HHS Public Access

Author manuscript

J Mol Biol. Author manuscript; available in PMC 2022 July 23.

Published in final edited form as:

J Mol Biol. 2021 July 23; 433(15): 167097. doi:10.1016/j.jmb.2021.167097.

Structural Insights into the Mechanism of Base Excision by MBD4

Lakshmi S. Pidugu[1], **Hilary Bright**[1], **Wen-Jen Lin**[2], **Chandrima Majumdar**[2], **Robert P. Van Ostrand**[2], **Sheila S. David**[2], **Edwin Pozharski**[1,3,*], **Alexander C. Drohat**[1,*]

[1] Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine, Baltimore, MD 21201, USA

[2] Department of Chemistry, University of California Davis, Davis, CA 95616, USA

[3] Center for Biomolecular Therapeutics, Institute for Bioscience and Biotechnology Research, Rockville, MD 20850, USA

Abstract

DNA glycosylases remove damaged or modified nucleobases by cleaving the *N*-glycosyl bond and the correct nucleotide is restored through subsequent base excision repair. In addition to excising threatening lesions, DNA glycosylases contribute to epigenetic regulation by mediating DNA demethylation and perform other important functions. However, the catalytic mechanism remains poorly defined for many glycosylases, including MBD4 (methyl-CpG binding domain IV), a member of the helix-hairpin-helix (HhH) superfamily. MBD4 excises thymine from G·T mispairs, suppressing mutations caused by deamination of 5-methylcytosine, and it removes uracil and modified uracils (e.g., 5-hydroxymethyluracil) mispaired with guanine. To investigate the mechanism of MBD4 we solved high-resolution structures of enzyme-DNA complexes at three stages of catalysis. Using a non-cleavable substrate analog, 2′-deoxy-pseudouridine, we determined the first structure of an enzyme-substrate complex for wild-type MBD4, which confirms interactions that mediate lesion recognition and suggests that a catalytic Asp, highly conserved in HhH enzymes, binds the putative nucleophilic water molecule and stabilizes the

[*]Correspondence to Edwin Pozharski and Alexander C. Drohat: corresponding authors. epozharskiy@som.umaryland.edu (E. Pozharski), adrohat@som.umaryland.edu (A.C. Drohat).

CRediT Author Statement

Lakshmi Pidugu: Investigation, Formal Analysis, Writing - Review & Editing. **Hilary Bright**: Investigation. **Wen-Jen Lin**: Resources. **Chandrima Majumdar**: Resources. **Robert Van Ostrand**: Resources. **Sheila David**: Funding acquisition, Resources, Supervision, Writing - Review & Editing. **Edwin Pozharski**: Investigation, Formal Analysis, Writing - Review & Editing. **Alex Drohat**: Conceptualization, Funding acquisition, Project Administration, Visualization, Writing – Original Draft, Review & Editing

Conflict of Interest

The authors declare no conflict of interest.

Declaration of interests

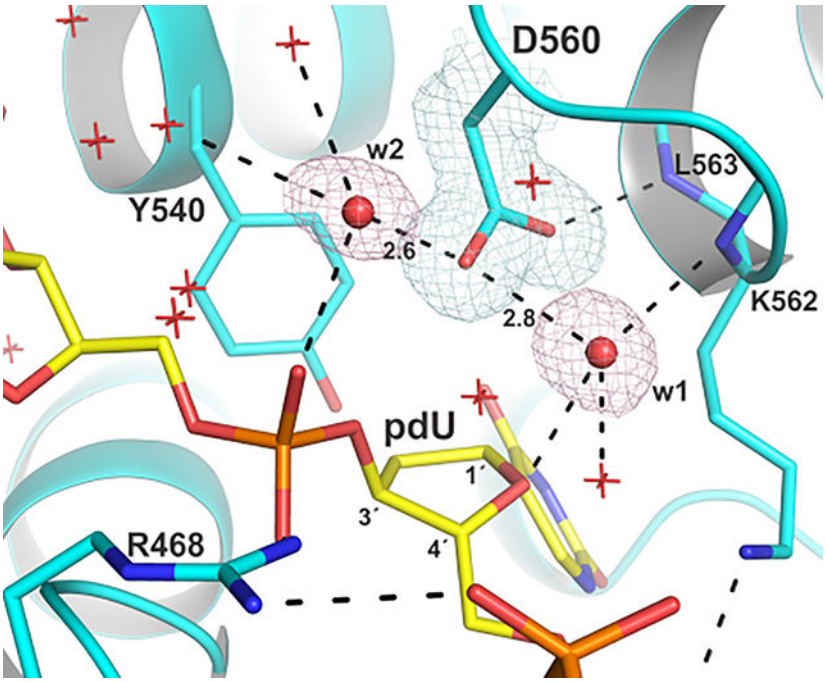
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary Material

Supplementary material can be found online

transition state. Observation that mutating the Asp (to Gly) reduces activity by 2700-fold indicates an important role in catalysis, but probably not one as the nucleophile in a double-displacement reaction, as previously suggested. Consistent with direct-displacement hydrolysis, a structure of the enzyme-product complex indicates a reaction leading to inversion of configuration. A structure with DNA containing 1-azadeoxyribose models a potential oxacarbenium-ion intermediate and suggests the Asp could facilitate migration of the electrophile towards the nucleophilic water. Finally, the structures provide detailed snapshots of the HhH motif, informing how these ubiquitous metal-binding elements mediate DNA binding.

Graphical Abstract



Keywords

base excision repair; DNA glycosylase; G/T mismatch; helix-hairpin-helix motif; 5-methylcytosine

Introduction

Base excision repair (BER) is initiated by DNA glycosylases, which find damaged or enzymatically modified DNA bases and remove them by cleaving the *N*-glycosyl bond.[1, 2] The resulting abasic (AP) site is excised and the correct nucleotide is incorporated by downstream BER enzymes. DNA glycosylases remove a broad range of lesions arising through alkylation, deamination, or oxidation of nucleobases.[3] These versatile enzymes also excise bulky nucleobase adducts, cleave interstrand cross links,[4–6] and modulate the levels of 5-methylcystosine (mC), an epigenetic mark in DNA, by excising mC or its derivatives.

[7, 8] Given their many critical cellular functions it is important to understand how these enzymes recognize substrates and catalyze base excision.

DNA glycosylases have been studied using a wide range of structural, biochemical, biophysical, and computational approaches and among the most rigorously characterized are uracil DNA glycosylase (UDG) and MutY.[9–11] UDG excises uracil from DNA and represents the UDG superfamily. MutY removes adenine mispaired with 8-oxoguanine and belongs to the Helix-Hairpin-Helix (HhH) superfamily.[12] UDG and MutY are the only DNA glycosylases subjected to transition state (TS) analysis using kinetic isotope effects (KIEs), and both studies found a direct-displacement hydrolytic reaction that follows a stepwise (S_N1) mechanism, where *N*-glycosyl bond cleavage yields a discrete oxacarbenium ion intermediate prior to nucleophile (water) addition (Figure 1a).[13, 14] KIE studies also indicated a stepwise mechanism for excision of adenine from DNA by ricin[15] and for non-enzymatic hydrolysis of deoxyadenosine monophosphate.[16] QM/MM studies of UDG found either a stepwise reaction[17] or one involving a single, highly dissociative oxacarbenium ion like TS (S_N2).[18, 19] Support for direct-displacement hydrolysis through a stepwise mechanism has been provided by structural, biochemical and biophysical studies of UDG. For example, a model for the second TS of a stepwise reaction (nucleophile addition) is offered by a crystal structure of DNA-bound UDG, with uracil and an oxacarbenium ion mimic, 1-azadeoxyribose (1-aza-dR, Figure 1c) in its active site.[20, 21] The structure shows a water molecule, thought to represent the nucleophile, contacting both the cationic sugar and the carboxylate of a catalytic Asp, and the uracil leaving group is bound as an anion (per NMR) at the opposite face of the sugar.[22, 23] A crystal structure of the UDG product complex shows the abasic sugar in an α-anomeric configuration, indicating a mechanism that leads to inversion of configuration at the anomeric center (Figure 1a).[24]

Intriguingly, studies of some HhH glycosylases have prompted suggestions that deoxynucleotide hydrolysis proceeds through a double-displacement mechanism that features a covalent glycosyl-enzyme intermediate involving an Asp residue that is highly conserved in the HhH superfamily (Figure 1b). Such a role for the Asp would differ starkly from its previously suggested roles in a direct-displacement reaction, which include positioning and possibly deprotonating the nucleophile (water) and stabilizing the TS and oxacarbenium ion intermediate.[14, 21, 25] The possibility that HhH enzymes catalyze a reaction involving a glycosyl-enzyme intermediate was initially raised by a structure of E coli AlkA bound to DNA with the 1-aza-dR mimic flipped into its active site.[26] Observation that the Asp contacted 1-aza-dR and that no water molecule (potential nucleophile) was found near 1-aza-dR prompted the idea that the Asp could either stabilize an oxacarbenium ion intermediate in a direct-displacement reaction or perform nucleophilic attack to form a covalent intermediate in a double-displacement reaction. A similar structure of DNA-bound MutY showed that N1 of 1-aza-dR is contacted on one face (3′) by the conserved Asp and on the opposite face by a water molecule, prompting the suggestion of a double-displacement mechanism. This was supported by NMR analysis of products resulting from MutY methanolysis reactions, which indicated retention of configuration.[27, 28] Finally, a structure of MBD4 (methyl-CpG binding domain IV), another HhH enzyme, was reported to feature a covalent bond between the conserved Asp and C1′ of the flipped deoxyribose. The crystallographic sample was generated by incubating MBD4 (catalytic domain) with a DNA

substrate, and it was proposed that the structure captured an enzyme-glycosyl intermediate in a double-displacement reaction.[29] However, the authors noted that additional studies would be needed to test their proposal, given the relatively low resolution of the structure (2.84 Å).

To investigate the mechanism of MBD4, we solved high resolution crystal structures of enzyme-DNA complexes that represent three potential stages of catalysis. A structure with the non-cleavable substrate analog 2′-deoxy-pseudouridine (pdU, Figure 1c) flipped into its active site provides the first glimpse of an enzyme-substrate (E·S) complex for wild-type MBD4 and is only the second structure of an E·S complex for a member of the HhH superfamily that features wild-type enzyme. A structure of MBD4 bound to 1-aza-dR DNA provides a model for a potential oxacarbenium ion intermediate. A structure of the enzyme product (E·P) complex addresses key questions raised by a previous structure of DNA-bound MBD4 solved at low resolution.[29] We also investigated the role of a catalytic Asp, which is highly conserved in HhH enzymes, using mutagenesis and single turnover kinetics experiments. Our new structures also provide detailed snapshots of the helix-hairpin-helix (HhH) motif, informing how these ubiquitous elements, together with a bound metal ion (typically K^+) mediate DNA binding. HhH motifs are found in a broad range of proteins and enzymes but are poorly defined in the large majority of protein structures. The structures also inform how MBD4 performs its key functions in BER.[30] It excises thymine from G·T mispairs arising through the deamination of mC, a modified base that mediates epigenetic regulation.[31–33] MBD4 also excises uracil and 5-substituted uracils (5-hydroxymethyluracil, 5-fluorouracil, etc.) with a strong preference for bases that are paired with guanine rather than adenine.[34, 35] The structures could facilitate biochemical, biophysical, and computational studies of MBD4, advance efforts to develop small molecule inhibitors targeting MBD4, and may help to inform how mutations that impair MBD4 function might be linked to carcinogenesis.[36–38]

Results and Discussion

Structure of the MBD4 enzyme-substrate complex

Solving the structure of an enzyme-substrate (E·S) complex requires modification of the enzyme or substrate to halt the reaction in a manner that minimally perturbs enzyme-substrate interactions, and this typically involves mutation of an essential catalytic residue of the enzyme or a non-reactive substrate analog. The former approach was used for prior structures of the E·S complex for MBD4 (catalytic domain), by mutating the conserved Asp (D560, human MBD4).[29, 39, 40] While these structures revealed much about how MBD4 recognizes lesions including G·T mispairs, they lack information about the role of the mutated catalytic Asp. Although the Asp does not contact the flipped base, the possibility that its mutation alters (indirectly) enzyme-base interactions, or perturbs other aspects of substrate recognition, could not be excluded in the absence of a structure of wild type enzyme bound to DNA with a non-cleavable substrate analog. Indeed, structures of MutY obtained with both approaches indicated that mutating the conserved Asp alters interactions with the target base.[25, 41]

The deoxynucleotide analogs used most frequently to preclude *N*-glycosyl bond cleavage feature a subtle 2′-fluoroarabino substitution that destabilizes the oxacarbenium ion like transition state (Figure 1c).[42–45] Such analogues have been used for studies of enzymes including MutY and thymine DNA glycosylase (TDG),[25, 46–49] and we investigated their utility for MBD4 using DNA with 2′-F-dT or 2′-F-dU, but found that the enzyme tended to precipitate in the presence of either substrate analog at the concentrations used for crystallization. We then considered 2′-deoxy-pseudouridine (pdU, Figure 1c), the dU analog employed for a structure of the UDG E·S complex,[50] and obtained a high-quality structure of the MBD4 E·S complex at 1.57 Å resolution (Figure 2, Table S1). The dU substrate and pdU analog differ at only two positions of the base; N1 and C5 of dU are replaced by C1 and N5 of pdU. Our structure shows that MBD4 does not contact either position of pdU, indicating that it accurately reflects interactions in the lesion-recognition complex. Regarding our finding that MBD4 precipitates at high concentration in the presence of DNA containing 2′-F-dT (or -dU), the structure suggests a potential steric clash between the 2′-F-arabino substituent and L466, an active-site residue that contacts the flipped nucleotide (Figure 2b).

The structure shows that pdU flips completely into the active site, with O2, N3, and O4 contacting the Q449 and Y540 side chains and the backbone nitrogen of V448 (Figure 2a). Each contact likely contributes to substrate recognition and helps to stabilize the flipped conformation of the deoxynucleotide, as needed for a productive E·S complex. The contacts involving O2 and O4 are also expected to facilitate *N*-glycosyl bond cleavage, by stabilizing negative charge developing on the leaving group in the TS. Consistent with this idea, the Q449A mutation depletes MBD4 activity and mutation of Y540 reduces its activity, as does mutation of the corresponding Tyr for MIG (mismatch glycosylase), a thermophilic HhH enzyme that also excises T from G·T mispairs.[39,51] The hydrogen bonds provided by Q449 to N3 and O4 likely confer specificity for excision of uracil, thymine, and other 5-substituted uracils, since they would be unfavorable for the corresponding positions of cytosine bases (N3, N4). Notably, the orientation of Q449 is stabilized by contacts to its backbone nitrogen and S544, so it is not likely to adopt an orientation (180° flip) that could facilitate the excision of cytosine bases. These observations likely account, at least in part, for findings that MBD4 does not excise cytosine or modified cytosines, including the epigenetic bases 5-formylcytosine or 5-carboxylcytosine.[39, 52]

Nucleotide flipping is likely to be stabilized by penetration of R468 through the DNA minor groove, where it occupies the space vacated by the flipped nucleotide, forms a salt bridge with two backbone phosphates, and provides a water-mediated contact to O6 of the mismatched Gua (Figure 2b). The mismatched Gua, which remains intrahelical, is also contacted (at N1, N2) by two backbone oxygens (R468, L506). We find no contacts to the Gua on the 3′ side of the flipped nucleotide, or its Cyt base-pairing partner, consistent with findings that the activity of MBD4 (glycosylase domain) does not depend on the nucleotide located 3′ of the mismatched thymine (it lacks specificity for a CpG context).[52] By contrast, the other mammalian G·T mismatch glycosylase, TDG, is highly specific for acting on G·T pairs in a CpG context.[47, 53, 54]

The electron density at 1.57 Å resolution shows that the sugar pucker of pdU is predominantly 3′-exo, with modest 2′-exo, when flipped into the MBD4 active site (Figure S1a). The same conformation is observed for the natural dT substrate in the active site of D560N-MBD4 and D560A-MBD4 (Figure S1b, c).[39, 40] Notably, TS analysis shows that the deoxyribose adopts a 3′-exo pucker in the TS for *N*-glycosyl bond cleavage, for nonenzymatic and enzymatic reactions (UDG, MutY).[11, 13, 14] These observations suggest that MBD4 utilizes ground-state strain to impose a deoxyribose conformation in the E·S complex which approaches that of the TS, similar to findings for UDG.[13, 50] It is of interest to compare the conformation the pdU substrate analog in structures of the E·S complex for MBD4 and UDG (Figure S2).[50] Previous studies found that UDG severely distorts the flipped pdU analog, rotating the uracil ring about its C1-N4 axis by about 90°, flattening the sugar to a slight 3′-exo pucker, and imposing a more tetrahedral geometry to the C1 position (N1 in uracil; normally trigonal planar), such that the plane of the nucleobase is bent down towards that of the sugar. MBD4 exerts far less distortion on the pdU analog, as judged by the three metrics noted above for UDG. Substrate distortion by UDG is thought to facilitate catalysis,[50] and observation of reduced distortion for MBD4 might reflect the fact that its rate enhancement (~10^8) is about four orders of magnitude lower than that of UDG (~10^{12}).[9,55] Turning our attention back to MBD4, we note that the enzyme-substrate interactions detailed above in our new structure of the E·S complex are similar to those observed in structures solved at lower resolution using DNA containing a natural G·T mispair and a catalytically inactive variant of MBD4 (D560N, D560A).[29, 39, 40] In addition, the overall protein conformation of our new structure is similar to the best of the prior structures, solved at 1.55 Å resolution using D560N-MBD4 (Figure S3).[40] Together, these observations indicate that the pdU analog captures the same binding mode as a natural dT substrate. They also indicate that the D560N/A mutations used in the previous structures of MBD4 do not substantially alter its interactions with the flipped nucleobase or the mismatched Gua, which is notable given the prior findings that mutating the conserved Asp of MutY alters its substrate interactions.[25, 41] Nevertheless, interactions involving the conserved Asp will be altered by its mutation and we next consider what our structure reveals about the role of this key catalytic residue.

As seen for other HhH enzymes, the conserved Asp of MBD4 resides at the N-terminus of an α-helix and it forms canonical helical-capping interactions with the backbone nitrogen of downstream residues (Figure 3).[56] Our structure reveals that the D560 carboxylate binds two water molecules, one of which (w1) is positioned 3.59 Å away from C1′ of pdU (representing the electrophile). The w1 water also contacts a backbone nitrogen (K562), the O4′ oxygen of pdU, and another water molecule. The position of w1 and its interaction with D560 make it a strong candidate to serve as the nucleophile in a direct-displacement reaction. In addition, the D560 carboxylate is positioned 3.6 Å from C1′ of pdU, suggesting it stabilizes positive charge developing on the sugar in the TS for *N*-glycosyl bond cleavage. Notably, the distance between the potential nucleophilic water (w1) and the electrophile (C1′) observed here for MBD4 (3.59 Å) is nearly equivalent to that found in a structure of the UDG E·S complex (3.54 Å), and UDG also binds the nucleophile using a highly conserved Asp.[50]

It is informative to compare our structure with previous structures of the E·S complex, which were solved using MBD4 with mutations of the conserved Asp.[29, 39] The structure of D560N-MBD4 with G·T DNA indicates the mutation alters interactions with the two D560-bound waters (w1, w2; Figure S4a). The oxygen of N560 that contacts w1 is also the oxygen which forms the helical cap, unlike the D560-w1 contact shown here, and w2 is not observed in the structure for D560N-MBD4. In addition, the carboxamide nitrogen of N560 is positioned near the face of the sugar for the flipped deoxynucleotide, suggesting it could destabilize positive charge in the TS. Moreover, the orientation of N560 is stabilized by a contact with the 3′-oxygen and the helical capping interaction. Regarding the D5650A mutation, TS stabilization provided by the D560 carboxylate would be lost, and neither of the D560-bound waters are observed in a structure of D560A-MBD4 with G·T DNA (Figure S4b).[39] Taken together, the observations above indicate that our structure provides the most accurate model of the E·S complex for MBD4.

Mutational interrogation of the conserved Asp

To further investigate the role of the catalytic Asp we sought to quantify the effect of removing the carboxylate group on the rate of base excision. Because the conserved Asp forms an N-terminal helix cap, we replaced it with Gly, which has a higher N-capping preference relative to Ala.[56] Initial studies indicated little activity for D560G-MBD4 on a G·T substrate in 24 h, so we tried a substrate anticipated to provide higher activity, 5-trifluoromethyluracil (aka trifluorothymine, T^{F3}), a thymine analog that is excised by MBD4.[57] The leaving-group quality is expected to be better for T^{F3} versus T, owing to the greater electron-withdrawing effect of CF_3 ($\sigma_m = 0.43$) relative to CH_3 ($\sigma_m = –0.07$),[58] and MBD4 activity was found to depend on leaving-group quality (for 5-substituted uracils),[34] as observed for TDG.[55] Accordingly, MBD4 excises T^{F3} from G·T^{F3} pairs with a rate constant of $k_{obs} = 7.3 \pm 0.8$ min^{-1}, some 66-fold faster than it removes T from G·T pairs, $k_{obs} = 0.11 \pm 0.01$ min^{-1} (Figure 4). D560G-MBD4 removes T^{F3} from a G·T^{F3} substrate with a rate of $k_{obs} = 0.0027 \pm 0.0003$ min^{-1}, revealing that removing the carboxylate of the catalytic Asp reduces MBD4 activity by 2700-fold. This finding indicates an important role(s) in catalysis for the Asp, but likely not one as a nucleophile that forms a covalent intermediate in a double-displacement reaction, as discussed below.[29]

Structure of the MBD4 enzyme-product complex

To further inform the mechanism of MBD4, we sought to determine a high-resolution structure of the enzyme-product complex. The need is driven in part by a prior structure, solved at 2.84 Å resolution, which was suggested to represent a stalled reaction intermediate (PDB ID 4EW4).[29] Seeking to obtain a structure of the product complex, the investigators generated crystallization samples by incubating MBD4 with G·T DNA substrate. While the electron density showed that the mismatched thymine was excised, it also appeared to suggest that the AP sugar lacked the expected hydroxyl group, and it was proposed that the complex could feature a stalled oxacarbenium ion intermediate, stabilized by the conserved Asp. An alternative interpretation, based on the observation of continuous electron density between the AP sugar and the Asp carboxylate, was that the complex featured a stalled covalent intermediate in a double displacement reaction whereby the Asp would attack C1′ of the substrate (Figure 1b). However, both proposals, based on a low-resolution structure,

seem unlikely as there is no reason to expect that the MBD4 reaction would spontaneously stall, indefinitely, at an intermediate state. Oxacarbenium ions are highly reactive and unlikely to be stabilized indefinitely in an active site.[11, 16] The authors noted that higher resolution structures would be needed to test their ideas.[29] Some clarification was offered by a recent structure of the MBD4 product complex, solved at 2.02 Å resolution, which shows the expected hydroxyl for the AP sugar, but ambiguity remained due to the observation of continuous electron density between the AP sugar and the Asp (PDB ID: 6VJW).[40] To help resolve the questions raised by these previous structures, we solved a structure of the MBD4 product complex, at 1.62 Å resolution, using crystals derived from samples generated by incubating MBD4 with a G·T DNA substrate (Figure 5, Table S1). The electron density confirms that the mismatched thymine was cleaved, as expected, and shows that it departs the active site of the product complex. Regarding the questions raised by a prior structure,[29] the electron density of our structure demonstrates clearly that the abasic sugar contains a hydroxyl group, as expected for the hydrolytic reaction. Moreover, we observe fully discrete electron density for the AP sugar and the catalytic Asp (D560). Thus, our findings do not support the proposed observation of a stalled reaction intermediate in a prior low-resolution structure of the MBD4 product complex.[29]

Our structure also shows that the AP sugar exists exclusively in the α-anomeric form, indicating inversion of configuration, the outcome expected for a direct displacement hydrolytic reaction (Figure 1a). Notably, the hydroxyl of the AP sugar is contacted by both carboxylate oxygens of D560 and by an ordered water molecule that is itself coordinated by two catalytic groups, Q449 and Y540. Together, these contacts to the AP sugar likely stabilize its flipped conformation and contribute to the tight binding of MBD4 to abasic DNA.[59]

Our structure of the product complex also reveals water molecules that contact catalytic groups, including two that bind the D560 carboxylate (w1, w2), as observed for the E·S complex. Notably, the structure also shows three water molecules in sites that would otherwise be occupied by O2, N3, and O4 of the flipped base (Ura, Thy) in the E·S complex, and the waters contact the same enzyme groups as those sites of the nucleobase. This suggests the water molecules could thermodynamically favor departure of the excised base from the product complex. Release of the product base could also be facilitated by the high degree of solvent exposure observed for the active site of MBD4 (Figure 2).

Structure of MBD4 with DNA containing 1-aza-dR

To obtain a model of MBD4 bound to a potential oxacarbenium ion intermediate, we solved a structure of the enzyme bound to DNA with 1-aza-dR (Figure 1c) flipped into its active site, at 1.68 Å resolution (Table S1). The overall structure and most of the protein-DNA contacts are similar to those found in the E·S and E·P complexes, so we focus on 1-aza-dR and observations in the active site (Figure 6a). The cationic amine (N1) of 1-aza-dR forms a close contact with the D560 carboxylate (2.67 Å), consistent with an important catalytic role for the conserved Asp, as discussed below. The azasugar is also contacted by a chloride anion (Cl^-), positioned on its opposite face relative to D560. The Cl^- anion also contacts two catalytic groups, Q449 and Y540, and a solvent molecule. The assignment of Cl^-, rather

than H_2O, is dictated by electron density, and the observed coordination parameters are consistent with those found for Cl^- in high resolution crystal structures.[60] Observation of Cl^- at this location is not particularly surprising given that a cationic nitrogen is a common ligand for Cl^-, and that the crystallization sample contained Cl^- at a concentration of 80 mM. Notably, the Cl^- anion could potentially serve as an electrostatic mimic for the anionic base of a stepwise intermediate (Figure 1a). The location of Cl^- and its distance from the cationic sugar are roughly similar to that which could be anticipated for the leaving-group nitrogen of the nucleobase in a stepwise intermediate, as suggested by studies of UDG and MutY,[14, 17, 21] and by aligning our structures of the E·S (pdU) and E·I (1-aza-dR) complexes for MBD4 (Figure 6b). Notably, we find that 1-aza-dR adopts a 3′-exo conformation (Figure S1d), one of two predominant puckers for oxacarbenium ions, and the conformation observed for the oxacarbenium ion intermediate in KIE studies of the MutY and UDG reactions.[13, 14]

The alignment of our structures also reveals a dramatic difference in the position of 1-aza-dR relative to the sugar of the pdU substrate analog, particularly at the 1′ position, which is displaced by 1.2 Å (Figure 6b). By contrast, small differences are seen for most other regions of the DNA and elements of the active site. Similarly, the distance between 1′ of the sugar and the D560 carboxylate is reduced from 3.6 Å (E·S) to 2.7 Å. This structural observation suggests that in the reaction catalyzed by MBD4, after *N*-glycosyl bond rupture, the Asp could lure the glycosyl cation intermediate towards the nucleophilic water molecule, reminiscent of findings for the catalytic Asp of UDG.[21] Such electrophile migration could potentially serve to separate the ribosyl cation from the nucleobase after C-N bond rupture, which could favor the addition of water over reformation of the *N*-glycosyl bond.

Regarding the nucleophilic water molecule, our structure of the E·P complex suggests that it adds to the opposite face of the deoxyribose relative to the leaving group, as discussed above. While the electron density for our structure of the E·I complex does not indicate an ordered water molecule that contacts N1 of 1-aza-dR, a water coordinated by D560 and the K562 backbone, corresponding to w1 in the E·S and E·P complexes, would likely to do so (Figure 6b). We speculate that the absence of an ordered water corresponding to w1 in our model of the E·I complex could be due in part to factors unique to the 1-aza-dR analog, including a weakened interaction with D560, which forms a close electrostatic contact with 1-aza-dR, and a potential steric clash with C4 of 1-aza-dR (Figure 6b). Taken together, our structures suggest that a water molecule coordinated by the conserved Asp (D560) and a backbone nitrogen (K562) serves as the nucleophile in the MBD4 reaction. While nucleophile activation is probably not required for a stepwise reaction, given the high reactivity expected for an oxacarbenium ion, D560 could potentially accept the proton liberated upon nucleophile addition.[11]

High resolution structures of the Helix-hairpin-helix motif

The helix-hairpin-helix (HhH) motif is a signature feature of HhH superfamily glycosylases and is also found in many other enzymes that act on DNA, including DNA polymerases.[61–65] These motifs mediate sequence-independent DNA binding through interactions with two backbone phosphates.[64] Notably, one of the phosphate contacts involves a metal ion that

binds to the HhH motif. Previous studies show that the metal is most likely potassium (K^+) under physiological conditions and that HhH motifs have little affinity for the metal in the absence of DNA.[26, 39, 65] Our structures show that the HhH motif contacts DNA phosphates downstream (3′) of the flipped nucleotide, at the +2 and +3 positions (Figure 7a), consistent with previous structures of MBD4 and other HhH glycosylases.[26, 29, 39, 52] While we will point to our structure of the E·S complex to illustrate these DNA contacts, the same set of interactions are observed for the E·P and E·I complexes though the metal ion differs, as noted below. The HhH motif provides a total of five contacts to two DNA phosphates (+2, +3) using backbone amides (G536, G538, K539, Y540, G541). It also forms two contacts to the +3 phosphate, through the metal ion and a metal-bound water molecule. Together, these seven interactions constitute half of those provided to the target DNA strand. One contact at the +2 site involves the backbone nitrogen of a key catalytic residue, Y540, suggesting a mechanism to couple DNA binding with the position of a catalytic group.

Our structures also inform the identity and coordination of the metal ion in HhH motifs. While dozens of crystal structures have been solved for a range of HhH glycosylases, including many at relatively high resolution, the identity and coordination of the metal in HhH motifs is found to be highly variable. The problem is exemplified by previous structures of MBD4, for which the metal site of the HhH motif has been modeled with nickel (Ni^{2+}), magnesium (Mg^{2+}), K^+, or a water molecule, and with a broad range of coordination distances and geometries.[29, 39, 40, 52] Similar variability is found for HhH motifs in other DNA glycosylases, modeled with these and other metals, including Ca^{2+}. Our structure of the E·S complex provides a high-resolution look at the HhH motif and shows that the metal is coordinated in octahedral geometry by six oxygen atoms, including three from the protein backbone, one from a DNA phosphate, and two from water molecules (Figure 7b). Taken together, the ligand type (oxygen), coordination geometry, and metal-oxygen distances indicate the HhH motif binds K^+, at least under the conditions used here for the E·S complex. Shorter metal-oxygen distances would be expected for other metals (Ca^{2+}, Na^+, Mg^{2+}). Moreover, electron density and structural parameters of the metal and its coordinating nuclei are consistent with K^+ but not with the other metals.[66] The K^+ assignment is supported by analysis of our structure using an online server (CMM) developed by Minor and colleagues to evaluate metal-binding sites in proteins.[67] Notably, the concentration of Na^+ and K^+ were roughly equivalent in the crystal conditions used for the E·S complex. Thus, our finding that K^+ occupies the HhH motif is consistent with previous findings that HhH motifs exhibit higher affinity for K^+ versus Na^+ (relative affinity: $K^+ > Na^+ > Ca^{2+} > Mg^{2+}$).[64, 65]

Meanwhile, similar analysis of the E·P and E·I complexes shows that Na^+ occupies the HhH motif (not shown), consistent with the presence of Na^+, and the absence of K^+, in the crystallization conditions for these structures. The Na^+ assignment is based on octahedral geometry, metal-oxygen distances for three backbone oxygens (2.31–2.39 Å), two water molecules (2.39–2.40 Å), and a DNA phosphate (2.57 Å), and structural parameters (*B*-factor, occupancy), and is supported by analysis using the CMM server. [67] Notably, in all three of our structures, the metal ion (K^+ or Na^+) is coordinated by the same set of ligands, including a carbonyl oxygen at three sites of the protein backbone (I532, L534, and I537), a DNA phosphate oxygen, and two water molecules. We explored the potential effect of K^+

relative to Na^+ on MBD4 activity (for a G·T substrate) using single turnover kinetics experiments and find that activity is only slightly elevated (1.3-fold) in buffer containing KCl versus NaCl (Figure 4). The location of metal-coordinating oxygens seen in our structures, vis-à-vis primary sequence (i, i+2, i+5), is identical to that observed for other proteins with a rigorously characterized HhH motif, including DNA polymerase beta and the AlkA DNA glycosylase.[26, 64, 65] Together, these observations point to a highly conserved mechanism for binding a metal ion, likely K^+, by HhH motifs *in vivo*.

Implications for the catalytic mechanism of MBD4 and other HhH glycosylases

The structural and biochemical findings reported here, together with previous studies, indicate that MBD4 hydrolyzes *N*-glycosyl bonds through a direct displacement mechanism (Figure 1a). Observation that the D560G mutation reduces activity by $10^{3.4}$-fold (for a $G{\cdot}T^{F3}$ substrate) demonstrates that the conserved Asp performs at least once central catalytic function, and our structures suggest several potential roles that could be consistent with this mutational effect. Our structure of the enzyme-substrate complex unveils interactions involving the conserved Asp, revealing that it binds a water molecule which could represent the nucleophile and that it likely stabilizes the TS for *N*-glycosyl bond cleavage. Our structure with the 1-aza-dR analog suggests the Asp could facilitate migration of an oxacarbenium ion intermediate towards the nucleophilic water molecule, which could be bound by the Asp and a backbone nitrogen. Notably, the Asp also seems poised to receive the proton released upon nucleophile addition. Our structural and biochemical results seem inconsistent with a role for the conserved Asp as the nucleophile in a double-displacement reaction. Indeed, our structure of the enzyme-product complex indicates a reaction that gives inversion of configuration, as would be expected for direct-displacement hydrolysis. A double-displacement mechanism also seems inconsistent with the finding that D560G-MBD4 retains some glycosylase activity. Previous studies of double-displacement reactions catalyzed by several different retaining glycosidases showed that detectible activity is completely eliminated (reduced by $>10^6$-fold) upon removal of the nucleophilic carboxylate group (Asp, Glu) that forms the covalent intermediate.[68] In addition to our finding for MBD4, other monofunctional HhH enzymes exhibit residual activity upon mutation of the conserved Asp, including MagI (*S. pombe*), MagIII (*H. pylori*),[3] and, most prominently, MIG, for which the Asp to Ala mutation gives a modest 16-fold loss in activity.[51] Findings that detectable activity is largely depleted when the Asp is converted to Asn, for MBD4 and other HhH glycosylases, could reflect unfavorable interactions involving the carboxamide nitrogen of Asn that are not present for Asp or other residues (Gly, Ala).[11] Indeed, structures of the Asn variant for MBD4 (D560N) indicate that the carboxamide nitrogen could destabilize an oxacarbenium ion like TS or oxacarbenium ion intermediate (Figure S4a). It should also be noted that retaining glycosidases feature a conserved general acid/base (Asp, Glu) that has two functions, one of which is to catalyze hydrolysis of the glycosyl-enzyme intermediate.[68] However, MBD4 lacks a residue that could serve this role in a double-displacement mechanism; the Glu proposed to catalyze deglycosylation for MutY (E43) is replaced by a Thr in MBD4 (T463). Similarly, while MIG features a Glu (E42) corresponding to E43 of MutY, the E42S mutation in MIG has a modest (26-fold) effect on activity. Together, these observations point to a direct-displacement hydrolysis mechanism for MBD4, MIG, and probably some other monofunctional HhH glycosylases. Evidence that

MutY employs a retaining mechanism, based on the methanolysis studies alluded to above, [27] raises the possibility that monofunctional glycosylases in the HhH family employ more than one mechanism for *N*-glycosyl hydrolysis. We anticipate that this possibility will be further explored in future studies of these remarkable enzymes.

Materials and Methods

Materials

The glycosylase domain of human MBD4, comprising the 155 C-terminal residues (426–580), hereafter termed "MBD4", was produced essentially as previously described,[52] as was the D560G variant. Standard oligodeoxynucleotides (ODNs) were obtained from IDT; ODNs containing 2'-deoxy-pseudouridine (pdU) or 5-trifluoromethyl-2'-deoxyuridine were synthesized at the Keck Foundation Biotechnology Resource Laboratory at Yale University, using phosphoramidites obtained from Glen Research (Sterling, VA) or ChemGenes (Wilmington, MA). These ODNs were purified by reverse phase HPLC.[69] Synthesis of (3R,4R)-3-hydroxy-4-(hydroxymethyl) pyrrolidine-1-ium (1-aza-dR) and its phosphoramidite was completed as described,[70] and ODNs containing 1-aza-dR were synthesized, deprotected, and purified by HPLC or denaturing PAGE as described.[27] ODNs were exchanged into 0.02 M Tris-HCl pH 7.5, 0.04 M NaCl, and quantified by absorbance. [71] The DNA included a 12mer target strand, 5′ CCA GCG xGC AGC, where x is dT, pdU, or 1-aza-dR, and a 12mer complementary strand, 5′ GCT GCG CGC TGG. For the MBD4 activity assays, the 28 bp duplex was comprised of a target strand, 5′-AGC TGT CCA TCG CTC AxG TAC AGA GCT G, where x is T or 5-trifluoromethyluracil (T^{F3}), and the complementary strand, 5′-CAG CTC TGT ACG TGA GCG ATG GAC AGC T.

X-ray crystallography

The samples used for crystallization contained 0.25 mM (5.2 mg/ml) MBD4 and 0.30 mM DNA in a buffer of 0.05 M Tris-HCl pH 7.5, 0.1 M NaCl, 0.2 mM DTT, and 10% glycerol. Samples of the enzyme-substrate (E·S) complex were produced by incubating MBD4 with DNA containing pdU (paired with G) for 2h at room temperature prior to crystallization. Samples of the enzyme-product (E·P) complex were produced by incubating MBD4 with DNA containing a G·T mispair for a sufficient time to ensure full conversion to AP-DNA product (2 h), as established by HPLC analysis.[55] Crystals were grown at room temperature by sitting drop vapor diffusion using 0.5 μl of the MBD4-DNA sample and either 0.5 μl or 1.0 μl of mother liquor, which varied as follows. Mother liquor was 0.2 M potassium sodium tartrate tetrahydrate pH 7.4 20% w/v PEG 3350 for the E·S complex, 0.1 M Tris HCl pH 8.0 20% w/v PEG 400 for the E·P complex, and 0.1 M HEPES pH 7.5 30% w/v PEG 1000 for the E·I complex. Crystals typically appeared within a few days and were cryo-protected using mother liquor supplemented with 18–20% glycerol prior to flash cooling in liquid nitrogen. X-ray diffraction data were collected at the Stanford Synchrotron Radiation Lightsource using beamlines 9–2, 11–1, and 12–2. Images were processed using XDS[72] and scaled using Aimless[73] from the CCP4 suite,[74] with the autoxds script developed by A Gonzalez and Y Tsai (http://smb.slac.stanford.edu/facilities/software/xds). The resolution cutoff was determined based on CC1/2 values.[75] The structures were solved by molecular replacement using Phaser,[76] and a previous structure of DNA-bound MBD4 as the search

model (PDB ID: 4DK9). Refinement was performed using BUSTER-TNT [77] or Phenix.Refine,[78] and model building was performed using Coot.[79] The TLS refinement utilized the TLSMD server,[80, 81] as described.[69] Structural figures were made with PyMOL (http://www.pymol.org).

Enzyme activity assays.

Glycosylase activity was determined for MBD4 or D560G-MBD4 acting on DNA containing a G·T or a G·T^{F3} pair using single turnover kinetics experiments performed under saturating enzyme conditions.[54, 82] Experiments were performed at 23 °C in a buffer comprising 0.02 M HEPES pH 7.5, 0.1 M NaCl, 0.2 mM EDTA (unless stated otherwise). Rate constants were determined by fitting progress curves (fraction product versus time) to eq. 1 using non-linear regression:

$$\text{fraction product} = A(1 - \exp(-kt)) \quad (1)$$

where A is the amplitude, k is the rate constant, and t is reaction time.

Accession Codes

Coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 7KZ0, 7KZ1, and 7KZG.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

The studies were supported by grants from the National Institutes of Health NIGMS (R01-GM072711, R35-GM136225, to ACD), and the National Science Foundation (CHE1905304, to SSD). X-ray diffraction data were collected at the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02–76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and the National Institutes of Health, National Institute of General Medical Sciences (P41GM103393). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH.

Abbreviations

1-aza-dR	1-azadeoxyribose
AP	apurinic/apyrimidinic
BER	base excision repair
CpG	cytosine-guanine dinucleotide
HhH	helix-hairpin-helix
KIE	kinetic isotope effect
mC	5-methylcytosine

MBD4	(methyl-CpG binding domain IV)
ODN	oligodeoxynucleotide
pdU	2′-deoxy-pseudouridine
TDG	thymine DNA glycosylase
TF^3	trifluorothymine
TS	transition state
UDG	uracil DNA glycosylase

References

1. Krokan HE, Bjoras M. Base excision repair. Cold Spring Harb Perspect Biol. 2013;5:a012583. [PubMed: 23545420]
2. Drohat AC, Maiti A. Mechanisms for enzymatic cleavage of the N-glycosidic bond in DNA. Org Biomol Chem. 2014;12:8367–8378. [PubMed: 25181003]
3. Brooks SC, Adhikary S, Rubinson EH, Eichman BF. Recent advances in the structural mechanisms of DNA glycosylases. Biochim Biophys Acta. 2013;1834:247–71. [PubMed: 23076011]
4. Mullins EA, Shi R, Parsons ZD, Yuen PK, David SS, Igarashi Y, Eichman BF. The DNA glycosylase AlkD uses a non-base-flipping mechanism to excise bulky lesions. Nature. 2015;527:254–8. [PubMed: 26524531]
5. Mullins EA, Rodriguez AA, Bradley NP, Eichman BF. Emerging Roles of DNA Glycosylases and the Base Excision Repair Pathway. Trends Biochem Sci. 2019;44:765–781. [PubMed: 31078398]
6. Semlow DR, Zhang J, Budzowska M, Drohat AC, Walter JC. Replication-Dependent Unhooking of DNA Interstrand Cross-Links by the NEIL3 Glycosylase. Cell. 2016;167:498–511. [PubMed: 27693351]
7. Drohat AC, Coey CT. Role of Base Excision “Repair” Enzymes in Erasing Epigenetic Marks from DNA. Chem Rev. 2016;116:12711–12729. [PubMed: 27501078]
8. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. Nature.2013;502:472–9. [PubMed: 24153300]
9. Stivers JT, Jiang YL. A mechanistic perspective on the chemistry of DNA repair glycosylases. Chem Rev. 2003;103:2729–59. [PubMed: 12848584]
10. Friedman JI, Stivers JT. Detection of Damaged DNA Bases by DNA Glycosylase Enzymes. Biochemistry. 2010;49:4957–4967. [PubMed: 20469926]
11. Berti PJ, McCann JA. Toward a detailed understanding of base excision repair enzymes: transition state and mechanistic analyses of N-glycoside hydrolysis and N-glycoside transfer. Chem Rev.2006;106:506–55. [PubMed: 16464017]
12. David SS, O’Shea VL, Kundu S. Base-excision repair of oxidative DNA damage. Nature.2007;447:941–50. [PubMed: 17581577]
13. Werner RM, Stivers JT. Kinetic isotope effect studies of the reaction catalyzed by uracil DNA glycosylase: evidence for an oxocarbenium ion-uracil anion intermediate. Biochemistry. 2000;39:14054–14064. [PubMed: 11087352]
14. McCann JAB, Berti PJ. Transition-state analysis of the DNA repair enzyme MutY. J Am Chem Soc.2008;130:5789–5797. [PubMed: 18393424]
15. Chen XY, Berti PJ, Schramm VL. Transition-state analysis for depurination of DNA by ricin A-chain. J Am Chem Soc. 2000;122:6527–6534.
16. McCann JAB, Berti PJ. Transition state analysis of acid-catalyzed dAMP hydrolysis. J Am Chem Soc. 2007;129:7055–7064. [PubMed: 17497857]
17. Dinner AR, Blackburn GM, Karplus M. Uracil-DNA glycosylase acts by substrate autocatalysis. Nature. 2001;413:752–755. [PubMed: 11607036]

18. Przybylski JL, Wetmore SD. A QM/QM investigation of the hUNG2 reaction surface: the untold tale of a catalytic residue. Biochemistry. 2011;50:4218–27. [PubMed: 21473605]
19. Naydenova E, Rossbach S, Ochsenfeld C. QM/MM Study of the Uracil DNA Glycosylase Reaction Mechanism: A Competition between Asp145 and His148. Journal of Chemical Theory and Computation. 2019;15:4344–4350. [PubMed: 31318548]
20. Makino K, Ichikawa Y. Synthesis of a 2-deoxy-ribose type 1-N-iminosugar. Tetrahedron Lett.1998;39:8245–8248.
21. Bianchet MA, Seiple LA, Jiang YL, Ichikawa Y, Amzel LM, Stivers JT. Electrostatic guidance of glycosyl cation migration along the reaction coordinate of uracil DNA glycosylase. Biochemistry.2003;42:12455–12460. [PubMed: 14580190]
22. Jiang YL, Drohat AC, Ichikawa Y, Stivers JT. Probing the limits of electrostatic catalysis by uracil DNA glycosylase using transition state mimicry and mutagenesis. J Biol Chem. 2002;277:15385–92. [PubMed: 11859082]
23. Drohat AC, Stivers JT. NMR evidence for an unusually low N1 pKa for uracil bound to uracil DNA glycosylase: Implications for catalysis. J Am Chem Soc. 2000;122:1840–1841.
24. Parikh SS, Mol CD, Slupphaug G, Bharati S, Krokan HE, Tainer JA. Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. EMBO J. 1998;17:5214–26. [PubMed: 9724657]
25. Lee S, Verdine GL. Atomic substitution reveals the structural basis for substrate adenine recognition and removal by adenine DNA glycosylase. Proc Natl Acad Sci U S A. 2009;106:18497–502. [PubMed: 19841264]
26. Hollis T, Ichikawa Y, Ellenberger T. DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, Escherichia coli AlkA. EMBO J. 2000;19:758–66. [PubMed: 10675345]
27. Woods RD, O'Shea VL, Chu A, Cao S, Richards JL, Horvath MP, David SS. Structure and stereochemistry of the base excision repair glycosylase MutY reveal a mechanism similar to retaining glycosidases. Nucleic Acids Res. 2016;44:801–10. [PubMed: 26673696]
28. Russelburg LP, O'Shea Murray VL, Demir M, Knutsen KR, Sehgal SL, Cao S, David SS, Horvath MP. Structural Basis for Finding OG Lesions and Avoiding Undamaged G by the DNA Glycosylase MutY. ACS Chem Biol. 2020;15:93–102. [PubMed: 31829624]
29. Hashimoto H, Zhang X, Cheng X. Excision of thymine and 5-hydroxymethyluracil by the MBD4 DNA glycosylase domain: structural basis and implications for active DNA demethylation. Nucleic Acids Res. 2012;40:8276–8284. [PubMed: 22740654]
30. Bellacosa A, Drohat AC. Role of base excision repair in maintaining the genetic and epigenetic integrity of CpG sites. DNA Repair (Amst). 2015;32:33–42. [PubMed: 26021671]
31. Hendrich B, Hardeland U, Ng HH, Jiricny J, Bird A. The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. Nature. 1999;401:301–304. [PubMed: 10499592]
32. Millar CB, Guy J, Sansom OJ, Selfridge J, MacDougall E, Hendrich B, Keightley PD, Bishop SM, Clarke AR, Bird A. Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. Science.2002;297:403–405. [PubMed: 12130785]
33. Bellacosa A, Cicchillitti L, Schepis F, Riccio A, Yeung AT, Matsumoto Y, Golemis EA, Genuardi M, Neri G. MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. Proc Natl Acad Sci U S A. 1999;96:3969–74. [PubMed: 10097147]
34. Turner DP, Cortellino S, Schupp JE, Caretti E, Loh T, Kinsella TJ, Bellacosa A. The DNA N-glycosylase MED1 exhibits preference for halogenated pyrimidines and is involved in the cytotoxicity of 5-iododeoxyuridine. Cancer Res. 2006;66:7686–93. [PubMed: 16885370]
35. Petronzelli F, Riccio A, Markham GD, Seeholzer SH, Genuardi M, Karbowski M, Yeung AT, Matsumoto Y, Bellacosa A. Investigation of the substrate spectrum of the human mismatch-specific DNA N-glycosylase MED1 (MBD4): fundamental role of the catalytic domain. J Cell Physiol.2000;185:473–80. [PubMed: 11056019]
36. Derrien AC, Rodrigues M, Eeckhoutte A, Dayot S, Houy A, Mobuchon L, Gardrat S, Lequin D, Ballet S, Pierron G, Alsafadi S, Mariani O, El-Marjou A, Matet A, Colas C, Cassoux N, Stern MH.

Germline MBD4 mutations and predisposition to uveal melanoma. J Natl Cancer Inst. 2020;113:80–87.

37. Johansson PA, Brooks K, Newell F, Palmer JM, Wilmott JS, Pritchard AL, Broit N, Wood S, Carlino MS, Leonard C, Koufariotis LT, Nathan V, Beasley AB, Howlie M, Dawson R, Rizos H, Schmidt CW, Long GV, Hamilton H, Kiilgaard JF, Isaacs T, Gray ES, Rolfe OJ, Park JJ, Stark A, Mann GJ, Scolyer RA, Pearson JV, van Baren N, Waddell N, Wadt KW, McGrath LA, Warrier SK, Glasson W, Hayward NK. Whole genome landscapes of uveal melanoma show an ultraviolet radiation signature in iris tumours. Nat Commun. 2020;11:2408. [PubMed: 32415113]
38. Howard JH, Frolov A, Tzeng CW, Stewart A, Midzak A, Majmundar A, Godwin A, Heslin M, Bellacosa A, Arnoletti JP. Epigenetic downregulation of the DNA repair gene MED1/MBD4 in colorectal and ovarian cancer. Cancer Biol Ther. 2009;8:94–100. [PubMed: 19127118]
39. Morera S, Grin I, Vigouroux A, Couve S, Henriot V, Saparbaev M, Ishchenko AA. Biochemical and structural characterization of the glycosylase domain of MBD4 bound to thymine and 5-hydroxymethyuracil-containing DNA. Nucleic Acids Res. 2012;40:9917–9926. [PubMed: 22848106]
40. Ouzon-Shubeita H, Jung H, Lee M, Koag MC, Lee S. Catalytic mechanism of the mismatch-specific DNA glycosylase methyl-CpG-binding domain 4. Biochem J. 2020;477:1601–1612. [PubMed: 32297632]
41. 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]
42. Scharer OD, Kawate T, Gallinari P, Jiricny J, Verdine GL. Investigation of the mechanisms of DNA binding of the human G/T glycosylase using designed inhibitors. Proc Natl Acad Sci U S A. 1997;94:4878–4883. [PubMed: 9144158]
43. Chepanoske CL, Porello SL, Fujiwara T, Sugiyama H, David SS. Substrate recognition by Escherichia coli MutY using substrate analogs. Nucleic Acids Res. 1999;27:3197–204. [PubMed: 10454618]
44. Stivers JT, Pankiewicz KW, Watanabe KA. Kinetic mechanism of damage site recognition and uracil flipping by Escherichia coli uracil DNA glycosylase. Biochemistry. 1999;38:952–63. [PubMed: 9893991]
45. Maiti A, Morgan MT, Drohat AC. Role of two strictly conserved residues in nucleotide flipping and N-glycosylic bond cleavage by human thymine DNA glycosylase. J Biol Chem. 2009;284:36680–36688. [PubMed: 19880517]
46. Barrett TE, Scharer OD, Savva R, Brown T, Jiricny J, Verdine GL, Pearl LH. Crystal structure of a thwarted mismatch glycosylase DNA repair complex. EMBO J. 1999;18:6599–6609. [PubMed: 10581234]
47. Coey CT, Malik SS, Pidugu LS, Varney KM, Pozharski E, Drohat AC. Structural basis of damage recognition by thymine DNA glycosylase: Key roles for N-terminal residues. Nucleic Acids Res. 2016;44:10248–10258. [PubMed: 27580719]
48. Pidugu LS, Flowers JW, Coey CT, Pozharski E, Greenberg MM, Drohat AC. Structural Basis for Excision of 5-Formylcytosine by Thymine DNA Glycosylase. Biochemistry. 2016;55:6205–6208. [PubMed: 27805810]
49. Pidugu LS, Dai Q, Malik SS, Pozharski E, Drohat AC. Excision of 5-Carboxylcytosine by Thymine DNA Glycosylase. J Am Chem Soc. 2019;141:18851–18861. [PubMed: 31693361]
50. Parikh SS, Walcher G, Jones GD, Slupphaug G, Krokan HE, Blackburn GM, Tainer JA. Uracil-DNA glycosylase-DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. Proc Natl Acad Sci USA. 2000;97:5083–5088. [PubMed: 10805771]
51. Mol CD, Arvai AS, Begley TJ, Cunningham RP, Tainer JA. Structure and activity of a thermostable thymine-DNA glycosylase: evidence for base twisting to remove mismatched normal DNA bases. J Mol Biol. 2002;315:373–84. [PubMed: 11786018]
52. Manvilla BA, Maiti A, Begley MC, Toth EA, Drohat AC. Crystal Structure of Human Methyl-Binding Domain IV Glycosylase Bound to Abasic DNA. J Mol Biol. 2012;420:164–175. [PubMed: 22560993]

53. Waters TR, Swann PF. Kinetics of the action of thymine DNA glycosylase. J Biol Chem. 1998;273:20007–20014. [PubMed: 9685338]
54. Dow BJ, Malik SS, Drohat AC. Defining the Role of Nucleotide Flipping in Enzyme Specificity Using (19)F NMR. J Am Chem Soc. 2019;141:4952–4962. [PubMed: 30841696]
55. Bennett MT, Rodgers MT, Hebert AS, Ruslander LE, Eisele L, Drohat AC. Specificity of Human Thymine DNA Glycosylase Depends on N-Glycosidic Bond Stability. J Am Chem Soc. 2006;128:12510–12519. [PubMed: 16984202]
56. Doig AJ, Baldwin RL. N- and C-capping preferences for all 20 amino acids in alpha-helical peptides. Protein Sci. 1995;4:1325–36. [PubMed: 7670375]
57. Suzuki N, Emura T, Fukushima M. Mode of action of trifluorothymidine (TFT) against DNA replication and repair enzymes. Int J Oncol. 2011;39:263–70. [PubMed: 21491084]
58. Hansch C, Leo A, Taft RW. A Survey of Hammett Substituent Constants and Resonance and Field Parameters. Chem Rev. 1991;91:165–195.
59. Petronzelli F, Riccio A, Markham GD, Seeholzer SH, Stoerker J, Genuardi M, Yeung AT, Matsumoto Y, Bellacosa A. Biphasic kinetics of the human DNA repair protein MED1 (MBD4), a mismatch-specific DNA N-glycosylase. J Biol Chem. 2000;275:32422–9. [PubMed: 10930409]
60. Carugo O Buried chloride stereochemistry in the Protein Data Bank. BMC Struct Biol. 2014;14:19. [PubMed: 25928393]
61. Thayer MM, Ahern H, Xing D, Cunningham RP, Tainer JA. Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. EMBO J. 1995;14:4108–20. [PubMed: 7664751]
62. Seeberg E, Eide L, Bjor~as M. The base excision repair pathway. Trends Biochem Sci. 1995;20:3917.
63. Doherty AJ, Serpell LC, Ponting CP. The helix-hairpin-helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA. Nucleic Acids Res. 1996;24:2488–97. [PubMed: 8692686]
64. Pelletier H, Sawaya MR, Wolfle W, Wilson SH, Kraut J. Crystal structures of human DNA polymerase beta complexed with DNA: implications for catalytic mechanism, processivity, and fidelity. Biochemistry. 1996;35:12742–61. [PubMed: 8841118]
65. Pelletier H, Sawaya MR. Characterization of the metal ion binding helix-hairpin-helix motifs in human DNA polymerase beta by X-ray structural analysis. Biochemistry. 1996;35:12778–87. [PubMed: 8841120]
66. Zheng H, Chruszcz M, Lasota P, Lebioda L, Minor W. Data mining of metal ion environments present in protein structures. J Inorg Biochem. 2008;102:1765–76. [PubMed: 18614239]
67. Zheng H, Chordia MD, Cooper DR, Chruszcz M, Muller P, Sheldrick GM, Minor W. Validation of metal-binding sites in macromolecular structures with the CheckMyMetal web server. Nat Protoc. 2014;9:156–70. [PubMed: 24356774]
68. Ly HD, Withers SG. Mutagenesis of glycosidases. Annu Rev Biochem. 1999;68:487–522. [PubMed: 10872458]
69. Malik SS, Coey CT, Varney KM, Pozharski E, Drohat AC. Thymine DNA glycosylase exhibits negligible affinity for nucleobases that it removes from DNA. Nucleic Acids Res. 2015;43:9541–9552. [PubMed: 26358812]
70. Chu AM, Fettinger JC, David SS. Profiling base excision repair glycosylases with synthesized transition state analogs. Bioorg Med Chem Lett. 2011;21:4969–72. [PubMed: 21689934]
71. Morgan MT, Bennett MT, Drohat AC. Excision of 5-halogenated uracils by human thymine DNA glycosylase: Robust activity for DNA contexts other than CpG. J Biol Chem. 2007;282:27578–27586. [PubMed: 17602166]
72. Xds Kabsch W. Acta Crystallogr D Biol Crystallogr. 2010;66:125–32. [PubMed: 20124692]
73. Evans PR. An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr D Biol Crystallogr. 2011;67:282–92. [PubMed: 21460446]
74. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011;67:235–42. [PubMed: 21460441]

75. Karplus PA, Diederichs K. Linking crystallographic model and data quality. Science. 2012;336:1030–3. [PubMed: 22628654]

76. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ. Likelihood-enhanced fast translation functions. Acta Crystallogr D Biol Crystallogr. 2005;61:458–464. [PubMed: 15805601]

77. Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Rovers iP, Sharff A, Smart OS, Vonrhein C, Womack TO. BUSTER version 2.10.2. Cambridge, United Kingdom: Global Phasing Ltd.; 2011.

78. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr. 2012;68:352–67. [PubMed: 22505256]

79. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004;60:2126–2132. [PubMed: 15572765]

80. Painter J, Merritt EA. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallogr D Biol Crystallogr. 2006;62:439–50. [PubMed: 16552146]

81. Painter J, Merritt EA. TLSMD web server for the generation of multi-group TLS models. J Appl Crystallogr. 2006;39:109–111.

82. Coey CT, Drohat AC. Kinetic Methods for Studying DNA Glycosylases Functioning in Base Excision Repair. Methods Enzymol. 2017;592:357–376. [PubMed: 28668127]

- DNA glycosylases remove damaged bases, initiating BER, but mechanisms for many are poorly defined
- MBD4, of the HhH superfamily, repairs mutagenic G·T mispairs arising from 5-methylcytosine deamination
- New structures of enzyme-DNA complexes, at three stages of catalysis, illuminate the MBD4 mechanism
- Our structural and biochemical findings inform the role of a catalytic Asp conserved in HhH glycosylases

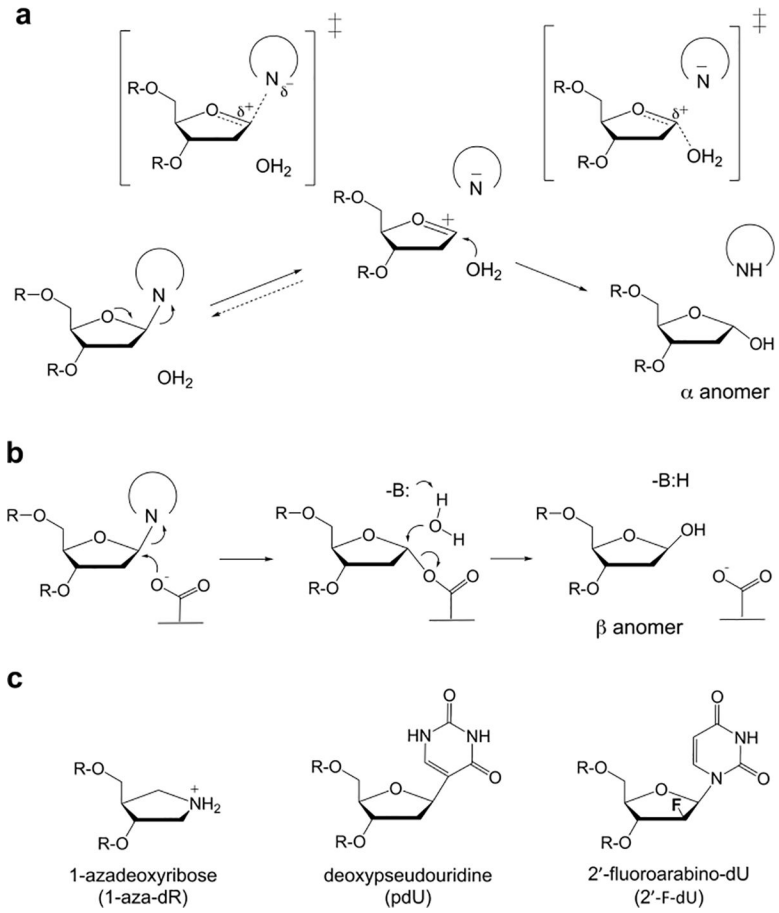


Figure 1.

Potential mechanisms for *N*-glycosyl bond hydrolysis and analogues used in the studies. (a) A direct-displacement reaction following a stepwise (S_N1) mechanism includes an oxacarbenium ion intermediate and two transition states (TS). The AP product exhibits inversion of configuration at the anomeric center (α-anomer). (b) A double-displacement reaction whereby a carboxylate (Asp or Glu) forms a glycosyl-enzyme intermediate that is hydrolyzed via general base catalysis, giving retention of configuration (β-anomer). (c) Structures of 1-azadeoxyribose and two deoxyuridine analogues.

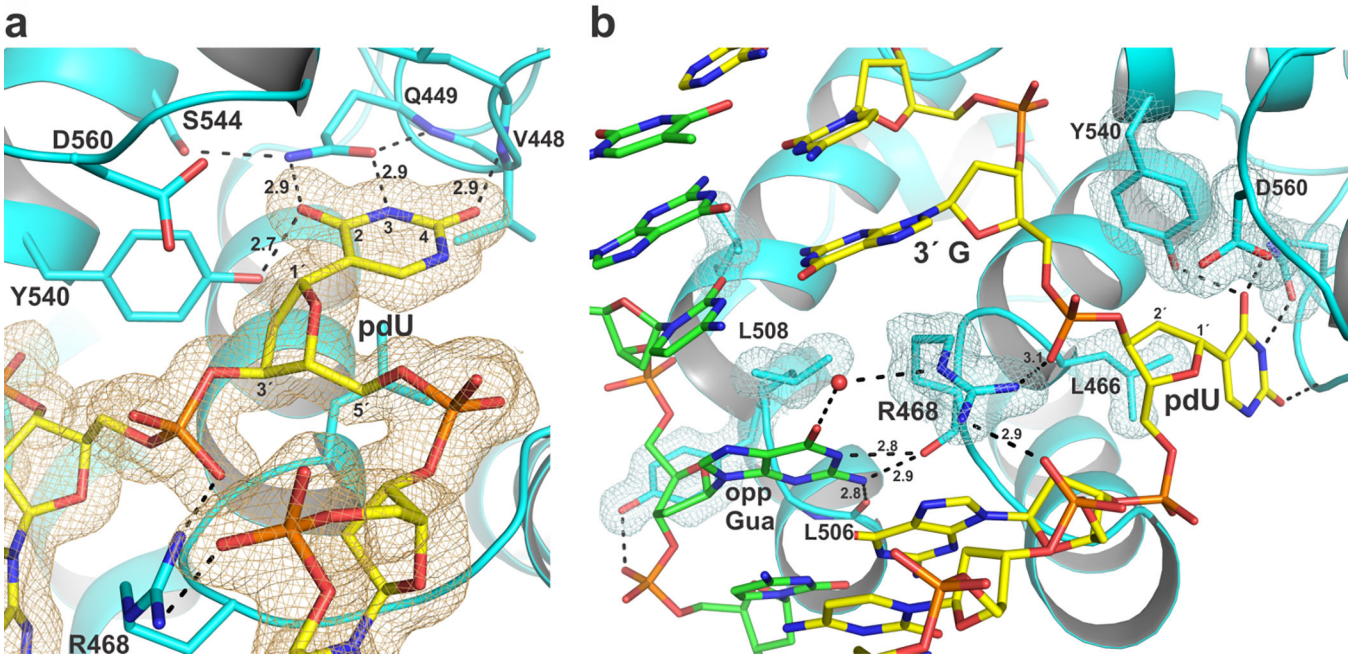


Figure 2.

Structure of the MBD4 enzyme-substrate complex with DNA containing pdU. (a) Closeup view of the pdU substrate analog flipped into the active site. The $2F_o$-F_c electron density map, contoured at 1.0 σ, is shown for DNA. Dashed lines represent hydrogen bonds; for some, interatomic distances are provided (Å). To enhance clarity, water molecules are not shown. (b) A different view shows R468 penetrating the space vacated by nucleotide flipping, and MBD4 contacts to the mismatched Gua ("opp Gua"). The $2F_o$-F_c electron density map, contoured at 1.0 σ, is shown for MBD4 side chains.

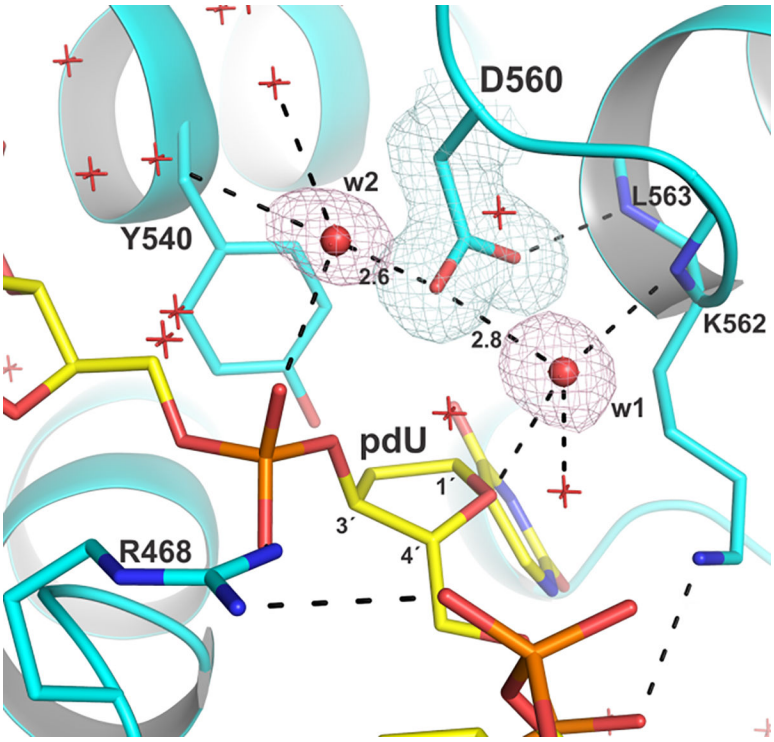


Figure 3.

Interactions involving the conserved Asp (D560) in the MBD4 enzyme-substrate complex with pdU. The $2F_o$-F_c electron density map, contoured at 1.0 σ, is shown for D560 and two water molecules that contact it (red spheres; w1, w2). The distance of these waters from C1′ of pdU is 3.6 Å (w1) and 5.6 Å (w2).

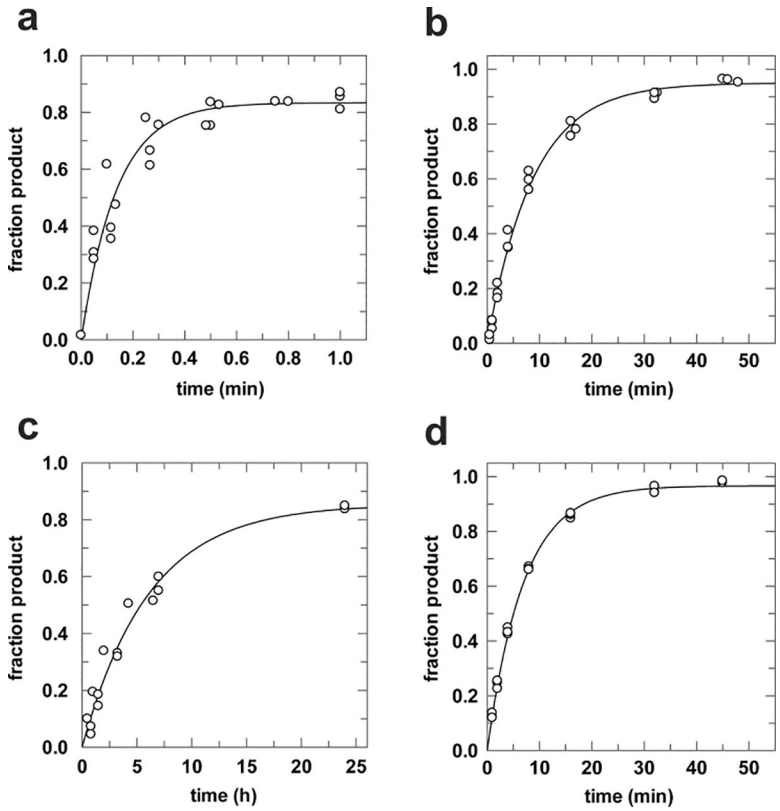


Figure 4.

Activity of MBD4 or D560G-MBD4 determined by single turnover kinetics experiments. (**a**) MBD4 excises trifluorothymine (T^{F3}) from a G·T^{F3} pair with a rate constant of k_{obs} = 7.3 ± 0.8 min^{-1}. (**b**) MBD4 excises T from a G·T pair with a rate constant of k_{obs} = 0.11 ± 0.01 min^{-1}. (**c**) D560G-MBD4 excises T^{F3} from a G·T^{F3} substrate with a rate constant of k_{obs} = 0.0027 ± 0.0003 min^{-1}. (**d**) MBD4 excises T from a G·T pair with a rate constant of k_{obs} = 0.14 ± 0.01 min^{-1} in reaction buffer for which NaCl is replaced with KCl. Experiments were collected with 0.5 μM substrate and 5.0 μM enzyme, at 23 ºC.

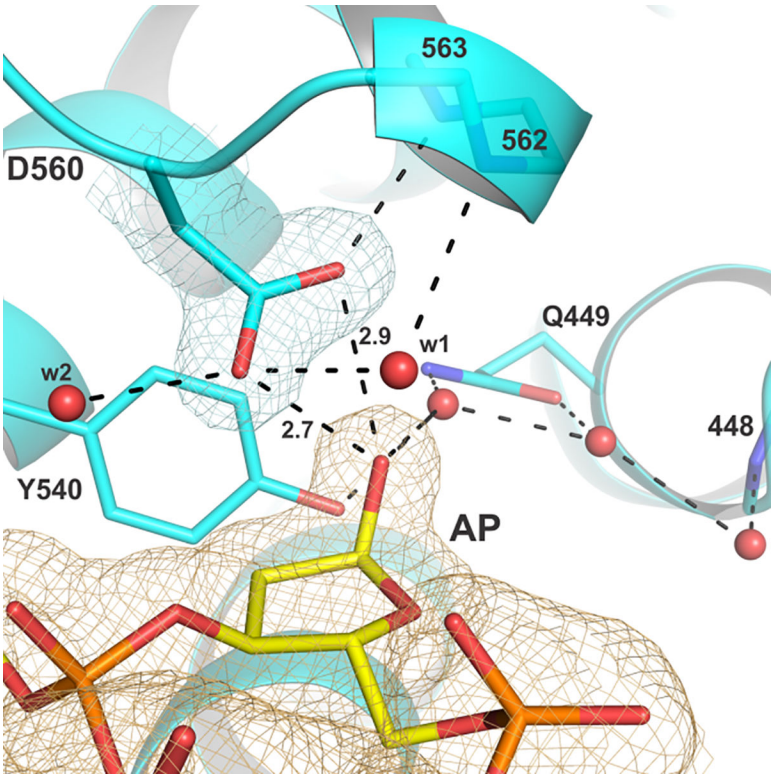


Figure 5.

New structure of the MBD4 enzyme-product complex with a focus on the abasic sugar and D560. The $2F_o$-F_c electron density map, contoured at 1.0 σ, is shown for D560 and DNA.

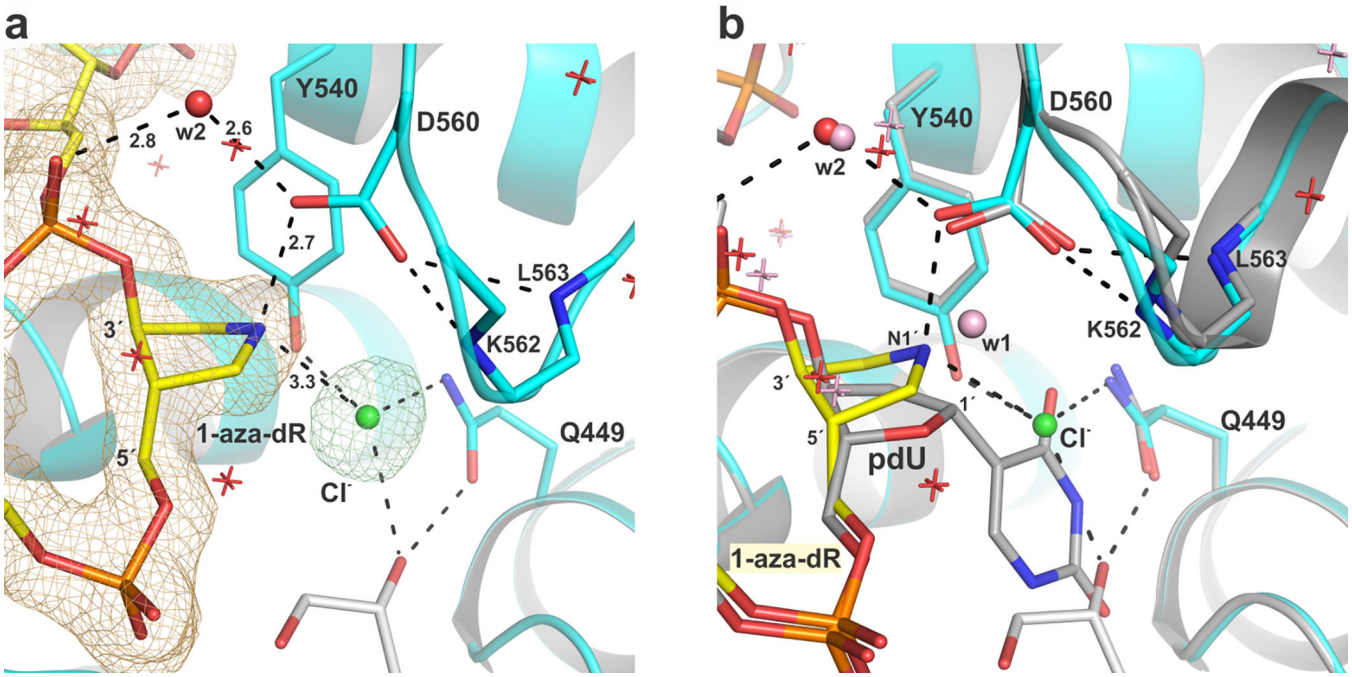


Figure 6.

Structure of DNA-bound MBD4 with 1-aza-dR flipped into its active site gives a mimic of the putative E·I complex (PDB ID: 7KZG). (a) Closeup view of active site showing interactions involving 1-aza-dR and D560. The water molecule (w2) contacting D560 and a backbone phosphate is shown as a red sphere and other waters are red crosses. The $2F_o$-F_c electron density map (contoured at 1.0 σ) is shown for DNA containing 1-aza-dR and the Cl ion (green). A solvent molecule is shown in white (modeled as glycerol but electron density suggests possibility for partial occupancy by water molecules). (b) Structural alignment of our models for the E·I complex (1-aza-dR) and the E·S complex (pdU). The electrostatic interactions and coloring for the E·I complex are as shown in panel 5a; coloring for the E·S complex is grey for MBD4 and DNA and pink for water molecules. Two water molecules (w1, w2) contact D560 in the E·S complex while only one of these (w2) does so in the E·I complex.

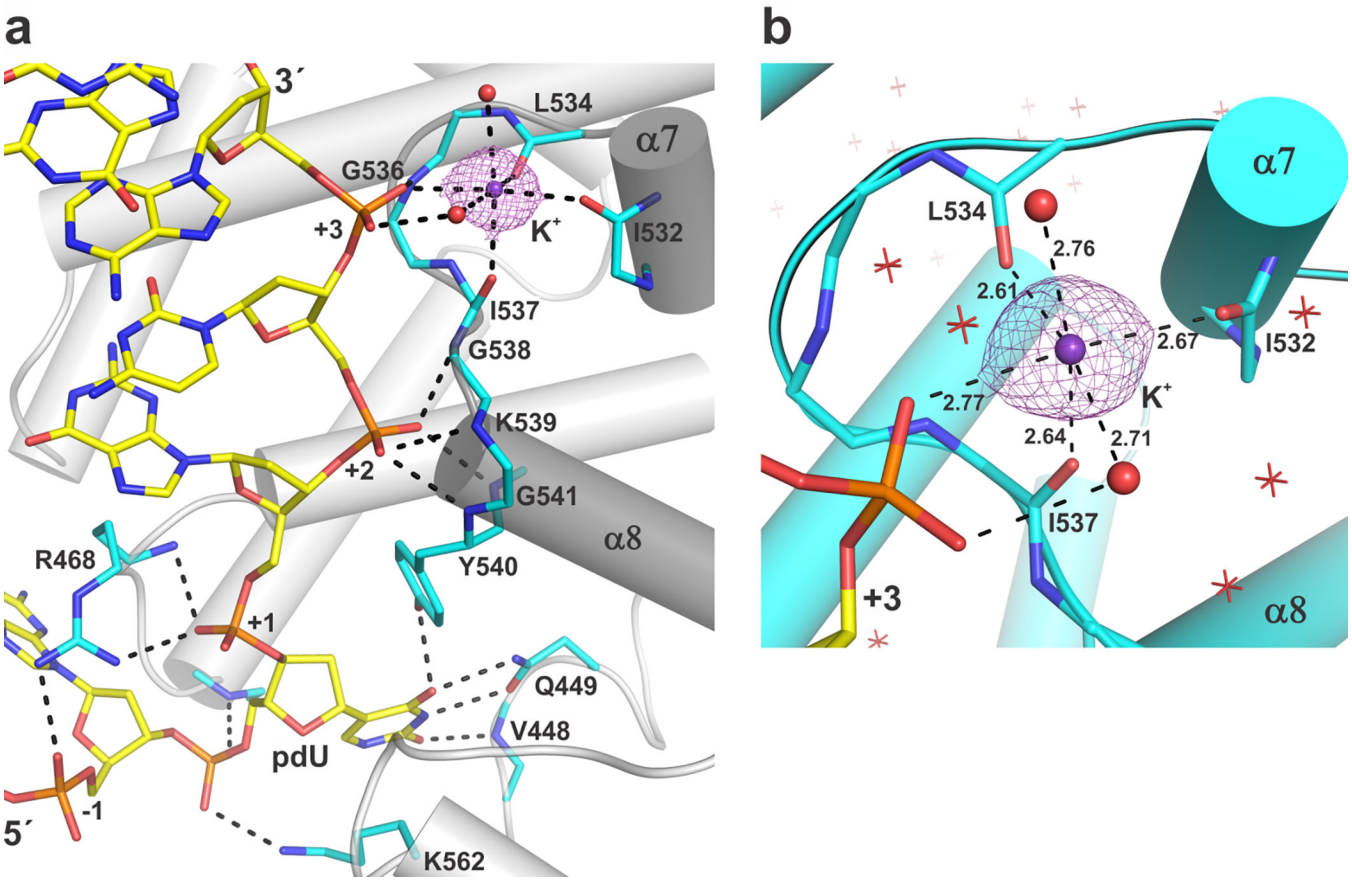


Figure 7.

Structure of the helix-hairpin-helix (HhH) motif as observed in the enzyme-substrate complex for MBD4 reported here. (a) Overview of the DNA contacts provided by the HhH motif and other regions of MBD4, for the DNA strand containing the flipped nucleotide. (b) Closeup view of metal ion (K^+) coordination by the HhH motif, with coordination distances shown (Å). In both panels, the two helices comprising the HhH motif are labeled (α7, α8), and the $2F_o$-F_c electron density map, contoured at 1.0 α, is shown for the K^+ ion. Two water molecules contacting K^+ are shown as red spheres; others waters are red crosses (panel b only). DNA backbone phosphates are numbered with respect to the flipped nucleotide.