

BEAMLINE

X25

PUBLICATION

A. Banerjee, W. Santos and G.L. Verdine, "Structure of a DNA glycosylase Searching for Lesions," *Science*, **311**, 1153-1157 (2006).

FUNDING

National Institutes of Health

FOR MORE INFORMATION

Gregory L. Verdine, Department of Chemistry, Harvard University
gregory_verdine@harvard.edu

Disulfide Trapped Structure of a Repair Enzyme Interrogating Undamaged DNA Sheds Light on Damaged DNA Recognition

A. Banerjee^{1,3}, W. Santos¹ and G.L. Verdine^{1,2}

¹Departments of Chemistry and Chemical Biology and ²Molecular and Cellular Biology, Harvard University; ³Current address: Rockefeller University

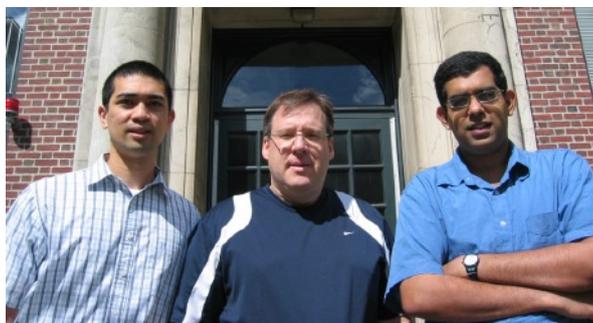
Spontaneous changes to the covalent structure of DNA are a threat that living systems battle constantly. Repairing the resulting lesions is initiated by enzymes known as DNA glycosylase enzymes, which extrude damaged nucleosides from the helical stack, insert them into an extrahelical active site pocket, and catalyze cleavage of their glycosidic bond. It is poorly understood how DNA glycosylases search the genome to locate the exceeding rare sites of damage embedded in a vast excess of normal DNA. Through the use of intermolecular disulfide crosslinking and x-ray crystallography, we have trapped and characterized in atomic detail the initial complex formed between a bacterial 8-oxoguanine (oxoG) DNA glycosylase, MutM, and undamaged DNA. The structures represent the earliest stage of DNA interrogation by a DNA glycosylase and suggest a kinetic mode of discriminating normal nucleobases from damaged ones.

The oxidation of guanine by escaped intermediates in aerobic respiration generates 8-oxoguanine, a potent endogenous mutagen that causes G:C to T:A transversion mutations. Repairing this lesion is initiated by 8-oxoguanine DNA glycosylase, MutM in bacteria, and the structurally unrelated OGG1 in eukaryotes. The mechanistic and structural details of oxoG recognition by MutM and human OGG1 (hOGG1) have been studied exhaustively (1, 2). Both enzymes bend DNA drastically at the site of damage and extrude the substrate oxoG from the helix into the extrahelical enzyme active site (**Figure 1a**); a similar strategy is used by DNA glycosylases. In previous work, we reported the use of intermolecular disulfide crosslinking (DXL) to trap and then characterize a complex in which hOGG1 was interrogating an undamaged extrahelical guanine nucleobase in DNA (3). Here we extend DXL technology to include crosslink attachment to the DNA backbone, and we use this chemistry to trap complexes of MutM with undamaged DNA having no extrahelical nucleobase. The trapping chemistry is based

on proximity-directed disulfide bond formation between backbone-modified DNA containing a disulfide linker attached via a phosphoramidite linkage and a cysteine residue engineered into the protein (**Figure 1b**). Of the three different positions we tested on MutM and DNA for crosslinking, the Q166C/p6 combination emerged as the one that underwent the most efficient crosslinking, hence we chose this position for further structural work. By varying the sequence of DNA and the position of crosslinking, we solved the structures of three different complexes of MutM crosslinked to undamaged DNA, namely interrogation complexes 1, 2, and 3 (IC1, IC2, and IC3, **Figures 2, 3**). The structures of all the complexes reveal, for the first time, a DNA

glycosylase interrogating a fully intrahelical base-paired duplex for damaged sites. Although the base-pairing in the duplex is intact, MutM severely bends the DNA and actively distorts the interrogated base-pair (also referred to as the sampled register) by inserting the aromatic ring of Phe114 in the duplex, thereby severely buckling the base-pair in question.

The sequence of the DNA in IC1 was designed to unravel the details of MutM interrogating a G:C base-pair being crosslinked to p6. However, when we solved the structure of IC1, it became clear that instead of sampling the G:C base-pair at position 8 (the +2 register), MutM jumps by one base-pair position along the duplex and instead samples the A:T base-pair at position 9 (register +3), evincing enough flexibility allowed by the disulfide crosslink to allow MutM to adopt the position of choice (**Figure 2**). In an attempt to coax MutM to sample a G:C base-pair, we moved the crosslinking position from p6 to p5. However, the structure of the resulting complex revealed that MutM continues to sample



Authors (from left): Webster Santos, Gregory L. Verdine, and Anirban Banerjee

the A:T base-pair at position 9 at the cost of inducing a pronounced distortion in the DNA helical structure (not described in this work). Next, by keeping the crosslinking position fixed at p5 and switching bp8 to an A:T, the resulting complex IC2 was found to adopt the same overall conformation seen in IC1, except now sampling the A:T base-pair at position 8 (**Figure 2**). Thus, MutM prefers to sample an A:T base-pair over G:C and the +3 sampling register over +2 or +4.

In IC1, IC2, and IC3, the target base-pair is intact but severely buckled, with Phe114 inserted deeply into the helical stack on the 3'-side of the nucleobase targeted for extrusion. The structure differs in other significant ways from that of MutM bound to an extrahelical oxoG lesion (lesion-recognition

complex, or LRC). In the LRC, Arg112 occupies part of the space vacated by the extruded oxoG and makes specific interactions with the partner C (**Figure 3b**). In the ICs, Arg112 curls down and makes non-specific backbone contacts with the DNA (**Figure 3b**). MutM makes numerous direct contacts to the DNA phosphate backbone in the LRC and a number of these contacts are replaced by water-mediated contacts in the ICs. It has now been shown by single-molecule studies that both MutM and hOGG1 rapidly slide along DNA with a low overall barrier for translocation from one base-pair to the next. Presumably these water molecules play the role of "lubricating" the protein-DNA interface to allow fast translocation of MutM along undamaged DNA, as is required by a fast search process. The fast

search process coupled to the severe distortion caused by the probe residue, Phe114 (**Figure 3b**), suggests a kinetic mode of discrimination between normal and damaged base-pairs where the distortion serves as a means for enhancing selective extrusion of damaged bases. Although structures of other DNA glycosylases with undamaged DNA are not available, a survey of existing structures of LRCs reveal a similar probe residue, mostly aromatic in nature, involved in severe local distortion of DNA by the disruption of local helical π -stacking. Thus, DNA interrogation by an intercalating probe residue akin to Phe114 may be a general strategy employed by diverse DNA glycosylases in their search for lesions in the genome.

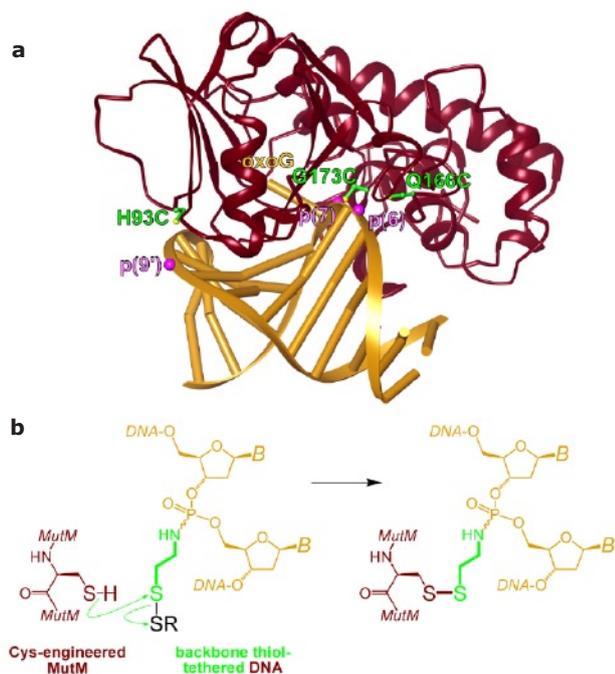


Figure 1. Experimental strategy employed to capture MutM interrogating undamaged DNA. **a.** Structure of the lesion-recognition complex (LRC) of MutM (crimson) bound to oxoG-containing DNA (gold). The side-chains of Cys residues introduced individually are shown in green, with nearby backbone phosphates shown in magenta. **b.** Structure of the N-thioalkyl phosphoramidate moiety substituted for a backbone phosphate (thiol-containing tether in green, DNA backbone in yellow); wavy line indicates a mixture of two diastereomers at phosphorus. Also shown is the chemical structure of the starting materials and products of the crosslinking reaction; curved arrows denote electron flow in the crosslinking reaction.

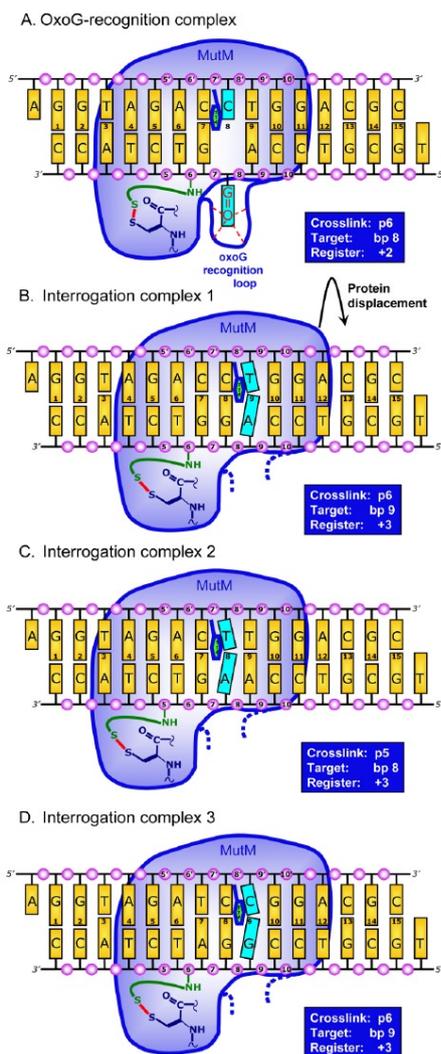


Figure 2. Schematic representation of Q166C MutM-DNA complexes. **a.** The MutM lesion-recognition complex (LRC) used as the basis for the design of the crosslinking system. **b. – d.** Interrogation complexes showing the positioning of MutM over the DNA duplex, with the target base-pair in aqua. The side-chain of the helix-probe residue Phe114 is indicated. The numbering system for the base-pairs and backbone phosphates is as indicated. The curved green line denotes the thiol-bearing tether engaged in a crosslink to Cys¹⁶⁶. Each blue box indicates the site of tether attachment to DNA, the position of the target base-pair, and the separation between them, here referred to as the register. Dashed blue lines indicate the lack of order in the oxoG-recognition loop.

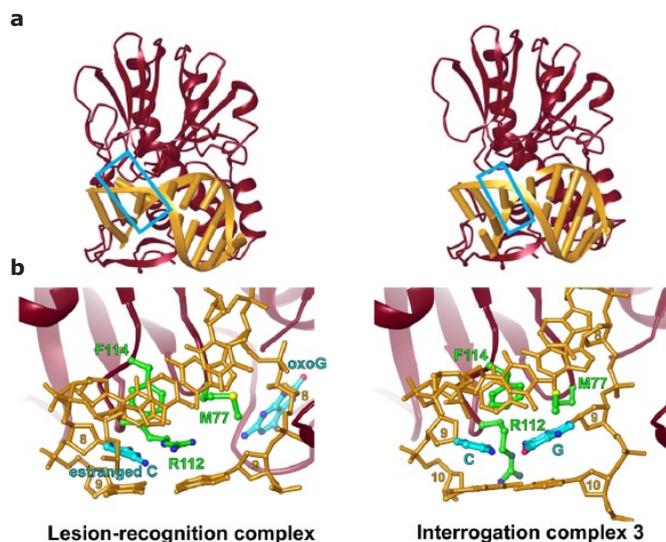


Figure 3. Comparison of the lesion recognition complex with interrogation complex 3. **a.** Overall structures of the complexes. **b.** Close-up view of key interactions between the target base-pair and MutM. The protein backbone is in a crimson ribbon representation, with the side-chain of key residues shown in green. The DNA is in gold. The target base-pair is shown with a box in **a** and is colored aqua in **b**.