

Articles

Interactions of Human O⁶-Alkylguanine-DNA Alkyltransferase (AGT) with Short Double-Stranded DNAs[†]Manana Melikishvili,[‡] Joseph J. Rasimas,^{§,||} Anthony E. Pegg,[§] and Michael G. Fried^{*,‡}

Department of Molecular and Cellular Biochemistry and Center for Structural Biology, University of Kentucky, Lexington, Kentucky 40536, and Department of Molecular and Cellular Physiology, Penn State University College of Medicine, Hershey, Pennsylvania 17033

Received September 2, 2008; Revised Manuscript Received October 23, 2008

ABSTRACT: O⁶-alkylguanine-DNA alkyltransferase (AGT) is a ubiquitous enzyme with an amino acid sequence that is conserved in Eubacteria, Archaea, and Eukarya. It repairs O⁶-alkylguanine and O⁴-alkylthymine adducts in single-stranded and duplex DNAs. In performing these functions, AGT must partition between adduct-containing sites and the large excess of adduct-free DNA distributed throughout the genome. Here, we characterize the binding of human AGT to linear double-stranded, adduct-free DNAs ranging in length from 11 bp to 2686 bp. Moderately cooperative binding ($22.6 \pm 3.7 \leq \omega \leq 145.0 \pm 37.0$) results in an all-or-nothing association pattern on short templates. The apparent binding site size S_{app} (mean = 4.39 ± 0.02 bp) oscillates with increasing template length. Oscillations in cooperativity factor ω have the same frequency but are of opposite phase to S_{app} , with the result that the most stable protein–protein and protein–DNA interactions occur at the highest packing densities. The oscillation period (4.05 ± 0.02 bp/protein) is nearly identical to the occluded binding site size obtained at the highest measured binding density (4 bp/protein) and is significantly smaller than the contour length (~ 8 bp) occupied in crystalline complexes. A model in which protein molecules overlap along the DNA contour is proposed to account for these features. High AGT densities resulting from cooperative binding may allow efficient search for lesions in the context of chromatin remodeling and DNA replication.

O⁶-alkylguanine and O⁴-alkylthymine are mutagenic and carcinogenic adducts that are formed in DNA as a result of exposure to alkylating agents (1–3). Both adducts are repaired by the ubiquitous repair protein, O⁶-alkylguanine-DNA alkyltransferase (AGT¹) (1, 4, 5). Ironically, AGT also protects cells against chemotherapeutic drugs that methylate or chloroethylate DNA (6, 7). Clinical trials of AGT inhibitors are underway in an attempt to increase the efficacy of such drugs in cancer chemotherapy (8–10). In addition, overexpression of AGT in hematopoietic stem cells is under study as a means to protect normal cells against the myelosuppressive effects of alkylating-agent chemotherapy

(10, 11). In spite of the interest focused on AGT as a result of its relevance to cancer, relatively little is known about its mechanisms of interaction with DNA and with proteins in its environment. A more complete characterization of these interactions is justified by the probability that they influence the repair activities of AGT *in vivo* as well as our ability to control DNA repair for therapeutic purposes.

Human AGT is a monomeric protein (207 amino acids, $M_r = 21,519$) that is expressed constitutively in normal cells (3, 5, 7) but may be overexpressed in some tumors that have been exposed to alkylating agents (12). The protein binds DNA with modest affinity and little sequence specificity (13–15). Studies with repair-inactive mutant proteins and with proteins in which the alkyltransfer function has been inactivated by prior alkylation of residue C145 (described below) suggest that binding affinities for alkylated DNA sites are only marginally greater than those for the corresponding nonalkylated sequences ((15) Fried, unpublished results). While these results do not rule out preferential binding to O⁶-alkylguanine or O⁴-alkylthymine containing sequences, they raise the intriguing possibility that the lesion-search mechanism of AGT may not depend strongly on preferential binding. In repair reactions, each protein molecule mediates the transfer of a single alkyl group from the O⁶ position of guanine or O⁴ position of thymine to an active-

[†] This work was supported by NIH grants GM-070662 (to M.G.F.) and CA-018137 and CA-097209 (to A.E.P.) and Medical Scientist Training Program Grant 5 T32 GM-08601-05 (J.J.R.).

* To whom correspondence should be addressed. Department of Molecular and Cellular Biochemistry University of Kentucky, 741 South Limestone, Lexington, KY 40536-0509. Tel: (859) 323-1205. Telefax: (859) 323-1037. E-mail: michael.fried@uky.edu.

[‡] University of Kentucky.

[§] Penn State University College of Medicine.

^{||} Current address: Office of the Clinical Director, National Institute of Mental Health, Hatfield Clinical Research Center, 10 Center Dr., Bethesda, MD 20892.

¹ AGT is also called O⁶-methylguanine DNA methyltransferase, MGMT; EC 2.1.1.63. Homologous proteins have been identified in Eubacteria, Archaea and Eukarya (3).

Table 1: Duplex Oligonucleotides

	Length, bp	Sequence
11	5'-TTT TTG TTT TT-3' 3'-AAA AAC AAA AA-5'	
12	5'-GAC TGA CTG ACT-3' 3'-CTG ACT GAC TGA-5'	
14	5'-GAC TGA CTG ACT GA-3' 3'-CTG ACT GAC TGA CT-5'	
16	5'-GAC TGA CTG ACT GAC T-3' 3'-CTG ACT GAC TGA CTG A-5'	
18	5'-GGA ACC TTG GAA CCT TGG-3' 3'-CCT TGG AAC CTT GGA ACC-5'	
21	5'-TGA AGT CCA AAG TTC AGT CCC-3' 3'-CT TCA GGT TTC AAG TCA GGG A-5'	
22	5'-CGC CAA CCC GCT GCC TAT CGT T-3' 3'-GCG GTT GGG CGA CGG ATA GCA A-5'	
24	5'-AAA AAA AAA AAA AAA AAA AAA AAA-3' 3'-TTT TTT TTT TTT TTT TTT TTT TTT-5'	
24	5'-GGG GGG GGG GGG GGG GGG GGG GGG-3' 3'-CCC CCC CCC CCC CCC CCC CCC CCC-5'	
26	5'-GAC TGA CTG ACT GAC TGA CTG ACT GA-3' 3'-CTG ACT GAC TGA CTG ACT GAC TGA CT-5'	
30	5'-GTG CCG CCA ACC CGC TGC CTA TCG TTA TAC-3' 3'-CAC GGC GGT TGG GCG ACG GAT AGC AAT ATG-5'	
41	5'-GCA ACG CAA TTA ATG TGA GTT AGC TCA CTC ATT AGG CAC CC-3' 3'-CGT TGC GTT AAT TAC ACT CAA TCG AGT GAG TAA TCC GTG GG-5'	

site cysteine (C145 in the human protein). This reaction returns the DNA base to an unmodified state and permanently inactivates the alkyl-acceptor activity of the enzyme, which is ultimately degraded (4, 16). As a result of this stoichiometric mechanism, the number of adducts that can be repaired at one time depends on the steady-state cellular concentration of the nonalkylated form of AGT (4, 5) and on its ability to partition between lesion sites and the large excess of genomic DNA in which they are embedded.

Several crystal structures of AGT–DNA complexes have been obtained in recent years. Those determined by Daniels et al. (PDB numbers 1T38 and 1T39) contain 1:1 AGT–DNA complexes (17) while that of Duguid et al. (PDB number 1YFH) contains two AGT molecules making separate contact with a single DNA molecule and a third binding a second DNA (18). The isolated binding of single protein molecules that is a recurring theme in these structures contrasts with the close-packed, cooperative pattern of binding observed in free solution with both single-stranded and duplex templates (15, 19, this work). These differences in binding mode suggest that binding in free solution may involve a different ensemble of molecular contacts than is present in the crystal structures. A more complete characterization of cooperative AGT–DNA interactions may allow us to account for this difference.

EXPERIMENTAL PROCEDURES

Enzymes and Reagents. T₄ polynucleotide kinase and *Eco*RI endonuclease were purchased from New England

Biolabs. [γ -³²P]ATP was from ICN Radiochemicals. Electrophoresis grade polyacrylamide was from Fisher. All other chemicals were reagent-grade or better.

Human AGT Protein. Recombinant human AGT protein (tagged with His₆ at its C-terminal end) was purified to apparent homogeneity according to published protocols (17). The purity of the protein was verified by SDS-gel electrophoresis followed by silver-staining (20, 21). Sedimentation equilibrium data were consistent with a single species of $M_r = 21,860 \pm 400$ for [AGT] $\leq 100 \mu\text{M}$ ((3) and Fried, M., unpublished results). This value is consistent with the value ($M_r = 21,614$) predicted from the sequence of this variant of the protein. The preparations used were >95% active in the transfer of methyl groups from O⁶-[³H]methyl guanine-labeled calf-thymus DNA to AGT and >95% active in debenzylating O⁶-benzylguanine, as previously described (22, 23). AGT concentrations were measured spectrophotometrically using $\epsilon_{280} = 3.93 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ (23). Samples were stored at -80°C until needed.

DNA Substrates. Oligodeoxyribonucleotides of 11, 12, 14, 16, 18, 21, 22, 24, 26, 30, and 41 residues and their complements (sequences shown in Table 1) were synthesized by the Macromolecular Core Facility of the Penn State College of Medicine or were purchased from Midland Certified Reagent Company. Some DNA sequences were previously used in studies of DNA binding and/or repair by AGT (c.f. 14, 24). Others were selected for particular sequence attributes (e.g., dG₂₄·dC₂₄ and dA₂₄·dT₂₄). Oligonucleotides were labeled at 5' termini with ³²P as described

by Maxam and Gilbert (25). Single-stranded DNA concentrations were measured spectrophotometrically, using extinction coefficients calculated by the nearest-neighbor method (26, 27). Depending on base composition and sequence, values of ϵ_{260} ranged from $8.45 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$ to $1.04 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ (per base). Duplex DNAs were obtained by mixing a labeled oligonucleotide with a 1.05-fold molar excess of its unlabeled complement. Samples were heated to 90 °C for 3 min and cooled to 20 °C over 2 h. Duplex formation was monitored by nondenaturing PAGE. Where necessary, samples were purified by preparative electrophoresis (28). Duplex 147 bp and 207 bp DNA fragments containing the sea-urchin nucleosome positioning sequence were isolated by restriction digestion and preparative electrophoresis from plasmid p5S-207-18 (29), kindly provided by Dr. Sergei Grigoryev, Penn State University. Plasmid pUC19 DNA (New England Biolabs) was linearized with *EcoRI* endonuclease. Restriction fragments were deproteinized by phenol extraction and ethanol precipitation (30). All DNAs were dialyzed against 10 mM Tris (pH 8.0), 0.1 mM EDTA prior to use or storage at -20 °C. Duplex DNA concentrations were measured by absorbance using the average molar extinction coefficient $\epsilon_{260} = 1.3 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ (per base pair).

Electrophoretic Mobility Shift Assays (EMSA). Mobility shift assays were carried out with DNAs of ≤ 41 bp. Under the conditions used here, complexes formed with DNAs of 148 bp and larger were not sufficiently resolved from free DNA to allow quantitative binding analysis. Binding reactions were carried out at 20 ± 1 °C in 10 mM Tris (pH 7.6), 100 mM NaCl, 1 mM dithiothreitol, and 10 $\mu\text{g/mL}$ bovine serum albumin. Mixtures were equilibrated at $20 (\pm 1)$ °C for 30 min. Duplicate samples incubated for longer periods gave identical results, indicating that equilibrium had been attained (result not shown). Electrophoresis was carried out in 10% polyacrylamide gels as previously described (19). Autoradiographic images were captured on storage phosphor screens (type GP, GE Healthcare) detected with a Typhoon phosphorimager and quantitated with Image-Quant software (GE Healthcare). In a few cases, autoradiographic data was captured on Kodak XL-1 film, and band intensities were measured by excision and scintillation counting as previously described (15).

The serial dilution method (31, 32) was used to obtain estimates of binding stoichiometry (n). For the binding of n protein molecules to 1 DNA, the association constant is $K_n = [P_n D]/[D][P]_{\text{free}}^n$. Here, P represents AGT protein, D represents DNA, and $P_n D$ the AGT–DNA complex of protein–DNA ratio n . Separating variables and taking logarithms gives

$$\log \frac{[P_n D]}{[D]} = n \log [P]_{\text{free}} + \log K_n \quad (1)$$

Dilution of an AGT–DNA mixture changes $[P_n D]/[D]$ by mass action while maintaining the ratio of $[P]_{\text{total}}$ to $[D]_{\text{total}}$. The free protein concentration at each dilution step can be estimated using $[P]_{\text{free}} = [P]_{\text{input}} - n[P_n D]$ starting with an initial value of n ; equation 1 was then used to calculate a new value of n from the DNA-binding distribution. This value was used to calculate a new estimate of $[P]_{\text{free}}$ and the calculations repeated recursively until values of n converged.

Analytical Ultracentrifugation. Human AGT protein and duplex DNAs were dialyzed against 10 mM Tris (pH 7.6), 1 mM DTT, 1 mM EDTA, and 100 mM NaCl. Analytical ultracentrifugation was performed at 20 ± 0.1 °C in a Beckman XL-A centrifuge using an AN60Ti rotor. Most scans were obtained at 252 nm to minimize protein contribution to the DNA signal (33), although some data sets were obtained at 260 nm. Approach to equilibrium was considered to be complete when scans taken 6 h apart were indistinguishable. Typically, equilibration times ≥ 24 h met this criterion for AGT–DNA mixtures. Five scans were averaged for each sample at each wavelength and rotor speed.

For DNAs with small numbers of protein-binding sites, cooperative binding can be described by the simple mechanism $nP + D \rightleftharpoons P_n D$ in which free DNA (D) is in equilibrium with saturated complex ($P_n D$), but intermediates with protein stoichiometries $< n$ are not present at significant concentrations. At sedimentation equilibrium, the radial distribution of absorbance for such a system is given by equation 2.

$$A(r) = \alpha_P \exp[\sigma_P(r^2 - r_o^2)] + \alpha_D \exp[\sigma_D(r^2 - r_o^2)] + \alpha_{P_n D} \exp[\sigma_{P_n D}(r^2 - r_o^2)] + \epsilon \quad (2)$$

Here, $A(r)$ is the absorbance at radial position r , and α_P , α_D , and $\alpha_{P_n D}$ are absorbances of protein, DNA, and the 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, DNA and protein–DNA complexes are given by $\sigma_P = M_P(1 - \bar{v}_P \rho)/(2RT)$, $\sigma_D = M_D(1 - \bar{v}_D \rho)/(2RT)$ and $\sigma_{P_n D} = (nM_P + M_D)(1 - \bar{v}_{P_n D} \rho)/(2RT)$. Here, M_P and M_D are the molecular weights of protein and DNA, n is the protein/DNA ratio of the complex; r is the solvent density, z , the rotor angular velocity, R the gas constant, and T the temperature (Kelvin). The partial specific volume of AGT ($\bar{v}_P = 0.744 \text{ mL/g}$) was calculated by the method of Cohn and Edsall (34), using partial specific volumes of amino acids tabulated by Laue and co-workers (35). The partial specific volume of double-stranded NaDNA at 0.1 M NaCl (0.540 mL/g) was estimated by interpolation of the data of Cohen and Eisenberg (36). Partial specific volumes of each protein–DNA complex were estimated using eq 3.

$$\bar{v}_{P_n D} = \frac{(nM_P \bar{v}_P + M_D \bar{v}_D)}{(nM_P + M_D)} \quad (3)$$

Equation 2 was used in global analysis of multiple data sets obtained at different macromolecular concentrations and rotor speeds (37). In this method, the values of α_P , α_D , $\alpha_{P_n D}$, and ϵ are unique to each sample, but the value of n must be common to all data sets. Terms accounting for nonideal effects were not included because there was no evidence of nonideality (results not shown).

For large DNAs with many protein-binding sites, the concentration of protein-free DNA molecules becomes negligible long before 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 reached. These systems were analyzed using equation 4, with parameters defined as they were for eq 2, above.

$$A(r) = \alpha_P \exp[\sigma_P(r^2 - r_o^2)] + \alpha_{P_n D} \exp[\sigma_{P_n D}(r^2 - r_o^2)] + \varepsilon \quad (4)$$

Determination of Equilibrium Constants and Cooperativity Parameters. Association constants and cooperativity parameters were evaluated by direct titration of DNA with protein, with binding detected by EMSA or sedimentation equilibrium analysis. The total concentration of protein binding sites on DNA was always much less than that of the protein allowing the approximation $[P]_{\text{total}} = [P]_{\text{free}}$ to be used. The dependence of binding density ν on the free protein concentration $[P]$ was given by the McGhee-von Hippel isotherm (38) as modified by Record et al. (39) to account for finite lattice size (eq 5).

$$\frac{\nu}{[P]} = K(1 - s\nu) \left(\frac{(2\omega - 1)(1 - s\nu) + \nu - R}{2(\omega - 1)(1 - s\nu)} \right)^{s-1} \times \left(\frac{1 - (s+1)\nu + R}{2(1 - s\nu)} \right)^2 \left(\frac{N - s + 1}{N} \right) \quad (5)$$

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

Here, ν is the binding density (protein molecules/nucleotide), K is the equilibrium association constant for binding a single site, ω is the cooperativity parameter, N is the length of the DNA in base pairs, and s is the occluded site size (the size of the site, in base pairs, that one protein molecule occupies to the exclusion of others)². The cooperativity parameter ω is equal to the equilibrium constant for moving a protein from an isolated site to a singly contiguous one or from a singly contiguous site to a doubly contiguous one (38). The model embodied in this equation is one in which proteins are assumed not to bind to fractional sites of length $< s$ base pairs located within or at the ends of DNA molecules (39).

RESULTS

Determination of Stoichiometries. Sedimentation equilibrium and EMSA assays were used to determine binding stoichiometries. When short double-stranded DNAs (dsDNAs) were titrated with AGT, the major DNA species present at equilibrium were free DNA and the saturated protein–DNA complex (Figure 1). As previously discussed, the formation of a saturated multiprotein complex from free DNA without accumulation of significant concentrations of lower-stoichiometry species is a consequence of positively cooperative binding (15). Thus, for sedimentation equilibrium analysis of binding to templates ≤ 207 bp (Fig 2), data were fit using the sedimentation equation corresponding to the concerted binding model ($nP + D \rightleftharpoons P_n D$; eq 2). For plasmid pUC19 DNA, the number of binding sites is much larger than the cooperativity parameter ω (described below); therefore, over a wide range of $[P]$, solutions contain only free protein and protein–DNA complexes, but no detectable free DNA. When this was the case, the data were fit using a two-species model (eq 4). The small, symmetrical residuals obtained from these fits (upper panels) indicate that these models account well for the mass distributions present in our samples. Molecular weights of protein–DNA complexes were obtained as fitting parameters, and the stoichiometry

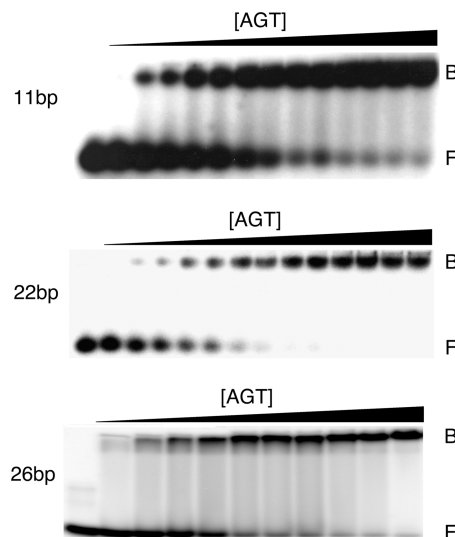


FIGURE 1: Titration of representative double-stranded DNAs with human AGT. Upper panel: 11-mer duplex DNA, 8.8×10^{-7} M with $[AGT]$ increasing from 0 M to 6.5×10^{-5} M from left to right. Middle panel: 22-mer duplex DNA, 3.6×10^{-7} M, with $[AGT]$ increasing from 0 M to 1.5×10^{-5} M from left to right. Lower panel: 26-mer duplex DNA, 3.9×10^{-7} M, with $[AGT]$ increasing from 0 M to 4.8×10^{-5} M, from left to right. Binding reactions were carried out at 20 ± 1 °C and samples were resolved on 10% polyacrylamide gels, as described in Experimental Procedures. Band designations B, bound DNA; F, free DNA. Although these images have been cropped and labeled for clarity, no additional bands were detectable between the origin of electrophoresis and the ionic front.

of each complex was inferred from the known molecular weights of the DNA and AGT protein. The results of these analyses are summarized in Table 2.

Additional stoichiometry values were determined by native gel electrophoresis (EMSA) using the serial dilution method (31, 32). A representative experiment showing mass-action dissociation of the AGT complex formed with the dA₂₄•dT₂₄ duplex is shown in Figure 3A. As with forward titrations, free DNA and an apparently saturated complex are the dominant species and stoichiometric intermediates are not detectably present. A graph of $\log[P_n D]/[D]$ as a function of $\log[P]$ is shown in Figure 3B. The slope of this graph near the midpoint of the binding reaction is a measure of the protein stoichiometry of the complex (40). Stoichiometry values determined in this way are summarized in Table 2. The agreement of stoichiometry values determined by EMSA with those obtained by sedimentation equilibrium indicates that the same complexes are detected by both methods and demonstrates that dissociation of complexes during electrophoresis is not a significant factor under the conditions that were used here. These data also show that increasing DNA length results in higher limiting stoichiometries, indicating that for the DNA concentration range explored here and under conditions of AGT excess, stoichiometry is limited by the number of binding sites available on each DNA.

Apparent Binding Site Size Depends on DNA Length. For a DNA of N bp, bound by a protein that occupies and occludes s bp/protein molecule, the maximum possible stoichiometry is the largest integer $\leq N/s$, which we designate n_{max} . If binding-site utilization is efficient, i.e., if saturated complexes contain n_{max} proteins, we can define an apparent binding site size $S_{\text{app}} = N/n_{\text{max}}$ bp/protein. For short DNAs,

² This is the occluded site size obtained for complexes with the highest binding densities (formed with DNAs of 12 bp, 16 bp, and 24 bp).

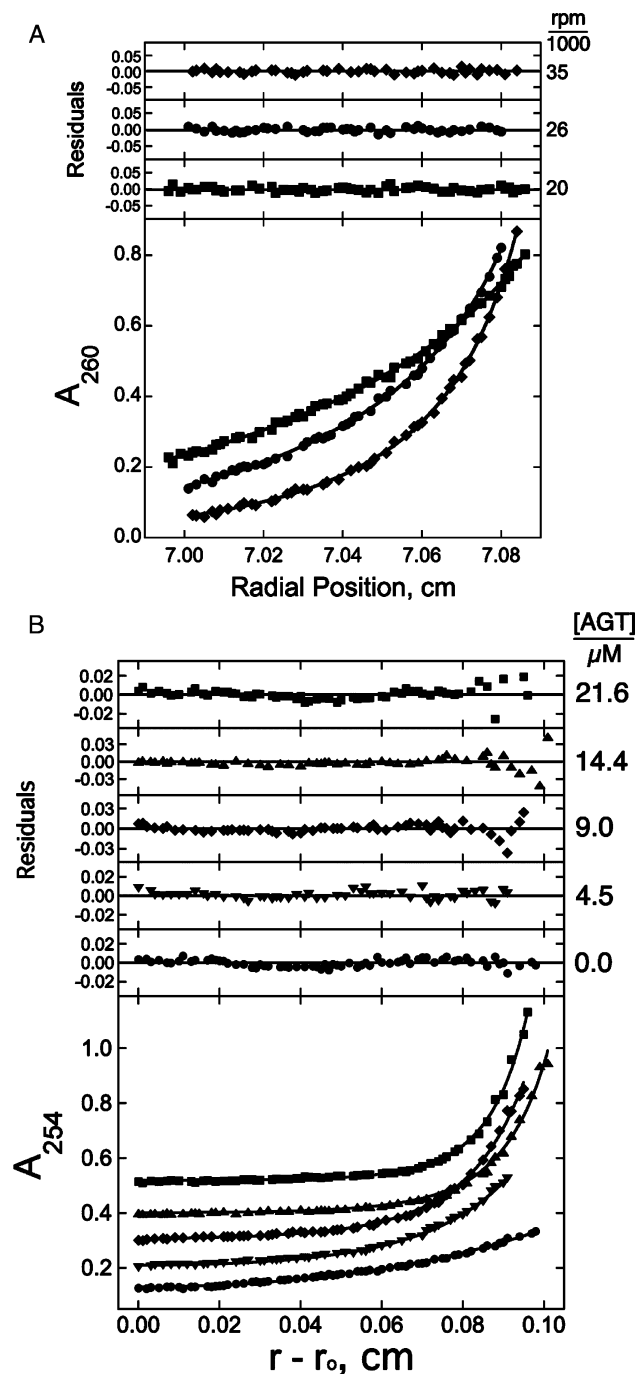


FIGURE 2: Sedimentation equilibrium of solutions containing AGT and double-stranded DNAs at $20 \pm 1^\circ\text{C}$. Panel A: Representative data for binding to 26 bp DNA. Samples contained DNA (5×10^{-7} M) and AGT (1.45×10^{-5} M) in buffer consisting of 10 mM Tris (pH 7.6), 1 mM DTT, 1 mM EDTA, and 100 mM NaCl. Radial scans taken at 20,000 rpm (■), 26,000 rpm (●), and 35,000 rpm (◆) are shown with vertical offsets for clarity. The smooth curves correspond to a global fit of eq 2 to a data set that includes these scans and scans obtained at 2 additional AGT concentrations (1.82×10^{-5} M and 2.18×10^{-5} M). The small residuals, nearly symmetrically distributed about zero (upper panels) indicate that the cooperative $nP + D \rightleftharpoons P_nD$ model is consistent with the mass distributions of DNA in these samples and that $n = 5.65 \pm 0.21$. Panel B: Representative data for binding to linear pUC19 DNA (2686 bp). Samples contained DNA (5×10^{-7} M) and AGT in buffer consisting of 10 mM Tris (pH 7.6), 1 mM DTT, 1 mM EDTA, and 100 mM NaCl. AGT concentrations were 0 M (●), 4.5×10^{-6} M (▼), 9.0×10^{-6} M (◆), 1.44×10^{-5} M (▲), and 2.16×10^{-5} M (■). Radial scans shown here were taken at 3,000 rpm. The smooth curves are fits of eq 4 to each data set.

values of S_{app} are expected to fall between s and $2s - 1$ and oscillate with increasing N . Because minima in S_{app} occur whenever N/s is an integer, this oscillation is expected to have a frequency equal to s . These properties, first described by Epstein (41), are characteristic of efficient, nonoverlapping binding with occluded site size of s bp. Below, we compare the properties of AGT–DNA complexes with the predictions of this simple model.

Values of S_{app} measured for AGT oscillate with increasing DNA length (Figure 4). This is most evident for short templates (≤ 30 bp); above this length, the separation between tested values is too great for oscillation to be easily seen. Shown in Figure 4B, experimental S_{app} values coincide with theoretical values calculated using $s = 4$ bp, the occluded site size observed with 12, 16, and 24 bp DNAs. The agreement between theoretical and experimental S_{app} values is best when $N/4$ is an integer. However, theoretical S_{app} maxima decrease with increasing DNA length as the proportion of DNA length within an unoccupied partial binding site decreases, and this decrease is not reproduced by the experimental data. This divergence may reflect increasingly degenerate binding as DNA length increases, which may result in population average stoichiometries $< n_{\text{max}}$. If this is the case, minimization of degeneracy whenever $N/4$ is an integer would still result in oscillations in S_{app} with a period of 4 bp. Fitting S_{app} to a cosine function³ shows that the period and phase of the oscillation are strongly retained over the range $11 \text{ bp} \leq N \leq 30 \text{ bp}$ (Figure 4C), with an oscillatory period of 4.05 ± 0.02 bp that is consistent with the $s = 4$ bp value inferred from minima in S_{app} .

For DNAs of 148, 207, and 2686 bp, values of S are significantly larger than those found for shorter DNAs. Since incomplete occupancy of DNA ends should be equally possible for all DNAs, the less-efficient packing of AGT on these DNAs may reflect the presence of gaps between groups of tightly packed protein molecules or a uniform but less tightly packed structure. The appearance of gaps between cooperative units is one of the features discussed by McGhee and von Hippel in their development of cooperative binding theory (38). At the other extreme in template length, the value of S is significantly greater for the 11 bp DNA than for any of the short duplexes (up to and including the 41 bp fragment). This is accounted for by the fact that with a site frequency of $\sim 4 \text{ bp/protein}$, only two AGT molecules can be accommodated on the 11-mer and that the 3 residual base pairs are a larger proportion of the length of this DNA than of any larger template. Between these extremes, the dominant picture is one of a protein array in which the average separation between the start of two adjacent binding sites is ~ 4 bp. This spacing is significantly shorter than the $\sim 8 \text{ bp/protein}$ occupied in crystalline AGT complexes (17, 18) and the optimum site size found for the rate of methyl-group transfer from single-stranded oligonucleotides (7 nt (42)). As discussed below, our working hypothesis is that binding sites for AGT molecules overlap *along the contour* of dsDNA.

DNA-Length Dependence of K and ω . Changes in binding density may be accompanied by changes in the intrinsic

³ This corresponds to the first terms of a Fourier cosine series (46). We chose this simple expression because it adequately models the oscillation and because the fit was not significantly improved when additional sine or cosine terms were added (results not shown).

Table 2: Properties of AGT–DNA Complexes

DNA length (bp)	M_r (saturated complex) ^a	stoichiometry	S_{app} bp/protein	K/M	ω	$K \cdot \omega$
11	$5.08 \pm 0.24 \times 10^4$	2.05 ± 0.11^a 1.99 ± 0.15^b	5.36 ± 0.31^a 5.52 ± 0.38^b	$1.69 \pm 0.19 \times 10^{3c}$	45.9 ± 6.0^c	$7.75 \pm 1.92 \times 10^4$
12		2.91 ± 0.14^b	4.13 ± 0.21^b	$4.08 \pm 1.12 \times 10^{3c}$	48.0 ± 5.5^c	$1.96 \pm 0.78 \times 10^5$
14		2.94 ± 0.10^b	4.75 ± 0.12^b	$6.11 \pm 0.96 \times 10^{3c}$	52.3 ± 5.1^c	$3.18 \pm 0.82 \times 10^5$
16	$9.71 \pm 0.17 \times 10^4$	3.94 ± 0.08^a 3.98 ± 0.01^b	4.06 ± 0.08^a 4.04 ± 0.02^b	$1.25 \pm 0.22 \times 10^{4c}$	105.2 ± 18.8^c	$1.31 \pm 0.47 \times 10^6$
18		3.82 ± 0.24^b	4.70 ± 0.18^b	$4.23 \pm 0.72 \times 10^{3c}$	85.7 ± 11.6^c	$3.59 \pm 1.10 \times 10^5$
21		5.06 ± 0.20^b	4.15 ± 0.17^b	$6.07 \pm 1.80 \times 10^{3c}$	92.6 ± 25.1^c	$5.62 \pm 3.21 \times 10^5$
22		4.58 ± 0.24^b	4.80 ± 0.24^b	$3.80 \pm 0.95 \times 10^{3c}$	68.3 ± 13.5^c	$2.58 \pm 1.16 \times 10^5$
24 ^e		5.93 ± 0.25^b	4.05 ± 0.18^b	$2.31 \pm 0.36 \times 10^{3c}$	83.1 ± 13.2^c	$1.92 \pm 0.61 \times 10^5$
24 ^f		6.01 ± 0.20^b	3.99 ± 0.14^b	$4.01 \pm 1.78 \times 10^{3c}$	110.1 ± 23.2^c	$4.41 \pm 2.89 \times 10^5$
26	$1.37 \pm 0.41 \times 10^5$	5.64 ± 0.18^a 5.50 ± 0.16^b	4.60 ± 0.18^a 4.72 ± 0.14^b	$8.18 \pm 2.52 \times 10^{3c}$	93.5 ± 26.5^c	$7.64 \pm 4.53 \times 10^5$
30	$1.49 \pm 0.34 \times 10^5$	6.08 ± 0.16^a 6.43 ± 0.13^b	4.93 ± 0.13^a 4.66 ± 0.11^b	$9.71 \pm 2.03 \times 10^{3c}$	64.7 ± 15.3^c	$6.28 \pm 2.79 \times 10^5$
41		9.01 ± 0.24^b	4.55 ± 0.13^b	$1.87 \pm 0.45 \times 10^{3c}$	145.0 ± 37.0^c	$2.71 \pm 1.35 \times 10^5$
148	$5.66 \pm 0.14 \times 10^5$	21.76 ± 0.65^a	6.80 ± 0.21^a	$8.23 \pm 1.51 \times 10^{3d}$	53.5 ± 22.1^d	$4.40 \pm 2.62 \times 10^5$
208	$8.05 \pm 0.33 \times 10^5$	31.04 ± 1.54^a	6.70 ± 0.35^a	$7.87 \pm 0.82 \times 10^{3d}$	63.6 ± 16.6^d	$4.95 \pm 1.82 \times 10^5$
2686	$1.05 \pm 0.51 \times 10^7$	406.3 ± 23.7^a	6.61 ± 0.41^a	$8.52 \pm 0.96 \times 10^{3d}$	22.6 ± 3.7^d	$1.92 \pm 0.53 \times 10^5$

^a Determined by sedimentation equilibrium analysis. The error ranges are 95% confidence limits. ^b Determined by EMSA-serial dilution analysis. The error ranges are 95% confidence limits. ^c Determined by scatchard analysis of EMSA data, using binding densities calculated from serial dilution analysis. ^d Determined by scatchard analysis of sedimentation equilibrium data, using binding densities calculated from M_w of the protein–DNA complex. ^e dG₂₄•dC₂₄. ^f dA₂₄•dT₂₄.

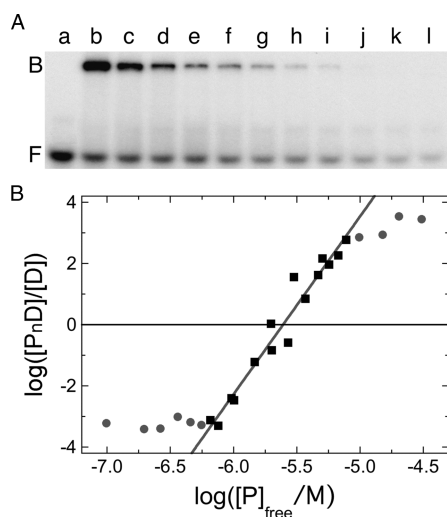


FIGURE 3: Serial dilution analysis of the AGT complex formed with dA₂₄•dT₂₄. Panel A: Binding detected by EMSA. Sample a: 24-mer DNA (1.10×10^{-7} M) only. Sample b: 24-mer DNA (1.10×10^{-7} M) plus AGT (5.36×10^{-6} M). Samples c–l are sequential 1.33-fold dilutions of sample b. All samples were equilibrated in buffer consisting of 10 mM Tris (pH 7.6), 100 mM NaCl, 1 mM DTT, and 0.05 mg/mL bovine serum albumin for 30 min at 20 ± 1 °C prior to resolution on native gels as described in Experimental Procedures. Although this image has been cropped and labeled for clarity, no additional bands were detectable between the origin of electrophoresis and the ionic front. Panel B: Graph of the dependence of $\log([P_nD]/[D])$ on $\log([P]_{free}/M)$ for the AGT complex formed with dA₂₄•dT₂₄. Data are from the experiment shown in A and others that provide additional [AGT] values. The line represents a least-squares fit to the data ensemble for the range about the midpoint of the reaction ($-6.18 \leq \log([AGT]/M) \leq -5.11$), with $[AGT]_{free}$ calculated as described in Experimental Procedures. Symbols: the points used in the fit are indicated by (■); other points in the data set are indicated by closed circles (●). The slope equals 5.81 ± 0.34 for this subset of the data.

association constant of AGT with DNA (K) or in the cooperativity with which AGT forms multiprotein complexes (ω). To explore these possibilities, we recast our binding data according to the short-lattice variant of the McGhee–von Hippel relation (eq 5). This relates binding density ν to free protein concentration $[P]$ in terms of K , ω , and the

occluded binding site size, s . Shown in Figure 5A are Scatchard plots of data obtained from serial dilution and direct titration experiments. The smooth curves represent fits of eq 5 to the data, in which values of s were calculated from stoichiometries determined independently (by sedimentation equilibrium or serial dilution as described above). The good correspondence of fits to data indicates that binding models that include neighbor exclusion and positive cooperativity can account for the binding activities in these systems.

The least-squares best values of K and ω obtained as fitting parameters are summarized in Table 2. Tests using fixed values of K show that K and ω are weakly anticorrelated (an increase in one parameter produces a slight decrease in the other). However, the quality of the fits measured by the χ^2 parameter decreased significantly if either parameter varied from its most probable value by more than the 95% confidence limits (results not shown). Thus, while the values of these parameters cannot be considered unique, the data quality is sufficient to allow us to parse binding affinity into reasonably well-defined estimates of K and ω .

Values of K and ω oscillate as DNA length increases (Figure 5B). The fact that these oscillations are larger than the confidence limits associated with individual parameters and the fact that they are not in phase argue that they are not consequences of fitting correlations between K and ω (as discussed above, K and ω are weakly anticorrelated). Oscillations of ω with DNA length are regular in amplitude and period; this oscillation has the same period but opposite phase as that for the oscillation of S_{app} with length. As a result, the most compact complexes (with smallest S_{app}) have relatively stronger protein–protein interactions than do less compact assemblies. One possible interpretation of these results is that optimal protein juxtaposition is lost as binding density decreases (S_{app} increases). It is noteworthy that the baseline for the oscillation of ω increases uniformly for $11 \text{ bp} \leq \text{DNA length} \leq 30 \text{ bp}$. This is an expected consequence of the growth of a cooperative array in which proteins bind adjacent sites on a linear substrate. Proteins at array ends

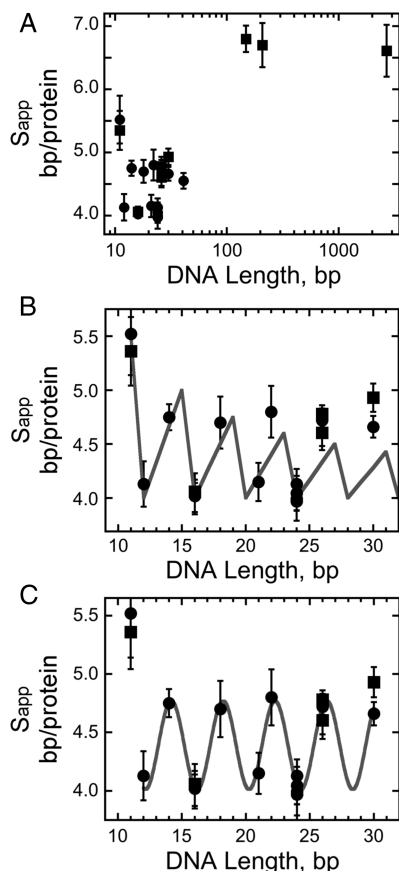


FIGURE 4: AGT forms a binding motif with a 4 bp periodicity. Panel A: Dependence of apparent binding site size (S_{app}) on template length. Data are from the entire set of AGT–dsDNA complexes. S_{app} was calculated using $S_{app} = N/n$, where N is DNA length in base pairs and n is the number of protein molecules bound to a DNA molecule. S_{app} values determined from sedimentation equilibrium data are indicated by filled squares (■), and values obtained from EMSA experiments are indicated by closed circles (●). The error bars correspond to 95% confidence limits. These data are also shown in Table 2. Panel B: Comparison of experimental and theoretical S_{app} values as a function of DNA length. Experimental S_{app} values determined from sedimentation equilibrium data are indicated by filled squares (■), and values obtained from EMSA experiments are indicated by closed circles (●). The error bars correspond to 95% confidence limits. Theoretical S_{app} values (solid line) were calculated using $S_{app} = N/n_{max}$, where n_{max} is the largest integer $\leq N/4$. Panel C: Data for the subset of AGT–DNA complexes formed with DNAs of 41 bp or less. Data symbols are defined as in panel A. The smooth curve is the least-squares fit of the equation $S_{app} = A \cos(BL) + C$ in which A is the amplitude of the oscillation, B the displacement angle in degrees/bp, and C an offset equal to the mean value of S_{app} . This fit returned $A = -0.37 \pm 0.04$, $B = 88.7 \pm 0.5^\circ$, and $C = 4.39 \pm 0.02$. This value of B indicates that successive binding sites are separated by $360/(88.7 \pm 0.5) = 4.05 \pm 0.02$ bp along the DNA contour.

interact with a single neighbor, while proteins in the center of the array interact with two or more neighbors. As arrays increase in length, the proportion of proteins in the center increases, with corresponding increase in ω averaged over the whole assembly. A model of a cooperative assembly that unites these features is proposed below.

In contrast to the regular pattern seen with ω , oscillations in K are irregular in amplitude and phase, suggesting that factors other than DNA length and binding density contribute significantly to the strength of the AGT–DNA contact. The difference in K values for dG₂₄•dC₂₄ and dA₂₄•dT₂₄ templates (Figure 5B and Table 2) suggests that DNA sequence has

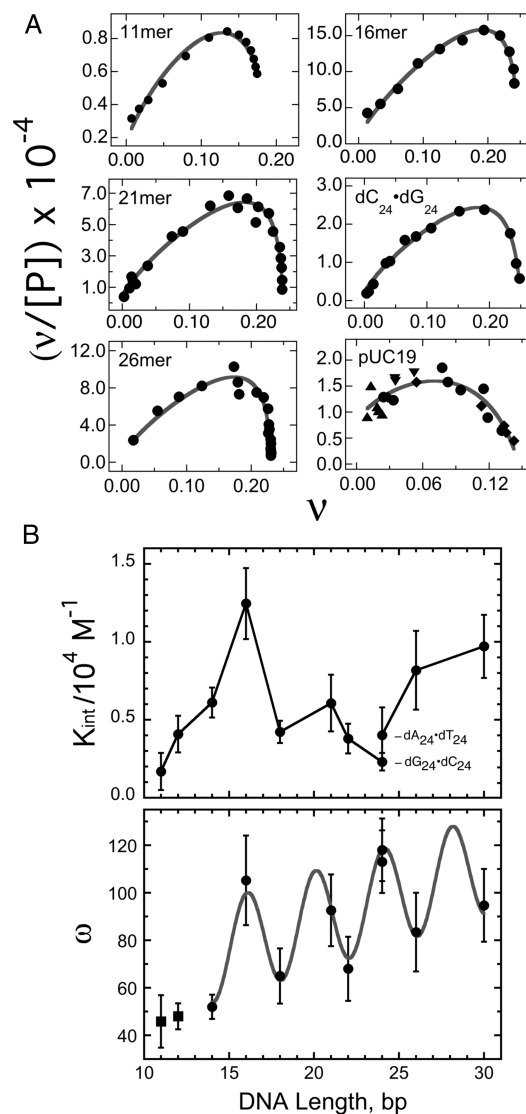


FIGURE 5: Analysis of binding affinities. Panel A: Scatchard plots. Forward titrations were carried out as described in Experimental Procedures, and binding was detected and quantitated by an electrophoretic mobility shift assay. Each data set is a composite derived from 2 or 3 independent titrations. The smooth curves correspond to nonlinear least-squares fits of eq 2 to the data. The values of K and ω obtained as parameters of these fits are compiled in Table 2. Panel B: Dependence of K and ω on DNA length (N). The data are derived in part from the experiments shown in panel A. The error bars correspond to 95% confidence limits estimated for each parameter. K data for dA₂₄•dT₂₄ and dG₂₄•dC₂₄ templates are labeled. Comparison with Figure 4B shows that the oscillations of ω with increasing N have the same period but phases opposite those of S_{app} .

an important influence on affinity. Consistent with this notion, the two DNAs that give local maxima in K (the 16-mer and the 26-mer) have repeating sequences that provide a guanine at every fourth residue on each strand (shown in Table 1). This frequency, which is the same as that of AGT in the cooperative complex, may allow the enzyme to make unusually favorable contacts with its substrate.

DISCUSSION

The data presented here are consistent with a binding mechanism in which weak affinity for duplex DNA ($1.7 \pm 0.2 \times 10^3 \text{ M}^{-1} \leq K \leq 1.3 \pm 0.2 \times 10^4 \text{ M}^{-1}$) is paired with

moderate binding cooperativity ($22.6 \pm 3.7 \leq \omega \leq 145.0 \pm 37.0$). On short DNAs, this cooperativity produces tight packing that results in oscillation in the apparent binding site size S_{app} with a period of 4 bp/protein molecule. On DNAs of length >41 bp, cooperativity is not sufficient to suppress binding degeneracy; therefore, S_{app} increases (Figure 4); however, our data for lengths greater than 41 bp is too sparse to allow us to determine whether oscillations in S_{app} occur in much larger templates. The correlation of maxima in ω with minima in S_{app} (Figures 4C and 5B) indicates that the optimal juxtaposition for cooperativity occurs when neighboring binding sites are 4 bp apart. Since a single molecule of AGT occupies ~ 8 bp with its center in the minor groove of double-stranded DNA (17, 18), this spacing requires that protein molecules overlap along the DNA contour, but it does not require overlap along the minor groove surface that provides the majority of DNA contacts⁴. Overlap along the DNA contour provides opportunities for protein–protein contacts that may contribute to cooperativity.

The cooperative binding mechanism described here for duplex DNA and previously for single-stranded DNAs (19) contrasts with the independent binding shown in currently available crystal structures of AGT–DNA complexes (17, 18). At present, we cannot account for this difference. To date, cooperative binding has been observed with wild-type human AGT over a range of temperature ($4^\circ\text{C} \leq T \leq 40^\circ\text{C}$), salt concentrations ($0 \text{ M} \leq [\text{KCl}] \leq 0.4 \text{ M}$), magnesium ion concentrations ($0 \text{ mM} \leq [\text{Mg}^{2+}]_{\text{free}} \leq 20 \text{ mM}$), and zinc availability (Zn-depleted and Zn-saturated proteins, and Zn-saturated protein with $0 \text{ mM} \leq [\text{Zn}^{2+}]_{\text{free}} \leq 10 \text{ mM}$). In addition, cooperative binding has been observed with active-site (C145) mutant- and C145-methylated and -benzylated proteins as well as C-terminal (His)₆-tagged, N-terminal (His)₆-tagged, and untagged proteins (15, 19, 23, 43). We conclude that cooperative binding is a *bona fide* activity of human AGT.

The pattern described here for double-stranded DNA binding is remarkably similar to that found for single-stranded substrates (19). Quantitative similarities include weak intrinsic affinity for substrate DNA ($1 \times 10^3 \text{ M}^{-1} \leq K$ (single stranded) $\leq 2 \times 10^4 \text{ M}^{-1}$), moderate cooperativity ($35 \leq \omega$ (single stranded) ≤ 90), and a statistical binding site size that oscillates on short DNAs with a period of ~ 4 nt/protein. Similar K values indicate that the free energies of protein–DNA interaction are almost the same for single-stranded and duplex substrates, while comparison of K and ω values indicates that protein–protein contacts contribute a similar, large proportion ($\sim 33\%$) of the free energy of addition of a protein monomer to cooperative assemblies. Together, these results suggest why similar stabilities are observed for AGT complexes formed on single-stranded and duplex DNAs (15): comparatively strong protein–protein interactions mask the contributions of relatively small differences in K to the overall product $K \cdot \omega$.

What role does cooperative binding play in sequence-specificity? Previous work has shown that overall affinity

($K \cdot \omega$) and DNA-repair efficiency are not strongly dependent on base composition or sequence of the substrate (14, 44, 45). For the DNAs described here, values of K and ω vary by factors of 7.4 and 6.4, respectively. However, because changes in K and ω are not perfectly in phase, the largest difference in their product is only a factor of 17. These results suggest that differences in K may be offset by differences in ω , resulting in a very small sequence-dependence of overall binding affinity. The absence of a large dependence of binding affinity on base composition or sequence may be important to ensure that repair activity is available to all sequences. In addition, cooperative binding may contribute to the efficiency of lesion search and repair by concentrating repair activity at a higher density on the DNA contour than would be available in noncooperative binding modes. With a new binding site starting every 4 bp, only modest protein displacement is needed for surveillance of every base-pair within a cooperative assembly. Preliminary kinetic results suggest that AGT complexes isomerize rapidly, as required for efficient surveillance according to this model (Fried, M., unpublished results). Although the moderate cooperativity detected here is insufficient to extend high-density binding over hundreds of base-pairs, it can generate densities approaching 1 protein every 4 bp for sections of template ≤ 50 bp long. Regions of this size are likely to become available between nucleosomes during the chromatin remodeling associated with transcription or DNA replication. The processive movements of transcription and replication complexes may allow systematic surveillance of large parts of the genome by AGT, while repair just upstream of a replication fork is likely to represent the last opportunity to prevent the conversion of a promutagenic lesion into an actual mutation.

How are AGT proteins arranged in the cooperative array? Our working model is based on the binding density of 1 protein/4 base-pairs characteristic of complexes formed on short DNAs. This frequency allows identical protein contacts with the minor groove of a B-DNA duplex if each protein is rotated clockwise approximately $(4 \text{ bp} \cdot \text{protein}^{-1} / 10.4 \text{ bp} \cdot \text{turn}^{-1}) \times 360 \text{ deg} \cdot \text{turn}^{-1} = 138.4 \text{ deg} \cdot \text{protein}^{-1}$ with respect to its immediate neighbor. This arrangement would allow each AGT protein to maintain the minor-groove contacts that are seen in the crystal structures of AGT–dsDNA complexes (17, 18), and it also produces a 3-start helical array of protein molecules, with the n th protein molecule positioned to interact with protein $n + 3$ (shown schematically in Figure 6). The ability to form an n -to- $n+3$ contact may account for the significantly greater cooperativity seen in the 4:1 complex formed with 16-mer DNA than in 2:1 and 3:1 complexes formed with 11, 12, and 14 bp DNAs (Figure 5B). The resulting tight-packing of proteins accounts for the dramatic phasing of S_{app} and cooperativity (ω) with DNA length that we have observed. It also allows us to make testable predictions about the protein interactions that maintain the complex *in vitro* and are likely to dominate its functions *in vivo*. A few of these are outlined below.

First, AGT is likely to bind relaxed-duplex DNA with greater affinity than it binds positively or negatively supercoiled molecules. If locally stiff duplex DNA constrains the protein–protein contacts to a suboptimal geometry compared to those formed on torsionally more-flexible single-stranded DNA, the free energy cost should be reflected in a reduction

⁴ In the structures of Duguid et al. (18) and Daniels et al. (17), AGT contacts 4 bp in the minor groove of slightly distorted B-DNA. The occluded site size $s = 4$ bp results in binding sites that do not overlap on the minor groove surface. This justifies the use of eq 5, which describes nonoverlapping cooperative binding, in our analyses.

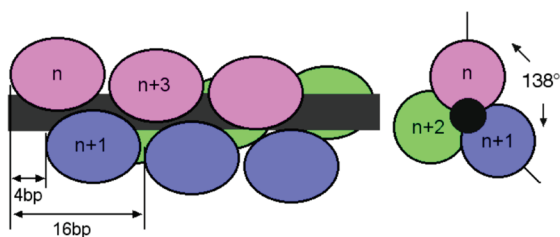


FIGURE 6: A model of binding topology for the cooperative array. Proteins are shown schematically as filled ovals and the DNA as a dark gray rod. Drawings are not to scale. Left, overlapping binding viewed from perpendicular to the DNA axis. Proteins are arranged as a three-start helix with individual arrays colored magenta, blue, and green. Each protein is rotated $\sim 138^\circ$ with respect to its nearest neighbors, as predicted if all AGT molecules make identical contacts with the minor groove of B-form DNA, and binding sites are separated by four base pairs. This view emphasizes the n -to- $n+3$ juxtaposition predicted for proteins along the same face of the DNA cylinder. Right: view down the long axis of the DNA. Only the first three proteins are shown for clarity. This view emphasizes the rotational juxtaposition of proteins n , $n+1$, and $n+2$.

in ω on duplex templates compared to the value on single-stranded DNAs. Since experimental values of ω are nearly the same for single-stranded and duplex DNAs (19) this work), it seems unlikely that the helical period of the DNA that is optimal for protein–protein contacts differs much from that of relaxed B-DNA. If correct, this prediction is likely to be functionally relevant. Since chromatin imposes topological constraints on genomic DNA and since transcription and DNA replication have the potential to impose torsional stress, we predict that AGT- and topoisomerase-activities will be found to cosegregate within chromosomal domains.

Second, cooperative interactions are likely to influence the rates of AGT binding and dissociation from target DNAs and hence, the kinetics of lesion search and repair. The portion of a cooperative array of bound proteins that is active in protein binding and/or dissociation is likely to influence the kinetics of these processes. Our current thinking is influenced by the notion that the addition or removal of a protein unit from the middle of a cooperative assembly is likely to be slow compared to the addition or removal of a unit from one end. If this is the case, the rate of transfer of AGT molecules between DNA segments may depend more strongly on the concentration of cooperative assemblies (and hence the concentration of end-monomers) than on the concentration of DNA-bound AGT monomers, themselves.

Finally, unless the rate of alkyl transfer greatly exceeds the rate of forming a cooperative complex, DNA-repair rates are likely to depend on the length of the cooperative complex. This is because repair requires the correct juxtaposition of an active AGT monomer and the damaged base. If the correct juxtaposition of AGT and lesion is not achieved during formation of the protein–DNA complex, repair may require repositioning of the entire protein array. If this is the case, productive binding and repair might be more rapidly achieved with short protein arrays than with long. Tests of these predictions are currently underway.

REFERENCES

1. Loveless, A. (1969) Possible relevance of O⁶ alkylation of dextroguanosine to the mutagenicity and carcinogenicity of nitrosamines and nitrosamides. *Nature (London)* 223, 206–207.
2. Pegg, A. E. (1984) Methylation of the O⁶ position of guanine in DNA is the most likely initiating event in carcinogenesis by methylating agents. *Cancer Invest.* 2, 223–231.
3. Pegg, A. E., Dolan, M. E., and Moschel, R. C. (1995) Structure, function and inhibition of O⁶-alkylguanine-DNA alkyltransferase. *Prog. Nucleic Acid Res. Mol. Biol.* 51, 167–223.
4. Pegg, A. E. (1990) Mammalian O⁶-alkylguanine-DNA alkyltransferase: regulation and importance in response to alkylating carcinogens and therapeutic agents. *Cancer Res.* 50, 6119–6129.
5. Margison, G. P., and Santibáñez-Koref, M. F. (2002) O⁶-alkylguanine-DNA alkyltransferase: role in carcinogenesis and chemotherapy. *BioEssays* 24, 255–266.
6. Gerson, S. L. (2002) Clinical relevance of MGMT in the treatment of cancer. *J. Clin. Oncol.* 20, 2388–2399.
7. Pegg, A. E. (2000) Repair of O(6)-alkylguanine by alkyltransferases. *Mutat. Res.* 462, 83–100.
8. Rabik, C. A., Njoku, M. C., and Dolan, M. E. (2006) Inactivation of O⁶-alkylguanine DNA alkyltransferase as a means to enhance chemotherapy. *Cancer Treat. Rev.* 32, 261–276.
9. Ranson, M., Middleton, M. R., Bridgewater, J., Lee, S. M., Dawson, M., Jowle, D., Halbert, G., Waller, S., McGrath, H., Gumbrell, L., McElhinney, R. S., Donnelly, D., McMurry, T. B., and Margison, G. P. (2006) Lomeguatrib, a potent inhibitor of O⁶-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.
10. Verbeek, B., Southgate, T. D., Gilham, D. E., and Margison, G. P. (2008) O⁶-Methylguanine-DNA methyltransferase inactivation and chemotherapy. *Br. Med. Bull.* 85, 17–33.
11. Zaboikin, M., Srinivasakumar, N., and Schuening, F. (2006) Gene therapy with drug resistance genes. *Cancer Gene Ther.* 13, 335–345.
12. Hunter, C., Smith, R., Cahill, D. P., Stephens, P., Stevens, C., Teague, J., Greenman, C., Edkins, S., Bignell, G., Davies, H., O'Meara, S., Parker, A., Avis, T., Barthorpe, S., Brackenbury, L., Buck, G., Butler, A., Clements, J., Cole, J., Dicks, E., Forbes, S., Gorton, M., Gray, K., Halliday, K., Harrison, R., Hills, K., Hinton, J., Jenkinson, A., Jones, D., Kosmidou, V., Laman, R., Lugg, R., Menzies, A., Perry, J., Petty, R., Raine, K., Richardson, D., Shepherd, R., Small, A., Solomon, H., Tofts, C., Varian, J., West, S., Widaa, S., Yates, A., Easton, D. F., Riggins, G., Roy, J. E., Levine, K. K., Mueller, W., Batchelor, T. T., Louis, D. N., Stratton, M. R., Futreal, P. A., and Wooster, R. (2006) A hypermutation phenotype and somatic MSH6 mutations in recurrent human malignant gliomas after alkylator chemotherapy. *Cancer Res.* 66, 3987–3991.
13. Chan, C. L., Wu, Z., Ciardelli, T., Eastman, A., and Bresnick, E. (1993) Kinetic and DNA-binding properties of recombinant human O⁶-methylguanine-DNA methyltransferase. *Arch. Biochem. Biophys.* 300, 193–200.
14. Fried, M. G., Kanugula, S., Bromberg, J. L., and Pegg, A. E. (1996) DNA binding mechanisms of O⁶-alkylguanine-DNA alkyltransferase: stoichiometry and effects of DNA base composition and secondary structures on complex stability. *Biochemistry* 35, 15295–15301.
15. 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.
16. Xu-Welliver, M., and Pegg, A. E. (2002) Degradation of the alkylated form of the DNA repair protein, O⁶-alkylguanine-DNA alkyltransferase. *Carcinogenesis* 23, 823–830.
17. 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.
18. 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.
19. 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.
20. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.
21. Nesterenko, M. V., Tilley, M., and Upton, S. J. (1994) A simple modification of Blum's silver stain method allows for 30 minute detection of proteins in polyacrylamide gels. *J. Biochem. Biophys. Methods* 28, 239–242.
22. Kanugula, S., Goodtzova, K., Edara, S., and Pegg, A. E. (1995) Alteration of arginine-128 to alanine abolishes the ability of human O⁶-alkylguanine-DNA alkyltransferase to repair methylated DNA

- but has no effect on its reaction with O⁶-benzylguanine. *Biochemistry* 34, 7113–7119.
23. 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 O⁶-alkylguanine-DNA alkyltransferase. *Biochemistry* 42, 980–990.
 24. Fang, Q., Kanugula, S., and AE, P. (2005) Function of domains of human O⁶-alkylguanine-DNA alkyltransferase. *Biochemistry* 44, 15396–15405.
 25. Maxam, A., and Gilbert, W. S. (1977) A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U.S.A.* 74, 560–565.
 26. Cantor, C. R., and Tinoco, I., Jr. (1965) Absorption and optical rotatory dispersion of seven trinucleoside diphosphates. *J. Mol. Biol.* 13, 65–77.
 27. Cavaluzzi, M. J., and Borer, P. N. (2004) Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. *Nucleic Acids Res.* 32, e13.
 28. Mandelkern, M., Elias, J. G., Eden, D., and Crothers, D. M. (1981) The dimensions of DNA in solution. *J. Mol. Biol.* 152, 153–161.
 29. Simpson, R. T., Thoma, F., and Brubaker, J. M. (1985) Chromatin reconstituted from tandemly repeated cloned DNA fragments and core histones: A model system for study of higher order structure. *Cell* 42, 799–808.
 30. Maxam, A. M., and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65, 449–560.
 31. Fried, M. G., and Crothers, D. M. (1981) Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* 9, 6505–6525.
 32. Zeman, S. M., and Crothers, D. M. (2001) Simultaneous measurement of binding constants and unwinding angles by gel electrophoresis. *Methods Enzymol.* 340, 51–68.
 33. Daugherty, M. A., and Fried, M. G. (2005) Protein-DNA Interactions at Sedimentation Equilibrium, in *Modern Analytical Ultracentrifugation: Techniques and Method* (Scott, D., Ed.) pp 195–209, Royal Society of Chemistry, Oxford.
 34. Cohn, E. J., and Edsall, J. T. (1943) Proteins, Amino Acids and Peptides As Ions and Dipolar Ions, in *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions* (Cohn, E. J., and Edsall, J. T., Eds.) pp 370–381 and 428–431, Reinhold, New York.
 35. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-Aided Interpretation of Analytical Sedimentation Data for Proteins, in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Harding, J. C., Eds.) pp 90–125, The Royal Society of Chemistry, Cambridge, England.
 36. 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.
 37. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) Analysis of data from the analytical ultracentrifuge by non-linear least squares techniques. *Biophys. J.* 36, 575–588.
 38. 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.
 39. Tsodikov, O. V., Holbrook, J. A., Shkel, I. A., and Record, M. T., Jr. (2001) Analytic binding isotherms describing competitive interactions of a protein ligand with specific and nonspecific sites on the same DNA oligomer. *Biophys. J.* 81, 1960–1969.
 40. Adams, C., and Fried, M. G. (2007) Analysis of Protein-DNA Equilibria by Native Gel Electrophoresis, in *Protein Interactions: Biophysical Approaches For The Study of Multicomponent Systems* (Schuck, P., Ed.) pp 417–446, Academic Press, New York.
 41. Epstein, I. R. (1978) Cooperative and non-cooperative binding of large ligands to a finite one-dimensional lattice. A model for ligand-oligonucleotide interactions. *Biophys. Chem.* 8, 327–339.
 42. Liem, L. K., Wong, C. W., Lim, A., and Li, B. F. L. (1993) Factors influencing the repair of the mutagenic lesion O⁶-methylguanine in DNA by human O⁶-methylguanine-DNA methyl-transferase. *J. Mol. Biol.* 231, 950–959.
 43. 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.
 44. Pegg, A. E., and Dolan, M. E. (1989) Investigation of Sequence Specificity in DNA Alkylation and Repair Using Oligodeoxynucleotide Substrates, in *DNA Repair Mechanisms and Their Biological Implications in Mammalian Cells* (Lambert, M. W., and Laval, J., Eds.) pp 45–59, Plenum Press, New York.
 45. Bender, K., Federwisch, M., Loggen, U., Nehls, P., and Rajewsky, M. F. (1996) Binding and repair of O⁶-ethylguanine in double-stranded oligodeoxynucleotides by recombinant human O⁶-alkylguanine-DNA alkyltransferase do not exhibit significant dependence on sequence context. *Nucleic Acids Res.* 24, 2087–2094.
 46. Burington, R. S. (1973) *Handbook of Mathematical Tables and Formulas*, 5th ed., McGraw-Hill, New York.

BI801666C