Structural studies of human alkyladenine glycosylase and *E. coli* 3-methyladenine glycosylase

Thomas Hollis, Albert Lau, Tom Ellenberger *

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

Abstract

Human alkyladenine glycosylase (AAG) and *Escherichia coli* 3-methyladenine glycosylase (AlkA) are base excision repair glycosylases that recognize and excise a variety of alkylated bases from DNA. The crystal structures of these enzymes have provided insight into their substrate specificity and mechanisms of catalysis. Both enzymes utilize DNA bending and base-flipping mechanisms to expose and bind substrate bases. Crystal structures of AAG complexed to DNA suggest that the enzyme selects substrate bases through a combination of hydrogen bonding and the steric constraints of the active site, and that the enzyme activates a water molecule for an in-line backside attack of the *N*-glycosylic bond. In contrast to AAG, the structure of the AlkA–DNA complex suggests that AlkA substrate recognition and catalytic specificity are intimately integrated in a S_N1 type mechanism in which the catalytic Asp238 directly promotes the release of modified bases. Published by Elsevier Science B.V.

Keywords: AlkA; AAG; Glycosidase mechanism; Protein–DNA complex; Base-flipping

1. Introduction

Maintaining the integrity of DNA is critical for proper cellular function as well as correct propagation of the genetic code. Cells rely on enzymes that can accurately detect, remove and repair damaged bases in DNA. The primary method for removal of single base aberrations is the base excision repair (BER) pathway. This process begins with *N*-glycosylases that locate damaged bases within a vast excess of normal DNA then expose the target nucleotide and cleave the glycosylic bond. The resulting abasic site is processed by a repair endonuclease followed by DNA polymerase and DNA ligase to restore the original DNA sequence. Although the crystal structures of several BER glycosylases are known, we still know little of how these enzymes search for damaged bases or the exact mechanism of distorting the DNA to capture the damage base in the enzyme active site.

Most BER glycosylases are highly specific, excising one type of modified base or a few closely related modifications. Some alkylation damage repair glycosylases such as *Escherichia coli* 3-methyladenine glycosylase (AlkA), *Saccharomyces cerevisiae* Mag and human alkyladenine glycosylase (AAG), however, are able to recognize a chemically varied and structurally diverse group of modified bases. All of these enzymes excise 3-methyladenine, 7-methyladenine, and 7-methylguanine from DNA

* Corresponding author.
  E-mail address: tome@hms.harvard.edu (T. Ellenberger).

[1], and AAG efficiently removes $1N^6$-ethenoadenine and hypoxanthine as well [2,3] (Fig. 1).

Although the alkylation repair glycosylases from various organisms have broadly overlapping functions, they fall into several discrete structural families. The *E. coli* AlkA protein is a member of the Helix-hairpin-Helix family of glycosylases that includes the endonuclease III, MutY, and 8-oxoguanine glycosylase (OGG) proteins. These proteins contain little amino acid sequence similarity except in the region of the HhH motif and a conserved catalytic aspartate, but share a very similar three-dimensional structure [4,5]. *S. cerevisiae* Mag shares the conserved HhH region and catalytic aspartate and there-fore is probably a member of this family as well. The human AAG glycosylase is in a separate family consisting of the highly homologous mammalian glycosylases that have a different structure [6]. Whereas AlkA is a mostly α-helical protein, AAG has a mixed α/β-structure that lacks an HhH motif. The structures of AAG and AlkA suggest that these enzymes might have very different catalytic mechanisms as well.

AAG and AlkA recognize very similar substrates and both enzymes utilize a base flipping mechanism to rotate the substrate base into the active site. Base flipping was first described for the cytosine methyl-transferases *Hae*III and *Hha*I [7, 8]. It has since
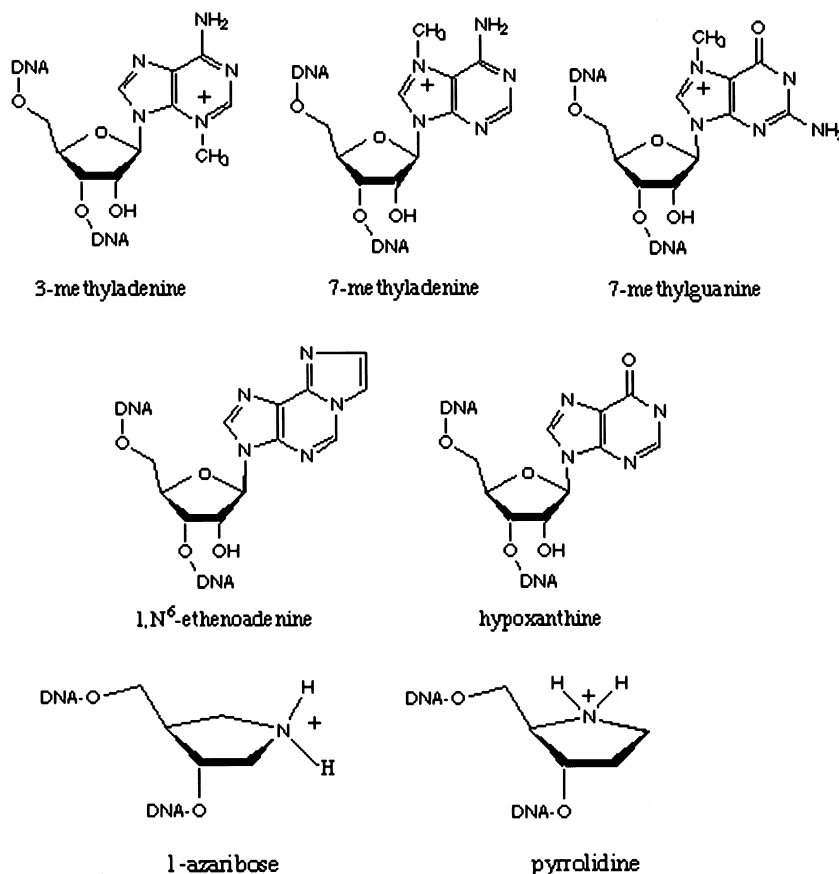

Fig. 1. Substrates and inhibitors of alkylation damage repair glycosylases. AAG and AlkA recognize a variety of alkylated bases including 3-methyladenine, 7-methyladenine, and 7-methylguanine. AAG also efficiently cleaves 1,$N6$-ethenoadenine and hypoxanthine, whereas AlkA does not. 1-azaribose and pyrrolidine are potent abasic inhibitors of AlkA and AAG, presumably acting as a transition state mimics [10,26]. These nucleotides were used for capturing the protein-DNA complex for the crystal structures of AlkA and AAG, respectively.

been reported for a number of other enzymes that catalyze nucleotide modifications within duplex DNA [9]. This scheme of rotating a base out of the DNA helix and into an active site pocket exposes the substrate and allows the tight control of substrate solvation and reaction stereochemistry. It remains to be proven, however, whether base flipping is an active process carried out by the enzyme upon binding or a passive event in which the enzyme captures a spontaneously rotated nucleotide. The recent proliferation of structural and biochemical information for DNA glycosylases and other base flipping enzymes has provided new insights into the mechanics of DNA repair. Here, we describe the structural progress made for the alkylation repair glycosylases AAG and AlkA, and we outline the implications to the biology and mechanisms of this very important process.

2. AAG structure

The crystal structure of AAG was initially determined in complex with duplex DNA containing a transition state mimic and potent inhibitor of the glycosylase reaction, a pyrrolidine abasic nucleotide (*pyr*) [6,10] (Fig. 2). In the complex, the *pyr* is flipped out of the DNA duplex and into a pocket on the enzyme surface. Tyr162 inserts into the minor groove of the DNA where its side chain fills the space vacated by the flipped-out nucleotide thus, the intercalating tyrosine acts as a surrogate base that presumably stabilizes the nucleotide targeted for repair in an extra-helical conformation. The substrate-binding pocket contains the enzyme active site, where a water molecule is positioned to attack the flipped-out nucleotide. This geometry is consistent with a direct, in-line displacement of a damaged base from the DNA backbone.

The fragment of AAG that was crystallized lacks residues 1–79 but it retains enzymatic activity and DNA binding specificity that are identical to those of longer protein constructs. A homologous truncation fragment of mouse MPG has similar kinetic properties and specific binding patterns as full-length MPG [11]. AAG folds into a single domain of mixed α/β structure that, overall, does not resemble any other
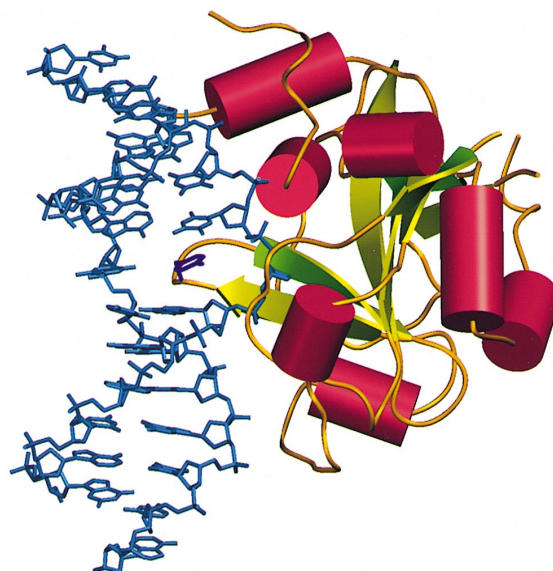

Fig. 2. Crystal structure of human AAG in complex with pyrrolidine-containing DNA. The protein fold is a single domain that binds DNA in the minor groove. Tyr162 (purple) fills the space left by the flipped-out pyrrolidine. The protein contacts the DNA mainly at the phosphate backbone. There are no base-specific interactions.

model in the PDB. On the surface of the protein, clusters of basic residues that flank the flipped-out *pyr* anchor the DNA phosphodiester backbone across the substrate-binding pocket. A β-hairpin juts out from this surface to insert the Tyr162 side chain into the minor groove of the DNA. AAG acts only on duplex DNA, and there are about the same number of protein-DNA backbone contacts for each DNA strand. Uracil DNA glycosylase (UDG), on the other hand, removes uracil from both single-and double-stranded DNA, and the crystal structure of a UDG-DNA complex shows that all of the protein-DNA contacts, except one, are with the uracil-containing strand [12]. The buried DNA surface in the AAG–DNA complex is 1034 Å^2 (measured with a 1.4 Å probe), which is similar to that in the UDG-DNA complex [12]. The buried DNA surface in these complexes is significantly less than that in the cytosine methyltransferase complexes [7,8]. On the whole, the bound DNA is a B-form helix that bends away from the protein. The DNA is bent by about 22° where Tyr162 intercalates between base-pairs. The modest AAG–DNA interaction surface, together with the absence of base sequence-specific interactions, might facilitate sliding of the protein along DNA in its search for substrate [13].

The pyrrolidine abasic nucleotide is rotated out of the DNA double helix into the protein active site. Here, a water molecule links the *pyr* N4′ to the side chains of Glu125 and Arg182 and the main-chain carbonyl of Val262 through hydrogen bonding interactions. This hydrogen bonding network likely contributes to the tight binding of *pyr*-containing DNA to AAG [10]. Furthermore, the position of the water with respect to the flipped-out *pyr* and the nearby amino acids strongly suggests that Glu125 acts as a general base to deprotonate the water for the backside attack of the *N*-glycosylic bond. The hydrogen bonding between Arg182 and the water might assist nucleophilic attack of the nucleotide by stabilizing the incipient hydroxide ion.

In order to gain insight into the structural basis for AAG's specific recognition of a diverse group of modified bases, we constructed an E125Q AAG point mutant protein and determined the crystal structure of this protein complexed to $1,N^6$-ethenoadenine (εA)-containing DNA (manuscript in preparation.). εA is formed by exposure of DNA to vinyl

chloride or chloroacetaldehyde [14] and this bulky adduct is efficiently removed from DNA by AAG [2]. The substitution of Glu125 with a glutamine
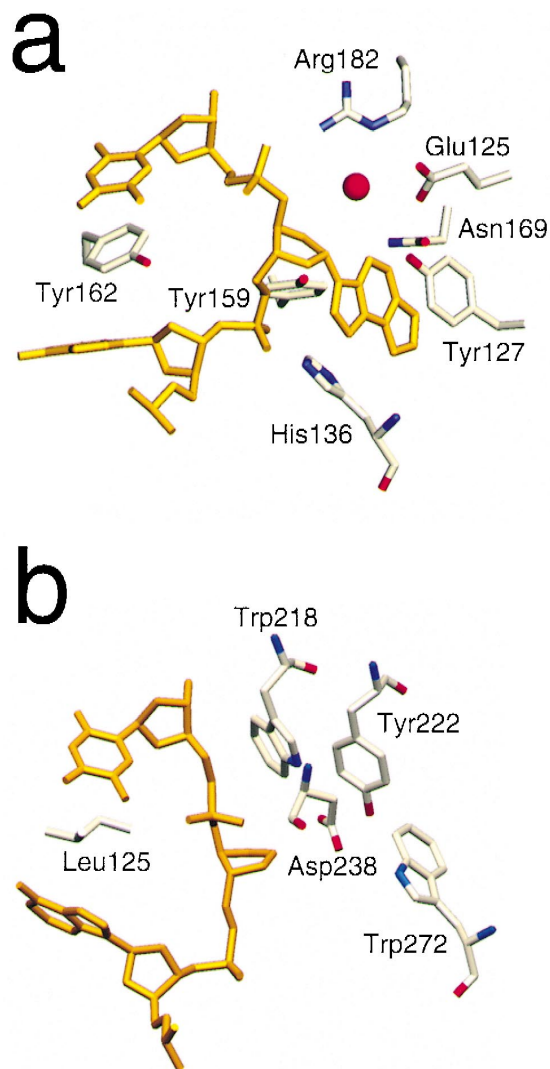

Fig. 3. A comparison of the active sites of AAG and AlkA. In the AAG–εA complex (a) the εA flips into the protein active site to stack snugly between Tyr127 on one side and His136 and Tyr159 on the other (Lau et al., manuscript in preparation). Glu125 is well positioned to deprotonate the bound water for a backside attack of the *N*-glycosylic bond. In the AlkA-DNA complex (b) the 1-azaribose is flipped out of the DNA helix into the active site such that the positively charged nitrogen of the 1-azaribose interacts with the carboxylate of the catalytic Asp238. Trp218 shields the substrate from solvent and leaves no room on the backside of the flipped-out ribose for nucleophilic water.

residue apparently prevents activation of the active site water and subsequent release of the modified base, allowing crystallization of protein–substrate complex.

The overall folds of the E125Q and wild-type proteins differ very little. In the E125Q–εA complex, the εA base is flipped into the enzyme active site where it stacks between Tyr127 on one side and His136 and Tyr159 on the other (Fig. 3a). The hydroxyl group of Tyr127 hydrogen bonds with Glu125, which could serve to position Glu125 in the active site. The εA has rotated about its glycosylic bond 85° away from its B-DNA anti conformation towards a high-anti conformation. The sugar pucker appears to be C2′-endo in the 2.4 Å resolution structure. A high-anti conformation for an unstacked deoxyadenosine nucleoside exhibiting a C2′-endo sugar pucker has been observed by NMR [15].

The bound water resides in a very similar location in the active sites of both complexes. The water's hydrogen bonding partners, however, are different. In the *pyr* complex, the water interacts with the *pyr* N4′, the side chains of Glu125 and Arg182, and the main-chain carbonyl of Val262. In the E125Q–εA complex, the water remains hydrogen bonded to the side chain of Arg182 and the main-chain carbonyl of Val262. In the absence of a *pyr* N4′ and with the replacement of Glu125 with Gln125, however, the other bonding partners of the water have shifted to the O3′ of εA and the terminal amine of Gln125.

3. AlkA structure

Crystal structures of *E. coli* AlkA [4,16] reveal a compact globular protein consisting of three subdomains. An amino-terminal subdomain (residues 1–88) forms a mixed α − β structure that is similar to the tandemly repeated fold of the TATA-binding protein, but it plays no direct role in DNA binding by AlkA [17]. The middle (residues 113–230) and carboxyl-terminal (residues 231–282) subdomains are α-helical bundles that are structurally homologous to MutY [5] and EndoIII [18]. The conserved HhH motif is located at the interface of these subdomains, adjacent to the enzyme active site. AlkA lacks the iron–sulfur cluster located near the carboxyl-terminus of EndoIII and MutY. AlkA's active site is located within a large hydrophobic cleft and it contains the catalytically essential Asp238.

The recent crystal structure of an AlkA–DNA complex has provided the first glimpse of DNA binding for a member of the HhH superfamily as well as some insight into AlkA's base recognition and catalytic mechanism [17]. The structure of AlkA complexed with DNA containing a 1-azaribose abasic nucleotide shows that AlkA induces a severe distortion in the bound DNA (Fig. 4). A sharp bend (66°) is centered around the flipped out abasic nucleotide where the minor groove is markedly widened. The enzyme uses a combination of polar and nonpolar interactions with the DNA to achieve this distortion. The mostly nonpolar DNA binding surface consists two loops flanking the nucleotide-binding pocket. These loops form a ''wedge'' that is inserted into the minor groove of DNA. At the tip of this wedge is Leu125 whose side chain occupies the position of the flipped out base, analogous to Tyr162 in AAG. Of the few polar interactions made with the DNA, most are contributed by the HhH motif that anchors the DNA to the protein on the 3′ side of the flipped out nucleotide. Residues 202–227, which form the HhH motif of AlkA, provide several hydrogen bonds and a metal mediated interaction with the DNA. The metal ion, modeled as sodium, is coordinated by the main-chain carbonyl oxygens of residues 210, 212, and 215 in the hairpin of the HhH, a DNA phosphate oxygen and a water molecule (Fig. 5). The DNA is apparently necessary for stable binding of the metal to the HhH motif because the metal is not present in the high-resolution structures of the AlkA, EndoIII lacking DNA, and MutY lacking DNA. Similar DNA binding motifs have been seen in enzymes such as DNA polymerase β [19] and more recently in interferon regulatory factor (IRF) [20]. Polymerase β has two HhH domains that are instrumental in binding to DNA. These HhH motifs ligate a metal ion only while DNA is bound and prefer monovalent sodium and potassium over divalent calcium and magnesium. In the IRF crystal structure, a related Helix-hairpin-Strand motif was found to ligate a potassium ion in its interaction with DNA. A number of other DNA binding proteins are predicted to contain this motif [19]. This method of DNA interaction has probably evolved as an efficient platform for tight binding, non-specific DNA recogni-
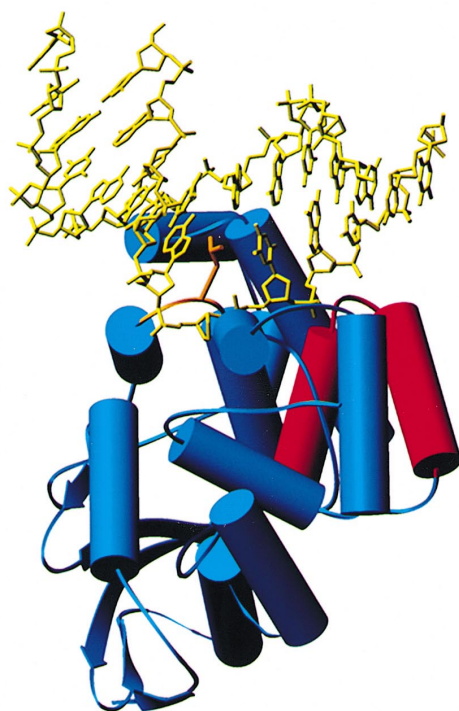
Fig. 4. AlkA bound to DNA. The protein induces a large distortion in the bound DNA. The minor groove of the protein is widened to ~15 Å and the DNA is bent by 66°. The target nucleotide is rotated out of the DNA helix and into the enzyme active site. Leu125 (shown in orange) intercalates into the DNA and fills the gap created by the flipped out base. The Helix-hairpin-Helix motif (shown in red) anchors the DNA to the protein via hydrogen bonds and a metal mediated interaction (cf. Fig. 5).

tion. Furthermore, the AlkA–DNA structure shows that the HhH motif serves as a stable support for the protein-induced DNA distortions but it does not directly participate in the flipping of a nucleotide substrate.

In the structure of the AlkA–DNA complex [17], the 1-azaribose is flipped out of the DNA helix and into the enzyme active site. The positively charged nitrogen of the 1-azaribose is situated about 3.2 Å from the carboxylate of Asp238, the catalytically essential residue. Leu125 is pushed into the minor groove where it occupies the position of the flipped out nucleotide. The geometry of the DNA distortion is quite different from that seen in the human UDG [12,21]. UDG creates a ''pinch'' in the DNA backbone, compressing the inter-phosphate distance

around the flipped out base, thereby producing a zig-zag conformation in the phosphodiester backbone. In the AlkA–DNA complex, the phosphate backbone around the 1-azaribose is fully extended and unlike the DNA distortion induced by UDG.

The crystal structures of AlkA, Endo III, and MutY reveal remarkably similar protein folds despite a very low overall sequence similarity [4,5,16,18,22]. Conservation of the HhH motif, as well as overall structure, implies that these proteins may interact with DNA in a similar fashion. Superposition of the DNA from the AlkA–DNA complex onto the unliganded structures of EndoIII and MutY reveals that the proposed DNA binding surfaces interact with the DNA without any major steric clashes and contain a similar nonpolar chemistry as the AlkA binding surface [17]. The residues of MutY (Gln42) and EndoIII (Gln41) that have been proposed to be involved in base flipping [5] are properly positioned to intercalate into the DNA in these modeled complexes.
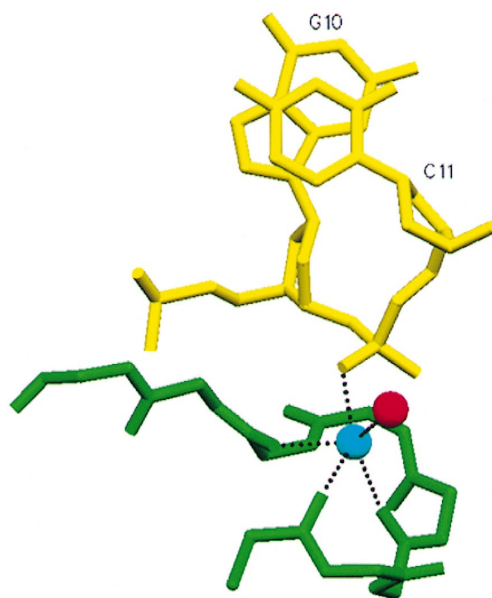

Fig. 5. Metal mediated protein-DNA interaction in AlkA HhH. The HhH motif of AlkA uses hydrogen bonding and a bound metal to create a tight binding sequence independent interaction with DNA. The metal ion (shown in blue) was modeled as sodium and is coordinated by a DNA phosphate oxygen, a water (shown in red) and by three main-chain carbonyl oxygens in the HhH motif.

4. Substrate recognition

Whereas most BER glycosylases are specific for a particular type of base, AlkA and AAG are unusual in that they recognize and excise a chemically and structurally diverse group of alkylated bases from DNA. The basis for this broad substrate specificity is not completely understood. Electron-deficient alkylated bases in DNA might be recognized and targeted for excision by virtue of their positive charge [4,6,23,24]. Partially charged bases would stack more tightly against the aromatic residues of the active site than uncharged normal bases thereby allowing the enzyme to distinguish alkylated bases from normal bases.

The substrate bases most efficiently cleaved by AlkA contain a delocalized positive charged and the structure of the AlkA–DNA complex has allowed new insight into how this enzyme might decipher between damaged and undamaged bases [17]. Superposition of a 3-methyladenine nucleoside substrate on the flipped out 1-azaribose abasic nucleotide in the active site of AlkA shows that the 3-methyladenine base stacks against Trp272 and that C1′ of the ribose is adjacent to Asp238. This stacking of a positively charged substrate against the aromatic tryptophan constitutes a cation–π interaction that could stabilize the extra-helical conformation of the substrate base. The open architecture of the binding site identified by modeling allows ample room to accommodate all alkylated substrate bases. Interestingly, the modeled base makes no detectable hydrogen bond contacts to the AlkA active site, suggesting that the enzyme's activity towards alkylated substrates may result from factors other than binding to a modified base. Additionally, AlkA's indifference to the base opposite of the flipped out nucleotide can be easily explained by the lack of any direct protein contacts to that base.

In contrast to AlkA, some of AAG's substrates are not positively charged, e.g., εA and inosine, suggesting that other types of interactions selectively stabilize bound substrates in AAG's active site. The active site accommodates the εA very snugly between Tyr127 on one side and His136 and Tyr159 on the other. This complementarity does not result from an induced fit to the εA because the residues of the *pyr* complex, with an unfilled active site, are

essentially superimposable on those of the complexes with the εA. An εA has 17 Å^2 more van der Waals contact surface with the sandwiching residues than an unmodified adenine. Also, a hydrogen bond forms between the main-chain amide of His136 and N6 of εA (Fig. 6). The N6 of a normal adenine could not accept this bond and would instead be repelled. AAG might therefore preferentially bind εA over an unmodified adenine by selective hydrogen bond formation and additional van der Waals interactions. How is a modified guanine recognized as being different from a normal guanine? The O6 of a normal guanine, when superimposed in the AAG active site, could accept the hydrogen bond from the main-chain amide of His136 like the N6 of εA, but the guanine's N2 would create a steric clash with Asn169, resulting in a repulsive interaction that might cause the guanine to be expelled from the active site (Lau et al., manuscript in preparation). The positive charge, however, on a substrate base such as 7-methylguanine might cause the base to be pulled into the active site to satisfy cation–π interactions with Tyr127 and His136 strongly enough so as to push
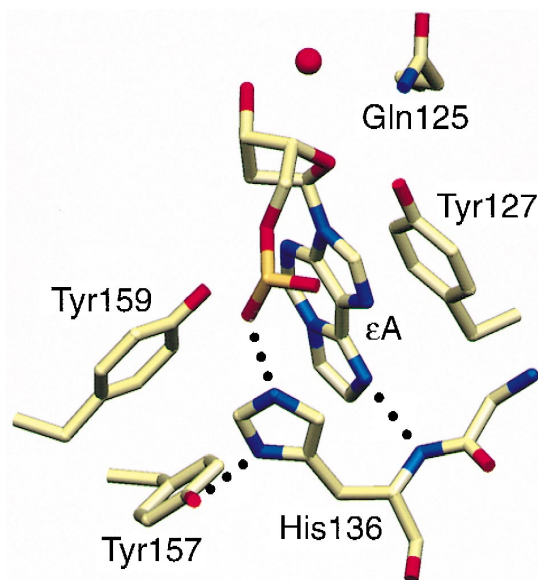

Fig. 6. The N6 of the εA hydrogen bonds with the main-chain amide of His136. The N6 of a normal adenine would not accept this hydrogen bond. This selective hydrogen bond formation might cause the protein to bind εA preferentially over a normal adenine. His136 also forms hydrogen bonds with Tyr157 and the 5′ phosphate of the εA.

aside Asn169 [23,24]. This hypothesis has yet to be tested. Like guanine, the O6 of an inosine substrate could accept the hydrogen bond from the main-chain amide of His136, but inosine lacks an N2, allowing the base to fit comfortably in the active site.

5. Catalysis by AlkA and AAG

It is thought that hydrolysis of *N*-glycosidic bonds by nucleoside hydrolase enzymes involves a positively charged transition state intermediate [25], and that this carbocation character builds up on the C1′–O4′ positions of the substrate ribose during bond cleavage. Presumably, the pyrrolidine and 1-azaribose abasic nucleotides, which have a positively charged N4′ and N1′, respectively, mimic some features of a transition state intermediate and thus, bind very tightly to AAG and AlkA [10,26]. Structures of AlkA and AAG in complex with these inhibitors have shed light on the possible mechanisms of enzymatic hydrolysis of substrate bases.

The crystal structure of the AAG–pyrrolidine complex reveals a water molecule positioned about 3.1 Å from the N4′ of the pyrrolidine and Glu125 in the enzyme active site [6] (Fig. 3a). The water is ideally positioned for an in-line displacement of the glycosidic bond via a backside attack of the ribose. Glu125 might deprotonate this water to form a strong nucleophile that could then attack the C1′ of the substrate ribose and subsequently release the base.

It was previously assumed that the catalytic mechanism of AlkA would also involve the activation of a water molecule by the conserved Asp238 [4,10,27], similar to the S_N2 mechanism proposed for AAG. The crystal structure of the AlkA–DNA complex, however, reveals that there is no room on the backside of the flipped out ribose for a nucleophilic water. The N1′ of the 1-azaribose is in direct contact with the carboxylate oxygen of Asp238 and shielded by Trp 218 from any solvent, thereby making the structure inconsistent with an S_N2 mechanism (Fig. 3b). Instead, an alternative mechanism is proposed involving the direct interaction of Asp238 with the anomeric position of the ribose, or an S_N1 mechanism [17]. AlkA's substrate bases are positively charged, or in a sense, activated for leaving. Proper positioning of the C1′ of the ribose of a substrate

base next to Asp238 may allow the stabilization of a carbocation intermediate, either through charge interaction or a covalent intermediate, and cleavage of the glycosidic bond. Unmodified and uncharged bases in the active site pocket would not be cleaved because they present poor leaving groups, thus providing AlkA with catalytic selectivity.

6. Implications and future directions

The crystal structures of DNA glycosylases have shed much light on the possible catalytic mechanisms of glycosylic bond cleavage, as well as substrate recognition in the active site. Some lingering questions remain, however, on the exact mechanism of locating damaged bases in a vast excess of normal DNA. It is still unclear if DNA glycosylases are able to find aberrations in normal B-form DNA or whether an abnormal base causes some sort of helix destabilization and conformational change, creating a signal for the glycosylase. Both crystallographic and NMR studies of DNA containing modified bases such as $1N^6$-ethenoadenine, 7-methylguanine, or inosine however, show little if any effect on the overall B-form structure of the DNA [28–30]. This may indicate that glycosylases do not depend on gross structural irregularities in DNA for locating a damaged base. Furthermore, it is unknown whether these enzymes use a processive mechanism to scan the DNA or distributive method of binding and releasing until it happens upon a damaged base. Experiments performed with uracil-DNA glycosylase and UV endonuclease suggest that at least these glycosylases operate with a processive mechanism [31–33]. It is conceivable that AAG and AlkA could operate in a similar manner. Both proteins make tight-binding, non-specific contacts with the DNA backbone, which may allow the protein to stay bound without flipping a base and while diffusing along the DNA. The minor groove reading elements of the respective proteins may act as a sensor or probe for damage as the protein moves along. Upon reaching a position of a weakened base–base interaction (i.e., methylated base, abasic site, base mismatch) the probe, Leu125 in AlkA or Tyr162 in AAG, could force itself into the helix, displacing the aberrant nucleotide into the active site pocket. This type of mechanism may also

be aided by the strain of DNA bending and the widening of the minor groove, which could act as a spring to help force the intercalating residue into the helix. Once the base is positioned in the active site, further selectivity for a substrate may take place. This could be hydrogen bonding and/or steric interference between base and active site residues, as proposed for AAG, or a physical property of the substrate such as charge and glycosylic bond strength as proposed for AlkA.

The crystal structures of DNA repair glycosylases have proven exceptionally valuable in answering some of the questions about base excision repair and they underscore the diversity of solutions to the problem of maintaining and repairing DNA. Unfortunately, the structures do not provide the full insight to function. Rather, they pose new questions that must be addressed before we can have a complete understanding of base excision repair.

References

[1] B. Singer, B. Hang, What structural features determine repair enzyme specificity and mechanism in chemically modified DNA?, Chem. Res. Toxicol. 10 (1997) 713–732.

[2] M. Saparbaev, K. Kleibl, J. Laval, *Escherichia coli*, *Saccharomyces cerevisiae*, rat and human 3-methyladenine DNA glycosylases repair 1,*N*6-ethenoadenine when present in DNA, Nucleic Acids Res. 23 (1995) 3750–3755.

[3] M. Saparbaev, J. Laval, Excision of hypoxanthine from DNA containing dIMP residues by the *Escherichia coli*, yeast, rat, and human alkylpurine DNA glycosylases, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 5873–5877.

[4] J. Labahn, O.D. Scharer, A. Long, K. Ezaz-Nikpay, G.L. Verdine, T.E. Ellenberger, Structural basis for the excision repair of alkylation-damaged DNA, Cell 86 (1996) 321–329.

[5] Y. Guan, R.C. Manuel, A.S. Arvai, S.S. Parikh, C.D. Mol, J.H. Miller, S. Lloyd, J.A. Tainer, MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily, Nat. Struct. Biol. 5 (1998) 1058–1064.

[6] A.Y. Lau, O.D. Scharer, L. Samson, G.L. Verdine, T. Ellenberger, Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision, Cell 95 (1998) 249–258.

[7] S. Klimasauskas, S. Kumar, R.J. Roberts, X. Cheng, HhaI methyltransferase flips its target base out of the DNA helix, Cell 76 (1994) 357–369.

[8] K.M. Reinisch, L. Chen, G.L. Verdine, W.N. Lipscomb, The crystal structure of HaeIII methyltransferase convalently complexed to DNA: an extrahelical cytosine and rearranged base pairing, Cell 82 (1995) 143–153.

[9] R.J. Roberts, X. Cheng, Base flipping, Annu. Rev. Biochem. 67 (1998) 181–198.

[10] O.D. Scharer, H.M. Nash, J. Jiricny, J. Laval, G.L. Verdine, Specific binding of a designed pyrrolidine abasic site analog to multiple DNA glycosylases, J. Biol. Chem. 273 (1998) 8592–8597.

[11] R. Roy, T. Biswas, T.K. Hazra, G. Roy, D.T. Grabowski, T. Izumi, G. Srinivasan, S. Mitra, Specific interaction of wild-type and truncated mouse *N*-methylpurine-DNA glycosylase with ethenoadenine-containing DNA, Biochemistry 37 (1998) 580–589.

[12] G. Slupphaug, C.D. Mol, B. Kavli, A.S. Arvai, H.E. Krokan, J.A. Tainer, A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA, Nature 384 (1996) 87–92.

[13] G.L. Verdine, S.D. Bruner, How do DNA repair proteins locate damaged bases in the genome?, Chem. Biol. 4 (1997) 329–334.

[14] Z. Matijasevic, M. Sekiguchi, D.B. Ludlum, Release of N2,3-ethenoguanine from chloroacetaldehyde-treated DNA by *Escherichia coli* 3-methyladenine DNA glycosylase II, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 9331–9334.

[15] D.J. Weber, G.P. Mullen, A.S. Mildvan, Conformation of an enzyme-bound substrate of staphylococcal nuclease as determined by NMR, Biochemistry 30 (1991) 7425–7437.

[16] Y. Yamagata, M. Kato, K. Odawara, Y. Tokuno, Y. Nakashima, N. Matsushima, K. Yasumura, K. Tomita, K. Ihara, Y. Fujii, Y. Nakabeppu, M. Sekiguchi, S. Fujii, Three-dimensional structure of a DNA repair enzyme, 3-methyladenine DNA glycosylase II, from *Escherichia coli*, Cell 86 (1996) 311–319.

[17] T. Hollis, Y. Ichikawa, T. Ellenberger, DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, *Escherichia coli* AlkA, EMBO J. 19 (2000) 101–109.

[18] C.F. Kuo, D.E. McRee, C.L. Fisher, S.F. O'Handley, R.P. Cunningham, J.A. Tainer, Atomic structure of the DNA repair [4Fe–4S] enzyme endonuclease III, Science 258 (1994) 434–440.

[19] H. Pelletier, M.R. Sawaya, Characterization of the metal ion binding helix-hairpin-helix motifs in human DNA polymerase beta by X-ray structural analysis, Biochemistry 35 (1996) 12778–12787.

[20] Y. Fujii, T. Shimizu, M. Kusumoto, Y. Kyogoku, T. Taniguchi, T. Hakoshima, Crystal structure of an IRF-DNA complex reveals novel DNA recognition and cooperative binding to a tandem repeat of core sequences, EMBO J. 18 (1999) 5028–5041.

[21] S.S. Parikh, C.D. Mol, G. Slupphaug, S. Bharati, H.E. Krokan, J.A. Tainer, Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA, EMBO J. 17 (1998) 5214–5226.

[22] M.M. Thayer, H. Ahern, D. Xing, R.P. Cunningham, J.A. Tainer, Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure, EMBO J. 14 (1995) 4108–4120.

[23] J.P. Gallivan, D.A. Dougherty, Cation–pi interactions in structural biology, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 9459–9464.

[24] G. Hu, P.D. Gershon, A.E. Hodel, F.A. Quiocho, mRNA cap recognition: dominant role of enhanced stacking interactions between methylated bases and protein aromatic side chains, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 7149–7154.

[25] V.L. Schramm, B.A. Horenstein, P.C. Kline, Transition state analysis and inhibitor design for enzymatic reactions, J. Biol. Chem. 269 (1994) 18259–18262.

[26] K. Makino, Y. Ichikawa, Synthesis of a 2-deoxy-ribose type 1-*N*-iminosugar, Tetrahedron Lett. 39 (1998) 8245–8248.

[27] B. Sun, K.A. Latham, M.L. Dodson, R.S. Lloyd, Studies on the catalytic mechanism of five DNA glycosylases. Probing for enzyme-DNA imino intermediates, J. Biol. Chem. 270 (1995) 19501–19508.

[28] J.C. Xuan, I.T. Weber, Crystal structure of a B-DNA dodecamer containing inosine, d(CGCIAATTCGCG), at 2.4 A resolution and its comparison with other B-DNA dodecamers, Nucleic Acids Res. 20 (1992) 5457–5464.

[29] G.A. Leonard, K.E. McAuley-Hecht, N.J. Gibson, T. Brown, W.P. Watson, W.N. Hunter, Guanine-1,*N*6-ethenoadenine base pairs in the crystal structure of d(CGCGAATT(epsilon dA)GCG), Biochemistry 33 (1994) 4755–4761.

[30] K. Ezaz-Nikpay, G.L. Verdine, The effects of *N*7-methyl-guanine on duplex DNA structure, Chem. Biol. 1 (1994) 235–240.

[31] S.E. Bennett, R.J. Sanderson, D.W. Mosbaugh, Processivity of *Escherichia coli* and rat liver mitochondrial uracil-DNA glycosylase is affected by NaCl concentration, Biochemistry 34 (1995) 6109–6119.

[32] M. Higley, R.S. Lloyd, Processivity of uracil DNA glycosylase, Mutat. Res. 294 (1993) 109–116.

[33] R.W. Hamilton, R.S. Lloyd, Modulation of the DNA scanning activity of the Micrococcus luteus UV endonuclease, J. Biol. Chem. 264 (1989) 17422–17427.