larger alkyls			
– poly(dA) treated with [^{14}C]EtI^{m}	??	HPLC/scintillation	(28)
– ss vector from 16-mer, synthesized, single ethyl-3C	bypass: 8 % → 98 %	CRAB/REAP	(5)

	mutagenesis: 30 % → 0 %	(*in vivo*)	
– M13 phage treated with EtI, PrI[n], iodoethanol, or propylene oxide	ethyl, propyl: +++ hydroxyalkyl: +	survival in *E. coli* ΔAlkB	(28, 38)

---

* denotes results from this study

[a] MeI, methyl iodide

[b] MNU, *N*-methyl-*N*-nitrosourea

[c] DMS, dimethylsulfate

[d] MMS, methyl methane sulfonate

[e] HPLC/scintillation, substrate and product nucleotides separated by reversed-phase HPLC and quantified by their absorbance peaks on the resulting chromatogram (HPLC) or by scintillation counting (HPLC/scintillation)

[f] MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight mass spectrometry measured loss of mass in repaired oligonuclotides

[g] GC-MS, pentafluorobenzyl hydroxylamine trapping reagent volatilized products released from DNA which were analyzed by gas chromatography-mass spectrometry (8)

[h] FDH-coupled, formaldehyde dehydrogenase-coupled reaction converts the formaldehyde produced by AlkB to fluorescent APADH (6)

[I] CRAB, Competitive Replication of Adduct Bypass measures the ability of DNA polymerase to traverse a lesion in *E.coli in vivo* (8)

[j] REAP, Restriction Endonuclease and Postlabeling Analysis of Mutation Frequency *in vivo* (8)

[k] ITC, isothermal titration calorimetry

[l] 1-[^{14}C]2OG, 1-[^{14}C]-2-oxoglutarate turnover by AlkB detected by measuring the release of $^{14}CO_2$ (37)

[m] EtI, ethyl iodide

[n] PrI, propyl iodide

Table S2. Data collection and refinement statistics.

Crystallization Preparation:				
Co-crystallization Ligands	Fe(II),2OG,TmAT	Mn(II),2OG,TmCT	Fe(II),2OG,TmCT	Mn(II),2OG,TεAT
Growth environment	Anaerobic	Aerobic	Anaerobic	Aerobic
Crystal Parameters:				
Space group	$P4_3$	$P4_3$	$P4_3$	P1
Cell dimensions:				
a, b, c (Å)	40.6, 40.6, 117.9	41.4, 41.4, 116.7	41.4, 41.4, 117.2	39.7, 39.2, 72.4
α, β, γ, (º)	90, 90, 90	90, 90, 90	90, 90, 90	77.2, 74.8, 63.6
Data Quality:				
Wavelength (Å)	0.97947	0.97947	0.97947	0.97947
Resolution (Å)	50.0-1.7 (1.73-1.70)	50.0-1.50 (1.53-1.50)	50.0-1.6 (1.63-1.60)	50.0-1.4 (1.42-1.40)
R_{sym} (%)	16.5 (0.0)	6.0 (43.8)	7.4 (65.6)	4.7 (59.0)
No. reflections	20972	31383	25984	71196
Mean Redundancy	8.45	3.73	7.89	3.96
Completeness, overall (%)	100.0 (100.0)	99.8 (99.9)	99.9 (100.0)	96.1 (93.6)
Completeness, $I \geq 2\sigma_I$ (%)	85.1 (58.9)	89.2 (72.2)	86.1 (59.2)	83.2 (52.5)
Mean I/σ_I	23.3 (3.2)	28.8 (5.2)	29.2 (3.4)	35.7 (3.6)
Refinement Residuals ($f \geq 2\sigma_f$):				
R_{free} (%)	21.5	21.4	20.9	20.8
R_{work} (%)	19.9	19.9	19.3	20.4
Model Quality:				
Rmsd bond lengths (Å)	0.011	0.005	0.018	0.012
Rmsd bond angles (º)	1.5	1.3	2.2	1.5
Ramachandran Plot:				
Core (%)	92.7	90.9	92.1	91.3
Allowed (%)	6.7	8.5	7.3	8.1
Disallowed (%)	0.6	0.6	0.6	0.6
B-factor (Å$_2$):	14.4	13.0	17.4	16.1
Model Contents:				
Protomers in asymmetric unit	1	1	1	2
Protein residues	15-213	15-214	15-214	A: 13-215 B: 12-214
Ligands	Fe(II),2OG,TmAT	Mn(II),2OG,TmCT	Fe(II),2OG,TmCT	Mn(II),2OG
No. atoms:				
Protein	1556	1565	1580	3232
Metal Ion	1	1	1	2
Ligands: 2OG	10	10	10	20
TmAT	63			
TmCT		61	61	
Water molecules	156	148	281	346
PDB accession code:	3I2O	3I3M	3I49	3I3Q

Standard definitions were used for all parameters (43). The abbreviations 2OG, TmAT, TmCT, TεAT, and Rmsd represent 2-oxoglutarate, dT-(1-methyl-dA)-dT, dT-(3-methyl-dC)-dT, dT-(1,*N*^{6}-etheno-dA)-

dT, and "root mean square deviation", respectively. No evidence of bound nucleotide was observed in electron density maps during refinement of crystals grown in the presence of TεAT. Entries in parentheses report data from the limiting resolution shell. All observations with $I \geq -3\sigma_I$ were merged and included in calculation of R_{sym}. Data collection and refinement statistics come from SCALEPACK (33) and CNS (9), respectively.

Table S3. Summary of *E. coli* AlkB crystal structures.

			Crystal parameters			Conformation			
PDB ID	Ligands	Crystallization conditions	Space group	Dimensions a, b, c (Å) α β γ°	Resolution (Å)	W178 / F156 Sidechains	O_2-binding cavity	O_2-diffusion tunnel	Reference
3I2O	Fe(II) 2OG TmAT	anaerobic	$P4_3$	40.6, 40.6, 117.9 90, 90, 90	1.7	flipped-out	open	open	*this study*
2FD8	Fe(II) 2OG TmAT	anaerobic	$P4_3$	40.7, 40.7, 118.3 90, 90, 90	2.3	flipped-out	open	open	(15)
3I49	Fe(II) 2OG TmCT	anaerobic	$P4_3$	41.4, 41.4, 117.2 90, 90, 90	1.6	flipped-in	closed	closed	*this study*
3I3M	Mn(II) 2OG TmCT	aerobic	$P4_3$	41.4, 41.4, 116.7 90, 90, 90	1.5	flipped-in	closed	closed	*this study*
3I3Q	Mn(II) 2OG --	aerobic	P1[a]	39.7, 93.2, 72.4 77.2, 74.8, 63.6	1.4	flipped-in	closed	closed	*this study*
2FDH	Mn(II) 2OG TmAT	aerobic	$P4_3$	40.7, 40.7, 118.3 90, 90, 90	2.1	flipped-out	open	open	(15)
2FDF	Co(II) 2OG TmAT	aerobic	$P4_3$	41.3, 41.3, 116.8 90, 90, 90	2.1	flipped-in	closed	closed	(15)
2FDG	Fe(II) sin TmAT	anaerobic	$P4_3$	40.5, 40.5, 117.4 90, 90, 90	2.2	flipped-out	open	partial[b]	(15)
2FDJ	Fe(II) sin --	anaerobic	P1	36.8, 38.7, 40.2 75.8, 75.0, 66.1	2.1	flipped-out	open	open	(15)
2FDI	Fe(II) 2OG TmAT	anaerobic (air, 3 hours)[c]	$P4_3$	40.7, 40.7, 117.9 90, 90, 90	2.1	flipped-out	open	open	(15)
2FDK	Fe(II) 2OG TmAT	anaerobic (air, 9 days)[c]	$P4_3$	40.6, 40.6, 117.0 90, 90, 90	2.1	flipped-out	open	partial[d]	(15)
3BKZ	Mn(II) 2OG DNA1[e]	aerobic	$P2_1$	41.4, 75.5, 46.8 90, 105.4, 90	1.6	flipped-in	closed	closed	(10)
3BI3	Mn(II) 2OG DNA5[e]	aerobic	$P2_1$	41.4, 75.6, 51.0 90, 107.7, 90	1.8	flipped-in	closed	closed	(10)
3BIE	Mn(II) 2OG DNA4[e]	aerobic	$P2_1$	41.4, 75.8, 52.2 90, 108.5, 90	1.6	flipped-in	closed	closed	(10)

---

[a] This crystal form had two protomers in the asymmetric unit. All other crystal forms had only one.

[b] A slight (~10°) rotation of Chi1 in L184 narrows the diffusion tunnel slightly such that it appears closed in surface representations. However, the movement of the L184 sidechain is less than in the structures with "closed" conformations and the backbone is similar to the "open" conformations.

[c] Crystals were exposed to air for the length of time indicated prior to data collection.

[d] The tunnel appears closed in surface representations because the sidechain of K134 extends across the tunnel opening. However, the electron density for this sidechain is absent in this structure. In all other structures, K134 is modeled in the opposite direction.

[e] Disulfide-tethered cytosine (C*) in double-stranded DNA substrates were cross-linked to the protein. DNA1, TAGGTAAC*ATCGT; DNA4, TAGGTAA(m1A)AC*CGT; DNA5, TAGGTAA(m1A)AC*CGT.

Table S4. Distances between the metal cofactor and ligating protein atoms.

		Distance (Å) from metal									
			protein		2-oxoglutarate		succinate				
PDB ID	Ligands	H131	D133	H187	2OG	2OG	sin	H_2O [a]	H_2O	Conformation	Reference
3I2O	Fe(II) 2OG TmAT	2.17	2.09	2.14	2.07	2.28			2.42	open	*this study*
2FD8	Fe(II) 2OG TmAT	2.29	2.41	2.19	2.10	2.42			2.90	open	(15)
3I49	Fe(II) 2OG TmCT	2.24	2.14	2.18	1.99	2.22			2.23	closed	*this study*
313M	Mn(II) 2OG TmCT	2.20	2.14	2.14	2.12	2.27			2.25	closed	*this study*
313Q A[b]	Mn(II) 2OG --	2.23	2.17	2.15	2.11	2.33			2.19	closed	*this study*
313Q B[b]	Mn(II) 2OG --	2.23	2.20	2.16	2.15	2.34			2.21	closed	*this study*
2FDH	Mn(II) 2OG TmAT	2.06	2.14	2.25	2.13	2.36			2.80	open	(15)
2FDF	Co(II) 2OG TmAT	2.05	2.28	2.11	2.16	2.27			2.73	closed	(15)
2FDG	Fe(II) sin TmAT	2.25	2.07	2.19			1.98	2.09	3.26	open	(15)
2FDJ	Fe(II) sin --	2.16	2.30	2.29			2.42	2.12	2.31	open	(15)
2FDI	Fe(II) 2OG TmAT (air, 3 hours)[c]	2.17	2.15	2.25	2.14	2.19	2.20	1.87	3.01	open	(15)
2FDK	Fe(II) 2OG TmAT (air, 9 days)[c]	2.02	2.36	2.27	2.12	2.36	2.26	2.46	2.67	open	(15)
3BKZ	Mn(II) 2OG DNA1 (xlink)[d]	2.16	2.22	2.17	2.30	2.32			2.30	closed	(10)
3BI3	Mn(II) 2OG DNA5 (xlink)	2.17	2.18	2.12	2.16	2.47			2.47	closed	(10)
3BIE	Mn(II) 2OG DNA4 (xlink)	2.18	2.18	2.17	2.14	2.42			2.41	closed	(10)

[a] In the presence of succinate (sin), a second water molecule fills the space previously occupied by the carboxylate of 2OG.

[b] This crystal form had two protomers in the asymmetric unit. Distances are given for both monomers A and B.

[c] Anaerobic crystals were exposed to air for the length of time indicated prior to data collection, resulting in partial turnover of 2OG to succinate.

[d] Disulfide-tethered cytosine (C*) in dsDNA substrates were cross-linked to the protein. DNA1, TAGGTAAC*ATCGT; DNA4, TAGGTAA(m1A)AC*CGT; DNA5, TAGGTAA(m1A)AC*CGT.

Supporting Figure S1

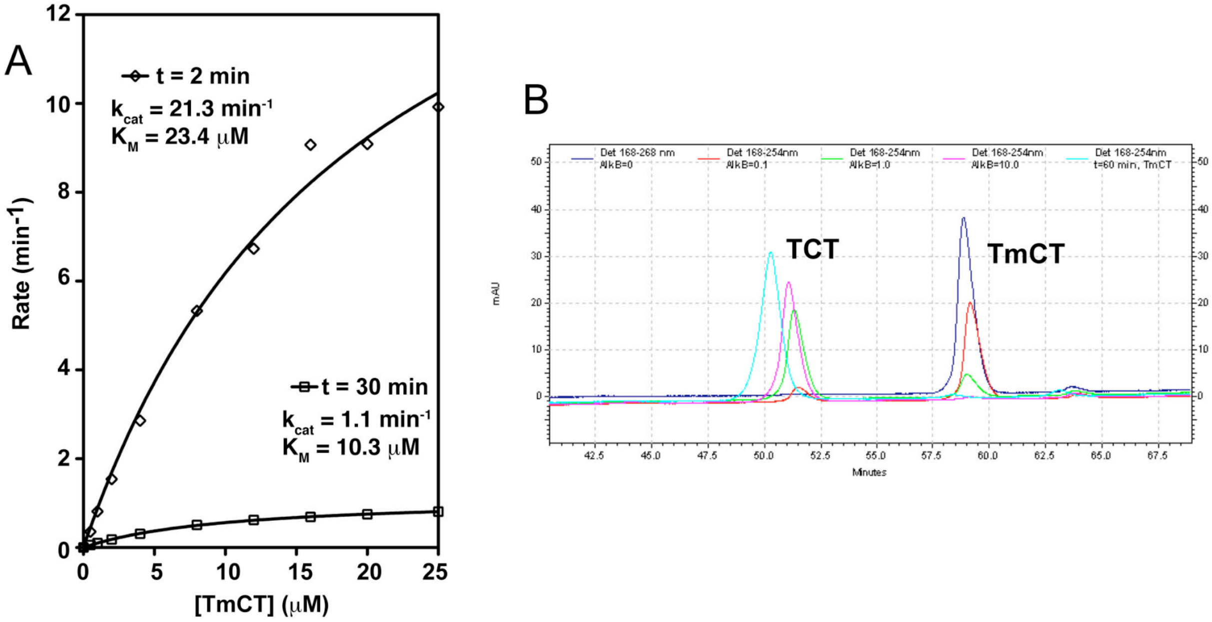


Figure S1. ***Repair of TmCT by E. coli AlkB-ΔN11.*** **(A)** Kinetic analysis of repair of TmCT by AlkB-ΔN11. The enzyme (0.1 μM) was incubated with TmCT (0-25 μM) at 37 °C for (◇) 2 or (☐) 30 minutes under standard aerobic reaction conditions containing saturating amounts of Fe(II) (75 μM) and excess 2OG (1 mM). Reactions quenched by addition of EDTA were analyzed by reversed-phase HPLC, and the dealkylation rates were determined by quantifying substrate and product peak areas in the resulting chromatograms. The curves represent non-linear regressions to the non-cooperative Michaelis-Menten equation. Depletion of the alkylated DNA substrate leads to systematic underestimation of the initial enzyme velocity in the reactions incubated for 30 minutes. Therefore, care was taken to use only reactions with less than 20% substrate-depletion for all Michaelis-Menten analyses presented in this paper. **(B)** Reversed-phase HPLC chromatograms of TmCT repair reactions. The trimer substrate TmCT (2 μM) was incubated with 0 (dark blue), 0.1 (red), 1.0 (green), or 10.0 (magenta) μM AlkB at 37 °C for 2 minutes and with 10.0 μM AlkB at 37 °C for 2 (magneta) or 60 (cyan) minutes. Reactions quenched by addition of EDTA were analyzed by reversed-phase HPLC. Complete repair of the TmCT lesion was observed with 10 μM AlkB.

Supporting Figure S2

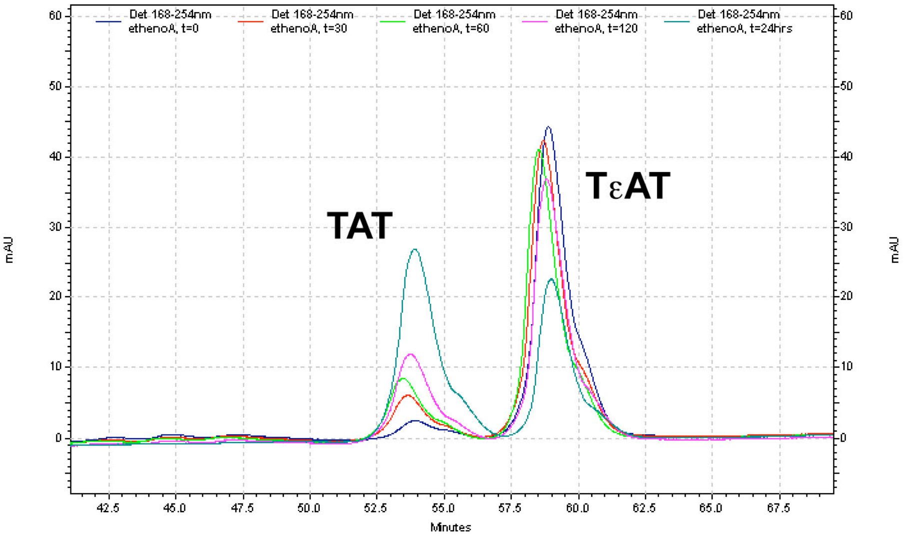


Figure S2. ***Reversed-phase HPLC chromatograms of TεAT repair reactions.*** *E. coli* AlkB-ΔN11 (1 μM) was incubated with TεAT (10 μM) at 37 °C for 0 (dark blue), 30 (red), 60 (green), or 120 (magenta) minutes or 24 hours (light blue) under standard aerobic reaction conditions containing saturating amounts of Fe(II) (75 μM) and excess 2OG (1 mM). Reactions quenched by addition of EDTA were analyzed by reversed-phase HPLC. Even after 24 hours of incubation, only 70% of the substrate was converted to product. There was no appearance of extra peaks that could correspond to reaction intermediates.

Supporting Figure S3

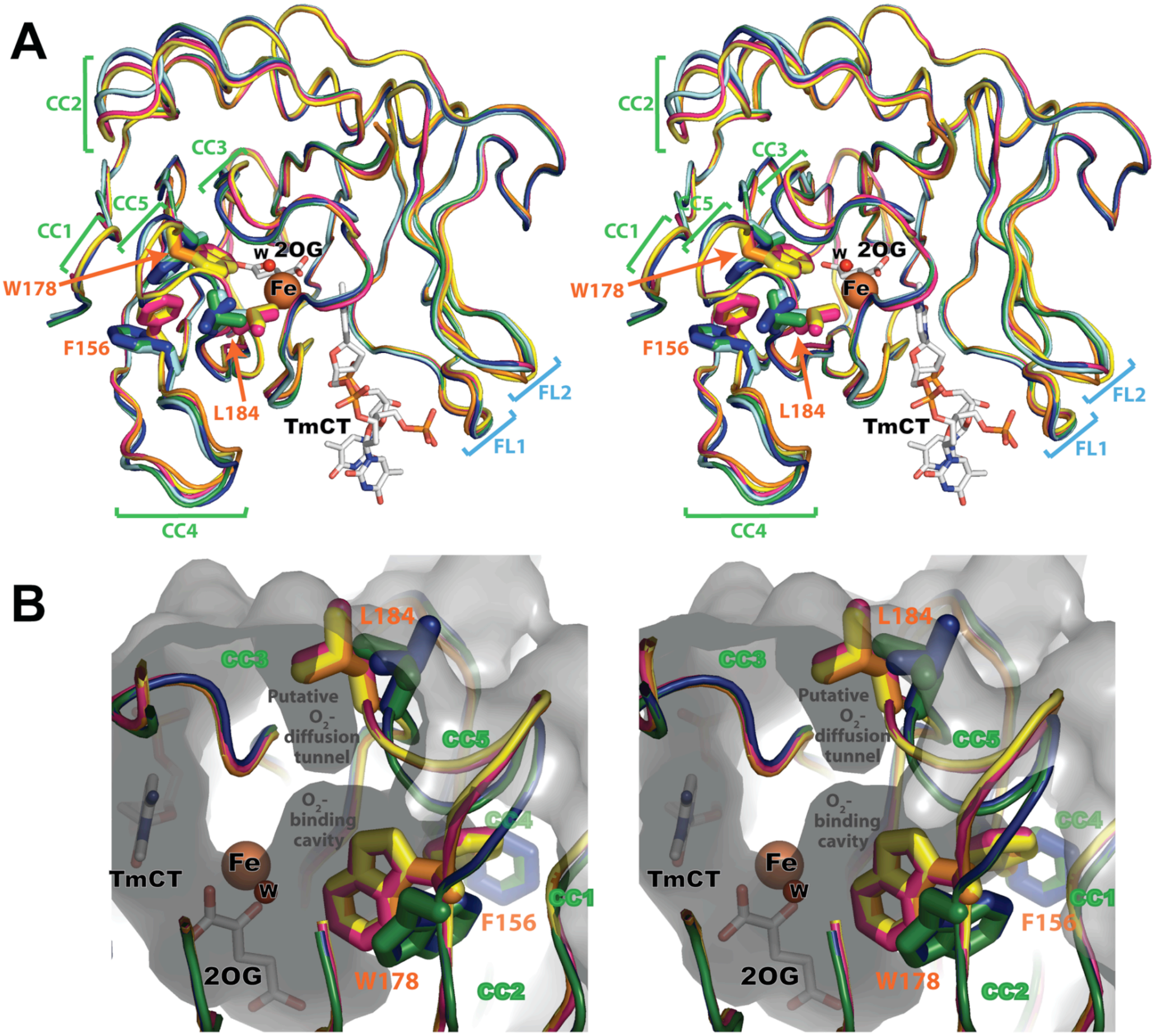
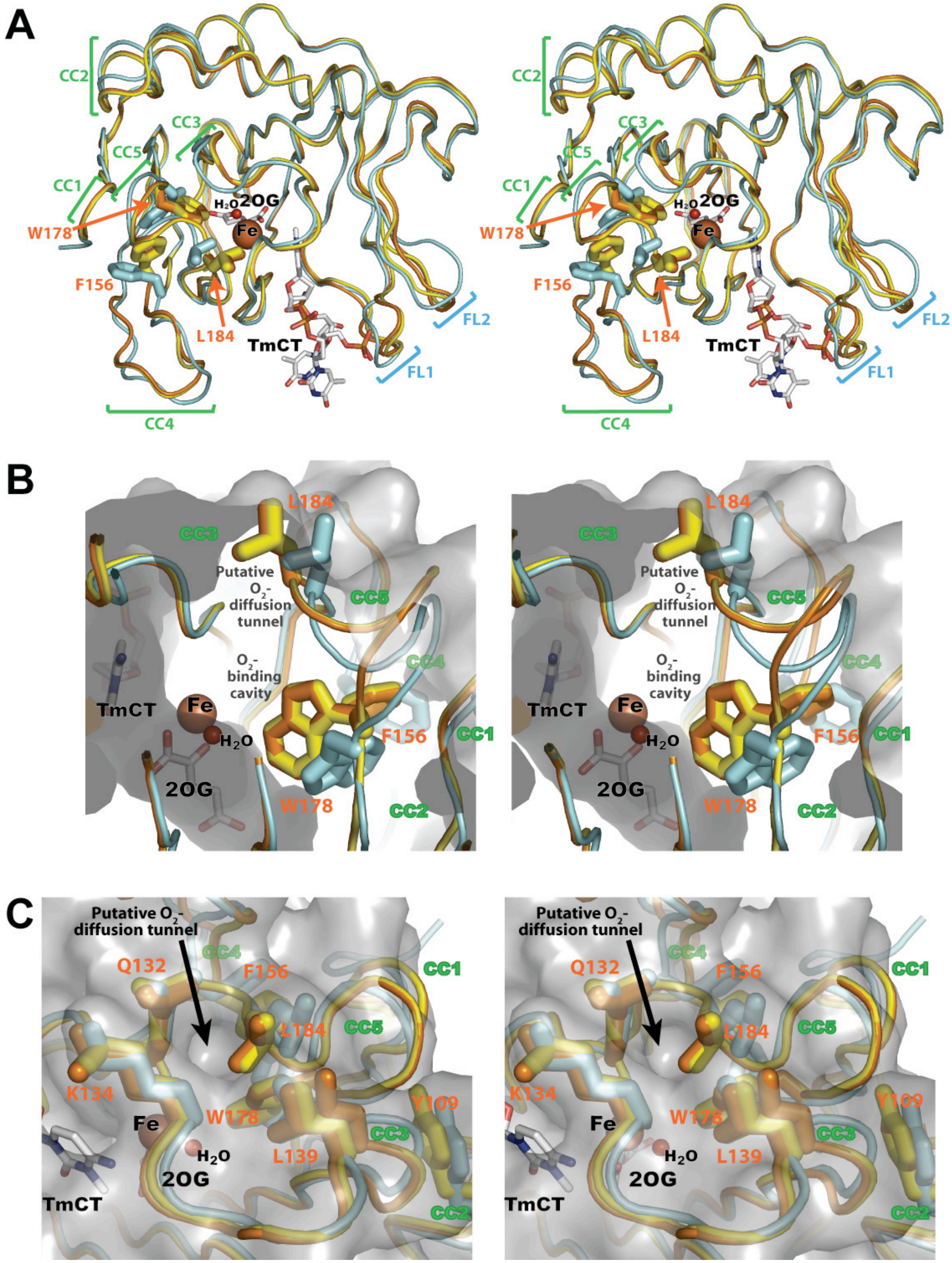


Figure S3. ***The ensemble of AlkB-ΔN11 crystal structures reveals a discrete change in global protein conformation.*** (This panel presents stereopairs of Figs. 3A-B in the main text.) **(A)** Nucleotide-bound structures after least-squares superposition of core β-strands in the Fe-2OG dioxygenase domain. (2OG, Fe(II), and TmCT ligands are shown from only a single structure.) Discrete local conformational changes are observed in five protein segments (CC1-CC5). Flexible loops (FL1 & FL2), both of which contact the backbone of the trinucleotide substrate, exhibit elevated backbone B-factors as well as conformational variability not correlated with the global conformational state of the enzyme (15). A water molecule, which is replaced by O_2 to initiate the oxidation reaction, occupies the sixth site on the octahedrally coordinated Fe ion. AlkB-Fe(II)-2OG-TmCT is shown in orange, AlkB-Mn(II)-2OG-TmCT in yellow, AlkB-Fe(II)-2OG-TmAT (at 1.7 Å) in green, AlkB-Mn(II)-2OG-TmAT (15) in blue, AlkB-Co(II)-2OG-TmAT (15) in magenta, and AlkB-Fe(II)-succinate-TmAT (15) in cyan. **(B)** The putative O_2-diffusion channel and O_2-binding cavity with structures colored as in panel A. The molecular surface of AlkB-Fe(II)-2OG-TmAT is shown in grey.

Supporting Figure S4



Supporting Figure S4 (continued)

Figure S4. ***Open and closed cavity conformations of E. coli AlkB-ΔN11 in nucleotide-free structures.*** **(A)** Stereopair showing nucleotide-free crystal structures after least-squares superposition of core β-strands in the Fe-2OG dioxygenase domain. Flexible loops (FL) and conformational changes (CC) are labeled according to Fig. 3 in the main text. For simplicity, only the Fe(II), 2OG and TmCT ligands from the AlkB-Fe(II)-2OG-TmCT structure are shown. A water molecule, which is replaced by O_2 to initiate the oxidation reaction, occupies the sixth site on the octahedrally coordinated Fe ion. AlkB-Fe(II)-succinate (15) is shown in cyan, AlkB-Mn(II)-2OG (subunit A from asymmetric unit) in orange, and AlkB-Mn(II)-2OG (subunit B) in yellow. **(B)** Stereopair showing the putative O_2-diffusion channel and O_2-binding cavity with structures colored according to panel A. The molecular surface of the AlkB-Mn(II)-2OG (subunit B) structure is shown in grey. The rotation of W178 towards the active site in the closed structure fills the putative O_2-binding cavity, while the coupled rotation of L184 closes the O_2-diffusion tunnel. **(C)** Stereopair showing the molecular surface of AlkB in the open conformation, with structures colored according to panel A. Panel B is rotated ~90° downward compared to panel A. In the structure of AlkB-Fe(II)-2OG-TmAT (PDB id 2FDK, Table S3), the sidechain of K134 appears to extends towards the right and block access to the O_2-diffusion tunnel. However, there is no electron density for this sidechain, so it does not stably adopt this conformation. In all other crystal structures, K134 adopts the conformation shown here.

Supporting Figure S5

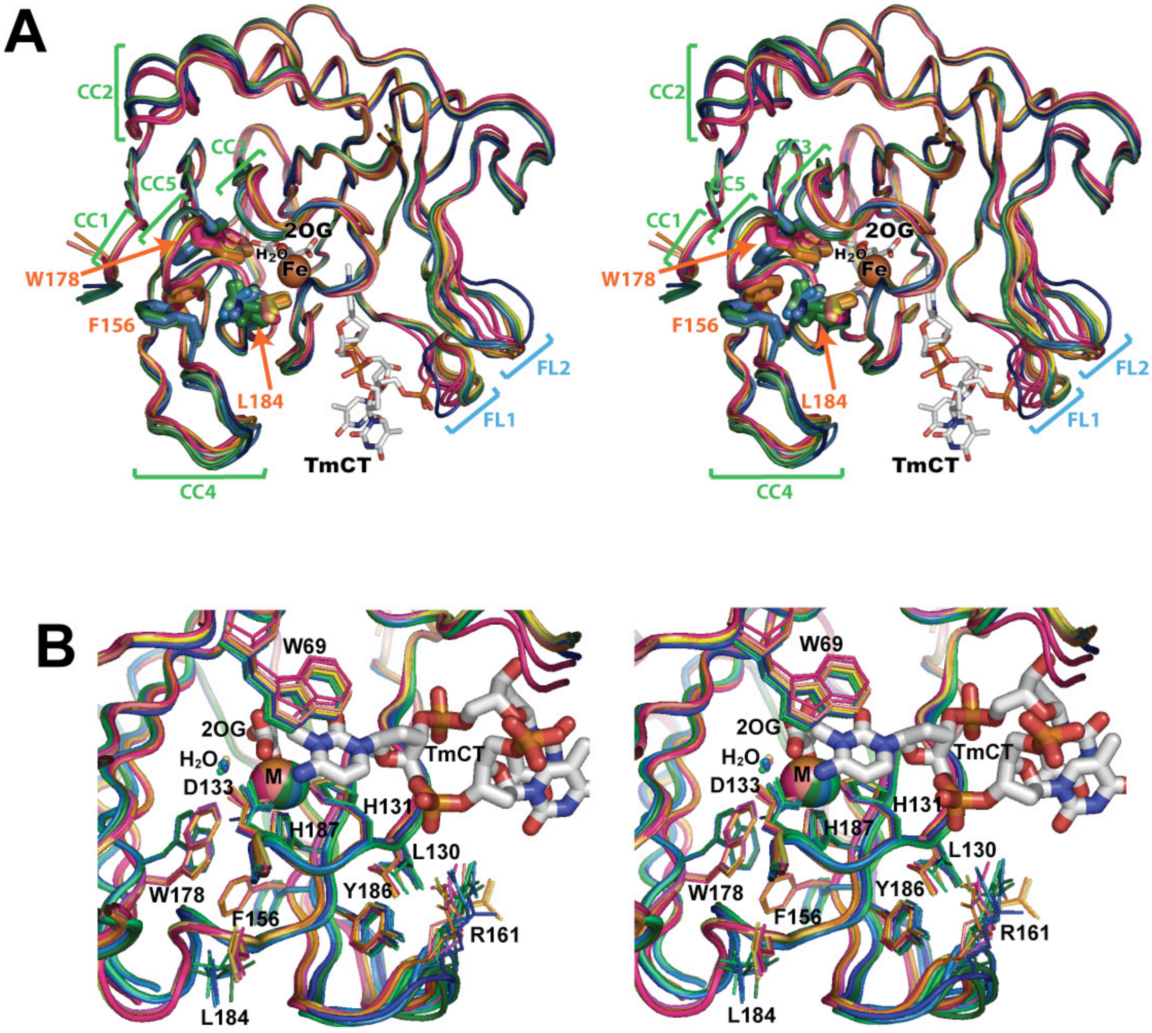


Figure S5. ***Open and closed conformations of E. coli AlkB-ΔN11.*** Stereopair showing all AlkB crystal structures in the PDB (10, 15) (see Table S3). Superposition of the protein backbone was based on core β-strands in the Fe-2OG dioxygenase domain. **(A)** Crystal structures of AlkB are found in two distinct global conformations. Flexible loops (FL) and conformational changes (CC) are labeled according to Fig. 3 in main text. Structures in the "open" conformation are shown in various shades of blue and green. Structures in the "closed" conformation are shown in various shades of pink, orange, and yellow. (See Fig. S6 for color legend.) For simplicity, the 2OG and TmCT ligands are shown from only the AlkB-Fe(II)-2OG-TmCT structure. A water molecule, which is replaced by O_2 to initiate the oxidation reaction, occupies the sixth site on the octahedrally coordinated Fe ion. **(B)** The structural alignment in panel A is rotated to show the conformational changes in the active site and the inter-residue interactions between R161 (in CC4) and Y186 and L130, which are adjacent to conserved Fe-ligating residues H187 and H131, respectively. The guanidino group of R161 contacts the nucleotide backbone in refined structures containing either the alkylated trinucleotide substrate (15) or the DNA duplex (not shown).

Supporting Figure S6

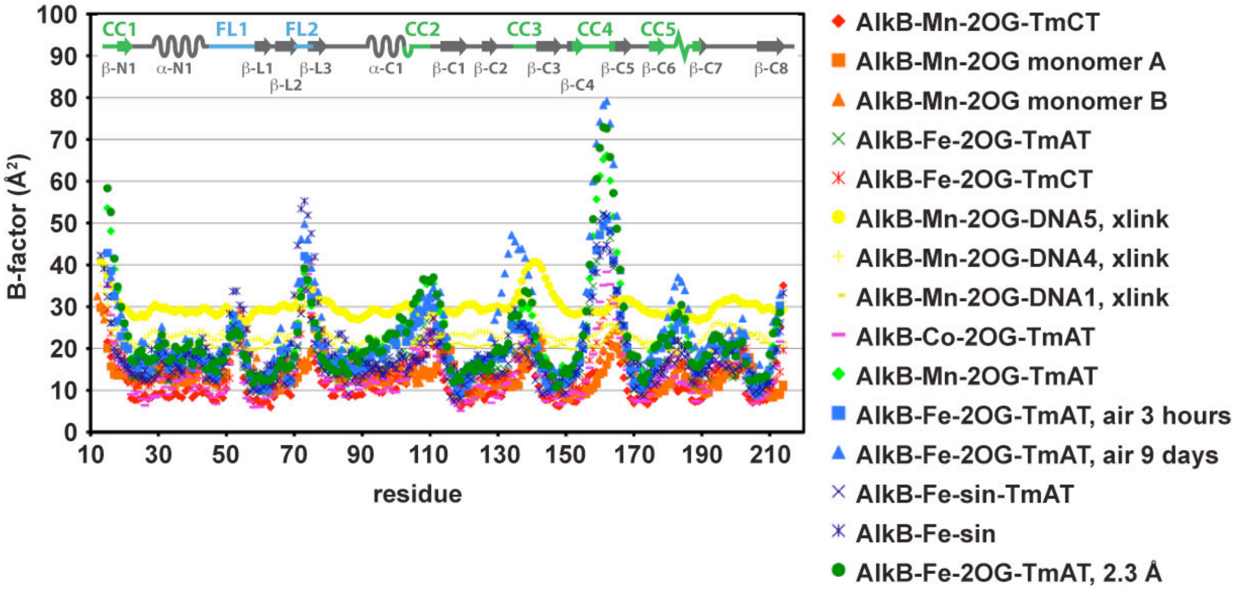
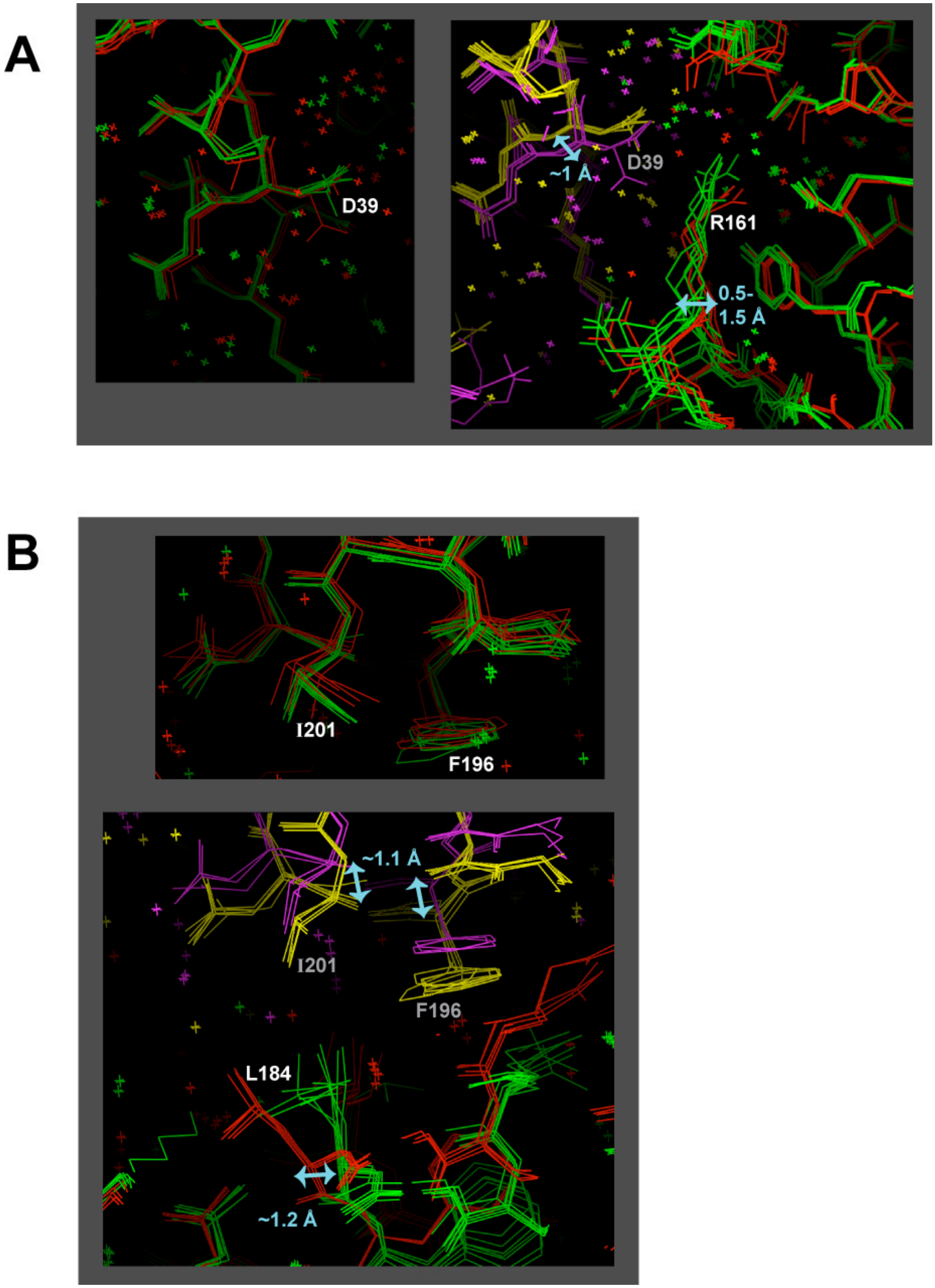


Figure S6. Mean backbone B-factors in refined AlkB-ΔN11 crystal structures. Structures in the closed conformation are shown in red, orange, yellow, and magenta, with the DNA-crosslinked (xlink) structures (10) in yellow. Structures in the open conformations are shown in green and blue. For reference, a secondary structural schematic is shown at the top of the graph indicating the locations of the flexible loops (FLs) and conformational changes (CCs).

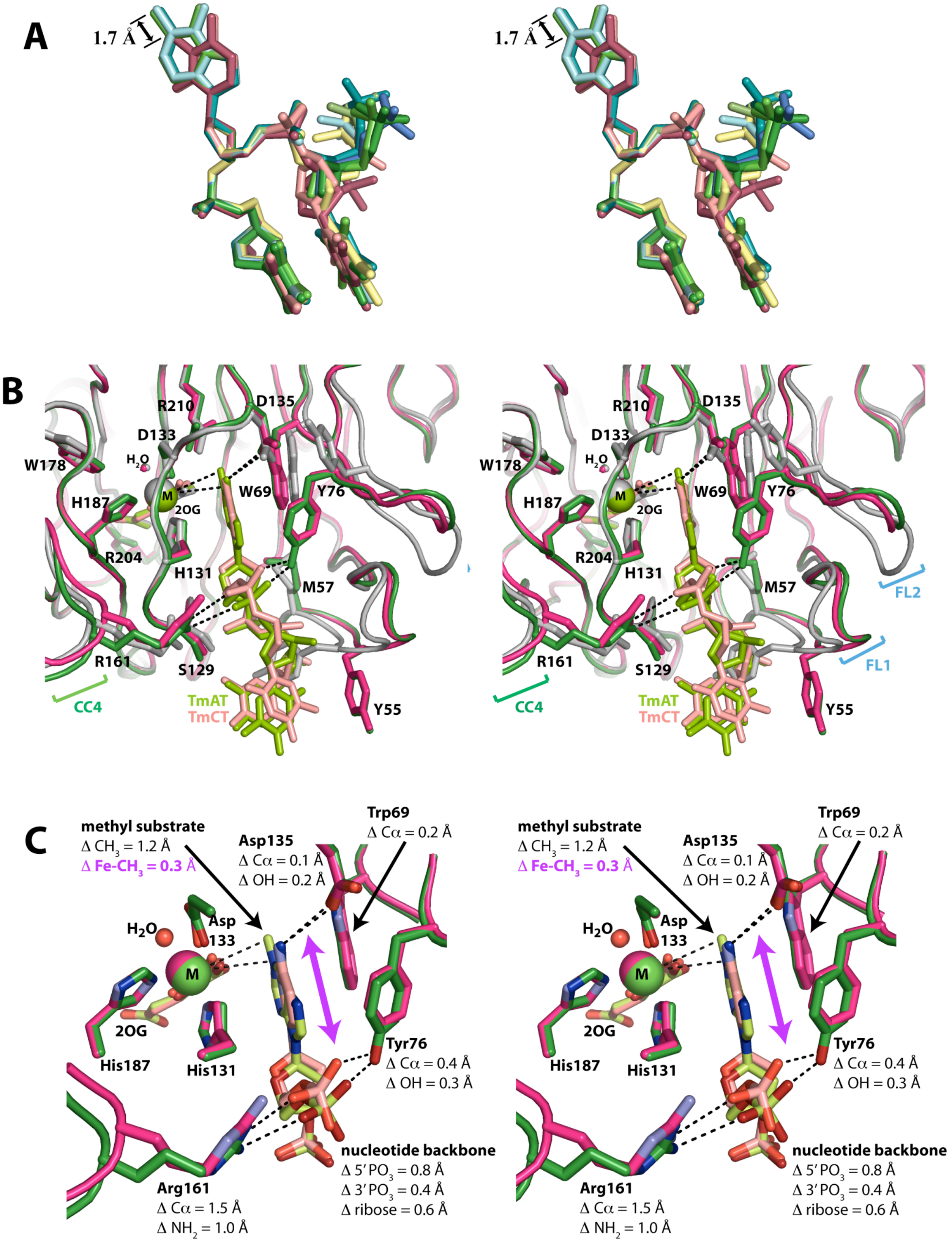
Supporting Figure S7



Supporting Figure S7 (continued)

Figure S7. ***Crystal packing of E. coli AlkB-ΔN11 in spacegroup $P4_3$.*** All trinucleotide-bound crystals of AlkB-ΔN11 were in spacegroup $P4_3$ with only one molecule per asymmetric unit. The open and closed conformations of the protomer in one unit cell are shown in green and red, respectively. The molecules in the adjacent unit cell are colored yellow for the open (green) structures and magenta for the closed (red) structures. **(A)** Closer crystal-packing contacts in the closed conformation. The Cα atom of residue D39 is ~1 Å closer to R161, which is part of CC4, in crystals of AlkB in the closed conformation (right). This change in intermolecular interactions results primarily from changes in lattice-packing interactions because D39 Cα atoms vary in position by < 0.5 Å when individual subunits are directly superimposed (left) (Fig. 3C in the main text). **(B)** Closer crystal-packing contacts in the open conformation. The Cα atoms of I201 and F196 are ~1 Å closer to L184 in crystals of AlkB in the open conformation (bottom). This change in intermolecular interactions also results primarily from changes in lattice-packing interactions because the corresponding Cα atoms vary in position by < 0.5 Å when individual subunits are directly superimposed (top) (Fig. 3C in the main text).

Supporting Figure S8



Supporting Figure S8 (continued)

Figure S8. ***Stereochemical comparison of 1-methyadenosine vs. 3-methylcytosine-bound structures.*** (Panels B and C present stereopairs of Figs. 4B-C in the main text) **(A)** Crystal structures of the alkylated trinucleotide substrates (colored as in Fig. S3) aligned based on least-squares superposition of the ribose ring in the central methylated base. The flanking thymine nucleotides show very similar conformations except for variations in the orientation of the ribose ring and phosphate group on the 5’ terminal thymine. This phosphate group has weak or absent electron density indicating local disorder. **(B)** Active site stereochemistry in the Co(II)-2OG-TmAT (green), Fe(II)-2OG-TmCT (magenta), and Mn(II)-2OG (grey) structures. Exclusively structures in the closed conformation are shown to avoid interference from the global conformational change opening the O_2-diffusion tunnel and thereby isolate conformational affects attributable to the identity of the bound nucleotide substrate. Invariant residues H131, D133, and H187 chelate the metal ion cofactor (M), while conserved residues R204 and R210 interact with the 2OG co-substrate. The partially disordered phosphates on the 5’ terminal thymine are omitted for clarity, as are water molecules other than that occupying the sixth coordination site on the metal ion. **(C)** Magnified view of the methylated DNA bases in the Co(II)-2OG-TmAT (green) and Fe(II)-2OG-TmCT (magenta) structures. Nitrogen and oxygen atoms are colored blue and red, respectively, while carbon atoms are colored like the protein backbone (identical to panel B). Only the methylated bases and flanking phosphates of the nucleotide substrates are shown. The purple arrow is aligned approximately parallel to the 0.6 Å displacement of the ribose group of the alkylated nucleotide in the m3C *vs.* m1A structures. This displacement effectively slides the alkylated base deeper into a binding slot that is roughly co-planar with the sidechain rings of W69 and H131 and terminates at the catalytic iron center.

Supporting Figure S9

]

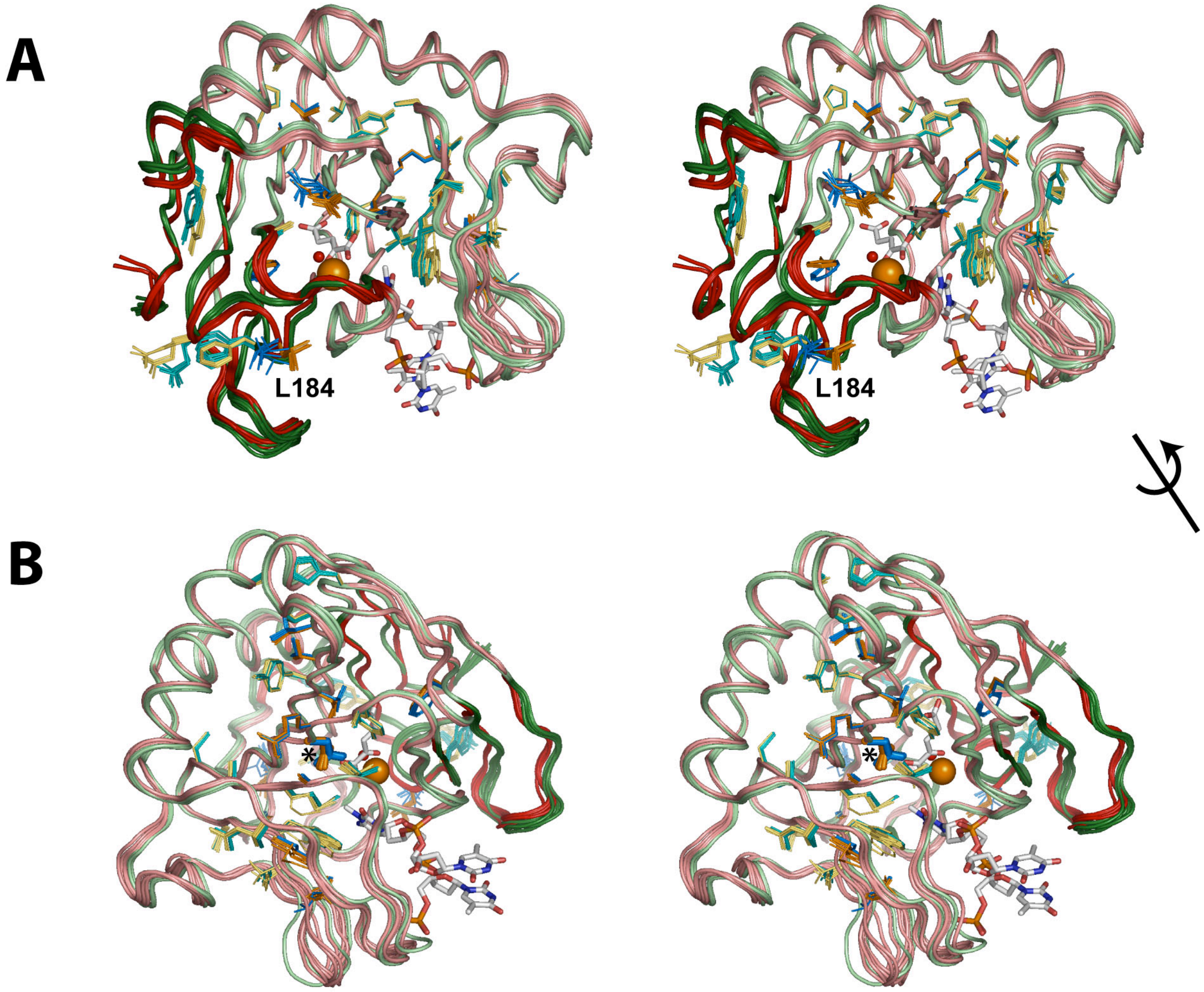


Figure S9. ***Sterochemical comparison of conformational changes in the crystal lattice compared to those observed in solution NMR experiments.*** **(A)** All crystal structures of *E. coli* AlkB in the PDB were aligned by least-squares superposition of core β-strands in the Fe-2OG dioxygenase domain. Crystal structures of AlkB structures are found in two distinct global conformations (Table S3). The backbone traces of structures in the "open" conformation are shown in green, while structures in the "closed" conformation are shown in red. The backbone of residues 14-22, 104-111, 136-141, 155-166, and 175-188, which exhibit the greatest conformational change from open to closed structures are colored dark green and dark red, respectively. NMR spectroscopy has documented a conformational change in AlkB between 2OG-bound and succinate-bound structures in solution (22). Residues M49, A60, T71, N120, R121, V146, L148, F154, G157, L184, N206, Q211, and A212 exhibit broadened resonances in the AlkB-Fe(II)-succinate complex, while residues A30, R47, V59, M61, T62, C64, L67, G68, W69, T70, Y109, Y122, A123, G125, A141, G149, H172, G180, E181, F185, G188, G195, T200, Y205, and G213 exhibit chemical shift changes in this complex compared to the AlkB-Fe(II)-2OG complex. The sidechains of residues with broadened resonances are shown in blue

Supporting Figure S9 (continued)

and orange in the open and closed crystal structures, respectively, while the sidechains of residues with different chemical shifts are shown in yellow and teal for the open and closed crystal structures, respectively. L184 gates opening and closing of the putative O_2-diffusion tunnel (Fig. S4B-C) and exhibits chemical exchange broadening in the succinate-bound but not 2OG-bound complexes of AlkB in solution (22). For simplicity, only the 2OG and TmCT ligands from the AlkB-Fe(II)-2OG-TmCT structure are shown. A water molecule, which is replaced by O_2 to initiate the oxidation reaction, occupies the sixth site on the octahedrally coordinated Fe ion. **(B)** Stereopair showing an alternative view of the same structural ensemble colored in the same manner but rotated around the axis shown between panels A and B. Mutation of residue N120 (indicated by the asterisk in panel B) to alanine reduced 2OG binding and exhibited similar NMR spectra to the succinate-bound complex (22).

Supporting Figure S10

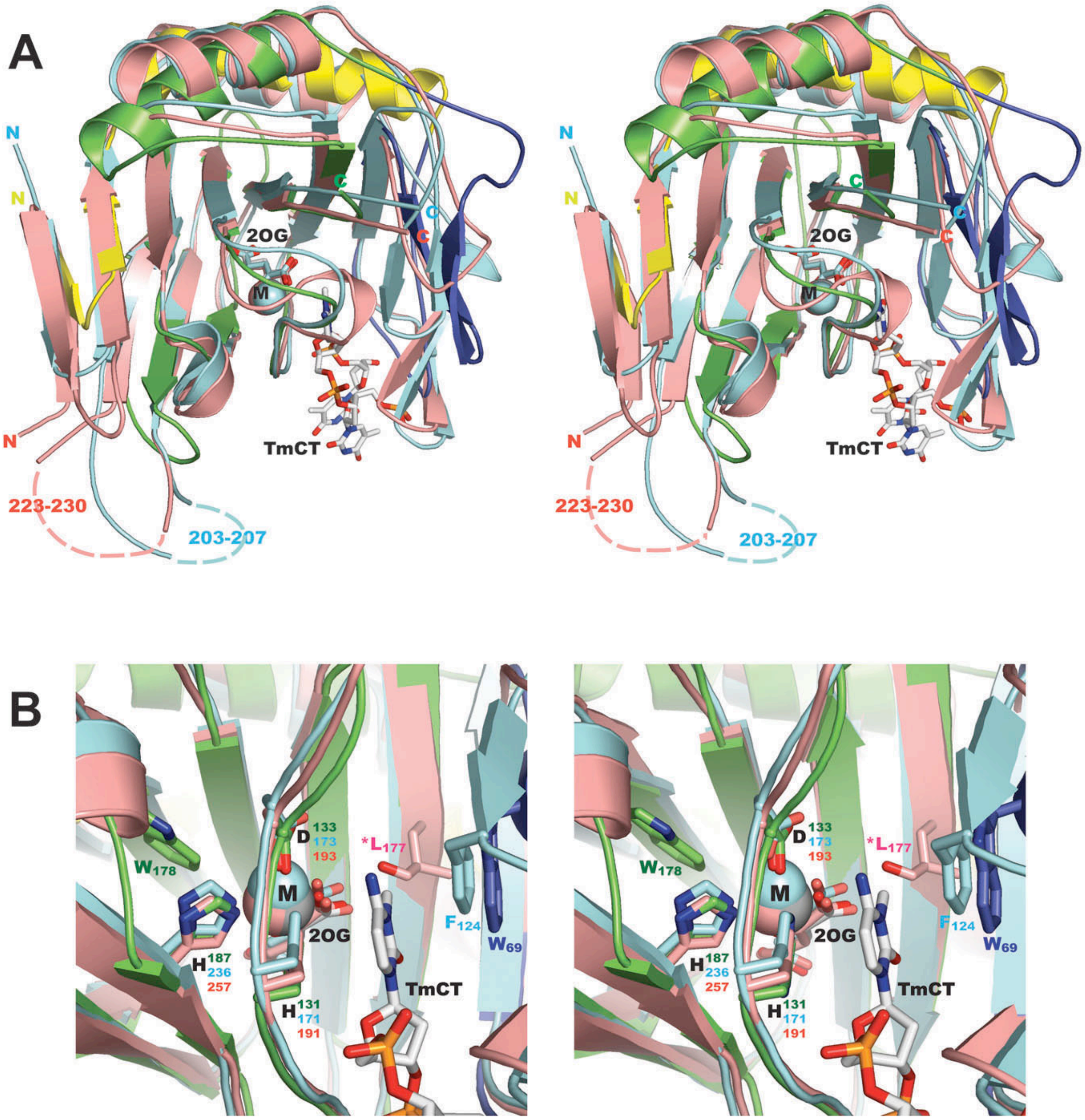


Figure S10. ***Crystal structures of E. coli AlkB-ΔN11, human ABH2-ΔN55, and human ABH3-ΔN69 aligned by least-squares superposition of their catalytic cores.*** **(A)** Superposition of the protein backbone of AlkB, ABH2, and ABH3 was based on the core β-strands in the Fe-2OG dioxygenase domain. The nucleotide-free complex of AlkB with Mn(II) and 2OG is colored according to subdomain organization as in Yu *et al.* (15) (with the N-terminal extension yellow, nucleotide-recognition lid blue, and catalytic core green). The complex of ABH2 with Mn(II), 2OG, and a crosslinked double-stranded DNA substrate is shown in cyan (10). The DNA molecule in this structure has been omitted for clarity. The complex of ABH3 with Fe(II) and 2OG is shown in pink (31). The location of the T-(1-methyl-C)-T trimer (TmCT) substrate from

Supporting Figure S10 (continued)

a nucleotide-bound AlkB structure is shown for reference. **(B)** Close-up view of the active sites in the structural alignment shown in panel A. The conserved metal-ligating residues are shown (H131, D133, and H187 in AlkB; H171, D173, and H236 in ABH2; and H191, D193, and H257 in ABH3). The label "M" indicates the bound metal cofactors. In AlkB, H131 and W69 form a base-stacking interaction with the nucleotide substrate, which helps position the target atom for catalysis. A homologous stacking interaction exists between H171 and F124 in ABH2. Two distinct conformations are observed for W178 in AlkB (see main text). In the absence of nucleotide, AlkB is reported to self-hydroxylate this residue although no evidence of this modification is observed in the nucleotide-free crystal structures. W178 is not conserved in ABH2 or ABH3. In the ABH3 crystal structure, conserved residue L177 (indicated by an asterisk) was observed to be post-translationally oxidized (31).

Supporting Figure S11

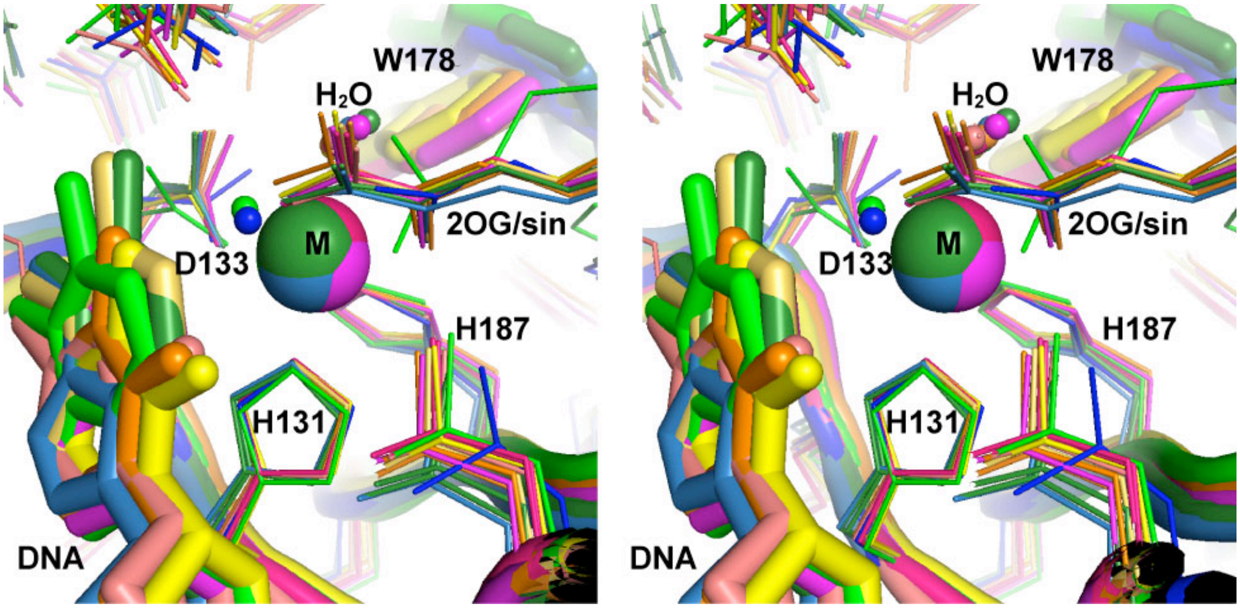


Figure S11. Active site stereochemistry in *E. coli* AlkB-ΔN11. All crystal structures in the PDB were aligned by least-squares superposition of their metal ion cofactor and its ligating atoms from the 2-oxoglutarate (2OG) co-substrate and protein residues H131, D133, and H187. Structures in the "open" conformation are shown in various shades of blue and green. Structures in the "closed" conformation are shown in various shades of pink, orange and yellow. (See Fig. S6 for color legend). In the presence of succinate, a second water molecule is bound in place of the carboxylate of 2OG. This alignment demonstrates that the coordination geometry is preserved for most of the atoms interacting directly with the Fe ion, including their distances to the Fe ion (Table S4), due to coupled backbone and sidechain rotations of the ligating residues. However, the open *vs.* closed structures show sub-angstrom differences in the positions of the interacting atoms on the 2OG co-substrate and the bound water molecule. Some of these geometric differences correlate consistently with the conformational state of the protein (*e.g.*, the position of the carboxylate group of the 2OG and the D133 sidechain) and could influence the steric accessibility of the O_2-binding orbital on the Fe(II) ion in addition to its redox potential. The blue and green D133 sidechains that appear to be outliers come from succinate-bound structures in the presence and absence of nucleotide substrate, respectively. Both of these structures are in the open protein conformation and the observed difference in the conformation of D133 could be attributable to DNA binding. However, similar sidechain motions are not observed in comparing nucleotide-bound *vs.* free structures in the presence of 2OG.

Supporting Figure S12

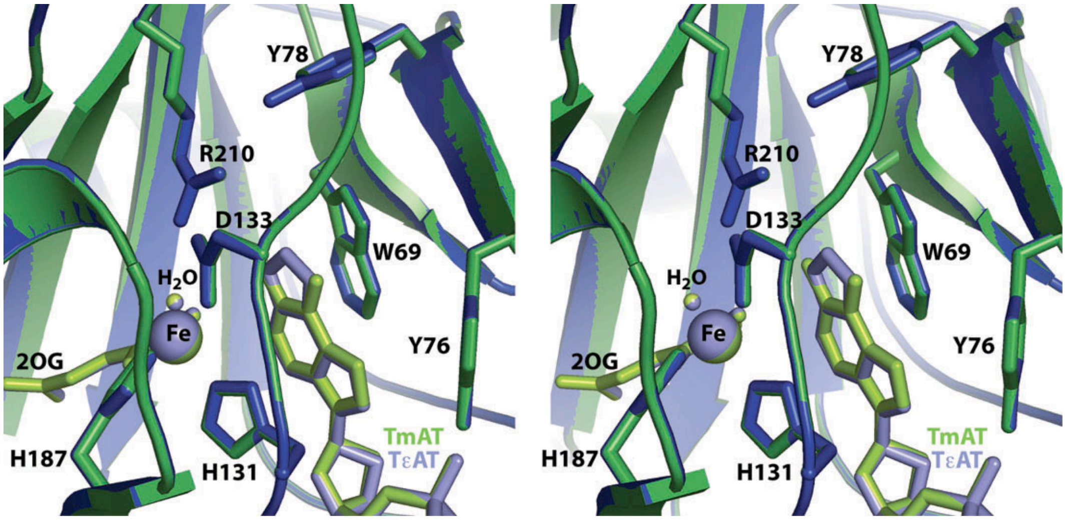


Figure S12. ***Stereopair of AlkB-ΔN11 modeled with a $1,N^6$-ethenoadenine substrate.*** The $1,N^6$-ethenoadenine base (blue) was modeled into the active site of AlkB-ΔN11 in place of the 1-methyladenine base. Refinement using diffraction data obtained for the Fe(II)-2OG-TmAT-bound structure (green) resulted in R_{free} and R_{work} values of 21.0 and 19.5 %, respectively, with negative density appearing for the extra carbon atom. Residues W69, D133, and R210 make the closest contacts to the extra carbon atom on $1,N^6$-ethenoadenine compared to 1-methyladenine but they are all at least 3.4 Å away.