Intercalating Residues Determine the Mode of HMG1 Domains A and B Binding to Cisplatin-Modified DNA†

Qing He, Uta-Maria Ohndorf, and Stephen J. Lippard*

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

Received July 21, 2000; Revised Manuscript Received September 12, 2000

ABSTRACT: Cisplatin exerts its anticancer activity by forming covalent adducts with DNA. High-mobility group (HMG)-domain proteins recognize the major 1,2-intrastrand cisplatin−DNA cross-links and can mediate cisplatin cytotoxicity. The crystal structure of HMG1 domain A bound to cisplatin-modified DNA, further analyzed here, reveals intercalation of a key Phe37 residue. Other published structures of HMG domains bound to DNA, including NHP6A and HMG-D, similarly indicate amino acid side chains intercalating into linear DNA to form a bend. To delineate the importance of such side chain intercalations and further to explore the binding modes of different HMG domains toward prebent DNA structures, site-directed mutagenesis was used to generate HMG1 domain A and domain B mutants. The affinities of these mutant proteins for cisplatin-modified DNA were determined in gel electrophoresis mobility shift assays. The results indicate that intercalating residues at positions 16 or 37 can both contribute to the binding affinity. The data further reveal that the length of the loop between helices I and II is not critical for binding affinity. Footprinting analyses indicate that the position of the intercalating residue dictates the binding mode of the domain toward platinated DNA. Both congruent and offset positioning of the HMG domain with respect to the locus of the cisplatin-induced bend in the DNA were encountered. Packing interactions in the crystal structure suggest how full-length HMG1 might bind to DNA by contacting more than one duplex simultaneously. Taken together, these results demonstrate that cisplatin modification of DNA provides an energetically favorable, prebent target for HMG domains, which bind to these targets through one or more side chain and favorable hydrophobic surface interactions.

cis-Diamminedichloroplatinum(II) (cisplatin)[1] is a widely used anticancer drug that forms covalent adducts on DNA (*1*−*3*). The major cisplatin−DNA adducts, 1,2-intrastrand cross-links, form distinctive bends and distortions in the DNA, which are recognized by a variety of cellular proteins including high-mobility group (HMG) domain proteins (*4*, *5*). HMG-domain proteins bind specifically to the major cisplatin−DNA adducts and form stable platinum−DNA−protein ternary complexes in vitro. Considerable evidence implicates the involvement of such complexes in cisplatin cytotoxicity (*6*−*10*). To delineate further the cisplatin mechanism of action and design better cancer therapies, it is important to understand thoroughly the interactions between HMG-domain proteins and cisplatin-modified DNA.

HMG-domain proteins are usually associated with chromatin (*11*, *12*) and can be categorized as sequence-specific or sequence-neutral (*13*). Sequence-specific HMG domains include transcription factors such as LEF-1 and SRY that are cell-type specific. Sequence-neutral HMG-domain proteins, such as HMG1, HMG2, HMG-D, and NHP6A, are present in many cell types and recognize distorted DNA independent of sequence (*13*). Both subfamilies share a minor groove DNA-binding motif and bind to distorted structures such as cisplatin-modified DNA (*14*), supercoiled DNA (*15*), and four-way junctions (*16*).

HMG1 is an architectural protein with two tandem HMG domains, A and B. Each domain alone specifically recognizes the major cisplatin−DNA adducts. The X-ray crystal structure of a 16-base pair double-stranded deoxyoligonucleotide containing a single *cis*-[Pt(NH_3)$_2${d(GpG)-N7(G_8)-N7(G_9)}] intrastrand cross-link in complex with HMG1 domain A has been described (*17*). A striking feature is the intercalation of the Phe 37 side chain into a hydrophobic notch in the minor groove formed by the drug−DNA adduct. The opening of the two platinum cross-linked GC base pairs involving guanine rings G_8 and G_9 allows for intermolecular stacking interactions between the Phe 37 phenyl ring and the G_9 purine heterocycle, interlocking the two components like pieces in a molecular jigsaw puzzle. This X-ray structure determination offered the first insight that cisplatin−DNA cross-links might provide a preformed intercalation site for minor groove DNA-binding proteins, a feature that is likely to contribute

† This work was supported by Grant CA34992 (S.J.L.) from the National Cancer Institute. Q.H. is a Howard Hughes Medical Institute Predoctoral Fellow, and U.-M.O. is the recipient of a graduate fellowship from the Fonds der Chemischen Industrie, Germany.

* Corresponding Author: Stephen J. Lippard, Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, Phone: (617) 253−1892. Fax: (617) 258−8150. E-mail: lippard@lippard.mit.edu.

[1] Abbreviations: cisplatin, *cis*-diamminedichloroplatinum (II); HMG, high-mobility group; domA, HMG1 domain A; domB, HMG1 domain B; EDTA, ethylenediaminetetraacetic acid; *E. coli*, *Escherichia coli*; BSA, bovine serum albumin; HEPES, *N*-[2-hydroxyethyl]piperazine-*N*′-[2-ethanesulfonic acid]; Tris, tris-[hydroxymethyl]aminomethane; 4WJ, 4-way junction DNA; HPLC, high performance liquid chromatography; CD, circular dichroism; rmsd, root-mean-square deviation; PBS, phosphate buffered saline.

10.1021/bi001700j CCC: $19.00
Published on Web 11/01/2000

to the genotoxic effects of the drug. Recent work from our laboratory demonstrated that TATA-binding protein interacts with cisplatin-modified TATA sequences in a similar manner (*18*).

In the present article, we describe a detailed structural analysis, combined with site-directed mutagenesis, that reveals the contributions of specific protein−DNA contacts to the composition and stability of the complexes. The binding orientations and affinities of HMG1 domain A and domain B mutants for cisplatin-modified DNA were determined. Gel-mobility shift assays and hydroxyl radical footprint analysis support a model in which sequence-neutral HMG-domain proteins utilize a two-pronged intercalation mode to bind to prebent DNA. We demonstrate that the intercalating residues are key to how the protein is positioned on the DNA relative to the platinum adduct. The binding mode can be dramatically changed with a single side chain mutation. The results are compared with those provided by the crystal and NMR structures of other HMG domain−DNA complexes (*19*−*22*).

MATERIALS AND METHODS

Site-Directed Mutagenesis. Mutagenesis of HMG1 domains A and B was carried out according to the Stratagene QuikChange site-directed mutagenesis kit protocol. Successful mutations were confirmed by DNA sequencing at the MIT Biopolymers Lab. HMG1 domain A and B mutants were expressed in *Escherichia coli* BL21(DE3) and purified as described (*23*), with the addition of an FPLC size-exclusion purification column (high-load Superdex 75, Pharmacia, 1 mL/min, 11.8 mM PBS, pH 7.4). The ability of protein mutants to fold correctly was confirmed by circular dichroism spectroscopy.

Circular Dichroism. CD spectra were recorded on an AVIV 62DS spectrophotometer. Proteins were diluted to 20 μM in 0.05× PBS buffer. The measurements were made at 4 °C with a 0.7-nm bandwidth and a 0.5-nm step size. Each spectrum is the average of four recordings. The raw data were smoothed by using AVIV software. Ellipticities are based on molar peptide rather than amino acid residue concentrations.

Gel Mobility Shift Assays. The oligonucleotide d(GGTTGGTCCAGAGAGG) was 5′ end-labeled and annealed to its complementary strand d(CCTCTCTG*G*ACCTTCC), where the asterisks represent a *cis*-diammineplatinum(II) cross-link. The duplex (0.4 nM) was incubated with increasing concentrations of proteins in 20-μL sample volumes containing 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. Samples were incubated on ice for 30 min, and then made 7% in sucrose and 0.017% in xylene cyanol prior to loading on prerun, precooled (4 °C) 10% native polyacrylamide gels (29:1 acrylamide/bis). Gels were electrophoresed for 3 h and vacuum-dried onto Whatman 3 MM chromatography paper. Gels were visualized by using a BioRad phosphorimager, and the bands were quantitated with the BioRad Multi-Analysist software.

Data Analysis. Apparent dissociation constants, K_d, were estimated from nonlinear least-squares fits of binding data to the Langmuir isotherm, eq 1 (*24*), where θ is the fraction of bound oligonucleotide probe and P is the total protein concentration. Each K_d is the average of at least two replicates.

$$\theta = P/(P + K_d) \quad (1)$$

Hydroxyl Radical Footprinting Assays. Labeled DNA samples with or without protein (20-fold excess protein over DNA) were incubated at 4 °C for 30 min in the same buffer described for the gel mobility shift assays. To a 20-μL aliquot was added a fresh mixture of sodium ascorbate (10 mM), $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ (2.5 mM), sodium EDTA (5 mM), and H_2O_2 (0.3%) to initiate the reaction. After 5 min at room temperature, 10 μL of thiourea (1 M) was added to stop the reaction. DNA was extracted with phenol/chloroform and then precipitated with ethanol. The DNA samples were electrophoresed on 20% denaturing polyacrylamide gels. Gels were dried and quantitated on a BioRad phosphorimager. The total counts in each lane were used to normalize the bands. The peak heights corresponding to each band were compared and, when the difference between the control and protein-added lanes was more than 10%, then the corresponding base was scored as being protected. Each area of protection was reproduced in at least two experiments.

Crystal Structure Analysis. DNA groove widths were calculated with the program Curves (*25*, *26*). The base and base-pair parameters were determined by using the program FREEHELIX (*27*). Solvent-accessible surfaces were calculated with customized XPLOR scripts (*28*). Calculations of the buried surface at the intercalation sites were performed with CNSsolve 1.0 (*29*). Structural comparison of the HMG domains was carried out through superposition of the Cα atoms 7−77 of HMG1 domain A (pdb accession number 1ckt) with the same Cα atoms in the NMR structure of HMG1 domain A (pdb accession number 1aab). For structural comparison of the DNA, the following phosphate atoms were used: HMG1 domain A complex (pdb accession number 1ckt): 106−113 and 205−213, or 109−116 and 202−210; HMG-D (pdb accession number 1qrv): 2−10 and 12−20; LEF-1 (pdb accession number 2LEF): phosphate atoms 5−12 of chain B and 5−13 of chain C. Insight97 was used for modeling the interactions between HMG mutants and DNA.

RESULTS

Protein−DNA Interactions in the Crystal Structure. The X-ray structure of HMG1 domain A complexed with cisplatin-modified DNA (*17*) allows comparison of the conformations of an HMG domain in the free and bound states. The structures of bound HMG1 domain A analyzed here and that of the free domain in solution (*30*, *31*) superimpose with an rmsd of 2.1 Å for all 70 Cα atoms. Conformational changes occur mainly in the loop region between helices I and II and at the N-terminus of helix II (Figures 1 and 2). To intercalate into the DNA, residue Phe37 rearranges from a position facing the hydrophobic core to one that is more exposed. In addition, the N-terminal strand and helix III, which form the longer arm of the L, are slightly displaced with respect to their positions in the NMR structure so as to optimize interactions with the sugar−phosphate backbone of the DNA. The tertiary structure of the protein, which adopts the typical L-shaped fold, undergoes a slight increase in α-helical content in the protein−DNA complex.

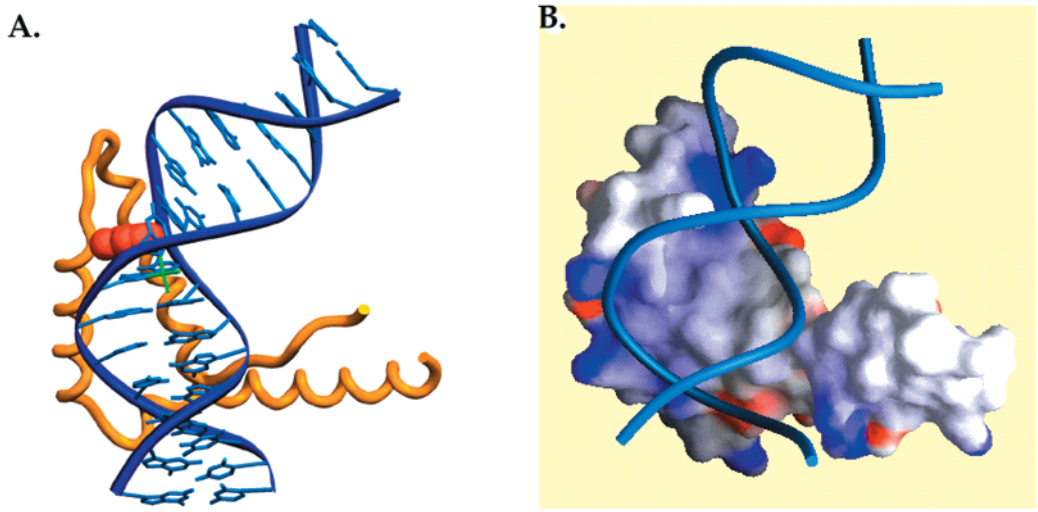

FIGURE 1: (A) Ribbon diagram of the protein and DNA backbone with an intercalating phenylalanine residue depicted as van der Waals spheres. (B) Molecular surface of the HMG domain color coded by electrostatic potential. Positive charge is depicted in blue, negative charge is shown in red, and uncharged areas are gray.

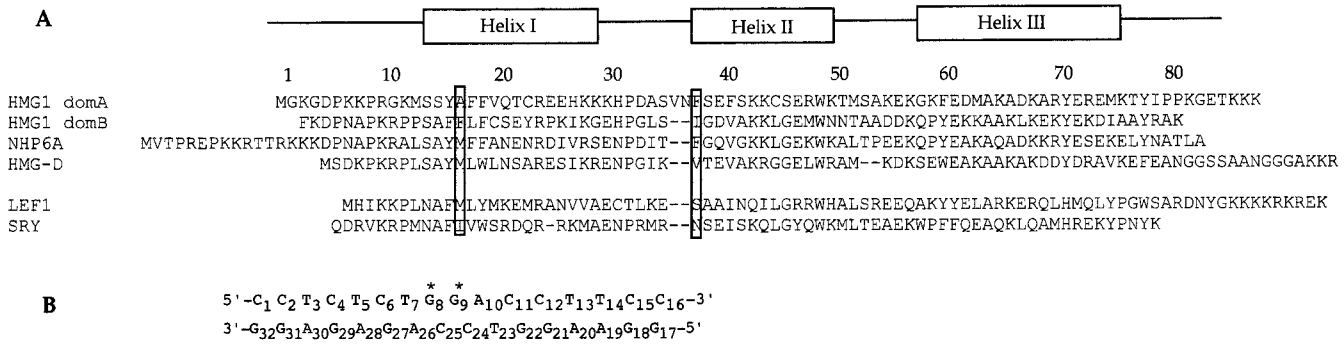

FIGURE 2: Sequence of HMG domain proteins and cisplatin-modified DNA. (A) Sequence alignment of HMG domain proteins. (B) Sequence of DNA used in this study; asterisks denote nucleotides cross-linked by cisplatin.

Helices I and II house two additional residues, with helix I spanning from 14 to 29 as compared to residues 14−27 in the NMR structure and helix II extending from residue 37−49; thus, the intercalating amino acid Phe37 becomes the first residue of the α-helix in the complex. The large hydrophobic core is maintained by Tyr15, Val19, His26, Trp48, and Phe59 and by the quartet Phe18, Cys22, Phe40, and Cys44. In contrast to the NMR structure, which was carried out with a C22S mutant to prevent protein oxidation during crystallization, this X-ray structure was determined with the wild-type protein. The high concentration of reducing agent added to the crystal setups most likely prevented cysteine oxidation. Our structure confirms that the hydrophobic core of the HMG domain remains largely unchanged in the C22S mutant used for the NMR structure. The distance between Cys18 and Cys44 in this wild-type structure was determined to be 4.3 Å, suggesting that in principle a disulfide bond might be formed. In contrast to Cys44, which is buried in the protein core, Cys22 points toward the back of helix II and is potentially solvent-accessible.

Analysis of the molecular interface between HMG domain A and cisplatin-modified DNA shows the protein to be well adapted for binding in the minor groove of this perturbed DNA structure. The binding site extends over five base pairs toward the 3′ side of the duplex with respect to the platinum adduct-containing strand, excluding a solvent accessible surface of 910 Å^2. Helices I and II, which together provide the majority of the protein binding surface, are spaced such that only a widened minor groove would support a stable interaction. There is a high degree of complementarity in shape, electrostatic, and chemical surface properties. This aspect is particularly noteworthy, considering that cisplatin-modified DNA is not a natural substrate for the HMG domain. A large hydrophobic core is enclosed at the floor of the minor groove with additional interactions of polar and charged residues with the DNA phosphate backbone (Figure 1). The N-termini of both helices are situated in the minor groove with the helix axes extending over the sugar−phosphate backbone of the two complementary DNA strands. The hydrophobic side chains of Tyr15, Ala16, Val19, Val35, and Phe40 function as a hydrophobic wedge by pressing against the furanose rings of C_{12}, G_{22}, T_{23}, C_{25}, and C_{24}, respectively, thereby stabilizing the widened minor groove. The polar side chains of Gln20, Arg23, Lys42, and Ser45 form hydrogen bonds and salt bridges to backbone phosphate oxygen atoms of C_{24}, C_{25}, A_{10}, and C_{12}, respectively. The residue Trp48 functions to support both protein and complex stability. It serves as the centerpiece of the hydrophobic core of the HMG domain and also forms a hydrogen bond through the Nϵ1 ring atom to the backbone phosphate oxygen of C_{12}.

The Platinated DNA Duplex. Analysis of the DNA conformation in this X-ray structure and comparison with the structures of unbound cisplatin-modified DNA allows identification of the recognition elements necessary for formation of sequence-neutral HMG domain−DNA com-

plexes. The general structural features of the DNA duplex in the complex and in unbound cisplatin-modified deoxyoligonucleotides are quite similar. The key element shared by the free and protein-complexed platinated DNA is the directionality of the DNA bend toward the major groove with a conserved DNA topology at the platination site and a widened, shallow minor groove. Despite these gross similarities, protein binding induces large variations in the magnitude of the DNA parameters underlying these distortions. The overall ~61° DNA bend in the protein−DNA complex, a result of a highly localized kink, is surprisingly symmetric with respect to its center. Superposition of the 3′ end onto the 5′ end of the DNA duplex yields an rmsd of 1.07 Å for 30 phosphate atoms. The site of the cisplatin−DNA cross-link forms the kink locus, with complete destacking of the cross-linked base pairs. An interbase dihedral angle of 75° occurs between the G_8 and G_9 purine residues. In the absence of protein binding, this opening would result in an energetically unfavorable exposure of a largely hydrophobic, solvent-accessible surface by 300 Å^2. Therefore, only a comparatively low dihedral angle of 30° occurs in the structure of the cisplatin-modified dodecamer, and the hydrophobic minor groove surface is stablilized by an A-DNA end-to-groove packing interaction in the crystal lattice (*32*). The helical twist angle of 9° indicates almost complete DNA unwinding in the complex. In comparison, the uncomplexed cisplatin−DNA structures exhibit a helical twist of 25−28° (*1*, *32*, *33*); the value for B-DNA is 36°. The buckle for both G−C base pairs involved in platinum coordination is 31° and −11°. Significant propeller twisting of ~20° occurs only at the G_9−C_{24} base pair. These distortions are accommodated without disruption of Watson−Crick hydrogen bonding at the modified bases. Bending of the double helix in combination with protein binding results in a minor groove even wider and shallower than observed in the complex of cisplatin-modified DNA alone, the width of the minor groove in the complex being 12 Å at the central four base-pair steps (*32*). Surprisingly, DNA perturbations remain localized to the central six base pairs of the duplex. The base pair and base pair step parameters for the last five base pairs on each end of the duplex strongly resemble those of B-DNA. Rmsd values for superposition of the bound duplex with canonical B-DNA are 1.1 Å at the 5′ end and 0.9 Å at the 3′ end. In summary, protein complexation of cisplatin-modified DNA stabilizes the bent DNA structure in a conformation more favorable than that of the free platinated duplex.

The Domain A and Domain B Mutants. Figure 2, panel A, presents the sequence alignment of HMG1 domains A and B together with several other HMG domains. The numbering scheme is based on that of HMG1 domain A unless specified otherwise. The sequence of the site-specifically platinated DNA used in this study is shown in Figure 2, panel B.

A variety of HMG1 domain A and domain B mutant proteins were expressed and purified. Circular dichroism spectroscopy showed that these mutants fold similarly to the native protein, with high α-helical content (Figure 3). The affinities of the HMG1 domain A and B mutants for site-specifically cisplatin-modified DNA were investigated by gel mobility shift assays. The K_d values are listed in Table 1.

Contribution of Helix I−II Spacing and Loop Length to Platinated DNA Protein Binding Affinity. Sequence alignment

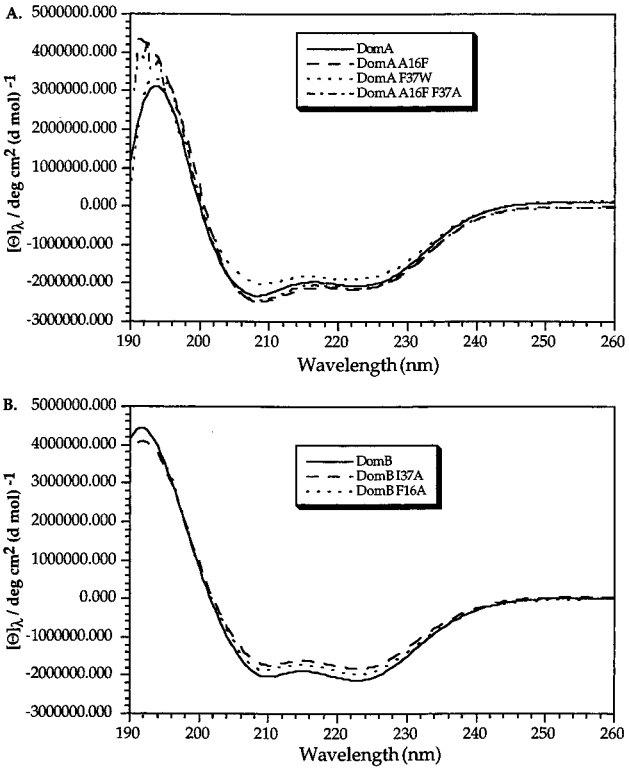

FIGURE 3: Circular dichroism spectra of HMG1 domain A and domain B mutants. (A) CD spectra of domA, domA F37W, domA A16F F37A, and domA A16F. (B) CD spectra of domB, domB F16A, and domB I37F.

Table 1: Affinities of HMG1 Domain A and Domain B Mutants toward Cisplatin-Modified DNA

protein	K_d (nM)	K_d relative[a]
DomA	1.5 ± 0.5	1
DomA ΔVN	5.0 ± 0.9	3.3
DomA F37W	9.9 ± 2.1	6.6
DomA A16F	17 ± 5.7	11
DomA A16F F37A	167 ± 32	111
DomA F37A	>1000	>667
DomA S41A	5.8 ± 0.7	3.9
DomB	39 ± 9.9	1
DomB iVN	39 ± 13	1
DomB I37F	19.5 ± 2.6	0.5
DomB F16A	119 ± 39	3
DomB I37A	85 ± 17	2.2
DomB F16A I37A	>1000	>25

[a] K_d/ K_d (DomA) or K_d/ K_d (DomB).

(Figure 2, panel A) shows that the HMG1 domain A has two amino acids, V35 and N36, inserted in the loop connecting helices I and II that are not present in HMG1 domain B and most other HMG-domain proteins (*34*). To determine whether these two amino acids are critical for protein binding, a deletion mutant of HMG1 domain A lacking V35 and N36, domA ΔVN, was designed. In a complementary experiment, these two amino acids were inserted in HMG1 domain B, domB iVN. DomA ΔVN has a 3.3-fold lower affinity for platinated DNA as compared to domA (Table 1). The affinity of domB iVN, however, is comparable to that of the wild-type domain B. These results suggest that the length of the intervening loop between helix I and II is not a determining factor of the platinated DNA binding affinity in these HMG domains.

Contribution of Protein-Base Hydrogen Bonding to Platinated DNA Binding Affinity. The crystal structure reveals only a single direct protein-base contact. Ser41 forms a hydrogen bond with the N3 atom of A_{10}, the base directly adjacent to the cisplatin−DNA cross-link. It had been suggested prior to the X-ray structure that the lower affinity of HMG1 domain B for cisplatin-modified DNA could be attributed to the lack of such an interaction (*35*). Domain B has an alanine residue at the equivalent position. Indeed, the S41A mutant of HMG1 domain A exhibits a 5-fold decrease in affinity for the cisplatin-modified substrate (Table 1).

Contribution of Intermolecular Stacking at Position 37 to Platinated DNA Binding Affinity. The crystal structure shows that Phe37 intercalates at the site of the cisplatin cross-link. It interacts through a $\pi-\pi$ stacking interaction with the cisplatin-modified G_9 purine ring and in an edge-to-face ring manner with the G_8 base. This intercalation stabilizes the unprecedented, large interbase dihedral angle of 75° at the site of cisplatin-modification by compensating for the missing base stacking interaction of the 5′ side of the G_9-C_{24} base pair. Furthermore, it reduces the area of solvent accessibility of the opened purine and pyrimidine ring faces of those base pairs by ∼100 Å^2.

The favorable stacking interactions between the intercalating residue Phe37 and the cisplatin-modified bases determine the orientation of the bound protein with respect to the DNA. Although the DNA phosphate backbone of the cisplatin-modified duplex is symmetric with respect to the bend, the HMG1 domain A binding site extends exclusively toward the 3′-side with respect to the cisplatin-modified strand. Comparison of the two possible binding orientations, 3′ versus 5′, reveals that a significant $\pi-\pi$ interaction between the aromatic systems of Phe37 and the G8 and G9 purine rings can occur only in the observed protein binding orientation.

Since not all sequence-neutral HMG-domain proteins utilize an aromatic residue for intercalation at position 37, mutagenesis experiments were designed to assess the extent to which favorable $\pi-\pi$ stacking interactions at this position contribute to binding affinity. Substitution of Ile37 in domain B for Phe increased the protein affinity toward cisplatin-modified DNA by 2-fold (Table 1), consistent with the conclusion that the intercalating Phe37 residue contributes to the higher affinity of HMG1 domain A for platinated DNA as compared to domain B. Trp has a larger aromatic ring and may form more favorable $\pi-\pi$ stacking interactions than Phe. The domA F37W mutant, however, displays diminished binding affinity (Table 1), presumably because of steric crowding between the exocyclic oxo atom of the C_8 base and the indole ring of Trp (data not shown).

An Intercalating Residue at Position 16 Influences DNA Affinity. Since the N-termini of both helices I and II lie within the minor groove, a potential intercalating residue could be positioned within either helix. The sequence-specific HMG domains LEF-1 and SRY intercalate into DNA by inserting Met or Ile, respectively, from helix I. These residues occupy the positions equivalent to that of 16 in domain A (*36*). Moreover, in other sequence-neutral HMG domain−DNA complexes (*19*−*22*) there are multiple intercalators situated at the N-termini of helices I and II. At position 16, HMG1 domain A has an alanine that is ill-suited to intercalate; instead, it forms a hydrophobic contact with the ribose ring of T_{23}. Loss of the Phe37 intercalation, as in the DomA F37A mutant, drastically diminishes DNA binding affinity (*17*). The DomA A16F F37A double mutant having a single potential intercalating residue at position 16, partially restores binding affinity (Table 1). The domA A16F mutant, which has two potential intercalators, however, displays decreased binding affinity possibly due to steric crowding of Phe16 with the adjacent Phe residues.

HMG1 domain B has two potential intercalating residues, Phe16 and Ile37. Mutation of either one to alanine decreases the affinity for cisplatin-modified DNA by about 2−2.5-fold (Table 1). The domain B double mutant, F16A I37A, has severely abrogated platinated-DNA affinity, similar to that of domain A F37A. Despite the presence of two intercalating residues, HMG1 domain B binds less well to cisplatin-modified DNA than domain A. These results confirm that favorable aromatic interactions provided by Phe37 and the overall complementarity of the protein and DNA binding surfaces are both important determinants of binding affinity.

Position of the Intercalator Determines the Binding Orientation. With a choice of two different positions for an intercalator, two different binding modes of the HMG domain with respect to the platination site are conceivable. These modes were probed by hydroxyl radical footprinting experiments. Figure 4, panel A shows the detailed analysis of footprints of various mutants in which the bottom strand of the probe was labeled. Corresponding footprints for labeled, platinated-top strand gave similar results (data not shown). The regions of protection from the hydroxyl radical cleavage reactions are shown schematically in Figure 4, panel B. Two different protection patterns are observed, either a symmetric footprint, in which the protein-binding site extends to both sides of the platination site, or an asymmetric footprint, in which the binding site extends toward the 3′ side of the platinum adduct. In proteins having a single intercalator at position 37, as in wild-type HMG1 domain A and domB F16A, an asymmetric footprint covering 5−6 bases is observed. On the other hand, the presence of a single intercalator at position 16, as occurs for domA A16F F37A and domB I37A, converts the footprint to a symmetric pattern extending over 7−8 bases. The footprint of domain B, which has two potential intercalators, covers about 8 base pairs and is centered around the platinum-adduct (Figure 4). By contrast, the footprint of the domA A16F mutant, which has two potential intercalators, is asymmetric and similar to that of domain A.

Comparison of HMG Domain Intercalation Sites. Minor groove binding of sequence-specific and sequence-neutral HMG-domain proteins is governed by intercalation of one or more hydrophobic residues between base pairs of the DNA duplex. The sequence-specific HMG-domain proteins LEF-1 and SRY employ an aliphatic Met or Ile residue at position 16, resulting in a DNA bend with its center two base pairs removed from the major kink site in the HMG1 domain A complex with cisplatin-modified DNA (*17*). Despite the disparate positioning of the protein with respect to the bend, comparison of the DNA structures in the proximity of these intercalation sites indicates a close conformational similarity with an rmsd value of 2.4 Å. The minor groove width in both complexes reaches a value of ∼11.5 Å.

By contrast, the sequence-neutral HMG-D protein induces DNA deformations through intercalation of multiple side

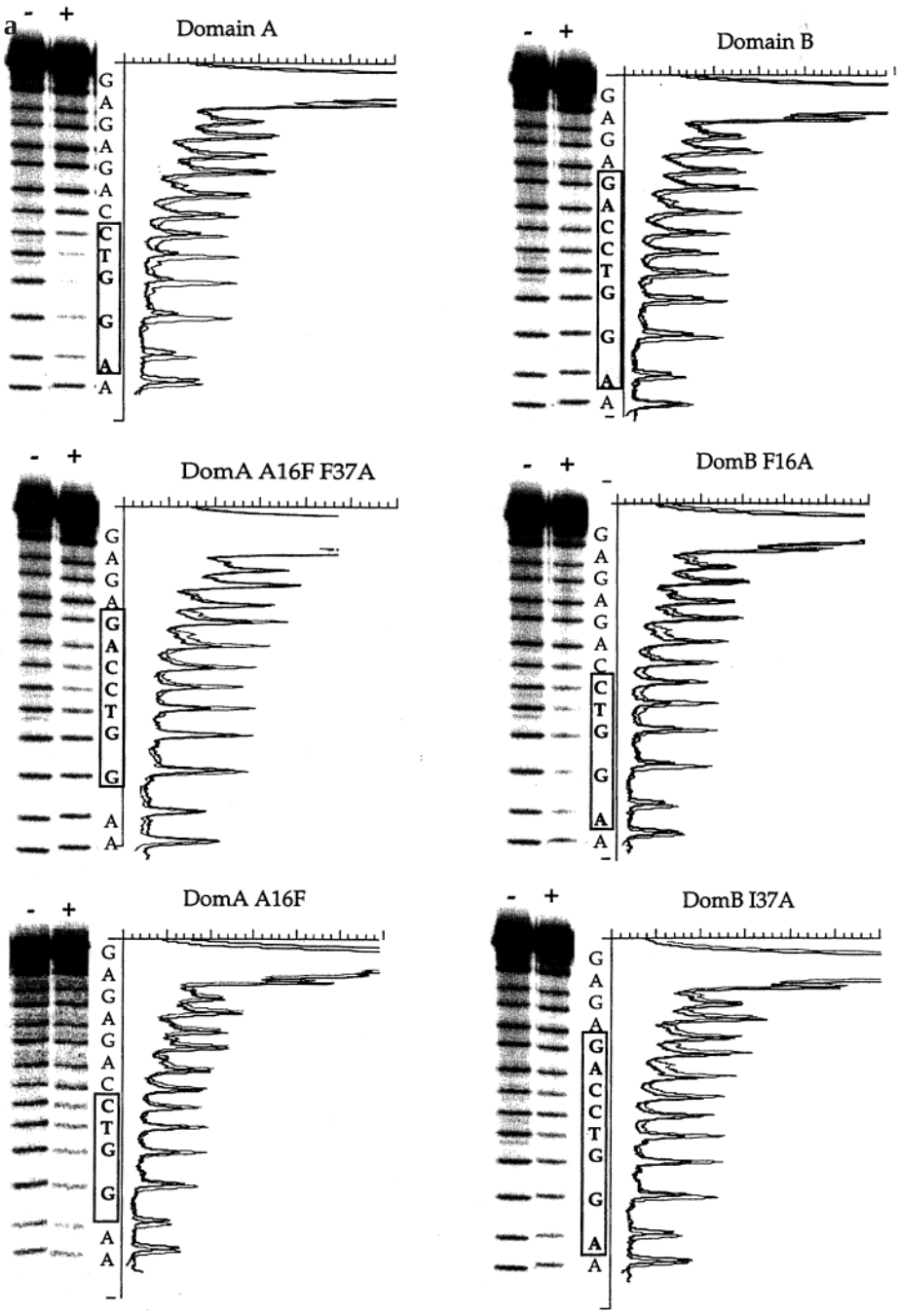

FIGURE 4: Hydroxyl radical footprint analysis. (A) Footprint analysis of the bottom strand DNA. Dotted lines are the traces for control samples without protein added. Solid lines are the traces for the footprint with protein added. Boxed letters indicate the protection areas. (B) Summary of footprint results. Bold letters indicate the protected areas and half circles above the GG bases indicate the site of platinum cross-linking.

chains (*22*). As in the sequence-specific LEF-1−DNA complex, a Met residue at position 16 is employed. An additional double intercalation of residues Val32 and Thr33 occurs at the site equivalent to position 37 in the sequence-neutral HMG1 domain A complex. Superposition of the latter sites reveals a dramatic difference in DNA structure manifest by an rmsd value of 4.5 Å. Despite the difference in protein positioning, the DNA structures at the intercalation sites of Met16 in HMG-D and Phe37 in HMG1 domain A more closely resemble each other with an rmsd value of 2.7 Å. From this similarity, it is predicted that HMG-D binding to cisplatin-modified DNA would occur mostly through Met16 intercalation similar to Phe37 in domain A and that the Val32/Thr33 residues are not as dominant as Met16. After the submission of this paper, an NMR structural study of HMG-D bound to disulfide cross-linked DNA appeared. The structure resembled that of an HMG domain bound to DNA that was not prebent (*37*). This result is probably because cisplatin-modified DNA has a highly localized bend whereas the disulfide-linked DNA does not (*37*).

DISCUSSION

Factors Affecting Protein Affinity for Cisplatin-Modified DNA. The affinity of HMG1 domain A for cisplatin-modified DNA is generally higher than that of domain B (*35*). Although the helix I−II loop of domain A is longer than that of domain B (Figure 2), the length of this loop is not critical for binding. Moreover, the specific base contacted by Ser41 of domain A contributes only modestly to the strength of the interaction. The intercalation mode of Phe37 observed in the crystal structure is important for DNA affinity (*17*), suggesting that if Ile37 in HMG1 domain B were replaced by a phenylalanine residue, the enhanced π-interactions would enhance the binding. Indeed, domB I37F has a higher affinity toward platinated DNA than wild type domain B.

The domA F37A mutant protein has extremely low binding affinity for the platinated DNA probe (*17*), owing to the loss of the Phe37 intercalator. The affinity is partially restored in the A16F F37A mutant, indicating the side chain at position 16 to be a potential intercalator. Instead of abolishing binding affinity, as occurs with domA F37A, mutating either Phe16 or Ile37 in domain B lowers the affinity by only 2−2.5-fold. When both residues are mutated, in domB F16A I37A, the affinity for the platinated probe is very low, as is the case for domA F37A. These results suggest that either Phe16 or Ile37 can intercalate in the domain B platinated-DNA complex, in support of previous reports on the contribution of intercalating residues to the affinity of HMG-domain proteins for DNA. A model of HMG-D binding to prebent bulged DNA indicates intercalations by both Met13 and Val32, equivalent to positions 16 and 37 in domain A (*21*). The HMG-D V32A mutant, which is similar to HMG1 domB I37A, binds to bulged DNA with 5−6-fold lower affinity (*21*). The affinity of several NHP6A mutants for linear DNA has been reported, and deleting Met29 or Phe48, equivalent to position 16 or 37 in domain A, causes a 2- or 4-fold reduction in affinity, respectively (*19*).

A recent study of HMG1 domain B mutants revealed that domB F16A and domB I37A bind more tightly to four-way-junction (4WJ) DNA (*38*). The authors propose that these mutations increase the flexibility of the angle between helices I and II, enhancing protein−4WJ DNA interactions. This result seems contrary to our findings. The binding mode to 4WJ of HMG domains, however, may be very different from that of cisplatin-modified DNA, and intercalating residues may not contribute to the strength of the interaction.

The DNA Binding Mode. The mutagenesis and footprinting experiments described here confirm that HMG domains can bind in more than one manner to duplex DNA, as postulated previously (*35*, *39*−*41*). Both symmetric and asymmetric binding modes are possible, as determined by the intercalator position in the protein. Furthermore, our results reveal that it is possible to alter rationally the protein binding orientation by changing the position of the intercalator within the protein scaffold. An HMG domain can be engineered to cover a specific region on the DNA with respect to the cisplatin lesion. If the intercalating residue is situated at position 37, as is the case in domA and domB F16A, the hydroxyl radical cleavage pattern is asymmetric with respect to the platination site. Coverage occurs exclusively to the 3′ side of the lesion. The footprints of the mutants domA A16F F37A and domB I37A, in which the potential intercalator is positioned in helix I, are symmetrically centered at the site of the major bend locus (Figure 4). These results indicate that Phe16 redirects the protein−DNA interactions and functions as the intercalating residue.

The domA A16F mutant has two potential intercalating residues. The binding mode remains the same as that of wild-type domain A, however, consistent with Phe37 being the dominating residue. The favorable hydrogen bond between Ser41 and A_{10}, in conjunction with other side chain−DNA interactions observed in the crystal structure, could explain the preference of the asymmetric over the symmetric orientation. In contrast, domain B, which like domA A16F has two possible intercalating residues, binds in a symmetric manner. The footprint region is slightly larger than that of domain A and extends over eight bases. This observation suggests that domain B may bind in more than one orientation. The following scenarios are conceivable: either Phe16 or Ile37 intercalate alternatively into the hydrophobic notch, possibly in a dynamic manner; domain B can pivot around the intercalating residue; or both might occur. Figure 5 portrays a model of how domain B might bind cisplatin-modified DNA symmetrically with Phe16 as the intercalating residue as compared with the structure of domain A binding to cisplatin-modified DNA. The ability of a protein to bind DNA in several different orientations has been reported previously. The engrailed homeodomain binds to certain DNA sequences in two different complexes as demonstrated by hydroxyl radical footprinting, although the crystal structure captured only one binding mode (*42*).

A model based on NMR data of the sequence-neutral yeast HMG-domain protein NHP6A complexed with DNA similarly proposes double intercalation of residues Met29 and Phe48, equivalent to positions 16 and 37 in domain A, respectively (*19*). Two adjacent base pair steps are proposed to be the intercalation sites. Superposition of the protein backbone of NHP6A with our X-ray structure indicates that, if Met29 were positioned between the cisplatin cross-linked base pair step, then Phe48 intercalation would occur at the base pair step below the cisplatin cross-link. However, the

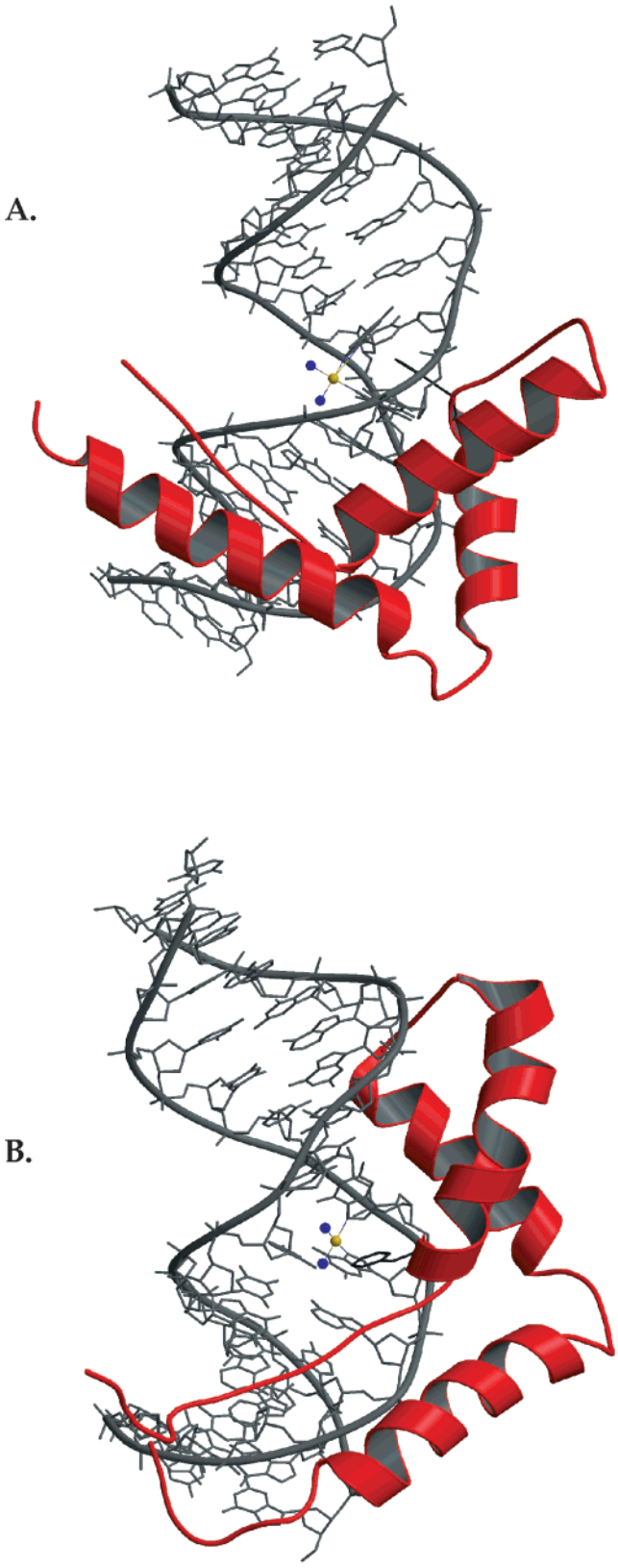

FIGURE 5: Comparison of two binding modes of HMG domains bind to cisplatin-modified DNA. **(**A) Domain A binding to cisplatin-modified DNA (*17*). (B) A model of domain B binding to Pt-DNA with Phe16 as the intercalating residue. The LEF-1−DNA complex binding orientation is used as the model, and platinated-DNA was superimposed with LEF-1 DNA (see Materials and Methods), where as domain B was superimposed with LEF-1.

aromatic ring of Phe48 would have to be oriented almost perpendicular to the base pair planes, leading to unfavorable stacking interactions. It is possible that intercalation into adjacent base pair steps is a feature specific for NHP6A binding to linear rather than prebent DNA substrates. The favorable π-stacking interactions at the intercalation site observed in our structure most likely reflect the presence of a preformed intercalation site.

The Recognition Mechanism. As compared to the major groove, the DNA minor groove offers few distinctive hydrogen bonding features. Recognition of specific sequences by DNA minor groove binding proteins is postulated to depend largely on the inherent flexibility of a sequence (*43*, *44*) and is often dominated by hydrophobic protein−DNA interactions (*45*, *46*). In addition to hydrophobic minor groove complementarity, sequence-specific HMG-domain proteins use an intercalating residue for binding site recognition. The protein−DNA complex structures of hSRY and LEF-1 reveal that residue 16 serves this function, whereas residue 37 is polar and makes a base-specific contact. Minor groove DNA-binding proteins often organize elements in chromatin structure by offering a DNA folding surface. Hydrophobic residues are employed as an intercalative wedge to pry open one or several base pair steps, thereby inducing a positive roll and bending the DNA away from the protein toward the major groove. In addition to the sequence-specific and sequence-neutral HMG-domain proteins, the TATA binding protein TBP (*47*, *48*), the hyperthermophile chromosomal proteins Sac7d (*49*) and Sso7d (*50*), the *E. coli* purine repressor protein PurR (*51*), and the prokaryotic architectural protein integration host factor IHF (*52*) belong to this group.

Many of these proteins also bind to DNA containing the cisplatin-1,2-intrastrand cross-link (*50*, *53*−*55*). It had been suggested previously that prebending of the DNA would be the recognition signal for protein binding (*43*, *44*). The structure of HMG1 domain A with cisplatin-modified DNA and the mutagenesis results presented here suggest, however, that preformation of a hydrophobic notch through cisplatin modification is at least of equal importance. Favorable intercalation interactions direct protein binding and orientation. The general mechanism of action of this drug is therefore likely to be based on its ability to provide an intercalation site for minor groove binding proteins at a much lower energetic cost, the energy penalty for base-pair destacking being paid upon *cis*-DDP modification. In further support of this hypothesis, the HMG-domain protein Rox1, which has polar residues at both positions 16 and 37, is unable to serve an intercalating function and is the only HMG-domain protein found to date that does not bind to the 1,2-intrastrand cisplatin−DNA cross-link (*8*).

Implications for Full-Length HMG1−DNA Binding. The mutagenesis results presented here show for the first time that the HMG domains in HMG1 possess two different DNA binding modes with respect to a defined intercalation site. It is striking that, in both wild-type domains, a Phe residue is known or presumed to interact with the hydrophobic notch created by the cisplatin−DNA adduct. As a result of the different positioning of these residues within the protein scaffold, the binding region of the two HMG domains differs by two base-pair steps with respect to the cisplatin-damaged site. Moreover, there is evidence that the two domains can be differentiated functionally. The isolated domain A of pig HMG2, which has 90% sequence identity to rat HMG1

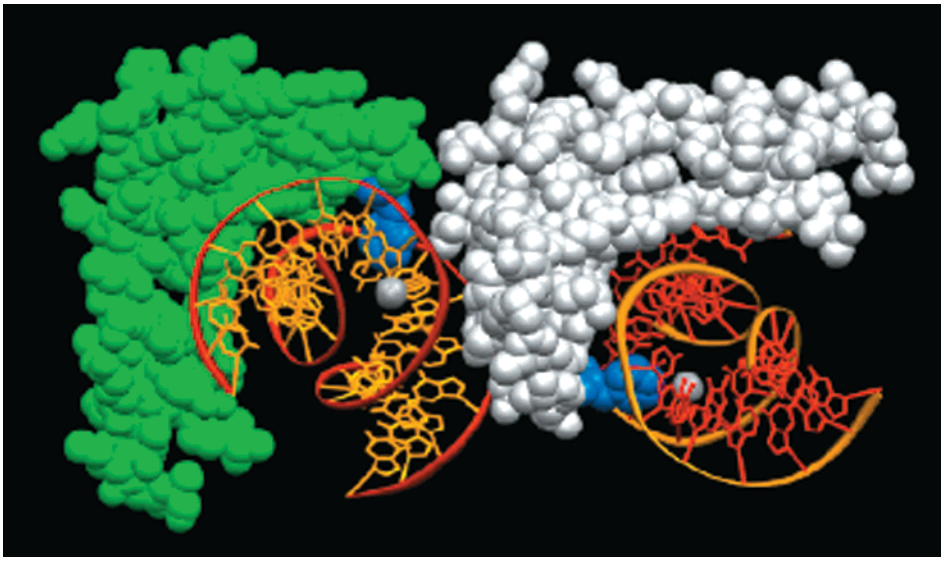

FIGURE 6: Crystal packing interactions of two adjacent complexes. Helix II of the HMG1 domain A contacts the minor groove of an adjacent duplex. The intercalating phenylalanine residue is shown in blue and the platinum atom as a gray sphere.

employed in our study, is more effective in DNA unwinding than domain B (*56*). Most sequence-neutral HMG-domain proteins comprise several HMG DNA-binding domains, with hUBF having as many as six. It is thus likely that fine tuning of HMG domain protein specificity during evolution has been achieved through sequential arrangement of several domains with different DNA recognition properties.

Members of the HMG1/2 subgroup can carry out a myriad of functions. In addition to stabilizing bent DNA structures, they have been shown to loop (*57*), wrap (*58*), supercoil (*40*), pair (*59*), and circularize (*53*, *60*) DNA. The spatial arrangement of these HMG domains within the full-length protein has not been adequately addressed. The ability to contact several DNA strands simultaneously may facilitate the performance of these tasks. In this respect it is noteworthy that packing in the crystals of HMG1 domain A complexed with cisplatin-modified DNA includes the interaction of helix II of the domain, which has a high number of polar amino acids, with the widened minor groove of a second DNA helix (Figure 6). In particular, the guanidinium group of Arg47 forms a hydrogen bond to the phosphate oxygen atoms of G_9, and the Lys43 amino group is within hydrogen-bonding distance to the G_8 phosphate oxygen. The main chain carbonyl group of Glu46 interacts with the exocyclic amino group and the N3 atom of G_8 through a water-mediated hydrogen bonding network. Lys42 maintains a water-mediated interaction with the exocyclic oxygen atom of T_7, and Thr50 is positioned right at the opened hydrophobic cleft of the cisplatin cross-link of the adjacent duplex. Furthermore, when studying the DNA looping activity of HMG1 domain A, it was discovered that this function was inhibited after treatment with the sulfhydryl-specific reagent *N*-ethylmaleimide (*61*). Most likely, this reagent formed a covalent adduct with the more solvent-exposed Cys22 residue at the back of helix II. It is this helix that is in proximity to the widened minor groove of an adjacent DNA duplex in this crystal structure. Modification of that cysteine could disturb the interaction with a second DNA strand and thereby prevent DNA looping.

The mechanism of full-length HMG domain proteins binding to cisplatin-modified DNA remains unresolved. In a recent study, it was found that domain A dominates the AB didomain in binding to 4WJ DNA (*62*) despite each individual domain binding to the 4WJ in a similar fashion. In addition, a stopped-flow fluorescence competition study of HMG1 domain A and domain B binding to cisplatin-modified DNA provides evidence that HMG1 domain A controls binding when both domain A and B are present (*63*). It is quite conceivable that domain A is the dominating domain in HMG1 that binds to the platinum site, while domain B facilitates binding by providing additional protein−DNA interactions. A structure of full-length HMG1 complexed with cisplatin-modified DNA will provide more insights into how HMG-domain proteins can mediate cisplatin cytotoxicity.

ACKNOWLEDGMENT

We thank Robert J. Kennedy for helping with the CD measurements and Prof. Dan Kemp for access to instrumentation. We greatly appreciate the collaboration and guidance of Drs. C. O. Pabo and M. A. Rould during the X-ray structural study, and we thank Drs. S. M. Cohen, S. S. Marla, C. Kneip, and Ms. A. T. Yarnell for helpful discussions. We thank Johnson Matthey for a gift of cisplatin.

REFERENCES

1. Gelasco, A., and Lippard, S. J. (1998) *Biochemistry 37*, 9230−9239.
2. Takahara, P. M., Rosenzweig, A. C., Frederick, C. A., and Lippard, S. J. (1995) *Nature 377*, 649−652.
3. Yang, D., van Boom, S. S. G. E., Reedijk, J., van Boom, J. H., and Wang, A. H.-J. (1995) *Biochemistry 34*, 12912−12920.
4. Zamble, D. B., and Lippard, S. J. (1999) in *Cisplatin-Chemistry and Biochemistry of a Leading Anticancer Drug* (Lippert, B., Ed.) pp 73−110, Verlag Helvetica Chimica Acta, Zurich.
5. Jamieson, E. R., and Lippard, S. J. (1999) *Chem. Rev. 99*, 2467−2498.
6. He, Q., Liang, C., and Lippard, S. J. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 5768−5772.
7. Brown, S. J., Kellett, P. J., and Lippard, S. J. (1993) *Science 261*, 603−605.
8. McA'Nulty, M. M., Whitehead, J. P., and Lippard, S. J. (1996) *Biochemistry 35*, 6089−6099.

9. Treiber, D. K., Zhai, X., Jantzen, H.-M., and Essigman, J. M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5672−5676.
10. Zamble, D. B., Mu, D., Reardon, J. T., Sancar, A., and Lippard, S. J. (1996) *Biochemistry 35*, 10004−10013.
11. Bustin, M., and Reeves, R. (1996) *Prog. Nucleic Acid Res. Mol. Biol. 54*, 35−100.
12. Read, C. M., Cary, P. D., Crane-Robinson, C., Driscoll, P. C., Carrillo, M. O. M., and Norman, D. G. (1995) in *Nucleic Acids and Molecular Biology* (Eckstein, F., and Lilley, D. M. J., Eds.) Vol. 9, pp 222−250, Springer-Verlag, Berlin.
13. Grosschedl, R., Giese, K., and Pagel, J. (1994) *Trends Genet. 10*, 94−100.
14. Whitehead, J. P., and Lippard, S. J. (1996) in *Metal Ions in Biological Systems* (Sigel, A., and Sigel, H., Eds.) pp 687−725, Marcel Dekker, Inc., New York.
15. Grasser, K. D., Teo, S. H., Lee, K. B., Broadhurst, R. W., Rees, C., Hardman, C. H., and Thomas, J. O. (1998) *Eur. J. Biochem. 253*, 787−795.
16. Lilley, D. M. J. (1992) *Nature 357*, 282−283.
17. Ohndorf, U. M., Rould, M. A., He, Q., Pabo, C. O., and Lippard, S. J. (1999) *Nature 399*, 708−712.
18. Cohen, S. M., Jamieson, E. J., and Lippard, S. J. (2000) *Biochemistry 39*, 8259−8265.
19. Allain, F. H. T., Yen, Y. M., Masse, J. E., Schultze, P., Dieckmann, T., Johnson, R. C., and Feigon, J. (1999) *EMBO J. 18*, 2563−2579.
20. Balaeff, A., Churchill, M. E. A., and Schulten, K. (1998) *Proteins: Structure, Function and Genetics 30*, 113−135.
21. Payet, D., Hillisch, A., Lowe, N., Diekmann, S., and Travers, A. (1999) *J. Mol. Biol. 294*, 79−91.
22. Murphy IV, F. V., Sweet, R. M., and Churchill, M. E. A. (1999) *EMBO J. 18*, 6610−6618.
23. Falciola, L., Murchie, A. I. H., Lilley, D. M. J., and Bianchi, M. E. (1994) *Nucleic Acids Res. 22*, 285−292.
24. Lohman, T. M., and Mascotti, D. P. (1992) *Methods Enzymol. 212*, 400−424.
25. Lavery, R., and Sklenar, H. (1988) *J. Biomol. Struct. Dyn. 6*, 63−91.
26. Lavery, R., and Sklenar, H. (1989) *J. Biomol. Struct. Dyn. 6*, 655−667.
27. Dickerson, R. E. (1998) *Nucleic Acids Res. 26*, 1906−1926.
28. Brünger, A. T. (1992) Yale University Press, New Haven, CT.
29. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., GrosseKunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
30. Broadhurst, R. W., Hardman, C. H., Thomas, J. O., and Laue, E. D. (1995) *Biochemistry 34*, 16608−16617.
31. 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.
32. Takahara, P. M., Frederick, C. A., and Lippard, S. J. (1996) *J. Am. Chem. Soc. 118*, 12309−12321.
33. Dunham, S. U., Dunham, S. U., Turner, C. J., and Lippard, S. J. (1998) *J. Am. Chem. Soc. 120*, 5395−5406.
34. Baxevanis, A. D., and Landsman, D. (1995) *Nucleic Acids Res. 23*, 1604−1613.
35. Dunham, S. U., and Lippard, S. J. (1997) *Biochemistry 36*, 11428−11436.
36. Love, J. J., Li, X., Case, D. A., Giese, K., Grosschedl, R., and Wright, P. E. (1995) *Nature 376*, 791−795.
37. Dow, L. K., Jones, D. N. M., Wolfe, S. A., Verdine, G. L., and Churchill, M. A. (2000) *Biochemistry 39*, 9725−9736.
38. Taudte, S., Xin, H., and Kallenbach, N. R. (2000) *Biochem. J. 347*, 807−814.
39. Pontiggia, A., Rimini, R., Harley, V. R., Goodfellow, P. N., Lovell-Badge, R., and Bianchi, M. E. (1994) *EMBO J. 13*, 6115−6124.
40. Teo, S. H., Grasser, K. D., and Thomas, J. O. (1995) *Eur. J. Biochem. 230*, 943−950.
41. Berners-Price, S. J., Corazza, A., Guo, Z. J., Barnham, K. J., Sadler, P. J., Ohyama, Y., Leng, M., and Locker, D. (1997) *Eur. J. Biochem. 243*, 782−791.
42. Draganescu, A., and Tullius, T. D. (1998) *J. Mol. Biol. 276*, 529−536.
43. Grove, A., Galeone, A., Mayol, L., and Geiduschek, E. P. (1996) *J. Mol. Biol. 260*, 120−125.
44. Lebrun, A., and Lavery, R. (1999) *Biopolymers 49*, 341−353.
45. Juo, Z. S., Chiu, T. K., Leiberman, P. M., Baikalov, I., Berk, A. J., and Dickerson, R. E. (1996) *J. Mol. Bol. 261*, 239−254.
46. Grove, A., Galeone, A., Yu, E., Mayol, L., and Geiduschek, E. P. (1998) *J. Mol. Biol. 282*, 731−739.
47. Kim, Y. C., Geiger, J. H., Hahn, S., and Sigler, P. B. (1993) *Nature 365*, 512−520.
48. Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993) *Nature 365*, 520−527.
49. Robinson, H., Gao, Y.-G., McCrary, B. S., Edmondson, S. P., Shriver, J. W., and Wang, A. H.-J. (1998) *Nature 392*, 202−205.
50. Gao, Y.-G., Su, S.-Y., Robinson, H., Padmanabhan, S., Lim, L., McCrary, B. S., Edmondson, S. P., Shriver, J. W., and Wang, A. H.-J. (1998) *Nature Struct. Biol. 5*, 782−786.
51. Schumacher, M. A., Choi, K. Y., Zalkin, H., and Brennan, R. G. (1994) *Science 266*, 763−770.
52. Rice, P. A., Yang, S., Mizuuchi, K., and Nash, H. A. (1996) *Cell 87*, 1295−1306.
53. Chow, C. S., Whitehead, J. P., and Lippard, S. J. (1994) *Biochemistry 33*, 15124−15130.
54. Trimmer, E. E., Zamble, D. B., Lippard, S. J., and Essigmann, J. M. (1997) *Biochemistry 37*, 352−362.
55. Coin, F., Frit, P., Viollet, B., Salles, B., and Egly, J.-M. (1998) *Mol. Cell. Biol. 18*, 3907−3914.
56. Yoshioka, K., Saito, K., Tanabe, T., Yamamoto, A., Ando, Y., Nakamura, Y., Shirakawa, H., and Yoshida, M. (1999) *Biochemistry 38*, 589−595.
57. Stros, M., Reich, J., and Kolíbalová, A. (1994) *FEBS Lett. 344*, 201−206.
58. Fisher, R. P., Lisowsky, T., Parisi, M. A., and Clayton, D. A. (1992) *J. Biol. Chem. 267*, 3358−3367.
59. Mishima, Y., Kaizu, H., and Kominami, R. (1997) *J. Biol. Chem. 272*, 26578−26584.
60. Read, C. M., Cary, P. D., Preston, N. S., Lnenicek-Allen, M., and Crane-Robinson, C. (1994) *EMBO J. 13*, 5639−5646.
61. Stros, M., Stokrova, J., and Thomas, J. O. (1994) *Nucleic Acids Res. 22*, 1044−1051.
62. Webb, M., and Thomas, J. O. (1999) *J. Mol. Biol. 294*, 373−387.
63. Jamieson, E. R., and Lippard, S. J. (2000) *Biochemistry 39*, 8426−8438.

BI001700J