DNA Sequence Context and Protein Composition Modulate HMG-Domain Protein Recognition of Cisplatin-Modified DNA†

Shari U. Dunham and Stephen J. Lippard*

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received April 23, 1997; Revised Manuscript Received June 30, 1997⊗

ABSTRACT: Proteins containing the high mobility group (HMG) DNA-binding domain form specific complexes with cisplatin-modified DNA which shield the major intrastrand d(GpG) and d(ApG) cross-links from excision repair. The molecular basis for the specificity of binding was investigated for the two isolated domains of HMG1 with a series of 15-bp oligonucleotides, d(CCTCTCN$_1$G*G*N$_2$TCTTC)·(GAAGAN$_3$CCN$_4$GAGAGG), where asterisks denote N$_7$-modification of guanosine with cisplatin. Alteration of the nucleotides flanking the platinum lesion modulated HMG1domA recognition in this series by over 2 orders of magnitude and revealed an unprecedented preference for N$_2$ = dA > T > dC. The flanking nucleotide preference for HMG1domB interaction with this oligonucleotide series was less pronounced and had only a 20-fold range of binding affinities. For the N$_1$ = N$_2$ = dA 15-bp probe, 100-fold stronger binding occurred with HMG1domA (K_d = 1.6 ± 0.2 nM) compared to HMG1domB (K_d = 134 ± 18 nM). The platinum-dependent recognition of the N$_1$ = N$_2$ = dA 15-bp probe saturates at 1 equiv of HMG1domA and is highly specific, as evidenced by the 1000-fold decrease in HMG1domA binding affinity for the corresponding unplatinated oligonucleotide. HMG domains were unable to bind specifically to cisplatin-modified DNA-RNA hybrids, revealing the need for a deoxyribose sugar backbone for specific complex formation with HMG-domain proteins. Protein−DNA contacts which may account for these observed binding preferences are proposed, and potential implications for the biological processing of cisplatin-DNA adducts are discussed.

Protein recognition of specific DNA sequences and structures triggers many cellular functions. The activity of the anticancer drug cisplatin has been attributed to its ability to bind to and modify the structure of DNA, forming adducts that are specifically recognized by cellular proteins (*1*). Cisplatin forms mainly 1,2-intrastrand cross-links at d(GpG) and d(ApG) sequences in duplex DNA (*2*, *3*), the formation and persistence of which correlate with its cytotoxicity and anticancer activity (*4*). The processing of cisplatin−DNA adducts by the cell involves their recognition by proteins including those required for replication, transcription, repair and apoptosis (*5*). Accordingly, considerable attention has been paid to identifying and characterizing human proteins which bind with high affinity and specificity to the major cisplatin intrastrand cross-links. The present study focuses on the DNA-binding domains of the HMG1 protein, which recognize cisplatin-modified DNA (*6*−*8*) and shield the major 1,2-intrastrand d(GpG) and d(ApG) adducts from excision repair (*9*).

Nearly all proteins in the HMG-1/-2 family bind specifically to cisplatin-induced 1,2-intrastrand cross-links in DNA (*10*). Proteins in this family contain at least one copy of a mildly conserved region of 80 amino acids known as the HMG[1] domain (*11*, *12*). These HMG-domain proteins exhibit the common characteristics of binding to distortable, usually bendable, motifs such as four-way junction DNA with little or no DNA sequence preference (*13*, *14*). Members of one class of HMG-domain proteins, including HMG1, HMG2, and hUBF, contain two or more consecutive HMG domains and exhibit structure-specific DNA interactions. Another class, which comprises many tissue-specific transcription factors such as the lymphoid enhancer-binding factor 1 (LEF-1) and the testis-determining factor (SRY), has the additional ability to recognize a specific DNA sequence and induce sharp bends (>100°) in the duplex (*11*, *12*, *15*).

Although only limited structural information is available for HMG-domain interactions with cisplatin-modified DNA (*8*, *16*−*18*), several high-resolution solution studies of isolated HMG domains (*14*, *19*−*22*) and of HMG domains bound to unmodified duplex DNA (*19*, *23*, *24*) have been carried out. These investigations reveal that, despite mild (∼30%) sequence conservation, the L-shaped fold of the α-helical domain is highly conserved, even across protein classes and in the presence of DNA. The concave surface of the HMG domain, including residues in the extended N-terminal region, in helices I and II, and at the C-terminus (Figure 1), contacts the oligonucleotide. A single binding mode in the DNA minor groove occurs for the sequence-specific HMG domains (*23*, *24*), but multiple binding modes may be present in solution for structure-specific HMG domains binding to duplex DNA (*19*).

HMG-domain protein recognition of cisplatin-modified DNA, like four-way junction DNA, is based on morphologi-

† Supported by Grant CA34992 from the National Cancer Institute.
* Author to whom correspondence should be addressed.
⊗ Abstract published in *Advance ACS Abstracts,* August 15, 1997.

[1] Abbreviations: HMG, high mobility group; cisplatin, *cis*-diammine-dichloroplatinum(II); HMG1domA, domain A of HMG1; HMG1domB, domain B of HMG1; bp, base pair; *E. coli*, *Eschericia coli*; Tris, (tris-[hydroxymethyl]aminomethane); EDTA, ethylenediaminetetraacetic acid; HEPES, (*N*-[2-hydroxyethyl]piperazine-*N*′-[2-ethanesulfonic acid]); BSA, bovine serum albumin; PBS, phosphate buffered saline; cpm, counts per minute.

S0006-2960(97)00945-8 CCC: $14.00 © 1997 American Chemical Society

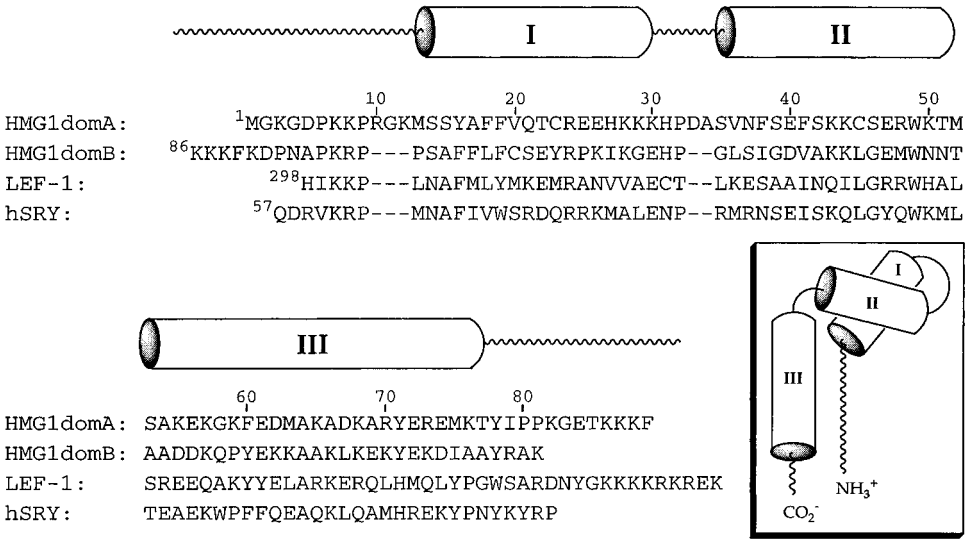

FIGURE 1: Sequence alignment and secondary structure of the HMG domains from several proteins. Numbers at the start of each amino acid sequence represent the residue numbering in the full length proteins.

cal changes in the duplex. The key platinum-induced distortions of DNA, which may facilitate HMG-domain protein recognition, include (i) destacking of the platinum-modified bases (*25*); (ii) bending of the helix toward the major groove with the bend locus at the platinum lesion (*16*, *26*); (iii) unwinding of the duplex (*27*); and (iv) widening and concomitant shallowing of the minor groove (*28*). A similarly widened, shallow minor groove is characteristic of A-type duplex oligonucleotides commonly encountered in RNA (*29*) and in regions of DNA−RNA hybrid helices, which are biologically active in transcription, replication, and reverse transcription (*30*−*34*). Since the platinum lesion can induce A-type helix characteristics in duplex DNA (*28*) and, in doing so, recruit minor-groove binding HMG-domain proteins, the possibility that cisplatin-modified DNA−RNA hybrids could have enhanced affinity for HMG-domain proteins was also investigated here.

In previous studies of HMG-domain protein binding to cisplatin-modified DNA, the platinated DNA probe was either globally modified and thus had a variety of adducts or site-specifically modified and contained a single 1,2-d(GpG), 1,2-d(ApG), or 1,3-d(GpTpG) cross-link embedded in a region of DNA of constant sequence context. Because footprinting studies reveal that the interaction of HMG-domain proteins with cisplatin-modified DNA extends over more than a turn of duplex (*8*, *35*, *36*), we were interested to determine whether the DNA sequence flanking the platinum lesion would influence the affinity and specificity of the resulting complex. The presence of a site-specific platinum lesion precludes the use of current combinatorial methodologies to optimize the flanking DNA sequence. A series of individual 15-bp oligonucleotides, d(CCTCTCN$_1$G*G*N$_2$TCTTC)·(GAAGAN$_3$CCN$_4$GAGAGG), where asterisks denote N$_7$-modification of guanosine with cisplatin, N$_1$ and N$_2$ = dA, dC or T, and N$_3$ and N$_4$ are their Watson−Crick complements, was therefore prepared to study the effects of DNA sequence on HMG-domain binding to cisplatin-modified DNA. Although 16 DNA sequences are possible by varying the sites flanking the central d(GpG) core, only nine were investigated. The seven sequences not examined contain three or more consecutive guanine residues, which renders difficult the isolation of pure N$_1$G*G*N$_2$ products. The information obtained here should be useful both for the design of potential anticancer drug candidates in the platinum family and for guiding experiments to determine the structures of the complexes by NMR or X-ray diffraction methods.

EXPERIMENTAL PROCEDURES

HMG-Domain Proteins. HMG1domB, K86-K165 of rat HMG1 (*37*), was expressed in *E. coli* (BL21-DE3) from a previously reported plasmid (*6*). The domain was isolated and purified as described (*6*) with the addition of a final FPLC size-exclusion purification step (high-load Superdex 75, Pharmacia, 1 mL/min, 11.8 mM PBS, pH 7.4). HMG1domA, M1-F89 of rat HMG1, was expressed in the same bacterial strain from the pT7HMG1bA plasmid (*38*). Cells were grown and harvested according to published procedures (*38*), and the protein was purified by using the same protocol as for HMG1domB.

Protein concentrations were determined by optical absorption at 278 nm using extinction coefficients of 14 000 M^{-1} cm^{-1} and 12 160 M^{-1} cm^{-1} for HMG1domA and HMG1domB, respectively. The extinction coefficient for each protein was determined from the combined measurements of optical absorbance at 278 nm and protein concentrations from amino acid analyses (MIT Biopolymers Lab). The isoelectric point (pI) for each protein sequence was determined by using the Pepsort algorithm available in the GCG program (*39*). The sequence alignment of isolated HMG domains (Figure 1) was performed by using the AMPS program (*40*).

Oligonucleotide Probes. Tables 1 and 2 list the oligonucleotides used in this study together with their abbreviations. Oligonucleotides were synthesized in 0.2 or 1.0 μmol quantities on a Cruachem PS250 DNA synthesizer by employing conventional solid support phosphoramidite chemistry (*41*). The fully protected oligoribonucleotide r(gaagaaccagagagg) and the d(GAAGA)r(accagagagg) chimera were purified by using reversed phase HPLC (C_{18} radial pak cartridge, Waters, nonlinear 0 to 100% methanol gradient in 200 mM NaOAc buffer, pH 7.2) prior to deprotection (*42*). These oligonucleotides were further purified on a 20% denaturing polyacrylamide gel (7.5 M urea, 19:1 acrylamide:bisacrylamide, 90 mM TRIS-borate, 1 mM EDTA, pH 8.3, 300 V, ~10 h) and desalted prior to use.

Table 1: Deoxyribonucleotide Sequences and Abbreviations[a]

abbreviation	duplex oligonucleotide sequence
AGGA	5'-CCTCTC**A**GG**A**TCTTC-3'
	3'-GGAGAG**T**CC**T**AGAAG-5'
AGGT	5'-CCTCTC**A**GG**T**TCTTC-3'
	3'-GGAGAG**T**CC**A**AGAAG-5'
AGGC	5'-CCTCTC**A**GG**C**TCTTC-3'
	3'-GGAGAG**T**CC**G**AGAAG-5'
TGGA	5'-CCTCTC**T**GG**A**TCTTC-3'
	3'-GGAGAG**A**CC**T**AGAAG-5'
TGGT	5'-CCTCTC**T**GG**T**TCTTC-3'
	3'-GGAGAG**A**CC**A**AGAAG-5'
TGGC	5'-CCTCTC**T**GG**C**TCTTC-3'
	3'-GGAGAG**A**CC**G**AGAAG-5'
CGGA	5'-CCTCTC**C**GG**A**TCTTC-3'
	3'-GGAGAG**G**CC**T**AGAAG-5'
CGGT	5'-CCTCTC**C**GG**T**TCTTC-3'
	3'-GGAGAG**G**CC**A**AGAAG-5'
CGGC	5'-CCTCTC**C**GG**C**TCTTC-3'
	3'-GGAGAG**G**CC**G**AGAAG-5'

[a] Bold face indicates basepair alterations.

Table 2: Ribose-Containing Oligonucleotide Sequences and Abbreviations[a]

hTGGT	5'-CCTCTCTGGTTCTTC-3'
	3'-ggagagaccaagaag-5'
cTGGT	5'-CCTCTCTGGTTCTTC-3'
	3'-GGAGAgaccaagaag-5'

[a] Upper case and lower case lettering denotes deoxyribonucleotides and ribonucleotides, respectively.

Deoxyribonucleotides were synthesized with the final trityl group cleaved, purified by using anion exchange HPLC (Dionex NucleoPac PA-100 9 × 250 mm column, 200 to 400 mM NaCl gradient in 25 mM NH_4OAc, pH 6.0, 10% acetonitrile), and desalted on a G25 Sephadex (Pharmacia) column prior to use. All oligonucleotides containing a single d(GpG) target site for platinum modification were allowed to react with 1.0−1.25 equiv of *cis*-$[Pt(NH_3)_2(H_2O)_2]^{2+}$ (*28*) in 10 mM sodium phosphate buffer, pH 6.8, for several hours at 37 °C. Platination reactions were monitored by anion exchange HPLC (Dionex NucleoPac PA-100 4 × 250 mm column, 50 to 400 mM NaCl gradient in 25 mM NH_4OAc, pH 6.0, 10% acetonitrile), and the major product was isolated in each case.

The modified oligonucleotides were determined by atomic absorption spectroscopy (Varian 1475 graphite furnace spectrometer operating in peak height mode at 265.9 nm) to have one platinum atom per strand (Table S1, Supporting Information). Formation of the desired 1,2-intrastrand adducts was confirmed by enzymatic digestion as described previously (*26*, *43*) (Figure S1, Supporting Information). Oligonucleotide strand concentrations were determined from A_{260} values and theoretically estimated extinction coefficients (*44*).

Single-stranded oligonucleotide probes (10−20 pmol) were 5′-end labeled with T4 polynucleotide kinase (20 units, NEB) and γ-^{32}P-ATP (50 μCi, NEN), purified on G25 Quick Spin columns (Boehringer Mannheim), and ethanol precipitated. Radiolabeled probe concentrations were determined as described (*45*). Radiolabeled duplexes were formed by annealing 10^6 cpm of one strand with 1.5−10 equiv of its complement in buffer I (10 mM HEPES, pH 7.5, 50 mM LiCl, and 10 mM $MgCl_2$) at 45 °C for 15 min followed by 10 h incubation at 4 °C. Unlabeled duplexes for competition experiments were annealed by combining equimolar amounts of the two complementary strands at concentrations of 250 μM to 1 mM in buffer I at 45 °C for 15 min followed by 10 h incubation at 4 °C.

Gel Mobility Shift Assays. Oligonucleotide duplexes (0.4−5.0 nM, 5000 cpm) were titrated with protein in 10 μL sample volumes in buffer II (10 mM HEPES, pH 7.5, 10 mM $MgCl_2$, 50 mM LiCl, 100 mM NaCl, 1 mM spermidine, 0.2 mg/mL BSA, and 0.05% Nonidet P40). For all gel mobility shift experiments, samples were incubated on ice for 1 h and made 7% in sucrose and 0.017% in xylene cyanol prior to loading on running, precooled (4 °C), prerun (300 V, 1−2 h) native polyacrylamide gels (29:1 acrylamide: bisacrylamide, 3.3% cross-linking, 22.5 or 45 mM Tris-borate, 1 mM EDTA, pH 8.3). Gels were electrophoresed for 1−2 h before drying under vacuum (80 °C, 1 h) and exposure to film (10−40 h). Quantitative measurements of bound and free oligonucleotide were performed by phosphorimager analysis (Molecular Dynamics PhosphorImager).

In order to determine the stoichiometry of binding, duplex oligonucleotides (100 nM, 5000 cpm) were titrated with 0−3 equiv of HMG1domA protein in buffer II. For competition assays, unlabeled competitor oligonucleotide probe (0−200 μM) was titrated against radiolabeled AG*G*A duplex (0.45 nM, 5000 cpm in buffer II) complexed with HMG1domA. By following a published protocol (*46*), both oligonucleotides were mixed prior to addition of the protein at a concentration which afforded 50−70% binding of the labeled probe in the absence of competitor.

Data Analysis. Apparent dissociation constants, K_d, were estimated from nonlinear least-squares fits of binding data to the Langmuir isotherm (eq 1; *47*), where θ is the fraction of bound oligonucleotide probe and P is the total protein concentration. Apparent relative dissociation constants, K_{rel},

$$\theta = \frac{P}{P + K_d} \tag{1}$$

were determined from nonlinear least-squares fits of the competition data to eq 2 (*48*), where K_{rel} is the ratio of K_t, the apparent dissociation constant of the labeled probe, to K_c, the apparent dissociation constant of the competitor, and P_t, T_t, and C_t are the concentrations of protein, radiolabeled

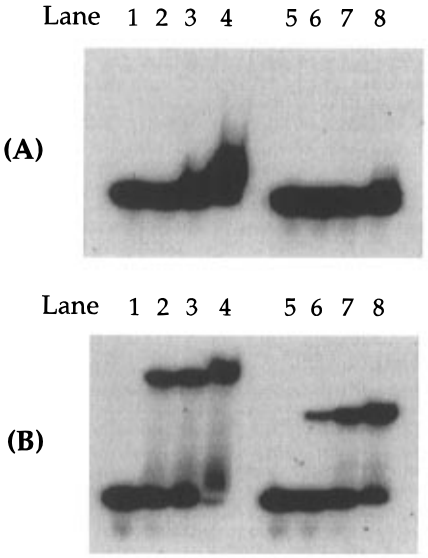

FIGURE 2: Gel mobility shift assays of (A) AGGA (0.45 nM) and (B) AG*G*A (0.45 nM) with HMG1domA (lanes 1−4) and HMG1domB (lanes 5−8). Protein concentrations in each set are 0, 10, 100, and 1000 nM.

probe, and competitor probe, respectively.

$$\theta = \frac{1}{2T_t}[K_t + K_{rel}C_t + P_t + T_t - \sqrt{(K_t + K_{rel}C_t + P_t + T_t)^2 - 4T_tP_t}] \quad (2)$$

The stoichiometry of binding was assessed from plots of θ versus protein equivalents. The point of intersection of two lines, one fit to presaturation and the other to postsaturation data, was taken as an indication of the stoichiometry of the protein−DNA complex.

RESULTS

Isolated HMG Domains of HMG1 Bind to a Cisplatin-Modified 15-bp DNA Duplex. A short duplex oligonucleotide, AGGA (Table 1), both with and without a single *cis*-$\{Pt(NH_3)_2\}^{2+}$ d(GpG) modification, was investigated for binding to HMG1domA and HMG1domB. The 15-bp length was chosen following preliminary gel mobility shift assays of HMG1domB with platinated 12-, 15- and 20-bp DNAs (data not shown). The 15-bp oligonucleotide was the minimal length that gave a good bandshift, consistent with the 14−15-bp footprint found by nuclease digestion of several HMG-domain proteins bound to cisplatin-modified DNA (*8*, *35*, *36*).

HMG1domA and HMG1domB exhibited minimal binding to the unmodified 15-bp AGGA duplex (Figure 2A). Smearing of the band corresponding to unplatinated probe occurred with increasing HMG1domA concentrations (Figure 2A, lanes 2−4) and was attributed to nonspecific (platinum-independent) interactions which are minimal for HMG1domB. This nonspecific binding may be electrostatic in origin, reflecting the greater positive charge of HMG1domA compared to HMG1domB at pH 7.5 (Table 3).

As indicated by the presence of a shifted band which increases with increasing protein concentration (Figure 2B), HMG1domA and HMG1domB both recognize the cisplatin-modified 15-bp duplex AG*G*A. Since only a single shifted band forms following incubation of dsAG*G*A with either HMG1domA or HMG1domB, detailed titration studies were carried out. The titration of dsAG*G*A with HMG1domB is shown at the top of Figure 3. Fits of the titration data for both HMG1domA and HMG1domB to eq 1 (Figure 3, bottom) afforded K_d values reported in Table 4. The dsAG*G*A probe was also titrated with stoichiometric amounts of HMG1domA. A plot of θ versus equivalents of protein (Figure S2) was linear at both pre- and postsaturation limits. The intersection of the best fit lines to the experimental data indicated that 1.0 ± 0.2 equiv of HMG1domA was required for saturation of binding to the dsAG*G*A probe.

Because a discrete, shifted band did not appear in gel mobility shift studies of unplatinated dsAGGA, dsCGGC, ssAGGA, or platinated ssAG*G*A with isolated HMG-domain proteins, nonspecific interactions formed by these probes could not be directly quantitated. Instead, platinum- and duplex-independent binding of HMG-domain proteins to oligonucleotide probes was investigated by using competition assays. As indicated in Figure 4 (top), addition of unlabeled AG*G*A duplex competes away the binding of HMG1domA to labeled dsAG*G*A. A plot quantitating these competition data is presented in Figure 4 (bottom),

Table 3: Selected Characteristics of Isolated HMG Domains

protein	pI[a]	net charge[b]	mass (kD)
HMG1domA	10.6	+12	10.4
HMG1domB	10.5	+9	9.2

[a] Calculated with the GCG Sequence Analysis Software Package, Version 7.0. [b] Charge calculated at pH 7.5.

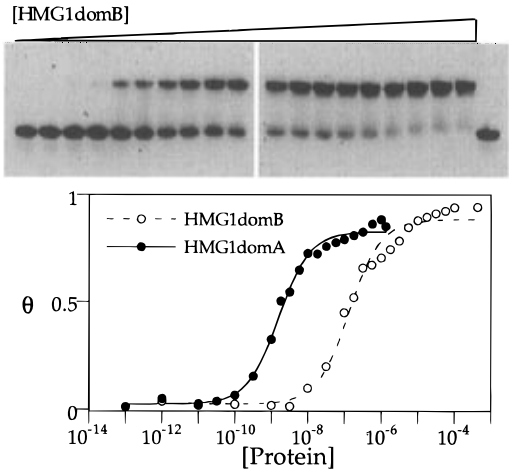

FIGURE 3: Gel mobility shift assay analysis of the titration of AG*G*A (0.45 nM) with HMG1domB (0.1 nM to 178 μM) (top). Plot of the fraction of bound DNA vs [HMG1domA] (closed circles) or [HMG1domB] (open circles) with the superimposed fits to eq 1 (bottom).

Table 4: $K_{d(app)}$ Values Determined for HMG-Domain Protein Interactions with Oligonucleotides[a]

oligonucleotide	HMG1domA	HMG1domB (nM)
dsAG*G*A	1.6 ± 0.2 nM 6.8 ± 0.8 nM[b]	134 ± 18
dsAG*G*T		48 ± 9
dsTG*G*C	127 ± 17 nM	
dsCG*G*A	8 ± 3 nM	1080 ± 235
dsCG*G*C	517 ± 60 nM	1300 ± 190
dsAGGA	1.6 ± 0.3 μM[b]	
dsCGGC	1.6 ± 0.2 μM[b]	

[a] Measured by direct titration unless otherwise specified. [b] Measured from competition experiments.

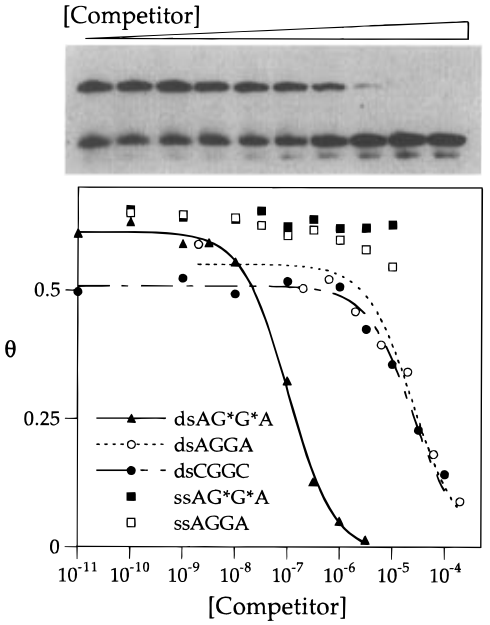

FIGURE 4: Gel mobility shift assay analysis of the self-competition of dsAG*G*A-HMG1domA complex with unlabeled competitor dsAG*G*A (top). Plot of the fraction of bound DNA vs competitor concentration for dsAG*G*A (triangles), dsAGGA (open circles), dsCGGC (closed circles), ssAGGA (open squares), and ssAG*G*A (closed squares) competitors with the dsAG*G*A−HMG1domA complex (bottom). Fits of the data to eq 3, where possible, are superimposed.

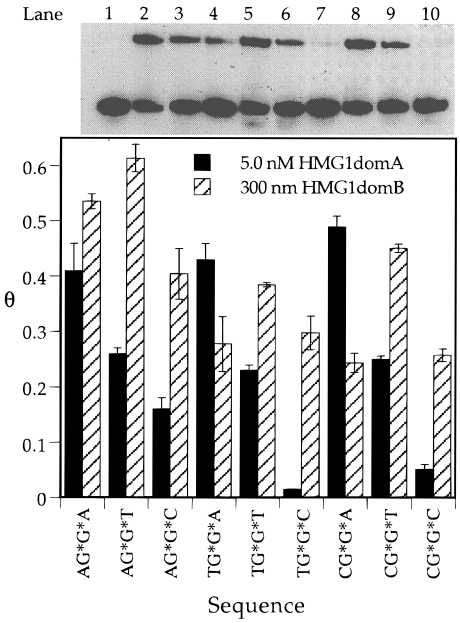

FIGURE 5: Gel mobility shift assay of each DNA sequence (N_1G*G*N_2, Table 1) with 5.0 nM HMG1domA (top). Lane 1 contains AG*G*A without protein. Lanes 2−10 contain the oligonucleotides in the order in which they appear along the ordinate below. Bar graph illustrating the fraction of bound DNA (θ) for each sequence (bottom) at 5.0 nM HMG1domA (black bars) and 300 nM HMG1domB (hatched bars). Values are the average of three independent experiments and error bars represent ±1 standard deviation.

together with competition data from analogous experiments in which dsAGGA, dsCGGC, ssAGGA, or ssAG*G*A were used as competitors. The dissociation constants for the dsAG*G*A, dsAGGA, and dsCGGC competitors, as determined from the best fits of these data to eq 2, are given in Table 4. Because the single-stranded oligonucleotides were such poor competitors, results for these oligonucleotides could not be accurately fit in the same manner.

Sequence Preference for HMG-Domain Binding to Cisplatin-Modified DNA Duplexes. The series of nine cisplatin-modified DNA duplexes (Table 1), which differ in base pair composition directly adjacent to the platinum lesion, were screened for binding by HMG1domA and HMG1domB. Figure 5 (top) shows the gel mobility shift assay of each oligonucleotide in this series with HMG1domA at a constant protein concentration. Qualitative comparisons of θ for each oligonucleotide, averaged from three independent experiments, with 5 nM HMG1domA or 300 nM HMG1domB are presented at the bottom of Figure 5. Titration experiments were carried out for the strongest and weakest binding oligonucleotides for both proteins (data not shown), and K_d values (Table 4) were determined from fits of these data to eq 1.

HMG1domA and HMG1domB Do Not Bind to DNA-RNA Heteroduplexes. None of the three duplexes, DNA duplex TGGT, DNA-RNA hybrid hTGGT, or DNA-RNA chimeric hybrid cTGGT (Table 2), was recognized specifically by either isolated domain of HMG1 (data not shown). The additional distortions of these duplexes by a site-specific, central platinum intrastrand d(GpG) cross-link imparted HMG-domain binding only for an all DNA duplex (Figure 6). The incorporation of ribose sugars abolished protein binding similarly for both HMG1domA (Figure 6A) and HMG1domB (Figure 6B). To exclude the possibility that the absence of protein-binding may be a result of hTG*G*T or cTG*G*T duplex instability, gels were run under more

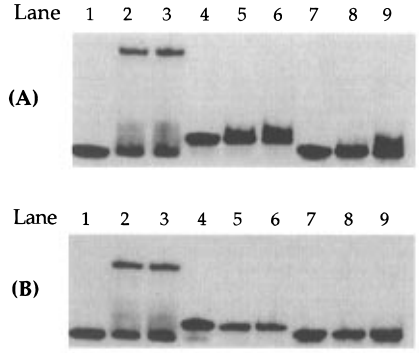

FIGURE 6: Gel mobility shift assays of 5 nM TG*G*T (lanes 1−3), hTG*G*T (lanes 4−6) and cTG*G*T (lanes 7−9) with (A) HMG1domA and (B) HMG1domB. Protein concentrations are 0 nM (lanes 1, 4, and 7), 250 nM (lanes 2, 5, and 8), and 500 nM (lanes 3, 6, and 9).

highly resolving conditions (data not shown). The TG*G*T, hTG*G*T and cTG*G*T duplexes exhibited decreased gel mobility compared to the labeled DNA single strand alone, indicating that double-stranded oligonucleotides were formed and stable under the experimental conditions.

DISCUSSION

*HMG1domA Binding to dsAG*G*A.* The present results demonstrate that a single, bifunctional platinum lesion can increase the affinity of an HMG-domain protein for a 15-bp DNA duplex by almost 3 orders of magnitude. In accord with previous studies of isolated HMG-domain interactions with cisplatin-modified DNA (*6*, *7*, *17*, *18*), HMG1domA formed a specific complex with dsAG*G*A (Figure 2B, lanes 1−4) but not with the corresponding unmodified duplex (Figure 2A, lanes 1−4). Under the conditions reported here, the interaction between HMG1domA and dsAG*G*A is very strong (K_d = 1.6 ± 0.2 nM), has 1:1 stoichiometry, and

decreases by a factor of ~1000 ($K_d = 1.6 \pm 0.3$ μM) when the bifunctional platinum lesion is not present (Figure 4). A 1000-fold specificity factor is unprecedented for the interaction between an isolated HMG domain and a site-specifically cisplatin-modified duplex and is greater than the specificity of full-length HMG1 binding to site-specifically cisplatin-modifed DNA (*49*).

*HMG1domA Binding to dsN_1G*G*N_2.* The interaction of HMG1domA with a site-specifically cisplatin-modified 15-bp DNA duplex is modulated by the nature of the base pairs flanking the platinum lesion (Figure 5). The apparent dissociation constants measured for HMG1domA binding to this series of cisplatin-modified duplexes range over 2 orders of magnitude from 1.6 ± 0.2 nM to 517 ± 60 nM (Table 4). For one of the weakest binding probes, dsCG*G*C, removal of the bifunctional platinum lesion resulted in only 2- to 3-fold loss in affinity ($K_d = 1.6 \pm 0.2$ μM, Table 4). This apparent dissociation constant, which represents the nonspecific binding of HMG1domA to dsCGGC, is the same as that observed for the interaction with dsAGGA (1.6 ± 0.3 μM, Table 4) and suggests that the platinum-independent binding of HMG1domA to duplex DNA is not affected by the DNA target sequence. These data support the proposal that structure-specific HMG domains may have more than one binding mode to duplex DNA (*18*, *50*). One mode, representing platinum-independent binding, is insensitive to duplex DNA sequence context, in agreement with all other studies of structure-specific HMG-domain proteins with unmodified duplex or four-way junction DNA (*15*). A second binding mode, which is specific for a cisplatin 1,2-intrastrand d(GpG) lesion in duplex DNA, is very sensitive to DNA sequence context surrounding the lesion site (this work). The *cis*-$\{Pt(NH_3)_2\}^{2+}$ 1,2-intrastrand d(ApG) cross-link is expected to behave similarly.

A striking trend was observed for the DNA sequence preference of HMG1domA binding to dsN_1G*G*N_2 oligonucleotides. The dominant base pair preference was apparent 3′ to the platinum lesion where, regardless of N_1, affinity decreased as N_2 = dA > T > dC (Figure 5). The base pair located 3′ to the platinum lesion has unique hydrogen-bonding and minor groove accessibility as shown from other studies of DNA site-specifically modified with cisplatin. In footprinting studies, irrespective of the nature of the base, the position immediately 3′ to the platinum lesion is one of only two positions in cisplatin-modified DNA which is hypersensitive to cleavage by the minor groove binding protein, DNase I (*8*, *51*, *52*). In addition, the solid state structure of a cisplatin-modified dodecamer (*28*) revealed disruption of hydrogen bonding at the base pair located 3′ to the lesion.

There are two likely explanations for the base pair preference of HMG1domA binding to cisplatin-modified DNA. The first is that the platinated DNA duplexes in solution in the absence of an HMG domain will be differentially flexible and/or bendable depending on the nature of the N_2-N_3 base pair (A-T, T-A, or C-G). Quantitation and comparison of duplex stabilities and high-resolution structures of site-specifically cisplatin-modified duplexes in which the nature of the 3′ base pair is varied are required to assess the extent of such contributions.

The second possibility is that the observed DNA sequence preferences may arise from base-specific protein−DNA contacts in the complex. The high-resolution structures available for hSRY (*23*) and LEF-1 (*24*) HMG domains in complex with consensus DNA binding sequences d(GCAC**AA**AC)·(GTTTGTGC) and d(GAGCTTC**AA**AGGGTG)·(CACCCTTTGAAGCTC), respectively, have many similar features and provide some clues about key protein−DNA contacts which may be responsible for the preferences observed at N_2 in the HMG1domA interaction with dsN_1G*G*N_2. In the SRY−DNA and LEF-1−DNA complexes, a hydrophobic residue at position 17 (Figure 1, HMG1domA numbering), an isoleucine in SRY, and a methionine in LEF-1, intercalates between two adjacent adenosine bases. These nucleotides, denoted by bold face in the above sequences, form the major bend locus of the duplex. The roll at the point of intercalation is comparable to that induced at the G*G* site of a cisplatin-modified duplex (*28*). An A-T base pair occurs in the SRY and LEF-1 DNAs just 3′ to the intercalated bases and corresponds to the A-T base pair which is preferred by HMG1domA at the N_2-N_3 position in the cisplatin-modified duplex. A hydrogen bond is formed between the keto oxygen atom at position 2 of the thymine base in this pair (T_{10} in SRY−DNA; T_{21} in LEF-1-DNA) and the protein residue at position 42 (HMG1domA numbering scheme, Figure 1) in both the SRY and LEF-1 domains. This protein residue is highly conserved as either serine or asparagine in all sequence-specific HMG domains. Although position 42 tends to be occupied by alanine, valine, or lysine in the structure-specific HMG domains, HMG1domA has a serine in this position. Like S94 of SRY and N330 of LEF-1, where numbering refers to the full-length proteins, S42 in HMG1domA may form a hydrogen bond to O_2 of this thymine base, accounting for the preference for dA at the N_2 position in dsN_1G*G*N_2.

*HMG1domB Binding to dsN_1G*G*N_2.* The second DNA-binding domain of HMG1, HMG1domB, also forms a specific complex with the cisplatin-modified 15-bp DNA, AG*G*A (Figure 2B, lanes 5−8), but not with the unmodified duplex AGGA (Figure 2A, lanes 5−8). The interaction of HMG1domB with dsAG*G*A is modest ($K_d = 134 \pm 18$ nM) but weaker than that measured for HMG1domA (Table 2). As observed with HMG1domA, the base pairs flanking the platinum lesion also modulate HMG1domB binding to a 15-bp cisplatin-modified duplex (Figure 5). The apparent dissociation constants measured for these interactions span more than 1 order of magnitude from 48 ± 9 nM to 1300 ± 190 nM (Table 4), with subtle but observable trends in base pair preferences. These preferences are not the same as those previously described for HMG1domA, however, since HMG1domB affinity decreased with N_1 = dA > T ≈ dC and with N_2 = T > dA ≥ dC (Figure 5, bottom).

Differences in DNA Binding of HMG1domA and HMG1domB. Factors such as protein size, net charge, and amino acid composition can contribute to the differences in binding affinities and DNA sequence preferences observed for HMG1domA and HMG1domB binding to platinated DNA. The gel mobility shift of the HMG1domB−dsAG*G*A complex is significantly less than that observed for the HMG1domA−dsAG*G*A complex (Figure 2B, compare lanes 4 and 8). Since gel mobility shift is a function of size and charge, the greater shift of the HMG1domA−AG*G*A complex may result from the larger size and higher net positive charge of HMG1domA as compared to HMG1domB (Table 3).

Although they have similar affinities for four-way junction DNA (*53*), HMG1domA and HMG1domB interact differ-

ently with the same cisplatin-modified duplex. Their K_d values for the duplex AG*G*A probe differ by almost 2 orders of magnitude, as do their affinities for the duplex CG*G*A probe (Table 4). These differences translate to ~2.5 kcal/mol of free energy of binding (Table S2) and correspond to those predicted by polyelectrolyte theory from the charges of HMG1domA and HMG1domB (*47*). HMG1domA also binds more tightly to the duplex CG*G*C probe when compared to HMG1domB, but the difference in affinities is less than a factor of 3 (~0.6 kcal/mol). This result indicates that protein net charge is not the only determinant and that protein composition must also modulate these DNA interactions.

In particular, the length of the basic C-terminal tail and the amino acid composition within the DNA-binding regions of these two HMG domains may affect their interactions with the platinum-modified DNA duplexes. The inclusion of a short, basic C-terminal region in HMG1domA may account for its greater affinity for dsN_1G*G*N_2 when compared to HMG1domB (Figure 1). Although this basic region did not affect the affinity of HMG1domA for globally platinated DNA (*7*), a similar but longer region in the LEF-1 domain spans the major groove of the consensus DNA target, establishing multiple contacts and increasing the affinity and the bend angle (*24*, *54*). Moreover, the basic region following the HMG domain in HMG-D, the *Drosophila* analog of HMG1, increases the protein affinity for four-way junction and linear DNA (*55*). The string of basic residues following HMG1domB in the full HMG1 sequence (*53*) was not included in the present HMG1domB construct, and its affect on binding or bending of cisplatin-modified DNA has not been reported.

Although the base pair preferences are not the same for HMG1domA and HMG1domB binding to dsN_1G*G*N_2, the best oligonucleotide probes for either domain have A/T base pairs flanking the platinum lesion (AG*G*A for HMG1domA and AG*G*T for HMG1domB). The fewer number of hydrogen bonds in an A/T base pair should increase the flexibility of the DNA near the platinum lesion, facilitating additional DNA bending caused by the protein (*16*). Increased bendability around the platinum lesion, however, does not explain all observed preferences, since HMG1domA prefers AG*G*A over TG*G*T and HMG1domB prefers AG*G*T over TG*G*A (Figure 5, bottom). Preferences for DNA bases surrounding the platinum lesion in a cisplatin-modified duplex were previously proposed to facilitate HMG-domain binding (*28*). Analysis of the solid-state structure of d(CCTCTG*G*TCTCC)·(GGAGACCAGAGG) suggested that, upon further bending, potential hydrogen-bonding interactions within the distorted duplex would be stabilized. In particular, it was suggested that duplex stability would be enhanced for N_1 = T > dA ≈ dC, which does not agree with the present experimental findings. It is clear that factors other than duplex flexibility and intraduplex hydrogen bonding are important in recognition of these cisplatin-modified DNAs by HMG domains.

The most notable difference between HMG1domA and HMG1domB preferences is at the base pair flanking the 3′ side of the platination site (N_2-N_3). Although both proteins prefer N_2 = T > dC (Figure 5), HMG1domA has a strong preference for N_2 = dA at this position, whereas HMG1domB does not. The previous argument highlighting the potential importance of the protein residue at position 42 (Figure 1, HMG1domA numbering), in addition to supporting the N_2 = dA preference of HMG1domA, also provides a potential explanation for the lack of such a preference with HMG1domB. HMG1domB, like many other structure-specific HMG domains, has an alanine residue at position 42, which is unlikely to form base-specific contacts at this apparently crucial base pair.

The composition of the isolated HMG domain does influence its ability to recognize a short, cisplatin-modified DNA duplex. If the DNA binding observed for isolated HMG1domA and HMG1domB is representative of their activity within the full-length HMG1 protein, then the present results suggest that these domains may not contribute equally to HMG1 recognition of cisplatin-modified DNA. Because HMG1domA has a higher affinity than HMG1domB for cisplatin intrastrand cross-links, independent of DNA flanking sequence, HMG1domA should dominate the interaction between HMG1 and cisplatin-modified DNA.

The observed HMG-domain preferences for DNA sequence context at a cisplatin intrastrand cross-link may need to be considered in proposed mechanisms of protein-mediated cisplatin antitumor activity. Until now, possible mechanisms that link HMG-domain proteins to the biological activity of cisplatin (reviewed in ref 1) have tacitly assumed that all cisplatin 1,2-intrastrand cross-links bind with similar affinity to a given HMG-domain protein. Since the DNA sequence context surrounding the cisplatin 1,2-intrastrand cross-link can modulate HMG-domain protein affinity by more than 2 orders of magnitude, a subset of such platinum lesions within optimum DNA sequence contexts may be more effectively shielded from repair and/or affect the natural functions of essential HMG-domain proteins. In this manner, the lethality of each platinum lesion on the genome may be dictated by the surrounding DNA sequence context as well as the particular bases to which platinum is coordinated. It is thus possible that other proteins, such as components of excision repair, mismatch repair, and the apoptotic apparatus, will recognize cisplatin 1,2-intrastrand cross-links differentially according to the DNA sequence context surrounding the platinum lesion.

DNA−RNA Hybrid Recognition. The presence of a cisplatin 1,2-intrastrand cross-link in a double-helical oligonucleotide is required for high-affinity binding of HMG-domain proteins. The lack of HMG-domain protein binding to a DNA−RNA hybrid (hTGGT) and to the chimeric hybrid (cTGGT) indicated that a propensity to form a more A-type helix conformation does not enhance protein-oligonucleotide affinity (Figure 6A). The substitution of ribose for deoxyribose sugars in these duplexes appears to offset any advantageous structural distortion caused by cisplatin modification which induces HMG1domA and HMG1domB recognition (Figure 6B). Although the 2′-hydroxyl groups of the ribose sugars, which line the minor groove of the double helix, have been implicated as important for protein recognition of DNA−RNA hybrids (*33*), they could well interfere with specific protein−oligonucleotide contacts which are essential for HMG-domain binding to the minor groove. If cisplatin modification of DNA−RNA hybrid or chimeric structures were important in the antitumor mechanism of the drug, then the present study suggests that HMG-domain proteins will not participate in such a mechanism.

Conclusion. This work provides the first evidence that variations in DNA and protein composition significantly affect the HMG-domain protein interaction with cisplatin-modified DNA in vitro. The large range in protein-binding

constants resulting from these modest changes in DNA sequence suggests that base variations farther from the platinum lesion, perhaps even those outside the 15-bp protein-binding region (*8*, *35*, *36*), may also affect HMG-domain protein binding (*56*). Sequence-dependent recognition of cisplatin−DNA adducts may also apply to cellular proteins outside of the HMG-1/-2 family, such as damage recognition proteins hMSH2 (*57*) and XPAC (*58*). In addition, alterations in flanking sequence may affect the binding of proteins to DNA damaged by agents other than cisplatin. Finally, these studies imply that, in vivo, the DNA sequence surrounding a platinum lesion in the genome, in addition to the nature of the platinum lesion, may be very important for recognition by cellular proteins. If cellular protein recognition and subsequent biological processing of cisplatin-modified DNA is active in the antitumor mechanism of this drug, then one may envision a generation of platinum drugs which optimally recruit cellular proteins by targeting not only GG sites but their specific sequence context as well.

ACKNOWLEDGMENT

We thank J. M. Morse and E. R. Jamieson for experimental assistance, R. Batey, S. Dunham, and E. Trimmer for helpful discussions and Johnson Matthey AESAR/Alfa Co. for a gift of cisplatin. We acknowledge M. E. Bianchi for providing the HMG1domA plasmid and G. L. Verdine and L. Chen for the *E. coli* BL21(DE3) strain.

SUPPORTING INFORMATION AVAILABLE

Measured platinum levels of oligonucleotides used in this study (Table S1); differences in free energies of binding for isolated HMG-domain interactions with cisplatin-modified 15-bp DNA duplexes (Table S2); HPLC traces of enzymatic digestion products of cisplatin-modified and unmodified oligonucleotides (Figure S1, panels A, B, and C); and titration binding data for dsAG*G*A and HMG1domA (Figure S2) (6 pages). Ordering information is given on any current masthead page.

REFERENCES

1. McA'Nulty, M. M., and Lippard, S. J. (1995) *Nucleic Acids Mol. Biol. 9*, 264−284.
2. Eastman, A. (1987) *Pharmacol. Ther. 34*, 155−166.
3. Fichtinger-Schepman, A. M. J., van der Veer, J. L., den Hartog, J. H. J., Lohman, P. H. M., and Reedijk, J. (1985) *Biochemistry 24*, 707−713.
4. Reed, E., Yuspa, S. H., Zwelling, L. A., Ozols, R. F., and Poirer, M. C. (1986) *J. Clin. Invest. 77*, 545−550.
5. Pil, P., and Lippard, S. J. (1997) in *Encyclopedia of Cancer* (Bertino, J. R., Ed.) Vol. 1, pp 392−410, Academic Press, San Diego, CA.
6. Chow, C. S., Barnes, C. M., and Lippard, S. J. (1995) *Biochemistry 34*, 2956−2964.
7. Farid, R. S., Bianchi, M. E., Falciola, L., Engelsberg, B. N., and Billings, P. C. (1996) *Toxic. Appl. Pharmac. 141*, 532−539.
8. Locker, D., Decoville, M., Maurizot, J. C., Bianchi, M. E., and Leng, M. (1995) *J. Mol. Biol. 246*, 243−247.
9. Zamble, D. B., Mu, D., Reardon, J. T., Sancar, A., and Lippard, S. J. (1996) *Biochemistry 35*, 10004−10013.
10. Whitehead, J. P., and Lippard, S. J. (1996) in *Metal Ions in Biological Systems* (Sigel, A., & Sigel, H., Eds.) pp 687−726, Marcel Dekker, Inc., New York.
11. Landsman, D., and Bustin, M. (1993) *Bioessays 15*, 539−546.
12. Grosschedl, R., Giese, K., and Pagel, J. (1994) *Trends Gen. 10*, 94−99.
13. Bianchi, M. E., Beltrame, M., and Paonessa, G. (1989) *Science 243*, 1056−1059.
14. Read, C. M., Cary, P. D., Crane-Robinson, C., Driscoll, P. C., and Norman, D. G. (1993) *Nucleic Acids Res. 21*, 3427−3436.
15. Read, C. M., Cary, P. D., Crane-Robinson, C., Driscoll, P. C., Carrillo, M. O. M., and Norman, D. G. (1995) *Nucleic Acids Mol. Biol. 9*, 222−250.
16. Chow, C. S., Whitehead, J. P., and Lippard, S. J. (1994) *Biochemistry 33*, 15124−15130.
17. Kane, S. A., and Lippard, S. J. (1996) *Biochemistry 35*, 2180−2188.
18. Berners-Price, S. J., Corazza, A., Guo, Z., Barnham, K. J., Sadler, P. J., Ohyama, Y., Leng, M., and Locker, D. (1997) *Eur. J. Biochem. 243*, 782−791.
19. Hardman, C. H., Broadhurst, R. W., Raine, A. R. C., Grasser, K. D., Thomas, J. O., and Laue, E. D. (1995) *Biochemistry 34*, 16596−16607.
20. Jones, D. N. M., Searles, M. A., Shaw, G. L., Churchill, M. E. A., Ner, S. S., Keeler, J., Travers, A. A., and Neuhaus, D. (1994) *Structure 2*, 609−627.
21. Weir, H. M., Kraulis, P. J., Hill, C. S., Raine, A. R. C., Laue, E. D., and Thomas, J. O. (1993) *EMBO J. 12*, 1311−1319.
22. Broadhurst, R. W., Hardman, C. H., Thomas, J. O., and Laue, E. D. (1995) *Biochemistry 34*, 16608−16617.
23. Werner, M. H., Huth, J. R., Gronenborn, A. M., and Clore, G. M. (1995) *Cell 81*, 705−714.
24. Love, J. J., Li, X., Case, D. A., Geise, K., Grosschedl, R., and Wright, P. E. (1995) *Nature 376*, 791−795.
25. Sherman, S. E., Gibson, D., Wang, A. H.-J., and Lippard, S. J. (1985) *Science 230*, 412−417.
26. Bellon, S. F., and Lippard, S. J. (1990) *Biophys. Chem. 35*, 179−188.
27. Bellon, S. F., Coleman, J. H., and Lippard, S. J. (1991) *Biochemistry 30*, 8026−8035.
28. Takahara, P. M., Frederick, C. A., and Lippard, S. J. (1996) *J. Am. Chem. Soc. 118*, 12309−12321.
29. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
30. Salazar, M., Champoux, J. J., and Reid, B. R. (1993) *Biochemistry 32*, 739−744.
31. Salazar, M., Fedoroff, O. Y., Miller, J. M., Ribeiro, S., and Reid, B. R. (1993) *Biochemistry 32*, 4207−4215.
32. Salazar, M., Fedoroff, O. Y., Zhu, L., and Reid, B. R. (1994) *J. Mol. Biol. 241*, 440−455.
33. Horton, N. C., and Finzel, B. C. (1996) *J. Mol. Biol. 264*, 521−533.
34. Fedoroff, O. Y., Salazar, M., and Reid, B. R. (1993) *J. Mol. Biol. 233*, 509−523.
35. Treiber, D. K., Zhai, X., Jantzen, H.-M., and Essigmann, J. M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5672−5676.
36. McA'Nulty, M. M., Whitehead, J. P., and Lippard, S. J. (1996) *Biochemistry 35*, 6089−6099.
37. Ferrari, S., Ronfani, L., Calogero, S., and Bianchi, M. E. (1994) *J. Biol. Chem. 269*, 28803−28808.
38. Falciola, L., Murchie, A. I. H., Lilley, D. M. J., and Bianchi, M. E. (1994) *Nucleic Acids Res. 22*, 285−292.
39. Genetics Computer Group. (1991) *Program Manual for the GCG Package*, Version 7.0, Madison, Wisconsin.
40. Barton, G. (1994) *Alignment of Multiple Protein Sequences*, Version D2.0, Oxford.
41. Gait, M. J. (1984) *Oligonucleotide Synthesis: A Practical Approach*, IRL Press Limited, Oxford.
42. Cruachem, Inc. (1994) in *Technical Bulletin No. 037R01*, Dulles, VA.
43. Eastman, A. (1986) *Biochemistry 25*, 3912−3915.
44. Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology* (Fasman, G. D., Ed.) pp 589−589, CRC Press, Cleveland.
45. Batey, R. T., and Williamson, J. R. (1996) *J. Mol. Biol. 261*, 536−549.
46. Long, K. S., and Crothers, D. M. (1995) *Biochemistry 34*, 8885−8895.
47. Lohman, T. M., and Mascotti, D. P. (1992) *Methods Enzymol. 212*, 400−424.
48. Lin, S.-Y., and Riggs, A. D. (1972) *J. Mol. Biol. 72*, 671−690.

49. Pil, P. M., and Lippard, S. J. (1992) *Science 256*, 234−237.
50. Teo, S.-H., Grasser, K. D., and Thomas, J. O. (1995) *Eur. J. Biochem. 230*, 943−950.
51. Visse, R., de Ruijter, M., Brouwer, J., Brandsma, J. A., and van de Putte, P. (1991) *J. Biol. Chem. 266*, 7609−7617.
52. Schwartz, A., and Leng, M. (1994) *J. Mol. Biol. 236*, 969−974.
53. Bianchi, M. E., Falciola, L., Ferrari, S., and Lilley, D. M. J. (1992) *EMBO J. 11*, 1055−1063.
54. Lnenicek-Allen, M., Read, C. M., and Crane-Robinson, C. (1996) *Nucleic Acids Res. 24*, 1047−1051.
55. Payet, D., and Travers, A. (1997) *J. Mol. Biol. 266*, 66−75.
56. Benight, A. S., Gallo, F. J., Paner, T. M., Bishop, K. D., Faldasz, B. D., and Lane, M. J. (1995) *Adv. Biophys. Chem. 5*, 1−55.
57. Mello, J. A., Acharya, S., Fishel, R., and Essigmann, J. M. (1996) *Chem. Biol. 3*, 579−589.
58. Jones, C. J., and Wood, R. D. (1993) *Biochemistry 32*, 12096−12104.

BI9709452