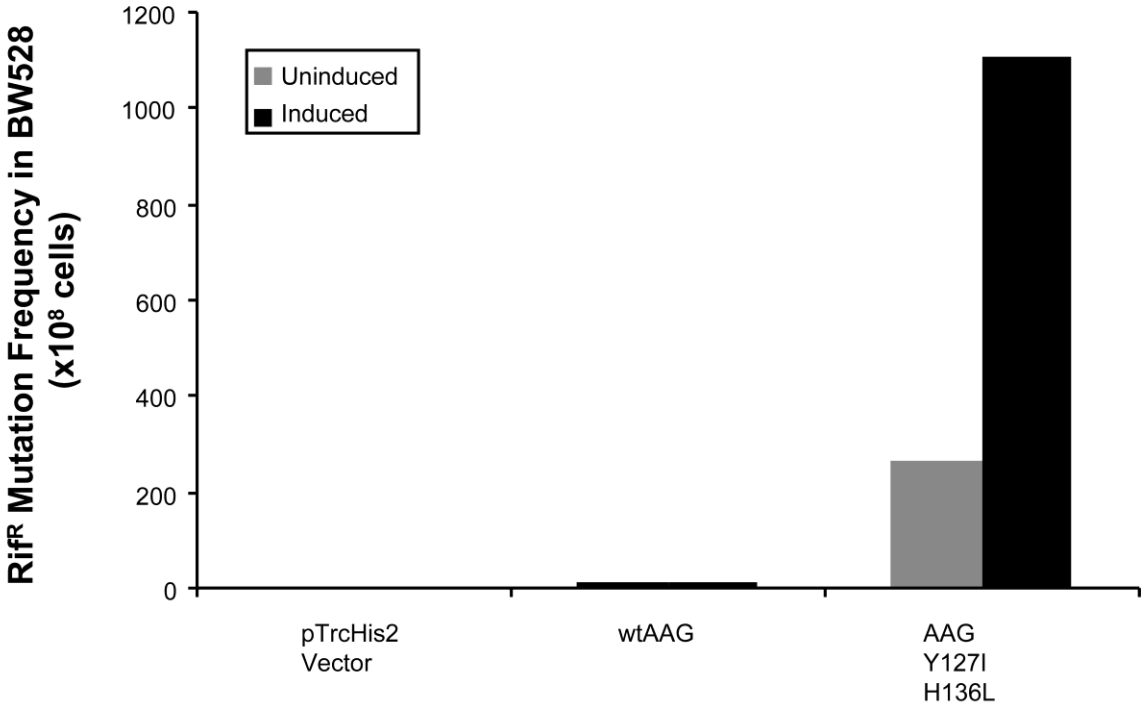
Klapacz et al., Inventory of Supplementary Information

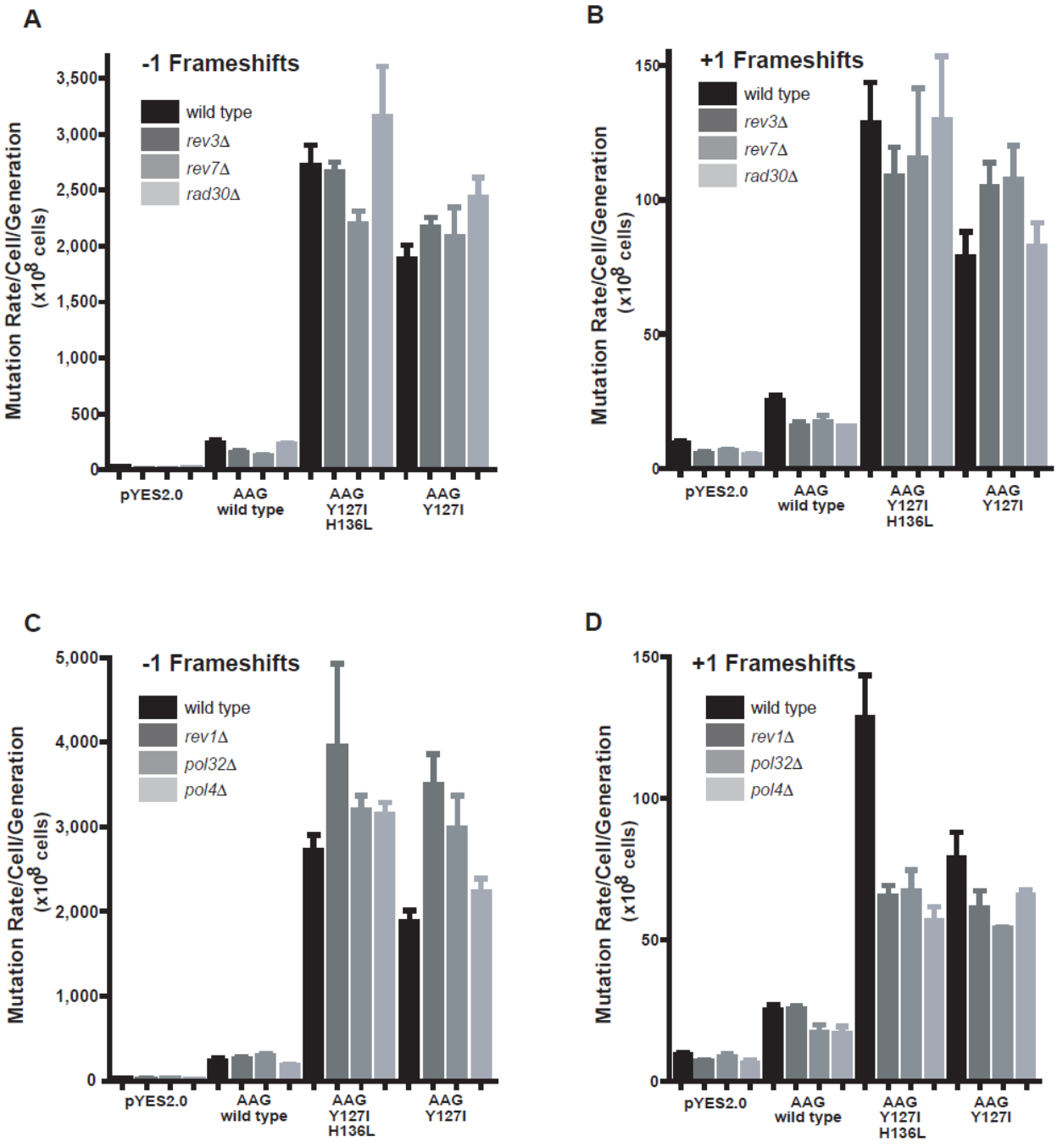
1. Supplementary Figure 1 presents Rif^R frequency in *E. coli* cells induced by the wild type hAAG and the isolated double mutant.
2. Supplementary Figure 2 investigates the effect of yeast DNA polymerases on hAAG-induced mutagenesis to supplement the mechanistic investigations found in Figure 2.
3. Supplementary Figure 3 contains the DNA structures of repetitive sequence oligos used in the study Figure 5.
4. Supplementary Figure 4 contains plots showing the calculation of hAAG binding affinity (K_d) to poly-A-loop-A oligos found in Table 2.
5. Supplementary Figures 1-4 Legends
6. Supplementary Table 1 contains expression vectors used in the study and their purpose of use.
7. Supplementary Table 2 lists microbial strains used in the study.
8. Supplementary Table 3 contains sequences of oligonucleotides, other than those in Supplementary Figure 3, used to clone, mutate and determine the expression of hAAGs cDNAs.
9. Supplemental Experimental Protocols describe cloning and mutagenesis of hAAG cDNAs, construction of yeast strains prepared for this study, purification of hAAG proteins and detailed biochemical protocols for glycosylase and EMSA assays.

Klapacz et al., Supplementary Data

Klapacz et al., Supplementary Figure 1

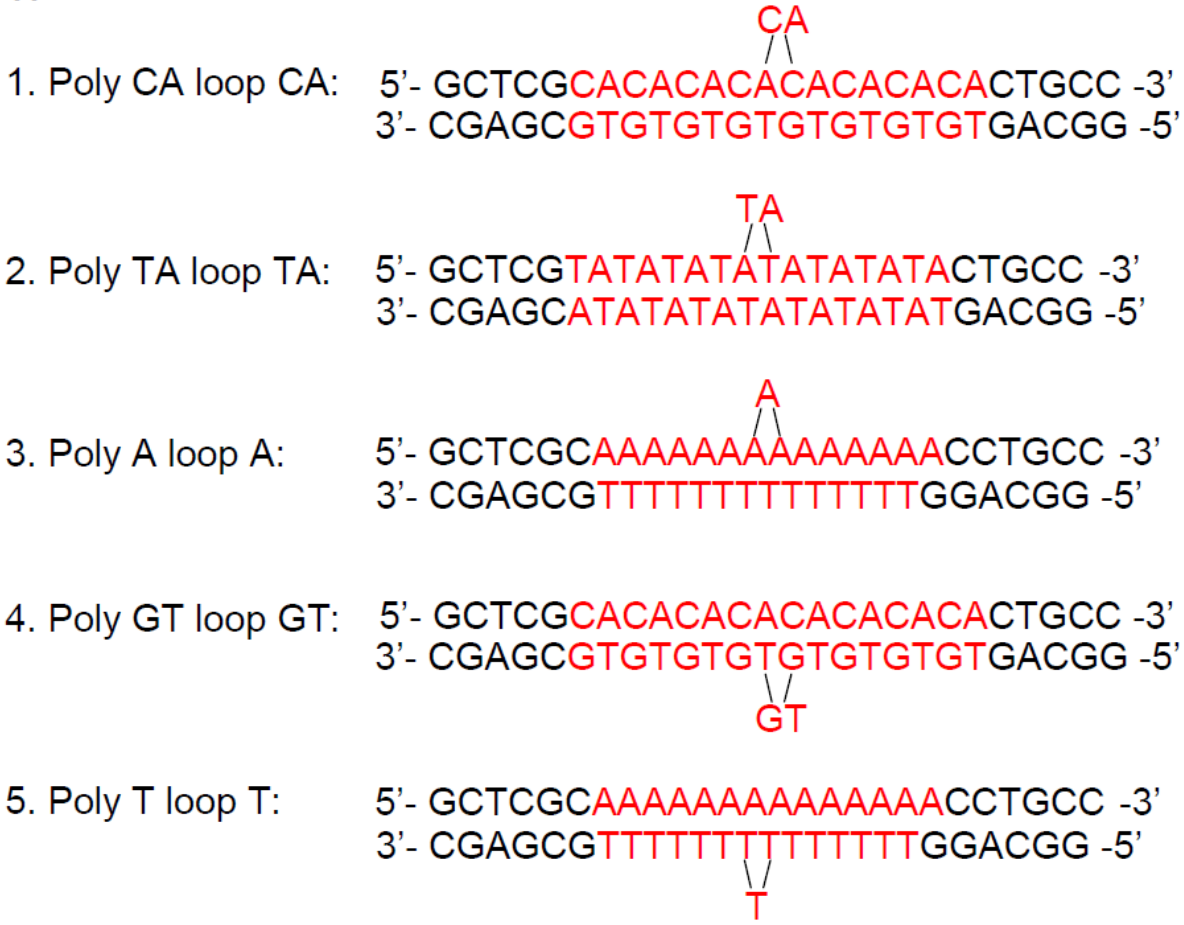


Klapacz et al., Supplementary Figure 2



Klapacz et al., Supplementary Figure 3

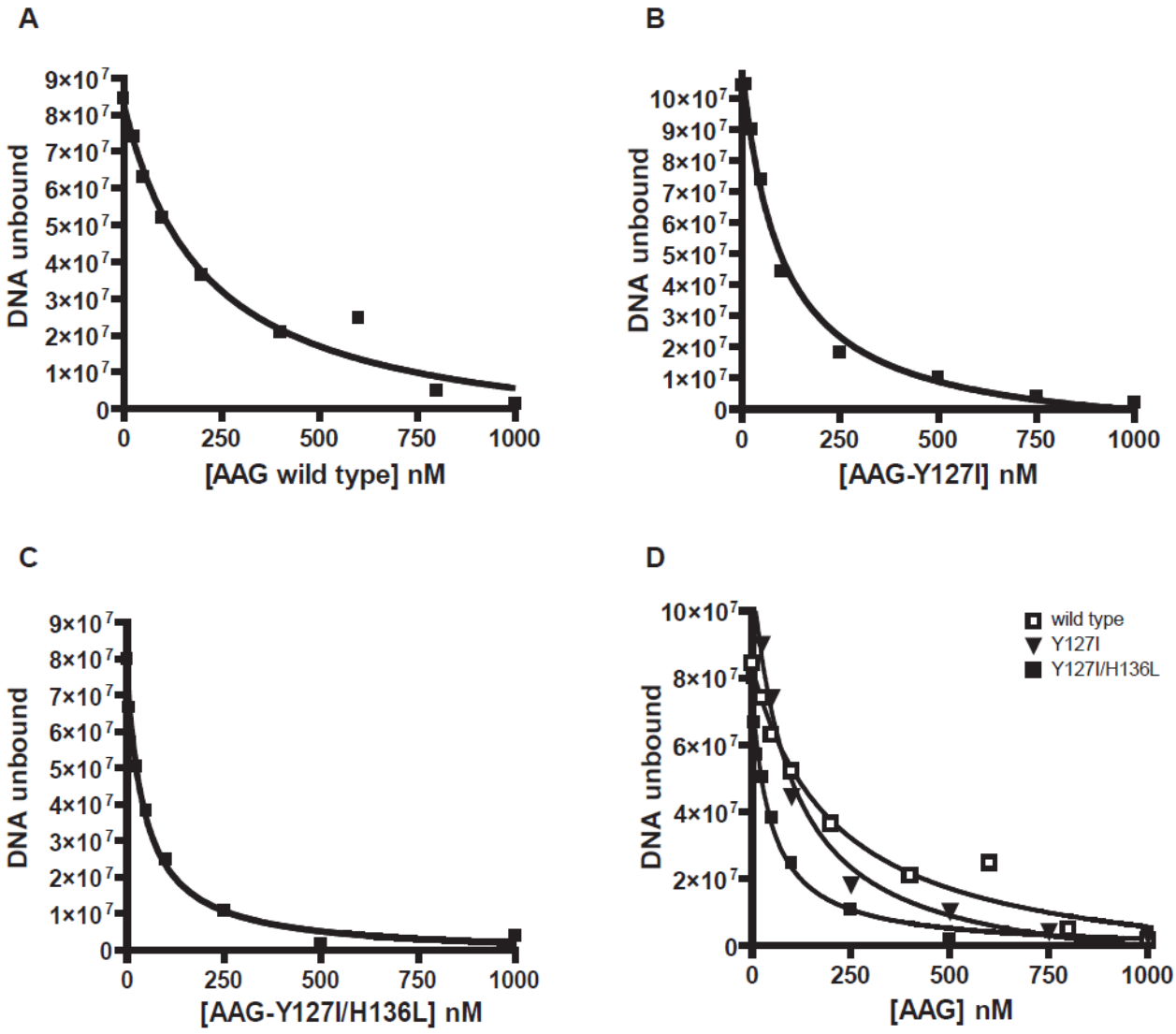
A



B

1. Poly CA:
   5'- GCTCGCACACACACACACACACACTGCC -3'
   3'- CGAGCGTGTGTGTGTGTGTGTGTGACGG -5'
2. Poly TA:
   5'- GCTCGTATATATATATATATATATACTGCC -3'
   3'- CGAGCATATATATATATATATATATGACGG -5'
3. Poly A:
   5'- GCTCGCAAAAAAAAAAAAAAACCTGCC -3'
   3'- CGAGCGTTTTTTTTTTTTTTTGGACGG -5'

Klapacz et al., Supplementary Figure 4



Supplementary Figure Legends

Supplementary Figure 1:

Rifampicin resistance of *E. coli* BW528 expressing hAAG proteins.

hAAG cDNAs were expressed from the pTrcHis vector in *E. coli* cells by adding IPTG (isopropyl β-D-thiogalactoside) to the final concentration of 1mM. After each selection round, Rif^R colonies were picked, plasmids were isolated and the individual clones were re-transformed and re-assessed for their ability to confer an increased Rif^R frequency. A double hAAG mutant with the substitutions Y127I and H136L was thus isolated. pTricHis2 plasmids with wild type and AAG-Y127I/H136L cDNAs were introduced into BW528 cells and expressed overnight. IPTG was added to the final concentration of 1mM where indicated. RifR cells were selected on LB plates with 150µg/ml Rif. Mutation frequency was calculated as total number of mutants counted per 10^8 cells plated. This cDNA was then re-cloned into eukaryotic expression vectors alongside its respective single mutants and wild type *AAG* cDNAs.

Supplementary Figure 2:

Effects of inactivation of DNA polymerases on hAAG-induced frameshift rates.

The mutation rates observed in wild-type yeast are represented in black and those observed in yeast deletion backgrounds are represented in different shades of gray. Mutation rates in wild-type, *rev3Δ*, *rev7Δ* and *rad30Δ* backgrounds for hAAG-induced (A) -1 frameshifts and (B) +1 frameshifts. Similarly, mutation rates in wild-type, *rev1Δ*, *pol32Δ* and *pol4Δ* backgrounds for hAAG-induced (C) -1 frameshifts and (D) +1

frameshifts. The n values for the most bars in all graphs are equal to 3 (n=3), except for bars with their respective n value is mentioned in parenthesis above.

Supplementary Figure 3:

Sequences of polynucleotide DNA substrates that either, (A) contain base pair-loops; or (B) that do not contain base pair-loops.

Supplementary Figure 4:

Quantitative binding of hAAG proteins to poly-A-loop-A containing DNA.
Plots showing the calculation of binding affinity (Kd) values presented in Table 2 for the binding of (A) wild type hAAG, (B) hAAG-Y127I, (C) hAAG-Y127I/H136L and (D) all three proteins combined in one plot to poly-A-loop-A containing oligonucleotides. The experiments were performed by gel shift assays, using different concentrations of hAAG proteins (0-1000 nM) and 32P-labeled poly-A-loop-A oligonucleotide substrate. The Kd values were calculated as indicated in Materials and Methods section by fitting the data to the equation (3).

Supplementary Tables

Supplementary Table 1. List of expression vectors used in the study.

Plasmid	**Purpose**	**Source**
pTrcHis	*E. coli* hAAG-mutant library creation and selection of Rif^R-producing clones	Invitrogen
pET19b-PPS	*E. coli* hAAG protein purification	Lau et al., 1998; Lau et al. 2000
pYES2.0	*S. cerevisiae* expression of 3MeA DNA glycosylases	Invitrogen; Glassner et al. 1998
pCAGGS	High-expression eukaryotic vector for hAAG protein expression in K562 cells	Niwa et al., 1991
pMACS-KkII	Co-transfection with pCAGGS and positive K562 clone isolation	Miltenyi Biotech

Supplementary Table 2. List of strains used in the study.

Strain	Genotype	Reference
S. cerevisiae		
E133	*MATα ade5-1 his7-2 leu2-3,112 trp-289 ura3-52 lys2::InsE-A12*	(Tran et al., 1997)
E134	*MATα ade5-1 his7-2 leu2-3,112 trp-289 ura3-5 lys2::InsE-A14*	(Tran et al., 1997)
E133*apn1Δ*	E133 *apn1-Δ1::HIS3*	(Rusyn et al., 2007)
E134*apn1Δ*	E134 *apn1-Δ1::HIS3*	(Rusyn et al., 2007)
E133*rev1Δ*	E133 *rev1-Δ::kanMX4*	This study
E133*rev3Δ*	E133 *rev3-Δ::HIS3*	(Rusyn et al., 2007)
E133*rev7Δ*	E133 *rev7-Δ::kanMX4*	This study
E134*rev1Δ*	E134 *rev1-Δ::kanMX4*	This study
E134*rev3Δ*	E134 *rev3-Δ::kanMX4*	This study
E134*rev7Δ*	E134 *rev7-Δ::HIS3*	(Rusyn et al., 2007)
E133*rad30Δ*	E133 *rad30Δ::kanMX4*	This study
E134*rad30Δ*	E134 *rad30-Δ::kanMX4*	This study
E133*pol32Δ*	E133 *pol32-Δ::kanMX4*	This study
E134*pol32Δ*	E134 *pol32-Δ::kanMX4*	This study
E133*pol4Δ*	E133 *pol4-Δ::kanMX4*	This study
E134*pol4Δ*	E134 *pol4-Δ::kanMX4*	This study
E133*msh2Δ*	E133 *msh2-Δ::kanMX4*	This study
E133*msh6Δ*	E133 *msh6-Δ::kanMX4*	This study
E133*msh3Δ*	E133 *msh3-Δ::kanMX4*	This study
E134*msh6Δ*	E134 *msh6-Δ::kanMX4*	This study
E134*msh3Δ*	E134 *msh3-Δ::kanMX4*	This study
BGY111	*MATα his3 leu2-Δ trp-289$_{\alpha}$ ura3-52*	(Glassner et al., 1998b)
BGY148	BGY111 *mag1Δ2::LEU2*	(Glassner et al., 1998b)
E. coli		
AB1157	*F$^-$ thr-1 leu-6 proA2 thi-1 argE lacY1 galK ara-14 xyl-5 mtl-1 tsx-33 strA sup-37*	Lab stock
BW528	*AB1157* Δ*xth-pncA nfo-1*::Kan	Lab stock
BL21(DE3)	*F$^-$ ompT hsdS*(rB$^-$ mB$^-$) *dcm*$^+$ Tet^r *gal* λ(DE3) *endA* Hte	Lab stock
XL10 Gold	Tet^r Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* F´ *proAB lac*IqZΔ*M15* Tn*10* Cam^r	Stratagene

Supplementary Table 3. List of oligonucleotides used in the study. For site directed mutagenesis, mutated codons are shown bold and underlined.

Primers	**Sequence (5'-3')**
pCAGGS cloning	Forward: AGTCCTCGAGACCATGGTCACCCCCGCTTTGCAG
	Reverse: AGTCCTCGAGTCAGGCCTGTGTGTCCTGCTC.
hAAG PCR	Forward: GGAAGGAAATGGGCGGGGAGGGCC
	Reverse: CGCAAGCCCCGTCCCCCTGGCTGGAG
hAAG RT-PCR	Forward: CCCATACCGCAGCATCTATT
	Reverse: CGGAGTTCTGTGCCATTAGG
hAAG mutant Y127I	Forward: CGCATCGTGGAGACTGAGGCA**ATC**CTGGGGCCAGAGG
	Reverse: CCTCTGGCCCCAG**GAT**TGCCTCAGTCTCCACGATGCG
hAAG mutant H136L	Forward: GATGAAGCCGCC**CTG**TCAAGGGGTGGCCG
	Reverse: CGGCCACCCCTTGA**CAG**GGCGGCTTCATC
hAAG mutant E125A (with Y127I mutants)	Forward: CGCATCGTGGAGACC**GCC**GCA**ATC**CTGGGGCCAG
	Reverse: CTGGCCCCAG**GAT**TGC**GGC**GGTCTCCACGATGCG
hAAG mutant E125A (with wild type)	Forward: CGCATCGTGGAGACT**GCC**GCATACCTGGGGCCAG
	Reverse: CTGGCCCCAGGTATGC**GGC**AGTCTCCACGATGCG

Supplementary Experimental Protocols

Description of plasmids used in the study:

For the expression of 3MeA DNA glycosylases in yeast, cDNAs were cloned into the galactose-inducible pYES2.0 expression vector (Invitrogen Corp., San Diego, CA) and confirmed by DNA sequencing (performed at Elim Biopharmeceuticals, San Francisco, CA). Wild-type *hAAG* and *MAG1* containing pYES vectors were from Glassner et al, 1998. For analysis in *E. coli*, *hAAG* cDNAs were cloned into pTrcHis (Invitrogen) and confirmed by sequencing. For protein purification, the expression plasmid was a modified version of pET19b containing an inserted precision protease site (PPS) in front of *hAAG* cDNAs (Lau et al., 1998; Lau et al., 2000). The hAAG-Y127I/H136L double mutant was isolated from a large library generated by oligonucleotide-replacement of the active site residues with random nucleotides. hAAG-Y127I and hAAG-H136L single mutants were created by site directed-mutagenesis of the wild-type cDNA, using the oligonucleotide primers (Invitrogen) that contained point mutations (Supplementary Table 3). The nucleotide sequence of all mutants was confirmed by DNA sequencing. For hAAG cloning into the high-expression eukaryotic pCAGGS plasmid (Niwa et al., 1991), cDNA was PCR amplified from pYES-hAAG containing plasmids (Supplementary Table 3), and PCR product and plasmid were digested with *Xho* I (NEB, MA), ligated and transformed into competent XL10 Gold cells (Stratagene). Isolated plasmids (Mini Kit, QIAGEN) were screened for inserts and for the right orientation with restriction digests and DNA sequencing. For transfection into K562 cells and clone selection, pMACS-Kk.II (Miltenyi Biotech, Auburn, CA) plasmids were co-transfected with pCAGGS

vectors with or without *hAAG* cDNAs. Positive, H-2Kk surface-marker expressing cells were isolated according to manufacturer's instructions, cloned and screened by PCR for pCAGGS-hAAG DNA using primers from Supplementary Table 3.

S. cerevisiae strain construction and scoring of frameshift mutants:

Strains E133*apn1Δ*, E134*apn1Δ*, E134*rev7Δ*, E133*rev3Δ* were constructed previously by disruption of the gene with the *HIS3* cassette (Glassner et al., 1998b; Hofseth et al., 2003; Rusyn et al., 2007). All other mutant alleles were derived from the *S. cerevisiae* Gene Deletion Library as follows. *kanMX4* cassette was PCR amplified from the genome of the appropriate strain in the deletion library and transformed into wild type E133 and E134 strains. The resulting deletion strains were confirmed by genomic PCR and by their sensitivity to MMS and other agents. Briefly, overnight cultures were serially diluted by 5-fold and stamped onto MMS-containing YPD plates, grown for 3 days at 30^{o}C and compared to wild-type E133/E134 strains. The *kanMX4* cassettes and pYES vectors were introduced into the yeast strains by way of G418-resistance or uracil auxotrophy, respectively, utilizing Geitz Transformation Kit and protocol. For determination of frameshift mutants, overnight cultures were diluted into 10 parallel cultures at an initial density of 4000 cells/ml. Cultures were grown to saturation at 30^{o}C for 4 days in SC-ura synthetic media containing 2% galactose (BioGene, CA). +1 and -1 frameshift revertants were scored on SC-lys/agar plates (Glassner et al., 1998), whereas base pair substation mutants were scored on SC-trp/agar plates (Glassner et al., 1998).

Purification of hAAG glycosylases from pET19b-PPS vector:

pET19b-PPS vectors containing recombinant *hAAG* cDNAs were introduced into BL21(DE3) expression cells by the calcium chloride method. Expression was induced by 0.5 μM IPTG in late-exponential cultures and overnight incubation at room temperature. The truncated hAAG constructs were deleted for the N-terminal 80 amino acids to facilitate purification and to improve the solubility as previously described (Lau et al., 1998). The soluble hAAG proteins tagged with hexahistidine (His_6) tag were purified to apparent homogeneity using three chromatographic steps: His-Trap Sepharose gel affinity column (Amersham Biosciences), and His_6 tag removal with PreScission Protease (Pharmacia) treatment overnight at 16^{o}C, His-Trap Sepharose gel affinity purification, and finally HiTrap sepharose cation exchange (Amersham Biosciences) and Superdex-75 gel filtration (Amersham Pharmacia Biotech) in gel filtration buffer (20mM HEPES, pH 7.5, 100mM NaCl, 1mM DTT, 5mM EDTA, and 10% glycerol). The final purified hAAG proteins were stored in the gel filtration buffer and were greater than 95% pure as evidenced by SDS-PAGE analysis.

DNA glycosylase activity assays:

DNA glycosylase assays were performed at 37^{o}C in a buffer containing 20 mM Tris-HCl pH 7.8, 100 mM KCl, 5mM β- mercaptoethanol, 6 mM EDTA, 4 nM double stranded ^{32}P-labeled DNA oligonucleotides containing hAAG substrates and 400 nM purified hAAG. The oligos were cleaved at AP-sites with 30 min 0.1 N NaOH treatment at 70ºC. Reaction products were separated on a 20% urea denaturing gel, imaged with Molecular Dynamics PhosphorImager and quantified using Kodak ID 1.0 software.

For excision of 3MeA and 7MeG, tritiated calf thymus DNA was incubated with hAAG glycosylases in glycosylase buffer (20 mM Tris pH 7.6, 100 mM KCl, 5 mM EDTA, 1 mM EGTA, 5 mM β-mercaptoethanol and BSA to 500 nmol of total protein per reaction) for indicated amounts of time. The reaction was terminated by the addition of 0.1 volume of sodium acetate buffer and DNA was precipitated with ethanol. Bases released by DNA glycosylase were spotted onto Whatman 3MM paper and descending paper chromatography was performed for 18 hrs in 70% isopropanol, 20% water and 10% ammonium hydroxide. Radioactivity was counted with Beckman LS6000IC scintillation counter.

Gel mobility shift assays:

Protein-DNA binding assays were carried out in buffer containing 4 mM Tris-HCl pH 7.8, 6 mM Hepes-KOH pH 7.8, 20 mM KCl, 30 mM NaCl, 0.5 mM EDTA, 1 mM β-mercaptoethanol, 10 ng Salmon Sperm DNA, 10% glycerol, and 0.5 nM ^{32}P-labeled DNA oligonucleotides (oligos) containing known or putative hAAG substrates. Purified hAAGs were incubated in the buffer at 4ºC for 30 min and separated on a 5% non-denaturing polyacrylamide gel (PAGE). The gel was run in 0.5x TBE buffer at 130 V and 4ºC for 210 min, vacuum dried onto 3 MM Whatman paper, imaged with Molecular Dynamics PhosphorImager and quantified using Kodak ID 1.0 software.