8296–8308 *Nucleic Acids Research, 2012, Vol. 40, No. 17* *Published online 22 June 2012*
doi:10.1093/nar/gks574

Cooperative cluster formation, DNA bending and base-flipping by O^6-alkylguanine-DNA alkyltransferase

Ingrid Tessmer[1,*], Manana Melikishvili[2] and Michael G. Fried[2,*]

[1]Rudolf Virchow Center for Experimental Biomedicine, Institute for Structural Biology, University of Würzburg, Josef-Schneider-Strasse 2, 97080 Würzburg, Germany and [2]Center for Structural Biology, Department of Molecular and Cellular Biochemistry, University of Kentucky, Lexington, KY 40536, USA

Received March 14, 2012; Revised May 8, 2012; Accepted May 21, 2012

ABSTRACT

O^6-Alkylguanine-DNA alkyltransferase (AGT) repairs mutagenic O^6-alkylguanine and O^4-alkylthymine adducts in DNA, protecting the genome and also contributing to the resistance of tumors to chemotherapeutic alkylating agents. AGT binds DNA cooperatively, and cooperative interactions are likely to be important in lesion search and repair. We examined morphologies of complexes on long, unmodified DNAs, using analytical ultracentrifugation and atomic force microscopy. AGT formed clusters of $\leq$11 proteins. Longer clusters, predicted by the McGhee–von Hippel model, were not seen even at high [protein]. Interestingly, torsional stress due to DNA unwinding has the potential to limit cluster size to the observed range. DNA at cluster sites showed bend angles ($\sim$0, $\sim$30 and $\sim$60°) that are consistent with models in which each protein induces a bend of $\sim$30°. Distributions of complexes along the DNA are incompatible with sequence specificity but suggest modest preference for DNA ends. These properties tell us about environments in which AGT may function. Small cooperative clusters and the ability to accommodate a range of DNA bends allow function where DNA topology is constrained, such as near DNA-replication complexes. The low sequence specificity allows efficient and unbiased lesion search across the entire genome.

INTRODUCTION

The genomes of living cells are constantly exposed to alkylating agents from endogenous and exogenous sources (1,2). O^6-Alkylguanine and O^4-alkylthymine are mutagenic and cytotoxic base modifications that result from this exposure. In human cells, these modifications are repaired by O^6-alkylguanine-DNA alkyltransferase (AGT; also known as methylguanine methyltransferase) (3,4). This enzyme is of clinical interest because, in addition to its native roles, it also protects tumor cells against drugs that methylate or chloroethylate DNA (5–8). Clinical trials of AGT inhibitors are underway in attempts to increase the efficacy of alkylating agents in cancer chemotherapy (9–11). In addition, transgenic expression of AGT in hematopoietic stem cells is under study as a means to reduce the myelosuppressive effects of alkylating chemotherapy (11–13).

Human AGT is a small, monomeric protein (M_r = 21 519) that binds DNAs with little base composition, sequence or lesion specificity (14–17). Crystal structures of complexes formed on short DNAs indicate that individual molecules of AGT occupy $\sim$8 bp on the minor groove face of B-form DNA and bend the DNA toward the major groove by $\sim$15 to 30° (3,18,19). This bending is accompanied by the displacement of a DNA base from its helical conformation into the active site of the enzyme. If the base is O^6-alkylguanine or O^4-alkylthymine, the alkyl group is transferred to an active site cysteine (C145 in the human enzyme), restoring the base to unmodified structure and inactivating the alkyltransfer function of the enzyme (4,20). Alkylated AGT retains its DNA-binding activity (17), but is not re-activated; instead, alkylated molecules are ubiquitinated and degraded (20,21).

Solution measurements made with short DNAs show that DNA binding is cooperative ($(25 \leq \omega \leq 150)$ with a limiting density of 1 protein/4 bp (or nt for single-stranded substrates) (15,16). This contrasts with the isolated binding seen in the currently available crystal structures (3,18,19). Cooperative binding has been detected *in vitro*, under a wide range of conditions

*To whom correspondence should be addressed. Tel: +49 931 318 0425; Fax: +49 931 318 7320; Email: ingrid.tessmer@virchow.uni-wuerzburg.de
Correspondence may also be addressed to Michael G. Fried. Tel: +1 859 323 1205; Fax: +1 859 323 1037; Email: michael.fried@uky.edu

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

(varying pH, salt concentration, temperature, divalent cation concentration as well as DNA sequence and secondary structure) (14–16,22). In addition, a series of AGT mutations located within the putative protein–protein interface and outside of the crystallographically defined DNA interface have been found to reduce binding cooperativity *in vitro* and to render cells sensitive to the DNA alkylating agent MNNG (23). Together, these results argue strongly that cooperative binding is a *bona fide* function of native AGT that plays an important role in DNA repair *in vivo*.

To date, most studies of DNA binding by AGT have made use of short synthetic or natural DNAs. These offer important advantages of sequence and secondary structure homogeneity, as well as ease of handling. However, short substrates also limit the potential size of cooperative binding units and the number of ways of placing a cooperative unit on a DNA (24). Structures such as DNA loops, which require long substrates for stability, will naturally be under-represented. Finally, in a short duplex DNA, a substantial fraction of protein-binding sites are close to DNA ends and thus experience structural and counterion environments that are not typical of the centers of long DNA molecules (25,26). Herein, we examine the binding of AGT to homogeneous linear DNAs of 1000 and 2686 bp, derived from pUC19 plasmid. As shown below, these are large enough to accommodate AGT binding without length-dependent packing constraint or significant contribution from end effects. As will also be shown, cooperative binding results in the formation of contiguously bound protein clusters. We use cluster-size analysis to compare the predictions of the homogeneous McGhee–von Hippel binding model with the properties of the AGT system and propose a novel mechanism for the limitation of cooperative cluster sizes. We quantify DNA bends associated with AGT clusters and compare the results to bends found in crystalline AGT–DNA complexes. Finally, we present evidence for an unexpected affinity for DNA ends. The results suggest ways that cooperative binding may contribute to AGT function *in vivo*.

MATERIALS AND METHODS

Reagents

Agar, yeast extract and tryptone broth were obtained from Midwest Scientific. T4 polynucleotide kinase was purchased from New England Biolabs and [γ-^{32}P]ATP was from ICN Radiochemicals. All other biochemicals were from Sigma.

Protein preparation

Human AGT, with wild-type sequence except for a C-terminal $(His)_6$-tag replacing residues 202–207, was encoded on plasmid pQE-hAGT (18). Protein was expressed in XL1-Blue *Escherichia coli* (Stratagene) and purified by Talon® chromatography as described (18). Minor protein contaminants were removed, and the protein was transferred into storage buffer (20 mM Tris (pH 8.0 at 20°C), 250 mM NaCl and 1 mM DTT) by chromatography on Sephadex G-50. Protein solutions were stored frozen at −80°C until needed. AGT concentrations were measured spectrophotometrically using $\varepsilon_{280} = 2.64 \times 10^4\,M^{-1}cm^{-1}$ (17). The samples of AGT used here were >95% active in DNA binding (16) and in repair of short DNAs containing O^6-methylguanine lesions (Supplementary Figure S1).

Table 1. Oligodeoxyribonucleotides used in this study

Oligo	Sequence	Purpose
1	5′-TTT GCA AGC AGC AGA TTA CG-3′	PCR primer
2	5′-CCA GAA ACG CTG GTG AAA GT-3′	PCR primer
3	5′-(2AP)GT CAG TCA GTC AGT C-3′	Binding substrate
4	5′-AGT CAG TC(2AP) GTC AGT C-3′	Binding substrate
5	5′-GAC TGA CTG ACT GAC T-3′	Binding substrate
6	5′-AGT CAG TCA GTC AGT C-3′	Binding substrate
7	5′-GGG TCA TTT GGC GCC TTT CGA TCC-3′ 123 456	Repair substrate[a]
8	3′-CCC AGT AAA CCG CGG AAA GCT AGG-5′	Repair substrate

[a]Residue numbers refer to the NarI recognition sequence. When residue 2 is O^6-methylguanine, the duplex of oligos 7 and 8 is refractory to cleavage.

DNA substrates

Synthetic DNAs (Table 1) were obtained from Invitrogen. Where needed, 5′-labeling with ^{32}P was performed by a standard method (27). Unincorporated [γ-^{32}P] ATP was removed by buffer exchange using Sephadex G-10 centrifuge columns (GE Healthcare) pre-equilibrated with 10 mM Tris and 50 mM KCl (pH 8.0 at 20°C). Plasmid pUC19 obtained from New England Biolabs was linearized by digestion with EcoR1 endonuclease. A 1000-bp fragment of pUC19 DNA (spanning residues 1414–2414) was obtained by polymerase chain reaction (PCR) amplification using primers 1 and 2 (28) and relaxed-circular pUC19 as template. The linear 1000-bp fragment was purified by agarose gel electrophoresis and recovered using the PCR clean-up kit from Qiagen. Stock DNA concentrations were measured spectrophotometrically, using $\varepsilon_{260} = 9.46 \times 10^3\,M^{-1}cm^{-1}$ (per base) for single-stranded DNA (ssDNA) and $\varepsilon_{260} = 1.31 \times 10^4\,M^{-1}cm^{-1}$ (per base pair) for double-stranded DNA (dsDNA).

Sedimentation equilibrium analysis

Human AGT protein and duplex DNAs were dialyzed against 10 mM Tris (pH 7.6 at 20°C), 1 mM DTT, 1 mM EDTA and 100 mM NaCl. Analytical ultracentrifugation was performed at 20°C in a Beckman XL-A centrifuge using an AN60Ti rotor. Scans were obtained at 260 nm. Equilibration was considered complete when scans taken 6 h apart were indistinguishable. At equilibrium, five scans were averaged for each sample at each rotor speed. For large DNAs, the concentration of protein-free DNA molecules becomes negligible long before all available binding sites are saturated. In such systems, one observes mixtures

of free protein and protein–DNA complex, with the weight-average molecular weight of the complex increasing smoothly until saturation is approached. These systems were analyzed with equation 1.

$$A(r) = \alpha_P \exp[\sigma_P(r^2 - r_o^2)] + \alpha_{P_nD} \exp[\sigma_{P_nD}(r^2 - r_o^2)] + \varepsilon. \quad (1)$$

Here, $A(r)$ is the absorbance at radial position r, α_P and α_{P_nD} are absorbances of protein and protein–DNA complex at the reference position, r_o, and ε is a baseline offset that accounts for radial position-independent differences in the absorbances of different cell assemblies. The reduced molecular weights of AGT protein and protein–DNA complexes are given by $\sigma_P = M_P(1 - \bar{\nu}_P\rho)\omega^2/(2RT)$ and $\sigma_{P_nD} = (nM_P + M_D)(1 - \bar{\nu}_{P_nD}\rho)\omega^2/(2RT)$. Here, M_P and M_D are the molecular weights of protein and DNA, n is the protein:DNA ratio of the complex, ρ the solvent density, ω the rotor angular velocity, R the gas constant and T the temperature (Kelvin). The partial specific volume of AGT and the density of sample buffer were calculated using the public-domain program SEDNTERP (available from http://www.rasmb.bbri.org/.) (29). The partial specific volume of duplex NaDNA at 0.1 M NaCl ($\bar{\nu}_D = 0.55$ ml/g) was estimated by interpolation of the data of Cohen and Eisenberg (30). Partial specific volumes of protein–DNA complexes were calculated using equation 2.

$$\bar{\nu}_{P_nD} = \frac{(nM_P\bar{\nu}_P + M_D\bar{\nu}_D)}{nM_P + M_D}. \quad (2)$$

Fluorescence analyses

Steady-state fluorescence measurements were acquired using a Perkin-Elmer LS55 spectrofluorometer. Sample temperatures were maintained at 20°C. Emission spectra (325–500 nm) were recorded with excitation wavelength $\lambda_{ex} = 320$ nm and excitation and emission bandwidths of 4 and 6 nm, respectively.

Atomic force microscopy

AGT–DNA complexes were incubated at protein concentrations of 150 nM to 29 μM and DNA concentrations of 20–480 nM DNA. Incubations were performed for 30 min at ambient temperature in atomic force microscopy (AFM) incubation buffer (10 mM Tris–HCl (pH 7.6 at 20°C), 100 mM NaCl, 1 mM EDTA and 1 mM DTT). The DNAs used in these experiments were linearized pUC19 plasmid and a 1000-bp PCR fragment of pUC19 (see above). Prior to incubation with AGT for AFM imaging, the DNA was heated to 65°C for 10 min to remove potential salt crystals from the DNA. For deposition, AGT–DNA samples were diluted 40- to 300-fold in AFM deposition buffer (25 mM HEPES (pH 7.5 at 20°C), 25 mM sodium acetate and 10 mM magnesium acetate). Immediately after dilution, 20 μl volumes were deposited on freshly cleaved mica (Grade V; SPI Supplies), rinsed with purified deionized water, dried in a stream of nitrogen and imaged using a molecular force probe 3D-Bio AFM (Asylum Research, Santa Barbara, CA). Imaging was performed in oscillating mode using OMCL-AC240TS (Olympus) non-contact/tapping mode silicon probes with spring constants of ~2 N/m and resonance frequencies of ~70 kHz. Images were captured at scan sizes of 2×2 μm and 1×1 μm, at scan speeds of 1.5–2.5 μm/s and at pixel resolutions of 512×512 and 1024×1024.

AFM data analysis

Measurements of DNA-bound protein segments, DNA bend angles and cluster distributions on DNA were performed using the program 'ImageJ' (http://rsbweb.nih.gov/ij/). AGT cluster lengths were measured along the DNA axis. Since AFM involves a mechanical scanning procedure, the resulting image represents a convolution of AFM tip and sample topography. To evaluate the effect of finite tip radii on measured dimensions, the tip radius r was estimated with a simple geometrical model, using the diameter of unoccupied DNA segments as a calibration standard (described in the Supplementary Data and Figure S2). Values of r measured in this way agreed well with tip radii measured by electron microscopy (Tessmer, unpublished data); these values were used as parameters in the same geometrical model to calculate 'corrected' cluster lengths from their uncorrected dimensions. Although simple and direct, this approach comes with a caveat. The DNA in air-dried AFM samples retains a tightly bound hydration layer that increases its apparent diameter (31,32). Thus, assuming the diameter of DNA is 2 nm (33,34) overestimates r and gives lower limit estimates of cluster dimensions. Similarly, because $r > 0$, the uncorrected values give upper limit estimates of cluster dimensions.

DNA bend angles were determined by measuring the angle between line segments drawn on the DNA contour on both sides of a protein cluster. Only isolated AGT clusters (with distances to their nearest neighboring clusters of ≥ 100 nm) were used in this analysis to avoid possible interference by adjacent clusters.

Statistics on protein cluster locations were obtained using the 1000-bp DNA fragment. Cluster locations were quantified using the length along the DNA contour between the center of a protein cluster and the DNA ends. As the unmodified DNA used in these experiments did not allow us to distinguish DNA ends, locations are reported in units of fractional DNA length ranging from 0% (at a DNA end) to 50% (at the DNA center). This interval was subdivided into 'bins', each representing a section containing 5% of the DNA length. From the ratio of total occupancy for a specific DNA site (A_{spec}, for instance DNA fragment ends) and that of non-specific DNA sites (i.e. all other positions on the DNA, A_{nsp}), the binding specificity S for this site can be calculated using equation 3 (35). In our analysis, we define A_{spec} as the fraction of binding occurrences at DNA ends (with 50 bp range as given by the bin size): $A_{spec} = n_{end}/n_{total}$. Background binding frequency is given by the average fraction of binding occurrences from the remaining bins ($\text{avg}(n_{int})/n_{total}$ from 5 to 50% of DNA length). A_{nsp} is the

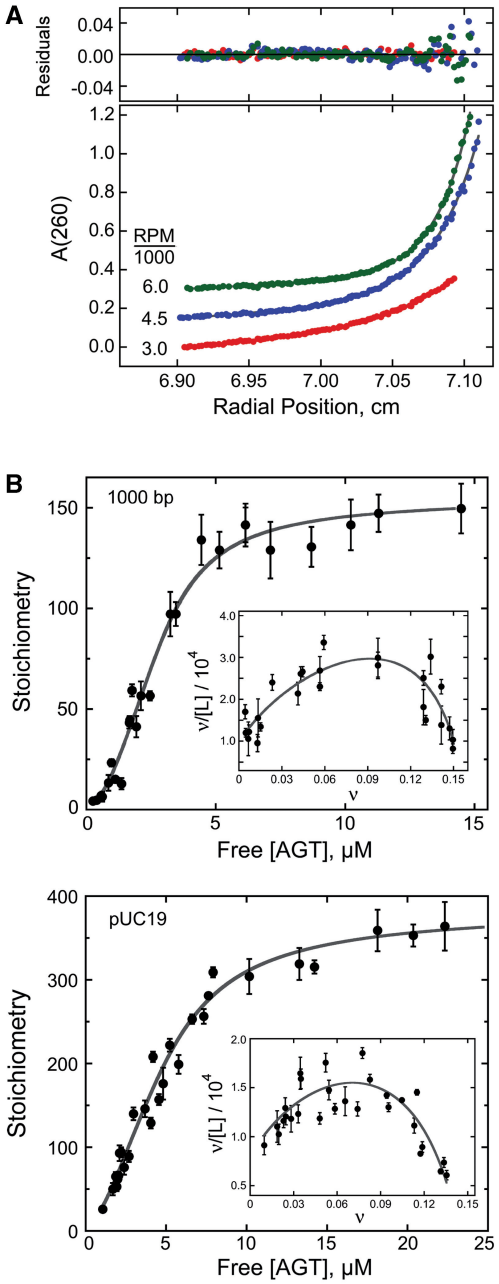

Figure 1. Analysis of AGT–DNA interactions by analytical ultracentrifugation. (**A**) Sedimentation equilibrium data for an AGT–DNA mixture obtained at 20 ± 0.1°C. This sample contained the 1000-bp DNA fragment (0.015 μM) and AGT protein (3.5 μM) in 10 mM Tris (pH 7.6 at 20°C), 1 mM EDTA, 100 mM NaCl and 1 mM DTT. Radial scans taken at 3000 rpm (red), 4500 rpm (blue) and 6000 rpm (green) are shown with vertical offsets for clarity. The smooth curves correspond to fits of equation 1 to these data. Small, symmetrically distributed residuals (upper panel) indicate that the two-species model represented by equation 1 is consistent with the mass distributions of DNA in these samples. (**B**) Dependence of binding stoichiometry on free AGT concentration for the 1000-bp fragment (upper panel) and linear pUC19 DNA (lower panel). Stoichiometries were inferred from weight-average molecular weights measured at sedimentation equilibrium. Error bars are 95% confidence limits for the individual parameters. The smooth curve is an isotherm calculated with equation 4 using parameters determined from the Scatchard plots shown in the insets. Insets: Scatchard plots for the data ensembles shown in the main panels. The solid curves are fits of equation 4, returning $K = 9667 \pm 1499$, $\omega = 35.9 \pm 6.8$ and $s = 6.32 \pm 0.12$ for binding the 1000-bp fragment and $K = 7960 \pm 916$, $\omega = 44.2 \pm 3.8$ and $s = 6.81 \pm 0.14$ for binding linear pUC19 DNA.

product of this average and the number of bins $A_{nsp} = \sum i \times \text{avg}(n_{int})/n_{total}$.

$$S = N\frac{A_{spec}}{A_{nsp}} + 1 \quad (3)$$

where N is the number of total available non-specific-binding sites (total number of DNA base pairs excluding ends, $N = 998$).

RESULTS

Cooperative binding by AGT and predicted cluster size

Two DNAs were used as binding substrates in these studies: linearized pUC19 DNA (2686 bp (36)) and a 1000-bp fragment derived from pUC19 by PCR. AGT–DNA mixtures were analyzed at sedimentation equilibrium (Figure 1A) and weight-average binding stoichiometries (n) were calculated from the reduced molecular weights of AGT–DNA complexes, as described in 'Materials and Methods'. Shown in Figure 1B are graphs of stoichiometry as functions of free [AGT] for the two DNA substrates. The large stoichiometries found in the limit of high [AGT] (~350 for pUC19 and ~150 for the 1000-bp fragment) are striking, although they correspond to smaller binding densities than those obtained with short duplexes (see below). Scatchard analyses are shown in the figure insets; the solid curves are fits of the infinite lattice form of the McGhee–von Hippel equation (equation 4)

$$\frac{\nu}{[P]} = K(1-s\nu)\left(\frac{(2\omega-1)(1-s\nu)+\nu-R}{2(\omega-1)(1-s\nu)}\right)^{n-1} \times \left(\frac{1-(s+1)\nu+R}{2(1-s\nu)}\right)^{2} \quad (4)$$

$$R = ((1-(s+1)\nu)^2 + 4\omega\nu(1-s\nu))^{1/2},$$

in which $\nu = n/(\text{number of base pairs per DNA})$, $[P] = [P]_{input} - n[\text{DNA}]_{input}$, K is the equilibrium association constant for binding a single site, ω is the cooperativity parameter and s is the number of base pairs per bound protein (an effective binding site size) (37). These analyses returned $K = 7960 \pm 916\,\text{M}^{-1}$, $\omega = 44.2 \pm 3.8$ and $s = 6.81 \pm 0.14$ for binding linear pUC19 and $K = 9667 \pm 1499\,\text{M}^{-1}$, $\omega = 35.9 \pm 6.8$ and $s = 6.32 \pm 0.12$ for binding the 1000-bp fragment. These values of K and ω are slightly smaller than those previously reported for binding to a double-stranded 16-mer oligonucleotide, while limiting binding site sizes are considerably larger (compare ~6.8 bp/protein monomer for linear pUC19 with ~4 bp/protein monomer for the 16-mers (16)). It seems likely that these differences reflect packing inhomogeneities that are expected when the binding substrate is large and which are largely absent near binding saturation, when short substrates are used.

When the DNA substrate is long, positive cooperativity results in 'clusters' of contiguously bound proteins. As might be expected, the distribution of the number of

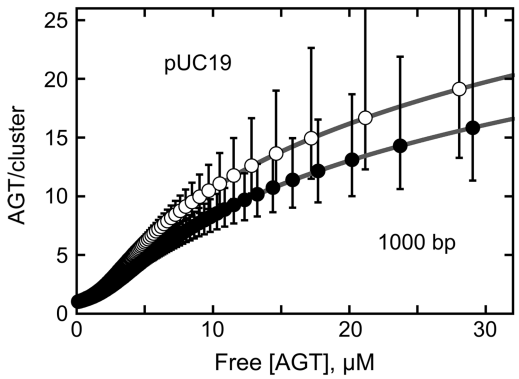

Figure 2. Predicted dependence of the mean cluster size $\bar{C}$ on AGT concentration. Values of $\bar{C}$ were calculated with equation 5 using measured values of ν as functions of [AGT] and values of ω and s from the Scatchard analyses shown in Figure 1. The confidence intervals (bars) were calculated by propagating 95% confidence limits for measured ω and s parameters through equation 5.

proteins in a population of clusters depends on the binding density ν and the values of s and ω at which complexes are formed. For cooperative binding according to the McGhee–von Hippel model (DNA long enough to allow neglect of end effects, cooperative interactions limited to contiguous neighbors and ω-values identical for all cooperative interactions (37)), Kowalczykowski *et al.* (38) showed that the mean cluster size $\bar{C}$ is given by

$$\bar{C} = \frac{2\nu(\omega - 1)}{(s-1)\nu - 1 + R}, \quad (5)$$

with ν, ω, s and R defined as described in equation 4. Shown in Figure 2 are values of $\bar{C}$ calculated as functions of [AGT], using the ν, ω and s parameters obtained from the analyses described above. These results predict a mean cluster size less than two molecules for [AGT] $\leq$1.5 μM; at such low protein concentrations, binding distributions will be dominated by singly bound AGT molecules. In contrast, at [AGT] = 10 μM, these calculations predict mean cluster sizes of ~8.5 and ~11 molecules of AGT for the 1000-bp fragment and pUC19, respectively, and still larger clusters at higher [AGT] (Figure 2) or under solution conditions that increase binding cooperativity (data not shown). These predicted cluster size distributions can be tested using AFM (39,40), as described below.

Measurement of cluster size distributions by AFM

AFM was used to characterize reaction mixtures containing AGT and long duplex DNAs (the 1000-bp fragment or linearized pUC19; Figure 3). These images contain particles with volumes consistent with that of monomeric AGT (Supplementary Figure S3) and elongated structures with contour lengths like those of the naked DNAs. Distributed along the contours of these structures are distinct segments with diameters and heights that are greater than that of naked DNA (Figure 3B arrows). The lengths of these segments along the DNA contour are greater than the diameter of free AGT protein (Supplementary Figure S3), consistent with the notion that they contain more than one protein molecule. The occupied segments are of relatively uniform length separated by gaps of varying lengths (Figure 3B and C). Segment length distributions (Figure 4A) were fitted using a Gaussian function to obtain mean lengths as functions of [AGT]. Mean values were corrected for the finite size of the AFM tips as described in 'Materials and Methods'. Assuming a binding density within the cooperative complex of 1 protein/4 bp, as seen with short fragments (16,17), these segment lengths correspond to ~3 AGT molecules/cluster at [AGT] ~2 μM and ~6 AGT molecules/cluster for [AGT] >10 μM (Figure 4B). Uncorrected lengths were taken as upper limit estimates as these will always overestimate cluster size. These ranged from ~6 molecules/cluster at [AGT] ~2 μM to ~8 molecules/cluster for [AGT] >10 μM. Similar results were obtained with the full-length pUC19 linear fragments, consistent with the absence of a strong dependence on DNA length or sequence. These cluster sizes coincide with the predictions of the McGhee–von Hippel binding model for [AGT] $\leq$10 μM, but they are significantly below the predicted range at higher AGT concentrations. Mechanisms that might limit cluster-size distributions are considered below.

DNA bending in clusters

The available crystal structures of AGT–DNA complexes show that the DNA is bent 15–30° toward the major groove (3,18,19). These structures contain single protein molecules bound to DNA and do not reflect any geometric contributions made by the protein–protein interactions that stabilize non-specific cooperative binding. To gain insight on protein–DNA structures induced by such cooperative AGT clusters, we measured DNA bend angles at the sites of complexes (see 'Materials and Methods' section). In the absence of protein, occasional peaks can still be observed on the DNA. These peaks occur at much lower frequency than those seen in the presence of AGT (Figure 5A, see also Figure 6) and are likely due to non-specific salt contamination. As these peaks cannot be unambiguously distinguished from protein peaks on the DNA, we also measured the bend angles at these positions for DNA deposited in the absence of AGT (gray bars in Figure 5A). The DNA bend-angle distribution at peaks found in the absence of protein can be fit by a broad, low-intensity half-Gaussian centered at ~0° (with width ± 90° and R^2 ~ 0.85, gray line in Figure 5A). In contrast, the bend-angle distribution at the site of bound AGT clusters was triphasic (Figure 5A). Gaussian decomposition returned maxima with angles of 0 ± 13° (~20% of the population), 27 ± 9° (~20%) and 58 ± 29° (~60%). When the population of complexes is subdivided by length (Figure 5B), the bend angles characteristic of short complexes ($\leq$5 proteins, mean bend angles of 0 ± 10°, 26 ± 13° and 60 ± 27°) were different from those of long complexes (>6 proteins, mean bend = 50 ± 35°). This trend is consistent with models in which the observed bend angle depends on the number of proteins present in the cooperative complex. A geometrical model that accounts for the distribution of DNA bend angles associated with AGT clusters is proposed below.

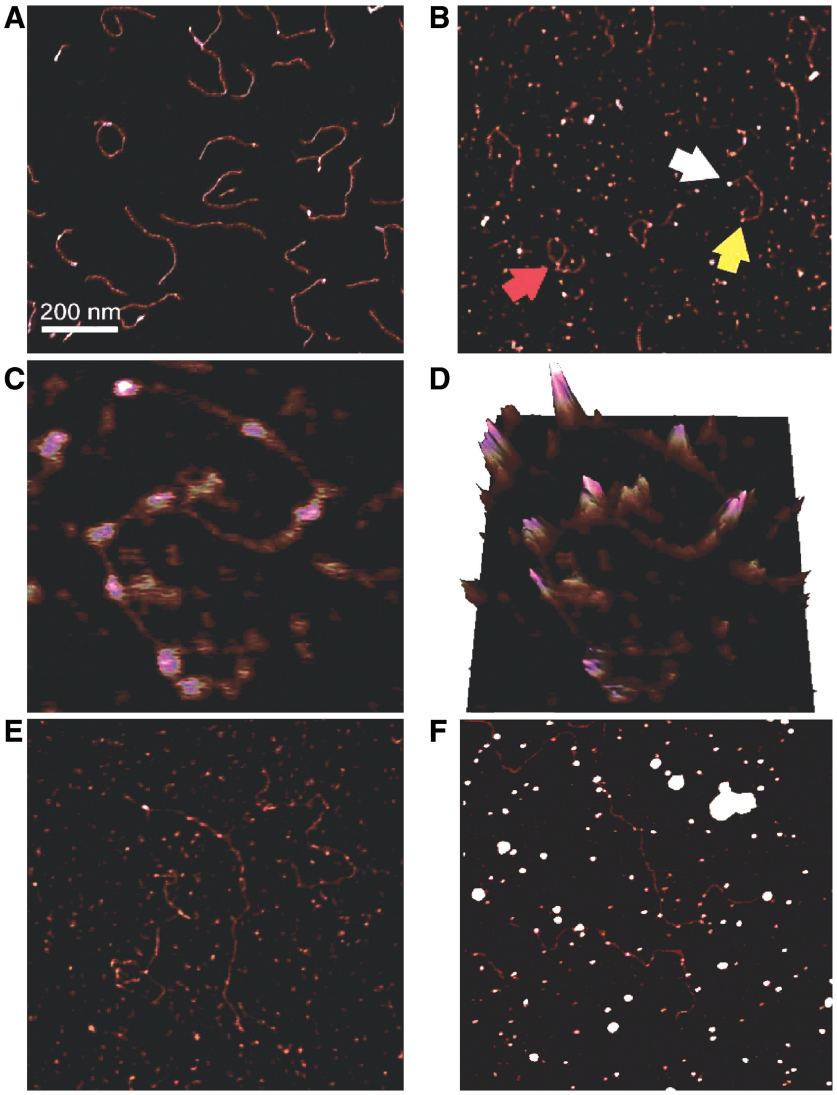

Figure 3. Visualization of AGT cooperative units on DNA. AFM images of the 1000-bp DNA fragment in the absence of protein (**A**) and after incubation with AGT at 6 μM at a protein:DNA ratio of 100:1 (**B**) or 12 μM AGT and protein:DNA ratio of 200:1 (**C**) show increasing numbers of AGT clusters on the DNA with increasing protein concentration. A three-dimensional projection of the data from (C) is shown in (**D**). Results are not dependent on the DNA substrate used: the image in (**E**) shows similar results for a linearized pUC19 plasmid substrate after incubation with 12 μM AGT. To test whether protein distributions were affected by the deposition process, samples containing linear pUC19 (60 nM) and AGT (12 μM) were crosslinked with glutaraldehyde (0.1% glutaraldehyde, 10 min at 37°C) and then applied to the mica substrate (**F**). Sodium dodecyl sulfate–polyacrylamide gel electrophoresis analysis (Supplementary Figure S4) showed that >50% of AGT molecules were crosslinked to a neighbor by this treatment. Arrows in (B) indicate AGT clusters on the DNA fragments. Images are 1 μm × 1 μm (A, B, E, F) and 170 nm × 170 nm (C, D).

Preferential binding near DNA ends

Distributions of AGT clusters along the DNA contour were obtained by measuring the contour lengths between cluster centers and DNA ends. A graph of occupation frequency as a function of position on the DNA showed no obvious preference for any internal position (Figure 6), consistent with previous reports that AGT binds DNA with little sequence specificity (14,41). However, the analysis showed a clear preference for binding DNA ends (an example is indicated by the white arrow in Figure 3B). We calculated binding specificities for DNA fragment ends from the relative frequencies of end binding and internal binding, using equation 3 (35). Data for a range of protein:DNA ratios (50:1 to 200:1) and concentrations ranging from 2 to 12 μM (Figure 6) give an average preference for fragment ends over internal sites of 258 ± 156 (from $n = 7$ independent experiments). Although a crystal structure has been obtained in which AGT bridges between two adjacent, stacked DNA ends (19), this is the first evidence, to our knowledge, that AGT binds to individual (unstacked) DNA ends with elevated affinity in free solution. Evidence suggesting a mechanism for this enhanced affinity is discussed below.

Enhanced base-flipping activity near DNA ends

As part of its repair mechanism, AGT flips bases out of the stacked conformations of free DNA and into its active site cleft (18,42). This process may be facilitated by the

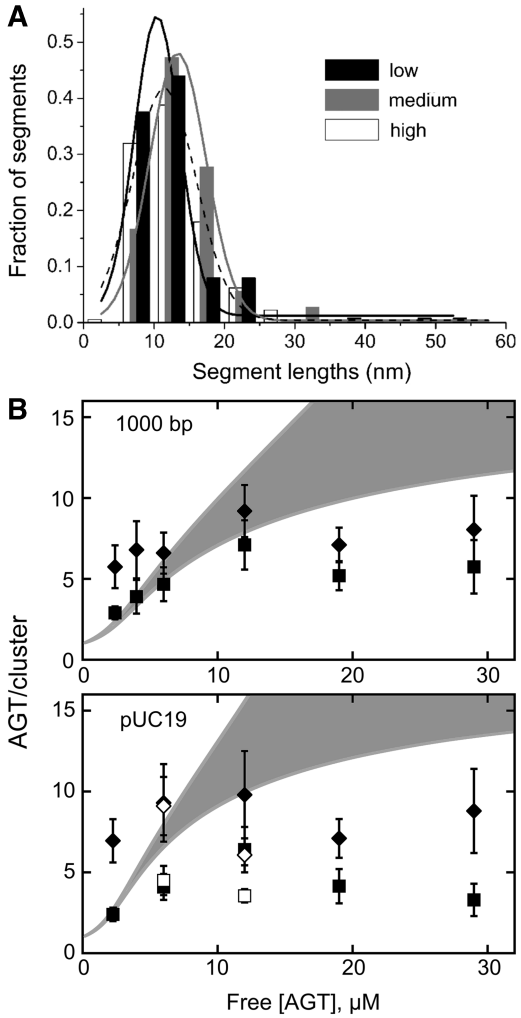

Figure 4. Comparison of measured and theoretical cluster length distributions. (**A**) Distribution of AGT cluster lengths on 1000-bp DNA fragments for incubations at 2 μM (black, $n = 125$), 6 μM (gray, $n = 36$) and 12 μM (white, $n = 178$) AGT. For comparison, data are shown as fractions of the total number of complexes for each protein concentration. Gaussian fits to the distributions give comparable cluster lengths of 10, 13 and 11 nm for 2 μM (black line), 6 μM (gray line) and 12 μM AGT (dashed line), respectively (all fits were characterized by $R^2 \geq 0.97$). (**B**) Comparison of measured cluster sizes, expressed as protein molecules/cluster (symbols), with cluster size predictions from the McGhee–von Hippel model (gray zones). Two values are given for each set of measurements; r_{tip}-corrected values (filled square) as lower limit estimates of the number of AGT monomers per cluster and uncorrected values (filled diamond) as upper limit estimates. Open symbols (open square, open diamond) denote measurements made after glutaraldehyde crosslinking. Error bars give the standard deviations of each sample population. Ranges for $\bar{C}$ predicted for the McGhee–von Hippel binding model (gray zones) were calculated with equation 5, using experimental values of n, ω and s and corresponding error ranges as determined in Figure 1B.

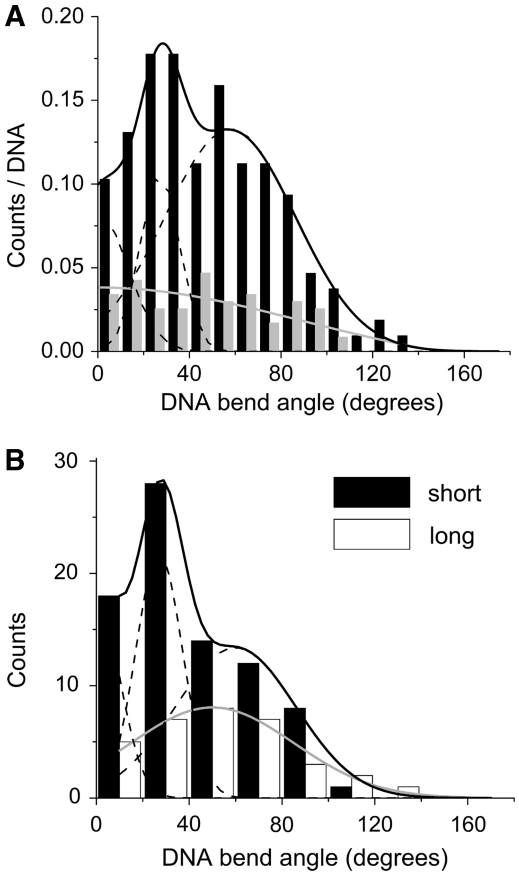

Figure 5. DNA bending associated with AGT clusters. (**A**) DNA bend-angle distributions were measured in the absence (gray, $n = 75$) or presence of AGT (black, $n = 139$). The dashed lines show the result of a Gaussian decomposition for AGT-induced bending that gives three peaks, centered around 0 ± 13°, 27 ± 9° and 58 ± 29° ($R^2 > 0.97$). In the absence of AGT, bend angles measured at salt contaminant peaks on the DNA show a broad distribution centered at 0°. In comparison, DNA bend angles measured at random positions along the DNA backbone in the absence of protein also display mean bending of 0° but with narrower distribution width (Supplementary Figure S5). (**B**) DNA bend-angle distributions separated for different cluster lengths: short complexes (≤5 proteins, black bars) and long complexes (>6 proteins; white bars). Gaussian fits to the short cluster population suggest mean bend angles of 0 ± 10°, 26 ± 13° and 60 ± 27° (black line with dashed lines showing the three Gaussian components; $R^2 > 0.99$, $n = 81$) and for the longer complexes a broad distribution centered at a mean bend angle of (50 ± 35) ° (light gray line; $R^2 > 0.97$, $n = 33$).

transient loss of base pairing and stacking (breathing) that takes place at DNA ends (25,43). We used 2-aminopurine (2AP)-substituted DNAs to test whether AGT is more effective at inducing extrahelical base conformations in the centers of ssDNA and duplex DNAs or near their ends. Reduction in base stacking can be detected as an increase in the fluorescence quantum yield of 2AP (44–46). Shown in Figure 7A are emission spectra for a single-stranded 16-mer containing a single 2AP at its 5′-end (oligo 3, Table 1), as a function of [AGT]. The emission maxima at 369 nm are similar to values reported for other 2AP-labeled DNAs (47,48), while the intensity increase with AGT binding is like that seen with other proteins that stabilize extrahelical base conformations in DNA (49,50).

Addition of AGT resulted in similar fluorescence increases from the ssDNA and from a duplex 16-mer with a 2AP residue located at the 5′-end of one strand (Figure 7B). Saturation of this effect required a somewhat greater AGT concentration for duplex DNA than for the single-stranded substrate (~15 μM for dsDNA and ~10 μM for ssDNA), possibly reflecting a

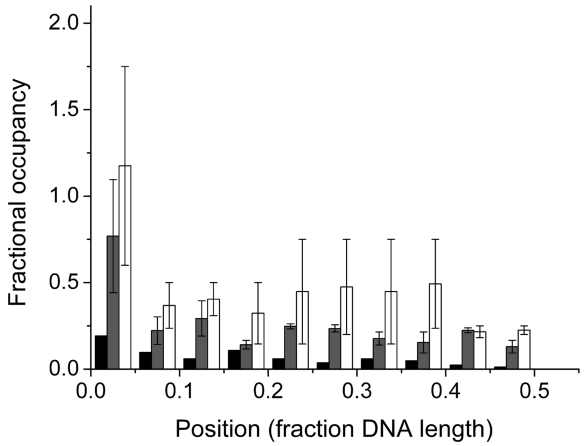

Figure 6. Distribution of AGT clusters along the DNA contour. Fractional occupancies for 50-bp-long sections of the 1000-bp DNA fragment, demonstrating increasing protein coverage of the DNA for increasing AGT concentrations as well as preferential DNA end binding. Because the unmodified DNA ends used in these experiments could not be distinguished, locations are reported in units of fractional DNA length ranging from 0% (at a DNA end) to 50% (at the DNA center). Low and high [AGT] correspond to incubations at 2 μM AGT (gray, 2.6 ± 0.3 peaks per DNA, $n = 220$) and 12 μM AGT (white, 4.6 ± 2.2 peaks per DNA, $n = 159$), respectively. Error bars represent the deviations of two independent experiments. For comparison, DNA in the absence of protein is shown (black, 0.7 peaks per DNA, $n = 54$ from one experiment). Peaks on the DNA in the absence of AGT likely stem from salt contaminations but were treated like AGT segments in the analyses because they are difficult to distinguish from protein peaks and are likely also present in samples after incubation with AGT.

modest difference in binding affinity. Increased fluorescence was also seen when ssDNA containing an internal 2AP residue was titrated with AGT. In marked contrast to these results, little increase in emission was detected when AGT was added to internally labeled, dsDNA (Figure 7C). Parallel CD measurements showed that both end-labeled and internally labeled dsDNAs were bound to similar extents at saturation (~4 AGT/DNA; Figure 7D), so the striking differences in emission intensities do not reflect differences in the limiting equilibrium binding density. One interpretation consistent with these results is that AGT promotes extrahelical base conformation(s) more readily at the end of a duplex DNA than it does at internal sites. Since contacts with the extrahelical base stabilize AGT binding, we propose that differences in the stabilities of helical base conformations at internal sites and DNA ends account for the preferential end binding found with AGT. If correct, this result predicts that AGT's binding and repair functions may be enhanced, not just at DNA ends, but wherever base stacking or hydrogen bonding is impaired.

DISCUSSION

The currently available crystal structures of AGT–DNA complexes show single protein molecules bound at isolated DNA sites (3,18,19). In contrast, solution binding and crosslinking studies indicate that AGT forms cooperative complexes on single- and double-stranded substrates (15,16,52) Cooperative binding is also seen with structurally similar alkyltransferase-like proteins, which have been proposed to bridge between direct base repair and nucleotide excision repair pathways (51). Mutations located in the protein–protein interface and far from the DNA-contact surface have strong effects on binding cooperativity *in vitro* and resistance to alkylating agents *in vivo* (23), suggesting that DNA-binding cooperativity plays a role in the physiological activities of the enzyme. To better understand the cooperative binding mechanism, we used ultracentrifugation and AFM to characterize the binding of AGT to 1000 and 2686 bp duplex DNAs. Binding affinities found for these DNAs were similar to ones previously measured, in the same buffer, for short (11–41 bp) duplexes (~2.8 μM) (16). This is consistent with the idea that protein–DNA and protein–protein interactions are similar on short and long substrates. On the other hand, the limiting binding site sizes were significantly greater on long DNAs than on short duplexes (~6.8 bp/protein monomer for linear pUC19 but ~4 bp/protein monomer for a duplex 16-mer (16)). It seems likely that these differences reflect packing inhomogeneities that are expected when the binding substrate is large and which are largely absent when short substrates are used.

The McGhee–von Hippel binding model is often used to characterize cooperative protein–DNA interactions (37,38). This model envisions a single binding mode characterized by a unique binding site size (s), a single class of protein–DNA contacts with a single equilibrium constant K (i.e. binding is not sequence- or DNA structure dependent) and a single class of cooperative interactions (the value of ω for a singly contiguous interaction is the same as each ω-value for a protein engaged in two or more interactions, and no other DNA sequence or structural factors influence the value of ω). Although this model may not capture all features of complicated biological systems, it provides a useful benchmark for the comparison of their properties.

In our AFM studies, AGT gave average cluster sizes that agree well with predictions of the McGhee–von Hippel model at low binding occupancies, but at high occupancies cluster sizes fell below the predicted range (Figure 4B). The presence of AGT molecules defective in protein–protein interaction could have limited the size of cooperative clusters. However, our preparations were >95% active in cooperative DNA binding (16) and in alkyltransferase reactions that require cooperative binding (Supplementary Figure S1). If the addition of an inactive AGT molecule ends cluster growth and if the probability of adding an inactive molecule is 0.05/step, the probability of terminating a cluster at length $\leq n$, $F(n) = 1 - (0.95)^{n-1}$. Under these conditions, the median cluster size (the number of steps required for $F(n)$ to reach 0.5) would be 15. As this is significantly greater than the cluster sizes that were detected by AFM, we conclude that the small fraction of inactive AGT molecules in our samples was unlikely to be the dominant factor limiting cluster size.

Cluster size could be limited by dissociation of protein or isomerization of complexes during deposition on the mica substrate for AFM analysis. To test this possibility,

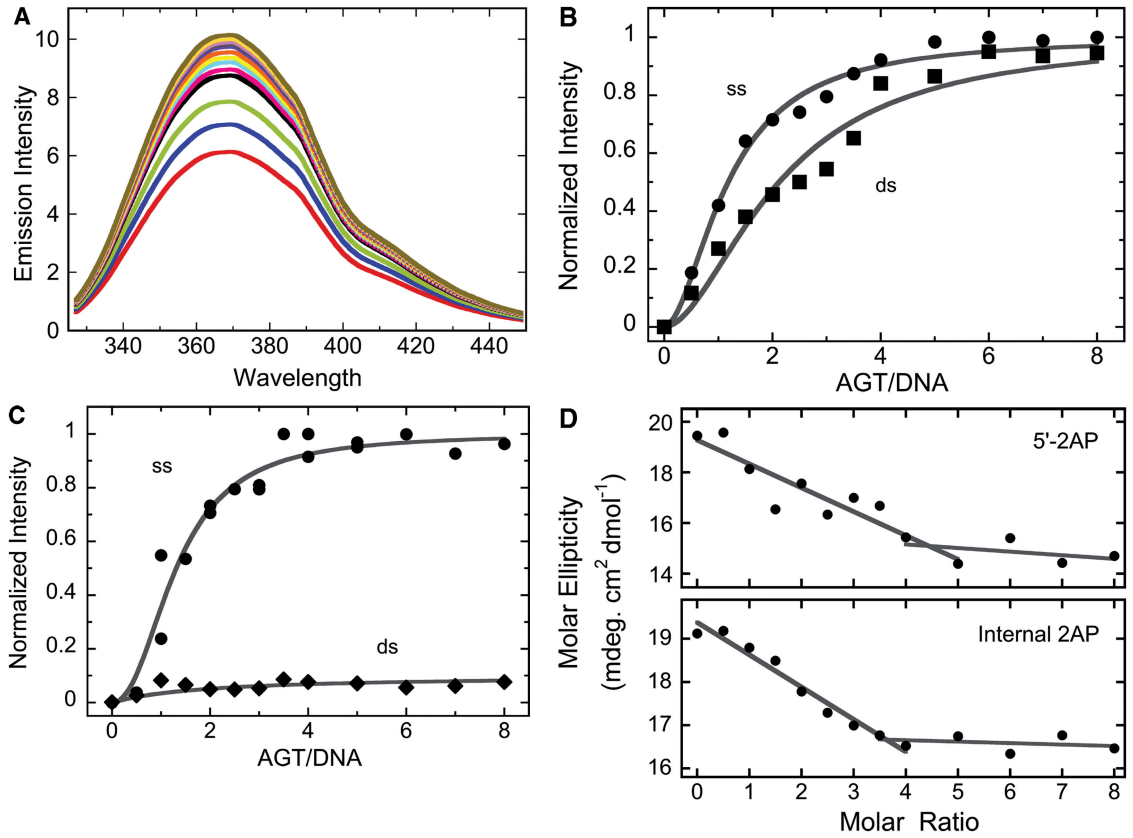

Figure 7. Base-flipping detected by 2AP fluorescence. (**A**) Steady-state emission spectra for a ssDNA containing a 5′-terminal 2AP residue and mixtures of this DNA with AGT. Solutions contained oligo 3 DNA (5.6 μM) and 0–44.9 μM AGT in 10 mM Tris, (pH 7.6 at 20°C), 1 mM EDTA, 100 mM NaCl and 1 mM DTT. Measurements were made at 20°C with excitation at 325 nm. Spectra, in order of lowest curve to highest, are for solutions containing AGT at final concentrations of 0, 2.8, 5.6, 8.4, 11.2, 14, 16.8, 19.7, 22.5, 28.1, 33.7, 39.3 and 44.9 μM, respectively. (**B**) Normalized emission intensities as functions of [AGT]/[DNA] ratio, for solutions containing single-stranded or duplex DNAs labeled with a 5′-terminal 2AP residue. DNAs were oligo 3 (5.6 μM) or the duplex containing oligos 3 and 5 (5.3 μM). Buffer and spectroscopy conditions were as described for Panel A. Labels ss and ds denote ssDNAs and dsDNAs, respectively. (**C**) Normalized emission intensities as functions of [AGT]/[DNA] ratio, for solutions containing single-stranded or duplex DNAs labeled at an internal position with a 2AP residue. DNAs were oligo 4 (5.2 μM) or the duplex containing oligos 4 and 5 (5.4 μM). Buffer and spectroscopy conditions were as described for Panel A. (**D**) Verification of binding by circular dichroism. Samples containing duplex DNA (aliquots of solutions used in the fluorescence experiments shown in Panels B and C) were subjected to CD spectroscopy. Molar ellipticity values (per nucleotide) at 260 nm are graphed as functions of [AGT]/[DNA] ratios for double-stranded 16-mers labeled with 2AP either internally or at the 5′-end. The lines are linear fits to subsets of the data. Break points at [AGT]/DNA] ~4 agree with previous stoichiometry measures for binding duplex 16-mer DNAs (16,17).

we used glutaraldehyde to covalently stabilize intermolecular contacts and then measured the dimensions of AGT clusters. Figure 3F shows representative AGT–DNA complexes after crosslinking by glutaraldehyde. As demonstrated in Supplementary Figure S4, more than half the AGT monomers were crosslinked to a neighbor by this treatment. Although some large spherical particles result from this treatment, individual DNA molecules decorated with clusters like those present in uncrosslinked samples are also abundant. As shown in Figure 4B, the distributions of cluster sizes for crosslinked complexes were similar to ones obtained without the crosslinker. This argues that protein rearrangements, if they occur during deposition, do not change the distribution of cluster sizes. A simpler interpretation, and one that we favor, is that protein rearrangements are slow compared with the deposition process and that deposition captures a close approximation of the distribution of proteins on DNA that prevailed in the sample before deposition.

A third mechanism that could limit cluster size is one in which protein, DNA or both are deformed as each AGT molecule is added to the cooperative unit. When ΔG(deformation) = $-\Delta G$(protein–protein interaction), binding should become non-cooperative (i.e. $\omega = 1$) and cluster growth will cease. This is an example of a mechanism that would cause binding cooperativity to change with binding density, in contrast to the uniform cooperativity specified by the McGhee–von Hippel model. In the AGT system, the deformational stress may be exerted on the DNA. Intermolecular contacts maintain spatial relationships between proteins within the cooperative unit, but each protein unwinds duplex DNA by ~7° (52). As a result, torsional stress should accumulate within growing cooperative units as long as protein–protein contacts remain intact. The dependence of ΔG(DNA twist) on the twist angle ϕ (in radians) is given by equation 6, in which C is the DNA torsional Hooke's constant

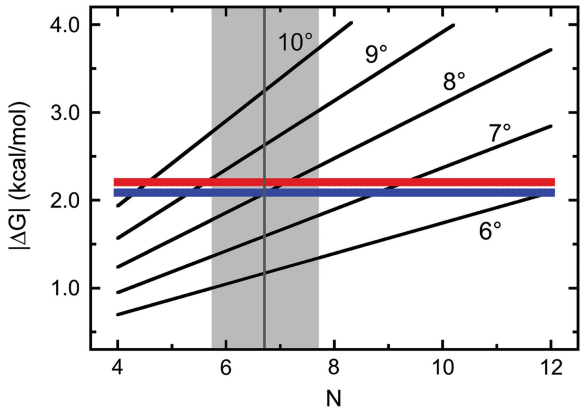

Figure 8. Comparison of predicted torsional free energies (ΔG(twist)) and cooperative free energies (ΔG(cooperative)) for clusters of 4–12 AGT proteins. ΔG is given as its absolute value. The quadratic model (equation 6) was used to calculate the cumulative ΔG(twist) as functions of proteins/cluster (N) for displacements of 6–10°/protein. The red and blue lines give $|\Delta G$(cooperative)| for the addition of a single protein molecule to complexes formed on linear pUC19 DNA and the 1000-bp fragment, respectively. The intersections of these functions indicate where ΔG(twist) = $-\Delta G$(cooperative). A vertical gray line is plotted for a cluster length of 6.7 proteins, the mean of all length estimates for [AGT] ≥ 6 μM, for the 1000-bp DNA. The lighter gray zone spans between means of minimum and maximum estimates of cluster length.

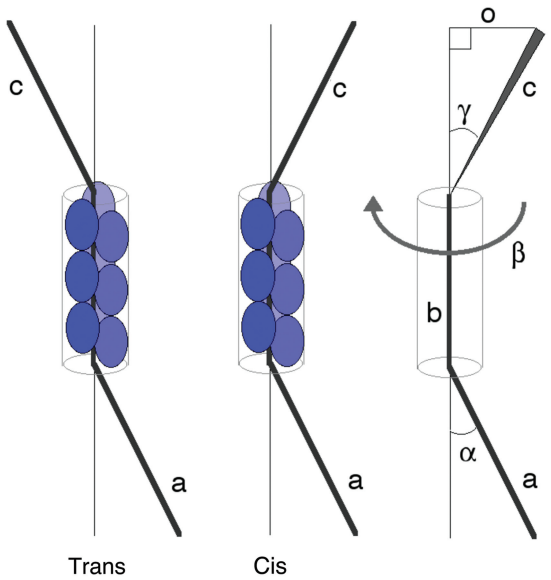

Figure 9. Paired DNA bends produce a triphasic angle distribution. In our model of the cooperative complex, each AGT monomer occupies 4 bp, adding 1.36 nm to the cluster length. Each protein is rotated ~138° with respect to its nearest neighbors; modest DNA bends associated with each protein cause the DNA to follow a gentle writhe about a central axis (51). In this diagram, the blue ovals represent AGT proteins, line segments a and c represent the axes of the free DNA entering and leaving the complex and segment b represents the writhe axis of DNA within the complex. Angles α and γ are the angles formed when the DNA enters and departs from this writhe, while angle β is the dihedral angle, which depends on the number of proteins in the complex. Complexes with dihedral angles near 0° will be captured on the planar AFM substrate in a *cis* arrangement, complexes with dihedral angles near 180° will be captured in a *trans* conformation and complexes in which $0° < \beta < 180°$ are seen in projection. For each complex, the bend angle apparent to AFM is that of a segment with respect to segment c and is thus the sum of angles α and γ or their projections, on the plane of the AFM matrix.

(~3×10^{-19} erg/cm (53,54)) and L is the DNA length over which the twist acts (55)

$$\Delta G_{\text{twist}} = \frac{1}{2} C \left(\frac{\phi}{L}\right)^2 L. \quad (6)$$

Graphs of this function are shown in Figure 8, together with lines indicating the cooperative free energies measured for pUC19 and 1000 bp DNAs. At the intersections of these functions, ΔG(twist) = $-\Delta G$(cooperative). For a cluster size of 6 (typical of the high [AGT] limits for both DNAs), this equivalence point is reached at a net unwinding of ~8.5°/protein, only slightly greater than the value (7.1 ± 0.3°/protein) measured by topoisomerase assay (52). Although this coincidence does not prove that torsional stress limits the size of AGT clusters, it shows that such a mechanism is capable of doing so and might account for the result that AGT clusters are smaller than those predicted by the McGhee–von Hippel model.

In our images, DNA bends often co-localize with AGT clusters. Analysis of bend angles associated with protein clusters revealed a triphasic angle distribution with maxima at (0 ± 13)°, (27 ± 9)° and (58 ± 28)° (Figure 5A). Simple models in which the DNA is bent just once give monophasic angle distributions (characterized by a single mean and standard deviation) that account poorly for the bend-angle distributions of samples containing AGT. On the other hand, a combination of two bends located close together can produce a triphasic angle distribution. Our current model of the AGT cooperative complex has features that might produce a double bend. Well-ordered AGT complexes contain one protein every 4 bp (16,24) in a three-start helical array (52). In this structure, small, uniform DNA bends, repeated every 4 bp, are partially cancelled by helical symmetry, and the DNA describes a gentle writhe about a straight axis. This cancellation is not available at the ends of the complex, so interfaces between occupied and free DNA segments will be bent, forming a dihedral (Figure 9). When such a structure is deposited on a planar surface, three different rotational states may be visible: one with the flanking DNA segments in a *trans* conformation, one with the segments in a *cis* conformation and one with a flanking segment rotated out of (or toward) the plane of the surface. If α is the bend angle at one end of the complex, the *cis* conformer will have an apparent net bend of ~2α and the *trans* conformer will have an apparent net bend of ~0. When one of the flanking segments is not in the plane of the substrate, it will be seen in projection and the sum of the apparent angles will be $\alpha + \gamma$, where $\gamma = \arcsin(\sin\alpha \cos\beta)$ and β is the angle by which the projected segment is rotated out of the plane. Since cosβ is symmetrical with values ranging from −1 to 1, γ will always be ≤α; if all values of β are equally likely, the population average $<\alpha + \gamma>$ will equal α and three maxima will be found at 0, α and 2α°. A model with α ~27° accounts reasonably well for the triphasic distribution observed in AGT–DNA complexes. A bend

of this magnitude lies within the range ($15^\circ \leq$ bend $\leq 30^\circ$) seen in the available crystal structures (18,19) and it strongly supports our helical model of cooperative AGT complexes (52). Taking the measured angular dispersion into account, these results suggest that AGT complexes can accommodate DNA structures with net bends ranging from 0 to $\sim 120^\circ$. Such plasticity may have functional value when the target DNA element lies within structures constrained by chromatin proteins or transcription or replication machinery.

As revealed by AFM and fluorescence data, AGT binds preferentially to DNA ends and has greater base-flipping activity against residues located at ends than in the centers of duplex DNAs (Figures 6 and 7). These effects may be related. Rupture of base pairing and stacking interactions during 'breathing' oscillations may reduce the free energy cost of forming the extrahelical base conformation(s) necessary for lesion search and repair. Breathing oscillations are more frequent at DNA ends than in the centers of correctly base-paired duplexes (25,43) and this may account for the end-binding preference of AGT. If this notion is correct, other structural perturbations that destabilize stacked base pairs (such as mis-pairing or the presence of a bulge or gap dislocation or double-strand break (46)) near the O^6-alkylguanine lesion may increase binding affinity and repair efficiency. In line with this view, one currently available AGT–DNA structure shows the protein spanning the gap where two DNA ends abut; in this structure, the 3′-terminal *T* is unstacked and partially inserted into the active site (19). At present, we know of no evidence that suggests that AGT might play a role in cellular responses to double-strand breaks, so this intriguing possibility remains to be tested.

We still have much to learn about how the structural features of AGT–DNA complexes relate to its functions. The small cooperative cluster size that we have observed ($n \leq 11$ under conditions tested here) may allow cooperative units to function wherever short stretches of free DNA are available. Under close-packing conditions, 11 proteins cover $\sim$44 bp, a length that fits easily into typical inter-nucleosome spacer regions (56,57) and may allow AGT to gain access to DNA in chromatin remodeling complexes or replication complexes. Short complexes are likely to confer kinetic benefits on the lesion search process, particularly if dissociation from cluster ends is more rapid than that from central positions. (Proteins in the center of a linear array have two nearest neighbors and make two cooperative contacts. Those at the ends of a linear array have just one neighbor and their binding is stabilized by only one cooperative contact. This difference may result in quite distinct dissociation rates.) The range of DNA bends imposed by AGT clusters indicates that cooperative complexes can form on substantially bent DNA, giving AGT the potential to occupy DNA segments that are bent or looped by flanking chromatin structures, chromatin remodeling or DNA-replication elements. Finally, AGT binds with highest affinity to DNAs that are only slightly under-wound (52), and data presented here suggest that DNA under-winding may limit the size of the cooperative unit. Together, these results are consistent with the idea that within the cell, AGT may interact preferentially with regions of torsionally relaxed DNA. Since the exposure of such DNA depends on the activities of chromatin remodeling and topoisomerase enzymes, and since these activities are required for DNA replication and transcription, the structures of AGT clusters predict that its activities will be concentrated near replication forks and regions of actively transcribed chromatin. These hypotheses can be tested using currently available technology.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–5, Supplementary Methods and Supplementary References [58,59].

FUNDING

Deutsche Forschungsgemeinschaft [Forschungszentrum FZ82 to I.T.]; National Institutes of Health (NIH) [GM-070662 to M.G.F.]. Funding for open access charge: NIH.

Conflict of interest statement. None declared.

REFERENCES

1. Fang,Q., Kanugula,S., Tubbs,J.L., Tainer,J.A. and Pegg,A.E. (2010) Repair of O^4-alkylthymine by O^6-alkylguanine-DNA alkyltransferases. *J. Biol. Chem.*, **285**, 8185–8195.
2. Shrivastav,N., Li,D. and Essigmann,J.M. (2010) Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation. *Carcinogenesis*, **31**, 59–70.
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 O^6-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. Pegg,A.E. (1990) Mammalian O^6-alkylguanine-DNA alkyltransferase: regulation and importance in response to alkylating carcinogens and therapeutic agents. *Cancer Res.*, **50**, 6119–6129.
6. Gerson,S.L. (2002) Clinical relevance of MGMT in the treatment of cancer. *J. Clin. Oncol.*, **20**, 2388–2399.
7. Margison,G.P. and Santibáñez-Koref,M.F. (2002) O^6-Alkylguanine-DNA alkyltransferase: role in carcinogenesis and chemotherapy. *BioEssays*, **24**, 255–266.
8. Kaina,B., Margison,G.P. and Christmann,M. (2010) Targeting O^6-methylguanine-DNA methyltransferase with specific inhibitors as a strategy in cancer therapy. *Cell. Mol. Life Sci.*, **67**, 3663–3681.
9. Rabik,C.A., Njoku,M.C. and Dolan,M.E. (2006) Inactivation of O^6-alkylguanine DNA alkyltransferase as a means to enhance chemotherapy. *Cancer Treat. Rev.*, **32**, 261–276.
10. Ranson,M., Middleton,M.R., Bridgewater,J., Lee,S.M., Dawson,M., Jowle,D., Halbert,G., Waller,S., McGrath,H., Gumbrell,L. *et al.* (2006) Lomeguatrib, a potent inhibitor of O^6-alkylguanine-DNA-alkyltransferase: phase I safety, pharmacodynamic, and pharmacokinetic trial and evaluation in combination with temozolomide in patients with advanced solid tumors. *Clin. Cancer Res.*, **12**, 1577–1584.
11. Verbeek,B., Southgate,T.D., Gilham,D.E. and Margison,G.P. (2008) O^6-Methylguanine-DNA methyltransferase inactivation and chemotherapy. *Br. Med. Bull*, **85**, 17–33.

12. Dasgupta,A., McCarty,D. and Spencer,H.T. (2010) Engineered drug-resistant immunocompetent cells enhance tumor cell killing during a chemotherapy challenge. *Biochem. Biophys. Res. Commun.*, **391**, 170–175.
13. Maier,P., Spier,I., Laufs,S., Veldwijk,M.R., Fruehauf,S., Wenz,F. and Zeller,W.J. (2010) Chemoprotection of human hematopoietic stem cells by simultaneous lentiviral overexpression of multidrug resistance 1 and O^6-methylguanine-DNA methyltransferase (P140K). *Gene Ther.*, **17**, 389–399.
14. Fried,M.G., Kanugula,S., Bromberg,J.L. and Pegg,A.E. (1996) DNA binding mechanisms of O^6-alkylguanine-DNA alkyltransferase: stoichiometry and effects of DNA base composition and secondary structures on complex stability. *Biochemistry*, **35**, 15295–15301.
15. Rasimas,J.J., Kar,S.R., Pegg,A.E. and Fried,M.G. (2007) Interactions of human O^6-alkylguanine-DNA alkyltransferase (AGT) with short single-stranded DNAs. *J. Biol. Chem.*, **282**, 3357–3366.
16. 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.
17. Rasimas,J.J., Pegg,A.E. and Fried,M.G. (2003) DNA-binding mechanism of O^6-alkylguanine-DNA alkyltransferase. Effects of protein and DNA alkylation on complex stability. *J. Biol. Chem.*, **278**, 7973–7980.
18. Daniels,D.S., Woo,T.T., Luu,K.X., Noll,D.M., Clarke,N.D., 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.
19. 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.
20. Liu,L., Xu-Welliver,M., Kanugula,S. and Pegg,A.E. (2002) Inactivation and degradation of O^6-alkylguanine-DNA alkyltransferase after reaction with nitric oxide. *Cancer Res.*, **62**, 3037–3043.
21. 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.
22. Rasimas,J.J., Kanugula,S., Dalessio,P.M., Ropson,I.J., Fried,M.G. and Pegg,A.E. (2003) Effects of zinc occupancy on human O^6-alkylguanine-DNA alkyltransferase. *Biochemistry*, **42**, 980–990.
23. Adams,C.A. and Fried,M.G. (2011) Mutations that probe the cooperative assembly of O^6-alkylguanine-DNA alkyltransferase complexes. *Biochemistry*, **50**, 1590–1598.
24. Melikishvili,M., Hellman,L.M. and Fried,M.G. (2009) Use of DNA length varition to detect periodicities in positively cooperative, nonspecific binding. *Methods Enzymol.*, **466**, 66–82.
25. Leroy,J.L., Charretier,E., Kochoyan,M. and Gueron,M. (1988) Evidence from base-pair kinetics for two types of adenine tract structures in solution: their relation to DNA curvature. *Biochemistry*, **27**, 8894–8898.
26. Olmsted,M.C., Anderson,C.F. and Record,M.T. (1989) Monte Carlo description of oligoelectrolyte properties of DNA oligomers: range of the end effect and the approach of molecular and thermodynamic properties to the polyelectrolyte limits. *Proc. Natl Acad. Sci. USA*, **86**, 7766–7770.
27. Maxam,A. and Gilbert,W.S. (1977) A new method for sequencing DNA. *Proc. Natl Acad. Sci. USA*, **74**, 560–565.
28. Innis,M.A. and Gelfand,D.H. (1990) *PCR Protocols: A Guide to Methods and Applications*. New York, Academic Press.
29. Laue,T.M., Shah,B.D., Ridgeway,T.M. and Pelletier,S.L. (1992) Computer-Aided Interpretation of Analytical Sedimentation Data For Proteins. In: Harding,S.E., Rowe,A.J. and Horton,J.C. (eds), *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. The Royal Society of Chemistry, Cambridge, England, pp. 90–125.
30. Cohen,G. and Eisenberg,H. (1968) Deoxyribonucleate solutions: sedimentation in a density gradient, partial specific volumes, density and refractive density increments and preferential interactions. *Biopolymers*, **6**, 1077–1100.
31. Tang,J., Li,J.W., Wang,C. and Bai,C.L. (2000) Enhancement of resolution of DNA on silylated mica using atomic force microscopy. *J. Vac. Sci. Technol. B*, **18**, 1858–1860.
32. Vesenka,J., Manne,S., Yang,G., Bustamante,C.J. and Henderson,E. (1993) Humidity effects on atomic force microscopy of gold-labeled DNA on mica. *Scanning Microsc.*, **7**, 781–788.
33. Watson,J.D. and Crick,F.H. (1953) The structure of DNA. *Cold Spring Harb. Symp. Quant. Biol.*, **18**, 123–131.
34. Mandelkern,M., Elias,J.G., Eden,D. and Crothers,D.M. (1981) The dimensions of DNA in solution. *J. Mol. Biol.*, **152**, 153–161.
35. Yang,Y., Sass,L.E., Du,C., Hsieh,P. and Erie,D.A. (2005) Determination of protein–DNA binding constants and specificities from statistical analyses of single molecules: MutS-DNA interactions. *Nucleic Acids Res.*, **33**, 4322–4334.
36. Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13 mp18 and pUC19 vectors. *Gene*, **33**, 103–119.
37. McGhee,J. and von Hippel,P.H. (1974) Theoretical aspects of DNA–protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. *J. Mol. Biol.*, **86**, 469–489.
38. Kowalczykowski,S.C., Paul,L.S., Lonberg,N., Newport,J.W., McSwiggen,J.A. and von Hippel,P.H. (1986) Cooperative and noncooperative binding of protein ligands to nucleic acid lattices: experimental approaches to the determination of thermodynamic parameters. *Biochemistry*, **25**, 1226–1240.
39. Hall,M.C., Wang,H., Erie,D.A. and Kunkel,T.A. (2001) High affinity cooperative DNA binding by the yeast Mlh1-Pms1 heterodimer. *J. Mol. Biol.*, **312**, 637–647.
40. Tessmer,I., Moore,T., Lloyd,R.G., Wilson,A., Erie,D.A., Allen,S. and Tendler,S.J. (2005) AFM studies on the role of the protein RdgC in bacterial DNA recombination. *J. Mol. Biol.*, **350**, 254–262.
41. Bender,K., Federwisch,M., Loggen,U., Nehls,P. and Rajewsky,M.F. (1996) Binding and repair of O^6-ethylguanine in double-stranded oligodeoxynucleotides by recombinant human O^6-alkylguanine-DNA alkyltransferase do not exhibit significant dependence on sequence context. *Nucleic Acids Res.*, **24**, 2087–2094.
42. Daniels,D.S., Mol,C.D., Arval,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 binding. DNA damage reversal revealed by mutants and structures of active and alkylated human AGT. *EMBO J.*, **19**, 1719–1730.
43. Peyrard,M., Cuesta-Lopez,S. and James,G. (2009) Nonlinear analysis of the dynamics of DNA breathing. *J. Biol. Phys.*, **35**, 73–89.
44. Rachofsky,E.L., Osman,R. and Ross,J.B. (2001) Probing structure and dynamics of DNA with 2-aminopurine: effects of local environment on fluorescence. *Biochemistry*, **40**, 946–956.
45. Martin,C.T., Ujvari,A. and Liu,C. (2003) Evaluation of fluorescence spectroscopy methods for mapping melted regions of DNA along the transcription pathway. *Methods Enzymol.*, **371**, 13–33.
46. Jose,D., Datta,K., Johnson,N.P. and von Hippel,P.H. (2009) Spectroscopic studies of position-specific DNA "breathing" fluctuations at replication forks and primer-template junctions. *Proc. Natl Acad. Sci. USA*, **106**, 4231–4236.
47. Holz,B., Klimasauskas,S., Serva,S. and Weinhold,E. (1998) 2-Aminopurine as a fluorescent probe for DNA base flipping by methyltransferases. *Nucleic Acids Res.*, **26**, 1076–1083.
48. Gray,R.D., Petraccone,L., Buscaglia,R. and Chaires,J.B. (2009) 2-Aminopurine as a probe for quadruplex loop structures. *Methods Mol. Biol.*, **608**, 121–136.
49. Malta,E., Moolenaar,G.F. and Goosen,N. (2006) Base flipping in nucleotide excision repair. *J. Biol. Chem.*, **281**, 2184–2194.
50. Neely,R.K., Tamulaitis,G., Chen,K., Kubala,M., Siksnys,V. and Jones,A.C. (2009) Time-resolved fluorescence studies of nucleotide flipping by restriction enzymes. *Nucleic Acids Res.*, **37**, 6859–6870.
51. 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.
52. Adams,C.A., Melikishvili,M., Rodgers,D.W., Rasimas,J.J., Pegg,A.E. and Fried,M.G. (2009) Topologies of complexes containing O^6-alkylguanine-DNA alkyltransferase and DNA. *J. Mol. Biol.*, **389**, 248–263.

53. Crothers,D.M., Drak,J., Kahn,J. and Levene,S.D. (1992) DNA bending, flexibility and helical repeat by cyclization kinetics. *Methods Enzymol.*, **212**, 3–29.
54. Delrow,J.J., Heath,P.J. and Schurr,J.M. (1997) On the origin of the temperature dependence of the supercoiling free energy. *Biophys. J.*, **73**, 2688–2701.
55. Bloomfield,V.A., Crothers,D.M. and Tinoco,I. Jr (2000) *Nucleic Acids: Structures, Properties, and Functions*. University Science Books, Sausalito, pp. 176–181.
56. Dorigo,B., Schalch,T., Kulangara,A., Duda,S., Schroeder,R.R. and Richmond,T.J. (2004) Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science*, **306**, 1571–1573.
57. Arya,G., Maitra,A. and Grigoryev,S.A. (2010) A structural perspective on the where, how, why, and what of nucleosome positioning. *J. Biomol. Struct. Dyn.*, **27**, 803–820.
58. Voigt,J.M. and Topal,M.D. (1990) O^6-methylguanine in place of guanine causes asymmetric single-strand cleavage of DNA by some restriction enzymes. *Biochemistry*, **29**, 1632–1637.
59. Roth,H.M., Romer,J., Grundler,V., Van Houten,B., Kisker,C. and Tessmer,I. (2012) XPB helicase regulates DNA incision by the *Thermoplasma acidophilum* endonuclease Bax1. *DNA Repair (Amst)*, **11**, 286–293.