

Structure-function relationships governing activity and stability of a DNA alkylation damage repair thermostable protein

Giuseppe Perugino^{1,†}, Riccardo Miggiano^{2,†}, Mario Serpe¹, Antonella Vettone¹, Anna Valenti¹, Samarпита Lahiri², Franca Rossi², Mosè Rossi¹, Menico Rizzi² and Maria Ciaramella^{1,*}

¹Institute of Biosciences and Bioresources, National Research Council of Italy, Via P. Castellino 111, 80125 Naples, Italy and ²DiSF-Dipartimento di Scienze del Farmaco, University of Piemonte Orientale 'A. Avogadro', Via Bovio 6, 28100 Novara, Italy

Received May 28, 2015; Revised July 09, 2015; Accepted July 18, 2015

ABSTRACT

Alkylated DNA-protein alkyltransferases repair alkylated DNA bases, which are among the most common DNA lesions, and are evolutionary conserved, from prokaryotes to higher eukaryotes. The human ortholog, hAGT, is involved in resistance to alkylating chemotherapy drugs. We report here on the alkylated DNA-protein alkyltransferase, SsOGT, from an archaeal species living at high temperature, a condition that enhances the harmful effect of DNA alkylation. The exceptionally high stability of SsOGT gave us the unique opportunity to perform structural and biochemical analysis of a protein of this class in its post-reaction form. This analysis, along with those performed on SsOGT in its ligand-free and DNA-bound forms, provides insights in the structure-function relationships of the protein before, during and after DNA repair, suggesting a molecular basis for DNA recognition, catalytic activity and protein post-reaction fate, and giving hints on the mechanism of alkylation-induced inactivation of this class of proteins.

INTRODUCTION

Alkylated DNA-protein alkyltransferases (AGTs, MGMTs or OGTs, EC 2.1.1.63) are conserved proteins that repair alkylation damage in DNA, mainly at position *O*⁶ of guanines. They use a peculiar single-step mechanism in which the direct repair of the alkylated base is coupled with irreversible alkylation of the catalytic cysteine in the protein active site. The trans-alkylated protein is permanently inactivated and prone to degradation both *in vivo* and *in vitro* (1–

4). Most knowledge these proteins comes from classic studies on Ada-C and OGT of *Escherichia coli*, as well as the human hAGT (1–4). In light of the observation that hAGT over-expression in tumor cells is frequently associated with resistance to alkylating agents, hAGT has received attention as a potential target for the development of treatments to be integrated into current chemotherapy protocols based on such drugs (5).

The current understanding of the AGTs molecular mechanism responsible for alkylated DNA recognition and repair is mainly based on the 3D structures of hAGT and its complex with double-stranded (ds) DNA molecules (6–9). The protein contacts the DNA minor groove through the helix-turn-helix (HTH) motif of its C-terminal domain. By adopting an extra-helical conformation stabilized by an 'arginine finger' (R128) of the HTH, the damaged base is deeply inserted in the protein active site, and the alkyl moiety is then transferred to the catalytic cysteine (C145). Data on structure and properties of alkylated AGTs are limited, because alkylation greatly destabilizes their folding; as for 3D structures, the only available are those of the methylated (hAGT^m) and benzylated (hAGT^b), which were obtained from native hAGT crystals flash frozen upon soaking in solutions containing *O*⁶-methyl- (*O*⁶-MG) and *O*⁶-benzyl-guanine (*O*⁶-BG), respectively (7). The hAGT^m and hAGT^b structures showed that C145 alkylation induces subtle conformational changes at the active site; however, these structures might not reflect the physiological conformation of the alkylated hAGT, since in the crystalline state the protein could not accurately display the conformation adopted in solution. Indeed, it is likely that protein movements are restricted in crystals, and larger rearrangements may lead to crystal decay (7).

Alkylation-induced instability of AGTs is interesting from both a biological and mechanistic point of view. AGTs

*To whom correspondence should be addressed. Tel: +390816132247; Fax: +39081632646; Email: maria.ciaramella@ibbr.cnr.it

†These authors contributed equally to the paper as first authors.

degradation is an important process in the organisms' response to alkylation damage; in human cells, hAGT physically associates and undergoes repair-mediated degradation with the BRCA2/FancD1 protein (10). Germline mutations of *Brca2* are associated with cancer prone syndromes and Fanconi anemia, and absence of a functional BRCA2 protein induces increased cell sensitivity to DNA crosslinking agents (10); thus, hAGT activity and degradation might also affect other DNA repair pathways. The relation between active site modification and protein unfolding/degradation has been difficult to study due to the instability of alkylated AGT forms. *In vivo*, conformational changes might expose residues, which are target for ubiquitination, thus triggering the protein degradation; on the other hand, alkylated hAGT is also intrinsically unstable *in vitro* (7,11). It has been suggested that alkylation-induced conformation modifications induce distortion of the DNA binding surface, facilitating the dissociation of alkylated hAGT from DNA, while destabilizing the protein native fold (7,12,13). Recently, two glycine residues (G131 and G132) were proposed to be implicated in the balance between stability and instability of hAGT, stabilizing the protein in the native form and triggering its destabilization upon alkylation, through a still unclear mechanism (14).

AGTs are present in organisms from the three living domains (Eucarya, Bacteria, Archaea). In thermophilic bacteria and archaea, living at >80°C, alkylation damage is a serious harm since alkylated bases are unstable at high temperature and induce DNA ruptures (15,16). We have previously reported on OGT from the archaeon *Sulfolobus solfataricus* (*Ss*OGT), a protein with outstanding stability at high temperature, which, upon alkylation, becomes unstable and undergoes degradation *in vivo*, suggesting that it follows the same fate as hAGT (17,18).

We present here a biochemical, structural and mutational analysis of *Ss*OGT. The crystal structure of the ligand-free protein and its complex with an *O*⁶-MG containing dsDNA revealed overall similarities with the corresponding structures of hAGT, but also peculiarities, which were found to have functional significance. Moreover, in contrast to the corresponding hAGT^m, the methylated form of *Ss*OGT (*Ss*OGT^m) was soluble and relatively stable, thus allowing in-deep analysis of the protein in its post-reaction form. Structural and biochemical analysis of *Ss*OGT^m, as well as of a mutant mimicking the presence of a bulkier adduct in the active site (C119L), suggested a possible mechanism of alkylation-induced *Ss*OGT unfolding and degradation. Based on our data, we suggest a general model for the mechanism of post-reaction AGTs destabilization.

MATERIALS AND METHODS

DNA mutagenesis and protein purification

Site-directed mutants were obtained by using the oligonucleotide 'mut' and 'rev' pairs (Supplementary Table S1) and the GeneTailorTM Site-Directed Mutagenesis System (Invitrogen); the template was the *S. solfataricus* ogt gene cloned in the pQE31TM vector (17). N-terminally His-tagged proteins were expressed in the *E. coli* ABLE-C strain and purified as described (17).

Fluorescent assays for *O*⁶-alkyl-guanine alkyltransferase activity

*Ss*OGT activity was determined by a fluorescent assay, by using the SNAP-Vista GreenTM (BG-VG) substrate under the standard reaction conditions as previously reported (17,19). After SDS-PAGE, fluorescent protein bands were visualized by gel imaging using the VersaDoc 4000TM system (Bio-Rad). The fluorescence intensity of each band was corrected for the amount of protein loaded by Coomassie Blue staining. Determination of pseudo-first-order reaction rate values was carried out under standard conditions, taking protein aliquots at different times. Second-order rate constants were then obtained by dividing values by the substrate concentration (17,20).

Direct DNA repair activity

The efficiency of repair of ds oligonucleotides containing a single *O*⁶-MG was determined by incubation of fixed amounts of protein (5 μM) for different timespans at 25°C in the presence of increasing concentrations of the ds-Fwd^{m26} oligonucleotide. Then, 10 μM of BG-VG was added, incubation was prolonged for 2 h and samples analyzed by SDS-PAGE. Gel imaging and Coomassie correction was performed as described above.

Competition assay with non-fluorescent *O*⁶-methyl-guanine containing DNA

The efficiency of repair of ds oligonucleotides containing a single *O*⁶-MG was determined by using the half maximal inhibitory concentration (IC₅₀) method, where IC₅₀ is the concentration of methylated DNA needed to reduce by 50% the fluorescence intensity of the *Ss*OGT band in reactions containing fixed concentrations of the fluorescent BG-VG molecule (Supplementary Figure S1A). Reactions were incubated at fixed temperatures with increasing concentrations (0–10 μM) of the appropriate methylated oligonucleotide. Gel imaging and Coomassie correction were performed as described above; corrected data of fluorescence intensity were fitted with the IC₅₀ equation. For competitive inhibition, the affinity of the inhibitor (K_{DNA}) is related to the IC₅₀ value by the adapted Cheng-Prusoff Equation (1) (21):

$$K_{DNA} = IC_{50} / [1 + ([VG]/K_{VG})] \quad (1)$$

where K_{DNA} is the binding affinity of the inhibitor, IC₅₀ is the functional strength of the inhibitor, [VG] and K_{VG} are the VG concentration and the concentration of the BG-VG substrate at which enzyme activity is half maximal, respectively (17).

Electrophoretic mobility shift assay

A tetramethylrhodamine labeled dsDNA probe was prepared by annealing the oligonucleotides A⁺ and D⁻, as reported (Supplementary Table S1; 17). Different amounts of protein (0.0–25.0 μM) were incubated at 37°C for 10 min with the probe (0.2 μM), in a total volume of 10.0 μl, as described (17). After loading the sample on an 8% polyacrylamide native gel run in 1X TBE (90.0 mM Tris-HCl,

90.0 mM boric acid, 2.0 mM EDTA, pH 8.3), signals were measured by gel imaging, using a green LED/605 band-pass filter as excitation/emission parameters, respectively. For DNA binding assay with unlabeled dsDNA oligonucleotides containing *O*⁶-MG, samples were prepared as described above, except that, after electrophoresis, gels were stained with ethidium bromide (10 mg/ml) for 15 min at RT.

Protein stability assays

Thermal stability was determined by two different methods. (i) temperature-induced aggregation: 30 μ l aliquots of 5.0 μ M (0.1 mg/ml) of each protein were incubated for 20 min at different temperatures in PBS 1X buffer, centrifuged for 20 min at 16 000 \times g, and 20 μ l of supernatant were immediately loaded on 15% SDS-PAGE. The relative intensity of Coomassie-stained SsOGT bands were plotted as a function of temperature, considering as 100% the intensity of each protein band incubated at 25°C. Reported data are the mean \pm SD of three independent experiments; (ii) Differential Scan Fluorimetry (adapting the protocol described in 22): each protein (25 μ M; 0.5 mg/ml) was incubated in PBS 1X buffer and SYPRO Orange dye 1X, in a total volume of 30 μ l. Samples were heated from 20 to 95°C in a Real-Time Light Cycler (Bio-Rad, Milan, Italy). Thermal stability scans were performed at 0.2°C/min (5 min/cycle with an increase of 1°C/cycle). Data were normalized to the maximum fluorescence value within each scan. Relative fluorescence intensities were plotted as a function of temperature; the obtained sigmoidal curve describes a two-state transition, where the T_m value represents the inflection point of the transition curve, as described by the Boltzmann Equation (2),

$$y = LL + \frac{(UL - LL)}{1 + \exp\left(\frac{T_m - x}{a}\right)} \quad (2)$$

where the values of minimum and maximum fluorescent intensities are LL and UL, respectively, and a represents the slope of the curve within T_m (22). Data are the mean \pm SD of three independent experiments.

Data analysis

Corrected data were fitted to appropriate equations by using GraFit 5.0 Data Analysis Software (Erithacus Software) or Prism Software Package (GraphPad Software) (23).

Preparation of SsOGT^m

20 ml of 10 μ M of the SsOGT wt protein were incubated at 25°C for 24 h in the presence of *O*⁶-MG (10 mM, protein:inhibitor ratio 1:1000). To test the efficiency of methylation, an aliquot containing 50 pmols (1.0 μ g) of SsOGT from the methylation reaction was incubated with an equimolar amount of BG-VG at 70°C for 30 min and subsequently subjected to SDS-PAGE and fluorescence imaging: no fluorescent signal was detected, confirming the complete methylation of the protein. Unreacted inhibitor was removed using a HiTrap DesaltingTM (GE, Healthcare) column pre-equilibrated in PBS 1X.

Preparation of SsOGT-C119A::dsDNA^m complex

To obtain the modified dsDNA molecule used in co-crystallization experiment (dsDNA^m), the *O*⁶-methylguanine-containing oligonucleotide (5'-GCCATG[*O*⁶-MG]CTAGTA-3', Primm, Milan, Italy) was annealed to the complementary oligonucleotide 5'-TACTAGCCATGGC-3' (Eurofins MWG Operon). The resulting sample was mixed with the C119A protein solution (7 mg/ml in 20 mM phosphate buffer, pH 7.3 and 150 mM NaCl) at a protein:DNA molar ratio of 1:1.2, and incubated 1 h at room temperature before crystallization trials.

Crystallization and data collection

Initial crystallization conditions for wild type, C119L and SsOGT^m proteins and for the C119A::dsDNA^m complex were identified by means of a robot-assisted (Oryx4; Douglas Instruments), sitting-drop-based sparse-matrix strategy using screen kits from Hampton Research and Qiagen. Wild-type SsOGT crystals grew at 4°C in 4 μ l drops obtained by mixing equal volumes of 9 mg/ml purified protein solution and precipitant 0.35 M potassium nitrate and 1.6 M ammonium sulfate in a final droplet volume of 1 μ l. A single crystal was cryo-protected in precipitant solution containing 15% glycerol, mounted in a cryo-loop, and flash-frozen in liquid nitrogen at 100 K for further X-ray diffraction analysis. Single crystal diffracted at 1.85 Å of resolution at the ID23 synchrotron radiation ($\lambda = 0.87$ Å) (European Synchrotron Radiation Facility [ESRF], Grenoble, France). The diffraction data were indexed with XDS program (24), whose indexing score assigned crystal to the trigonal space-group R3 with the cell dimension $a = 94.72$ Å $b = 94.72$ Å $c = 76.70$ Å. Crystals of C119L mutant (8 mg/ml) were identified in the initial crystallization trials in the condition containing 4 M sodium formate as reservoir solution in a protein to reservoir ratio of 1:1, in a final droplet volume of 1 μ l. Single crystal suitable for X-ray diffraction was manipulated as previously describe for wild-type protein and it diffracted at 2.6 Å of resolution at 100K at the ID23 beam line ($\lambda = 1.89$ Å) (ESRF, Grenoble, France). Indexing process with XDS program assigned the crystal to the cubic I432 space-group with the dimension $a = b = c = 140.56$ Å. SsOGT^m, obtained as described above, was concentrated to 7 mg/ml. Well-shaped single crystals grew in a 1 μ l droplet containing equal volumes of protein and reservoir solution (0.1 M Bis-Tris pH 5.5, 0.1 M ammonium acetate, 17% w/v PEG 17000). Data collection was performed at BM30 synchrotron radiation ($\lambda = 0.979$ Å) (ESRF, Grenoble, France), under cryogenic condition using 15% glycerol as cryo protectant in reservoir solution. The best crystal diffracted at 2.8 Å of resolution and it was assigned to the orthorhombic space group P212121 with the cell dimensions $a = 48.49$ Å $b = 50.22$ Å $c = 142.02$ Å, containing 2 molecule for asymmetric unit, with a Matthews parameter and a solvent fraction of 2.28 Å³ Da⁻¹ and 45.98%, respectively. C119A::*O*⁶-MG-DNA single crystals grew in approximately 1 week in a 1 μ l droplet of equal volume of protein-DNA complex solution and reservoir solution composed by 0.2 M Na/K tartrate 20% (w/v) PEG 3350. Diffraction experiments were conducted at 100 K using synchrotron radiation ($\lambda = 0.972$

Å) at the ID29 (ESRF, Grenoble, France). Data collection was performed up to 2.7 Å of resolution. Indexing score with XDS assigned crystal to the orthorhombic space group P2₁2₁2₁ with the cell dimensions a = 41.76 Å b = 65.88 Å c = 97.52 Å. For all data sets described above, further data manipulations were carried out using COMBAT and SCALA from the CCP4 program suite (25). The data statistics of the solved structures are summarized in Table 1.

Structure determination, model building and refinement

The initial phases for wild-type SsOGT, C119L mutant and SsOGT^m structures were generated by molecular replacement with the program PHASER (26) of the PHENIX software suite (27) using *Sulfolobus tokodaii* OGT structure (PDB ID code:1WRJ) as the search model. The starting search model for C119A::O⁶-MeG-DNA complex structure consisted of *Sulfolobus tokodaii* OGT structure and the double stranded, methylated DNA molecule as crystallized in complex with hAGT (PDB ID code: 1T38) for the protein and DNA component, respectively. Initial model building was performed using AUTOBUILD of the PHENIX suite (28) followed by manual model building with the program COOT (29). Solvent molecules were added by ARP/wARP SOLVENT program from CCP4 program suite followed by structure refinement that was done with PHENIX (27). In the refined SsOGT^m structure we noticed a large difference between R_{work} and R_{free} values. We deeply investigated all possible space groups suggested by the XDS analysis, and acceptable statistics could only be obtained by processing data in the space group P2₁2₁2₁. Similarly an unambiguous solution of molecular replacement could only be found in P2₁2₁2₁ space group. All figures illustrating structural analyses were generated with PyMol (<http://www.pymol.org/>) (30).

Protein structure accession numbers

The atomic coordinates and structure factors of the SsOGT wild type, C119L mutant, SsOGT^m and C119A::dsDNA^m have been deposited in the Protein Data Bank (<http://www.rcsb.org>) under the accession codes 4ZYE, 4ZYH, 4ZYG, and 4ZYD, respectively

RESULTS

Structure-function analysis of free SsOGT

The crystal structure of SsOGT was solved at 1.8 Å resolution (Table 1). As observed for all AGTs structures present in the Protein Data Base (6–9,19,31–34), SsOGT folds in two domains joined by a long connecting loop (a.a. 54–69) (Figure 1A). The N-terminal domain (a.a. 1–53) consists of an anti-parallel β-sheet, connected to a conserved α-helix (H1) by a random-coiled region, which is stabilized at its N-side by a disulphide bridge established between the C29 and C31. The C-terminal domain (a.a. 70–151) houses the functional elements required for DNA binding and repair: (i) the catalytic C119 residue within the conserved PCHR signature; (ii) the helix-turn-helix motif (HTH), which binds the DNA minor groove and holds the arginine finger (R102) that participates to the modified base flipping out from the

DNA base stacking; (iii) the ‘asparagine hinge’ that, together with the helix H4 of the HTH, defines one wall of the ligand-binding pocket; and (iv) the active site loop on the H4-facing side of the active site (Figure 1A). Previously, we demonstrated that mutation of the R102 residue reduces DNA binding efficiency, whereas mutation of five residues in the HTH motif abolishes the protein ability to form stable complexes with DNA, although both mutants are normally proficient in the trans-alkylation reaction (17).

An interesting feature, not previously reported for other AGTs, is the C29-C31 S-S bridge of the N-terminal domain. To test its role in the SsOGT activity and stability, we produced the C29A mutant, which resulted as active as the wild-type protein in repair of a O⁶-MG-containing ds oligonucleotide at 50°C (Supplementary Figure S1B). In contrast, the C29A mutant was significantly less thermostable than the wild-type (Figure 1B): incubation at increasing temperatures showed that, whereas the wild-type protein remained 100% soluble after 20 min at 70°C, the C29A protein aggregated above 60°C. Quantitative analysis by Differential Scanning Fluorimetry (DSF) allowed calculation of a T_m of 80°C for SsOGT and 60°C for C29A (Figure 1C). Thus, the S-S bond is not involved in DNA repair activity, but is an important structural element contributing to the impressive thermal stability of SsOGT.

Alkyltransferase-like proteins (ATLs) share structural similarity with AGTs, but are catalytically inactive and are believed to act as DNA damage sensors (35). Although the general fold is shared by hAGT and *Schyzosaccharomyces pombe* ATL1, important differences were found in the catalytic loop and Asn hinge, resulting in larger size of the lesion-binding pocket in ATL1, which might account for its broad lesion recognition range (36). Superimposition of SsOGT, ATL1 and hAGT structures showed the same differences, confirming the higher similarity of SsOGT with hAGT as compared with ATL1 (Supplementary Figure S2).

Structure-function analysis of dsDNA-bound SsOGT

Different strategies have been adopted to trap the AGT-alkylated dsDNA complex and solve its structure: the wild-type hAGT has been crystallized crosslinked to oligonucleotides containing alkylated guanine analogues (8,9), and the hAGT C145S inactive mutant was co-crystallized with a more physiologic, O⁶-MG-containing, 13 base pair long oligonucleotide (dsDNA^m) (8). These DNA-bound hAGT structures were found essentially superposable to each other and to ligand-free hAGT, suggesting that binding to DNA does not affect the protein overall structure (8,9).

To solve the SsOGT DNA-bound crystal structure, we obtained the C119A mutant, carrying a substitution of the catalytic C119 to prevent protein alkylation and subsequent dissociation from the repaired DNA substrate (data not shown). The crystal structure of the C119A::dsDNA^m (Figure 2A) was solved at 2.7 Å resolution (Table 1), revealing that a single SsOGT monomer is able to occlude 4 base-pairs (bp) on dsDNA substrate, paralleling what observed for DNA-bound hAGT (37–39). The crystal structure of the ligand-free protein and of the C119A::dsDNA^m complex can be superposed with a 0.3 Å average root-mean-square deviation (r.m.s.d.) of Cα positions. Assuming that

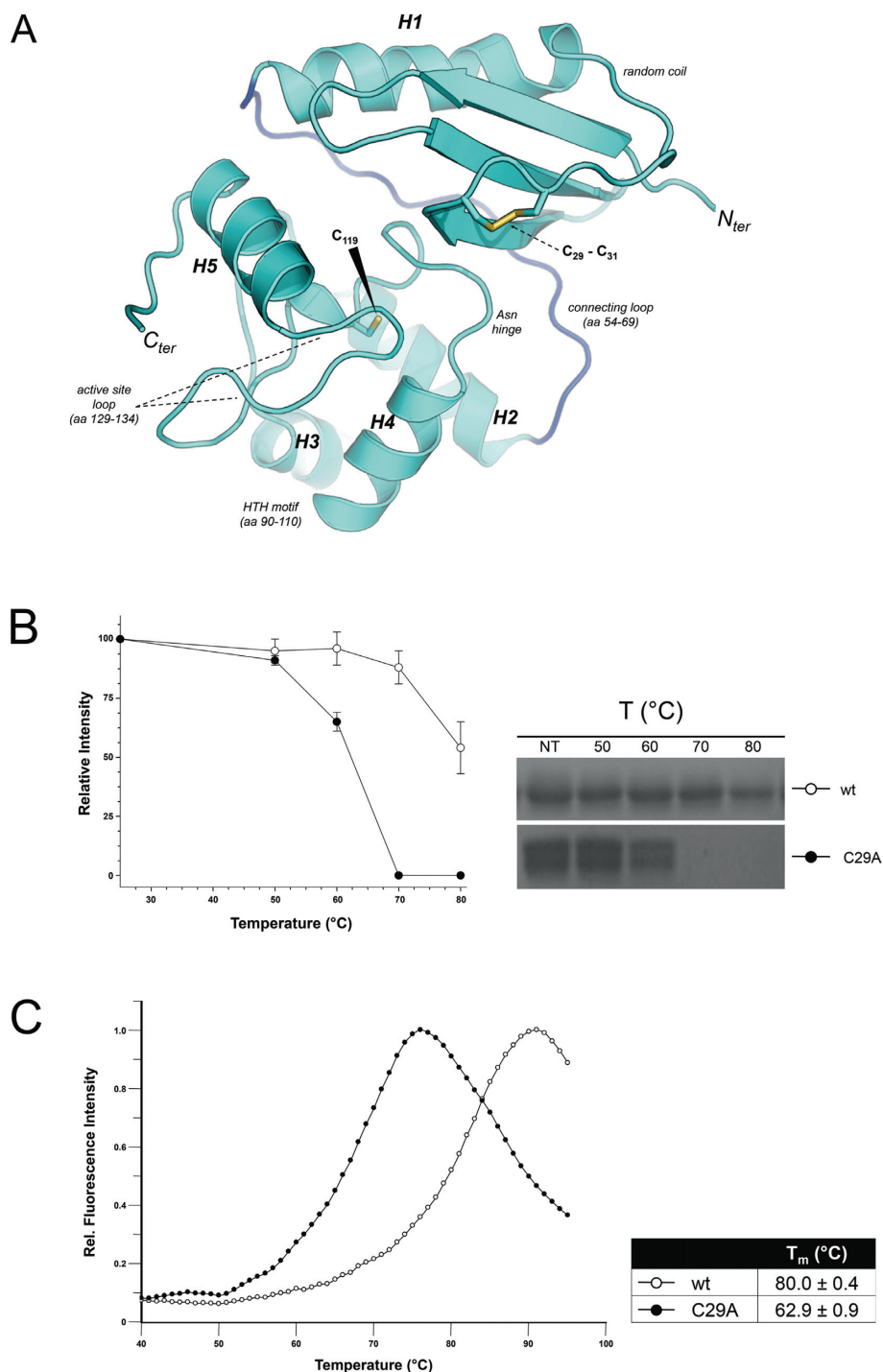


Figure 1. Structure–function of *SsOGT*. (A) Cartoon representations of the crystal structure of wild-type ligand-free *SsOGT*. Secondary structure elements and functional domains/motifs are indicated; residues commented through the text appears as sticks. (B) Thermal stability of the C29A mutant. The indicated proteins (0.1 mg/ml, 5 μ M) were incubated for 20 min at the indicated temperatures; after incubation, samples were centrifuged for 20 min at $16\,000 \times g$ and supernatants were analyzed by SDS-PAGE (left); the relative intensity of the Coomassie-stained bands were plotted as a function of temperature (right), considering as 100% the reference value at 25°C for each protein; shown is the mean \pm SD from three independent experiments. (C) Differential Scanning Fluorimetry (DSF). The relative fluorescence intensity values of each protein as a function of temperature were measured and used to obtain the T_m values. Data were obtained from three independent experiments.

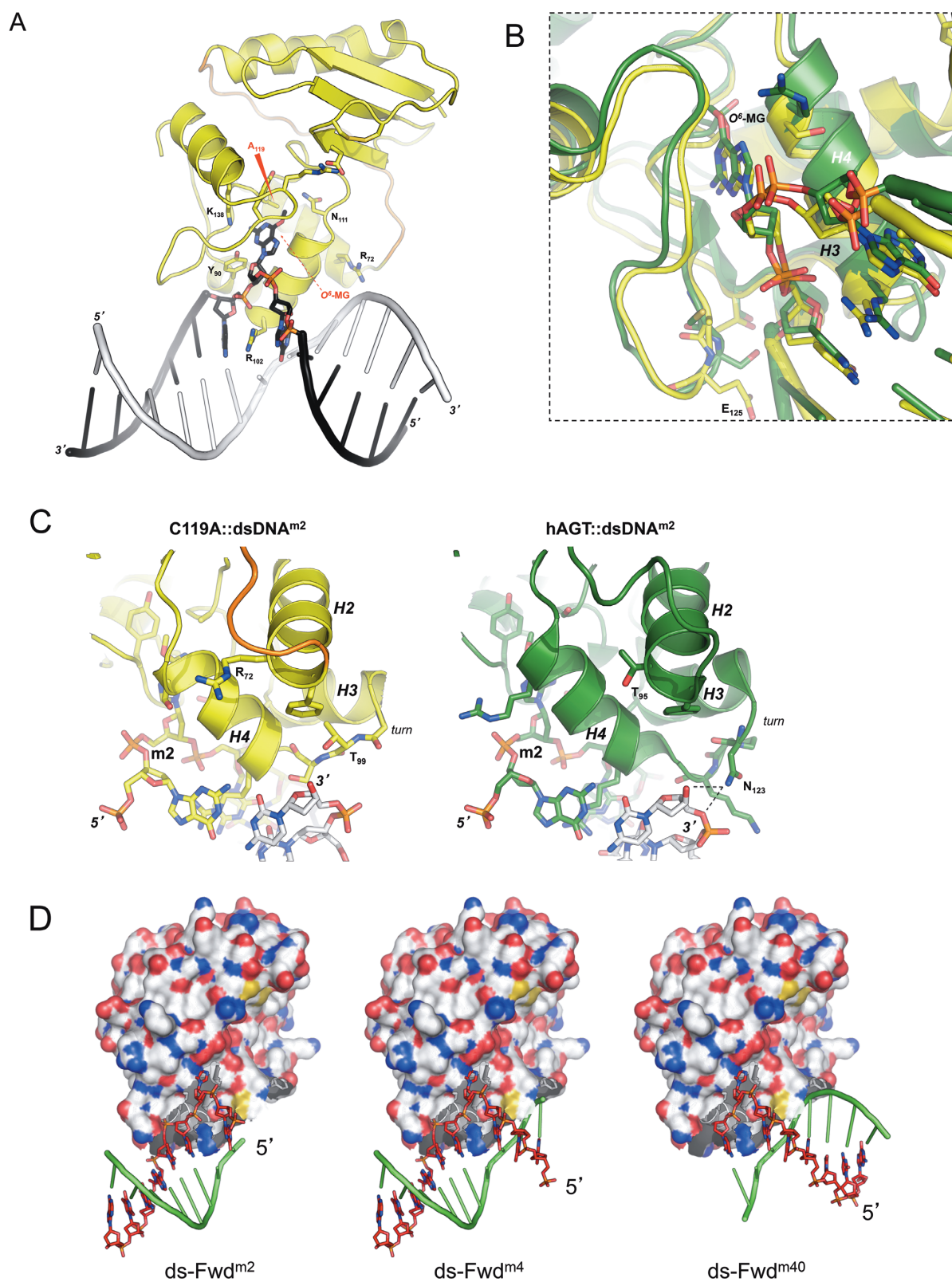


Figure 2. Structure-function of *SsOGT*-DNA complex. (A) Cartoon representations of the crystal structure of the C119A mutant in complex with dsDNA^m. (B) Detail of the interaction of superposed C119A (yellow) and hAGT (green) (PDB ID: 1T38) with the *O*⁶-MG ds DNA at the 3' side of the lesion; (C) left, detail of the C119A::ds DNA^{m2} interactions at the 5' side of the lesion; right, detail of the hAGT::ds DNA^{m2} interactions at the 5' side of the lesion; in both images the damaged DNA strand is colored as the corresponding protein chain. (D) Structure-based models of C119A::ds oligonucleotides carrying the lesion in different positions (see Supplementary Table S1).

Table 1. Data collection, phasing, and refinement statistics

| | SsOGT | SsOGT ^m | SsOGT-C119L | C119A::dsDNA ^m |
|--|---|--------------------------------------|--------------------------------------|--------------------------------------|
| Data collection | | | | |
| Space Group | R3 | P212121 | I432 | P212121 |
| Unit cell (Å) | a = b = 94.72 | a = 48.49 b = 50.22 c = 142.02 | a = b = c = 140.56 | a = 41.76 b = 65.88 c = 97.52 |
| | c = 76.70 $\alpha = \beta = \gamma = 90^\circ$ | $\alpha = \beta = \gamma = 90^\circ$ | $\alpha = \beta = \gamma = 90^\circ$ | $\alpha = \beta = \gamma = 90^\circ$ |
| Wavelength (Å) | 0.873 | 0.979 | 1.89 | 0.972 |
| Resolution (Å) | 1.8 | 2.8 | 2.6 | 2.7 |
| Total reflections | 83580 | 35418 | 78410 | 41645 |
| Unique reflections | 21897 | 8764 | 7600 | 7960 |
| Mean(I)/sd(I) | 11.4 (2.43) ^a | 10.1 (3.1) ^a | 25.6 (5.3) ^a | 10.62 (1.7) ^a |
| Completeness (%) | 99.9 (99.8) ^a | 97.5 (98.7) ^a | 99.9 (100) ^a | 98.5 (90.9) ^a |
| Multiplicity | 3.8 (3.7) ^a | 4.0 (4.1) ^a | 10.3 (10.3) ^a | 5.2 (5.2) ^a |
| R _{merge} (%) | 6.7 | 12.9 | 7.3 | 8.1 |
| R _{meas} (%) | 7.8 | 14.8 | 7.7 | 9.0 |
| Refinement | | | | |
| R _{factor} /R _{free} (%) | 16.0/18.2 | 19.0/29.0 | 17.5/24.2 | 21.3/27.1 |
| Protein/DNA Atoms | 1204 | 2380 | 1175 | 1707 |
| Ligand atoms | 24 | - | - | - |
| Water molecules | 211 | 12 | 45 | 6 |
| R.m.s.d. bonds (Å) | 0.007 | 0.010 | 0.009 | 0.012 |
| R.m.s.d. angles (°) | 0.98 | 1.27 | 1.31 | 1.29 |
| Average B (Å ²) | | | | |
| Protein | 25.3 | 17.9 | 35.3 | 63.8 |
| Solvent | 38.9 | 11.0 | 39.4 | 49.6 |

^aValues in parentheses refer to the highest resolution shell.


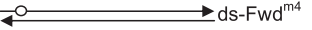
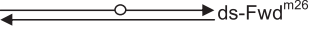

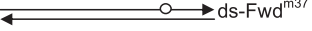
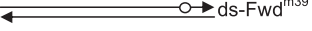
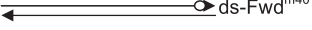
the structure of the C119A::dsDNA^m complex is similar to that of the physiologic complex formed by the wild-type protein, our analysis indicates that DNA binding does not substantially alter the protein architecture, as also shown for hAGT (8).

The C119A::dsDNA^m crystal structure was largely, but not completely, similar to that described for hAGT in complex with the same dsDNA^m substrate. The interaction network established by both proteins with the region of the modified strand at the 3' side of the methylated base appears conserved. Indeed, in the C119A::dsDNA^m complex the phosphodiester bond between bases +1 and +2 downstream of O⁶-MG is clamped between the positive dipole momentum at the N-side of the H3 helix of the HTH domain and the backbone nitrogen of E125, at a mean 3.0 Å distance (Figure 2B). From this analysis we predict that at least 2 bases downstream to the O⁶-MG are needed to establish correct interactions, as shown for hAGT (40). In contrast, the protein-dsDNA interactions at the 5' end of the O⁶-MG base appear less tight in the C119A::dsDNA^m structure compared to the human counterpart. Indeed, we found important differences in structure-based models of SsOGT and hAGT in complex with a dsDNA^{m2}, an oligonucleotide carrying the O⁶-MG lesion located one base from the 5' end of the damaged strand (Supplementary Table S1). In the C119A::dsDNA^{m2} complex, the phosphodiester bond joining the last two bases at the 3' end of the complementary intact strand is not engaged in contacts with the protein (Figure 2C, left panel), while it appears kept in place by the side chain of N123 in an equivalent model of the hAGT::dsDNA^{m2} complex (Figure 2C, right panel).

hAGT-catalyzed repair of dsDNA molecules containing a single O⁶-alkylguanine is directionally biased, as alkylated

bases at positions near the 3' end of the modified strand are repaired less efficiently than those located in the middle or at the 5' end (40). We reasoned that, if this behavior reflects the protein-DNA complex architecture, SsOGT might show a different bias toward the position of the lesion. To test this hypothesis we modified our previously developed competitive kinetic assay for O⁶-BG sensitive AGTs activity (17,19). Briefly, AGT proteins become covalently labeled when incubated with the fluorescent competitive inhibitor SNAP-Vista GreenTM (hereafter BG-VG); the fluorescence intensity of the protein bands is a direct measure of the protein trans-alkylation activity. Moreover, the fluorescence intensity obtained in competition assays with BG-VG and unlabeled alkylated DNA is an indirect measure of the efficiency of DNA repair (17). This method was successfully applied to determination of kinetic constants for DNA trans-alkylation reaction and DNA repair activity by SsOGT and *Mycobacterium tuberculosis* OGT (17,19). We have modified this assay to allow determination of DNA repair activity by SsOGT by measuring the protein fluorescence intensity in competition assays performed in the presence of fixed amounts of BG-VG and increasing amounts of ds oligonucleotides containing a single O⁶-MG. This method allows rapid determination of an IC₅₀, which can be converted to K_{DNA}, giving an indirect measure of the efficiency of O⁶-MG repair by SsOGT (Supplementary Figure S1A). The method was validated by determining the K_{DNA} of SsOGT for the ds-UP^m oligonucleotide (Supplementary Table S1), which was comparable to the value previously reported (17) (Supplementary Figure S1A). We thus applied this method to determine the efficiency of repair by SsOGT of 41 bp dsDNA substrates, each containing a single O⁶-MG at different posi-

Table 2. Polarity of lesion recognition by SsOGT

| oligonucleotide | IC ₅₀ (μ M) | K _{DNA} (μ M) |
|---|--------------------------------|--------------------------------|
|  ds-Fwd ^{m2} | 6.38 ± 1.20 | 4.29 ± 0.39 |
|  ds-Fwd ^{m4} | 1.01 ± 0.08 | 0.68 ± 0.03 |
|  ds-Fwd ^{m26} | 1.23 ± 0.07 | 0.83 ± 0.02 |
|  ds-Fwd ^{m33} | 0.73 ± 0.11 | 0.49 ± 0.04 |
|  ds-Fwd ^{m37} | 0.89 ± 0.11 | 0.59 ± 0.04 |
|  ds-Fwd ^{m39} | 1.10 ± 0.12 | 0.74 ± 0.04 |
|  ds-Fwd ^{m40} | 84.3 ± 73.2 | 56.6 ± 23.8 |

The ds oligonucleotides are named according to the position of the O⁶-MG with respect to the 5' end of the upper strand. Arrows indicate the 3' end of each strand. Each IC₅₀ value was obtained from three independent experiments.

tions along the molecule (Supplementary Table S1; see also Figure 2D). The efficiency of repair was comparable when the methylated base was located in the middle or within 3 bases from either the 5' or the 3' end. In contrast, the efficiency of repair was significantly reduced (about 6-fold) when the damaged base was located one base from the 5' end, and dropped dramatically (more than 50-fold) when the lesion was located one base from the 3' end (Table 2). Solution binding and AFM studies clearly indicated that hAGT binds both single- and double-stranded substrates in a cooperative fashion (38,39,41–44); the wild-type SsOGT also showed cooperative binding when analyzed in EMSA assays with short oligonucleotides (17). The C119A protein showed similar behavior and bound all methylated oligonucleotides with comparable efficiency (Supplementary Figure S3), thus suggesting that the observed differences reflect the efficiency of O⁶-MG recognition and/or removal, rather than of unspecific DNA binding.

These results show that SsOGT, like hAGT, is highly inefficient in repairing lesions near the 3' end of the molecule, in line with the conserved network of interactions formed with the DNA strand on the 3' side of the lesion by both proteins (Figure 2B). In contrast, whereas two bases from the 5' end of the molecule are sufficient for efficient repair by hAGT, SsOGT requires at least 4 bases between O⁶-MG and the 5' end for optimal activity (a situation illustrated by the central miniature in Figure 2D); in this latter case, positively charged residues of both helix H2 (e.g. R72 in Figure 2A) and helix H4 could efficiently contact the sugar-phosphate backbone of the complementary strand 3' end, fully restoring the protein-DNA association potential.

Effect of alkylation on SsOGT stability

Methylated and benzylated forms of hAGT are highly unstable (37), and soaking of hAGT crystals in O⁶-alkyl-guanine substrates leads to crystals destabilization (7). Considering that SsOGT is extremely stable at high temperature, but is also active at room temperature (17), we wondered whether it might be alkylated and remain relatively stable if kept at temperatures below its optimum. Consis-

tently with this assumption, we obtained a homogeneously methylated protein form (SsOGT^m) by incubating purified SsOGT with O⁶-MG at room temperature. Since the benzylated SsOGT form (SsOGT^b) was not soluble, to test the effect of the presence of larger adducts in the protein active site, we constructed two site-directed mutants carrying substitutions of the catalytic C119 with an F or an L residue, mimicking a benzylated or isopropylated protein, respectively. Whereas the corresponding C145F and C145L mutants of hAGT were extremely unstable when expressed in *E. coli* cells (7), both SsOGT mutants resulted stable in the same expression system and could be purified to homogeneity (data not shown).

When tested in EMSA analysis, SsOGT^m and mutants were able to bind dsDNA, although with slightly reduced affinity as compared with the wild-type protein (Figure 3A), in line with results reported for hAGT, whose dsDNA binding affinity decreases to little extent when the protein is alkylated *in vitro* (37). We then analyzed the stability of the mutants and SsOGT^m at different temperatures. Whereas, as shown above, the wild-type protein was 100% soluble after 20 min incubation at 70°C, complete aggregation was observed for SsOGT^m at 70°C and for C119F at 50°C (Figure 3B). DSF analysis demonstrated the instability of these proteins, showing a *T_m* value of 60°C for SsOGT^m and ca. 45°C for both the C119L and C119F mutants, compared with the wild-type protein (*T_m* = 80°C) (Figure 3C). Thus, the presence of alkylated groups bound to the catalytic C119 leads to *T_m* destabilization, whose extent is dependent on the size of the adduct in the active site.

Effect of alkylation on SsOGT structure

In order to elucidate the structural basis of the dramatic effect of alkylation on protein stability, we solved the crystal structure of the SsOGT^m and SsOGT-C119L proteins at 2.8 and 2.6 Å resolution respectively, whereas we failed to obtain suitable crystals of the C119F mutant, likely as a consequence of its intrinsic instability.

The analysis of the optimally superposed crystal structures of the ligand-free and SsOGT^m proteins revealed conformational changes that occur in discrete regions of the molecule, upon the C119 methylation (Figure 4A). The only comparable structures available so far are those obtained by alkylating hAGT after crystallization (7); similarly to what observed in the human counterpart, in the SsOGT^m structure we observed a 1.0 Å movement of the C-side of the recognition helix H4 and following Asn-hinge, a conformational change responsible for a modest increase of the ligand binding site size (Figure 4A). In addition, in the structure of SsOGT^m the distance between the C-side of the conserved H2 helix and the H4 recognition helix is increased of approximately 3 Å with respect to the observed distance between the same structural elements in the ligand-free wild-type protein. Similarly, the distance between the active site loop and the H4 is increased of 1.9 Å in the SsOGT^m structure. Alkylation of hAGT crystals induce shift of 0.5–1.5 Å C α of the recognition helix away from the N-terminal domain (7).

An unbiased difference distance matrix plot of SsOGT^m versus SsOGT revealed that many interactions between

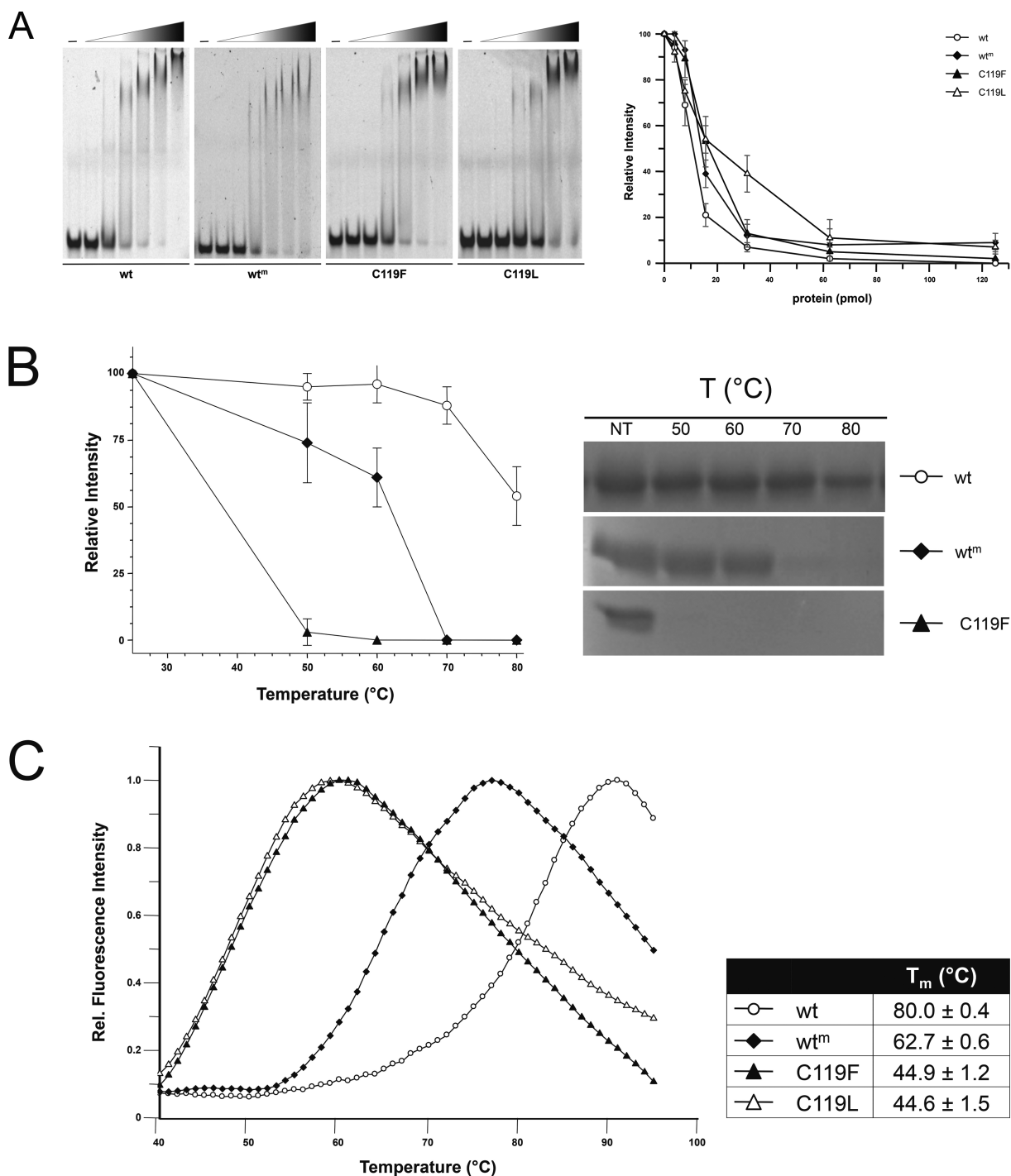


Figure 3. Effect of active site alkylation on *Ss*OGT activity and stability. **(A)** EMSA. Reactions (10.0 μ l) containing increasing amounts of each indicated protein (0–25.0 μ M) were incubated with the TAMRA-labeled ds A^+/A^- oligonucleotide (0.2 μ M; Supplementary Table S1) for 10 min at 37°C. Native polyacrylamide gels were analyzed by gel fluorescence imaging. The first lane of each gel is the no protein control. For quantification (right), the relative intensity of each band was plotted as a function of the protein concentration (pmol); data are from three independent experiments. **(B)** Thermal stability. Reactions were set and analyzed and quantified as described in the legend to Figure 1B. **(C)** DSF. Data were obtained from three independent experiments as described in the legend to Figure 1C.

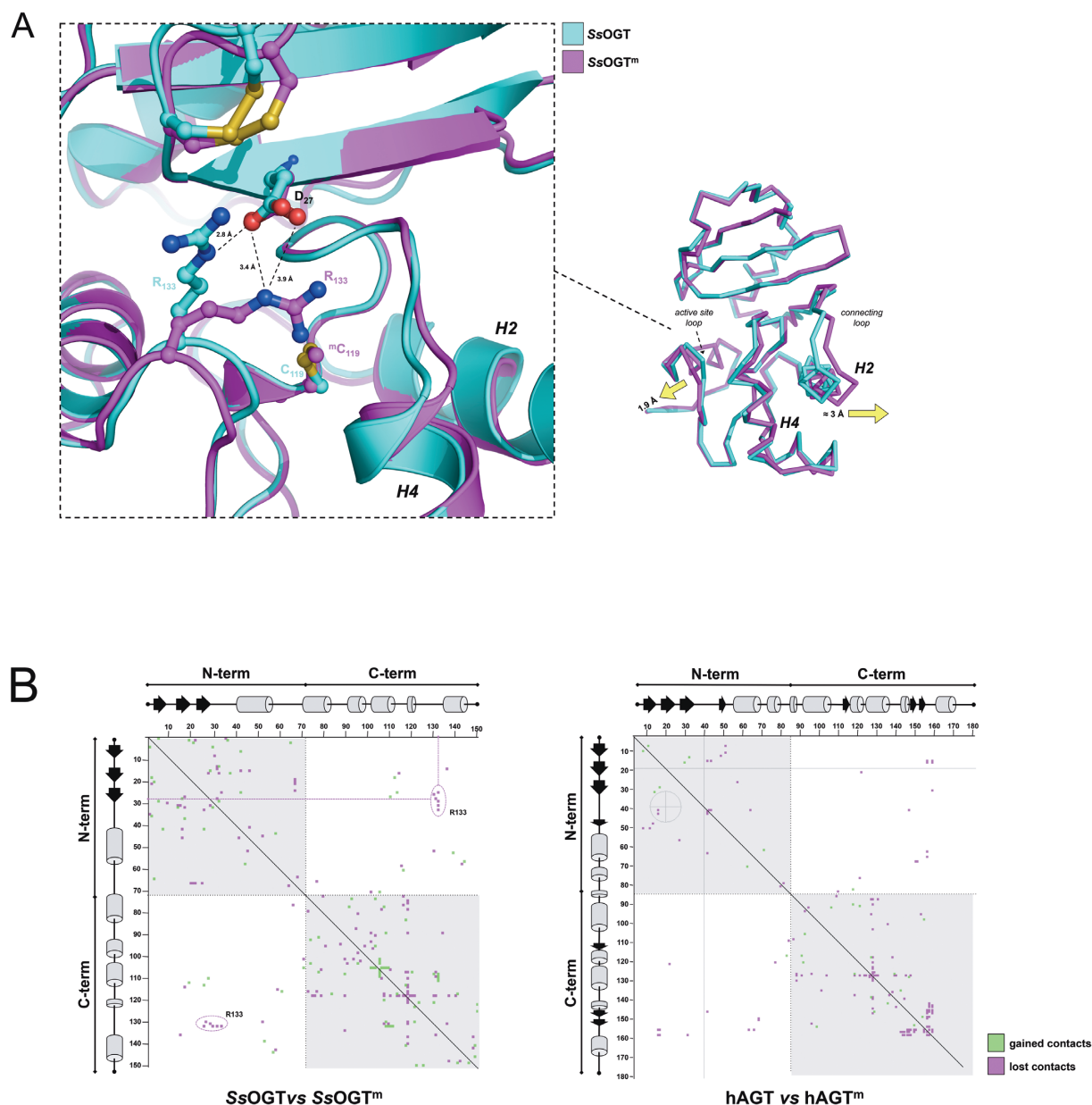


Figure 4. Comparative analysis of structures of free and methylated *SsOGT*. **(A)** Ribbon representation of the crystal structures of the two proteins upon optimal overlaying. Each protein chain is uniformly colored following the legend on the right of the picture. The arrows indicate the movements described in the text. An enlarged view of the D27 and R133 residues (drawn as sticks) appears in the inset on the left. **(B)** Distance difference matrix plot, obtained by the computation of Needleman-Wunsch sequence alignment by using the CMView v1.1.1 freeware, choosing amino acid side chain contact type and 8 Å cutoff. (left panel) methylated versus native *SsOGT*; (right panel) methylated (PDB ID: 1EH7) versus native hAGT (PDB ID: 1EH6).

amino acid residues were lost or gained, suggesting overall reshaping of the molecule upon methylation (Figure 4B, left panel). Compared with analogous analysis performed on alkylated versus free hAGT (Figure 4B, right panel), many more interactions were affected upon *SsOGT* methylation, thus suggesting that crystallization of the protein in its reacted form allow us to observe the repositioning of higher number of residues than that observed in the hAGT methylated in crystallized form. In both proteins most rearrangements occurred within the N- and C-terminal domain, respectively (Figure 4B, gray quarters) rather than between domains (white quarters); however, in *SsOGT*^m a

group of inter-domain interactions were lost in correspondence of the R133 residue (Figure 4B). Interestingly, in the *SsOGT* structure, the R133 residue at the C-side of the active site loop of the C-terminal domain and the D27 residue at the C-side of the third β -strand of the N-terminal domain form an interaction at 2.6 Å distance (Figure 4A, inset). In the *SsOGT*^m structure the distance between the carboxylic group of D27 and the guanidinium group of R133 is increased by 0.6 Å and the R133 side chain is rotated by ca 60° with respect to its position in the unmodified protein on the same plane, overall resulting in a 1.9 Å movement of the active site loop toward the solvent (Figure 4A). Distance

increase from the D27 residue and rotation of the R133 residue was also observed in the superposition of the C119L structure with that of the wild type, although at lower extent (Supplementary Figure S4, inset), suggesting again that the modification of the D27–R133 interaction is a direct consequence of the active site modification. This observation prompted us to hypothesize that the C119 alkylation could have a role in the conformational change of the active site loop, which, in turn, might perturb the D27/R133 interaction, or alter the interaction geometry, ultimately triggering protein destabilization.

The active site expansion is not observed in the crystal structure of the C119L mutant, which rather shows a 1 Å movement of the H2 helix toward the Asn hinge (Supplementary Figure S4); therefore the C119L mutant is not as informative as the *Ss*OGT^m protein with respect to the description of structural rearrangements taking place upon trans-alkylation reaction.

Alkylation-induced structural changes affecting the connection between the two domains impair *Ss*OGT stability and activity

In order to directly test whether the D27/R133 interaction plays a role in *Ss*OGT stability, we prepared two mutants carrying substitutions of the D27 residue, namely D27A and D27K, whose positive charge should eventually enhance the mutation destabilizing effect. In thermal stability assays, the D27A mutant was indeed destabilized with respect to the wild type, showing complete aggregation at 70°C; moreover, the D27K protein was dramatically less stable and aggregated completely above 50°C (Figure 5A); the T_m determined by DSF was of 72°C for D27A and 44.7°C for D27K, respectively (Figure 5B). Thus, these data confirmed our prediction that the interaction between D27 and R133 plays a crucial role in *Ss*OGT stability.

We then characterized the D27A and D27K activity. Both mutants were effective in dsDNA binding, although the D27K mutant showed slightly reduced binding efficiency as compared with the wild-type and the D27A proteins (Figure 5C). In addition, mutated proteins were tested in their DNA repair activity by using the fluorescent assay with *O*⁶-MG containing ds oligonucleotides described in Supplementary Figure S1. While the D27A mutant showed slight reduction of repair efficiency, as compared with the wild-type (IC_{50} of 2.26 ± 0.13 versus 1.23 ± 0.07 μM at 50°C), we could not calculate an IC_{50} value for the D27K protein at either 50 or 25°C, thus suggesting that the D27K protein is unable to repair DNA, or this activity is greatly impaired. A slow DNA repair activity by the D27K protein could be detected in prolonged (6 h) incubation at 25°C with methylated oligonucleotides (Figure 5D; note that under the same conditions the wild-type *Ss*OGT completes the reaction within few minutes; data not shown); thus, the D27K protein is not completely inactive, yet its DNA repair activity is strongly compromised. This result suggests that the disruption of the D27–R133 interaction not only affects the protein stability, but also its DNA repair activity even at low temperature, thus not as a consequence of thermal induced denaturation. To understand the reason of the D27 mutants DNA repair defect, we then determined

the efficiency of the alkyl-transfer reaction independently of DNA binding and lesion recognition, using our previously developed assay based on the use of the synthetic substrate BG-VG (17). Interestingly, only marginal reduction of the catalytic efficiency of the covalent modification reaction at 25°C was observed for the D27K mutant, and no significant changes for the D27A, as compared with the wild-type protein (Figure 5E). Thus, the D27/R133 interaction does not play an important role in the alkyl-transfer reaction from free alkylated bases to the catalytic cysteine, while it greatly affects the efficiency of de-methylation of *O*⁶-MG bases in the context of dsDNA.

DISCUSSION

*Ss*OGT proved to be a convenient model to unravel the structure-function relations of AGTs. Indeed, the overall architecture of the free and DNA-bound protein (this work), as well as details of its reaction mechanisms (17) are conserved between *Ss*OGT and mesophilic counterparts. Most important, the peculiar thermal stability of *Ss*OGT allowed us to obtain the protein in a post reaction form suitable to both biochemical analysis and crystallization, overcoming the restrictions to experimental manipulation imposed by the extreme instability of alkylated hAGT forms.

One peculiar feature of *Ss*OGT is the disulphide bond in its N-terminal domain; this structural element is not found in the corresponding domain of hAGT, which instead contains a zinc atom coordinated by two cysteine and two histidine residues (7). Previously, it was reported that nitric oxide synthase undergoes a conformational change with release of a Zn ion coordinated by four cysteines, coupled to formation of a disulphide bond involving two of such cysteines (45), thus suggesting that an exchange between Zn ion coordination and S-S bond is in principle possible. Although in the nitric oxide synthase this change appears to have regulatory function, it might occur in *Ss*OGT as an artifact due to Zn depletion during purification of the His-tagged protein through metal-chelating columns, inducing formation of the S-S bond. However, this is unlikely to be the case for two reasons: (i) the S-S bond is also present in the OGT protein from the strictly related *Sulfolobus tokodai* species, showing 68% aminoacid sequence identity with *Ss*OGT; although no functional data are available for this protein, its crystal structure (PDB entry 1WRJ) was obtained from the protein without any tag (http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=1wrj&template=header_records.html&r=getheader); (ii) the *Ss*OGT structure does not reveal any possible canonical tetrad for Zn coordination around the C29–C31 residues. Thus, the S-S bond appears to be a structural peculiarity of the thermophilic protein, and we showed that its disruption significantly destabilizes *Ss*OGT; it is possible that the N-terminal domain is involved in protein activity, as previously suggested (19,46), as well as in the overall stability of AGTs, although by means of different structural elements in different proteins.

The efficiency of DNA repair by *Ss*OGT is highly dependent on the *O*⁶-MG position, and its optimal activity requires at least three bases at both sides of the lesion, consis-

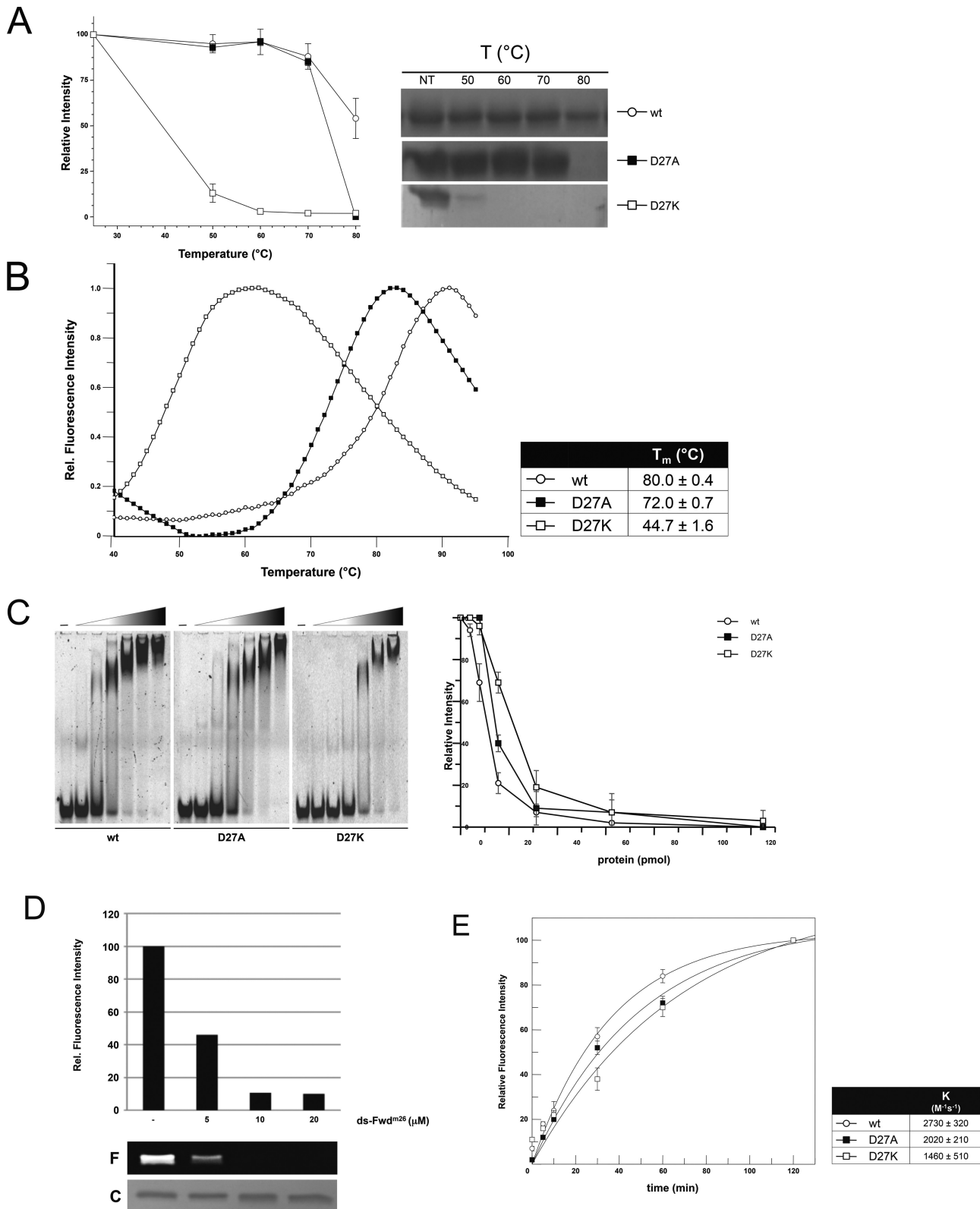


Figure 5. Effect of D27 residue mutations on *SsOGT* stability and activity. **(A)** Thermal stability (see legend to Figure 1B; mean ± SD from three independent experiments). **(B)** DSF. (see legend to Figure 1C; data from three independent experiments). **(C)** EMSAs were performed and quantified as in the legend to Figure 3A. The first lane of each gel is the no protein control. **(D)** Direct DNA repair activity of *SsOGT* D27K mutant. The protein (5 μM) was incubated for 6 h at 25°C in the presence of increasing amounts of the ds-Fwd^{m26} oligonucleotide, as indicated. Then, 10 μM of BG-VG was added and the incubation was prolonged for 2 h. The histogram reports the reduction of the fluorescence intensity of the protein band (shown in the gel labeled F) depending on the methylated oligonucleotide concentration in pre-incubation before BG-VG labeling. Equal amounts of protein were loaded in each lane, as shown by the Coomassie staining of the gel (C). **(E)** Trans-alkylation activity of D27 mutants. Second-order rate constants were obtained incubating the proteins at 25°C with BG-VG for increasing time spans; data were fitted as previously reported (17).

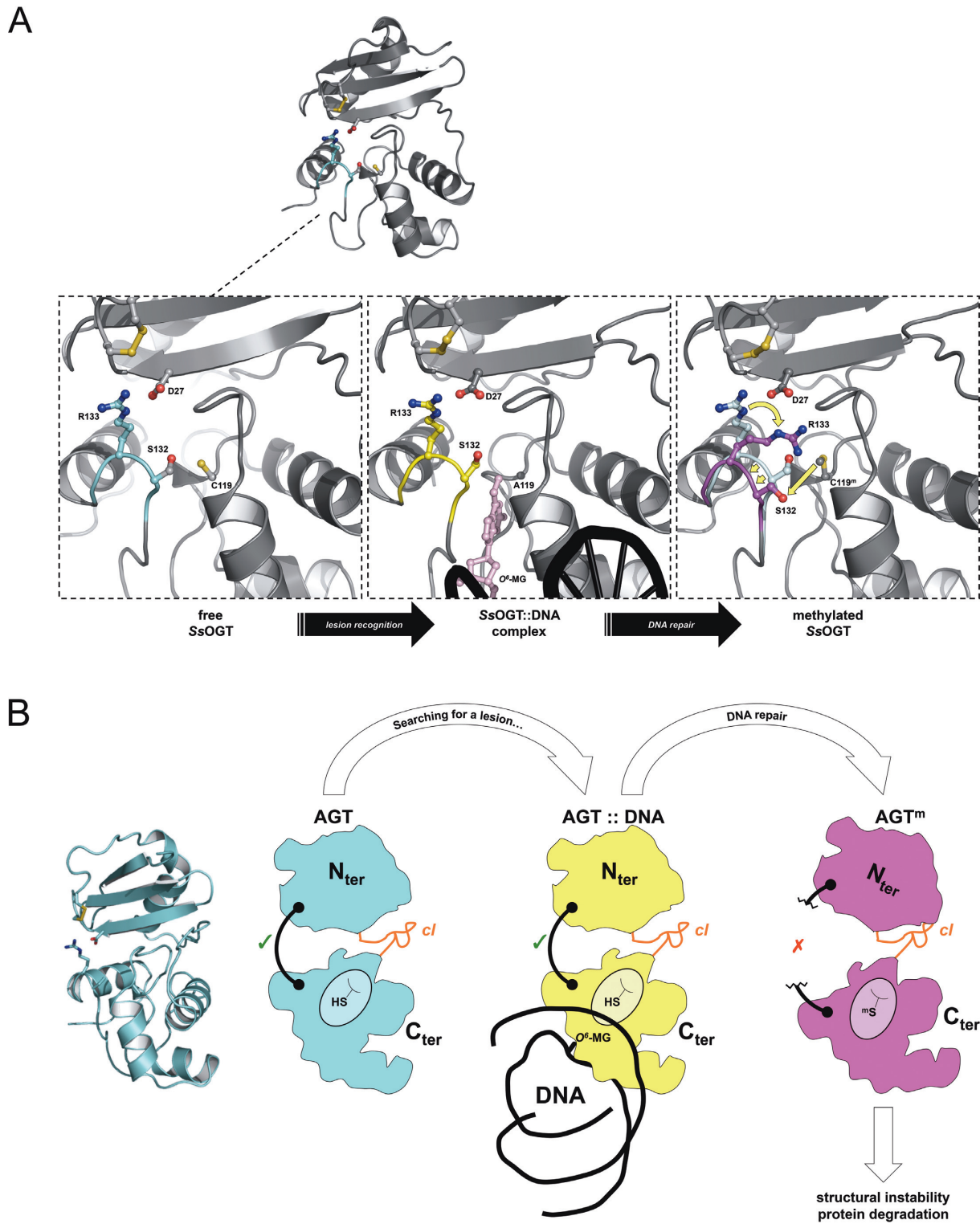


Figure 6. (A) Details of the D27–R133 interaction during the reaction cycle and its methylation-induced modification. (B) Model of the conformation modification induced by alkylation-induced perturbation of inter-domain interactions.

tent with the contacts between the protein and both DNA strands in crystal structure. Interestingly, whereas contacts formed with the DNA strand at the 3' side of the lesion are important for both hAGT and SsOGT, those formed with bases at the 5' side seem more important for the latter. One possible explanation for this difference is that SsOGT needs to form stronger contacts at both sides of the DNA substrate to assist stabilization of the DNA-protein complex at higher temperatures, and/or hold the bound DNA in a correct ds conformation, counteracting thermal denaturation and facilitating lesion removal.

Alkylation-induced instability of hAGTs is a well known process whose molecular mechanism is, however, poorly understood, mainly due to the instability of alkylated hAGT and its C145F and C145L mutants (7,37). Thanks to the relative stability of SsOGT^m, and of C119L and C119F mutants at mild temperatures, we could obtain direct and quantitative data on the protein stability in correlation with the active site status, as well as insights in the structural modifications occurring upon methylation in solution, thus in conditions more physiological than those in which the structures of alkylated hATG could be obtained (7). Indeed, although structural rearrangements were observed upon alkylation of hAGT crystals, Daniels and coworkers predicted that such rearrangements would be even larger in solution (7). Consistently, we found extensive remodeling of interactions between aminoacid residues upon methylation and larger movements in the backbone structure of SsOGT^m, as compared with alkylated hAGT. These data support the correlation between active site alkylation, conformational changes and protein unfolding (7). Our data are also in line with the observation that alkylated SsOGT undergoes degradation after treatment of *S. solfataricus* cells with alkylating agents (17): at the physiological growth temperature (75–80°C) alkylated SsOGT is destabilized, which might target the protein to degradation pathways, either directly or after some still unidentified post-translational modification.

Our structural and biochemical data show that the D27 residue of the N-terminal domain plays an important role in both SsOGT activity and stability, through the formation of an interaction with the R133 residue of the catalytic C-terminal domain (Figure 6A). Intriguingly, the remarkable extent of D27K destabilization is strikingly similar to that of the C119F and C119L mutants, showing that the same effect on protein stability is obtained by acting on two completely different residues. Moreover, the D27/R133 interaction is important not only to maintain SsOGT folding at high temperature, but also to allow its activity even at low temperature. Exploiting our different assays, which allow dissection of the SsOGT reaction, we showed that the D27–R133 interaction is not involved in the trans-alkylation reaction *per se*, whereas it is required for the protein to repair efficiently the alkylated base in the DNA context. An attractive hypothesis is that C119 alkylation-induced perturbation of the D27–R133 interaction weakens the contacts between the two domains, impairing optimal co-ordination between the N- and C-terminal protein domains, which in turn might be crucial both for DNA repair and maintenance of the correct protein folding.

We could also obtain the first crystal structure of an AGT mutant carrying a substitution of the catalytic cysteine with a leucine, mimicking an isopropyl group; the hAGT corresponding mutant was extremely unstable (7). The crystal structure of the C119L protein showed that the conformational modifications occurring in this mutant at the level of the D27–R133 interaction recall those occurring in the methylated protein. However, it is also important to note that not all the structural modifications observed in SsOGT^m are also present in the C119L protein, thus suggesting that the bulky adduct in the active site inserted by mutation may not completely reproduce the effect of the conformational changes triggered by alkylation. Indeed, alkylation takes place in the correctly folded protein, whereas the mutant accommodates the large adduct during its folding, which might affect the protein conformation in a different manner. These results suggest that structural and functional data obtained with substitution mutants should be carefully considered.

The inter-domain D27–R133 interaction is not conserved in the structures of other AGTs, as a consequence of poor sequence conservation at level of the N-terminal domain of AGTs from different species. However, in the case of *Thermococcus kodakaraensis* Tk-MGMT the two domains are connected by an ion-pair network formed by the R50–E93–R132 residues. Although no data are available on the effect of alkylation on this interaction, mutation to alanine of the E93 residue, located at the center of the network, destabilized the protein (32,47). In hAGT, alkylation causes disruption of the active site hydrogen-bond network and perturbs the hydrophobic packing between the N-hinge and the N-terminal domain, which gives an important contribution to the interface between the two protein domains (7). Upon alkylation, movements of the HTH helix H6 and collision of the N137 residue with the alkyl adduct disrupt three H-bonds formed by the N137 residue. Most important, mutation of N137 to alanine resulted in dramatic hAGT destabilization (48). Thus, although in three AGTs the connection between the two domains are provided by different structural elements, these observations suggest a common theme for alkylation-induced destabilization through perturbation of the connection between the two domains, triggering protein destabilization (Figure 6B).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

FIRB-Futuro in Ricerca RBFR12001G_002-Nematic and Merit RBNE08YFN3—Molecular Oncology. Funding for open access charge: FIRB-Futuro in Ricerca RBFR12001G_002-Nematic and Merit RBNE08YFN3—Molecular Oncology.

Conflict of interest statement. None declared.

REFERENCES

1. Pegg, A.E. (2000) Repair of O(6)-alkylguanine by alkyltransferases. *Mutat. Res.*, **462**, 83–100.

2. Daniels, D.S. and Tainer, J.A. (2000) Conserved structural motifs governing the stoichiometric repair of alkylated DNA by O(6)-alkylguanine-DNA alkyltransferase. *Mutat. Res.*, **460**, 151–163.
3. Tubbs, J.L., Pegg, A.E. and Tainer, J.A. (2007) DNA binding, nucleotide flipping, and the helix–turn–helix motif in base repair by O6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *DNA Repair*, **6**, 1100–1115.
4. Pegg, A.E. (2011) Multifaceted roles of alkyltransferase and related proteins in DNA repair, DNA damage, resistance to chemotherapy, and research tools. *Chem. Res. Toxicol.*, **24**, 618–639.
5. Sabharwal, A. and Middleton, M.R. (2006) Exploiting the role of O6-methylguanine-DNA-methyltransferase (MGMT) in cancer therapy. *Curr. Opin. Pharmacol.*, **6**, 355–363.
6. Wibley, J.E., Pegg, A.E. and Moody, P.C. (2000) Crystal structure of the human O(6)-alkylguanine-DNA alkyltransferase. *Nucleic Acids Res.*, **28**, 393–401.
7. Daniels, D.S., Mol, C.D., Arvai, A.S., Kanugula, S., Pegg, A.E. and Tainer, J.A. (2000) Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding. *EMBO J.*, **19**, 1719–1730.
8. Daniels, D.S., Mol, C.D., Arvai, A.S., Kanugula, S., Pegg, A.E. and Tainer, J.A. (2004) DNA binding and nucleotide flipping by the human DNA repair protein AGT. *Nat. Struct. Mol. Biol.*, **11**, 714–720.
9. Duguid, E.M., Rice, P.A. and He, C. (2005) The structure of the human AGT protein bound to DNA and its implications for damage detection. *J. Mol. Biol.*, **350**, 657–666.
10. Philip, S., Swaminathan, S., Kuznetsov, S.G., Kanugula, S., Biswas, K., Chang, S., Loktionova, N.A., Haines, D.C., Kaldis, P., Pegg, A.E. *et al.* (2008) Degradation of BRCA2 in alkyltransferase-mediated DNA repair and its clinical implications. *Cancer Res.*, **68**, 9973–9981.
11. Xu-Welliver, M. and Pegg, A.E. (2002) Degradation of the alkylated form of the DNA repair protein, O(6)-alkylguanine-DNA alkyltransferase. *Carcinogenesis*, **23**, 823–830.
12. Kanugula, S., Goodtzova, K. and Pegg, A.E. (1998) Probing of conformational changes in human O6-alkylguanine-DNA alkyltransferase protein in its alkylated and DNA-bound states by limited proteolysis. *Biochem. J.*, **329**, 545–550.
13. Rasimas, J.J., Dalessio, P.A., Ropson, I.J., Pegg, A.E. and Fried, M.G. (2004) Active-site alkylation destabilizes human O6-alkylguanine DNA alkyltransferase. *Protein Sci.*, **13**, 301–305.
14. Brunk, E., Mollwitz, B. and Rothlisberger, U. (2013) Mechanism to trigger unfolding in O(6)-alkylguanine-DNA alkyltransferase. *Chembiochem.*, **14**, 703–710.
15. Valenti, A., Napoli, A., Ferrara, M.C., Nadal, M., Rossi, M. and Ciaramella, M. (2006) Selective degradation of reverse gyrase and DNA fragmentation induced by alkylating agent in the archaeon *Sulfolobus solfataricus*. *Nucleic Acids Res.*, **34**, 2098–2108.
16. Valenti, A., Perugino, G., Nohmi, T., Rossi, M. and Ciaramella, M. (2009) Inhibition of translesion DNA polymerase by archaeal reverse gyrase. *Nucleic Acids Res.*, **37**, 4287–4295.
17. Perugino, G., Vettone, A., Illiano, G., Valenti, A., Ferrara, M.C., Rossi, M. and Ciaramella, M. (2012) Activity and regulation of archaeal DNA alkyltransferase: conserved protein involved in repair of DNA alkylation damage. *J. Biol. Chem.*, **287**, 4222–4231.
18. Vettone, A., Perugino, G., Rossi, M., Valenti, A. and Ciaramella, M. (2014) Genome stability: recent insights in the topoisomerase reverse gyrase and thermophilic DNA alkyltransferase. *Extremophiles*, **18**, 895–904.
19. Miggiano, R., Casazza, V., Garavaglia, S., Ciaramella, M., Perugino, G., Rizzi, M. and Rossi, F. (2013) Biochemical and structural studies of the *Mycobacterium tuberculosis* O6-methylguanine methyltransferase and mutated variants. *J. Bacteriol.*, **195**, 2728–2736.
20. Gautier, A., Juillerat, A., Heinis, C., Corrèa, I.R. Jr, Kindermann, M., Beauvais, F. and Johnsson, K. (2008) An engineered protein tag for multiprotein labeling in living cells. *Chem. Biol.*, **15**, 128–136.
21. Cheng, Y. and Prusoff, W.H. (1973) Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50 per cent inhibition (I₅₀) of an enzymatic reaction. *Biochem. Pharmacol.*, **22**, 3099–3108.
22. Niesen, F.H., Berglund, H. and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.*, **2**, 2212–2221.
23. Leatherbarrow, R.J. (2004) *GraFit, version 5.0*, Erithacus Software Ltd., Staines.
24. Kabsch, W. (2010) XDS. *Acta Crystallogr. D*, **66**, 125–132.
25. Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A. *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D*, **67**, 235–242.
26. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
27. Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D*, **66**, 213–221.
28. Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Zwart, P.H., Hung, L.W., Read, R.J. and Adams, P.D. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D*, **64**, 61–69.
29. Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D*, **66**, 486–501.
30. DeLano, W.L. (2002) *The PyMOL molecular graphics system*. DeLano Scientific, Palo Alto, CA.
31. Hashimoto, H., Nishioka, M., Inoue, T., Fujiwara, S., Takagi, M., Imanaka, T. and Kai, Y. (1998) Crystallization and preliminary X-ray crystallographic analysis of archaeal O6-methylguanine-DNA methyltransferase. *Acta Crystallogr. D Biol.*, **54**, 1395–1396.
32. Hashimoto, H., Inoue, T., Nishioka, M., Fujiwara, S., Takagi, M., Imanaka, T. and Kai, Y. (1999) Hyperthermostable protein structure maintained by intra and inter-helix ion-pairs in archaeal O6-methylguanine-DNA methyltransferase. *J. Mol. Biol.*, **292**, 707–716.
33. Roberts, A., Pelton, J.G. and Wemmer, D.E. (2006) Structural studies of MJ1529, an O6-methylguanine-DNA methyltransferase. *Magn. Reson. Chem.*, **44**, 71–82.
34. Moore, M.H., Gulbis, J.M., Dodson, E.J., Demple, B. and Moody, P.C. (1994) Crystal structure of a suicidal DNA repair protein: the Ada O6-methylguanine-DNA methyltransferase from *E. coli*. *EMBO J.*, **13**, 1495–1501.
35. Tubbs, J.L. and Tainer, J.A. (2010). Alkyltransferase-like proteins: Molecular switches between DNA repair pathways. *Cell. Mol. Life Sci.*, **67**, 3749–3762.
36. Tubbs, J.L., Latypov, V., Kanugula, S., Butt, A., Melikishvili, M., Kraehenbuehl, R., Fleck, O., Marriott, A., Watson, A.J., Verbeek, B. *et al.* (2009) Flipping of alkylated DNA damage bridges base and nucleotide excision repair. *Nature*, **459**, 808–813.
37. Rasimas, J.J., Pegg, A.E. and Fried, M.G. (2003) DNA-binding mechanism of O6-alkylguanine-DNA alkyltransferase. Effects of protein and DNA alkylation on complex stability. *J. Biol. Chem.*, **278**, 7973–7980.
38. Rasimas, J.J., Kar, S.R., Pegg, A.E. and Fried, M.G. (2007) Interactions of human O6-alkylguanine-DNA alkyltransferase (AGT) with short single-stranded DNAs. *J. Biol. Chem.*, **282**, 3357–3366.
39. Melikishvili, M., Rasimas, J.J., Pegg, A.E. and Fried, M.G. (2008) Interactions of human O(6)-alkylguanine-DNA alkyltransferase (AGT) with short double-stranded DNAs. *Biochemistry*, **47**, 13754–13763.
40. Luu, K.X., Kanugula, S., Pegg, A.E., Pauly, G.T. and Moschel, R.C. (2002) Repair of oligodeoxyribonucleotides by O(6)-alkylguanine-DNA alkyltransferase. *Biochemistry*, **41**, 8689–8697.
41. Adams, C.A. and Fried, M.G. (2011) Mutations that probe the cooperative assembly of O⁶-alkylguanine-DNA alkyltransferase complexes. *Biochemistry*, **50**, 1590–1598.
42. Tessmer, I., Melikishvili, M. and Fried, M.G. (2012) Cooperative cluster formation, DNA bending and base-flipping by O6-alkylguanine-DNA alkyltransferase. *Nucleic Acids Res.*, **40**, 8296–8308.
43. Tessmer, I. and Fried, M.G. (2014) Insight into the cooperative DNA binding of the O⁶-alkylguanine DNA alkyltransferase. *DNA Repair*, **20**, 14–22.

44. Melikishvili, M. and Fried, M.G. (2012) Lesion-specific DNA-binding and repair activities of human O⁶-alkylguanine DNA alkyltransferase. *Nucleic Acids Res.*, **40**, 9060–9072.
45. Crane, B.R., Rosenfeld, R.J., Arvai, A.S., Ghosh, D.K., Ghosh, S., Tainer, J.A., Stuehr, D.J. and Getzoff, E.D. (1999) N-terminal domain swapping and metal ion binding in nitric oxide synthase dimerization. *EMBO J.*, **18**, 6271–6281.
46. Fang, Q., Kanugula, S. and Pegg, A.E. (2005) Function of domains of human O⁶-alkylguanine-DNA alkyltransferase. *Biochemistry*, **44**, 15396–15405.
47. Nishikori, S., Shiraki, K., Yokota, K., Izumikawa, N., Fujiwara, S., Hashimoto, H., Imanaka, T. and Takagi, M. (2004) Mutational effects on O(6)-methylguanine-DNA methyltransferase from hyperthermophile: contribution of ion-pair network to protein thermostability. *J. Biochem.*, **135**, 525–532.
48. Crone, T.M., Goodtzova, K. and Pegg, A.E. (1996) Amino acid residues affecting the activity and stability of human O⁶-alkylguanine-DNA alkyltransferase. *Mutat. Res.*, **363**, 15–25.