# Structural and Biochemical Analysis of DNA Helix Invasion by the Bacterial 8-Oxoguanine DNA Glycosylase MutM

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 288, NO. 14, pp. 10012–10023, April 5, 2013


# Structural and Biochemical Analysis of DNA Helix Invasion by the Bacterial 8-Oxoguanine DNA Glycosylase MutM*[S]

Received for publication, August 31, 2012, and in revised form, January 26, 2013 Published, JBC Papers in Press, February 12, 2013, DOI 10.1074/jbc.M112.415612

**Rou-Jia Sung**[‡§1], **Michael Zhang**[¶1,2], **Yan Qi**[‖3], **and Gregory L. Verdine**[‡§¶**4]

*From the Departments of [‡]Molecular and Cellular Biology, [¶]Chemistry and Chemical Biology, and [§]Stem Cell and Regenerative Biology, Harvard University, Cambridge, Massachusetts 02138, the **Chemical Biology Initiative and Program in Cancer Chemical Biology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, and the [‖]Graduate Program in Biophysics, Harvard Medical School, Boston, Massachusetts 02115*

**Background:** The role of helix-invading residues during MutM-catalyzed oxoG repair remains unclear.
**Results:** Mutation of Phe$^{114}$ and Met$^{77}$ to alanine destabilizes the extrahelical conformation of oxoG and decreases efficiency of oxoG repair.
**Conclusion:** Helix invasion by Phe$^{114}$ and Met$^{77}$ is important for promoting and stabilizing extrusion of oxoG during repair.
**Significance:** Helix invasion constitutes a critical step during the base extrusion process.

**MutM is a bacterial DNA glycosylase that serves as the first line of defense against the highly mutagenic 8-oxoguanine (oxoG) lesion, catalyzing glycosidic bond cleavage of oxoG to initiate base excision DNA repair. Previous work has shown that MutM actively interrogates DNA for the presence of an intrahelical oxoG lesion. This interrogation process involves significant buckling and bending of the DNA to promote extrusion of oxoG from the duplex. Structural snapshots have revealed several different highly conserved residues that are prominently inserted into the duplex in the vicinity of the target oxoG before and after base extrusion has occurred. However, the roles of these helix-invading residues during the lesion recognition and base extrusion process remain unclear. In this study, we set out to probe the function of residues Phe$^{114}$ and Met$^{77}$ in oxoG recognition and repair. Here we report a detailed biochemical and structural characterization of MutM variants containing either a F114A or M77A mutation, both of which showed significant decreases in the efficiency of oxoG repair. These data reveal that Met$^{77}$ plays an important role in stabilizing the lesion-extruded conformation of the DNA. Phe$^{114}$, on the other hand, appears to destabilize the intrahelical state of the oxoG lesion, primarily by buckling the target base pair. We report the observation of a completely unexpected interaction state, in which the target base pair is ruptured but remains fully intrahelical; this structure vividly illustrates the disruptive influence of MutM on the target base pair.**

Oxidative damage to DNA, caused primarily by escaped reactive oxygen intermediates of aerobic respiration, is among the most prevalent and deleterious forms of spontaneous genetic change (1). Numerous DNA lesions arise from reactive oxygen intermediate-induced damage, but that which has received the most attention is 7,8-dihydro-8-oxoguanine (8-oxoguanine (oxoG)),[5] a nearly perfect mutagen because of its strong preference to mispair with adenine during genome replication, which gives rise to G:C → T:A transversion mutations (2, 3). In all organisms, oxoG lesions are repaired through the action of lesion-specific DNA glycosylases that conduct surveillance of the genome to locate and excise damaged nucleobases (4, 5). oxoG repair in bacteria is carried out by the DNA repair enzyme MutM, a bifunctional DNA glycosylase/lyase that recognizes oxoG:C base pairs, catalyzing removal of both the oxoG nucleobase and its sugar moiety (Fig. 1*B*) (3). Our current understanding of oxoG recognition and repair MutM indicates that it takes place via four major sequential steps: 1) encounter between MutM and an intrahelical, fully base-paired oxoG nested within the DNA duplex; 2) extrusion of the oxoG nucleoside from the duplex and insertion into the extrahelical enzyme active site; 3) excision of the *N*-glycosidic bond via nucleophilic displacement by a nucleophilic proline residue on the enzyme (Fig. 1*B*); and 4) conjugate elimination on both sides of enzyme-linked abasic site to remove the sugar moiety, leaving behind a nick bracketed by 5′- and 3′-phosphate groups (6, 7). Considerable intrigue and interest surrounds the first two steps, because oxoG differs by only two atoms from G (Fig. 1*A*), because the lesion has negligible effects on duplex DNA structure and only modest effects on duplex stability (8–11), and because oxoG residues are embedded among a million-fold excess of undamaged G residues.

The development of structure- and mechanism-based trapping technologies has recently made it possible to characterize, at atomic resolution, the interactions of MutM with lesion-containing DNA at multiple stages of the repair cycle, including lesion searching, encounter, recognition, and extrusion. For

* This work was supported, in whole or in part, by National Institutes of Health Grants GM044853 and CA100742.
[S] This article contains supplemental text, Table S1, and Figs. S1–S5.
*The atomic coordinates and structure factors (codes 4G4R and 4G4O) have been deposited in the Protein Data Bank (http://wwpdb.org/).*
[1] These authors contributed equally to this work.
[2] Supported by a grant by the Harvard College Research Program.
[3] Supported by a predoctoral fellowship from the National Science Foundation.
[4] To whom correspondence should be addressed: 12 Oxford St., Cambridge, MA 02138. Fax: 617-495-8755; E-mail: gregory_verdine@harvard.edu.

[5] The abbreviations used are: oxoG, 8-oxoguanine; RMSD, root mean square deviation; LRC, lesion recognition complex; OCL, oxoG-capping loop; EC, encounter complex; DNA, deoxyribonucleic acid.

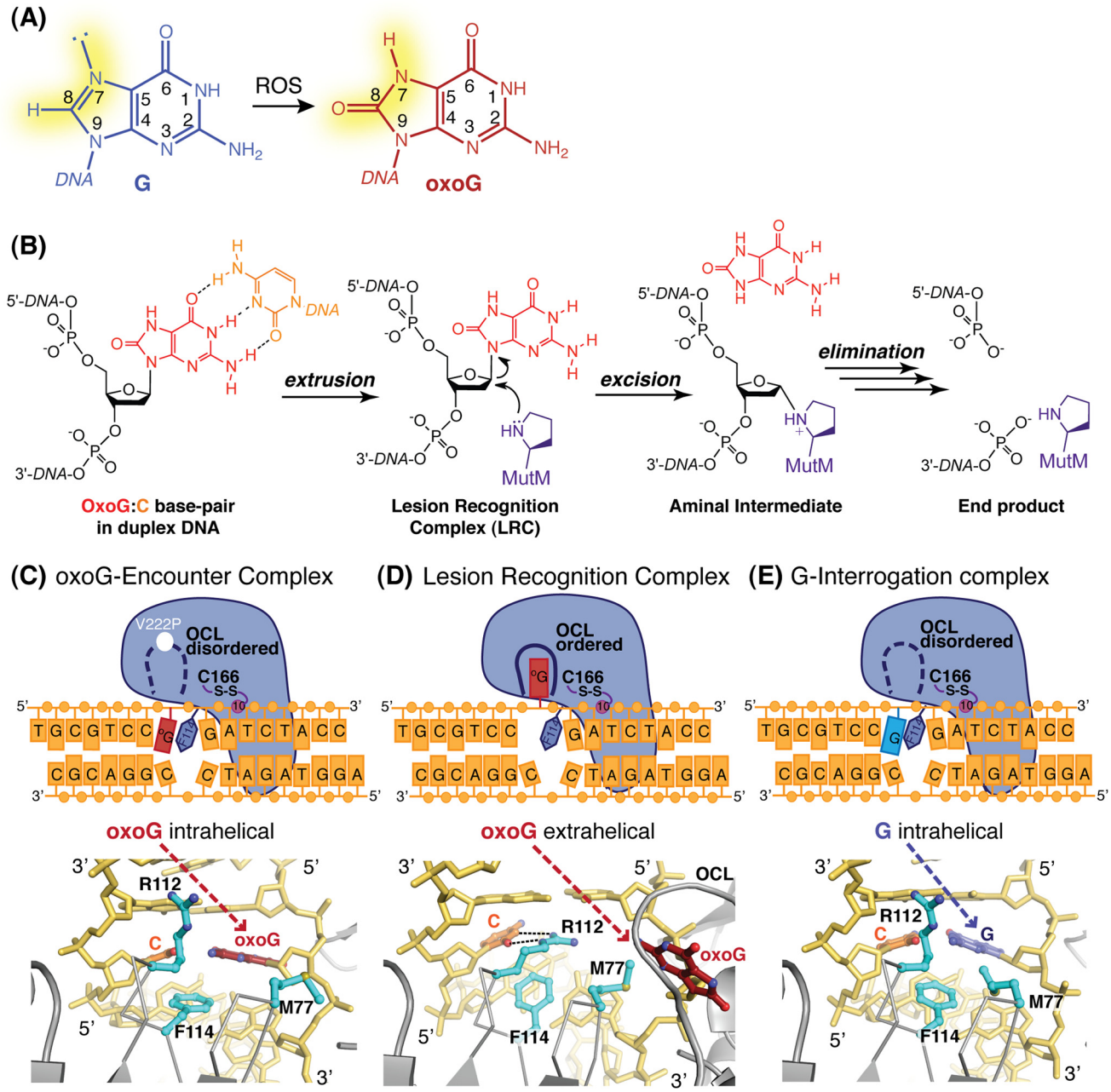



FIGURE 1. **Formation and repair of oxoG.** *A*, diagram showing differences at $N^7$ and $C^8$ for guanine *versus* oxoG. *B*, oxoG repair by MutM. First step, encounter between MutM and a fully base-paired oxoG in the DNA duplex. Second step, extrusion of oxoG into the active site and cleavage of the *N*-glycosydic bond. Third and fourth steps, Schiff base formation followed by β,δ-elimination to yield the end product-bound complex. Chemical structures of relevant intermediates are shown. *C–E*, positions of residues $Met^{77}$, $Arg^{112}$, and $Phe^{114}$ relative to the target base are shown for structures of MutM bound to an intrahelical oxoG, (*C*) oxoG that has been extruded into the active site (known as the lesion recognition complex, or LRC) (*D*), and intrahelical G (*E*). The side chains are shown as *cyan sticks*, and the rest of the protein chain is shown as *gray ribbon*. Hydrogen bonds are indicated by the *dashed lines*. The OCL is indicated by a *blue loop*.

example, mutation of the critical catalytic residue $Glu^3$ enabled structural characterization of a state poised for catalysis, specifically an enzyme-DNA complex (lesion recognition complex (LRC); Fig. 1*D*) having a fully extrahelical oxoG lesion inserted into the enzyme active site (12). More recently, disulfide cross-linking technology has been employed to stabilize and thereby structurally elucidate ordinarily fleeting intermediates at earlier stages of the repair pathway, including that of initial encounter by MutM of a fully base-paired, intrahelical oxoG, and also during interrogation of undamaged DNA (Fig. 1, *C* and *E*) (7, 12–14). Detailed comparison of the structures at these three states has revealed that MutM can detect the presence of an oxoG lesion even at the earliest stage of initial encounter, with the intrahelical oxoG inducing an altered DNA backbone conformation that encourages extrusion of the lesion from the helical stack (14). Computational simulations have revealed that this "extrudogenic conformation" indeed lowers the free energy barrier for extrusion of oxoG *versus* G. It is noteworthy, moreover, that MutM shows no thermodynamic preference to bind DNA containing an intrahelical oxoG (14, 15). Taken together, these data have led to formulation of a model wherein the enzyme locates oxoG residues by a kinetic mechanism involv-

ing accelerated extrusion of the lesion from DNA following the initial encounter (14, 16).

These structures revealed that base extrusion is mediated by three highly conserved residues: $Met^{77}$, $Arg^{112}$, and $Phe^{114}$, that form important contacts at the protein-DNA interface from the minor groove side. Each of these residues occupies a distinct position in close proximity to the target base pair. In the LRC structure (Fig. 1*D*), $Phe^{114}$ wedges into the DNA duplex between the target base pair and the base pair 3′ to the target base, inducing a bend in the global DNA conformation. $Arg^{112}$ invades the duplex to fill the space left by oxoG, forming hydrogen bonding contacts to the opposing C (Fig. 1*D*). $Met^{77}$ is positioned in the intervening space between the extruded base and the resulting gap in the duplex (Fig. 1*D*) (12). In contrast, these residues take on a decidedly different organization in the presence of an intrahelical oxoG-containing substrate (Fig. 1*C*) (14). The fully base-paired nature of the target base pair blocks the entry of $Arg^{112}$ and $Met^{77}$ into the duplex, leaving them to take up alternative positions in the minor groove. $Phe^{114}$ occupies a similar position as in the LRC state; however, insertion of the aromatic side chain into a fully duplex substrate severely buckles the aforementioned base pairs, significantly destabilizing the target base (Fig. 1*C*).

The striking changes in the arrangement of these three residues in response to the state of the target base (whether intrahelical or extrahelical) have strong implications for their role in the base extrusion process. Molecular dynamics simulations have suggested a role for $Arg^{112}$ in promoting base extrusion through direct competition with oxoG for hydrogen bonding contacts to the opposing C (14). The function of $Arg^{112}$ has been further characterized by new work observing the structural effects of an alanine substitution at $Arg^{112}$ on the interaction between MutM and DNA (17). The absence of $Arg^{112}$ results in the entrapment of a MutM-DNA complex with protein-DNA contacts distinct from those of other intrahelical structures. These structures suggest an additional role for $Arg^{112}$ in DNA translocation during lesion search, providing a novel mechanism for the movement of MutM along DNA as the enzyme probes undamaged bases in the DNA duplex. $Phe^{114}$ has also been implicated in the lesion search process, most directly by the aforementioned structural studies but also by recent biochemical results indicating that removal of the intercalating $Phe^{114}$ side chain increases the rate of diffusion of the protein along DNA (18).

The ability to trap MutM bound to oxoG- and G-containing DNA at different steps in the repair cycle has provided a powerful tool for investigating the functions of these three residues in promoting changes at the target base pair that facilitate base extrusion. Here, we report the structural and biochemical characterization of *Bacillus stearothermophilus* MutM in which the two helix-invading side chains, those on $Phe^{114}$ and $Met^{77}$, have been individually removed by mutation to Ala. These structures provide direct evidence that MutM actively interrogates the DNA duplex during the search for lesions, and they shed light on the means by which MutM discourages reannealing of the extruded oxoG lesion.

## EXPERIMENTAL PROCEDURES

*Protein Expression and Purification*—Point mutations of *Bst* MutM were made using the QuikChange site-directed mutagenesis kit (Stratagene), and the mutations were confirmed by sequencing. Expression and purification of the mutant proteins were carried using the wild-type protocol.

Mutant MutM proteins were overexpressed and purified as previously described (14). MutM was expressed in *Escherichia coli* BL21 (DE3) plysS cells and grown at 37 °C until $A_{600}$ reached 0.5–0.7. Protein expression was induced by addition of 0.5 mM isopropyl-β-D-thiogalactopyranoside (Invitrogen) and 0.05 mM $ZnCl_2$. The cells were allowed to grown for 4–5 h at 30 °C, harvested by ultracentrifugation, resuspended in 50 mM $NaPO_4$, pH 8.0, 500 mM NaCl, and 0.1% β-mercaptoethanol, flash frozen in liquid nitrogen, and stored at −80 °C.

Thawed cells were supplemented with 1 mM PMSF and 1 Complete, EDTA-free Protease inhibitor mixture tablet (Roche Applied Science), lysed by sonication, and clarified by centrifugation at 14,000 rpm for 20 min. The clarified lysate was diluted 1:5 with buffer A (20 mM Tris, pH 7.4, and 0.1% β-mercaptoethanol) and loaded onto a 20-ml Sepharose Fast Flow column (GE Healthcare). The protein was purified using a linear gradient from 0.1–0.5 M NaCl over 10 column volumes, followed by size exclusion chromatography (Superdex75; GE Healthcare) into 1 M NaCl, 10 mM Tris, pH 7.4, 5 mM β-mercaptoethanol. The unused protein was stored by adding glycerol to 20% final concentration, flash frozen, and stored at −80 °C.

*DNA Synthesis and Purification*—All of the DNA substrates were synthesized using solid phase synthesis on an ABI 392 DNA synthesizer using standard reagents and protocols (Applied Biosystems and Glen Research). For the cross-linker-modified oligonucleotides, the protocol was modified to include a 3′ H-phosphonate (Glen Research) in the coupling step at the site of the backbone modification. Following coupling, the disulfide tether was incorporated onto the backbone by an oxidation step with carbon tetracholoride (Sigma-Aldrich) and the diamine disulfide (free base). All of the oligonucleotides were purified according to the same protocol. The oligonucleotides were deprotected and cleaved from the CPG resin using ammonium hydroxide at 55 °C for 8–10 h. The ammonium hydroxide was removed using a SpeedVac, and the oligonucleotides were purified using the crush and soak method (19). The oligonucleotides were desalted using hydrophobic chromatography (Sep-Pak columns; Waters Chromatography) and then evaporated to dryness using a SpeedVac. The oligonucleotides were dissolved in 10 mM Tris, pH 8.0, and their appropriate masses were verified using MALDI-TOF mass spectrometry. The lesion-containing strand was annealed to the complementary strand in a 1:1.2 ratio in 1× NaTE (50 mM NaCl, 10 mM Tris-HCl, 1 mM EDTA, pH 8.0). The following DNA oligomers were used for the structures discussed in this paper, with G° designating oxoG and *x* representing the location of the phosphate group bearing the nonbridging *N*-ethylthio-tether used in generating the cross-linked complexes: oxoG-containing strand, 5′-TGCGTCCG°$AG_x$TCTACC-3′; G-containing strand, 5′-TGCGTCCG$AG_x$TCTACC-3; and the complementary strand, 5′-AGGTAGACTCGGACGC-3′.

*Preparative Cross-linking Reactions*—The complexes were made by cross-linking the corresponding DNA substrates to either Q166C MutM (for the G-containing complexes) or E3Q Q166C MutM (for the oxoG-containing complexes). Purified protein was concentrated to roughly 400 $\mu$M, and the reactions were set up using 10 $\mu$M cross-linker-containing duplex substrate with 20 $\mu$M protein in degassed cross-linking buffer (50 mM NaCl and 20 mM Tris, pH 7.4). The reactions were purged with argon and kept rotating at 4 °C for 2–3 days.

The reactions were purified using a MonoQ column (GE Healthcare) with a linear gradient from 0.1–0.6 M NaCl in buffer A (20 mM Tris, pH 7.4, degassed) over 30 column volumes. The fractions were pooled, buffer-exchanged into degassed cross-linking buffer, and concentrated to ~200–250 mM for crystallization experiments. The concentration was determined by the $A_{260}$ of the DNA in the complex.

*Crystallization, Data Collection, and Structure Determination*—Crystallization drops were set up at 4 °C using the hanging drop vapor diffusion method, in a 1:1 or 1:2 ratio of protein:DNA complex to reservoir solution. The complexes crystallized in 12–18% PEG 8000, 100 mM sodium cacodylate, pH 7.0, and 5% glycerol. Crystals appeared with 2–3 days and were allowed to grow to size over 1–2 weeks before briefly soaking in cryoprotectant solution containing 18% PEG 8000, 100 mM sodium cacodylate, pH 7.0, and 25% glycerol before flash freezing in liquid nitrogen. Diffraction data were collected at the 24ID-E beamline at Argonne Photon Source and the X29 beamline at the National Synchrotron Light Source. The data were processed using the HKL2000 program suites (20).

Protein coordinates from the isomorphous structure of MutM cross-linked to undamaged DNA (Protein Data Bank code 3GPX) were used as the initial model for refinement (14). Rigid body fitting, energy minimization, and simulated annealing in PHENIX resulted in a partial model (21). The DNA (including the disulfide tether) was built into the $F_o$-$F_c$ map, using the strong density for the cross-link to determine the register. Manual readjustment of protein side chains and DNA bases were done in COOT, followed by successive rounds of simulated annealing, energy minimization, and individual B-factor refinement (22, 23). After all visible protein and DNA atoms were built and the $R_{free}$ dropped below 28%, water molecules were added to the model using automated water picking and manual inspection of the difference map in PHENIX. Translation Liberation Screw (TLS) refinement was also included during refinement (two groups total: the protein and then both DNA strands were treated as one TLS group) (24, 25). Electron density for residues 217–237 (comprising the oxoG-capping loop) was not visible, and these residues were omitted from the model. Protein side chains were truncated in instances where density was not observed. Data collection and refinement statistics are shown in Table 1. Supplemental Table S1 lists the abbreviations used for the protein:DNA complexes described in the manuscript.

Model statistics and validation were carried out using PROCHECK and Molprobity (26, 27). Analysis of the DNA conformation including base step parameters, sugar puckers, and torsion angles were done using 3DNA (28). The figures were made using PyMOL v1.3 (29).

*oxoG and Abasic Site Cleavage Assays*—The duplex DNA substrate used for the cleavage assays was identical in sequence to the oligonucleotide used for crystallization but did not contain a *N*-ethylthio tether. The oxoG-containing strands were 5′-end-labeled using T4 polynucleotide kinase (New England Biolabs) and [$\gamma$-$^{32}$P]ATP (PerkinElmer Life Sciences) and then annealed with a 1.1-fold excess of the complementary strand. Single turnover cleavage reactions were carried out using 100 nM DNA duplex and 500 nM wild-type, F114A, or M77A MutM in a standard reaction buffer of 40 mM NaCl, and 50 mM Tris-HCl, pH 7.4, at room temperature. The data were fit using a single exponential equation to determine the first order rate constant as previously described (GraphPad Prism 5 Software) (30). Abasic site substrates were prepared by pretreating a 5′-$^{32}$P-labeled uracil-containing 16-mer duplex DNA (5′-AGCGTCCAUGTCTACC-3′) with uracil DNA glycosylase (New England Biolabs) at 37 °C for 1 h. The abasic site cleavage reaction was carried out using 100 nM pretreated DNA duplex and 100 nM wild-type, F114A, or M77A MutM in a reaction buffer of 40 mM NaCl and 50 mM Tris-HCl, pH 7.4, at room temperature. Aliquots of the reaction were removed periodically and quenched with an equal amount of 100 mM dithiothreitol in 95% formamide and 1× Tris borate/EDTA buffer, subjected to denaturing urea-PAGE, and then visualized on a phosphorimaging plate.

## RESULTS

*Biochemical Characterization of the M77A and F114A Variants of MutM*—We expressed and purified variants of MutM containing individually either the F114A or M77A mutation. These variants showed chromatographic behavior indistinguishable from that of the wild-type protein, indicating that the mutations caused no gross conformational abnormality. To probe more definitively the activities of these proteins, we measured their ability to catalyze base excision DNA repair of oxoG in a duplex oligonucleotide substrate. As shown in Fig. 2 (*A* and *B*), the rate of oxoG repair was significantly diminished for both F114A and M77A MutM. This loss of activity could in principle result either from a generalized defect in the protein or from a specific defect in catalysis of glycosidic bond cleavage, the first chemical step in the multistep reaction cascade catalyzed by MutM (Fig. 1*B*). To resolve this ambiguity, we took advantage of the fact that it is possible to circumvent the base excision step and synthetically generate a downstream intermediate in the repair cascade—the aminal intermediate—by providing MutM with a substrate containing an abasic site (Fig. 2*D*). Because all steps of the entire reaction cascade are catalyzed by a single active site (7), preservation of enzyme activity on an abasic site indicates preservation of overall active site architecture, with the catalytic defect being limited to the upstream steps of in the repair process. When the catalytic competence of the mutant enzymes was analyzed in this manner, it was found to be virtually unaffected by either the F114A or M77A mutation (Fig. 2*C*). Taken together with the oxoG repair data, we conclude that the F114A and M77A mutations specifically impair the early events in oxoG repair, namely extrusion of the lesion from DNA and excision of the oxoG glycosidic bond.

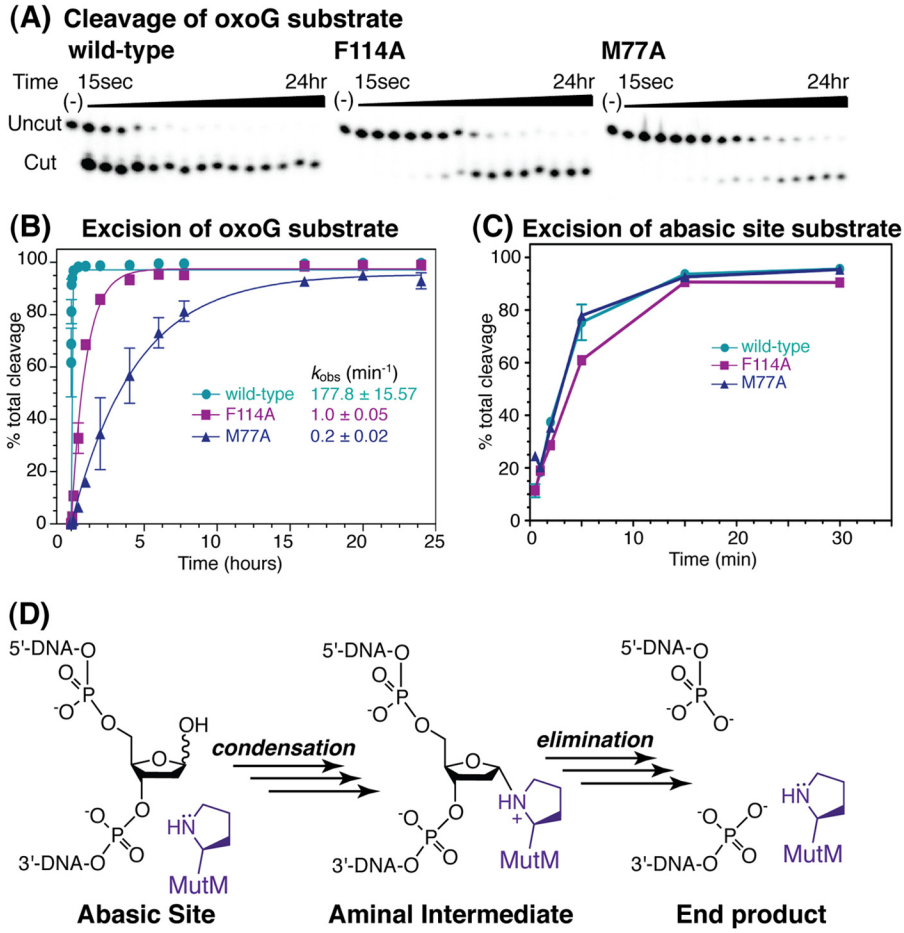



FIGURE 2. **Biochemical characterization of F114A and M77A mutations on oxoG repair.** *A*, oxoG repair assay monitoring end product formation using an oxoG-containing substrate (following the reaction pathway in Fig. 1*B*) showing significant increase in the length of time necessary for the reaction to complete. *B*, plot following product formation from *A*. Data points were measured in triplicate, with indicated *error bars*. The data were fit to a single exponential equation as previously described to determine the first order rate constant (30). *C*, quantification of end product formation using an abasic site-containing DNA substrate. Data points were measured in duplicate, with indicated *error bars*. *D*, alternative route toward the aminal intermediate, beginning with an abasic site DNA substrate.

To determine whether the removal of these key helix-invading side chains affects the strength of DNA binding by MutM, we analyzed this binding interaction using fluorescence polarization assays. Neither mutant enzyme shows any thermodynamic preference for binding to oxoG-containing DNA, suggesting that neither forms a thermodynamically stable lesion recognition complex. Previous work has shown that whereas recognition of an extrahelical oxoG lesion by MutM is driven by a thermodynamic preference to attain that state, there is no such thermodynamic preference for MutM to bind an intrahelical oxoG lesion (6, 14–16). The affinity of MutM for non-lesion-containing DNA was unaffected by introduction of the M77A mutation but was enhanced severalfold by F114A (supplemental Fig. S1, Supplemental Methods). The latter results are consistent with prior suggestions that the insertion of the Phe$^{114}$ side chain into DNA is likely to be energetically expensive (13, 14) and that this insertion is an intrinsic feature of the MutM-DNA interaction, occurring even when the protein slides along non-lesion-containing DNA (18).

*Structural Characterization of DNA-bound Complexes of F114A and M77A MutM*—To gain atomic level insights into the functions of Met$^{77}$ and Phe$^{114}$, we crystallized a total of four complexes of MutM containing either the F114A or M77A mutation, in complex with duplex DNA containing either a single, centrally located oxoG lesion or an undamaged G at the same position. All of the proteins used in these studies bore a mutation of Glu$^{3}$ to Gln (E3Q mutation), which is known to abrogate base excision while preserving lesion recognition (12). To stabilize these otherwise ephemeral complexes and restrict the roaming range of the mutant glycosylases to the desired interrogation site on DNA, we employed a well established intermolecular disulfide-cross-linking strategy, which has previously been used to capture several otherwise elusive states of the MutM-DNA interaction (13, 14, 17, 31, 32). All complexes utilized position 166 as the cross-linking site on the protein, with the partner thiol on DNA being introduced as an *N*-ethylthiol tether at phosphate 10 (p10); the target base (either oxoG or G) was located at p8. Save for the identity of the target base (either G or oxoG), all four structures employed an otherwise identical DNA molecule.

Whether a target nucleobase held in the grasp of MutM prefers to remain in the intrahelical state or transition to an extrahelical state depends upon the balance of multiple energetic influences that stabilize or destabilize each of the two states. For

**TABLE 1**
**Data collection and refinement statistics**

| | M77A-G | M77A-oG | F114A-G | F114A-oG |
|---|---|---|---|---|
| **Data collection** | | | | |
| Radiation source | APS-24-IDE | APS-24-IDE | NSLS-X29 | APS-24-IDE |
| Resolution (Å) | 50-1.85 | 50-1.95 | 50-1.86 | 50-1.95 |
| Unique reflections | 38,803 | 33,477 | 38,633 | 33,409 |
| Completeness (%)[a] | 99.9 (100.0) | 99.9 (99.9) | 99.9 (99.9) | 99.9 (99.9) |
| Redundancy[a] | 7.3 (7.3) | 7.2 (7.3) | 6.9 (6.9) | 4.8 (4.9) |
| $R_{merge}$[a,b] | 0.077 (0.381) | 0.062 (0.408) | 0.049 (0.411) | 0.056 (0.405) |
| $<I/\sigma>$[a] | 28.8 (5.7) | 25.0 (5.4) | 38.9 (3.5) | 19.9 (3.6) |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ |
| Unit cell dimensions | $a = 45.40$, $b = 93.45$, $c = 104.76$ | $a = 45.19$, $b = 93.54$, $c = 104.65$ | $a = 45.37$, $b = 95.50$, $c = 103.10$ | $a = 45.39$, $b = 93.14$, $c = 105.52$ |
| **Refinement and model** | | | | |
| Resolution (Å) | 32.7-1.85 | 32.7-1.95 | 32.7-1.86 | 37.9-1.95 |
| $R_{work}$ (%)[a,c] | 16.9 (18.1) | 17.5 (19.1) | 17.9 (18.3) | 17.2 (17.7) |
| $R_{free}$ (%)[a,c] | 19.6 (23.4) | 21.1 (24.4) | 19.8 (18.2) | 21.4 (20.9) |
| Mean *B*-factors | | | | |
| Protein | 23.96 | 28.23 | 31.08 | 27.05 |
| Water | 35.98 | 38.32 | 40.95 | 37.06 |
| RMSD from ideality | | | | |
| Bond lengths (Å) | 0.007 | 0.007 | 0.007 | 0.007 |
| Bond angles (°) | 1.092 | 1.107 | 1.062 | 1.085 |
| Ramachandran plot (%)[d] | | | | |
| Most favored | 94.4 | 93.9 | 93.0 | 94.8 |
| Additionally allowed | 5.1 | 5.6 | 6.5 | 4.7 |
| Generously allowed | 0.5 | 0.5 | 0.5 | 0.5 |
| **Protein Data Bank code** | 4G4N | 4G4O | 4G4Q | 4G4R |

[a] The values in parentheses refer to the highest resolution shell.
[b] $R_{merge} = \Sigma|I - <I>|/\Sigma<I>$, where $I$ is the observed intensity.
[c] $R_{work} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where $F_o$ and $F_c$ are the observed and calculated structure factor amplitudes, respectively. $R_{free}$ was calculated based on 5% data randomly selected and omitted throughout structure refinement (34).
[d] The values were calculated using PROCHECK (27).

example, the so-called oxoG-capping loop (OCL; Fig. 1*D*) on MutM contributes extensive protein-DNA contacts to the Hoogsteen face of extrahelical oxoG; hence the OCL is considered to stabilize the extrahelical state (12). Consistent with this notion, mutations in the OCL, such as V222P, that alter even subtly its interaction with extrahelical oxoG abolish crystallographic observation of the extrahelical state and furnish instead an oxoG encounter complex (EC) (hereafter referred to as V222P-oxoG EC), in which MutM is bound to an intact albeit severely buckled target oxoG:C base pair (Fig. 1, *C* and *E*) (14). The present studies utilized only protein constructs containing a wild-type OCL, and therefore the structures reflect only the influence of the residues under study, $Phe^{114}$ and $Met^{77}$ on the DNA conformational state.

All four complexes yielded crystals that diffracted to high resolution (1.8–1.9 Å; Table 1), and the structures were determined using molecular replacement followed by standard refinement. The target G in both non-lesion-containing structures is fully intrahelical, forming a Watson-Crick base pair with the complementary C (Fig. 3*A*). The structure of the M77A mutant protein bound to non-lesion-containing DNA (hereafter referred to as the M77A-G structure) was virtually identical to the corresponding V222P-G complex (Protein Data Bank code 2F5O), with a C$\alpha$ RMSD between the two of 0.19 Å, thus indicating that the $Met^{77}$ thioether-bearing side chain does not play a prominent role in directing the specific DNA conformation at the site of intrahelical interrogation by MutM. The structure of the F114A-G complex also bore marked overall similarity to the structure of V222P MutM bound to a target G in DNA (V222P-G; C$\alpha$ RMSD = 0.27 Å), with one prominent exception. Previously, all 26 intrahelical MutM-DNA complexes reported to date had revealed a highly unusual and characteristic buckling of the target base pair, a feature ranging in magnitude from 28° to 44° (buckle angles measured using 3DNA). The structure of the F114A-G complex, by contrast, altogether lacks this otherwise invariant and distinctive feature, having a virtually negligible buckle angle of −0.8° (*versus* −36.2° in the corresponding wild-type structure; Fig. 3*A*). This observation provides definitive evidence supporting prior suggestions that DNA intercalation by the $Phe^{114}$ side chain is responsible for buckling the target base pair (12–14, 17, 18, 32, 33). These unusual changes in DNA conformation cannot be ascribed to crystal packing, because the wild-type and F114-G mutant structures have the same packing arrangement but different conformations, because buckling of the target intrahelical base pair has been observed in three distinct crystal forms, and because no crystal contacts occur within 17 Å of residues 77 and 114 (supplemental Fig. S3).

The structures of the M77A and F114A mutant proteins interrogating a target oxoG in DNA proved particularly intriguing and informative. In the M77A structure, the target oxoG is fully intrahelical and engaged in Watson-Crick base pair with its complement C (Fig. 3*B*, *middle structure*), a state termed the intrahelical lesion EC, which prior to this work had been captured only in MutM structures having OCL mutations. The target base pair in M77A-oxoG is buckled to roughly the same extent as in the corresponding EC structure of V222P-oxoG (Fig. 3*B*, compare *left* and *middle structures*). It thus appears that the identity of the residue at position 77 is sufficient to convert between the two major states of lesion recognition, with the wild-type residue $Met^{77}$ crystallizing as an extrahelical LRC (as in Fig. 1*D*) and the mutant form $Ala^{77}$ crystallizing as an intrahelical lesion EC; an analogous situation obviously pertains for single point mutations in the OCL, for example at position 222 (14). It is worth noting that we have determined a large number of additional structures having $Met^{77}$ intact but in

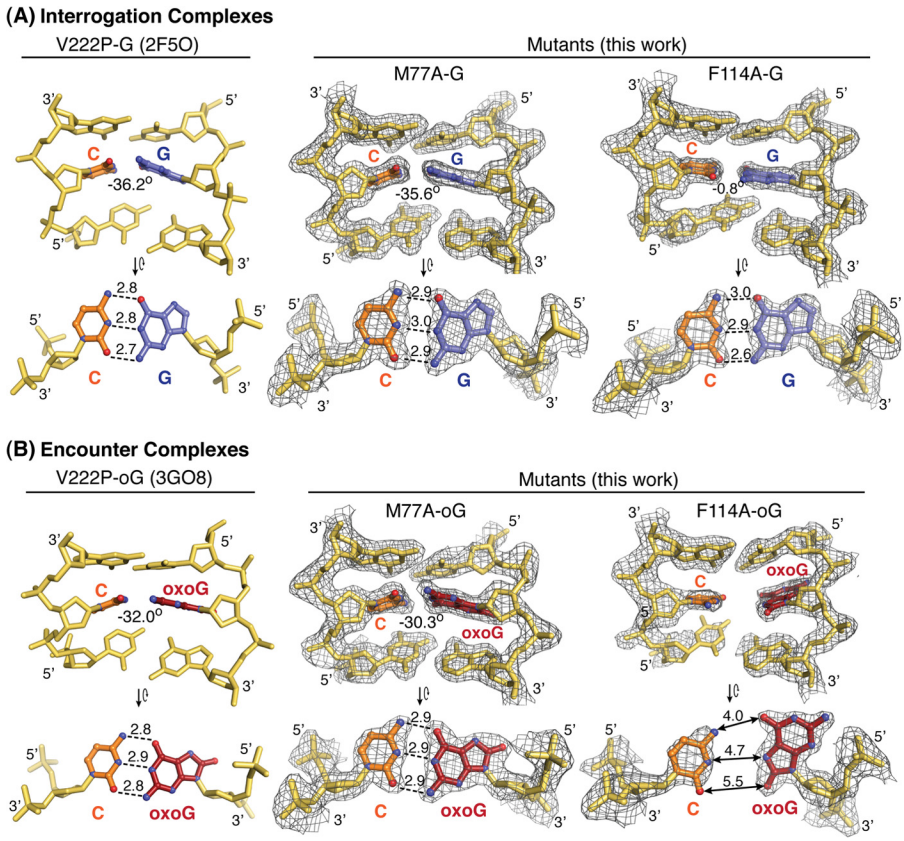



FIGURE 3. **The DNA conformation at the target site, with either G (*A*) or oxoG (*B*) as the target base.** The corresponding V222P structure for each set is shown on the *left*, the M77A structure are shown in the *middle*, and the F114A structures are shown on the *right*. The *top panels* show the buckle angle for the target base pair (with the exception of the V222P-oG and F114A-oG structures), whereas the *bottom panels* show a view down the helical axis for the target base pair. All of the distances were measured in Å, and the buckle angles were calculated using 3DNA. $2F_o$-$F_c$ maps were contoured to 1.0 $\sigma$.

which the cross-linking site or the sequence was permuted (Protein Data Bank codes 3GPY, 3GQ4, and 3JR5) (14, 31, 32); without exception, those structures also represent extrahelical lesion recognition complexes, lending further support to the notion that that the truncation of the Met$^{77}$ side chain alone is sufficient to shift the energetic balance for base extrusion so as to prefer the intrahelical state, at least in the crystal environment. Although the absolute free energies of the complexes in our crystals cannot be assumed to equal those in solution, it is safe to assume that free energy differences between states, or at least the rank order of free energy differences, are the same in the crystal and in solution. This strongly suggests that under physiologic conditions, the Met$^{77}$ side chain significantly stabilizes the extrahelical state of the target oxoG.

Truncation of the Phe$^{114}$ side chain yields a state of MutM-DNA interaction never before observed in the dozens of structures available for MutM and its orthologs bound to DNA. In this state, the target oxoG is fully intrahelical and is both nearly coaxial and coplanar with the complementary C (Fig. 3, *A* and *B*, *right panels*). However, the target oxoG in the F114A-oxoG complex is swiveled ~180° about its glycosidic bond so as to adopt the *syn* glycosidic configuration, as opposed to the typical *anti* configuration observed in the corresponding wild-type complex (Fig. 3*B*, compare *right* and *left panels*). Although the *syn*-oxoG and the complementary C bear an unmistakable resemblance to a Watson-Crick base pair, the two nucleobases are separated by too great a distance to form hydrogen bonds, and indeed, the interaction between O$^8$-oxoG and O$^2$-C would be repulsive if the two were brought much closer together. The F114A-oxoG structure thus represents a complex in which the target base pair has been disrupted, but in which the oxoG retains its engagement in stacking interactions with the 5′-neighbor and consequently remains nested within the helical stack.

*Detailed Analysis of the Role of Met$^{77}$*—As mentioned above, OCL-perturbing mutations such as V222P invariably result in crystals of the EC state (Fig. 1*C*) (14). Here we revealed that this state can also be accessed by introduction of the M77A mutation. Although V222P and M77A both alter the energetic landscape of the base extrusion pathway so as to favor the intrahelical encounter state, these mutations are structurally distinct; whereas V222P resides far from the locus of helix invasion on a loop that is disordered in EC structures, M77A lies at the epicenter of helix invasion. Comparison of the EC structures produced by V222P *versus* M77A will therefore lend insight into the influence of the Met$^{77}$ side chain on the DNA structure in the EC state. Comparison of the M77A-oG EC structure with the V222P-oxoG EC structure (Protein Data Bank code 3GO8) reveals that the two EC structures are nearly identical, with a C$\alpha$ alignment resulting in an RMSD of 0.17 Å. Most importantly, the three other residues that make intimate contact with the duplex at the site of invasion: Phe$^{114}$, Arg$^{112}$, and Arg$^{76}$, adopt essentially identical positions in the two complexes, and the details of DNA conformation are also fully conserved

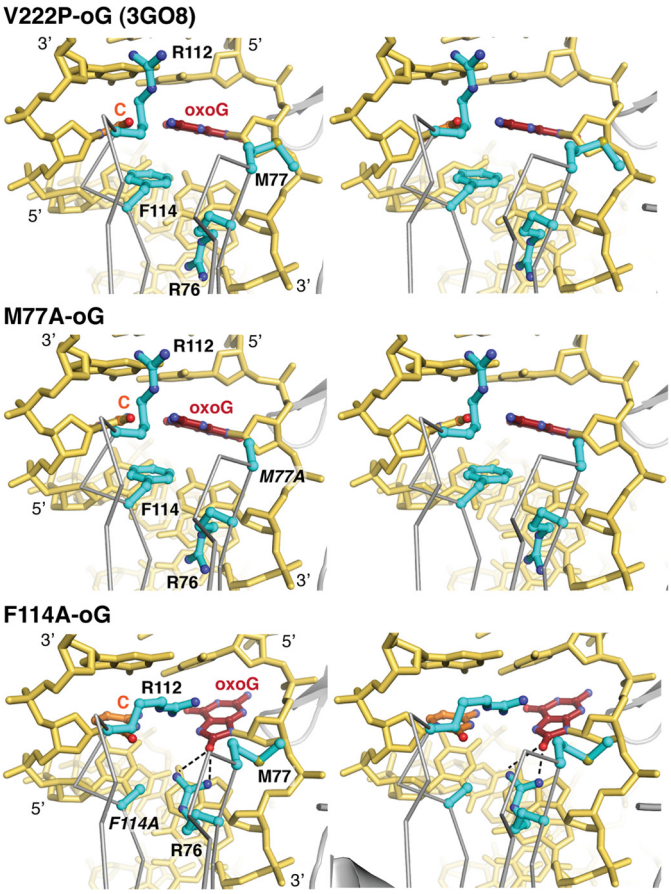



FIGURE 4. **Stereo views of side chain interactions at the target site.** The side chains are colored in *cyan*, the protein chain is *gray*, the target oxoG is *red*, and the opposing C is *orange*. The mutated residues are indicated in *italics*. Shown in the *right panels*, in the F114A-oG structure $Arg^{76}$ shifts upward to form two new hydrogen bonds (at distances of 2.9 and 2.7 Å) with the $O^8$.

(RMSD 0.14 Å; Fig. 4). Intrahelical oxoG encounter has previously been found to result in a structural rearrangement of the DNA backbone at the site of the target base, as compared with corresponding G structure, and the nature of the rearrangement was found to promote base extrusion. The driving force behind the formation of this "extrudogenic conformation" with oxoG is provided by a severe steric clash within the oxoG moiety, between its nucleobase C8-carbonyl) and both the 2′-C and 5′-phosphate (14). Virtually identical rearrangements leading to formation of the extrudogenic conformation are also observed in comparing M77A-G with M77A-oxoG (Fig. 5*A*). Taken together, these structural comparisons led us to conclude that even though the M77A mutation perturbs the energetic landscape of the protein-DNA complex, it does not perturb the fundamental structural nature of intrahelical oxoG encounter.

How does $Met^{77}$ stabilize the extrahelical state? High level computational simulations will ultimately be required to answer this question, but the x-ray structures do provide some noteworthy hints. As a pretext to this analysis, we assume that the x-ray structures represent the products of free equilibration, at least in solution prior to crystallization. Thus, although it is tempting to speculate that $Met^{77}$ acts as a steric block toward reannealing of the oxoG lesion, and indeed that could well be the case, our structures provide no information on the existence of any such kinetic effect. As for thermodynamic stabilization of the extrahelical state by $Met^{77}$, we note that this residue has a greatly restricted range of positions in the extrahelical (LRC) state than intrahelical (EC) state and also that extrusion of oxoG is accompanied by retraction of $Met^{77}$ into the space vacated by departure of oxoG (Fig. 6*B*). In all of the LRC structures to date, the side chain of $Met^{77}$ makes several direct van der Waals contacts with the sugar moiety of the extrahelical oxoG and the phosphate backbone flanking the target oxoG (Fig. 6*A*). In contrast, the van der Waals contact repertoire made by $Met^{77}$ in the EC state is much less extensive and limited to only the sugar moiety of the intrahelical G and the phosphate backbone 3′ to the target base compared with the LRC state. These observations suggest that $Met^{77}$ thus appears to stabilize the extrahelical state through the establishment of nonpolar contacts that are uniquely tailored to the extrahelical (LRC state).

*Detailed Analysis of the Role of $Phe^{114}$*—Whereas deletion of the $Met^{77}$ side chain has no obvious effect on the structure of the MutM complexes, deletion of the $Phe^{114}$ side chain has pronounced effects on DNA structure, and these effects provide insight into the unique function of this residue. One of the most striking features of intrahelical DNA helix interrogation by MutM is the pronounced buckling of the target base pair. Buckling occurs at the site of DNA duplex insertion by $Phe^{114}$, during which the base pair buckles away from $Phe^{114}$ as if to avoid it (Fig. 4). This behavior led to suggestions that $Phe^{114}$ directly promotes rupture of the target base pair by putting it under conformational and steric duress (12–14, 17, 32). The hypothesis that $Phe^{114}$ is directly and singularly responsible for buckling of the target base pair is fully supported by the present structural data. Precisely, with undamaged DNA (F114A-G structure), removal of the $Phe^{114}$ side chain results in complete unbuckling of the target base pair, a 35° reduction in the buckle angle (calculated using 3DNA (28); Fig. 3*A*), and restoration of a typical Watson-Crick base pair. Apart from the degree of buckling of the target base pair, no other significant structural adjustment is attributable to the presence or absence of the $Phe^{114}$ side chain.

The present work reveals that the nature of the encounter between MutM and an intrahelical oxoG is surprisingly and dramatically dependent upon the presence or absence of the $Phe^{114}$ side chain. With $Phe^{114}$ intact, MutM forms an intrahelical lesion encounter complex, the hallmarks of which are buckling of a fully intact target oxoG:C base pair and adoption of an extrudogenic conformation in the DNA backbone (Fig. 4 and supplemental Fig. S2). Mere removal of the $Phe^{114}$ side chain drastically changes the nature of target recognition, with the oxoG lesion being swiveled ~180° about its glycosidic bond to adopt the *syn* conformation and with the target nucleobase being intrahelical and unbuckled but also unpaired. The unusual conformation of oxoG appears to be stabilized through a hydrogen-bonding interaction between its $O^8$ and the side chain of $Arg^{76}$, an interaction not seen previously in any structure of MutM bound to DNA (Fig. 4). The conformation of $Arg^{76}$ required to establish this hydrogen bonding interaction with oxoG would be expected to clash sterically with $Phe^{114}$

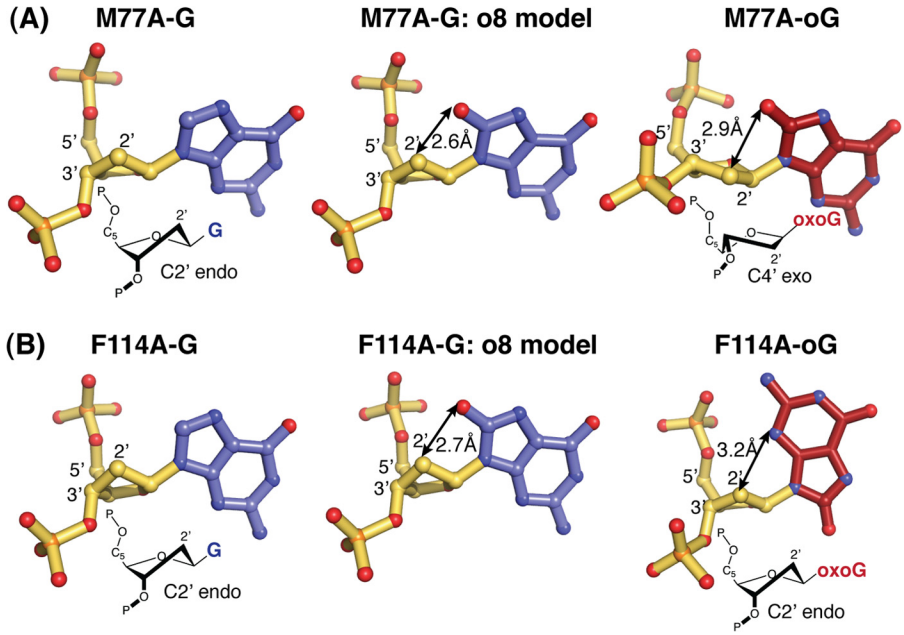
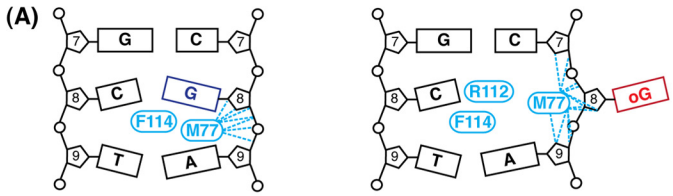



FIGURE 5. **Conformation of the DNA backbone at the target base for M77A (*A*) and F114A (*B*).** *Double-headed arrows* show the distance between the $O^8$ atom and $C^{2'}$ carbon when modeled into the G structures (*middle panels*). In the M77A-oG structure, the sugar pucker switches to $C^{4'}$ exo to avoid repulsive interactions between the $O^8$ atom and the $C^{2'}$ carbon, whereas the rotation of the target oxoG to syn in the F114A-oG structure allows the sugar to remain in the ground state C2′ endo pucker (the distance between $C^{2'}$ and the $N^3$ atom, which is now the closest nucleobase atom to the sugar, is shown).




| EC state | | | LRC state | | |
|---|---|---|---|---|---|
| Met77-Sδ | G8-C1′ | 4.0 Å | Met77-Cε | C7-C2′ | 3.4 Å |
| Met77-Sδ | G8-C2′ | 3.9 Å | Met77-Cε | C7-O3′ | 3.6 Å |
| Met77-Sδ | G8-C3′ | 4.1 Å | Met77-Cε | oG8-O5′ | 3.7 Å |
| Met77-Sδ | G8-C4′ | 3.8 Å | Met77-Cε | oG8-C5′ | 4.0 Å |
| Met77-Sδ | G8-O3′ | 4.0 Å | Met77-Cε | oG8-C3′ | 3.9 Å |
| Met77-Cε | A9-O3′ | 4.1 Å | Met77-Cε | oG8-C2′ | 3.9 Å |
| Met77-Cε | A9-OP1 | 4.5 Å | Met77-Sδ | oG8-C3′ | 3.8 Å |
| Met77-Cε | A9-C5′ | 3.8 Å | Met77-Sδ | oG8-C2′ | 3.9 Å |
| | | | Met77-Sδ | A9-O5′ | 3.5 Å |
| | | | Met77-Sδ | A9-C5′ | 3.7 Å |
| | | | Met77-Sδ | A9-O4′ | 3.9 Å |
| | | | Met77-Cγ | A9-O4′ | 4.0 Å |

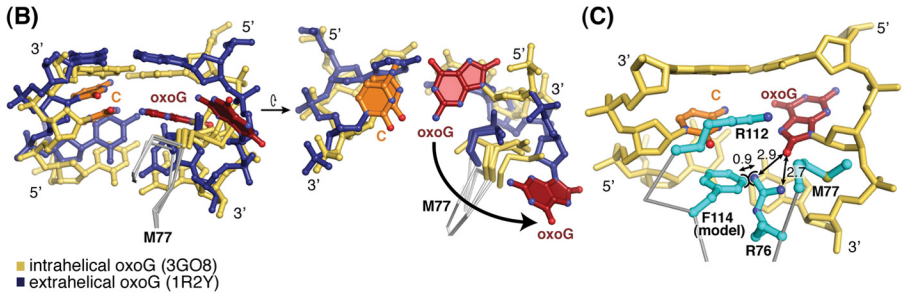



FIGURE 6. **Role of residues Met$^{77}$ and Phe$^{114}$ in oxoG repair.** *A*, van der Waals contacts between $Met^{77}$ and neighboring DNA backbone in the EC and LRC states. *B*, the $Met^{77}$ side chains from a Cα alignment of all available MutM-DNA complexes are shown with the indicated intrahelical oxoG and intrahelical G structure. A total of six intrahelical oxoG (Protein Data Bank codes 3GO8, 3GP1, 3GPP, 3GPU, and 3GQ3 and F114A-oG) and six intrahelical G (Protein Data Bank codes 2F5O, 3GPX, 3GQ5, 2F5N, and 2F5P and F114A-G) structures were used for the alignment. In the *right panel*, the base pair 5′ to the oxoG has been removed for clarity. The *black arrow* represents the path of base extrusion. The position of the side chain varies widely in the intrahelical oxoG state, whereas in the extrahelical state the side chain occupies a firm position between the extruded oxoG and the DNA. *C*, modeling the $Phe^{114}$ side chain into the F114A-oG structure shows a distinct steric clash between $Phe^{114}$ and $Arg^{76}$, which forms two stabilizing hydrogen bonding interactions with the $O^8$ from the syn oxoG. The distances are shown in Å.

when the latter is present (Fig. 6*C*); however, this clash appears as though it could be alleviated through rotation about the $Phe^{114}$ Cα-Cβ bond. In an unpublished work, we have observed $Arg^{76}$ taking on a similar role in structures of MutM (containing $Phe^{114}$) bound to oxoG-containing DNA in which the oxoG has been repositioned onto the opposing "complementary"

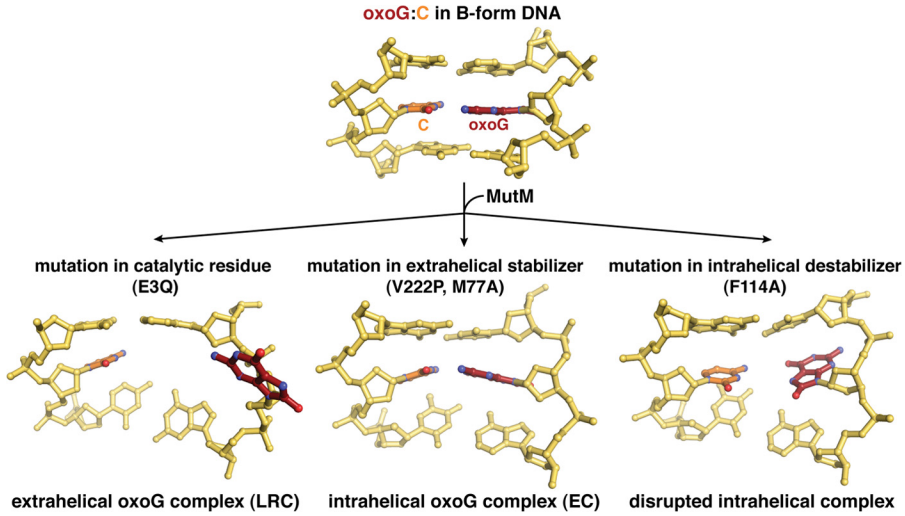



FIGURE 7. **Effects of various mutations on oxoG conformation in duplex DNA during lesion recognition by MutM.** Protein Data Bank codes (clockwise from *top*) were 183D, 4G4R (this work), 1R2Y, and 4G4O (this work).

strand; in those complexes, $Arg^{76}$ forms stabilizing hydrogen bonds to a *syn* undamaged guanine base located on the opposing strand, directly 5′ to the base occupying the position of the opposing C in our mutant structures.[6] Preliminary mutagenesis studies into the role of $Arg^{76}$ in oxoG repair show that replacement of the arginine side chain with alanine significantly impairs base excision DNA repair of a duplex oxoG-containing substrate (supplemental Fig. S4). However, substitution to methionine or lysine also failed to rescue oxoG repair activity, indicating that side chain length is not the determining factor for the function of residue 76 (supplemental Fig. S4). Further investigation will be necessary to fully explore the role of this residue in oxoG recognition and repair.

The large separation of the oxoG nucleobase from its complementary C is associated with a significant shift of the DNA backbone toward MutM at the site of the lesion, one that is even more extrudogenic than the corresponding shift seen in the EC structures (supplemental Fig. S2*B*, shifting an average 2.2 Å from the G-bound to oxoG-bound state with the F114A mutation *versus* 0.92 Å with $Phe^{114}$ intact (14)). We conclude that the extrudogenic backbone conformation alone is not sufficient to induce extrusion of oxoG and that indeed $Phe^{114}$ plays this role of enforcer under physiologic conditions with wild-type MutM. The substantial decrease in base excision activity that accompanies the F114A mutation (Fig. 2*A*) thus appears to arise from an inability of the mutant protein to populate the extruded state that is required for catalysis of base excision.

## DISCUSSION

An abundance of structural, biochemical, and computational evidence indicates that MutM actively promotes the extrusion of oxoG lesions from DNA. This extrusion is of course a prerequisite to execution of a catalytic DNA repair reaction cascade by MutM, because the single enzyme active site that accelerates all steps of the repair cascade is capable of recognizing only a fully extrahelical nucleobase. Interestingly, evidence suggests that MutM also accelerates the extrusion of undamaged bases from DNA but is ineffective at catalyzing their excision. Computational studies have indicated that MutM lowers the barrier for base extrusion at oxoG sites by ~7 kcal/mol more than at sequence-matched G sites, thus providing a kinetic basis for discrimination of lesions from undamaged DNA even at the earliest stages of encounter (14). Wresting a nucleobase from DNA and stabilizing it in the extrahelical state does not occur spontaneously to any appreciable extent; hence MutM must invest energy in the form of interfacial contacts to enforce presentation of an extrahelical lesion to the enzyme active site. *A priori*, these contacts would be expected to consist of two broad classes: those that serve to destabilize the intrahelical state and those that stabilize the extrahelical state.

At present, MutM is the only base excision DNA repair enzyme for which the structural features of encounter with an intrahelical lesion have been elucidated in detail. These EC structures indeed provide evidence suggesting that MutM destabilizes the target base pair, with that pair being completely destacked from its 3′-neighbor and being severely buckled toward the 5′-neigbor (Fig. 1*C*). The structures also provide evidence suggesting stabilization of the extrahelical LRC state, namely acquisition of multiple interfacial contacts to the extrahelical oxoG (including the oxoG-capping loop, OCL), insertion of $Arg^{112}$ into the helical stack to hydrogen bond with the complementary C, and establishment of van der Waals contacts with $Met^{77}$. Here and in previous work, we have used site-directed mutagenesis to probe the structural and (by inference) energetic influence of residues in MutM that prominently figure in the protein-DNA interface.

Prior to the present work, two distinct classes of mutations were known. Mutations that affect catalysis but not substrate recognition, typified by E3Q, afford extrahelical lesion recognition complexes, LRCs (Fig. 7, *left structure*). Mutations that deprive the enzyme of its ability to stabilize the extrahelical state, typified by the OCL mutant V222P, afford intrahelical lesion ECs (Fig. 7, *middle structure*). In the present work, we have found that the M77A mutation also leads to entrapment of an EC state; hence we conclude that $Met^{77}$ stabilizes the extrahelical state. Here we have also observed a third state of MutM

[6] R.-J. Sung, M. Zhang, Y. Qi, and G. L. Verdine, manuscript in preparation.

interaction with a lesion, one characterized by complete rupture of the target base pair but failure to extrude the target oxoG nucleobase (Fig. 7, *right structure*). The observation of such a structure highlights the fact that rupture of a base pair with at least partial preservation of stacking interactions is less energetically expensive than extrusion of a nucleobase from DNA. Enforcement of extrusion requires additional repulsive interactions beyond those seen in the F114A-oxoG structure, and those repulsive interactions are provided in the wild-type enzyme by the $Phe^{114}$ side chain being jammed into the DNA helix so forcefully as to buckle the target base pair. The very observation of such an unusual structure attests to the extraordinary ability of MutM to pry open the DNA duplex upon encounter with an oxoG lesion. Although F114-oxoG can hardly be considered a physiologic complex, we believe it provides insight into an exceedingly important but poorly understood physiologic property of MutM. When base-paired against A, MutM is nearly devoid of base excision activity; evolution provides strong pressure against MutM processing A:oxoG lesions, because such repair would have the pro-mutagenic effect of accelerating C:G to T:A transversion mutations. In the A:oxoG base pair, the A and oxoG adopt a strong, stable Hoogsteen pairing configuration, with A adopting the *anti* glycosidic bond configuration and oxoG adopting the *syn* configuration, the same as that observed in the F114A-oxoG structure. Although no structure is yet available for MutM bound to an intrahelical A:oxoG base pair, the F114A-oxoG structure provides a basis for extrapolation to A:oxoG. As shown in supplemental Fig. S5, if the complementary C in the F114A-oxoG structure were replaced with A, the A, by virtue of its larger ring size, would span the entire gap between the two target bases and would be perfectly positioned to Hoogsteen pair with oxoG. oxoG in turn would be hydrogen-bonded to $Arg^{76}$. Taken together, the unusual features of the F114A structure leaves the DNA perfectly poised to accept an A:oxoG and form a stable intrahelical recognition complex with it. This stabilization of A:oxoG in the intrahelical state would increase the kinetic barrier to extrusion of oxoG, thereby rendering A:oxoG a poor substrate for base excision. What is difficult to predict is the effect of $Phe^{114}$ on the A:oxoG complex, but based on the available information, it seems likely that $Phe^{114}$ buckles the A:oxoG base pair, but the destabilizing effect of buckling is more than compensated by the intrinsically high stability of the A:oxoG base pair, the $Arg^{76}$-oxoG hydrogen bond, and the other intermolecular contacts that stabilize the gap in the F114A-oxoG complex.

$Phe^{114}$ is inserted into DNA at the site of a pronounced DNA bend that is localized to 3′-side of the target base pair and that completely destacks the target base pair from its 3′-neighbor. This association suggests the possibility of causation, specifically that $Phe^{114}$ might actually be responsible for formation of the bend. Our results with the F114A-G complex rule out that scenario conclusively: the bend is just as pronounced in this complex as in other intrahelical recognition complexes, although the target base pair is completely unbuckled. $Met^{77}$ is also inserted into the duplex at the kinked site, yet it also is not required to produce the bend, as judged by inspection of the M77A-G structure. Hence we conclude that interfacial contacts by residues other than $Phe^{114}$ and $Met^{77}$ produce the signature bend seen thus far in all MutM/DNA complexes. The observation that the bend is retained upon removal of side chains that invade DNA having only intact base pairs suggests that bending is an intrinsic feature of MutM/DNA binding, in which case the bend would be expected to migrate along with MutM as the protein slides along DNA while searching for lesions. By the same token, during every base pair-wise translocation event, the side chain of $Phe^{114}$ would have to clear the obstruction presented by the 5′-base base pair, suggesting that $Phe^{114}$ is in nearly constant repulsive contact with DNA during the lesion search. Consistent with this notion, deleting the $Phe^{114}$ side chain modestly increases the affinity of the protein for DNA (supplemental Fig. S1) and also increases the rate of processive sliding (18).

*Acknowledgments—We are grateful to staff from the 24-ID beamline of the Argonne Photon Source and at the X29 beamline of the National Synchrotron Light Source for assistance in data collection. We thank Kwangho Nam and Danaya Pakotiprapha for constructive feedback during the writing of this manuscript. We also thank members of the Verdine group for helpful advice.*

## REFERENCES

1. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature* **362,** 709–715
2. Barnes, D. E., and Lindahl, T. (2004) Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu. Rev. Genet.* **38,** 445–476
3. Michaels, M. L., and Miller, J. H. (1992) The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7,8-dihydro-8-oxoguanine). *J. Bacteriol.* **174,** 6321–6325
4. Nash, H. M., Bruner, S. D., Schärer, O. D., Kawate, T., Addona, T. A., Spooner, E., Lane, W. S., and Verdine, G. L. (1996) Cloning of a yeast 8-oxoguanine DNA glycosylase reveals the existence of a base-excision DNA-repair protein superfamily. *Curr. Biol.* **6,** 968–980
5. van der Kemp, P. A., Thomas, D., Barbey, R., de Oliveira, R., and Boiteux, S. (1996) Cloning and expression in *Escherichia coli* of the OGG1 gene of *Saccharomyces cerevisiae*, which codes for a DNA glycosylase that excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine. *Proc. Natl. Acad. Sci. U.S.A.* **93,** 5197–5202
6. Fedorova, O. S., Nevinsky, G. A., Koval, V. V., Ishchenko, A. A., Vasilenko, N. L., and Douglas, K. T. (2002) Stopped-flow kinetic studies of the interaction between *Escherichia coli* Fpg protein and DNA substrates. *Biochemistry* **41,** 1520–1528
7. Fromme, J. C., and Verdine, G. L. (2002) Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM. *Nat. Struct. Biol.* **9,** 544–552
8. Oda, Y., Uesugi, S., Ikehara, M., Nishimura, S., Kawase, Y., Ishikawa, H., Inoue, H., and Ohtsuka, E. (1991) NMR studies of a DNA containing 8-hydroxydeoxyguanosine. *Nucleic Acids Res.* **19,** 1407–1412
9. Lipscomb, L. A., Peek, M. E., Morningstar, M. L., Verghis, S. M., Miller, E. M., Rich, A., Essigmann, J. M., and Williams, L. D. (1995) X-ray structure of a DNA decamer containing 7,8-dihydro-8-oxoguanine. *Proc. Natl. Acad. Sci. U.S.A.* **92,** 719–723
10. Plum, G. E., Grollman, A. P., Johnson, F., and Breslauer, K. J. (1995) Influence of the oxidatively damaged adduct 8-oxodeoxyguanosine on the conformation, energetics, and thermodynamic stability of a DNA duplex. *Biochemistry* **34,** 16148–16160
11. Bowman, B. R., Lee, S., Wang, S., and Verdine, G. L. (2008) Structure of the *E. coli* DNA glycosylase AlkA bound to the ends of duplex DNA. A system for the structure determination of lesion-containing DNA. *Structure* **16,** 1166–1174

12. Fromme, J. C., and Verdine, G. L. (2003) DNA lesion recognition by the bacterial repair enzyme MutM. *J. Biol. Chem.* **278,** 51543–51548
13. Banerjee, A., Santos, W. L., and Verdine, G. L. (2006) Structure of a DNA glycosylase searching for lesions. *Science* **311,** 1153–1157
14. Qi, Y., Spong, M. C., Nam, K., Banerjee, A., Jiralerspong, S., Karplus, M., and Verdine, G. L. (2009) Encounter and extrusion of an intrahelical lesion by a DNA repair enzyme. *Nature* **462,** 762–766
15. Ishchenko, A. A., Vasilenko, N. L., Sinitsina, O. I., Yamkovoy, V. I., Fedorova, O. S., Douglas, K. T., and Nevinsky, G. A. (2002) Thermodynamic, kinetic, and structural basis for recognition and repair of 8-oxoguanine in DNA by Fpg protein from *Escherichia coli*. *Biochemistry* **41,** 7540–7548
16. Nam, K., Verdine, G. L., and Karplus, M. (2009) Analysis of an anomalous mutant of MutM DNA glycosylase leads to new insights into the catalytic mechanism. *J. Am. Chem. Soc.* **131,** 18208–18209
17. Qi, Y., Nam, K., Spong, M. C., Banerjee, A., Sung, R.-J., Zhang, M., Karplus, M., and Verdine, G. L. (2012) Strandwise translocation of a DNA glycosylase on undamaged DNA. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 1086–1091
18. Dunn, A. R., Kad, N. M., Nelson, S. R., Warshaw, D. M., and Wallace, S. S. (2011) Single Qdot-labeled glycosylase molecules use a wedge amino acid to probe for lesions while scanning along DNA. *Nucleic Acids Res.* **39,** 7487–7498
19. Maxam, A. M., and Gilbert, W. (1977) A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U.S.A.* **74,** 560–564
20. Otwinowski, Z., and Minor, W. (1997) *Methods in Enzymology,* **276,** 307–326
21. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX. A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221
22. Emsley, P., and Cowtan, K. (2004) Coot. Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
23. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2256–2268
24. Painter, J., and Merritt, E. A. (2006) TLSMD web server for the generation of multi-group TLS models. *J. Appl. Crystallogr.* 109–111
25. Painter, J., and Merritt, E. A. (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D Biol. Crystallogr.* **62,** 439–450
26. Davis, I. W., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2004) MOLPROBITY. Structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.* **32,** W615–W619
27. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (2011) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26,** 283–291
28. Lu, X.-J., and Olson, W. K. (2003) 3DNA. A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **31,** 5108–5121
29. Schrodinger, LLC. (2010) The PyMOL Molecular Graphics System, Version 1.3r1
30. Porello, S. L., Leyes, A. E., and David, S. S. (1998) Single-turnover and pre-steady-state kinetics of the reaction of the adenine glycosylase MutY with mismatch-containing DNA substrates. *Biochemistry* **37,** 14756–14764
31. Qi, Y., Spong, M. C., Nam, K., Karplus, M., and Verdine, G. L. (2010) Entrapment and structure of an extrahelical guanine attempting to enter the active site of a bacterial DNA glycosylase, MutM. *J. Biol. Chem.* **285,** 1468–1478
32. Sung, R.-J., Zhang, M., Qi, Y., and Verdine, G. L. (2012) Sequence-dependent structural variation in DNA undergoing intrahelical inspection by the DNA glycosylase MutM. *J. Biol. Chem.* **287,** 18044–18054
33. Koval, V. V., Kuznetsov, N. A., Ishchenko, A. A., Saparbaev, M. K., and Fedorova, O. S. (2010) Real-time studies of conformational dynamics of the repair enzyme E. coli formamidopyrimidine-DNA glycosylase and its DNA complexes during catalytic cycle. *Mutat. Res.* **685,** 3–10
34. Brunger, A. (1993) *Acta Crystallogr. D* **49,** 24–36

# Structural and Biochemical Analysis of DNA Helix Invasion by the Bacterial 8-Oxoguanine DNA Glycosylase MutM

Rou-Jia Sung, Michael Zhang, Yan Qi and Gregory L. Verdine

*J. Biol. Chem. 2013, 288:10012-10023.*
*doi: 10.1074/jbc.M112.415612 originally published online February 12, 2013*