Structure, Vol. 13, 87–98, January, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.str.2004.10.011

A DNA Glycosylase from *Pyrobaculum aerophilum* with an 8-Oxoguanine Binding Mode and a Noncanonical Helix-Hairpin-Helix Structure

Gondichatnahalli M. Lingaraju,[1] Alessandro A. Sartori,[2,3] Dirk Kostrewa,[1] Andrea E. Prota,[1] Josef Jiricny,[2] and Fritz K. Winkler[1,*]

[1]Biomolecular Research
Paul Scherrer Institut
CH-5232, Villigen
Switzerland

[2]Institute of Molecular Cancer Research
University of Zurich
August Forel-Strasse 7
CH-8008, Zurich
Switzerland

Summary

Studies of DNA base excision repair (BER) pathways in the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum* identified an 8-oxoguanine-DNA glycosylase, *Pa*-AGOG (archaeal GO glycosylase), with distinct functional characteristics. Here, we describe its crystal structure and that of its complex with 8-oxoguanosine at 1.0 and 1.7 Å resolution, respectively. Characteristic structural features are identified that confirm *Pa*-AGOG to be the founding member of a functional class within the helix-hairpin-helix (HhH) superfamily of DNA repair enzymes. Its hairpin structure differs substantially from that of other proteins containing an HhH motif, and we predict that it interacts with the DNA backbone in a distinct manner. Furthermore, the mode of 8-oxoguanine recognition, which involves several hydrogen-bonding and π-stacking interactions, is unlike that observed in human OGG1, the prototypic 8-oxoguanine HhH-type DNA glycosylase. Despite these differences, the predicted kinked conformation of bound DNA and the catalytic mechanism are likely to resemble those of human OGG1.

Introduction

One of the principal and most deleterious products of oxidative DNA damage is 8-oxoguanine (GO) (Grollman and Moriya, 1993). In the *anti*-conformation, GO forms a Watson-Crick base pair with cytosine (GO:C), but in the *syn*-conformation it can form a stable Hoogsteen pair with adenine (GO:A). Erroneous incorporation of dAMP opposite GO during DNA replication can thus lead to G:C→T:A transversion mutations (A-Lien et al., 2001; Grollman and Moriya, 1993; Michaels and Miller, 1992; Pearson et al., 2004; Shimizu et al., 2003).

Like most single-base modifications in DNA, GO residues are repaired predominantly by the base excision repair (BER) pathway (reviewed in Hazra et al., 2001; Scharer and Jiricny, 2001), which is initiated by specific DNA glycosylases. These enzymes can be grouped into two major mechanistic classes (Fromme et al., 2004b). Monofunctional DNA glycosylases remove the aberrant base through the single-step hydrolysis of the glycosidic bond and thus generate apurinic or apyrimidinic (AP) sites in the DNA. Bifunctional DNA glycosylases/AP lyases deploy an amino function of the enzyme to form a Schiff's base intermediate, which undergoes β-elimination through a multistep reaction cascade that leads to DNA strand scission 3′ from the lesion (McCullough et al., 1999, 2001).

Structural studies of DNA glycosylases have thus far identified four superfamilies (Fromme et al., 2004b), two of which comprise members capable of GO-specific base excision. In bacteria, removal of GO from GO:C base pairs is mediated by the Fpg/MutM enzymes (Boiteux et al., 1987). The more recently identified eukaryotic Fpg homologs known as Neil (Nei-like) proteins (Zharkov et al., 2003) preferentially recognize oxidized pyrimidines (Wallace et al., 2003). In humans, GO is processed by a structurally unrelated enzyme, OGG1, a bifunctional DNA glycosylase (Radicella et al., 1997) originally identified in the yeast *Saccharomyces cerevisiae* (Nash et al., 1996; van der Kemp et al., 1996). OGG1 belongs to a superfamily of DNA repair enzymes that process a wide variety of substrates. Although members of this family have little sequence identity, they share a conserved two-domain fold containing a helix-hairpin-helix DNA binding motif, followed by a glycine/proline-rich stretch and an invariant aspartate (HhH-GPD motif) (Nash et al., 1996; Thayer et al., 1995). Several members of this superfamily such as the bacterial endonuclease III (*Nth*), AlkA, and MutY enzymes as well as and the human OGG1 (hOGG1) have been structurally characterized (Fromme et al., 2004b).

OGG1 displays a clear preference for GO:C substrates (Bjoras et al., 1997; Girard et al., 1998), which was explained by the finding that the protein interacts not only with the modified purine that is flipped out of the helix, but also with the widowed C in the other strand (Bruner et al., 2000). The crystal structure of a catalytically inactive mutant of hOGG1 bound to a double-stranded DNA fragment containing a GO:C base pair shows that the enzyme discriminates GO from guanine with the help of a single hydrogen bond formed between a main chain carbonyl oxygen and the protonated N^7-position of the aberrant purine (Bruner et al., 2000). The same kind of GO discrimination was observed also in the bacterial MutM enzyme (Fromme and Verdine, 2003b). Interestingly, the 8-oxo carbonyl function is devoid of interaction partners in both these crystal structures. The bifunctional, DNA glycosylase/β-lyase HhH-GPD glycosylases use a conserved lysine as catalytic nucleophile to generate a covalently linked enzyme-DNA adduct, which undergoes a series of subsequent transformations resulting in DNA strand scission on the 3′ side of the lesion. The mechanism has been investigated using modified substrates and dif-

*Correspondence: fritz.winkler@psi.ch
[3]Present address: The Wellcome Trust/Cancer Research, UK Gurdon Institute, Tennis Court Road, Cambridge CB2 1QR, United Kingdom.

ferent mechanistic schemes have been considered and discussed (McCullough et al., 2001; Zharkov et al., 2000). More recent studies of the catalytic mechanism of hOGG1 yielded two unexpected results. First, it appears that cleavage of the glycosidic bond proceeds through a dissociative S_N1-type mechanism and that the ϵ-amino group of the catalytic lysine is not activated by the catalytic aspartate as previously assumed (Norman et al., 2003). Second, it was shown that the excised base may act as a cofactor in the β-lyase reaction (Fromme et al., 2003).

Pyrobaculum aerophilum is a facultatively aerobic hyperthermophilic crenarchaeon growing at an optimum temperature of 100°C (Volkl et al., 1993). Its genome encodes several uracil DNA glycosylases that protect it from the deleterious effects of cytosine deamination (Sartori et al., 2001, 2002; Yang et al., 2000). It also encodes a highly-active GO-specific DNA glycosylase, *Pa*-AGOG (Sartori et al., 2004), which represents the founding member of a functional class within the HhH superfamily of DNA repair enzymes. This enzyme is unique for one other reason: unlike the other characterized GO DNA glycosylases, *Pa*-AGOG efficiently excises GO from all substrates, irrespective of whether the DNA is single stranded, or whether the oxidized guanine is situated opposite T, C, G, or A. Database searches failed to identify structurally or functionally characterized *Pa*-AGOG homologs, but secondary structure predictions grouped it with members of the HhH-GPD superfamily of DNA glycosylases (Sartori et al., 2004). We now describe the high resolution crystal structure of *Pa*-AGOG, both alone and in complex with 8-oxoguanosine, which reveals that *Pa*-AGOG indeed belongs to the HhH superfamily of DNA repair enzymes despite the absence of the otherwise highly conserved hairpin consensus sequence. It also recognizes and binds GO very differently from hOGG1. Structural comparisons with other members of the HhH superfamily reveal distinct features that define *Pa*-AGOG as the founding member of an oxoguanine-specific, archaeal DNA glycosylase/lyase family of proteins.

Results and Discussion

Pa-AGOG Is a Member of the HhH-GPD DNA Glycosylase Superfamily

As illustrated in Figure 1, the overall fold of *Pa*-AGOG shares striking similarity with the catalytic core structures of the DNA repair enzymes hOGG1, EndoIII, and MutY. The structure of *Pa*-AGOG consists of two α-helical domains, with the 8-oxoguanine binding site located in a cleft at their interface (Figure 1A). The larger domain, referred to as the six-helix barrel or HhH domain, consists of six topologically conserved helices (α4–α9) and comprises additional nonconserved helices in some cases. The smaller domain, referred to as the four-helix or N/C domain (because it contains both the N and C termini), differs somewhat from those of the other proteins, in as much as only three of its helices (α10, α12, and α13) appear topologically conserved. The large and small domains are linked by two polypeptide chains, one being the loop that connects α9, the second helix of the HhH motif, to the α10 helix that carries the catalytic aspartate. Spatial alignment of the various HhH and loop domains (Figure 1C) revealed that the positions of the catalytic Lys and Asp residues that are conserved among the bifunctional DNA glycosylases/lyases of the HhH-GPD family (Nash et al., 1996) (Figure 1D) are superimposable. This alignment identified K140 and D172 as the likely catalytic residues of *Pa*-AGOG, a prediction that could be substantiated by site-directed mutagenesis studies (Sartori et al., 2004).

The HhH motif is found in many DNA-interacting proteins (Doherty et al., 1996), but is considered the hallmark of the HhH superfamily of DNA repair glycosylases (Denver et al., 2003). In the known structures, the hairpin, which is highly conserved in sequence and structure (canonical hairpin structure, Figures 1C and 1D), interacts with two DNA phosphates primarily via main chain N-H groups. In the case of EndoIII (Fromme and Verdine, 2003a) and AlkA (Hollis et al., 2000a), the hairpin also contributes three main chain carbonyl oxygen atoms to a cation binding site that becomes occupied in the presence of substrate DNA. In *Pa*-AGOG, the hairpin sequence motif hPGhG (h indicating a hydrophobic residue) is not conserved (Figure 1D) and it contains a two-residue insertion. This produces quite a different hairpin structure which is stabilized by several hydrogen bonds and nonpolar interactions and appears unlikely to undergo a major conformational change upon DNA binding. The 3-methyladenine DNA glycosylase I (TAG) enzyme of *E. coli* is a unique member of the HhH superfamily whose hairpin sequence also differs from the consensus and, like *Pa*-AGOG, it also has a two residue insertion. However, closer inspection of its NMR structure (Drohat et al., 2002) reveals that, in contrast to *Pa*-AGOG, its altered hairpin structure could still form the characteristic interactions with the DNA.

The second (α9) helix of the HhH motif is about twice as long in *Pa*-AGOG as in the other enzymes. In order to maintain the geometry around the catalytic cleft, a longer peptide sequence is needed to connect it to the helix (α10) carrying the catalytic aspartate (Figures 1C and 1D). This linker peptide resembles the Gly/Pro-rich GPD motif (Nash et al., 1996) of the other enzymes of this family. The three hydrophobic residues N-terminal to the conserved aspartate are in an extended conformation and can be superposed rather closely with the other structures, with the exception of AlkA, which has an insertion in this motif. Notably, when the structurally-conserved residues of the HhH-GPD motifs of these enzymes are superposed, the structures align very closely also for the first two turns of the helix following the catalytic aspartate. As noted earlier (Norman et al., 2003), the α10 helix of *Pa*-AGOG is an integral component of the active site architecture and the conserved aspartate acts as its N-terminal helix cap. We will therefore refer to it as the aspartate-cap helix.

The Base Binding Pocket of Pa-AGOG Is Highly GO Specific

Our attempts to cocrystallize *Pa*-AGOG with a variety of DNA substrates have so far met with failure. In order to determine the mode whereby the enzyme recognizes

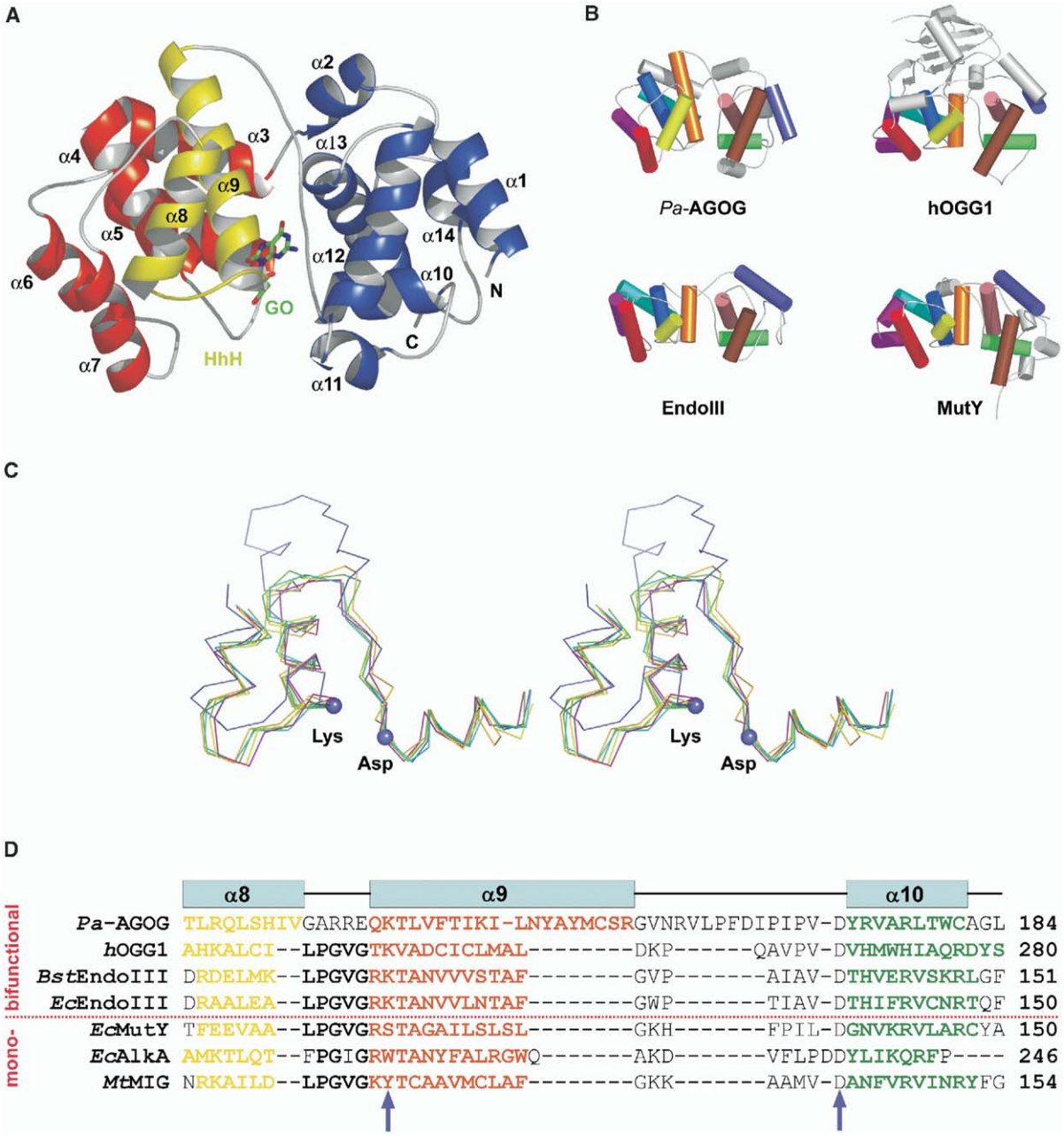

Figure 1. Overall Structure and HhH Motif of *Pa*-AGOG

(A) Ribbon drawing of the *Pa*-AGOG monomer. The HhH- and the N/C-domains are shown in red and blue, respectively. The HhH motif is highlighted in yellow and the bound 8-oxoguanosine is shown in green stick representation.

(B) Comparison of *Pa*-AGOG and other HhH-type DNA glycosylase structures. Helices are shown as cylinders. The topologically conserved helices can be superimposed with an rmsd of 3.3 Å (58 C_α atoms, hOGG1, PDB code 1ebm [Bruner et al., 2000]), 2.8 Å (50 C_α atoms, *Ec*EndoIII, 2abk [Thayer et al., 1995]), and 3.3 Å (56 C_α atoms, *Ec*MutY, 1muy [Guan et al., 1998]), and are highlighted using the same color code. Nonconserved structural elements are shown in gray.

(C) Superposition, in stereo, of the HhH C_α-traces of *h*OGG1 (green, 26 C_α, rmsd 1.1 Å), *Bst*EndoIII (yellow, 26 C_α, rmsd 1.2 Å, 1p59 [Fromme and Verdine, 2003a]), *Ec*MutY (magenta, 26 C_α, rmsd 1.0 Å), *Ec*AlkA (orange, 24 C_α, rmsd 1.3 Å, 1diz, [Hollis et al., 2000a, 2000b]), and *Mt*MIG (cyan, 23 C_α, rmsd 1.1 Å, 1kea [Mol et al., 2002]) on *Pa*-AGOG (slate). The catalytic residues Lys and Asp are indicated as slate spheres.

(D) Structure-based sequence alignment of the HhH motifs shown in (C). The sequences are highlighted using the same color code as in (B). *Pa*-AGOG secondary structure elements are shown schematically on top. The arrows indicate the positions of the catalytic aspartate (conserved in all members) and that of the catalytic lysine (conserved only in the bifunctional enzymes).

GO, we soaked the *Pa*-AGOG crystals in crystallization buffer containing 10 mM 8-oxoguanosine. The difference electron density derived from data, collected to 1.7 Å resolution (see Experimental Procedures and Table 1), showed clear density for 8-oxoguanosine and some accompanying structural adjustments in the active site region. Far from the active site, a second large density difference was observed between two

Table 1. Data Collection and Refinement Statistics

	Native[a]	GO-Complex[b]	SeMetλ1 (Infl)[c]	SeMetλ2 (Peak)[d]	SeMetλ3 (Remote)[e]
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Cell constants (Å)	a = 70.9, b = 97.7, c = 36.1	a = 72.2, b = 98.1, c = 36.2	a = 69.1, b = 96.9, c = 36.0	a = 69.1, b = 96.9, c = 36.0	a = 69.1, b = 96.9, c = 36.0
Source	SLS-PX	Enraf-Nonius FR591	SLS-PX	SLS-PX	SLS-PX
Wavelength (Å)	0.9000	1.5418	0.9796	0.9794	0.9717
Resolution (Å)[g]	40–1.03 (1.1–1.03)	40–1.69 (1.8–1.69)	40–1.15 (1.2–1.15)	40–1.15 (1.2–1.15)	40–1.15 (1.2–1.15)
No. total observations	587,271	194,367	632,108	633,381	632,480
No. unique observations	116,782	28,623	166,339	166,338	166,338
Completeness (%)[g]	93.7 (78.1)	96.7 (90.5)	100 (100)	100 (100)	100 (100)
$\langle I/\sigma(I)\rangle$[g]	11.6 (2.5)	17.0 (3.5)	11.6 (2.3)	10.5 (2.6)	11.6 (1.9)
R_{sym} (%)[f,g]	5.8 (46.8)	5.4 (57.8)	5.7 (53.6)	6.6 (48.5)	5.3 (65.8)
No. selenium sites			2	2	2
Phasing power[h] (acentric/centric)					
Isomorphous			1.40/1.01	1.08/0.76	—/—
Anomalous			1.13/—	2.10/—	0.98/—
R_{cullis} [i] (acentric/centric)					
Isomorphous			0.60/0.63	0.71/0.75	—/—
Anomalous			0.79/—	0.58/—	0.84/—
Figure of merit[j]					
Acentric/centric reflections					0.57/0.45
After density modification (all refl.)					0.92
No. water molecules	346	227			
R_{work}[k]	0.16	0.18			
R_{free}[k]	0.18	0.23			
Rmsd bonds (Å)	0.015	0.016			
Rmsd angles (°)	1.7	1.6			
Average B factors (Å^2); protein/water/GO ligand	17.4/29.2/—	32.3/39.2/24.4			

[a] *Pa*-AGOG native.
[b] *Pa*-AGOG-GO complex.
[c] Selenomethionine-derived *Pa*-AGOG at inflection.
[d] Selenomethionine-derived *Pa*-AGOG at peak.
[e] Selenomethionine-derived *Pa*-AGOG at remote.
[f] $R_{sym} = \Sigma|I_{hkl} - \langle I_{hkl}\rangle|/\Sigma I_{hkl}$, where I is the intensity of a reflection hkl and <I> is the average over symmetry-related reflections of hkl.
[g] Values in parentheses refer to the highest resolution shell.
[h] Phasing power = $\langle|F_H|\rangle/\langle||F_{PH}| - |F_P + F_H||\rangle$ in which F_H is the heavy atom structure factor amplitude, F_P is the protein structure factor amplitude, and F_{PH} is the structure factor amplitude of the heavy atom derivative.
[i] $R_{cullis} = \Sigma||F_{PH}| - |F_P + F_H||/\Sigma|F_{PH} - F_P|$
[j] Figure of merit = $\langle\Sigma P(\alpha)e^{i\alpha}/\Sigma P(\alpha)\rangle$, in which α is the phase, $P(\alpha)$ is the phase probability distribution, and α ranges from 0 to 2π.
[k] $R_{work} = \Sigma|F_o - F_c|/\Sigma|F_o|$ in which F_o and F_c are the observed and calculated structure factor amplitudes, respectively. R_{free} is calculated from 5% of the reflections not used in the model refinement.

neighboring molecules across the crystallographic 2-fold axis and was interpreted as a nonspecific 8-oxoguanosine binding site generated in this particular crystal lattice.

Like enzymes that chemically modify bases in double-stranded DNA (Roberts and Cheng, 1998), DNA glycosylases bind their respective substrates in an extrahelical conformation (Fromme et al., 2004b). In agreement with this expectation, GO is seen bound in a deep pocket of *Pa*-AGOG (Figures 2A and 2B), where it is sandwiched between the aromatic side chains of Phe144 and Trp222. The protein contacts all H-bond donor and acceptor positions of GO, with the exception of N^3. Of greatest interest is the recognition of the urea system of GO, which discriminates it from guanine. The carboxamide oxygen of Gln31 recognizes the N^7-H donor function and the $N^{\epsilon 1}$ of Trp69 forms a hydrogen bond with the O^8 carbonyl oxygen. The remaining interactions of *Pa*-AGOG with GO are guanine specific and clearly exclude A, C, and T from the binding pocket. The O^6 oxygen accepts one hydrogen bond from the carboxamide nitrogen of Gln31 and one from a water molecule interacting with Lys147 and Asp29. This water molecule is part of an extended hydrogen bond network involving the side chains of Gln31, Lys147, Asp 218, and Thr219 (Figure 2B) and the more distant Glu26, Asp29, Tyr151, and His216. These residues are conserved in all the archaeal homologs (Figure 5). This network is evident also in the free enzyme structure, with the difference that the GO binding pocket contains several well defined water molecules, some of which occupy positions of hydrogen bond donor and acceptor sites (N^2, N^3, O^6, and O^8) of GO in the complex. The tight and cooperative nature of this network is illustrated by the fact that the Lys147Gln mutation (Sartori et al., 2004) reduces both the activity and the thermostability of the enzyme.

GO Binding Induces a Number of Structural Changes

Pa-AGOG undergoes a number of apparently concerted conformational changes in both main chain and side chain conformations upon nucleoside binding. Significant changes were observed in the immediate vicinity of the GO binding pocket (Figure 3B), in an adja-

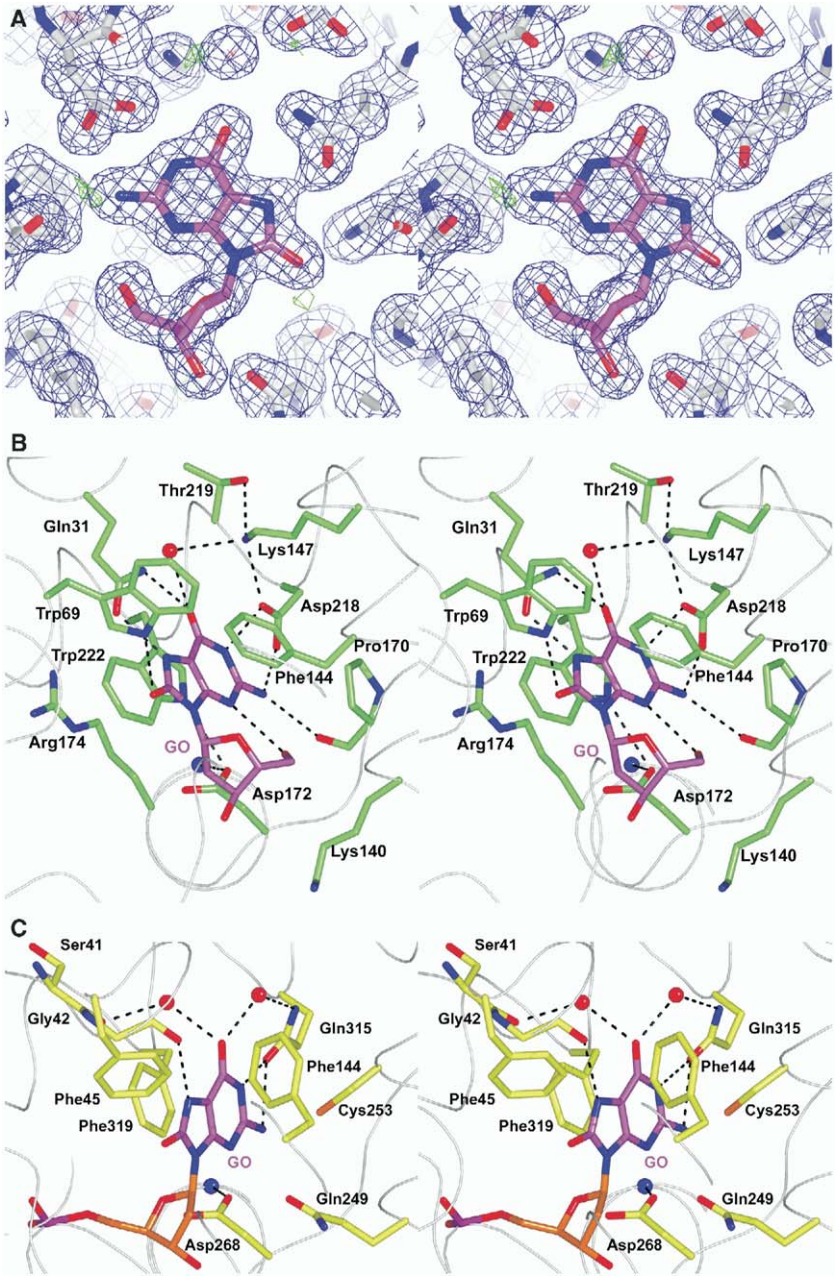

Figure 2. Structure of the *Pa*-AGOG Active Site and Comparison of GO Recognition with hOGG1

(A) Final electron density map at 1.7 Å resolution centered at GO. The $2F_{obs} - F_{calc}$ map is contoured in blue at 1.0 σ, the $F_{obs} - F_{calc}$ map is in red and green at –3 / +3 σ, respectively.

(B) Stereo view into the active site of *Pa*-AGOG showing the specific GO recognition formed by a tight hydrogen bond network and additional stacking interactions.

(C) Similar stereo view (as in [B]) into the active site of hOGG1. Three of the five residues that are in direct contact with the aberrant base in *Pa*-AGOG (Phe144, Asp218, and Trp222) have structurally equivalent residues in conserved α helices of hOGG1 (Cys253, Gln315, and Phe319) (Bruner et al., 2000).

cent surface loop, predicted to be involved in the interaction with substrate DNA, and in a more remote surface segment (Figure 3A). Overall, we observed a small but significant closure of the two domains over the substrate binding cleft.

As discussed above, the GO binding pocket is largely preformed in the free enzyme structure, with Phe144 and the in-plane hydrogen bonding network almost ready to accommodate the aberrant base after the expulsion of the place holding water molecules. The notable exception is Trp222, which appears to be intrinsically mobile, as indicated by the presence of two distinct conformations of its side chain in the free enzyme structure. In the complex, the indole group is seen swung into a stacking position on the π face of GO (Figure 3B). This change is accompanied by a number of other structural rearrangements. The guanidinium group of Arg174, previously involved in a hy-

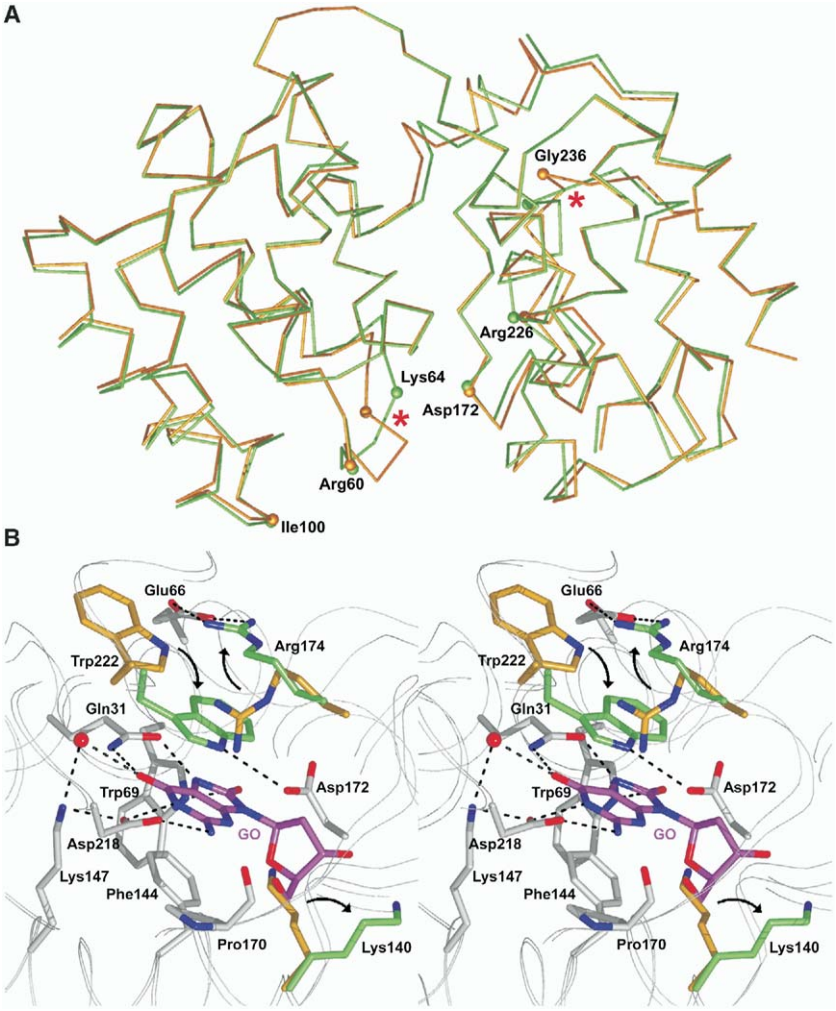

Figure 3. Conformational Changes upon GO Binding

(A) Superposition of the free (orange) and complexed structure of *Pa*-AGOG with 8-oxoguanosine (green). The two segments that undergo major main chain structural changes are highlighted by a red asterisk.

(B) Structural rearrangements within the *Pa*-AGOG active site. The side chains of the active site residues that undergo structural rearrangements upon ligand binding are highlighted in orange (free enzyme) and green (nucleoside/enzyme complex). The directions of the movements are indicated by black arrows.

drogen bond to the catalytic Asp172, becomes stacked over the indole ring of Trp222, where it forms a strong salt bridge/hydrogen bonding contact with Glu66. In turn, the indole N-H of Trp222 replaces Arg174 in hydrogen bonding to the catalytic Asp172 side chain. The repositioning of Glu66 appears coupled to a major change in the conformation of segment Leu61-Lys64 (Figure 3A) and some less dramatic adjustments in adjacent residues. The likely interaction of this segment with DNA is discussed below. Additional changes in side chain conformations appear to propagate along the exposed surface of helix α13. The reorientation of the side chain of Lys64 correlates with the repositioning of the Arg226 and Glu232 side chains on α13 and this probably provokes the observed change in the main chain conformation of the loop connecting α13 to α14 (Glu232 to His235).

The second binding site of GO on the surface of the protein is remote from the active site and, apart from a change in the side chain conformation of Arg86, does not induce significant structural changes. Crystals soaked with 2′-deoxyguanosine show no binding to the GO recognition site (G.M.L., unpublished data). However, the nonspecific binding site becomes similarly occupied, strengthening the notion that it is not biologically relevant.

Pa-AGOG Recognizes GO Differently from hOGG1 and MutM

The structures of two GO-specific DNA glycosylases, hOGG1 (Bruner et al., 2000) and MutM (Fromme and Verdine, 2003b), bound to GO-containing duplexes have been determined using catalytically-inactive mutants. Interestingly, the 8-oxo-carbonyl function is devoid of a hydrogen bonding partner in both these enzymes, such that GO is distinguished from G only by nature of the N^7 position (an H-bond acceptor in G but a donor in GO) which is probed by a main chain carbonyl

group. Apart from this common feature, the two enzymes interact with GO very differently (Fromme and Verdine, 2003b). This is perhaps not too surprising, given that MutM does not belong to the HhH-GPD superfamily. As might be anticipated from the structural similarity of the folds of hOGG1 and *Pa*-AGOG, three of the five residues that are in direct contact with the aberrant base in *Pa*-AGOG (Phe144, Asp218, and Trp222), have functional analogs in hOGG1 (Cys253, Gln315, and Phe319), which occupy structurally equivalent positions in the conserved helical framework. However, the two enzymes interact very differently with the urea function of GO, in as much as the amino acid residues involved in its recognition, Gln31 and Trp69 in *Pa*-AGOG, and Gly42 in hOGG1, are not conserved (Figure 2).

The glycosidic bond of the bound GO nucleoside is in the *syn*-conformation, which was reported to be thermodynamically favored also in free 8-oxoguanosine (Uesugi and Ikehara, 1977). The intramolecular hydrogen bond between the 5′-OH group and N^3 that helps stabilize this conformation is maintained in the bound nucleoside (Figures 2 and 3). In the hOGG1 complex, the extrahelical base is observed in the *anti* conformation while it is bound in the *syn* conformation in the complexes of MutY (Fromme et al., 2004a) and MutM (Fromme and Verdine, 2003b). We are certain that the observed binding mode of the GO base with its many precisely matched interactions must be essentially the same for a DNA substrate. In contrast, the location and orientation of the nucleoside sugar moiety cannot be representative for a substrate complex. It is completely different from that observed in the complexes of hOGG1 and EndoIII with their respective DNA substrates (Bruner et al., 2000; Fromme and Verdine, 2003a) and extensions at its 3′ and 5′ ends to build a DNA strand are sterically impossible. As the GO recognition pockets of *Pa*-AGOG and hOGG1 are similarly positioned and oriented with respect to the active site and as a very similar binding mode is indicated for the lesion carrying DNA strand (see below), we anticipate that *Pa*-AGOG will bind the aberrant base also in the *anti* conformation. That the observed nucleoside sugar conformation and location cannot be functionally relevant is also indicated by the fact that sugar and catalytic Lys140 (Figure 3B) have become arranged in a way that excludes an attack of Lys140 on C1′.

Predicted Interaction with DNA

The overall mode of interaction of the HhH type DNA glycosylases hOGG1 (Bruner et al., 2000), EndoIII (Fromme and Verdine, 2003a), MutY (Fromme et al., 2004a), and AlkA (Hollis et al., 2000a) with their respective DNA substrates appears to be very similar. In all complexes, the bound DNA is highly distorted, with a bend of 55°–70° centered at the site of the flipped-out base (Fromme et al., 2004a). The widened minor groove faces the protein and the flipped-out base is bound in the recognition pocket. All enzymes interact primarily with the strand carrying the aberrant base, which runs through the deep cleft at the domain interface and forms direct hydrogen bonds between the enzymes and four or five contiguous DNA phosphates.

To examine whether and how *Pa*-AGOG could interact with DNA distorted in this manner we have carried out structural superpositions and modeling studies (Figure 4). Indeed, a rather convincing fit of the lesion strand backbone between phosphates P_{-3} to P_1 to the surface of *Pa*-AGOG is observed and a number of very plausible interactions can be postulated (Figure 4A). Only small structural adaptations suffice to join the GO base as observed in the active site of *Pa*-AGOG to the DNA backbone as observed in the hOGG1-DNA complex (Figures 4B and 4C). In the known complexes, several hydrogen bonds to the five contacted phosphates are provided by main chain N-H groups and most of these appear equally possible in *Pa*-AGOG. The most notable exception is the interaction with P_{-3} which is contacted by the backbone of the N-terminal part of the HhH hairpin in all known DNA HhH motif complexes. In *Pa*-AGOG, the different structure and location of the hairpin does not permit an analogous direct interaction. Instead, the side chain of Arg108, which is conserved in the archaeal homologs, is optimally positioned to directly contact P_{-3} (Figure 4B). Arg174, another strictly conserved residue in the archaeal homologs, is predicted to form a salt bridge to P_0. Its side chain has been observed to become involved in a salt bridge with the equally conserved Glu66 across the domain interface upon GO binding and only in this conformation is it suitably positioned for DNA interaction (Figure 4C).

A characteristic feature of the observed protein/DNA interaction interfaces is that two loops from the HhH domain insert in a wedge-like fashion into the minor groove, widened at the bend. The first, connecting the helices equivalent to α4 and α5 of *Pa*-AGOG (Figures 1A and 1B) inserts a residue deeply into the minor groove at the position of the flipped-out base. On the basis of the structural alignments, we propose this residue to be Arg60 in *Pa*-AGOG (Figure 3A). Notably, the segment containing amino acid residues 61–64 as seen in the structure of the free enzyme would cause a severe steric clash with the GO-containing strand. However, this enzyme segment undergoes a major conformational change upon nucleoside binding, which suggests that it is flexible and thus able to accommodate DNA. Interestingly, DNA binding brought about the rearrangement of the corresponding chain segment also in hOGG1 (Bjoras et al., 2002). The second loop entering the widened minor groove connects the two helices equivalent to α6 and α7 of *Pa*-AGOG, and inserts a poorly conserved residue into the open space between the widowed base and the strongly tilted base on its 5′ side. The structural alignments position Ile100 or Gly101 of *Pa*-AGOG closest to the residues that occupy this space in the mesophilic enzymes. The weak and poorly conserved interactions observed in this region of the known complexes preclude any reliable detailed prediction.

In summary, the structural superpositions predict that the distorted conformation of bound DNA observed for hOGG1 and particularly that of the GO-containing DNA strand will be similar for *Pa*-AGOG. In addition, a number of the interactions with the DNA phosphates that are conserved in the known complexes appear similarly possible through analogous structural ele-

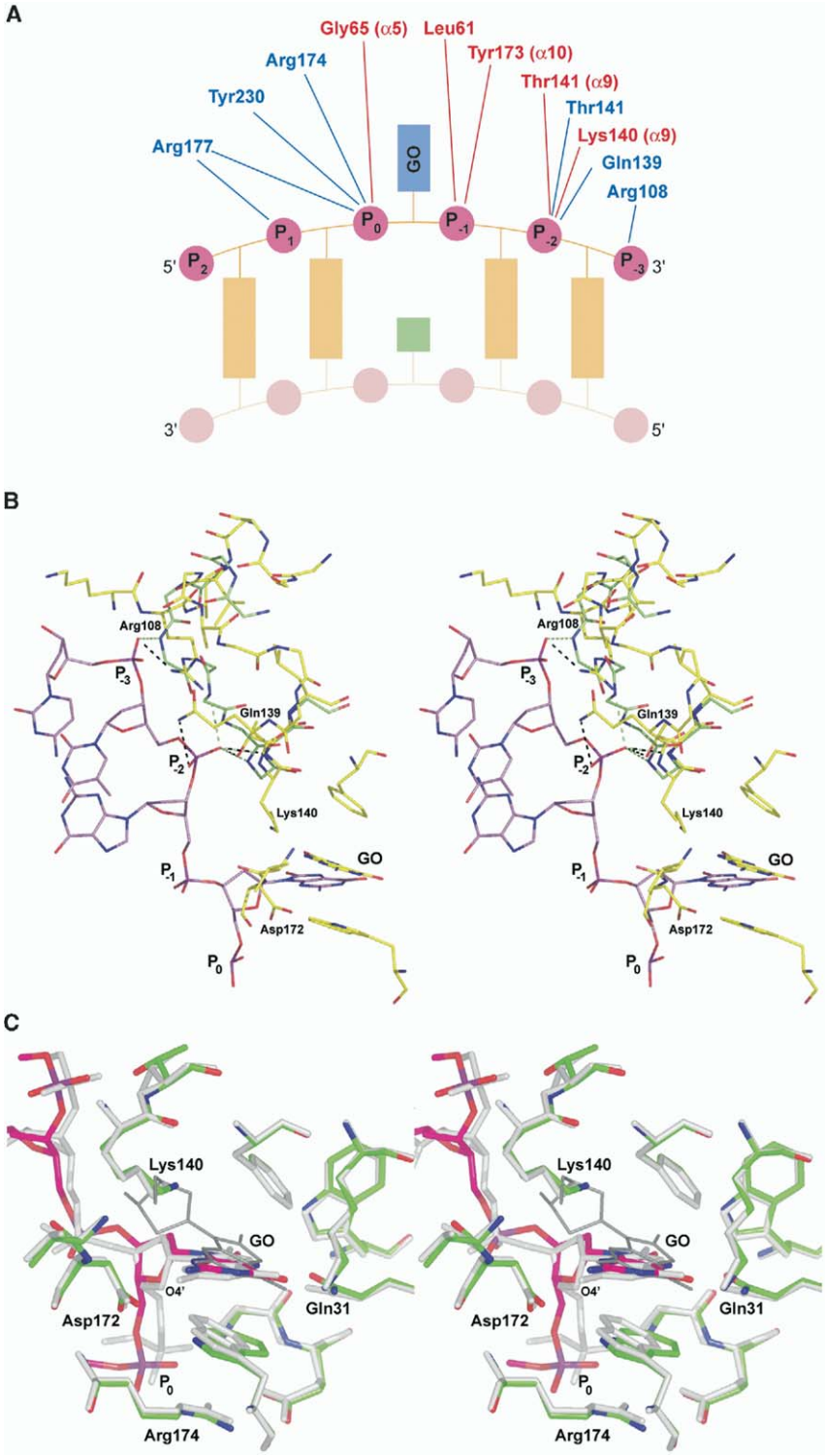

Figure 4. Predicted Protein-DNA Interactions

(A) *Pa*-AGOG main chain (N-terminal helix locations are indicated in brackets) and side chain interactions with the modeled phosphodiester groups (magenta) of the lesion strand are shown schematically in red and blue, respectively. The sugar/phosphate group attached to the aberrant base is marked as P_0 and numbers are increasing in the $3' \rightarrow 5'$ direction. The predicted main chain interactions are analogous to similar conserved interactions in the known complexes of hOGG1, EndoIII, MutY, and AlkA with DNA. The putative interactions of side chains Arg108 and Gln139 are unique for the *Pa*-AGOG family and have no equivalent interactions in the other enzymes. To examine possible DNA-backbone-*Pa*-AGOG interactions, we have first superposed the structurally characterized complexes using the contiguous phosphates (P_{-3} to P_1) and their protein hydrogen bonding partners (only conserved main chain nitrogens). Subsequently, the structurally conserved α-carbon atoms from the HhH-GPD motif of *Pa*-AGOG were superposed onto those of the other four enzymes.

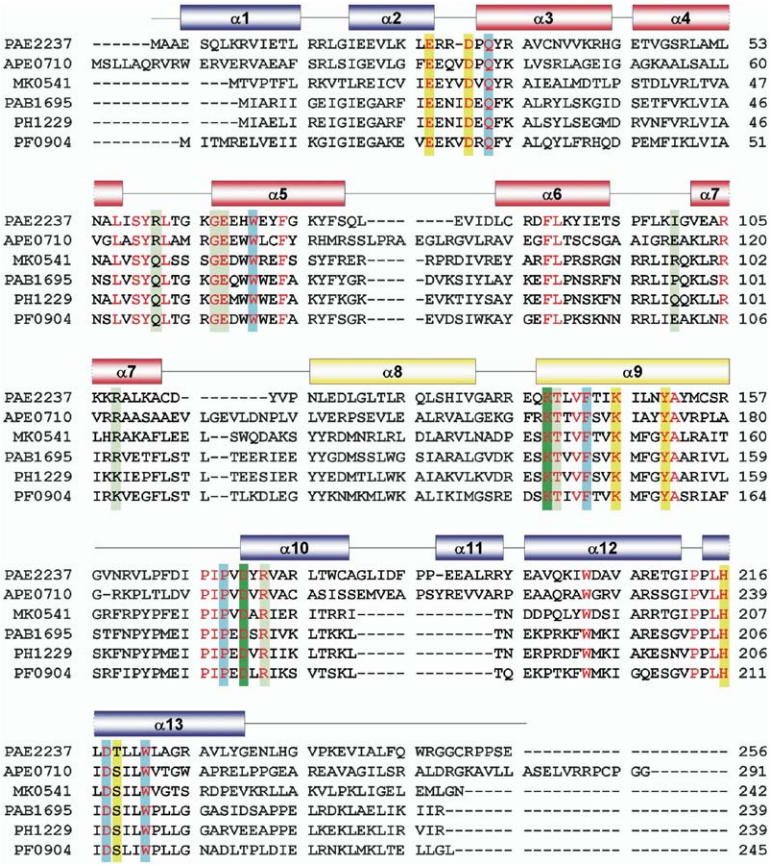

Figure 5. Structure-Based Sequence Alignment of *Pa*-AGOG with Presently Known Homologs

Pa-AGOG secondary structure elements are shown schematically on top of the sequences using the same color codes as in Figure 1A. Sequence identities/similarities between *Pa*-AGOG (PAE2237) and the listed homologs are 31/56 (APE0710, hypothetical protein from *Aeropyrum pernix K1*), 32/54 (MK0541, uncharacterized protein from *Methanopyrus kandleri AV19*), 28/53 (PAB1695, hypothetical protein from *Pyrococcus abyssi GE5*), 29/54 (PH1229, hypothetical protein from *Pyrococcus horikoshii OT3*), and 26/55 (PF0904, hypothetical protein from *Pyrococcus furiosus DSM 3638*) percent. Strictly conserved residues are highlighted in red. Residues involved in defined functions or interactions are boxed as follows: catalysis (green), aberrant base recognition (cyan), DNA phosphate contact and minor groove wedge candidates (lime), internal H-bond network extending from aberrant base (yellow).

ments of *Pa*-AGOG. However, we also predict distinct differences in the way *Pa*-AGOG interacts with the strand carrying the aberrant base, most notably for part of the HhH motif, and these appear characteristic for the new archaeal family.

Archaeal Homologs

The HhH superfamily of DNA repair glycosylases was recently subdivided into six families, which suggested that the genes diverged very early in evolution (Denver et al., 2003). The functional (Sartori et al., 2004) and structural characterization establishes *Pa*-AGOG as the founding member of an additional, seventh family. Structurally, it differs from the other six subgroups of the HhH-GPD DNA glycosylase superfamily by its unique mode of GO recognition and by its distinct HhH hairpin structure resulting in a locally modified DNA contact.

Based on the structure of *Pa*-AGOG, we could modify the sequence-based alignment of the current members of this seventh family in two regions, by moving insertions/deletions to surface segments (Figure 5). One such segment is the loop between helices α5 and α6, the other the connection between α10 and α12, where the short helix α11 appears to be replaced by a much shorter connection in some of the archaeal enzymes. Our structure shows that many of the 32 strictly conserved residues play key roles in catalysis, aberrant base recognition and (putative) DNA interactions. One example of a residue conserved for structural reasons is Arg105 in α7. Its side chain is directed toward the wedge-forming loop between α4 and α5 and its guani-

(B) Stereo view of the predicted *Pa*-AGOG HhH hairpin DNA backbone interactions based on the structural superposition with the hOGG1 DNA complex. *Pa*-AGOG (structure as observed in the complex with GO) is shown in yellow, the DNA and HhH hairpin backbone in violet and green (thin lines), respectively. Observed H bonds (hOGG1 complex) to P_{-3} and P_{-2} are in green, predicted H bonds between P_{-3} and P_{-2} and *Pa*-AGOG in black. The side chains of Gln139 and Lys140 of *Pa*-AGOG were reoriented to avoid steric clashes with the DNA strand (assumed to be very similarly positioned as that of the shown hOGG1 complex). The conformation of the reoriented Lys140 side chain is actually very close to that observed in the uncomplexed enzyme. Note that the space occupied by the main chain of the canonical hairpin (as present in hOGG1) is taken by the side chains of residues Arg108 and Gln139 in *Pa*-AGOG.

(C) Energy-minimized model of *Pa*-AGOG DNA complex (in green and magenta, respectively) in the active site region. Using the Moloc united atom force field (Gerber and Muller, 1995), the starting model, *Pa*-AGOG plus the GO carrying DNA strand (P_{-3} to P_{+1}) from hOGG1 (both shown in gray), was optimized. Overall, only small adjustments are observed. P_0 moves by about 1.5 Å to avoid a close contact with the chain around Lys64 (not shown here, see Figure 3A) and makes a very favorable interaction with Arg174. The plane of the GO base is seen to be tilted by about 20° compared to its orientation in the complex with 8-oxoguanosine (shown in thin lines).

dinium group makes hydrogen bonds to three main chain carbonyls of segments 57–60, which stabilize the main chain conformation around the predicted wedge residue Arg60.

Catalytic Mechanism

Before structural information on substrate-like complexes of bifunctional HhH-GPD-type DNA glycosylases was available, it was generally assumed that the aspartate acts as a general base in the glycosylase reaction by activating the lysine ϵ-nitrogen for nucleophilic attack at the anomeric carbon C1′ with displacement of the base in an S_N2-type reaction. In the hOGG1 and EndoIII complexes, the ϵ-nitrogen of the catalytic lysine is about 6 Å from the nearest aspartate oxygen and substantial side chain reorganization would be required for a direct proton exchange between these residues (Norman et al., 2003). Moreover, the ϵ-nitrogen appears too far from a trajectory required for in-line attack on the glycosidic bond and such an attack appears sterically impossible with the observed sugar conformation (Figure 4B). An alternative S_N2-type reaction mechanism, in-line attack on the C1′-O4′ bond, has been proposed for Fpg and Nei (Zharkov et al., 2003). While much more compatible with the active site stereochemistry observed in the hOGG1 and EndoIII complexes, it equally requires initial proton transfer from the catalytic lysine to the catalytic aspartate.

Based on structural studies and mutagenesis experiments, hOGG1 has recently been suggested to act via a dissociative mechanism (Norman et al., 2003), with the catalytic aspartate (Asp268) being suitably positioned to stabilize an incipient positive charge on $O^{4'}$ (Figure 2B). Earlier, structural studies of the monofunctional AlkA glycosylase (Hollis et al., 2000a) had lead to the conclusion that an S_N1-type mechanism is stereochemically much more plausible and that the conserved aspartate might help stabilize the positive oxocarbenium ion transition state that arises after dissociation of the base. A nucleophilic S_N1-type substitution mechanism has also been proposed for uracil DNA glycosylases (Dinner et al., 2001; Werner and Stivers, 2000). When complexed with 8-oxoguanosine, the carboxylate of the catalytic Asp172 of *Pa*-AGOG, shown to be essential for catalysis by site-directed mutagenesis (Sartori et al., 2004), is fixed by the helix capping interactions in essentially the same conformation as observed in the hOGG1/DNA complex. The striking similarity between the active site geometries of hOGG1 and *Pa*-AGOG as apparent from their structural comparison and our DNA modeling studies (Figure 4C) strongly indicates that *Pa*-AGOG shares a common mechanism with hOGG1 for the glycosylase reaction independent of whether this mechanism is S_N1 or S_N2 like.

The subsequent multistep β-lyase reaction is much more complex and the groups participating in the different deprotonation/protonation steps have not been identified with certainty. Detailed mechanistic studies of the mode of action of hOGG1 (Fromme et al., 2003) implied that the excised base assists in the β-lyase reaction. However, this does not appear to be the case for all HhH-GPD DNA glycosylases (Fromme and Verdine, 2003a). For *Pa*-AGOG, β elimination is indicated to be rather inefficient (Sartori et al., 2004) and product assistance as observed for hOGG1 is unlikely to occur.

Conclusions

Pa-AGOG represents, to our knowledge, the first functionally and structurally characterized member of a new HhH-GPD DNA glycosylase family. Structure analysis showed that these new enzymes share with the other families the overall fold and the general active site architecture, but that they differ most significantly by the trace of the hairpin loop of the HhH motif. The location and shape of the GO recognition pocket is very similar to that observed in hOGG1, but the detailed recognition mode is unique. Indeed, the two enzymes do not share a single identical side chain involved in direct hydrogen bonding or stacking interactions with the aberrant base. Unlike hOGG1 and MutM, *Pa*-AGOG contacts the 8-oxo function via a direct hydrogen bond. Binding of the aberrant base is coupled to concerted conformational changes of the free enzyme, which also appear to be required to generate an interaction surface for binding substrate DNA. The high conservation of all residues implicated in the interaction with the GO substrate suggests that all members of the seventh family of HhH-GPD-type DNA glycosylases will be GO specific.

Modeling studies based on the conserved binding mode of DNA to the HhH-GPD-type DNA glycosylases suggest that *Pa*-AGOG binds DNA in a similar strongly kinked conformation and in an overall similar manner. A distinct difference concerns its HhH hairpin structure which cannot form some of the characteristic HhH motif-DNA phosphate interactions. We predict that an arginine side chain acts as a substitute to stabilize the same conformation of bound DNA. Clearly, the structural analysis of a complex with substrate-like DNA will be needed to verify the proposed detailed interactions and to resolve the complete protein DNA interface.

Experimental Procedures

Protein Purification and Crystallization

Native *Pa*-AGOG was expressed in *E. coli* and purified in three steps using nickel affinity chromatography followed by ion exchange and gel filtration (Sartori et al., 2004). The protein contains an additional 19 amino acids at its N terminus which include the hexahistidine tag ($GSSH_6SSGLVPRGSH$). The selenomethionyl variant of *Pa*-AGOG was expressed following standard methods (Hendrickson et al., 1990), and purified under highly reducing conditions using the same purification scheme. The yields of purified protein were 8 and 7 mg per liter of culture for native and selenomethionine-labeled protein, respectively.

Crystals of native *Pa*-AGOG were grown in sitting drops by mixing 3 µl of 50 mg/ml protein solution: 20 mM HEPES-NaOH (pH 7.5), 100 mM NaCl, 1 mM DTT, and 3 µl of reservoir solution: 0.1 M HEPES-NaOH (7.5) and 15% PEG 4000. Crystals of Se-Met derived *Pa*-AGOG were grown in sitting drops by mixing 2.5 µl of *Pa*-AGOG at 35 mg/ml protein solution: 20 mM HEPES-NaOH (pH 7.5), 100 mM NaCl, 1 mM DTT, and 2.5 µl of reservoir solution: 0.1 M Tris-HCl (pH 8.0), 20% PEG 4000. In both cases, crystals grew to their full size in 3 to 4 weeks at 22°C. Crystals of *Pa*-AGOG complexed with 8-oxoguanosine were prepared by overnight soaking of native crystals in reservoir solution containing 10 mM 8-oxoguanosine (2′-deoxy-7,8-dihydro-8-oxoguanosine, Berry and Associates, Inc.) and 5% DMSO.

Data Collection and Structure Determination

Data sets of crystals of the native and selenomethionine-labeled enzymes were collected to 1.0 and 1.2 Å resolution, respectively at 100 K on the X06SA beamline at the Swiss Light Source (SLS, PSI Villigen, Switzerland) using a MAR CCD detector. The data of the *Pa*-AGOG/8-oxoguanosine complex were collected to 1.7 Å resolution using a MAR 345 image plate and CuK_α radiation generated by an Enraf-Nonius FR591 rotating anode X-ray generator equipped with Osmic

mirrors. The data were processed using XDS and XSCALE (Kabsch, 1993) and statistical figures are given in Table 1.

Matthews coefficient calculations suggested the presence of one molecule in the asymmetric unit and a solvent content of 37%. Using SHELXD (Schneider and Sheldrick, 2002), two of the three selenium atoms expected in the asymmetric unit were located using the anomalous differences of the λ_{peak} data set. Using SHARP, the two sites were refined and initial MAD phases to 1.2 Å resolution were generated (De La Fortelle and Bricogne, 1997). These initial phases were further improved by solvent flattening using Solomon, which produced an electron density of excellent quality. Automatic model building using ARP/wARP 6.0 (Morris et al., 2002) produced a correct main chain model, but failed to correctly fit a significant fraction of the side chains. These side chains were manually fitted into the electron density map using the computer graphics program Moloc (Gerber and Muller, 1995). This model was refined against the 50–1.0 Å resolution native data set using Refmac5.2 (CCP4, 1994; Murshudov et al., 1997) after an initial rigid body refinement. The model was completed by the addition of bound water molecules and further improved through iterative rounds of model building and refinement. Additional refinements using TLS parameters, individual isotropic B factors and individual anisotropic B factors were very effective to improve the quality of the fit, which reduced the *R* factors by more than 5% from the initial round of refinement. For 11 residues (E20, C36, T45, S48, S77, E80, S130, T141, C155, L221, and W222) there was clear indication of multiple side chain conformations and they were thus included as two or three (T141) equally populated conformations in the model. In addition, the peptide group between G249 and G250 was modeled in two conformations. The final model converged at an *R* factor of 0.16 (R_{free} = 0.18) with very good stereochemistry (Table 1). Quality checks of the final model were done using PROCHECK (Laskowski et al., 1993), WHAT IF (Vriend, 1990), and Rotamer (CCP4, 1994).

The structure determination of the complex with 8-oxoguanosine was initiated by rigid body refinement of the native model, which was followed by restrained refinement against the data of the complex. A subsequent σ_a-weighted difference electron density map contoured at 3σ showed strong positive density for the nucleoside in the binding pocket. In addition, a second binding site was observed across the crystallographic 2-fold axis, which was included in the model with half occupancy. The model of the *Pa*-AGOG complex was subjected to the same type of refinement as described for the free enzyme except for refining individual isotropic B-factors. Double conformations were included for the side chains of residues L17, T45, and T141, and for the peptide link G249/G250. For both structures, no residues were observed in the generous and disallowed regions of the Ramachandran plot. Final statistical values of structure solution and refinement are given in Table 1.

The native model comprises residues 2–254 (full-length sequence 1–256, see Figure 5) and residues 3–254 in the case of the nucleoside complex. In both structures, the additional 19 N-terminal residues originating from the tag are disordered.

Figures were prepared with PyMOL (http://www.pymol.org [DeLano, 2002]).

Acknowledgments

Data were collected at beamline X06SA of the Swiss Light Source, Paul Scherrer Institut, Villigen, Switzerland and we gratefully acknowledge the excellent support of the beamline staff. The project was supported by the Swiss National Science Foundation, Project 3100-062027. The generous support of the UBS Stiftung to A.A.S. is also acknowledged.

Received: September 14, 2004
Revised: October 25, 2004
Accepted: October 28, 2004
Published: January 11, 2005

References

A-Lien, L., Xianghong, L., Yesong, G., Patrick, M.W., and Dau-Yin, C. (2001). Repair of oxidative DNA damage: mechanisms and function. Cell Biochem. Biophys. *35*, 141–170.

Bjoras, M., Luna, L., Johnsen, B., Hoff, E., Haug, T., Rognes, T., and Seeberg, E. (1997). Opposite base-dependent reactions of a human base excision repair enzyme on DNA containing 7,8-dihydro-8-oxoguanine and abasic sites. EMBO J. *16*, 6314–6322.

Bjoras, M., Seeberg, E., Luna, L., Pearl, L.H., and Barrett, T.E. (2002). Reciprocal "flipping" underlies substrate recognition and catalytic activation by the human 8-oxo-guanine DNA glycosylase. J. Mol. Biol. *317*, 171–177.

Boiteux, S., O'Connor, T.R., and Laval, J. (1987). Formamidopyrimidine-DNA glycosylase of Escherichia coli: cloning and sequencing of the fpg structural gene and overproduction of the protein. EMBO J. *6*, 3177–3183.

Bruner, S.D., Norman, D.P., and Verdine, G.L. (2000). Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. Nature *403*, 859–866.

CCP4 (Collaborative Computational Project, Number 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

De La Fortelle, E., and Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. In Methods in Enzymology, C.W. Carter and R.M. Sweet, eds. (New York: Academic Press), pp. 472–494.

DeLano, W.L. (2002). The Pymol Molecular Graphics System (San Carlos,CA: DeLano Scientific).

Denver, D.R., Swenson, S.L., and Lynch, M. (2003). An evolutionary analysis of the helix-hairpin-helix superfamily of DNA repair glycosylases. Mol. Biol. Evol. *20*, 1603–1611.

Dinner, A.R., Blackburn, G.M., and Karplus, M. (2001). Uracil-DNA glycosylase acts by substrate autocatalysis. Nature *413*, 752–755.

Doherty, A.J., Serpell, L.C., and Ponting, C.P. (1996). The helix-hairpin-helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA. Nucleic Acids Res. *24*, 2488–2497.

Drohat, A.C., Kwon, K., Krosky, D.J., and Stivers, J.T. (2002). 3-Methyladenine DNA glycosylase I is an unexpected helix-hairpin-helix superfamily member. Nat. Struct. Biol. *9*, 659–664.

Fromme, J.C., and Verdine, G.L. (2003a). Structure of a trapped endonuclease III-DNA covalent intermediate. EMBO J. *22*, 3461–3471.

Fromme, J.C., and Verdine, G.L. (2003b). DNA lesion recognition by the bacterial repair enzyme MutM. J. Biol. Chem. *278*, 51543–51548.

Fromme, J.C., Bruner, S.D., Yang, W., Karplus, M., and Verdine, G.L. (2003). Product-assisted catalysis in base-excision DNA repair. Nat. Struct. Biol. *10*, 204–211.

Fromme, J.C., Banerjee, A., Huang, S.J., and Verdine, G.L. (2004a). Structural basis for removal of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase. Nature *427*, 652–656.

Fromme, J.C., Banerjee, A., and Verdine, G.L. (2004b). DNA glycosylase recognition and catalysis. Curr. Opin. Struct. Biol. *14*, 43–49.

Gerber, P.R., and Muller, K. (1995). MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. J. Comput. Aided Mol. Des. *9*, 251–268.

Girard, P.M., D'Ham, C., Cadet, J., and Boiteux, S. (1998). Opposite base-dependent excision of 7,8-dihydro-8-oxoadenine by the Ogg1 protein of Saccharomyces cerevisiae. Carcinogenesis *19*, 1299–1305.

Grollman, A.P., and Moriya, M. (1993). Mutagenesis by 8-oxoguanine: an enemy within. Trends Genet. *9*, 246–249.

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

Hazra, T.K., Hill, J.W., Izumi, T., and Mitra, S. (2001). Multiple DNA glycosylases for repair of 8-oxoguanine and their potential in vivo functions. Prog. Nucleic Acid Res. Mol. Biol. *68*, 193–205.

Hendrickson, W.A., Horton, J.R., and LeMaster, D.M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. EMBO J. *9*, 1665–1672.

Hollis, T., Ichikawa, Y., and Ellenberger, T. (2000a). DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, Escherichia coli AlkA. EMBO J. *19*, 758–766.

Hollis, T., Lau, A., and Ellenberger, T. (2000b). Structural studies of human alkyladenine glycosylase and E. coli 3-methyladenine glycosylase. Mutat. Res. *460*, 201–210.

Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Crystallogr. *26*, 795–800.

Laskowski, R., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

McCullough, A.K., Dodson, M.L., and Lloyd, R.S. (1999). Initiation of base excision repair: glycosylase mechanisms and structures. Annu. Rev. Biochem. *68*, 255–285.

McCullough, A.K., Sanchez, A., Dodson, M.L., Marapaka, P., Taylor, J.S., and Lloyd, R.S. (2001). The reaction mechanism of DNA glycosylase/AP lyases at abasic sites. Biochemistry *40*, 561–568.

Michaels, M.L., and Miller, J.H. (1992). The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7,8-dihydro-8-oxoguanine). J. Bacteriol. *174*, 6321–6325.

Mol, C.D., Arvai, A.S., Begley, T.J., Cunningham, R.P., and Tainer, J.A. (2002). Structure and activity of a thermostable thymine-DNA glycosylase: evidence for base twisting to remove mismatched normal DNA bases. J. Mol. Biol. *315*, 373–384.

Morris, R.J., Perrakis, A., and Lamzin, V.S. (2002). ARP/wARP's model-building algorithms. I. The main chain. Acta Crystallogr. D Biol. Crystallogr. *58*, 968–975.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Nash, H.M., Bruner, S.D., Scharer, O.D., Kawate, T., Addona, T.A., Spooner, E., Lane, W.S., and Verdine, G.L. (1996). Cloning of a yeast 8-oxoguanine DNA glycosylase reveals the existence of a base-excision DNA-repair protein superfamily. Curr. Biol. *6*, 968–980.

Norman, D.P., Chung, S.J., and Verdine, G.L. (2003). Structural and biochemical exploration of a critical amino acid in human 8-oxoguanine glycosylase. Biochemistry *42*, 1564–1572.

Pearson, C.G., Shikazono, N., Thacker, J., and O'Neill, P. (2004). Enhanced mutagenic potential of 8-oxo-7,8-dihydroguanine when present within a clustered DNA damage site. Nucleic Acids Res. *32*, 263–270.

Radicella, J.P., Dherin, C., Desmaze, C., Fox, M.S., and Boiteux, S. (1997). Cloning and characterization of hOGG1, a human homolog of the OGG1 gene of Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA *94*, 8010–8015.

Roberts, R.J., and Cheng, X. (1998). Base flipping. Annu. Rev. Biochem. *67*, 181–198.

Sartori, A.A., Schar, P., Fitz-Gibbon, S., Miller, J.H., and Jiricny, J. (2001). Biochemical characterization of uracil processing activities in the hyperthermophilic archaeon Pyrobaculum aerophilum. J. Biol. Chem. *276*, 29979–29986.

Sartori, A.A., Fitz-Gibbon, S., Yang, H., Miller, J.H., and Jiricny, J. (2002). A novel uracil-DNA glycosylase with broad substrate specificity and an unusual active site. EMBO J. *21*, 3182–3191.

Sartori, A.A., Lingaraju, G.M., Hunziker, P., Winkler, F.K., and Jiricny, J. (2004). Pa-AGOG, the founding member of a new family of archaeal 8-oxoguanine DNA-glycolases. Nucleic Acid Res. *32*, 6531–6539.

Scharer, O.D., and Jiricny, J. (2001). Recent progress in the biology, chemistry and structural biology of DNA glycosylases. Bioessays *23*, 270–281.

Schneider, T.R., and Sheldrick, G.M. (2002). Substructure solution with SHELXD. Acta Crystallogr. D Biol. Crystallogr. *58*, 1772–1779.

Shimizu, M., Gruz, P., Kamiya, H., Kim, S.R., Pisani, F.M., Masutani, C., Kanke, Y., Harashima, H., Hanaoka, F., and Nohmi, T. (2003). Erroneous incorporation of oxidized DNA precursors by Y-family DNA polymerases. EMBO Rep. *4*, 269–273.

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

Uesugi, S., and Ikehara, M. (1977). Carbon-13 magnetic resonance spectra of 8-substituted purine nucleosides. Characteristic shifts for the syn conformation. J. Am. Chem. Soc. *99*, 3250–3253.

van der Kemp, P.A., Thomas, D., Barbey, R., de Oliveira, R., and Boiteux, S. (1996). Cloning and expression in Escherichia coli of the OGG1 gene of Saccharomyces cerevisiae, which codes for a DNA glycosylase that excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine. Proc. Natl. Acad. Sci. USA *93*, 5197–5202.

Volkl, P., Huber, R., Drobner, E., Rachel, R., Burggraf, S., Trincone, A., and Stetter, K.O. (1993). Pyrobaculum aerophilum sp. nov., a novel nitrate-reducing hyperthermophilic archaeum. Appl. Environ. Microbiol. *59*, 2918–2926.

Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. *8*, 52–56.

Wallace, S.S., Bandaru, V., Kathe, S.D., and Bond, J.P. (2003). The enigma of endonuclease VIII. DNA Repair (Amst.) *2*, 441–453.

Werner, R.M., and Stivers, J.T. (2000). Kinetic isotope effect studies of the reaction catalyzed by uracil DNA glycosylase: evidence for an oxocarbenium ion-uracil anion intermediate. Biochemistry *39*, 14054–14064.

Yang, H., Fitz-Gibbon, S., Marcotte, E.M., Tai, J.H., Hyman, E.C., and Miller, J.H. (2000). Characterization of a thermostable DNA glycosylase specific for U/G and T/G mismatches from the hyperthermophilic archaeon Pyrobaculum aerophilum. J. Bacteriol. *182*, 1272–1279.

Zharkov, D.O., Rosenquist, T.A., Gerchman, S.E., and Grollman, A.P. (2000). Substrate specificity and reaction mechanism of murine 8-oxoguanine-DNA glycosylase. J. Biol. Chem. *275*, 28607–28617.

Zharkov, D.O., Shoham, G., and Grollman, A.P. (2003). Structural characterization of the Fpg family of DNA glycosylases. DNA Repair (Amst.) *2*, 839–862.

Accession Numbers

The coordinates and structure factors of free and complexed structure were deposited with the Protein Data Bank (http://www.rcsb.org/pdb/) with accession codes 1XQO and 1XQP respectively.