Solution Structure and Base Perturbation Studies Reveal a Novel Mode of Alkylated Base Recognition by 3-Methyladenine DNA Glycosylase I*  

Chunyang Cao, Keehwan Kwon, Yu Lin Jiang, Alexander C. Drohat, and James T. Stivers‡  
From the Department of Pharmacology and Molecular Sciences, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205-2185

The specific recognition mechanisms of DNA repair glycosylases that remove cationic alkylpurine bases in DNA are not well understood partly due to the absence of structures of these enzymes with their cognate bases. Here we report the solution structure of 3-methyladenine DNA glycosylase I (TAG) in complex with its 3-methyladenine (3-MeA) cognate base, and we have used chemical perturbation of the base in combination with mutagenesis of the enzyme to evaluate the role of hydrogen bonding and π-cation interactions in alkylated base recognition by this DNA repair enzyme. We find that TAG uses hydrogen bonding with heteroatoms on the base, van der Waals interactions with the 3-Me group, and conventional π-π stacking with a conserved Trp side chain to selectively bind neutral 3-MeA over the cationic form of the base. Discrimination against binding of the normal base adenine is derived from direct sensing of the 3-methyl group, leading to an induced-fit conformational change that engulfs the base in a box defined by five aromatic side chains. These findings indicate that base specific recognition by TAG does not involve strong π-cation interactions, and suggest a novel mechanism for alkylated base recognition and removal.

DNA glycosylases are a remarkable enzyme class that recognize and remove damaged DNA bases as the first step in the DNA base excision repair pathway (1). This important pathway serves as the primary cellular defense against the accumulation of unwanted and toxic DNA base lesions (2). The removal of purine bases that have been alkylated at electronegative heteroatoms is the province of a highly specialized subgroup of these enzymes that specifically recognize and remove these unusual cationic DNA bases. One evolutionary solution to this problem in biological systems is found in the human alkyladenine DNA glycosylase (AAG), which has an α/β fold that is unique for DNA glycosylases (3). However, many more alkylated purine-specific DNA glycosylases, such as 3-methyladenine (3-MeA) DNA glycosylase II (AlkA) from Escherichia coli, contain a highly conserved α-helical domain with a “helix-hairpin-helix” (HHH) DNA binding motif that is found in other DNA glycosylases and DNA-binding proteins (4, 5). Both AAG and AlkA recognize a broad range of alkylated bases (3-methyladenine, 7-methyladenine, and 7-methylguanine), and the substrate range of AAG extends to nonalkylated bases (hypoxanthine and 1-NH2-ethenoadenine) (4, 6). One recently described addition to the HHH DNA repair superfamily is 3-methyladenine DNA glycosylase I (TAG) from E. coli, which is unique in its high specificity for 3-MeA and 3-MeG (7).

Enzymatic recognition and removal of cationic DNA bases such as 3-MeA represents a unique problem in DNA repair. However, repair of these lesions need not involve profound mechanisms for leaving group activation because these bases are electron-deficient and prone to spontaneous hydrolysis at 103- to 104-fold faster rates than neutral purine bases (8, 9). Consistent with this viewpoint, the active site pockets of AlkA and AAG possess no obvious polar groups capable of forming hydrogen bonds to the electronegative acceptor groups on the base to facilitate glycosidic bond cleavage, as might be expected from their abilities to remove a variety of bases (4, 5). AlkA and AAG also possess no obvious binding pockets for the alkyl modification, presumably because of the relatively broad substrate specificity of these enzymes (10). In fact, structural studies on AlkA and AAG revealed that their active sites are lined with conserved tryptophan and tyrosine residues that form stacking and edgewise interactions with the damaged cationic base (3, 11, 12). On the basis of the aromatic character of their active sites, it has been proposed that these enzymes use aromatic π-cation interactions to attract the cationic-base into the active site, and thus discriminate between cationic-damaged purines and neutral undamaged purines (3).

Previous structural studies of TAG suggested both similarities and differences in base recognition and catalysis as compared with AAG and AlkA. From NMR chemical shift perturbation studies, the 3-MeA binding pocket of TAG was also assigned to an extremely aromatic-rich pocket, reminiscent of AAG and AlkA (7). However, TAG also has an absolutely conserved glutamic acid residue that was suggested to form hydrogen bonds with the N6 and N7 positions of 3-MeA in a similar fashion as an analogous group observed in the adenine binding pocket of MutY, an adenine-specific DNA glycosylase of 3-MeA, 3,8-DMP, 3,8-dimethylpurine; r.m.s.d., root mean square deviation; HMQC, heteronuclear multiple quantum coherence.
Enzymatic Recognition of 3-MeA

Shigemi NMR tubes.

The superfamily (13). In contradiction with expectations from the aromatic π-cation hypothesis, it was found that the neutral 3-methyladalenine base bound specifically and tightly to TAG, whereas binding of the normal base adenine could not be detected (7). Thus, the TAG active site can specifically recognize a 3-methyladenine base even when its positive charge has been ablated, suggesting that other unique features of the base are detected by this enzyme.

A deeper understanding of the catalytic mechanisms of alkylpurine-specific glycosylases is hampered by the absence of structural and mechanistic studies on the complexes of these enzymes with cationic and neutral alkylpurine bases. In this study we present the solution structure of TAG in complex with the neutral 3-MeA base. To complement the structural information, we have also synthesized a series of neutral and cationic 3-methyladenine base analogues to probe the role of charge and hydrogen bonding in specific base recognition. In addition, we have performed extensive mutagenesis of the enzyme to examine the energetic interactions of selected enzyme side chains with 3-MeA, as well as each of the 3-MeA base analogues. This comprehensive data set strongly indicates that TAG recognizes 3-MeA using specific hydrogen bonding, conventional π–π stacking interactions, and by van der Waals interactions with the 3-methyl group that lead to an induced fit conformational change in the enzyme. These results may provide evidence for a catalytic mechanism involving weak binding of the cationic substrate base in the ground state and tighter binding of the neutral base in the transition state.

EXPERIMENTAL PROCEDURES

Sample Preparation—Samples of unlabeled and isotope-labeled TAG were prepared as described in Drohat et al. (7). The TAG mutants, Y13A, Y16A, H17A, D18A, W21A, E38A, and W46A were prepared using the QuikChange double-stranded mutagenesis kit from Stratagene (La Jolla, CA) using the pET28b overexpression vector, which were prepared as described in Drohat et al. (7). The TAG mutants, Y13A, Y16A, H17A, D18A, W21A, E38A, and W46A were prepared using the QuikChange double-stranded mutagenesis kit from Stratagene (La Jolla, CA) using the pET28b overexpression vector, which were prepared as described in Drohat et al. (7).

Ultrafiltration Binding Assay—The binding affinities of wild-type and E38A for 3-MeA and 3,6-DMP were determined by ultrafiltration of the free and bound 3-MeA base, planar hydrogen bond restraints between the 3-MeA and the Oγ-H3 of Tyr-16, and the C6-O′-15N atoms of Gln-38 were then added as indicated by the initial structures and the biochemical data.

Fluorescence Spectroscopy—The binding affinities of wild-type and E38A for 3-MeA and 3,6-DMP were determined by ultrafiltration of the equilibrated mixture of enzyme and base analogue. The reaction mixtures (50 μl), consisting of 100 μM of each base analogue and a series of enzyme concentrations (0–550 μM), were incubated for 2 min and filtered for 5 min using microcon-3 (Millipore, Bedford, MA). The free base analogue in the filtrate was separated using a Phenomenex C-18 HPLC column using isocratic buffer elution (10 mM phosphate, 300 mM NaCl) and stirred for 30 min, followed by more methyl iodide (70 ng) with stirring overnight at room temperature. The temperature was increased by 2 °C increments, and the samples were equilibrated for 1 min prior to measuring the fluorescence. A buffer containing 20 mM NaH2PO4, 100 mM NaCl (pH 7.5) was used, with excitation at 280 nm and fluorescence emission monitored at 343 nm using a Spex Fluoromax-3.

Ultrafiltration Binding Assay—The binding affinities of wild-type and E38A for 3-MeA and 3,6-DMP were determined by ultrafiltration of the equilibrated mixture of enzyme and base analogue. The reaction mixtures (50 μl), consisting of 100 μM of each base analogue and a series of enzyme concentrations (0–550 μM), were incubated for 2 min and filtered for 5 min using microcon-3 (Millipore, Bedford, MA).

RESULTS AND DISCUSSION

Solution Structure of the TAG Complex with 3-MeA—We have previously solved the solution structure of free TAG in the absence of any ligands (7). Here we have determined the solution structure of TAG in complex with the 3-MeA product using

1H NMR (D2O, ppm): 8 6.41 (s, H1, H2); 8.14 (s, H1, H3); 8.28 (s, H1, 3-CH3); 7.46 (s, H5, 9-CH3). UV–VIS nm. HRMS (matrix-assisted laser desorption ionization-Fourier transform mass spectroscopy) calc. for C8H10N5 (M + H) 184.093; found, 184.093.

NMR Spectroscopy—NMR experiments were performed at 20 °C on Bruker DMX 500-, 600-, and 750-MHz NMR spectrometers, or Varian Unity Plus 600-MHz NMR spectrometers, all of which were equipped with four channels and pulse-field gradients. The standard suite of experiments for assigning 1H, 13C, and 15N backbone and side-chain chemical shifts and for obtaining NOE-based distance restraints were collected as previously described (7). Intermolecular NOE restraints between the 3-MeA base and 75% deuterated, 13N-labeled, and selectively Tyr-Trp protonated TAG (see above) were obtained using the 2D 1H–15N NOESY pulse sequence with a mixing time of 200 ms. Data were processed using NMRPipe (17) and analyzed using SPARKY version 3 (T. D. Goddard and D. G. Kneller, University of California, San Francisco).

The resonance assignments have been submitted to the BioMagResBank in Madison, WI (accession code 7834). The 1H chemical shifts were referenced to 2.2-dimethyl-2-silapentane-5-sulfonate, and 13C and 15N resonances were indirectly referenced to 2.2-dimethyl-2-silapentane-5-sulfonate (18). pKw of 3-MeA—To obtain the pKw value of the free and bound base, 2D 1H–15N HMQC experiments were collected using a Varian Unity Plus 500-MHz spectrometer for both free and TAG-bound 3-[15N]methyladenine at several pH values. For each experiment sweep widths of 2514 Hz (13C) and 4999 Hz (1H) were used with the carbon and proton carriers set at 130 and 4.82 ppm, respectively. The sample of the complex contained 0.75 mM unlabeled TAG and 0.65 mM 3-[15N]-methyladenine in NMR buffer, such that about 95% of the base was enzyme-bound.

Structural Calculations—Calculations were carried out as previously described (7). Interproton distance restraints obtained from NOE experiments and backbone (ϕ/ψ) dihedral angle and hydrogen bond restraints derived from the observed Hα, HCO, Cα, Cβ, and Cγ chemical shifts, were used as inputs in the torsion angle dynamics simulated annealing protocol in CNS (19). A set of 200 structures was calculated, starting from a high quality CNS-derived structure, using the torsion angle dynamics simulated annealing protocol in the XPLOR-NIH program,2 including a torsion angle base data base of neutral potential forces (20). From this set, 25 low energy structures were selected that exhibited no distance restraint violations of >0.5 Å and no dihedral angle restraint violations of >5.0° (Table 1). After structural calculations using just intermolecular NOE restraints to the base, the 3-MeA and the Oγ-H3 of Tyr-16, and the C6-O′-15N atoms of Gln-38 were then added as indicated by the initial structures and the biochemical data.

Fluorescence Spectroscopy—The binding affinities of wild-type and E38A for 3-MeA and 3,6-DMP were determined by ultrafiltration of the equilibrated mixture of enzyme and base analogue. The reaction mixtures (50 μl), consisting of 100 μM of each base analogue and a series of enzyme concentrations (0–550 μM), were incubated for 2 min and filtered for 5 min using microcon-3 (Millipore, Bedford, MA). The free base analogue in the filtrate was separated using a Phenomenex 5 μC-18 HPLC column using isocratic buffer elution (10 mM phosphate, 300 mM NaCl) and stirred for 30 min, followed by more methyl iodide (70 ng) with stirring overnight at room temperature. The temperature was increased by 2 °C increments, and the samples were equilibrated for 1 min prior to measuring the fluorescence. A buffer containing 20 mM NaH2PO4, 100 mM NaCl (pH 7.5) was used, with excitation at 280 nm and fluorescence emission monitored at 284 nm using a Spex Fluoromax-3.
a similar suite of heteronuclear NMR experiments as previously employed to obtain the structure of the free enzyme (7), and more recently, to refine the structure of the zinc binding site of TAG (21). One special approach that was key to obtaining resonance assignments for the aromatic residues in the active site of the complex was a selective labeling experiment in which the \(^{15}\text{N}\)-labeled enzyme was partially deuterated by overexpression in 75% D\(_2\)O, and just prior to induction, unlabeled Trp and Tyr were added to the minimal growth media to selectively protonate these groups. This approach reduced spin diffusion pathways and provided much more robust 2D NOESY spectra that allowed assignment of all key active site aromatic side chains and unambiguous measurements of intermolecular NOE values between these groups and the 3-MeA base. The overall quality of the structure for the complex is very similar to that of the free enzyme (r.m.s.d. = 0.64 Å, Fig. 1A), and the complete structural statistics are reported in Table I. Binding of the 3-MeA base (which is neutral at the pH of the NMR studies, see below) induces only small changes in the overall structure as compared with the free enzyme but alters the environment of Trp-6 (see below), giving rise to the large tryptophan fluorescence decrease that accompanies 3-MeA binding (7). A similar induced fit mechanism has been previously described for uracil DNA glycosylase upon binding to uracil (22). Mapped onto the structure shown in Fig. 1B are the side chains of six active site amino acid residues that are completely conserved in all TAG family members (Trp-6, Tyr-13, Tyr-16, Trp-21, Glu-38, and Trp-46) (23). In the studies described below, these residues and two others, His-17 and Asp-18, were mutated to alanine.

**Structural Basis for 3-MeA Recognition**—The position of the 3-MeA base in the active site was well defined by 27 intermolecular NOE restraints from the enzyme to the H8, H2, and 3-N-methyl hydrogens of the base. An overlay of the 3-MeA base and several surrounding residues that comprise the binding pocket, using the 10 lowest energy structures, is shown in Fig. 2A, and a molecular model of the lowest energy structure is shown in Fig. 2B. This structure reveals that the base is stacked with a single conserved tryptophan (Trp-46), and that Glu-38 serves as a bifunctional hydrogen bond donor and acceptor to N7 and the 6-amino group of 3-MeA. The hydroxyl of Tyr-16 is positioned to donate a hydrogen bond to the N1 position of the base, and Trp-6 appears to serve as a lid to encapsulate the 3-MeA base in an aromatic box defined by itself and Tyr-13, Tyr-16, and Trp-46 (Fig. 2B). Additional compelling evidence for a hydrogen bond between Glu-38 and the exocyclic amino group of 3-MeA is provided by the observation of a broad proton resonance in the complex with a chemical...
Table I

| Parameter                                         | (25) |
|---------------------------------------------------|------|
| Backbone atoms residues 11–174                    | 0.69 ± 0.10 |
| All heavy atoms residues 11–174                   | 1.27 ± 0.14 |
| Backbone atoms for all residues (1–188)           | 1.04 ± 0.24 |
| All heavy atoms for all residues (1–188)          | 1.51 ± 0.19 |
| Backbone torsion angle restraints (°)              | 0.0028 ± 0.0002 |
| Angles (°)                                         | 0.497 ± 0.004 |
| Improper (°)                                       | 0.419 ± 0.02 |
| Lennard-Jones potential energy (kcal mol⁻¹)       | 699 ± 23 |
| Bad contacts per 100 residues                      | 17.9 ± 2.8 |
| Ramachandran analysis (%)                         | 86.1 ± 1.4 |
| Most favored                                       | 9.6 ± 1.6 |
| Additionally allowed                               | 3.4 ± 0.9 |
| Generously allowed                                 | 0.8 ± 0.5 |
| Disallowed                                         | 0.2 ± 0.05 |

a The 25 structures calculated and energy minimized in CNS.
b None of the ensemble of structures displayed a distance violation of >0.5 Å.
c Backbone torsion angle restraints were derived using TALOS (29). None of the structures exhibit a dihedral angle violation of >5°.
d Lennard-Jones potential energy calculated using the CHARMM parameters.
e Calculated using the program PROCHECK (32).

Our previous finding that TAG could discriminate exquisitely between neutral 3-MeA and adenine suggested that the enzyme could sense the one difference between these bases: the 3-methyl group. As shown in the molecular surface representation of Fig. 2C, the methyl group is nestled in a pocket defined by the aromatic rings of Trp-6, Tyr-13, and the side chain methylene group of Trp-46. Favorable van der Waals interactions of the methyl group in this nonpolar pocket may be used to drive the induced-fit clamping of the enzyme, which creates the aromatic box around the 3-MeA base. Indeed, it is difficult to envision binding of 3-MeA in this aromatic box without it first being presented in an open conformation.

The structure of free TAG is consistent with this idea, because the Trp-6 lid is “ajar,” allowing improved access of 3-MeA to the binding pocket (6). This induced-fit mechanism involving nonpolar interactions with the 3-methyl group may explain the exceptional discrimination against adenine. Two other groups shown in Fig. 2C that are in van der Waals contact with 3-MeA are Ala-168 and Ser-164. Both of these groups are highly conserved in the TAG family, as would be expected from their close proximity to the bound base.

**The Bound 3-MeA Is Neutral**—We have previously shown that TAG recognizes neutral 3-MeA, but not adenine, suggesting that positive charge was not required for specific binding (7). One caveat to this previous conclusion was that it could not be excluded that the enzyme protonated the neutral base upon binding, and that specific recognition did indeed involve favor-
able π-cation interactions between the enzyme and base. To directly address this question we incorporated a site-specific 13C label at the 8-carbon of the purine ring of 3-MeA and determined the pKₐ value of the free base using a 2D 1H-13C HMQC experiment (Fig. 3A). From this experiment, the pKₐ (3-MeA) = 5.6 ± 0.1, which is 1.9 log units lower than the pH value employed in the previous 3-MeA binding studies. To establish that 3-MeA remained neutral upon binding to the TAG active site, we measured the 8-13C shift of 3-MeA while it was bound in the active site pocket at pH 6.5 and 7.5 (Fig. 3B). The chemical shifts at both pH values were similar to free 3-MeA (Fig. 3A, closed circles), providing strong evidence that the enzyme does not significantly change the pKₐ value of the base, and that the 3-MeA binding measurements at pH 7.5 reflect binding of the neutral base. These findings establish the previous conclusion that specific recognition of 3-MeA need not involve positive charge on the base.3

Overview of Base Perturbation and Mutagenesis Approach—To further illuminate the energetic basis of 3-MeA recognition by TAG, we employed a combined approach in which mutations in the enzyme and alterations in the 3-MeA base were systematically introduced. Our strategy was 2-fold. First, we wished to determine whether introduction of a stable positive charge on the base by alkylation or protonation at N9 resulted in tighter or weaker binding (see chemical structures below in Fig. 5). Second, we sought to probe the importance of the observed hydrogen bonds between Glu-38 and the N7 and 6-NH₂ groups of 3-MeA. This second issue was approached using a double perturbation method in which the energetic effect of adding a methyl group at N7, or substituting a methyl group for the 6-amino group, was assessed for both the wild-type enzyme and the E38A mutant. We surmised that if Glu-38 interacted with both of these positions, then the damaging effect of these base perturbations on binding would be greater for the wild-type enzyme than for the E38A mutant. Of course, definitive conclusions from such studies can be compromised by indirect effects, but when combined with high resolution structural information, such approaches can provide strong collaborative mechanistic information.

Enzymatic Recognition of 3-MeA

3 An issue that is beyond the scope of this report is the tautomeric state of 3-MeA in DNA and in the TAG active site. Small molecule crystallographic studies of 3-methyl2'-deoxyadenosine (30) have indicated that the base is in the N6-iminium tautomer, with an estimated pKₐ of 12 based on comparison with the analogous iminium tautomer of 4-amino-N3-methylpyrimidine (31). The nearly identical proton NMR shifts of free and bound 3-MeA measured in the present work provide no evidence to support a change in tautomeric state upon binding. Therefore, the free and bound states are likely to be the same tautomer. Addressing these issues will require heteronuclear NMR approaches and isotopically labeled 3-MeA and its analogues.
sensing the effects of alanine mutagenesis on 3-MeA binding, we first measured the heat denaturation profiles for the Y13A, Y16A, D18A, W21A, E38A, and W46A mutations by following the decrease in tryptophan fluorescence at 45 °C, whereas the other mutants were shifted to T^{m}_{D} values that were 5 to 10 °C lower. We speculate that the multiphasic melting profiles may reflect loss of the tightly held zinc ion, followed by melting of the remaining secondary structure. Consistent with this viewpoint, removal of the His-17 side chain, which is one of the four zinc ligands, resulted in a protein that was insoluble and impossible to obtain even in small quantities. Given these results, we performed the base binding measurements at 15 °C to ensure that the mutant enzymes were in the fully folded state.

Of the six mutations that were characterized, only Y16A, E38A, and W46A had a significant effect (13- to 14-fold) on the decrease in tryptophan fluorescence (Fig. 4A). One general observation from these studies was that wtTAG and all of the mutants are relatively unstable, and that the denaturation profiles displayed complex non-two-state behavior. Wild-type TAG, Y13A, W46A, and E38A showed half-maximal decreases in tryptophan fluorescence at about 45 °C, whereas the other mutants were shifted to T^{m}_{D} values that were 5 to 10 °C lower. We speculate that the multiphasic melting profiles may reflect loss of the tightly held zinc ion, followed by melting of the remaining secondary structure. Consistent with this viewpoint, removal of the His-17 side chain, which is one of the four zinc ligands, resulted in a protein that was insoluble and impossible to obtain even in small quantities. Given these results, we performed the base binding measurements at 15 °C to ensure that the mutant enzymes were in the fully folded state.

Of the six mutations that were characterized, only Y16A, E38A, and W46A had a significant effect (13- to 14-fold) on 3-MeA binding (Fig. 4B). These mutagenesis findings are fully consistent with the structural findings and indicate that the other conserved residues Tyr-13, Asp-18, and Trp-21 are not involved in stabilizing the bound base, and may therefore be involved in other steps of the overall reaction. We then proceeded to probe the interactions of Tyr-16, Glu-38, and Trp-46 using the combined approach of mutagenesis and base perturbation.

Base Perturbation Studies—One approach to address the role of charge in 3-MeA binding is to methylate the N9 position of the base, thereby locking the base into a cationic charge state. If aromatic n-cation interactions are important for TAG recognition, it would be expected that 3,9-dimethyladenine (3,9-DMA) would bind more tightly than neutral 3-MeA. This expectation is reasonable, because the 9-methyl group can be sterically accommodated within the structure of the active site (Fig. 2C), and it is less bulky than the sugar ring in the intact 3-methyladenine nucleotide. Nevertheless, a steric effect of the methyl group on binding cannot be completely excluded, so as an alternative approach, we also measured binding of 3-MeA at pH 5.7, a pH value in which the base is ~44% protonated at N9 (i.e. the potentially interfering methyl group is replaced with a proton). As shown in the fluorescence-based binding measurements in Fig. 5 (A and B), the cationic base 3,9-DMA binds 7.5-fold more weakly to wtTAG than neutral 3-MeA, and the 44% protonated 3-MeA base binds about 1.7-fold more weakly (K_{D} = 70 ± 7 μM). Correcting for the partial protonation of 3-MeA suggests that the fully protonated base binds about 4-fold more weakly than the neutral base, confirming that steric effects are small for binding of 3,9-DMA, and that positive charge is detrimental to binding.

Removal of the aromatic groups Trp-46 and Tyr-16 also provides support that charge is not beneficial to 3-MeA binding. We found that the damaging effect of the W46A mutation is only 3-fold for binding of the cationic base 3,9-DMA (K_{D} = 1000 ± 400 μM), as compared with 13-fold for binding of neutral 3-MeA (K_{D} = 500 ± 200 μM), indicating that Trp-46 interacts more strongly with the neutral base then with the cationic base (Fig. 5C). Interestingly, for the Y16A mutation, binding of the cationic base 3,9-DMA is 6-fold tighter than 3-MeA (Fig. 5D). This result is more difficult to interpret with confidence but may indicate, in addition to contributing a hydrogen bond to N1 (Fig. 2B), that Tyr-16 contributes to the nonpolar character of the active site by exclusion of water. Thus removal of Tyr-16 may diminish the hydrophobic character of the site by allowing water to enter and, therefore, favor binding of the cationic base over the neutral base. Regardless of the detailed interpretation, the most important conclusion from these findings is that positive charge does not lead to enhanced binding of 3-MeA, and in contrast with expectations from the n-cation recognition hypothesis, the nonpolar, aromatic-rich active site favors binding of the neutral base.

To establish the energetic importance of the hydrogen bonds between Glu-38 and the N7 and 6-amino groups of 3-MeA, we measured the binding affinities of 3,7-dimethyladenine (3,7-DMA) and 3,6-dimethyluridine (3,6-DMP) to wtTAG and E38A using fluorescence and ultrafiltration binding methods. Like the 9-methyl substituent, the 7-methyl group introduces a positive charge (see structure in Fig. 6E) but also introduces steric bulk that would be expected to collide with the side chain of Glu-38 and disrupt binding (Fig. 2B). On the other hand, the 6-methyl group is an excellent nonpolar steric mimic for the exocyclic 6-amino group (24) and would be expected to disrupt binding if the hydrogen bond between Glu-38 and the 6-amino group is important. Consistent with the above expectations, no detectable binding of 3,7-DMA or 3,6-DMP was observed for wtTAG (Figs. 5A and 6B). In contrast, the E38A mutant bound to both analogues with a similar affinity as 3-MeA (Fig. 6, A and B), providing strong confirmatory evidence that Glu-38
interacts with both the N6 and N7-positions of the base. The difference free energies for binding of 3,7-DMA and 3,6-DMP as compared with 3-MeA are shown in Fig. 6C for wtTAG and E38A (where \( \Delta G = \Delta G_{\text{base analogue}} - \Delta G_{\text{3-MeA}} \)). The binding affinities of 3,7-DMA and 3,6-DMP to wtTAG were too weak to be determined, so only lower limit difference free energies are shown. D, binding of wtTAG and E38A to 3-MeA at pH 7.5 and 9.0. The \( K_D \) values for wtTAG were 42 ± 8 \( \mu \)M at pH 7.5 and 179 ± 7 \( \mu \)M at pH 9.0. For E38A, the \( K_D \) values were 600 ± 200 \( \mu \)M at pH 7.5 and 400 ± 100 \( \mu \)M at pH 9.0.

Fig. 6. The importance of the N7 and 6-amino groups of 3-MeA in specific recognition. A, binding of 3-MeA and 3,7-DMA to E38A. B, an ultrafiltration binding assay was used to obtain the binding constants of 3,6-DMP to wild-type TAG (○) and E38A (●). The \( K_D \) value for E38A was 700 \( \mu \)M, whereas the affinity for wtTAG was too weak to be measured. C, difference free energies for binding of TAG or E38A to 3,7-DMA and 3,6-DMP as compared with 3-MeA (\( \Delta G = \Delta G_{\text{base analogue}} - \Delta G_{\text{3-MeA}} \)). The binding affinities of 3,7-DMA and 3,6-DMP to wtTAG were too weak to be determined, so only lower limit difference free energies are shown. D, binding of wtTAG and E38A to 3-MeA at pH 7.5 and 9.0. The \( K_D \) values for wtTAG were 42 ± 8 \( \mu \)M at pH 7.5 and 179 ± 7 \( \mu \)M at pH 9.0. For E38A, the \( K_D \) values were 600 ± 200 \( \mu \)M at pH 7.5 and 400 ± 100 \( \mu \)M at pH 9.0.

Glu-38 is protonated in the presence of 3-MeA. Supporting this proposal, the \( K_D \) value of wtTAG for 3-MeA was 4.5-fold weaker at pH 9 (180 ± 7 \( \mu \)M, Fig. 6D), which is nearly one-half the damaging effect observed upon complete removal of the Glu-38 side chain (Fig. 4). This result suggests that an enzyme group involved in binding becomes deprotonated in the pH range 7.5–9. Evidence that Glu-38 is the protonated group is provided by the observation that the binding affinity of E38A for 3-MeA was similar at pH 7.5 (\( K_D = 600 \pm 200 \)) and at pH 9 (\( K_D = 400 \pm 100 \)) (Fig. 6D). We attempted to directly measure the \( pK_a \) of Glu-38 using the 2D H\textsubscript{2}Y(C/O)CO NMR experiment that correlates the H\textsubscript{y} hydrogens of glutamates with their side-chain carboxyl carbons (25). However, the correlation peaks were
very weak in this experiment, and no correlations corresponding to the expected chemical shifts of Glu-38 were observed. The high pK_a suggested for Glu-38 may arise from a multitude of effects: its location in a nonpolar environment wedged between helices α2 and α9 (Fig. 1B), hydrogen bonding to 3-MeA, and its position near the carboxyl terminus of α2. A New Role for a Carboxylic Acid Group in Alkylation Base Repair—Our previous NMR structure of free TAG, and the current structure of its complex with 3-MeA, have revealed mechanistic surprises with respect to water activation, glycosylation stabilization, and alkyl base recognition. Previous to structural analysis of the free enzyme, it was believed that all HhH superfamily glycosylases possessed a conserved aspartate residue that served to deprotonate water or stabilize the positive charge that develops on the sugar during glycosidic bond cleavage (26). On the contrary, TAG was not found to have an analogous aspartate group (12, 13), and therefore, could not employ a similar mechanism to facilitate glycosidic bond cleavage. Instead, TAG appears to use a high pK_a glutamic acid residue (Glu-38) to promote recognition and removal of 3-MeA by hydrogen bonding to the 6-amino and N7 positions. Although detailed kinetic studies of TAG have yet to be completed, the present results suggest that Glu-38 serves to selectively stabilize the base in the transition state. This proposal is reasonable because the pK_a of N7 would increase in moving from the ground state to the transition state, leading to stronger hydrogen bonding to Glu-38 as the transition state is approached (Fig. 7A) (27). It seems that the recognition mechanism of TAG is closely related to that of MutY, an adenine glycosylase that uses a conserved glutamic acid residue to recognize and activate the leaving adenine when it is opposite to the oxidized base 8-oxoguanine in DNA (13).

Implications for Catalysis—A hallmark feature of alkyl purine-specific DNA glycosylases is their exceedingly low rate enhancements as compared with other DNA glycosylases (9). For the most powerful DNA glycosylases, such as uracil DNA glycosylase and human 8-oxoguanine DNA glycosylase, rate enhancements of 10^6 to 10^12 are observed as compared with the spontaneous hydrolysis reactions of the free deoxynucleosides (9), whereas TAG and AAG have meager enhancements around 10-fold (9). We conclude from this analysis that alkylpurine-specific enzymes need not employ extraordinary mechanisms to accelerate these reactions and that specificity may simply arise from a passive mechanism that exploits the intrinsic differences in reactivity of alkylated and neutral nucleotides. The relatively small 10- to 15-fold damaging effects of removing active site residues Trp-46, Tyr-16, and Glu-38 are therefore not unexpected given the small catalytic prowess required of TAG. Indeed, each of these damaging effects is sufficient to account for the rate enhancement provided by TAG. It should not be noted that the small catalytic power of TAG makes it difficult to assign catalytic roles to specific groups because of the small energetic effects of mutations, which may be only modestly greater than nonspecific effects of these perturbations. Nevertheless, the combined structural, mutagenesis, and base perturbation approaches have provided a consistent view for the roles of these active site groups.

One simple mechanism for accelerating the removal of an alkylated base but not a normal base would be to place the damaged base in a nonpolar binding pocket where it is destabilized relative to aqueous solution (Fig. 7). Thus, the nonpolar active sites of these alkylated base-specific enzymes could provide an environment that promotes ground-state electrostatic strain, thereby selectively lowering the activation barrier for glycosidic bond cleavage of cationic nucleobases. In this mechanism, substrate binding energy is used to drive the cationic damaged base into an active site of low dielectric, where it is then destabilized. As the glycosidic bond is broken, and the electrons are released into the aromatic system of the base, the electrostatic strain decreases and the departing base, which is now neutral, binds more tightly as the transition state is approached. Thus the nonpolar environment lowers the activation barrier by straining the charged ground state, and by binding tightly to the neutral base in a dissociative transition state. This would provide a plausible mechanism to promote specificity at the catalytic step, because neutral bases would bind tightly in the ground state and more weakly as the anionic transition state is approached, which is anticytotoxic.

Although this mechanistic proposal is based upon the binding of a product base, it is more consistent with the available data than the alternative π-cation recognition mechanism. Indeed, similar NMR studies of the chemical properties and ionization state of the uracil product of the uracil DNA glycosylase reaction provided key insights into the nature of the enzymatic transition state not obvious from crystallographic studies of the UDGS-substrate complex (22, 28, 29). TAG provides an excellent experimental system to further investigate the mechanistic basis for alkylated base removal.

Summary—We have discovered a new mode of alkylpurine base recognition by a DNA glycosylase that differs significantly from previously described recognition strategies. TAG shares the hydrogen bonding recognition modes of its cousin MutY, which detects adenine in mismatched base pairs with 8-oxoG, and the aromatic stacking interactions of its other relative AlkA, which acts on alkylpurines. Thus the TAG recognition mechanism represents an evolutionary hybrid of its two closely related HhH family members. We postulate that the preference for binding the neutral form of 3-MeA of TAG reflects a catalytic strategy involving ground-state electrostatic strain upon binding of the cationic base of the DNA substrate and selective stabilization of the neutral base in the transition state. The unique base recognition mode of TAG described here accounts for its exquisite specificity for 3-MeA and 3-MeG and its discrimination against other purine bases.

REFERENCES

1. Mol, C. D., Parikh, S. S., Putnam, C. D., Lo, T. P., and Tainer, J. A. (1999) *Annu. Rev. Biochem.* 68, 1–102
2. Lindahl, T., and Wood, R. A. (1999) *Science* 286, 1497–1505
3. Lau, A. Y., Wyatt, M. D., Glassner, B. J., Samson, L., and Ellenberger, T. (2000) *Proc. Natl. Acad. Sci. U. S. A.* 97, 13573–13578
4. Hollis, T., Lau, A., and Ellenberger, T. (2000) *Mol. Cell.* 6, 201–210
5. Hollis, T., Lau, A., and Ellenberger, T. (2001) *Proc. Natl. Acad. Sci. U. S. A.* 98, 305–314
6. Asada, A., Ito, H., Asagoughi, K., Matsuyama, S., Tano, K., Murakami, A., Takamori, Y., and Kish, K. (2000) *Biochemistry* 39, 1897–1905
7. Drohat, A. C., Kwon, K., Krosky, D. J., and Stivers, J. T. (2002) *Nat. Struct. Biol.* 9, 659–664
8. Lindahl, T. (1992) *Annu. Rev. Biochem.* 61, 61–87
9. Stivers, J. T., and Jiang, Y. L. (2003) *Chem. Rev.* 103, 2729–2760
10. Berdal, K. G., Johansen, R. F., and Seeberg, E. (1998) *EMBO J.* 17, 363–367
11. Lau, A. Y., Scharer, O. D., Samson, L., Verdine, G. L., and Ellenberger, T. (1998) *Cell* 95, 249–258
12. Hollis, T., Ichikawa, Y., and Ellenberger, T. (2000) *EMBO J.* 19, 758–768
13. Guia, Y., Manuel, R. C., Arvai, A. S., Parikh, S. S., Mol, C. D., Miller, J. H., Lloyd, S., and Tainer, J. A. (1998) *Nat. Struct. Biol.* 5, 1058–1064
14. Fujii, T., Saito, T., Kira, H., Hayashibara, H., Kumaizawa, Y., and Nakajima, S. (1986) *Heterocycles* 24, 2449–2454
15. Lam, F. L., and Parham, J. C. (1978) *Heterocycles* 9, 287–291
16. Ogivie, K. K., Beaucage, S. L., and Gillen, M. F. (1978) *Tetrahedron Lett.* 35, 3203–3206
17. Delaglio, F., Grzesiek, S., Vuister, W. G., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* 6, 277–283
18. Wishart, D. S., Bigam, C. G., Yao, J., Ahlhaberg, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) *J. Biomol. NMR* 6, 135–140
19. Brungger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Goo, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuzezowski, J., Nilges, M., Pannu, N. S., Read, R. J., Read, I. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr.* Sect. D: Biol. Crystallogr. 54, 805–921
20. Kuzezowski, J., Gronenborn, A. M., and Clore, G. M. (1996) *Protein Sci.* 5, 1087–1090
21. Kwon, K., Cao, C., Jiang, Y. L., and Stivers, J. T. (2003) *J. Biol. Chem.* 278, 19442–19446
22. Drohat, A. C., and Stivers, J. T. (2000) *Biochemistry* 39, 11865–11875
Enzymatic Recognition of 3-MeA

23. Bujnicki, J. M., and Rychlewski, L. (2002) DNA Repair (Amst.) 1, 391–395
24. Schweitzer, B. A., and Kool, E. T. (1994) J. Organic Chem. 59, 7238–7242
25. Kay, L. E. (1993) J. Am. Chem. Soc. 115, 2055–2057
26. Nash, H. M., Bruner, S. D., Scharer, O. D., Kawate, T., Addona, T. A., Spooner, F., Lane, W. S., and Verdin, G. L. (1996) Curr. Biol. 6, 968–980
27. Shan, S. O., and Herschlag, D. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 14474–14479
28. Drohat, A. C., and Stivers, J. T. (2000) J. Am. Chem. Soc. 1840–1841
29. Parikh, S. S., Walcher, G., Jones, G. D., Slupphaug, G., Krokan, H. E., Blackburn, G. M., and Tainer, J. A. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 5083–5088
30. Fujii, T., Saito, T., and Date, T. (1989) Chem. Pharm. Bull. 37, 1208
31. Brown, D. J., Hoerger, E. N., and Mason, S. F. (1955) J. Chem. Soc. Abstracts, 4035–4040
32. Laskowski, R. A., McArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
33. Koradi, R., Billeter, M., and Wuthrich, K. (1996) J. Mol. Graph. 14, 29–32, 51–55