Frameshift Mutagenesis and Microsatellite Instability Induced by Human Alkyladenine DNA Glycosylase

Joanna Klapacz,1,3 Gondichatnahalli M. Lingaraju,1,3 Haiwei H. Guo,2,3 Dharini Shah,1 Ayelet Moar-Shoshani,1 Lawrence A. Loeb,2 and Leona D. Samson1,4

1Departments of Biological Engineering and Biology and Center for Environmental Health Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
2Department of Pathology, The Joseph Gottstein Memorial Cancer Research Laboratory, University of Washington, Seattle, WA 98195, USA
3These authors contributed equally to this work
*Correspondence: lsamson@mit.edu
DOI 10.1016/j.molcel.2010.01.038

SUMMARY

Human alkyladenine DNA glycosylase (hAAG) excises alkylated purines, hypoxanthine, and etheno bases from DNA to form abasic (AP) sites. Surprisingly, elevated expression of hAAG increases spontaneous frameshift mutagenesis. By random mutagenesis of eight active site residues, we isolated hAAG-Y127I/H136L double mutant that induces even higher rates of frameshift mutation than does the wild-type hAAG; the Y127I mutation accounts for the majority of the hAAG-Y127I/H136L-induced mutator phenotype. The hAAG-Y127I/H136L and hAAG-Y127I mutants increased the rate of spontaneous frameshifts by up to 120-fold in S. cerevisiae and also induced high rates of microsatellite instability (MSI) in human cells. hAAG and its mutants bind DNA containing one and two base-pair loops with significant affinity, thus shielding them from mismatch repair; the strength of such binding correlates with their ability to induce the mutator phenotype. This study provides important insights into the mechanism of hAAG-induced genomic instability.

INTRODUCTION

DNA continuously suffers damage as a consequence of normal cellular metabolism and as a result of attack by environmental agents (Friedberg, 2006; Lindahl, 1993). Base excision repair (BER) is the major pathway that defends cells against DNA base damage such as alkylation, oxidation, and deamination (Fan and Wilson, 2005), and BER is initiated by one of a collection of DNA glycosylases, each of which excises specific modified bases leaving abasic (AP) sites in the DNA backbone. AP endonuclease or AP lyase activities nick DNA at the AP site, and the 5’ and 3’ ends are processed such that DNA polymerase can fill the gap and the DNA ends can be ligated to complete the repair process.

A variety of modified DNA bases are recognized and processed by the so-called 3-methyladenine (3MeA) DNA glycosylases, including positively charged N3- and N7-alkylated purines, deaminated adenine (hypoxanthine, Hx) and guanine (xanthine, X), and cyclic etheno adducts of adenine (1,N6-etheno adenine, εA) and guanine (3,N2-ethenoguanine, εG) induced under conditions of lipid peroxidation (O’Brien and Ellenberger, 2004a, 2004b). These enzymes exhibit significant differences in their ability to excise lesions within polynucleotide repeats (Lingaraju et al., 2008; Wyatt and Samson, 2000) and can also excise normal, undamaged purines from DNA, albeit with very low efficiency (Berdal et al., 1998; Wyatt et al., 1999). The crystal structures of 3MeA DNA glycosylases from E. coli (AikA and Tag) and from humans (hAAG, also called MPG or ANPG) reveal that these enzymes position the flipped out damaged base in the electron-rich, aromatic active site (Drohat et al., 2002; Hollis et al., 2000; Lau et al., 1998). When hAAG is complexed with an oligo containing the εA substrate, εA snugly stacks between Tyr127 on one side and His136 plus Tyr159 on the other (Lau et al., 2000) (Figure 1A). This triplet of residues captures and stabilizes the base in the active site through π-stacking interactions and a hydrogen bond donated from the main-chain amide of His136 to the N6 of εA (Lau et al., 2000).

We have proposed that, under certain circumstances, increased, rather than decreased, levels of DNA glycosylase can lead to elevated spontaneous mutation and genomic instability (Glassner et al., 1998a, 1998b; Guo and Loeb, 2003; Hofseth et al., 2003; Posnick and Samson, 1999; Xiao and Samson, 1993). We showed that higher than normal levels of S. cerevisiae 3MeA DNA glycosylase (Mag1) induce elevated rates of both base-pair substitutions and frameshift mutations in yeast (Glassner et al., 1998b; Hofseth et al., 2003; Rusyn et al., 2007). Ulcerative colitis (UC), a chronic inflammatory condition of the large intestine, is frequently associated with microsatellite instability (MSI) and exhibits a 19-fold increased risk of developing colorectal cancer (Gillen et al., 1998). We reported that chronic inflammation associated with UC is accompanied by an increase in hAAG levels and activity in the inflamed tissues (Hofseth et al., 2003) and positively correlated with increased MSI compared to uninfamed epithelia of UC patients. Indeed, increasing hAAG levels in cultured human cells induced an MSI phenotype and increased spontaneous frameshift mutation rates in yeast cells (Hofseth et al., 2003). We inferred that inappropriate expression of yeast and human 3MeA DNA glycosylases can increase the accumulation of BER intermediates and that these BER intermediates promote genomic instability.
In this study, we explore the mechanism by which hAAG induces genomic instability. We isolated hAAG mutants that induce even stronger mutator phenotypes in *E. coli*, *S. cerevisiae*, and human cells. Biochemical characterization of wild-type and hAAG mutant enzymes and determination of their ability to induce the mutator phenotype in a range of *S. cerevisiae* strains provides important insights into the mechanism of hAAG-induced genomic instability. It turns out that wild-type and mutant hAAGs specifically bind to the DNA containing one and two base-pair loops within repetitive sequences, and the strength of such binding correlates with the extent of frameshift mutagenesis and MSI.

**RESULTS**

**Selection of hAAG Mutant Whose Expression Increases Spontaneous Mutation**

In order to identify mutants of hAAG that can significantly elevate spontaneous mutation, wild-type AAG cDNA was cloned into the pTrcHis plasmid and randomly mutated at eight of the nine residues lining the active site pocket with the exception of Glu125 residue required for catalysis (Figure 1). The random library contained one or more amino acid substitutions per molecule. The mutated hAAG cDNAs were expressed in *E. coli* and enriched for mutator variants using sequential rounds of RifR selection. RifR colonies were isolated from the population with the notion that RifR cells expressing mutator glycosylases would be preferentially recovered from this population. Among the different mutants isolated, an AAG cDNA with two amino acid substitutions, Y127I and H136L, was shown to significantly elevate RifR mutation frequency above wild-type levels (Figure S1 available online). The cDNA of the Y127I/H136L double mutant was subsequently cloned into different expression vectors for further analysis.

Wild-Type AAG and AAG-Y127I/H136L Increase Spontaneous Frameshift Mutations in *S. cerevisiae*

We tested the effects of expressing both wild-type AAG and AAG-Y127I/H136L on the rate of spontaneous mutation in *S. cerevisiae* (Figures 2A and 2B). hAAG cDNAs were cloned into the galactose-inducible pYES vector, and proteins were expressed in various *S. cerevisiae* strains to monitor spontaneous mutation. In addition to base-pair substitutions monitored at *trp1-289* of the BGY strains, we also monitored /C01 and +1 frameshift mutations within (A)12 and (A)14 runs in the lys2 locus, respectively (Tran et al., 1997). In our previous work, the expression of hAAG in wild-type *S. cerevisiae* did not significantly increase base-pair substitution mutation rates (Glassner et al., 1998b). Similarly, expression of AAG-Y127I/H136L also failed to elevate the base-pair mutation rates at the *trp1-289* locus (data not shown). However, in agreement with a previous report (Hofseth et al., 2003), we observed a modest increase in −1 and +1 frameshift mutation rates by the expression of wild-type AAG in wild-type *S. cerevisiae*. Upon expression of AAG-Y127I/H136L, we observed more than a 100-fold increase in the rate of −1 frameshifts (Figure 2A) and ~14-fold increase in the rate of +1 frameshifts (Figure 2B) over the background.

hAAG-Induced Spontaneous Mutation Differs Mechanistically from that of Mag1

The elevated expression of the yeast Mag1 DNA glycosylase was previously shown to increase the rates of spontaneous base-pair substitutions at the *trp1-289* locus by about 20-fold in wild-type *S. cerevisiae* and by more than 300-fold in the...
absence of AP endonuclease (Apn1) (Glassner et al., 1998b). Further, the coexpression of Apn1 with Mag1 suppressed the elevated mutation rate, suggesting that it is the unprocessed AP sites generated by Mag1 that increase base-pair substitutions (Glassner et al., 1998b; Rusyn et al., 2007). We tested whether hAAG-induced frameshift mutations might also result from unrepaired AP sites by expressing the wild-type AAG and AAG-Y127I/H136L proteins in the \textit{apn1} \textit{D} background (Figures 2A and 2B). In this case, the absence of Apn1 did not amplify the mutator phenotype induced by hAAG proteins for both \(-1\) frameshifts (Figure 2A) and +1 frameshifts (Figure 2B); this agrees with a previous observation that the co-expression of Apn1 with hAAG does not suppress frameshift mutagenesis (Hofseth et al., 2003). Taken together, these results suggest that Apn1 sites are not the critical intermediates of hAAG-induced frameshift mutagenesis and that hAAG-induced spontaneous mutagenesis differs mechanistically from that of Mag1.

**Y127I Substitution Confers a Strong Frameshift Mutator Phenotype**

In order to understand the relative contributions of each amino acid substitution to the observed mutator phenotype conferred by AAG-Y127I/H136L, the single mutants, AAG-Y127I and AAG-H136L, were each expressed in \textit{S. cerevisiae}, and both \(-1\) and +1 frameshift mutation rates were monitored. Expression of the AAG-H136L only increased spontaneous frameshift mutation to the same level as that of wild-type AAG (Figures 2C and 2D). In contrast, AAG-Y127I induced a much stronger mutator phenotype compared to that of both wild-type AAG and AAG-H136L, increasing the rates of spontaneous \(-1\) frameshifts by \(-80\)-fold and +1 frameshifts by more than 8-fold over the background (Figures 2C and 2D). Thus, the AAG-Y127I amino acid substitution accounts for the majority of the increased frameshift mutator phenotype induced by AAG-Y127I/H136L.

**Expression of Mutant hAAG Glycosylases Increases Microsatellite Instability in Cultured Human K562 Cells**

Having established that the expression of wild-type AAG, AAG-Y127I/H136L, and AAG-Y127I in \textit{S. cerevisiae} gives rise to increased frameshift mutation, we further tested the ability of these AAG proteins to induce genomic instability in cultured human K562 cells (Table 1) after \(-40\) population doublings. Genomic DNA was isolated after this expansion and subjected to MSI genotyping analysis at five canonical, NCI-recommended microsatellite sites (Table 1). In addition, the \textit{BAT40} locus located within the TGF-\(\beta\) type II receptor gene was monitored for microsatellite instability (MSI). We previously demonstrated that increased levels of wild-type AAG in human cells result in low-level MSI (MSI-L) with instability at the \textit{D2S123} locus (one out of the five NCI loci) and additionally at the \textit{BAT40} locus (Hofseth et al., 2003). We found the same MSI-L phenotype in this study (Table 1). In agreement with results from \textit{S. cerevisiae}, the expression of AAG-Y127I/H136L and AAG-Y127I in K562 cells induced an even stronger mutator phenotype, the high microsatellite instability (MSI-H) with changes at the \textit{D5S346} and \textit{D2S123} loci and additionally at \textit{BAT40} (Table 1). Taken together, our results suggest that the expression of hAAG proteins in both \textit{S. cerevisiae} and human K562 cells leads to genomic instability.
The Mutant hAAG Enzymes Have Reduced DNA Glycosylase Activity

As one measure of the ability of hAAG and its mutants to repair alkylated DNA base lesions, we assessed their ability to protect *S. cerevisiae* mag1Δ cells against methylmethanesulfonate (MMS)-induced cytotoxicity, using in vivo complementation (Figure 3A). This assay primarily provides an estimation of glycosylase-mediated repair of the cytotoxic 3MeA DNA lesion. It should be noted that hAAG can also excise 7MeG lesions but that 7MeG is not considered to be cytotoxic. Upon MMS exposure, both wild-type AAG and AAG-H136L provided the same level of protection in the mag1Δ strain, indicating that these enzymes have equivalent capacities for 3MeA repair. In contrast, cells expressing AAG-Y127I showed only partial rescue relative to that of cells expressing wild-type AAG, and AAG-Y127I/H136L completely failed to rescue mag1Δ yeast from the MMS-induced cytotoxicity. Given that AAG-H136L behaves similarly to wild-type AAG for both glycosylase activity and mutator activity, we did not biochemically analyze this protein any further.

The results from the in vivo complementation assays for MMS sensitivity were supported by in vitro DNA glycosylase activity

---

Table 1. Expression of hAAG Glycosylases in K562 Human Cells Induced MSI

<table>
<thead>
<tr>
<th>Microsatellite Marker</th>
<th>BAT25</th>
<th>BAT26</th>
<th>D17S250</th>
<th>D5S346</th>
<th>D2S123</th>
<th>Phenotype</th>
<th>BAT40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empty vector</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>MSS</td>
<td>–</td>
</tr>
<tr>
<td>AAG wild-type</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>MSI-L</td>
<td>+</td>
</tr>
<tr>
<td>AAG-Y127I/H136L</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>MSI-H</td>
<td>+</td>
</tr>
<tr>
<td>AAG-Y127I</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>MSI-H</td>
<td>+</td>
</tr>
</tbody>
</table>

Five NCI-recommended markers were analyzed for MSI (BAT25 through D2S123). If these markers were not found changed from control after the expression period, the phenotype was classified as microsatellite stable (MSS). If one of these five markers was observed with altered nucleotide lengths, the phenotype was classified as low microsatellite instability (MSI-L; wtAAG at the D2S123). Finally, if two or more markers were found changed from controls, the phenotype was classified as high microsatellite instability (MSI-H; AAG-Y127I and AAG-Y127I/H136L at D5S346 and D2S123). In addition, BAT40 was analyzed, as it is analogous to lys2 poly-A runs of *S. cerevisiae*.

---

**Figure 3. In Vivo Characterization of AAG-Y127I and AAG-Y127I/H136L**

(A–C) (A) In vivo complementation assay showing the ability of hAAGs to rescue the mag1Δ yeast against different concentrations of MMS incorporated in YPD plates. Percent growth across MMS gradient obtained with empty vector in mag1Δ is represented as filled diamond (◦), mag1Δ yeast complemented with wtAAG are represented as filled squares (■), AAG-H136L as inverted empty triangles (▲), AAG-Y127I as filled triangles (▲), and AAG-Y127I/H136L as empty squares (□). E125A substitution significantly diminishes, but does not abolish, the ability of AAG-Y127I and AAG-Y127I/H136L to elevate (B) –1 frameshifts and (C) +1 frameshifts. Mutation rates in wild-type yeast generated by AAG-Y127I/E125A and AAG-Y127I/H136L/E125A are statistically significant compared to rates with an empty vector. The n values for most bars in both (B) and (C) are equal to 3, except for one bar in (C). Error bars represent SEM.
assays. Wild-type AAG efficiently removed 3MeA and 7MeG at the rates of 0.27 ± 0.05 and 0.06 ± 0.01 fmol/min, respectively (Table 2). In contrast, the activity of AAG-Y127I/H136L (Table 2) for both 3MeA and 7MeG removal was not detected, accounting for its total failure to protect against MMS-induced cytotoxicity (Figure 3A). In contrast, AAG-Y127I showed intermediate activity for the removal of 3MeA at the rate of 0.19 ± 0.04 fmol/min and showed no detectable activity for 7MeG removal (Table 2). This explains the partial rescue provided by AAG-Y127I mutant against MMS cytotoxicity.

We further studied the ability of hAAG and its mutants to excise Hx and αA and their ability to bind AP site-containing DNA (Table 2). The excision rate for the removal of Hx was robust for the wild-type AAG, with $k_{\text{obs}} = 0.42 ± 0.05 \text{ min}^{-1}$; in comparison, AAG-Y127I and AAG-Y127I/H136L completely failed to excise this lesion (Table 2). Whereas wild-type AAG removed αA at the rate of $k_{\text{obs}} = 0.11 ± 0.01 \text{ min}^{-1}$, AAG-Y127I showed significantly reduced activity for αA, with $k_{\text{obs}} = 0.032 ± 0.007 \text{ min}^{-1}$, and the activity of AAG-Y127I/H136L for αA lesions was undetectable (Table 2). Surprisingly, in spite of their difference in αA excision activity, both wild-type AAG and AAG-Y127I bound the DNA duplex containing an αA lesion with roughly equal affinity, with $K_d$ of 23.9 and 31.5 nM, respectively. Even though the excision of αA by AAG-Y127I/H136L was not detectable, it bound the αA containing DNA with a significant affinity, $K_d = 114.8 ± 32.3 \text{ nM}$ (Table 2). Similar to αA, both wild-type AAG and AAG-Y127I bound the AP site containing DNA with roughly equal affinity. However, compared to wild-type AAG, AAG-Y127I/H136L showed ~18-fold reduced affinity for binding the AP site containing DNA.

### The Glu125 Catalytic Residue Is Required for hAAG’s Frameshift Mutator Phenotype

The fact that the ability of the hAAG mutants to induce frameshift mutagenesis increased with decreasing DNA glycosylase activity led us to question whether base excision is actually required for the observed mutator phenotype. We therefore mutated the catalytic residue Glu125 (to Ala) in the wild-type and mutant hAAG enzymes. Previous studies showed that E125Q and E125A mutants of hAAG show undetectable glycosylase activity and that both bind efficiently to lesion-containing DNA duplex (Abner et al., 2001; O’Brien and Ellenberger, 2003). We also observed that, like AAG-Y127I (Table 2), AAG-Y127I/E125A efficiently binds αA, but, unlike AAG-Y127I, it fails to remove this lesion from the DNA duplex (data not shown). In order to understand the role of Glu125 in the hAAG-induced mutator phenotype, AAG-E125A, AAG-Y127I/E125A, and AAG-Y127I/H136L/E125A mutants were expressed in S. cerevisiae, and their ability to increase frameshift rates was monitored. The results in Figures 3B and 3C show that the introduction of the E125A substitution abolished the ability of wild-type AAG to increase both −1 and +1 frameshift mutations. Moreover, the mutator ability of both AAG-Y127I/H136L and AAG-Y127I mutants was also reduced upon replacement of the catalytic Glu125, albeit not completely abolished, indicating that catalytic activity is not the only factor contributing to the observed mutator phenotype. The residual increase in mutation rates for AAG-Y127I/H136L/E125A compared to the empty vector was ~17-fold for −1 frameshifts (Figure 3B) and ~3-fold for +1 frameshifts (Figure 3C). AAG-Y127I/E125A behaved in a similar manner to the triple mutant, being largely suppressed for mutator activity but still increasing the rate of −1 frameshifts by more than 12-fold and +1 frameshift rates by ~2-fold, compared to the empty vector. These results clearly show that the catalytic Glu125 residue is essential for the mutator phenotype induced by wild-type AAG and that it is required for most, but not all, of the mutator phenotypes induced by the hAAG mutants.

### The Role of DNA Polymerases in the hAAG-Induced Mutator Phenotype

We tested whether different translesion synthesis (TLS) and replicative DNA polymerases modulate hAAG-induced frameshifts. In S. cerevisiae, there are two major TLS polymerases: Pol ε′, composed of Rev3 and Rev7, and Polη/Rad30. Rev3 is the catalytic subunit of Pol ε′, and Rev7 is its accessory factor required for bypass activity; Pol ε′ together with Rev1 forms a complex required for TLS (Acharya et al., 2006; Prakash et al., 2005). hAAG proteins were expressed in rev3Δ, rev7Δ, and rad30Δ yeast strains, and frameshift rates were monitored (Figures S2A and S2B). We previously showed that Pol ε′ and Rev1 are required for Mag1-induced base-pair substitutions by facilitating replication and elongation past Mag1-induced AP

---

### Table 2. The Excision Rate Constants and Binding Affinity of hAAG with Standard Deviation Tested for Different DNA Lesions Using Purified hAAG

<table>
<thead>
<tr>
<th>Substrates</th>
<th>$k_{\text{obs}}$ ± SD (min$^{-1}$)</th>
<th>$K_d$ ± SD (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3MeA$^a$</td>
<td>AAG Wild-Type AAG-Y127I AAG-Y127I/H136L</td>
<td>AAG Wild-Type AAG-Y127I AAG-Y127I/H136L</td>
</tr>
<tr>
<td>0.27 ± 0.05</td>
<td>0.19 ± 0.04 ND**</td>
<td>NA</td>
</tr>
<tr>
<td>7MeG$^a$</td>
<td>0.06 ± 0.01 ND**</td>
<td>23.9 ± 2.4</td>
</tr>
<tr>
<td>Hx</td>
<td>0.42 ± 0.05 ND**</td>
<td>31.5 ± 3.9*</td>
</tr>
<tr>
<td>αA</td>
<td>0.11 ± 0.01 0.032 ± 0.007** ND**</td>
<td>114.8 ± 32.3**</td>
</tr>
<tr>
<td>AP site</td>
<td>NA NA NA NA</td>
<td>376.9 ± 98.3**</td>
</tr>
<tr>
<td>Poly A loop A</td>
<td>251.3 ± 64.2 98.5 ± 18*</td>
<td>43.95 ± 8.2**</td>
</tr>
</tbody>
</table>

Refer to Figure S4 for raw data used to calculate $K_d$ constants presented in this table. ND, no detectable activity (corresponds to $k_{\text{obs}}$ ≤ 0.00002 min$^{-1}$ and < 0.0006 fmol/min); *, **, significantly different compared to WT at $p < 0.05$ or $p < 0.01$, respectively; NA, the experiment was not performed.

$^a$ Measurement units: fmol/min.
sites (Glassner et al., 1998b). Here, the inactivation of Pol\textsubscript{I} had no significant effect on the −1 and +1 frameshift rates generated by hAAG proteins. These results once again indicate that there is a clear mechanistic difference between the hAAG- and Mag1-induced spontaneous mutator phenotypes in S. cerevisiae. Of interest, in the absence of Rad30, there was a nonsignificant trend toward increasing the rate of −1 frameshifts (Figure S2A) and no significant change in +1 frameshifts. Together, the results showed that Pol\textsubscript{I} and Pol\textsubscript{I} do not significantly contribute to the observed hAAG-induced mutator phenotype.

Because inactivation of the major TLS polymerases did not significantly affect hAAG-induced frameshifts, we investigated the role of other possible polymerases. Three candidates in S. cerevisiae are Rev1, Pol\textsubscript{I}’s Pol32 subunit, and Pol4 (Gibbs et al., 2005; Kalinowski et al., 1995; Shimizu et al., 1993). Rev1 is a deoxyctydyl transferase that forms a complex with Pol\textsubscript{I} and interacts with ubiquitinated PCNA present at stalled replication forks (Garg and Burgers, 2005). Pol32 is the smallest subunit of Pol\textsubscript{I} that is thought to play a role in Pol\textsubscript{I}’s processivity and DNA repair activity, linking DNA replication with error-prone, Rad6-dependent translesion bypass (Huang et al., 2000). Finally, Pol4 has been shown to have an intrinsic 5’-deoxyribose phosphate (5’-dRP) lyase activity and is a homolog of mammalian Pol\textsubscript{I} that has a role in base excision repair (Bebenek et al., 2005; Shimizu et al., 1993). Upon expression of AAG-Y127I/H136L and AAG-Y127I/H136L, Pol\textsubscript{I}, Pol\textsubscript{I}, and Pol\textsubscript{I} each suppress hAAG-induced +1 frameshift rates, compared to wild-type (Figure S2C). In stark contrast, expression of AAG-Y127I/H136L and AAG-Y127I/H136L in the absence of Rev1, Pol32, and Pol4 displayed reduced +1 frameshift rates, compared to wild-type yeast (Figure S2D). Overall, our results suggest that Rev1, Pol4, and Pol32 each suppress hAAG-induced −1 frameshifts but contribute to +1 frameshift mutations in vivo.

**Role of Postreplicative DNA Mismatch Repair in hAAG-Induced Frameshifts**

The accumulation of frameshifts in vivo is known to be influenced by the postreplication DNA mismatch repair (MMR) pathway (Greene and Jinks-Robertson, 2001; Tran et al., 1997). MMR is initiated either by the MutS\textsubscript{a} complex (the MSH2-MSH6 heterodimer) binding single base-pair (bp) mismatches or 1 bp loops or by the MutS\textsubscript{i} complex (the MSH2-MSH3 heterodimer) binding bp loops ≥ 1 (Kunkel and Erie, 2005). It seemed plausible that hAAG expression in yeast might somehow interfere with MMR activity and that such interference was responsible for hAAG’s frameshift mutator phenotype. We reasoned that, if hAAG-induced frameshifts relied solely on inhibiting or depleting MMR activity, then the expression of hAAG proteins in MMR-deficient strains should not be able to further affect spontaneous frameshift mutagenesis. The hAAG proteins were thus expressed in msh\textsubscript{6}Δ (MutS\textsubscript{a}-deficient), msh\textsubscript{3}Δ (MutS\textsubscript{i}-deficient), and msh\textsubscript{2}Δ (MutS\textsubscript{a}- and MutS\textsubscript{i}-deficient) yeast strains, and frameshift rates were monitored (Figure 4). In agreement with previous studies (Clark et al., 1999; Tran et al., 1997), the msh\textsubscript{3}Δ deletion had a modest effect on spontaneous frameshifts, msh\textsubscript{6}Δ had a greater effect, and msh\textsubscript{2}Δ had the greatest effect, with ~1700-fold increase in the rate of +1 frameshifts over wild-type. When expressed in the msh\textsubscript{3}Δ and msh\textsubscript{6}Δ yeast strains, wild-type AAG induced a moderate increase in both −1 and +1 frameshift rates, whereas AAG-Y127I and AAG-Y127I/H136L induced even higher increases (Figure 4). Thus, in the absence of either MutS\textsubscript{a} or MutS\textsubscript{i}, the hAAG proteins are still capable of inducing robust −1
and +1 frameshift mutator phenotypes with the same relative strengths. However, because MutS\(\alpha\) and MutS\(\beta\) overlap in their ability to bind 1 bp loop substrates, we went on to monitor hAAG-induced +1 frameshift mutation in the msh2\(\Delta\) (MutS\(\alpha\)- and MutS\(\beta\)-deficient) yeast strain. (Note that, because the msh2\(\Delta\) strain displayed a 50,000-fold increase in \(C_0\)1 frameshift mutagenesis compared to wild-type [Tran et al., 1997], it was not included in this part of the study.) Despite the high background of +1 frameshift mutation rate in the msh2\(\Delta\) strain of \(\approx 15,000 \times 10^{-8}/\)generation, expression of the hAAG proteins further enhanced this rate by an extra 7,000 to 12,000 \(\times 10^{-8}/\)generation (Figure 4D). These results demonstrate that simple inhibition or depletion of the major MMR pathways is not sufficient to explain the hAAG-induced mutator phenotype and, more importantly, that there exists an MMR-independent mechanism that contributes to hAAG-induced frameshift mutagenesis and microsatellite instability.

**Interaction of hAAG and Its Mutants with Different One and Two Base-Pair Loops**

We went on to test whether hAAG and its mutants could recognize bp loops that are thought to arise from polymerase slippage in microsatellite sequences. We chose substrate sequences analogous to those that showed the MSI-H phenotype in human K562 cells: poly-CA repeats in D5S346 and D2S123, poly-TA repeats in D2S123, and poly-A/T repeats present in BAT-40 as well as in the S. cerevisiae lys2 loci. We tested the interaction of hAAG proteins with oligos containing CA, TA, A, GT, or T bp loops located within their respective repeats (Figures 5 and S3). Binding assays showed that all three hAAG proteins specifically bind CA, TA, A, and GT bp loop-containing oligos (Figure 5B), displaying no binding to the same oligos without loops (Figure 5A). The hAAG proteins also weakly bound to T bp loops (data not shown). The quantification of the binding affinity to the critical poly-A loop A oligo showed that AAG-Y127I/H136L binds this loop with a high affinity of 43.95 ± 8.2 nM, whereas AAG-Y127I binds the loop with an affinity of 98.5 ± 18 nM that is 2-fold reduced (Table 2). In contrast, wild-type AAG bound poly-A loop A oligo with significantly reduced affinity of 251.3 ± 64.2 nM that is 6-fold reduced compared to AAG-Y127I/H136L (Table 2). The binding affinity of hAAG proteins is thus...
positively correlated with their ability to induce frameshift mutations in vivo; in other words, the higher the binding affinity to the poly-A loop A oligo in vitro, the higher the induced frameshift mutator phenotype. We tested whether the hAAG proteins can remove normal bases from the one and two bp loops to which they bind; all three hAAG proteins failed to remove normal bases from loop-containing oligos (data not shown); note that we can not eliminate the possibility that base excision may occur in vivo. Importantly, we found that AAG-E125A and AAG-Y127I/E125A failed to bind the poly-A loop A-containing oligo (data not shown). It is important to point out here that the E125A mutant does not lose the ability to flip εA into its active site but does lose the ability to excise this base; however, E125A (and Y127I/E125A) simply does not bind the DNA-containing poly-A loop A. This explains, at least in part, the observation that E125A mutants of AAG and AAG-Y127I were compromised in their ability to elevate frameshift mutation (Figures 3B and 3C). Taken together, these results suggest that hAAG proteins specifically bind the one or two bp loops that are expected to form at DNA microsatellites as a result of polymerase replication slippage and that such binding is responsible for the frameshift mutator and MSI phenotype conferred by these proteins (Figure 5C).

**DISCUSSION**

DNA microsatellites are comprised of repetitive nucleotide sequences that constitute about 30% of the human genome. Microsatellites have been reported to be hot spots for recombination, DNA polymerase slippage, and random DNA integration. Changes in repetitive sequences can result in deleterious effects on gene function leading to adverse phenotypic effects (Atkin, 2001; Boland et al., 2008). Microsatellite instability results from small deletions or expansions within polynucleotide repeats in genomic DNA, and frameshift mutations (−1 or +1) represent the most common mutations associated with MSI. Frameshifts are thought to result from DNA polymerase slippage and strand misalignment during replication (Debrauwere et al., 1997; Kunkel and Bebenek, 2000). An increase or decrease in the repeat length could occur if the slippage is on the daughter or on the template strand, respectively (Boland et al., 2008). Such replication slippage intermediates would have to either be misrepaired or persist until the next round of DNA replication in order for frameshift mutations to be fixed. Proteins that specifically bind and stabilize replication slippage intermediates, thus interfering with their repair, could clearly enhance both frameshift mutagenesis and MSI (Figure 5C).

In humans, MSI has been closely associated with many diseases, including cancer and other disorders that exhibit a dominant mode of inheritance, such as spinal and bulbar muscular dystrophy, Huntington’s disease, spinocerebellar ataxia type I, insulin-dependent diabetes mellitus, Friedrich’s ataxia, Myotonic dystrophy, Fragile X syndrome, Jacobsen 11q-deletion syndrome, and so on (Debrauwere et al., 1997). It has been shown that the majority of tumors in hereditary nonpolyposis colorectal cancer (HNPCC), sporadic colorectal, and gastric and endometrial cancers are associated with microsatellite length alterations, termed high-frequency microsatellite instability (MSI-H) (Imai and Yamamoto, 2008). HNPCC results primarily from germline mutations in one of the DNA MMR genes, primarily in MLH1 or MSH2 in HNPCC and ovarian tumors and less so in MSH6 and PMS2 (Imai and Yamamoto, 2008; Parsons et al., 1995a; Plotz et al., 2006). The key function of MMR is to eliminate base mismatches and insertion/deletion loops arising during DNA replication from polymerase slippage (Kunkel and Erie, 2005). The genetic and epigenetic inactivation of MMR genes predisposes to cancer development. Correspondingly, in yeast, frameshift errors at repetitive sequences increase by 10- to 100-fold in MMR-deficient cells relative to controls (Harfe and Jinks-Robertson, 2000; Tran et al., 1997). In this study, we find evidence for an alternative pathway for the induction of frameshift mutations and microsatellite instability that is independent of MMR.

We previously found that elevated expression of hAAG increases spontaneous frameshift mutagenesis in yeast and confers a low-frequency MSI (MSI-L) phenotype in human cells. Here, we report the isolation and characterization of an hAAG mutant AAG-Y127I/H136L with an enhanced ability for spontaneous mutation in E. coli, yeast, and human cells. Tyr127 and His136 are highly conserved and important residues in hAAG’s active site. Previous studies have assessed the tolerance of hAAG’s active site residues for random substitutions and the essentiality of these residues for survival after DNA methylation damage (Chen et al., 2008; Guo et al., 2004). When selecting for MMS resistance in E. coli, Tyr127 and His136 show very low substitutability, underscoring their importance in hAAG-mediated excision of cytotoxic 3MeA DNA lesions. Another study tested the importance of active site residues using genetic complementation assay for the rescue of MMS sensitivity in mag1Δ yeast cells (Lau et al., 2000); whereas AAG-Y127F provided a partial resistance to MMS, AAG-H136Q provided complete resistance to MMS. Similarly, we find that AAG-Y127I provided partial resistance to MMS and that AAG-H136L provided almost complete resistance to MMS in mag1Δ yeast (Figure 3A). Of interest, AAG-Y127I/H136L completely failed to rescue mag1Δ cells from MMS-induced sensitivity (Figure 3A), presumably due to its inability to remove MMS-generated 3MeA lesions (Table 2).

The effects of Y127I and H136L substitutions on hAAG’s damage base removal can be understood from structural studies of hAAG-DNA complexes (Lau et al., 2000). In the crystal structure of hAAG-DNA complexes, εA base is stacked between Tyr127 and His136 residues (Figure 1). Tyr127 also donates a hydrogen bond to the catalytic Glu125. His136 forms hydrogen bonds from its main chain amide to the N6 of εA, through its side chain to the 5’ phosphate of εA, and with the Tyr157 side chain (Figure 1). When the Y127I and H136L mutations are modeled into the hAAG active site, Ile127 significantly lacks the ability to stack εA and fails to hydrogen bond to Glu125. Leu136 maintains the possibility of hydrogen bonding to N6 of εA through its mainchain amide but fails to form hydrogen bonds to 5’ phosphate of εA and Tyr157 side chain (Figure 1).

Previous biochemical studies showed that hAAG recognizes purine and pyrimidine mismatches that represent one set of MMR substrates and can excise the undamaged purines, but not pyrimidines, from these mismatches (Biswas et al., 2002;
O’Brien and Ellenberger, 2003; O’Brien and Ellenberger, 2004a). Our study now reports that hAAG recognizes another set of MMR substrates, namely one and two bp loops that are generated at microsatellites by DNA polymerase slippage and are prime substrates for MMR. We now know that hAAG and its variants can bind these bp loops and greatly increase the chance that the base-pair loops will be fixed into frameshift mutations. It seems likely that bp loop repair could be inhibited simply by shielding them from MMR. However, hAAG and its mutants are able to induce frameshift mutagenesis even in the absence of MMR, albeit to a limited extent, indicating that whereas shielding bp loops from MMR is highly likely to be involved, it can not solely account for hAAG’s mutator activity. We propose that hAAG binding stabilizes bp loops such that they persist until the next round of DNA replication and that such stabilized loops induce frameshift mutations and microsatellite instability. Indeed, the increase in the rate of +1 and −1 frameshift mutations by the hAAG proteins corresponds well with their binding affinity for the 1 bp loop in poly-A-containing DNA (Figures 4 and 5, Table 2). Of interest, hAAG-induced frameshifts are influenced by Rev1, Pol4, and PolS2; in the absence of these polymerases, mutant hAAG-expressing cells exhibit increased −1 frameshifts and decreased +1 frameshifts. We infer that these polymerases are involved in either facilitating or suppressing strand elongation from the looped out primer-template structure (see Figure 5C); when present, the polymerases presumably suppress elongation from daughter strands paired with a slipped parental strand (thus decreasing −1 frameshifts) and facilitate elongation from a daughter strand that has slipped relative to its template (thus increasing +1 frameshifts). Despite their binding to one and two bp loops, hAAG and its mutants were unable to remove normal bases from these loops to produce abasic sites; this perhaps explains the fact that modulating AP-endonuclease activity and eliminating translesion DNA polymerase activities did not influence hAAG-induced frameshifts (Figure 2).

Taken together, our results provide evidence for an additional mechanism for hAAG-induced genomic instability. Elevated expression of hAAG and its variants induces frameshift mutagenesis and MSI by binding to and stabilizing bp loops and shielding them from repair both in the presence and the absence of the MMR pathway. Such hAAG-induced frameshift mutagenesis and MSI is highly likely to contribute to an increased risk of cancer, such as that observed in ulcerative colitis patients.

**EXPERIMENTAL PROCEDURES**

**Plasmids**

Table S1 shows all plasmids used in this study.

**Strains**

Table S2 shows all E. coli and S. cerevisiae strains used in this study. Most yeast strains were derived from E133 and E134 (Tran et al., 1997) and created by introduction of appropriate kanMX4 cassette according to standard protocols (Supplemental Experimental Procedures).

**Selection of hAAG-1Y127I/H136L Mutant in E. coli**

A large library of mutant hAAG cDNAs with substitutions targeted to eight active site residues (Tyr127, Ala134, Ala135, His136, Tyr159, Cys167, Asn169, and Leu180) was expressed in E. coli (Figure S1). Mutant cDNAs were isolated after several rounds of rifampicin resistance (Rif') selection as described in the Supplemental Experimental Procedures. This forward mutation assay mostly scores for base-pair substitutions but also for frameshift mutations (Takechi et al., 2006). A double hAAG mutant with the substitutions Y127I and H136L was isolated.

**Determination of Mutation Rates**

Spontaneous frameshift mutation rates were determined as previously described (Rusyn et al., 2007), using Drake’s method of the median (Rosche and Foster, 2000). Trp' revertants that score for base-pair substitutions were assayed in the BGY11 strain.

**Complementation Assays with hAAG Glycosylases**

The mag1Δ strain was used to test the ability of the mutant hAAG cDNAs to provide alkylation resistance on MMS gradient plates. Overnight cultures were divided into two separate cultures and grown for additional 3 hr at 30°C either uninduced, in 2% glucose, or induced with 2% galactose. Following this initial expression period, cells were mixed with top agar (1%) and stamped onto gradient MMS plates containing either glucose or galactose (Chen et al., 1994). After 3 day incubation at 30°C, growth across the MMS gradient plate was measured for different hAAG cDNAs and compared to growth across the gradient for the mag1Δ strain carrying the empty vector.

**DNA Glycosylase Activity Assays**

DNA glycosylase assays were performed essentially as described (Kartalou et al., 2000). Excision of 3MeA and 7MeG from tritiated calf thymus DNA was previously described (Chen et al., 2008). For detailed protocols, refer to the Supplemental Information. The rate constants were calculated by fitting the data to the one-phase exponential function (equation 1) using GraphPad Prism analysis software (GraphPad Software, Inc., La Jolla, CA).

\[
Y = Y_{\text{max}} \left(1 - e^{-k_{\text{obs}} t}\right)
\]

in which \(Y\) is the amount of substrate cleaved at any given time point, \(Y_{\text{max}}\) is the maximum amount of substrate cleaved, \(t\) is the time, and \(k_{\text{obs}}\) is the observed rate constant.

**Gel Mobility Shift Assays**

Protein-DNA-binding assays were carried out as previously described with slight modifications (Kartalou et al., 2000). For detailed protocol, refer to the Supplemental Information. The binding constants for AAG binding to A+ and AP site DNA were calculated by fitting the data to one-site binding (equation 2), using GraphPad Prism analysis software (GraphPad Software, Inc., La Jolla, CA).

\[
Y = K_{\text{d}} \frac{X}{(K_{\text{d}} + X)}
\]

in which \(Y\) is the total bound substrate, \(K_{\text{d}}\) is the maximum specific binding, \(X\) is the concentration of the protein, and \(K_{\text{d}}\) is the apparent binding constant. The \(K_{\text{d}}\) values for AAG binding to poly-A loop-A substrates were calculated by fitting the data to the following equation (3) for quadratic binding isotherm.

\[
F = \frac{1}{2} \left(\frac{L_{0} + 1}{L_{0} - f_{1}}\right) \left[(P_{o} + L_{0} + K_{d}) - \sqrt{(P_{o} + L_{0} + K_{d})^{2} - 4P_{o} L_{0}}\right]
\]

in the above equation, \(f_{1}\) is the signal in the absence of protein, and \(f_{2}\) is the signal at the maximum protein concentration. \(P_{o}\) is the total protein concentration, \(L_{0}\) is the total ligand concentration, and \(K_{d}\) is the dissociation constant (\(K_{d}\)).

**Selection of hAAG-Expressing Clones in K562 Cells**

pCAGGS vectors with or without hAAG cDNAs and selective pMACS-Kk.II (Miltenyi Biotech, Auburn, CA) plasmids were linearized with Sc I, ethanol precipitated, and cotransfected into K562 cells with lipofectamine 2000 (Invitrogen) according to manufacturer’s instructions. Positive, H-2Kk surface marker-expressing cells were isolated according to manufacturer’s instructions and cloned at 0.5 cells per well in 96-well plates. pCAGGS-hAAG clones were screened by PCR for hAAG and assayed for hAAG-expression by western blots and qRT-PCR. For westerns, polyclonal rabbit antibody raised against full-length AAG was used (Covance Research Company, Denver, CO).
PA). mRNA was isolated from ~1 × 10^6 cells with QiAshredder Kit, and cDNA was obtained with Omniscript Reverse Transcriptase Kit (QIAGEN). qRT-PCR was performed with SYBR green method (Applied Biosystems, Foster City, CA). Primers for hAAG PCR and RT-PCR are described in Table S3.

Analysis of MSI
K562 cell lines expressing hAAG glycosylases were cultured as previously described for ~40 population doublings (Hofseth et al., 2003). Genomic DNA was isolated using High Pure PCR Template Preparation Kit (Roche, Burlington, NC) according to manufacturer’s instructions. Forward PCR primers (Invitrogen) were fluorescently labeled with FAM dye, and Platinum GenoTYPE Taq DNA Polymerase (Invitrogen) was used. PCR reactions were performed according to manufacturer’s instructions with 20 ng template DNA per reaction. Reaction products were run on ABI Prism 3738xl analyzer at SeqWright facility (Houston, TX), and resulting data from the genotyping chromatograms were aligned and scored for significant changes in microsatellite band length distribution. A set of five markers for MSI was chosen and analyzed with accordance to NCI recommendations (Boland et al., 1998). Additionally, poly(A) BAT40 locus within the TGF-β type II receptor gene was included (Parsons et al., 1995d).

SUPPLEMENTAL INFORMATION
Supplemental Information includes Supplemental Experimental Procedures, four figures, and three tables and can be found with this article online at doi:10.1016/j.molcel.2010.01.038.

ACKNOWLEDGMENTS
This work was supported by the following grants: CA055042, CA115802, and ES02109. L.D.S. is an American Cancer Society Research Professor.

Received: August 3, 2009
Revised: December 28, 2009
Accepted: January 21, 2010
Published: March 25, 2010

REFERENCES


