Title

Structural Basis for Finding OG Lesions and Avoiding Undamaged G by the DNA Glycosylase MutY

Journal

ACS Chemical Biology, 15(1)

ISSN

1554-8929

Authors

Russelburg, L Peyton
Murray, Valerie L O'Shea
Demir, Merve
et al.

Publication Date

2020-01-17

DOI

10.1021/acschembio.9b00639

Peer reviewed

HHS Public Access

Author manuscript

ACS Chem Biol. Author manuscript; available in PMC 2021 January 17.

Published in final edited form as:

ACS Chem Biol. 2020 January 17; 15(1): 93–102. doi:10.1021/acschembio.9b00639.

Structural basis for finding OG lesions and avoiding undamaged G by the DNA glycosylase MutY

L. Peyton Russelburg[*,1], **Valerie L. O'Shea Murray**[*,2,3], **Merve Demir**[*,2], **Kyle R. Knutsen**[1], **Sonia L. Sehgal**[1], **Sheng Cao**[2], **Sheila S. David**[2,‡], **Martin P. Horvath**[1,‡]

[1]School of Biological Sciences, University of Utah, 257 S 1400 E, Salt Lake City, Utah, 84112

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

[3]Department of Chemistry, University of Utah, 315 S 1400 E, Salt Lake City, Utah, 84112

Abstract

The adenine glycosylase MutY selectively initiates repair of OG:A lesions and, by comparison, avoids G:A mispairs. The ability to distinguish these closely related substrates relies on the C-terminal domain of MutY which structurally resembles MutT. To understand the mechanism for substrate specificity, we crystallized MutY in complex with DNA containing G across from the high-affinity azaribose transition state analog. Our structure shows that G is accommodated by the OG site and highlights the role of a serine residue in OG *versus* G discrimination. The functional significance of Ser308 and its neighboring residues was evaluated by mutational analysis, revealing the critical importance of a beta-loop in the C-terminal domain for mutation suppression in cells, and biochemical performance *in vitro.* This loop comprising residues Phe307, Ser308, and His309 (*Geobacillus stearothermophilus* sequence positions) is conserved in MutY but absent in MutT and other DNA repair enzymes, and may therefore serve as a MutY-specific target exploitable by chemical biological probes.

Aberrant DNA modifications that arise from chemical reactions with exogenous and endogenous agents are considered "DNA damage" since these modifications put biological systems at risk. DNA repair enzymes mitigate this risk by counteracting chemical damage that otherwise would erode information content of DNA.[1] Guanine is particularly prone to oxidative damage due to its low redox potential.[2] Oxidation of G results in 8-oxo-7,8-dihydroguanine (OG) which differs from G by only two atoms (Figure 1). The OG lesion is especially problematic because the *syn* conformer mispairs with adenine during DNA replication. The guanine oxidation (GO) repair pathway prevents mutations that otherwise would arise from OG template ambiguity (Figure 2). The GO repair pathway features

‡Corresponding authors: martin.horvath@utah.edu; ssdavid@ucdavis.edu.
*These authors contributed equally to this work

Accession Codes

Atomic coordinates for TSAC-G:1N have been deposited with the Protein Data Bank with PDB ID 6Q0C. The updated refinement for TSAC-OG:1N has been deposited with PDB ID 6U7T.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acschembio.9b00639.

• Supplemental figures and tables, X-ray crystallography statistics, experimental methods (PDF).

enzymes MutT, MutM/Fpg, and MutY.[3] MutT (MTH1 in humans) prevents misincorporation of OG across A by hydrolyzing OGTP to remove it from the nucleotide pool.[4,5] Fpg (the MutM gene product) in bacteria and its human ortholog hOGG1 initiate *b*ase *e*xcision *r*epair (BER) by removing OG from OG:C base pairs. When replication proceeds before Fpg/hOGGI-mediated repair, an OG:A mispair may arise. The BER enzyme MutY (MUTYH in humans) removes A from OG:A mispairs and is thus a last defense to prevent G:C to T:A transversion mutations (see Figure S1 for mechanism of adenine excision).

Clinical variants of each of these GO-pathway enzymes have been identified in association with cancer.[6] Inherited variants of MUTYH lead to increased frequency of somatic G:C to T:A mutations which can accelerate colon cancers by disabling genes, especially the APC tumor suppressor gene, a situation now designated *M*UTYH *a*ssociated *p*olyposis (MAP),[7] reviewed recently.[8,9] The GO-pathway enzymes are of additional biomedical interest as validated targets for cancer therapy.[10] Cancer cells experience higher oxidative stress compared to quiescent cells and are often reliant on a particular DNA repair enzyme, providing opportunities for therapeutic intervention. Indeed, knockdown of MUTYH expression favorably impacted the proliferative potential and apoptotic rate of cells derived from pancreatic cancers indicating inhibitors of MUTYH may provide a route to pancreatic cancer cell-killing drugs.[11]

MutY is unique among BER glycosylases as it targets an undamaged base mispaired with a chemical lesion on the opposite DNA strand. Locating the OG:A mispair is therefore a more complex molecular recognition assignment than finding and removing a damaged base on one strand of DNA (e.g. deoxy-U by UDG, or alkylated A by AlkA). Domain-deletion analysis previously indicated that OG recognition is a function provided by the C-terminal domain (CTD), a domain which is not found in other protein superfamily members.[12] When the CTD was removed, the catalytic N-terminal domain (NTD) processed OG:A and G:A with diminished preference for the cognate OG:A lesion, conferring a mutator phenotype for *E. coli* expressing MutY-NTD.[13–15] This observation led to the view that the OG-recognition site of MutY resides within the CTD by analogy with MutT, which is homologous to MutY-CTD and which also recognizes the OG base moiety.[13,16]

MutY acts on OG:A mispairs,[17,18] avoids undamaged bases and mismatches such as G:T, yet shows activity towards the G:A mismatch.[19] Indeed, differences in the degree of product inhibition experienced by MutY processing G:A compared to OG:A mismatches can lead to the impression that G:A substrates are preferred.[20] However, OG:A lesions are the primary substrate of MutY as evidence by in-cell DNA repair assays.[21] It makes sense that MutY evolved with OG:A preference and, by comparison, G:A aversion since adenine removal in the later context is mutagenic. Unlike the mismatch repair system, MutY does not distinguish the template parental DNA strand from the newly synthesized daughter DNA strand. By contrast, adenine removal from OG:A mispairs suppresses mutations, a situation ensured by MutT which minimizes the likelihood of incorporating OG in the daughter DNA strand. Curiously, MutY substrate preference does not completely exclude G:A substrates. MutY-dependent BER converts G:A sites to G:C *in vitro*.[19] Also, MutY acting on G:A mismatches contributes to hyper-recombination,[22] and MutY seems to cooperate with MutS

to process G:A sites in DNA appropriately.[23] MutY thus solves an intriguing molecular-recognition puzzle with subtle, not exclusive preference for OG:A lesions.

The first x-ray structures of MutY in complex with DNA revealed MutY interacts with the OG base in an *anti* conformation and intra-helical position, largely through contacts with hydrogen-bonding residues and an intercalating tyrosine provided by the NTD, none of which are expected to be OG-specific.[24] One residue of the CTD, Ser308, provided an OG-specific hydrogen bond to O8 and an ambiguous hydrogen bond to N7 suggesting a mechanism for OG versus G discrimination but also leaving unanswered questions as to how MutY preferentially attacks OG:A lesions.[24] The same OG interactions were noted in a recent structure of MutY engaged in a Transition State Analog Complex (TSAC),[25] created by incorporating OG on one DNA strand across from the DNA strand containing the transition state analog (3R,4R)-4-(hydroxymethyl)pyrrolidin-3-ol, hereafter referred to as 1N (Figure 1), which mimics shape and charge properties of the oxacarbenium ion.[26–28]

To extend the structural and chemical basis for OG recognition, we report here a crystal structure of MutY from *Geobacillus stearothermophilus* (*Gs* MutY) in complex with DNA containing undamaged G across from 1N. We will refer to this structure as the TSAC-G:1N to distinguish it from the previously described TSAC-OG:1N structure.[25] The TSAC-G:1N structure reveals that G is accommodated in the OG-recognition site, implying that MutY does not contain an alternative site to exclude G. Ser308 in the CTD of *Gs* MutY changes hydrogen bonding partners in response to the OG-to-G perturbation, yet all of the other molecular interactions with DNA including electrostatic interaction between 1N and Asp144 at the active site remain intact. Altering or deleting residues Phe307, Ser308 and His309 within a conserved FSH loop reduced the mutation suppression function of MutY, impaired DNA-binding stability and slowed the kinetics of adenine removal, with an overall reduction in OG:A versus G:A selectivity. These results identify the FSH loop as a structural element within the CTD of MutY critical for DNA repair efficiency, which is significant as this can guide development of chemical biology probes specifically targeting sites within MutY that are remote from yet mechanistically coupled to the active site.

Results and Discussion

Structure determination.

In our effort to understand the chemical mechanisms and molecular recognition events underlying DNA repair by MutY, we determined its structure in complex with DNA containing undamaged G paired with the 1N transition state mimic. Crystals were grown as previously described,[25] except the OG DNA lesion was replaced with a normal G nucleotide across from 1N. MutY-DNA disulfide crosslinking (DXL) has previously been applied to stabilize lesion scanning, substrate and product analog complexes;[24,29,30] however, DXL was not needed for crystallization of MutY with DNA containing 1N,[25] probably a consequence of exceptionally high affinity for the TS analog.[25,26] Diffraction data to the 2.0 Å resolution limit were collected at the Advanced Light Source SIBYLS beamline.[31] Initial phases were obtained by molecular replacement using the TSAC-OG:1N structure (PDB ID 5DPK).[25] Refinement by multiple rounds of energy minimization with PHENIX[32] and model rebuilding with Coot[33] yielded a final structure with excellent stereochemistry

(r.m.s.d. bond length = 0.003 Å; bond angle = 0.7 degrees), 98% of residues within the favorable region and no residues in the excluded region of the Ramachandran plot, a MolProbity all atom clash score of 2.1, and R-work = 0.247/R-free = 0.271. See Supplemental Table S1 for additional data collection and model refinement statistics. The final model of *Gs* MutY includes residues 6–360 with the exception of residues 275–276 and 288–291; both regions of missing electron density are found at the tips of beta-loops pointed away from the DNA. The DNA structure comprises 21 nucleotides many of which are base paired. The DNA strand containing the critical G base has a 5' terminal, unpaired overhanging base, and the DNA strand containing the 1N TS mimic has a 5' terminal nucleotide that was not visible in electron density maps.

Overview of the structure.

The helical bundles of *Gs* MutY-NTD position the helix-hairpin-helix motif into the exaggerated minor groove of the DNA, bending it by a characteristic ~45-degree angle (Figure 3), as previously observed for MutY structures of the lesion recognition complex,[24] the fluorinated lesion recognition complex,[29] and the 1N transition state analog complex.[25] An extended linker (residues 226–235) passes across the major groove but makes no significant contact with DNA. The CTD comprises a highly twisted beta sheet that assembles with two alpha helixes, one alpha helix contouring each face of the sheet. Of these, the longer C-terminal alpha helix (residues 345–360) points its amino terminus at the 3'-phosphate of the G nucleotide to orient the helix dipole moment favorably with the negatively charged DNA backbone.

A pattern of elevated temperature factors indicates that the CTD is more flexible compared with the NTD and DNA. The average B-value for the CTD is 75 $Å^2$, comparable to the average B-value for the extended inter-domain linker region, and substantially higher than the average B-values for the NTD and DNA which are each 50 $Å^2$ (Figure S2). A similar pattern is apparent for previously determined MutY structures complexed to DNA containing OG, suggesting that CTD flexibility is an inherent property, not the consequence of substrate substitution.

Many of the residues with direct interactions to the G nucleotide are provided by the NTD, as described further below. Residues at the tip of a loop connecting two beta strands within the CTD, come in close van der Waals (VDW) contact with the G base, and make salt bridges to phosphate groups of the DNA. This beta-loop in the CTD, comprising residues Phe307, Ser308 and His309 (abbreviated hereafter as the FSH loop), inserts itself between the major groove of the G nucleotide and the extended inter-domain linker (Figure 3).

G fits into the OG site.

Undamaged G fits into the same site that recognizes OG. Evidence comes from initial difference maps viewed after molecular replacement with the TSAC-OG:1N structure (PDB ID 5DPK) that showed the O8 atom highlighted by negative density (red), consistent with placement of G in the OG site (Figure S3A). This negative peak along with other features in the electron density maps assured us that the x-ray data were able to define subtle differences in the new structure and that phases were not unduly influenced by model bias.

Replacement of OG with G during refinement eliminated the negative density, and the final simulated annealing (SA) composite omit maps fit perfectly with G at the OG site in TSAC-G:1N (Figure S3B) and OG at the OG site in TSAC-OG:1N (Figure S3C). BER glycosylases commonly isolate anti-substrate bases away from their active sites in an "exo-site." Other structures determined for MutY in complex with OG:C[34] and for hOGG1 in complex with undamaged DNA[35] highlight the use of exo-sites that bind with the off-target nucleobase and thus prevent non-cognate sites from being processed. For each of these examples, the base being engaged by the active site or intercepted by the exo-site must flip out of the DNA helix. By contrast, in the case of OG recognition by MutY the base maintains an intra-helical disposition, which probably precludes the exo-site strategy, necessitating a different mechanism to avoid G:A mismatches.

Altered hydrogen bonding for Ser308 in TSAC-G:1N and TSAC-OG:1N.

Ser308 is of particular interest because in the previously described MutY-DNA structures this residue hydrogen bonds with O8 and H7,[24,25,29] the two atoms that distinguish OG from G. Expectedly, in the TSAC-G:1N structure the C8 position of G no longer hydrogen bonds to the peptide amide of Ser308. It was unexpected, however, that MutY would forfeit both hydrogen bonds involving OG-specific atoms. The sidechain of Ser308 has surprisingly switched hydrogen bonding partners from the N7 position to a solvent molecule. The altered Ser308 rotamer in the TSAC-G:1N structure is strongly supported by electron density maps (Figure 4A). When Ser308 of TSAC-G:1N was modeled with the same rotamer as found in TSAC-OG:1N, strong positive and negative peaks emerged in the |Fo| − |Fc| difference map (Figure 4B) indicating incorrect rotamer placement and eliminating the null-hypothesis that Ser308 is unaffected by OG-to-G substitution. The converse is also true; when Ser308 of TSAC-OG:1N was modeled with the rotamer observed in TSAC-G:1N, the |Fo| − |Fc| difference map showed inverted positive and negative peaks (Figure 4C). By contrast, the |Fo| − |Fc| difference maps were essentially featureless when the structure-specific rotamers for Ser308 were re-instated (Figure 4A and Figure 4D).

The Ser308 sidechain is ambivalent concerning hydrogen bond polarity and therefore replacement of N7-H with N7: lone pair electrons would not, *per se*, preclude hydrogen bonding interactions with Ser308 since the hydroxyl group can both donate and accept hydrogen bonds. When describing the Lesion Recognition Complex (LRC) Fromme *et al.* point out the significance of Ser308 in OG *versus* G discrimination and suggest that polarity reversal of its hydrogen bond with N7 could impact an extended molecular interaction network involving Ser308, Tyr88 and the OG base.[24] Our TSAC-G:1N structure supports and extends these ideas, as we now clearly see that the Ser308 disengages from N7: of G. The hydrogen bond breaks and Ser308 finds a new hydrogen-bonding partner in a solvent molecule. Ser308 retains a hydrogen bond with Tyr88 in both structures, and in TSAC-G:1N Ser308 maintains contact with C8 and N7 of G through van der Waals interactions. In this way, the hydrogen-bonding network connecting the OG recognition site to the active site appears highly preserved between the TSAC-G:1N and TSAC-OG:1N, even though hydrogen bonds involving Ser308 are different (Figure S4).

Molecular interactions connecting the OG site to the active site.

The current structure with G retains many of the molecular interactions seen previously for OG at the OG site.[24,25] Figure 5 maps the DNA-protein contacts involving the G nucleotide. Aromatic stacking interactions fix the G base within the helical stack of base pairs in the 3' direction and intercalated Tyr88 in the 5' direction. The nucleobase adopts an *anti* conformation just as seen for OG in other MutY-DNA structures. The Watson Crick face and minor groove edge of G primarily interact with MutY's NTD. There are no hydrogen bonding interactions with the Hoogsteen face of G, unlike the TSAC-OG:1N structure where Ser308 hydrogen bonds with N7 and O8 of OG. Residues from both domains make salt bridges with the 5' and 3' phosphate groups of G (Figure 5). The Watson Crick face of G interacts with a conserved, extended structure involving residues Gln48 and Thr49 that connects two alpha-helices found within the 8-helix cluster comprising the catalytic domain of MutY. Specifically, the central amine (atom N1) of the Watson-Crick face hydrogen bonds with the carbonyl oxygen of the peptide connecting Gln48 and Thr49, and the exocyclic amine (atom N2) hydrogen bonds with the hydroxyl group of Thr49. Together, these interactions partially replace hydrogen bonds found for G:C base pairs. Moreover, these hydrogen bonds involving G and the residues Gln48 and Thr49 are preserved for OG in the TSAC-OG:1N structure and build a more extensive network of covalent and hydrogen bonds that connect the OG-recognition site to the catalytic active site (Figure S4).

Active site interactions with 1N are unaltered.

Plasmid-based DNA repair assays in cells and kinetic characterization *in vitro* determined that MutY strongly prefers OG:A over G:A substrates.[20,21] This functional difference motivated us to search for structural differences at the active site comparing TSAC-G:1N and TSAC-OG:1N. To fairly compare the two models, we updated the refinement of TSAC-OG:1N by extending the resolution of included data to 2.0 Å and applied the same parameters and protocols as applied for refinement of TSAC-G:1N (see Experimental Methods in Supporting Information). This refinement update increased the number of reflections by 25% and improved the MolProbity clash score for TSAC-OG:1N from 8.1 to 2.8 with only subtle adjustment to atom positions (R.M.S.D 0.14 Å; less than the coordinate error 0.35 A). The position and 1'-exo sugar pucker of 1N and its interactions with catalytic residues (Tyr126, Asp144 & Asn146) appear highly comparable for the two TSAC structures (see Figure S4 and Figure S5). Absence of significant structural differences at the active site suggests that the late transition state (TS2) mimicked by 1N is stabilized to the same degree regardless of whether the OG site is occupied by G or OG, implying also that selection for OG:A and avoidance of G:A is determined at an earlier point during substrate engagement and TS1 stabilization.

Mutation suppression of FSH variants.

Ser308 sits at the tip of the FSH loop that connects two beta strands in the CTD of MutY. This residue is part of a HXF(S/T)H motif that is conserved among MutYs, but not found in other BER enzymes (Figure 6). We investigated the importance of Ser308 and neighboring positions for mutation suppression function by replacing or deleting FSH residues and measuring the frequency of rifampicin resistant cells in overnight cultures. The FSH variants

were created in the context of a chimera protein with N-terminal domain derived from MutY of *E. coli* and the C-terminal domain derived from MutY of *G. stearothermophilus (EcNGsC).* This chimera MutY strategy made it possible to correlate functional changes to experimentally determined structural features in the C-terminal domain of *Gs* MutY. The C-terminal domain of *Ec* MutY has eluded structural biologists, and expression of full-length *Gs* MutY proved to be toxic to the reporter strain (see Experimental Methods in Supporting Information).

Table 1 reports the median frequency of rifampicin resistant (Rif^R) mutants, which is a proxy for the mutation frequency. As expected, negative control cultures lacking MutY expression (*null* in Table 1) displayed a 29-fold increase in Rif^R frequency compared to reference cultures expressing *Ec*N*Gs*C. Of note, cultures expressing *Ec* MutY from the pKK223 plasmid had a small but statistically noticeable 3.4-fold increase in median Rif^R frequency compared to reference cultures suggesting that the *Ec*N*Gs*C chimera MutY performed as well as and was, if anything, better at tracking down and removing promutagenic lesions than the version of MutY found naturally in *E. coli.*

Cultures expressing single amino-acid substitution variants of *Ec*N*Gs*C, YSH, ASH, FAH, and FVH, produced Rif^R frequencies similar to reference cultures, indicating that functional performance was not impacted by replacing Phe307 with smaller or slightly bigger side chains and also not impacted by removing the hydroxyl group of Ser308. Cultures expressing the double amino-acid substitution variants of *Ec*N*Gs*C, FPD and AAH, displayed a greater than 5-fold increase in Rif^R frequency. Cultures expressing *Ec*N*Gs*C with the three FSH residues deleted (Del FSH in Table 1) experienced an even larger 11-fold increase in Rif^R frequency, attaining a mutation frequency within 3-fold of the *null* cultures that did not express MutY.

The FPD sequence is found in the structural homolog MutT (see Figure 6). That some mutation suppression function was retained for the FPD variant of *Ec*N*Gs*C is consistent with the hypothesis that MutY was created through fusion of a MutT-like domain to an ancestral adenine glycosylase.[13] The alanine substitution variants of *Ec*N*Gs*C, ASH, FAH and AAH, complete a double-replacement cycle with apparent non-additivity since performance was retained following single-alanine replacement, yet a strong reduction in performance was produced by double replacement. This outcome, along with the graded response for the more drastic deletion variant, suggest that function of the FSH loop is distributed among several residues with redundancy and not completely dependent on a single position. Moreover, these results convince us that this highly conserved region, located remote from the site of catalysis, is critical for MutY's core function to suppress mutations in cells.

DNA-binding performance of selected FSH variants.

To understand the molecular basis for FSH loop function we pursued in vitro characterization of DNA binding and glycosylase reaction kinetics, measuring dissociation constants (K_d) and adenine removal rate constants (k_2) described by a minimal kinetic scheme (Scheme 1). DNA binding affinity for the *Ec*N*Gs*C chimera MutY and selected FSH loop variants was evaluated by monitoring electrophoretic mobility of a non-cleavable DNA

with fluorine-substituted adenine (FA-DNA) as a function of enzyme concentration (Figure 7A and Figure S6). The data were fit to a single-site binding model to derive apparent dissociation constants (K_d) for the FA-DNA–$_{Ec}$N*Gs*C complex (Table 2). The *Ec*N*Gs*C chimera MutY bound OG:FA-DNA tightly with apparent K_d = 0.46 (±0.05) nM at 25 °C and 0.1 M sodium chloride, comparable to the dissociation constant measured for *Ec* MutY by this method, K_d = 0.12 (±0.05) nM.[38] The complex with G:FA-DNA [K_d = 3 (±2)] was noticeably less stable compared to the complex with OG:FA-DNA indicating OG versus G specificity derives partly from DNA-binding energy in the case of *Ec*N*Gs*C, although not to the same degree as previously reported for *Ec* MutY.[38]

Changes in the FSH loop destabilized the DNA-enzyme complex as evidenced by a shift to higher apparent K_d in the binding titrations. The *Ec*N*Gs*C MutY variant with Ala replacing Ser308 (FAH) bound DNA less tightly, with an ~8-fold increase in the apparent K_d for both OG:FA-DNA and G:FA-DNA. The molecular interactions involving Ser308 seen in the TSAC crystal structures apparently contribute to DNA-binding stability. The FAH variant retains a preference for OG:FA-DNA over G:FA-DNA, a finding that was reinforced by patterns in the kinetic data, as discussed in the next section. Replacement of two residues (AAH) or deletion of the FSH residues (Del FSH) had a more drastic effect on OG:FA-DNA binding with K_d climbing to 30 nM. Significantly, these variants no longer showed a preference for DNA containing OG versus G. This study corroborates the idea that the FSH loop and its molecular interactions with DNA are critical for lesion DNA-binding stability and OG recognition.

Kinetic performance of selected FSH variants.

We next turned our attention to the second step of the adenine removal reaction and measured the rate of adenine cleavage, k_2, under single-turnover conditions as previously described (Scheme 1).[20] Figure 7B shows reaction time courses, and Table 3 reports k_2 measured with DNA substrates containing OG:A and G:A mismatches. We tested *Ec*N*Gs*C chimera MutY at both 37 °C and 60 °C, and selected 60 °C as the standard condition for evaluating FSH variants since the adenine removal rate and specificity were improved by increasing the temperature. Apparently, the thermo-preference of MutY is determined, at least partially, by adaptations in the C-terminal domain.

Notably, the features of adenine removal and substrate preference measured for *Ec*N*Gs*C were highly comparable to that of *Ec* MutY. The chimera exhibited “burst” kinetics under conditions of multiple turnover (not shown) similar to *Ec* MutY consistent with high affinity for the product, confirming appropriateness of a similar minimal kinetic scheme.[20] Together with the mutation suppression phenotype in cultures and high affinity for DNA (Table 2), this outcome validates the chimera strategy since *Ec*N*Gs*C was demonstrated to function in cells and retained biochemical behavior consistent with an OG:A-specific adenine glycosylase. The impact of replacing Ser308 with Ala on adenine removal rate was measurable but modest (see FAH *Ec*N*Gs*C in Figure 7B and Table 3). As was the case for DNA binding, this single-alanine replacement had no impact on specificity assessed here as k_2-OG/k_2-G. These outcomes are consistent with no measurable impact on mutation suppression function in cells observed for FAH *Ec*N*Gs*C. The hydrogen bond between the

sidechain of Ser308 and N7-H of OG observed in crystal structures of *Gs* MutY and discussed as the determinant of OG specificity is apparently not as important for function in cells nor a major determinant of OG specificity, although it does contribute to DNA-binding stability and efficiency of adenine removal. We believe that the hydrogen bond between Ser308's peptide amide and O8 of OG, which is expected to remain intact upon substitution with alanine, may play a more significant role in OG versus G specificity.

Double-alanine substitution (AAH *Ec*N*Gs*C) and deletion (Del FSH *Ec*N*Gs*C) severely impacted kinetic performance and specificity for the OG:A substrate. The adenine removal rate for these variants of *Ec*N*Gs*C MutY was ~10-fold diminished for OG:A, but only ~2-fold diminished for G:A containing DNA substrate with a net reduction in OG versus G specificity (Figure 7 and Table 3). Indeed, the vestigial specificity remaining for AAH and Del FSH variants is less than specificity determined for a truncated *Ec*N MutY(residues 1-224), also measured by k_2 ratios.[39] This pattern mirrors the mutator phenotype observed for cultures expressing the AAH and Del FSH variants in the rifampicin resistance assay (Table 1), which retained some mutation suppression function, but were within 5-fold and 3-fold, respectively, of the mutation rates measured for *null* cultures lacking any MutY. Cultures expressing the truncated *Ec*N MutY(1-224) also exhibited a mutator phenotype with Rif^R frequencies within 2.3-fold of *null* cultures.[14] These outcomes tell us that much, if not all, of the function provided by the entire CTD of MutY hinges on the FSH residues.

Finding OG lesions and avoiding undamaged bases involves multiple steps.

The structure presented here shows stabilizing interactions for 1N when opposed by G that appear identical to those seen when the OG site is occupied by its cognate base, explaining how G:A substrates can be processed by MutY. The differing interactions of Ser308 with G and OG, with two hydrogen bonds for the later and an absence of hydrogen bonds for the former, also explains why product inhibition is less for G:A substrates. The complete picture for MutY's aversion for undamaged DNA, including G:A sites, almost certainly involves intermediates which precede the late state mimicked by IN-containing structures. Structure-activity relationships previously revealed the critical importance of the 2-amino group for OG lesion identification.[42] The 2-amino group makes similar molecular interactions with the NTD of MutY for both TSAC-G_{anti}:1N and TSAC-OG_{anti}:1N (see Supplemental Figure S4), but its location in DNA prior to late-stage engagement with MutY will be different and therefore distinguishing. Upon initial encounter with MutY the 2-amino group is expected to be in the minor groove for G_{anti}:A_{anti} and in the major groove for OG_{syn}:A_{anti}. The idea emerging from SAR analysis and structural models of the early MutY-lesion encounter is that the 2-amino group bumps into and engages with an element of MutY located in the CTD.[30,42] Our herein mutational, DNA-binding, and kinetic analysis of Ser308 and its neighbors illustrate the functional importance of the FSH loop and highly suggests this initial 2-amino OG major groove-specific encounter involves the FSH loop of MutY's C-terminal domain. A caveat here is that at this early point in lesion identification, the FSH structure may be different from the one observed in any previously determined structure. Indeed, flexibility in the CTD, deduced from elevated B-values (Figure S2), and also from absence of electron density in the lesion scanning complex,[30] may be essential for structural adjustments in the FSH loop that support a transition from an early-stage encounter to the

late-stage engagement needed to bend the DNA, convert OG_{syn} to OG_{anti} and extrude adenine into the catalytic site.

Significance for targeting MUTYH with chemical biology probes.

Insights from our structural, biological and chemical investigation are significant for guiding the design of MUTYH-specific chemical biology probes. There are several motives for the development of such probes. MUTYH is implicated in colorectal and pancreatic cancer.[7,11] Additionally, inhibition of BER sensitizes cancer cells to chemotherapy, especially in cancers deficient in other DNA damage response pathways.[10,43] Small molecule inhibitors identified by high-throughput screening against the BER enzyme hOGG1 increase the OG burden in cells and are currently being tested in animal models for inflammation and diseases such as cancer.[44,45] Glycosylase inhibitors that mimic the early TS1 transition state, synthesized through an innovative application of copper assisted “click” chemistry to DNA, succeeded in creation of ultra-high affinity ligands and identified design principles that can favor inhibition of a particular glycosylase.[46] These TS1 mimics necessarily target the active site of MUTYH, which was found to be “druggable” as judged by computational analyses.[47] While highly promising, allosteric inhibitors may have an advantage over TS mimics, as the former are expected to be less susceptible to competition with DNA. Our work suggests small molecules targeting allosteric sites found in the CTD of MUTYH may succeed in chemical biology applications. An allosteric inhibitor targeting the FSH loop, for example, is expected to impair glycosylase function, especially altering substrate specificity, and may leave intact other non-canonical signaling functions of MUTYH (reviewed in Raetz and David 2019).[9]

Conclusions

We set out to uncover key elements of OG recognition with a structure of *G. stearothermophilus* MutY-DNA complex containing G:1N. In our TSAC-G:1N structure we found that G occupies the OG site with different hydrogen bonding partners for Ser308 depending upon base identity in the OG detection site. We evaluated the importance of Ser308 and neighboring residues in the FSH loop and found that these residues are critical for function, yet, interestingly, no single residue was irreplaceable. Altering two or more of the FSH residues compromised mutation suppression function *in vivo* and severely reduced biochemical performance *in vitro*, including DNA-binding affinity, efficiency of adenine removal, along with OG:A versus G:A substrate specificity. The FSH loop appears to be critical for most, if not all, of the function provided by the C-terminal domain of MutY and thus defines a highly localized site, remote from, yet mechanistically coupled, to the active site that could be targeted in drug discovery.

Materials and Methods

Detailed methods are described in the Supporting Information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank SIBYLS beamline staff, especially S. Classen and J. Tanamachi, for assistance with data collection. We thank the University of Utah undergraduate students enrolled in Molecular Biology of DNA Lab (BIOL3525, fall 2018) for their contribution to measuring Rif^R mutant frequency data. Technical assistance in early stages of rifampin assay standardization by E. Drage is gratefully appreciated.

Funding Sources

This work was supported by NSF Chemistry of Life Processes (CLP) awards to SSD (award 1610721) and MPH (award 1608934), and by NIH/NCI grant to SSD (CA067985). VLO was supported at the University of Utah by NIH pre-doctoral training fellowships (GM08537 and CA093247). Diffraction data were collected at the Advanced Light Source (ALS), a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the Department of Energy, Office of Basic Energy Sciences, through the Integrated Diffraction Analysis Technologies (IDAT) program, supported by DOE Office of Biological and Environmental Research. Additional support comes from the National Institute of Health project ALS-ENABLE (P30 GM124169) and a High-End Instrumentation grant (S10OD018483).

References

(1). Lindahl T (1993) Instability and Decay of the Primary Structure of DNA. Nature 362, 709–715. [PubMed: 8469282]

(2). Steenken S, and Jovanovic SV (1997) How Easily Oxidizable Is DNA? One-Electron Reduction Potentials of Adenosine and Guanosine Radicals in Aqueous Solution. J. Am. Chem. Soc 119, 617–618.

(3). Michaels ML, and Miller JH (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. [PubMed: 1328155]

(4). Ito R, Hayakawa H, Sekiguchi M, and Ishibashi T (2005) Multiple Enzyme Activities of Escherichia Coli MutT Protein for Sanitization of DNA and RNA Precursor Pools. Biochemistry 44, 6670–6674. [PubMed: 15850400]

(5). Maki H, and Sekiguchi M (1992) MutT Protein Specifically Hydrolyses a Potent Mutagenic Substrate for DNA Synthesis. Nature 355, 273–275. [PubMed: 1309939]

(6). Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Jang W, Karapetyan K, Katz K, Liu C, Maddipatla Z, Malheiro A, Mcdaniel K, Ovetsky M, Riley G, Zhou G, Holmes JB, Kattman BL, and Maglott DR (2018) ClinVar: Improving Access to Variant Interpretations and Supporting Evidence. Nucleic Acids Res. 46, D1062–D1067. [PubMed: 29165669]

(7). Al-Tassan N, Chmiel NH, Maynard J, Fleming N, Livingston AL, Williams GT, Hodges AK, Davies DR, David SS, Sampson JR, and Cheadle JP (2002) Inherited Variants of MYH Associated with Somatic G:C → T:A Mutations in Colorectal Tumors. Nat. Genet 30, 227–232. [PubMed: 11818965]

(8). Banda DM, Nuñez NN, Burnside MA, Bradshaw KM, and David SS (2017) Repair of 8-OxoG:A Mismatches by the MUTYH Glycosylase: Mechanism, Metals and Medicine. Free Radic. Biol. Med 107, 202–215. [PubMed: 28087410]

(9). Raetz AG, and David SS (2019) When You're Strange: Unusual Features of the MUTYH Glycosylase and Implications in Cancer. DNA Repair (Amst). 80, 16–25. [PubMed: 31203172]

(10). Curtin NJ (2012) DNA Repair Dysregulation from Cancer Driver to Therapeutic Target. Nat. Rev. Cancer 12, 801–817. [PubMed: 23175119]

(11). Sharbeen G, Youkhana J, Mawson A, Mccarroll J, Nunez A, Biankin A, Johns A, Goldstein D, and Phillips P (2017) MutY-Homolog (MYH) Inhibition Reduces Pancreatic Cancer Cell Growth and Increases Chemosensitivity. Oncotarget 8, 9216–9229. [PubMed: 27999205]

(12). Gogos A, Cillo J, Clarke ND, and Lu A-L (1996) Specific Recognition of A/G and A/7,8-Dihydro-8-Oxoguanine (8-OxoG) Mismatches by Escherichia Coli MutY: Removal of the C-Terminal Domain Preferentially Affects A/8-OxoG Recognition. Biochemistry 35, 16665–16671. [PubMed: 8988002]

(13). Noll DM, Gogos A, Granek JA, and Clarke ND (1999) The C-Terminal Domain of the Adenine-DNA Glycosylase MutY Confers Specificity for 8-Oxoguanine•Adenine Mispairs and May Have Evolved from MutT, an 8-Oxo-DGTPase. Biochemistry 38, 6374–6379. [PubMed: 10350454]

(14). Li X, Wright PM, and Lu A-L (2000) The C-Terminal Domain of MutY Glycosylase Determines the 7,8-Dihydro-8-Oxo-Guanine Specificity and Is Crucial for Mutation Avoidance. J. Biol. Chem 275, 8448–8455. [PubMed: 10722679]

(15). Chmiel NH, Golinelli M-P, Francis AW, and David SS (2001) Efficient Recognition of Substrates and Substrate Analogs by the Adenine Glycosylase MutY Requires the C-Terminal Domain. Nucleic Acids Res. 29, 553–564. [PubMed: 11139626]

(16). Volk DE, House PG, Thiviyanathan V, Luxon BA, Zhang S, Lloyd RS, and Gorenstein DG (2000) Structural Similarities between MutT and the C-Terminal Domain of MutY. Biochemistry 39, 7331–7336. [PubMed: 10858279]

(17). Michaels ML, Cruz C, Grollman AP, and Miller JH (1992) Evidence That MutY and MutM Combine to Prevent Mutations by an Oxidatively Damaged Form of Guanine in DNA. Proc. Natl. Acad. Sci. U. S. A 89, 7022–7025. [PubMed: 1495996]

(18). Michaels ML, Tchou J, Grollman AP, and Miller JH (1992) A Repair System for 8-Oxo-7,8-Dihydrodeoxyguanine. Biochemistry 31, 10964–10968. [PubMed: 1445834]

(19). Au KG, Clark S, Miller JH, and Modrich P (1989) Escherichia Coli MutY Gene Encodes an Adenine Glycosylase Active on G-A Mispairs. Proc. Natl. Acad. Sci. U. S. A 86, 8877–8881. [PubMed: 2682664]

(20). Porello SL, Leyes AE, and David SS (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. [PubMed: 9778350]

(21). Livingston AL, O'Shea VL, Kim T, Kool ET, and David SS (2008) Unnatural Substrates Reveal the Importance of 8-Oxoguanine for in Vivo Mismatch Repair by MutY. Nat. Chem. Biol 4, 51–58. [PubMed: 18026095]

(22). Samrakandi MM, and Pasta F (2000) Hyperrecombination in Streptococcus Pneumoniae Depends on an Atypical MutY Homologue. J. Bacteriol 182, 3353–3360. [PubMed: 10852864]

(23). Bai H, and Lu A-L (2007) Physical and Functional Interactions between Escherichia Coli MutY Glycosylase and Mismatch Repair Protein MutS. J. Bacteriol 189, 902–910. [PubMed: 17114250]

(24). Fromme JC, Banerjee A, Huang SJ, and Verdine GL (2004) Structural Basis for Removal of Adenine Mispaired with 8-Oxoguanine by MutY Adenine DNA Glycosylase. Nature 427, 652–656. [PubMed: 14961129]

(25). Woods RD, O'Shea VL, Chu A, Cao S, Richards JL, Horvath MP, and David SS (2016) Structure and Stereochemistry of the Base Excision Repair Glycosylase MutY Reveal a Mechanism Similar to Retaining Glycosidases. Nucleic Acids Res. 44, 801–810. [PubMed: 26673696]

(26). Chu AM, Fettinger JC, and David SS (2011) Profiling Base Excision Repair Glycosylases with Synthesized Transition State Analogs. Bioorg. Med. Chem. Lett 21, 4969–4972. [PubMed: 21689934]

(27). Makino K, and Ichikawa Y (1998) Synthesis of a 2-Deoxy-Ribose Type 1-N-Iminosugar. Tetrahedron Lett. 39, 8245–8248.

(28). Schramm VL (2013) Transition States, Analogues, and Drug Development. ACS Chem. Biol 8, 71–81. [PubMed: 23259601]

(29). Lee S, and Verdine GL (2009) Atomic Substitution Reveals the Structural Basis for Substrate Adenine Recognition and Removal by Adenine DNA Glycosylase. PNAS 106, 18497–18502. [PubMed: 19841264]

(30). Wang L, Chakravarthy S, and Verdine GL (2017) Structural Basis for the Lesion-Scanning Mechanism of the MutY DNA Glycosylase. J. Biol. Chem 292, 5007–5017. [PubMed: 28130451]

(31). Classen S, Hura GL, Holton JM, Rambo RP, Rodic I, Mcguire PJ, Dyer K, Hammel M, Meigs G, Frankel KA, and Tainer JA (2012) Implementation and Performance of SIBYLS: A Dual Endstation Small-Angle X-Ray Scattering and Macromolecular Crystallography Beamline at the Advanced Light Source. J. Appl. Crystallogr 46, 1–13.

(32). Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, and Zwart PH (2010) PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. Acta Crystallogr. Sect. D Biol. Crystallogr 66, 213–221. [PubMed: 20124702]

(33). Emsley P, Lohkamp B, Scott WG, and Cowtan K (2010) Features and Development of Coot. Acta Crystallogr. Sect. D Biol. Crystallogr 66, 486–501. [PubMed: 20383002]

(34). Wang L, Lee SJ, and Verdine GL (2015) Structural Basis for Avoidance of Promutagenic DNA Repair by MutY Adenine DNA Glycosylase. J. Biol. Chem 290, 17096–17105. [PubMed: 25995449]

(35). Banerjee A, Yang W, Karplus M, and Verdine GL (2005) Structure of a Repair Enzyme Interrogating Undamaged DNA Elucidates Recognition of Damaged DNA. Nature 434, 612–618. [PubMed: 15800616]

(36). Pei J, Kim BH, and Grishin NV (2008) PROMALS3D: A Tool for Multiple Protein Sequence and Structure Alignments. Nucleic Acids Res. 36, 2295–2300. [PubMed: 18287115]

(37). Pei J, and Grishin NV (2014) PROMALS3D: Multiple Protein Sequence Alignment Enhanced with Evolutionary and 3-Dimensional Structural Information. Methods Mol. Biol 1079, 263–271. [PubMed: 24170408]

(38). Chepanoske C. Lou Golinelli M-P, Williams SD, and David SS (2000) Positively Charged Residues within the Iron-Sulfur Cluster Loop of E. Coli MutY Participate in Damage Recognition and Removal. Arch. Biochem. Biophys 380, 11–19. [PubMed: 10900127]

(39). Francis AW, and David SS (2003) Escherichia Coli MutY and Fpg Utilize a Processive Mechanism for Target Location. Biochemistry 42, 801–810. [PubMed: 12534293]

(40). Chepanoske C. Lou Lukianova OA, Lombard M, Golinelli-Cohen M-P, and David SS (2004) A Residue in MutY Important for Catalysis Identified by Photocross-Linking and Mass Spectrometry. Biochemistry 43, 651–662. [PubMed: 14730969]

(41). Richards JL Recognition and Repair of DNA Damage by Bacterial Adenine Glycosylases, The University of Utah.

(42). Manlove AH, McKibbin PL, Doyle EL, Majumdar C, Hamm ML, and David SS (2017) Structure-Activity Relationships Reveal Key Features of 8-Oxoguanine: A Mismatch Detection by the MutY Glycosylase. ACS Chem. Biol 12, 2335–2344. [PubMed: 28723094]

(43). Visnes T, Grube M, Hanna BMF, Benitez-Buelga C, Cázares-Körner A, and Helleday T (2018) Targeting BER Enzymes in Cancer Therapy. DNA Repair (Amst). 71, 118–126. [PubMed: 30228084]

(44). Donley N, Jaruga P, Coskun E, Dizdaroglu M, McCullough AK, and Lloyd RS (2015) Small Molecule Inhibitors of 8-Oxoguanine DNA Glycosylase-1 (OGG1). ACS Chem. Biol 10, 2334–2343. [PubMed: 26218629]

(45). Tahara YK, Auld D, Ji D, Beharry AA, Kietrys AM, Wilson DL, Jimenez M, King D, Nguyen Z, and Kool ET (2018) Potent and Selective Inhibitors of 8-Oxoguanine DNA Glycosylase. J. Am. Chem. Soc 140, 2105–2114. [PubMed: 29376367]

(46). Yuen PK, Green SA, Ashby J, Lay KT, Santra A, Chen X, Horvath MP, and David SS (2019) Targeting Base Excision Repair Glycosylases with DNA Containing Transition State Mimics Prepared via Click Chemistry. ACS Chem. Biol 14, 27–36. [PubMed: 30500207]

(47). Michel M, Visnes T, Homan EJ, Seashore-Ludlow B, Hedenstrom M, Wiita E, Vallin K, Paulin CBJ, Zhang J, Wallner O, Scobie M, Schmidt A, Jenmalm-Jensen A, Berglund UW, and Helleday T (2019) Computational and Experimental Druggability Assessment of Human DNA Glycosylases. ACS Omega 4, 11642–11656. [PubMed: 31460271]

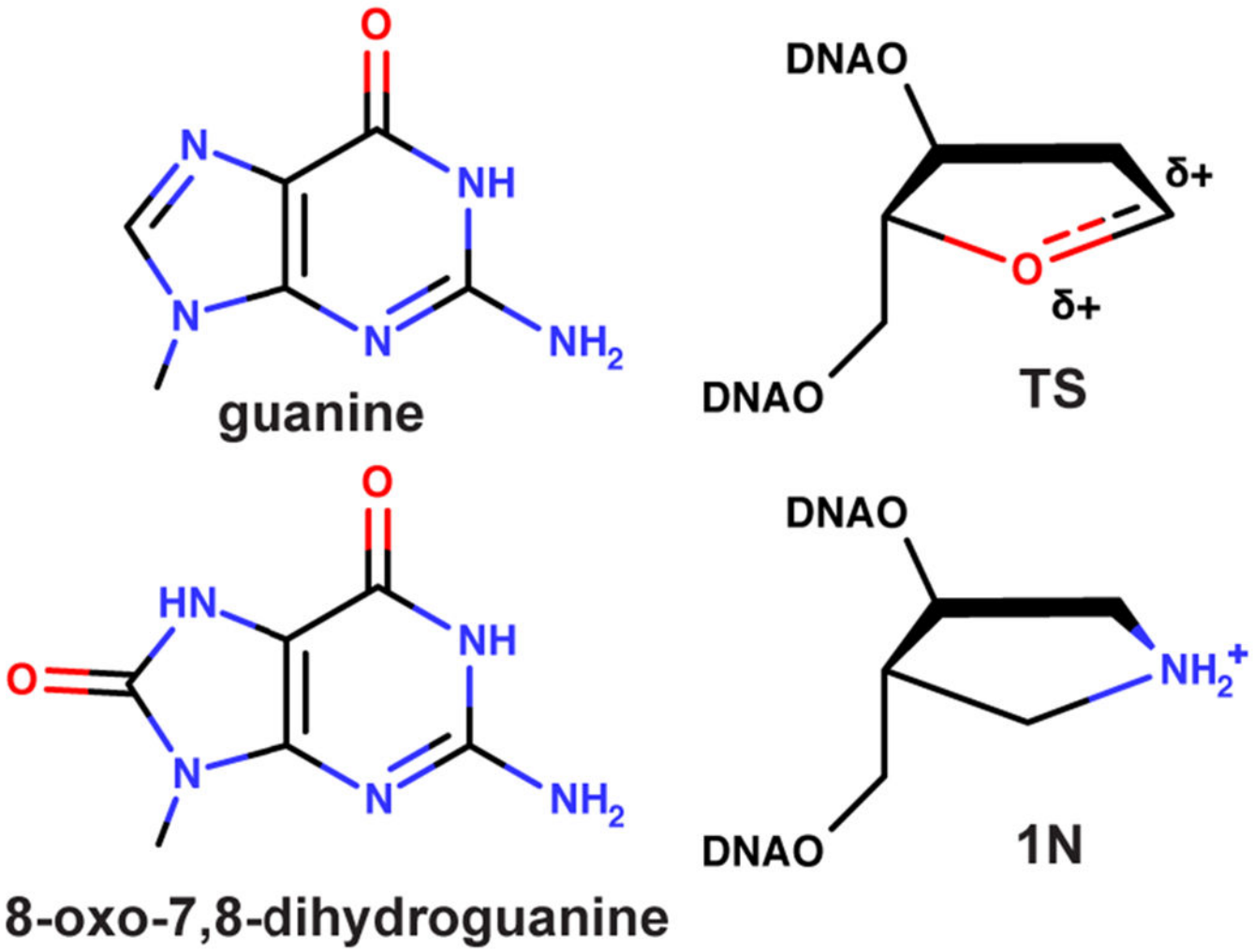


Figure 1.

Overview of OG, TS and 1N analog structures. The chemical structure of guanine (G) and 8-oxo-7,8-dihydroguanine (OG) differ only slightly by changes in atoms attached at positions 7 and 8. The positively charged and extremely unstable oxacarbenium ion transition state is mimicked in terms of shape and charge by the inert 1N TS analog.

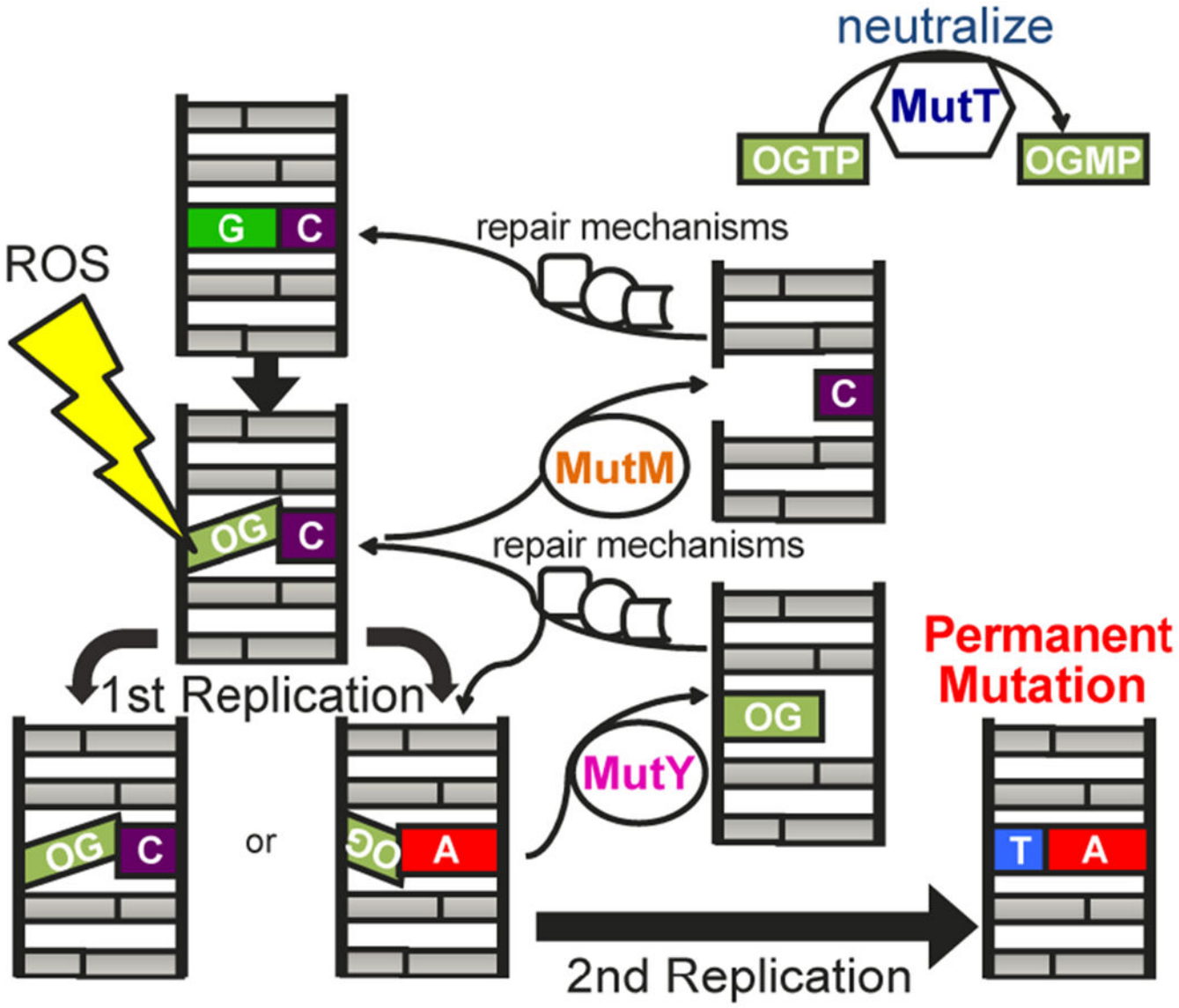


Figure 2.

The GO repair pathway. MutY acts as a final line of defense to prevent G:C to T:A mutations that otherwise would accumulate consequent to reactive oxygen species (ROS).

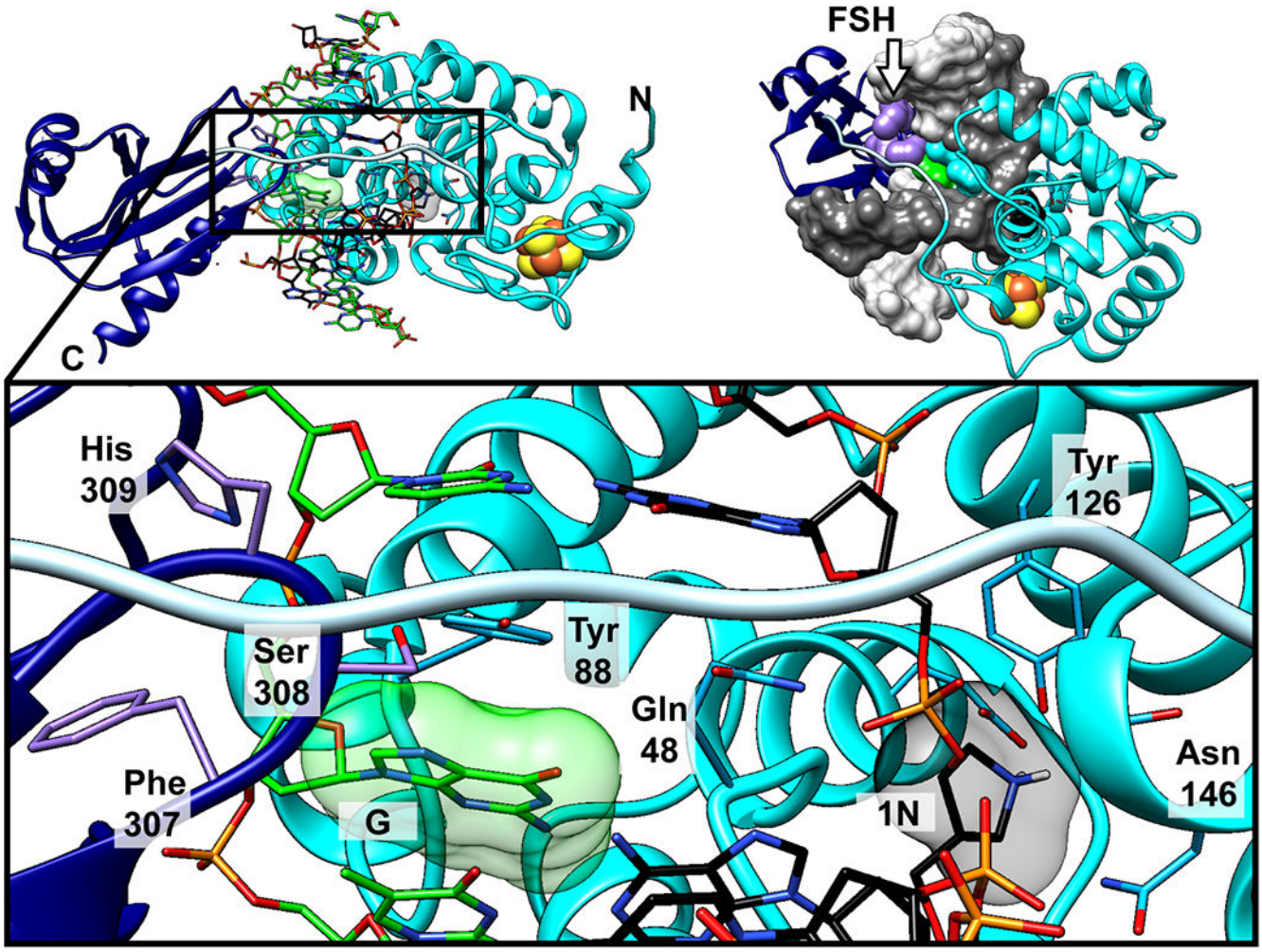


Figure 3.

Overview of Gs MutY in complex with DNA containing G paired with 1N. The N-terminal domain (cyan) and C-terminal domain (navy) surround the DNA (all atom stick model in the left-hand and expanded view; surface in the right-hand view). The 1N TS analog (black transparent surface) engages with catalytic residues of the active site found in the NTD. On the opposite DNA strand, the G base (green transparent surface) adopts the *anti* conformation stacked between a DNA base pair and Tyr88. FSH residues (purple) contact the DNA on the major groove side. The overall architecture of TSAC-G:1N shown here is highly comparable to previously determined structures of MutY in complex with DNA, especially the TSAC-OG:1N structure.

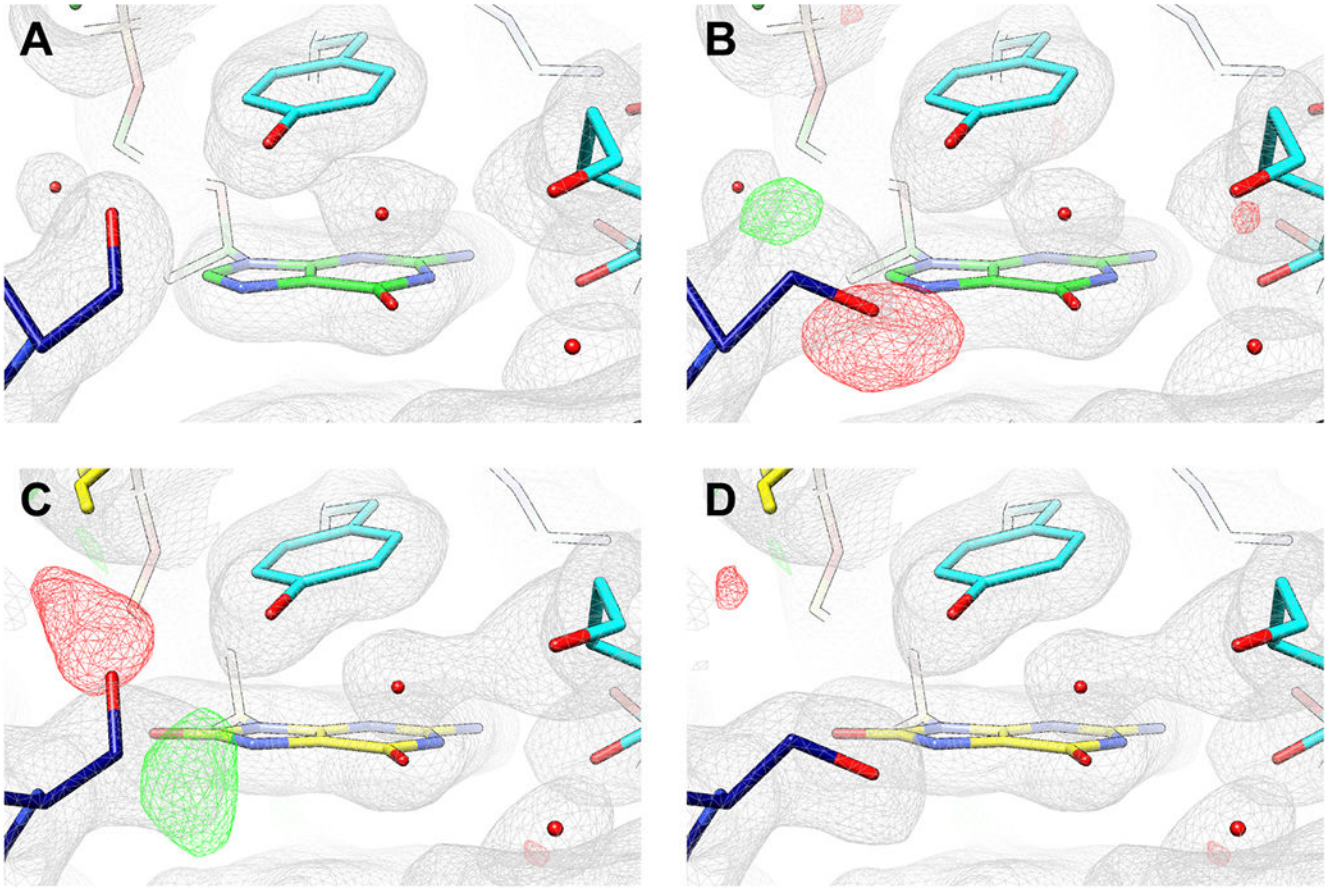


Figure 4.

Difference maps define Ser308 side chain rotamer selection. Ser308 assumes different interactions and altered rotamer positions when G is present in the OG site compared to when OG is present in the OG site. (A) G (green) occupies the OG recognition site as seen in the TSAC-G:1N structure. (B) Ser308 (navy blue) in TSAC-G:1N showed a 5.0σ positive (green) peak and an 8.2σ negative (red) peak when modeled with the same rotamer as found in TSAC-OG:1N indicating incorrect model placement. (C) Ser308 of TSAC-OG:1N showed a 6.4σ positive peak and a 5.2σ negative peak when modeled with the same rotamer as found in TSAC-G:1N. (D) OG (yellow) as found in the TSAC-OG:1N structure. The 2|Fo| −|Fc| difference maps (gray) are contoured at 1.0σ. Positive and negative peaks in the |Fo|−|Fc| difference maps are contoured at 3.5σ.

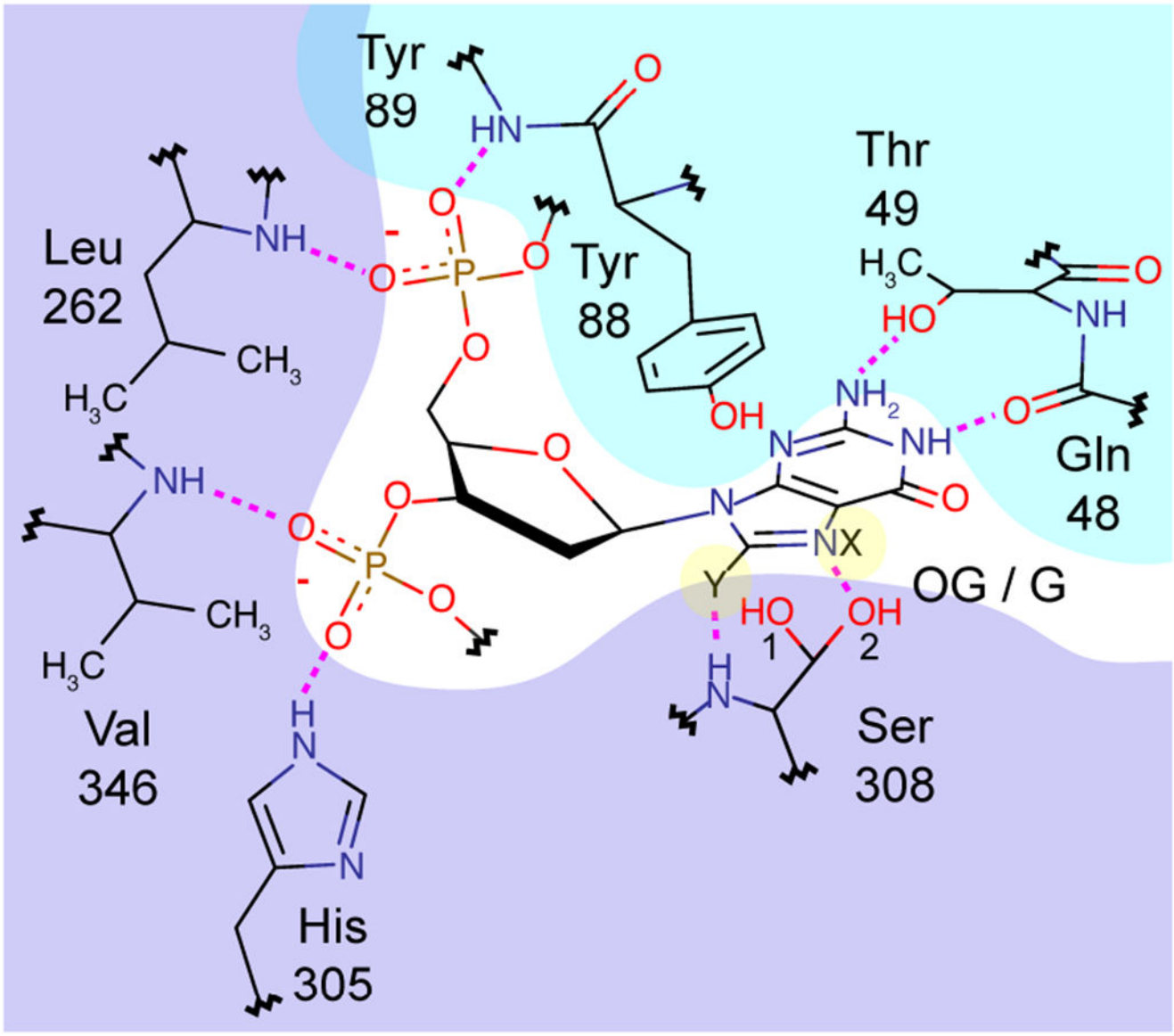
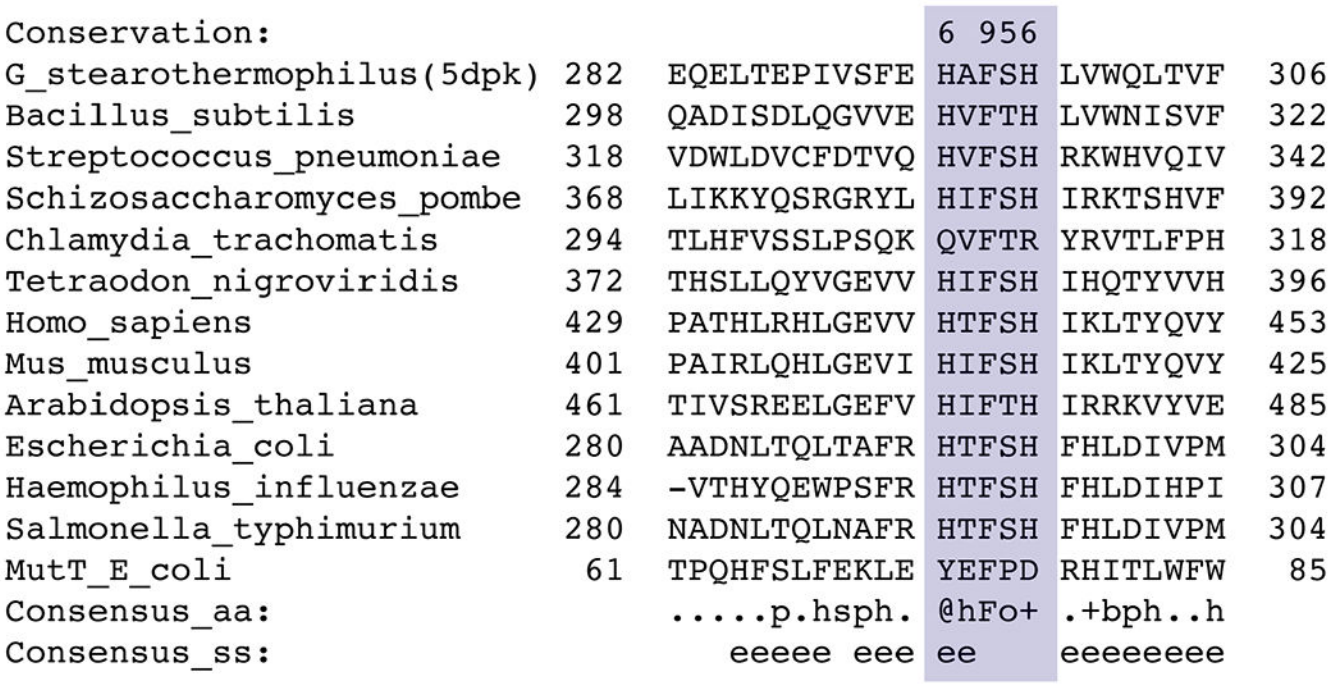


Figure 5.

Molecular interactions at the OG detection site. The OG site is made up of residues from both the NTD (cyan shading) and the CTD (violet shading). The G base interacts with rotamer “1” of Ser308 and OG interacts with rotamer “2”. For the G base X represents a lone-pair electron and Y represents a hydrogen atom. For the OG base X represents a hydrogen and Y represents an oxygen. Ser308 in the CTD is the only residue that makes interactions with OG-specific atoms, and these interactions are altered upon replacement of OG with G.



Conservation:			6 956		
G_stearothermophilus(5dpk)	282	EQELTEPIVSFE	HAFSH	LVWQLTVF	306
Bacillus_subtilis	298	QADISDLQGVVE	HVFTH	LVWNISVF	322
Streptococcus_pneumoniae	318	VDWLDVCFDTVQ	HVFSH	RKWHVQIV	342
Schizosaccharomyces_pombe	368	LIKKYQSRGRYL	HIFSH	IRKTSHVF	392
Chlamydia_trachomatis	294	TLHFVSSLPSQK	QVFTR	YRVTLFPH	318
Tetraodon_nigroviridis	372	THSLLQYVGEVV	HIFSH	IHQTYVVH	396
Homo_sapiens	429	PATHLRHLGEVV	HTFSH	IKLTYQVY	453
Mus_musculus	401	PAIRLQHLGEVI	HIFSH	IKLTYQVY	425
Arabidopsis_thaliana	461	TIVSREELGEFV	HIFTH	IRRKVYVE	485
Escherichia_coli	280	AADNLTQLTAFR	HTFSH	FHLDIVPM	304
Haemophilus_influenzae	284	-VTHYQEWPSFR	HTFSH	FHLDIHPI	307
Salmonella_typhimurium	280	NADNLTQLNAFR	HTFSH	FHLDIVPM	304
MutT_E_coli	61	TPQHFSLFEKLE	YEFPD	RHITLWFW	85
Consensus_aa:		.....p.hsph.	@hFo+	.+bph..h	
Consensus_ss:		eeeee eee	ee	eeeeeeee	

Figure 6.

MutY alignment. The HXF(S/T)H region (shaded) is highly conserved among MutY from different species. Prepared with PROMALS3D.[36,37]

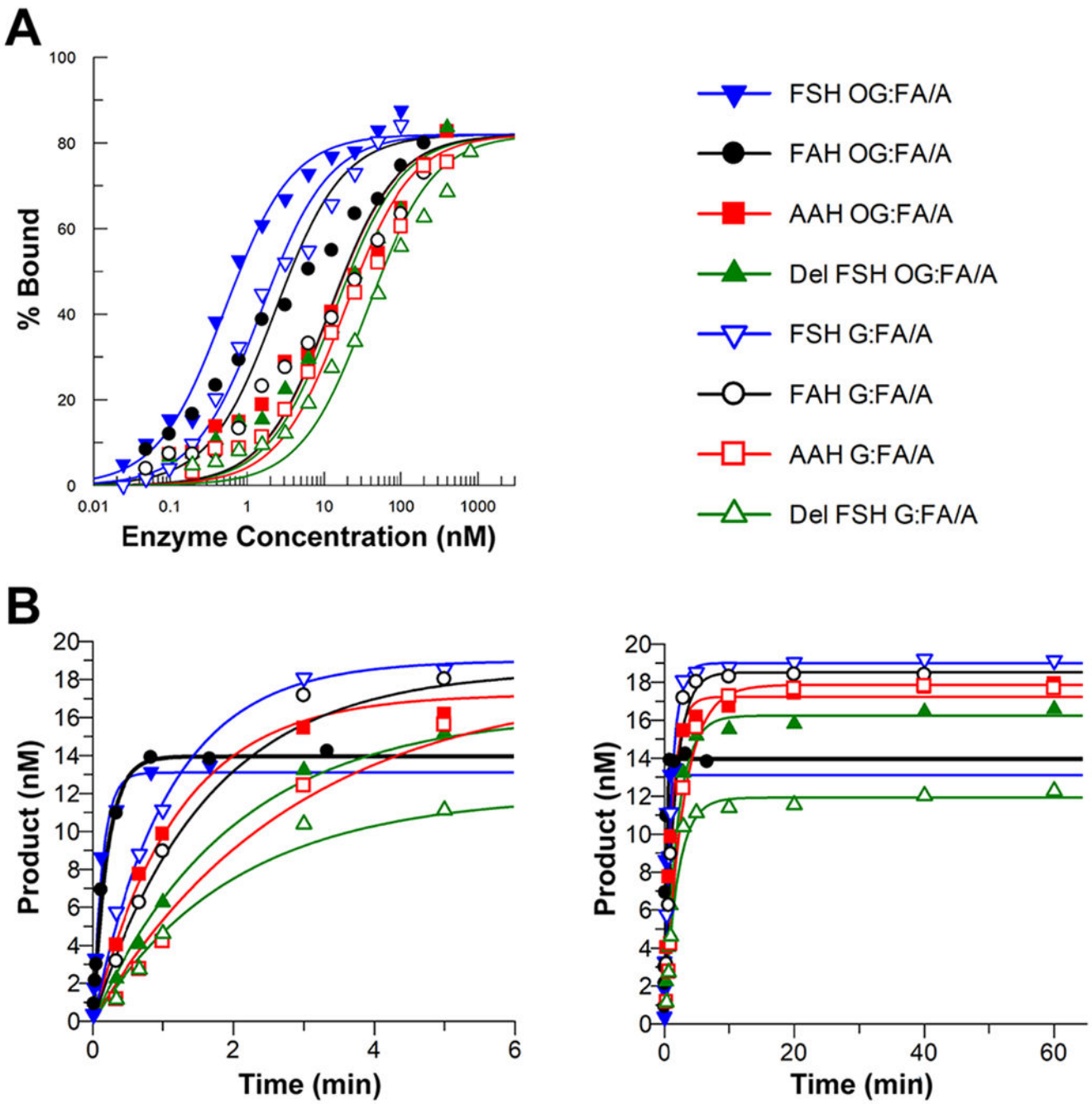
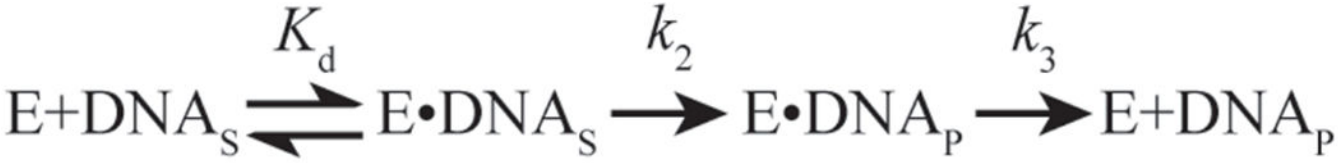


Figure 7. DNA-binding and kinetics.

(A) DNA-binding isotherms measured by EMSA. *Ec*N*Gs*C MutY with FSH residues intact (FSH, blue triangles) and the S308A variant (FAH, black circles) each bound OG:FA-DNA (filled symbols) and G:FA-DNA (open symbols) with markedly different midpoints (see Table 2 for K_d values). The variants AAH (red squares) and Del FSH (green triangles) appear to all have very similar curves for the two substrates indicating a loss of OG versus G specificity. (B) Adenine glycosylase reactions. The left-hand panel highlights early timepoints and the right-hand panel shows full time-courses. *Ec*N*Gs*C MutY and the FAH

variant processed OG:A and G:A with markedly different rates under single-turnover conditions at 60 °C (Table 3). Notably, this specificity indicator diminished or vanished for the variants with double-alanine replacement (AAH) and FSH residues deleted (Del FSH).



Scheme 1.
Minimal Kinetic Scheme for Glycosylase Reaction

Table 1.

Rif^R mutant frequency

MutY Variant	Median Rif^R (95% CI) *	Mutation frequency (per 10^8 cells) ‡	N cultures	Fold change compared to *Ec*NG*s*C (95% CI) *
null	210 (200 – 260)	46	91	29 (17 – 42)
Ec MutY	24 (18 – 36)	8.8	45	3.4 (1.7 – 5.2)
*Ec*N*Gs*C	7.0 (5.5 – 13)	1.2	144	1
YSH *Ec*N*Gs*C	9 (6 – 11)	1.3	51	1.3 (0.67 – 1.8)
ASH *Ec*N*Gs*C	8 (6 – 10)	2.3	75	1.1 (0.61 – 1.7)
F**V**H *Ec*N*Gs*C	8 (5 – 12)	2.4	55	1.1 (0.50 – 1.8)
F**A**H *Ec*N*Gs*C	8 (5 – 13)	1.5	81	1.1 (0.61 – 2.0)
F**PD** *Ec*N*Gs*C	40 (33 – 50)	4.8	50	5.7 (3.2 – 8.3)
AAH *Ec*N*Gs*C	38 (30 – 61)	9.4	68	5.4 (2.7 – 9.1)
Del FSH *Ec*N*Gs*C	80 (54 – 130)	16	68	11 (5.5 – 19)

* 95% confidence intervals obtained by bootstrap estimation.

‡ Mutation frequency reported as the median number of Rif^R cells per 100 million Amp^R cells (see also Methods provided with Supporting Information).

Table 2.

DNA-enzyme complex stability[a]

MutY	K_d (nm) OG:FA	Fold Change relative to *Ec*NGsC	K_d (nm) G:FA	Fold Change relative to *Ec*NGsC	Specificity
Ec MutY	0.12 (±0.05) [b]		5.8 (±0.6) [b]		50 (±20)
*Ec*N*Gs*C	0.46 (±0.05)	1	3 (±2)	1	6 (±4)
FAH *Ec*N*Gs*C	4 (±2)	8 (±4)	24 (±3)	9 (±6)	6 (±3)
AAH *Ec*N*Gs*C	30 (±20)	50 (±40)	18 (±6)	6 (±5)	0.7 (±0.5)
Del FSH *Ec*N*Gs*C	30 (±20)	60 (±40)	27 (±9)	10 (±7)	1.0 (±0.7)

[a]Dissociation constant K_d was measured at 25 °C using 10 pM radiolabeled DNA in reaction mixtures containing 0.1 M sodium chloride. The K_d values are reported as the average of at least three trials. Uncertainties are one standard deviation. Specificity was calculated as the ratio of K_d measured for G:FA and OG:FA DNAs. Uncertainties for fold change and specificity were derived by propagation of error.

[b]K_d values previously reported.[38]

Table 3.

Adenine cleavage rate and specificity[a]

MutY	T (°C)	k_2 (min^{-1}) OG:A	Fold reduction	k_2 (min^{-1}) G:A	Fold reduction	Specificity
Ec MutY	37	12 (±2) [b]		1.80 (±0.05) [b]		7 (±1)
EcN MutY (1-224)	37	0.39 (±0.07) [c]		0.15 (±0.01) [c]		2.6 (±0.5)
Gs MutY	60	54 (±4) [d]		0.44 (±0.04) [e]		120 (±10)
*Ec*N*Gs*C	37	1.6 (±0.1)		0.33 (±0.02)		4.8 (±0.4)
	60	8 (±1)	1	1.0 (±0.1)	1	8 (±1)
FAH *Ec*N*Gs*C	60	5.1 (±0.2)	1.6 (±0.2)	0.6 (±0.1)	1.7 (±0.3)	8 (±1)
AAH *Ec*N*Gs*C	60	1.0 (±0.2)	8 (±2)	0.4 (±0.1)	2.5 (±0.7)	2.5 (±0.8)
Del FSH *Ec*N*Gs*C	60	0.6 (±0.1)	13 (±3)	0.5 (±0.1)	2.0 (±0.4)	1.2 (±0.3)

[a] Rate constant k_2 was measured using 20 nM DNA and excess enzyme in reaction mixtures containing 0.05 M sodium chloride. The k_2 values are reported as the average of at least three trials. Uncertainties are one standard deviation. Specificity was calculated as the ratio of k_2 rate constants measured for OG:A and G:A DNAs. Uncertainties for fold reduction and specificity were derived by propagation of error.

[b] k_2 values previously reported.[40]

[c] k_2 values previously reported.[15]

[d] k_2 values previously reported.[25]

[e] k_2 values previously reported.[41]