Recognition of DNA Interstrand Cross-link of Antitumor Cisplatin by HMGB1 Protein†

Jana Kasparkova,‡ Olivier Delalande,‡,§ Michal Stros,‡ Miguel-Angel Elizondo-Riojas,§,|| Marie Vojtiskova,‡ Jiri Kozelka,§ and Viktor Brabec*,‡

Institute of Biophysics, Academy of Sciences of the Czech Republic, Kralovopolska 135, CZ-61265 Brno, Czech Republic, and Laboratoire de Chimie et Biochimie Pharmacologiques et Toxicologiques, Universite Rene Descartes, UMR 8601 CNRS, 45 rue des Saints-Peres, F-75006 Paris 06, France

Received August 21, 2002; Revised Manuscript Received November 20, 2002

ABSTRACT: Several proteins that specifically bind to DNA modified by cisplatin, including those containing HMG-domains, mediate antitumor activity of this drug. Oligodeoxyribonucleotide duplexes containing a single, site-specific interstrand cross-link of cisplatin were probed for recognition by the rat chromosomal protein HMGB1 and its domains A and B using the electrophoretic mobility-shift assay. It has been found that the full-length HMGB1 protein and its domain B to which the lysine-rich region (seven amino acid residues) of the A/B linker is attached at the N-terminus (the domain HMGB1b7) specifically recognize DNA interstrand cross-linked by cisplatin. The affinity of these proteins to the interstrand cross-link of cisplatin is not very different from that to the major 1,2-GG intrastrand cross-link of this drug. In contrast, no recognition of the interstrand cross-link by the domain B lacking this region or by the domain A with or without this lysine-rich region attached to its C-terminus is noticed under conditions when these proteins readily bind to 1,2-GG intrastrand adduct. A structural model for the complex formed between the interstrand cross-linked DNA and the domain HMGB1b7 was constructed and refined using molecular mechanics and molecular dynamics techniques. The calculated accessible areas around the deoxyribose protons correlate well with the experimental hydroxyl radical footprint. The model suggests that the only major adaptation necessary for obtaining excellent surface complementarity is extra DNA unwinding (∼40°) at the site of the cross-link. The model structure is consistent with the hypothesis that the enhancement of binding affinity afforded by the basic lysine-rich A/B linker is a consequence of its tight binding to the sugar−phosphate backbone of both DNA strands.

Sufficient evidence has accumulated to identify DNA as the relevant pharmacological target of antitumor *cis*-diamminedichloroplatinum(II) (cisplatin)[1] (*1*−*3*). Cisplatin binds to DNA, forming mainly intrastrand cross-links (CLs) between adjacent purine residues (∼90%) (*4*, *5*). The minor adducts are monofunctional lesions, 1,3-GNG intrastrand CLs (N = A, C, T) and interstrand CLs (for instance ∼6% in linear DNA in cell-free media) formed preferentially between guanine residues in the 5′-GC/5′-GC sequence (*6*−*8*). In all adducts, cisplatin is coordinated to the N7 atom of purine residues. Several discoveries (*9*) have suggested that the toxicity of cisplatin originates from the 1,2 intrastrand CLs although relative antitumor efficacy of these and other CLs formed on DNA by this drug still remains unknown (*8*, *10*). Interestingly, in negatively supercoiled DNA the frequency of interstrand CLs is noticeably higher than that in the corresponding relaxed or linearized forms (*11*). The mechanism by which interstrand CLs formed by cisplatin in DNA elicit their biological responses has not been clarified. To shed light on this question, these interstrand lesions have been intensively analyzed, although in a lesser extent than more frequent DNA intrastrand CLs.

A structure of the interstrand CL formed in DNA by cisplatin has been derived on the basis of ^{1}H NMR and X-ray crystal structure analyses (*12*−*14*). In this structure the cross-linked deoxyriboguanosine residues are not paired with hydrogen bonds to the complementary deoxyribocytidines, which are located outside the duplex and not stacked with other aromatic rings. An entirely unforeseen feature is that the *cis*-diammineplatinum(II) bridge resides in the minor groove. Also surprising is the finding that the double helix is locally reversed to a left-handed, Z-DNA-like form. The change of the helix sense and the extrusion of deoxyribocytidine residues (complementary to the platinated deoxyriboguanosine residues) from the duplex result in the helix unwinding by approximately 80° (*12*, *13*). The helix axis is

† Supported by the Grant Agency of the Czech Republic (Grant Nos. 305/02/1552 and 305/01/0418), the Internal Grant Agency of the Ministry of Health of the Czech Republic (Grant No. NL6058-3/2000), and the Grant Agency of the Academy of Sciences of the Czech Republic [A5004105 (to M. S.)].

* Corresponding author. Tel, 420−541517148; fax, 420−541240499; e-mail: brabec@ibp.cz.

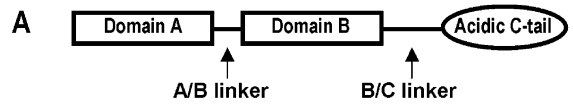
‡ Institute of Biophysics.

§ Universite Rene Descartes.

|| Present address: Centro Universitario Contra el Cancer, Hospital Universitario "Dr. Jose Eleuterio Gonzales" Universidad Autonoma de Nuevo Leon, Monterrey, N.L., Mexico.

[1] Abbreviations: bp, base pair; cisplatin, *cis*-diamminedichloroplatinum(II); CL, cross-link; EMSA, electrophoretic mobility-shift assay; GB, Generalized Born; HMG, high mobility group; MD, molecular dynamics; SAS, solvent-accessible-surface.

10.1021/bi026695t CCC: $25.00
Published on Web 01/17/2003



B

Domain A (HMGB1a):	^{1}MGKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFS KKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGE^{84}
Domain A' (HMGB1a'):	HMGB1a+^{85}TKKKF^{89} (residues 1-89)
Region of A/B linker:	^{85}TKKKFKD^{91}
Domain B (HMGB1b):	^{92}PNAPKRPPSAFFLFCSEYRPKIKGEHPGLSIGDVAKKLGEMW NNTAADDKQPYEKKAAKLKEKYKEDIAAY^{162}
B/C linker:	^{163}RAKGKPDAAKKGVVKAEKSKKKK^{185}
Acidic C tail:	^{186}EEEDDEEDEEDEEEEEEEEDEDEEEDDDDE^{215}
Domain B' (HMGB1b'):	^{85}TKKKFKD^{91}+HMGB1b (residues 85-162)
Domain B3 (HMGB1b3):	^{89}FKD^{91}+HMGB1b+^{163}RAKGKPDAAKKGVVKAEK^{180} (residues 89-180)
Domain B7 (HMGB1b7):	HMGB1b'+^{163}RAKGKPDAAKKGVVKAEK^{180} (residues 85-180)

FIGURE 1: The domain organization of the rat HMGB1 protein (A) and sequences of the HMGB1 domains and linker regions from this protein as used in the present work with their abbreviations (B). Numbers at the start and end of each amino acid sequence represent the residue numbering in the full-length protein (see GenBank accession no. Y00463).

bent by 20−45° toward the minor groove, the minor groove is enlarged to >1.0 nm, and the square planar platinum coordination remains preserved.

Gel mobility shift assays have revealed the presence of proteins in mammalian cellular extracts that bind specifically to DNA modified by antitumor cisplatin and its direct analogues (*3*, *15*, *16*). It has been suggested (*15*, *16*) that these proteins mediate cisplatin cytotoxicity. One class of these proteins has been identified as proteins containing a high mobility group (HMG) domain, HMGB1 and HMGB2 proteins. These structure-specific DNA-binding proteins belong to architectural chromatin proteins that play some kind of structural role in the formation of functional higher order protein−DNA or protein−protein complexes (*17*). The HMGB1 protein contains two tandem HMG-box domains, A and B (HMGB1a and HMGB1b), and an acidic C-terminal tail (schematic representation of the domains of HMGB1 protein as used in the present work is shown in Figure 1) (*18*). The domains A and B in the HMGB1 protein are linked by a short sequence (A/B linker) containing a short lysine-rich segment of seven amino acids at the N-terminus of the domain B, whereas the domain B and C tail are linked by a longer chain (containing 23 amino acids, B/C linker region) (Figure 1).

The initial studies were focused on recognition by these proteins of 1,2- and 1,3-intrastrand CLs and monofunctional adducts. It has been found (*17*) that full-length HMGB1 protein and each domain HMGB1a and HMGB1b (both lacking the lysine-rich region of the A/B linker) bind selectively to the 1,2-GG or AG adducts of cisplatin, but not to the 1,3-intrastrand CLs or monofunctional adducts. At the same time, HMG-domain proteins have been reported not to bind to 1,3-intrastrand CLs and monofunctional adducts of clinically ineffective transplatin [*trans*-diamminedichloroplatinum(II)]. The affinity of the domain HMGB1a for DNA containing 1,2-GG intrastrand CL is generally higher than that of the domain HMGB1b, and it has been proposed (*19*, *20*) that HMGB1a is the dominating domain in the full-length HMGB1 protein that binds to the site of the intrastrand CL, while the domain HMGB1b facilitates binding providing additional protein−DNA interactions. Also, interestingly, the binding affinity of HMGB1a and HMGB1b domains to the duplex containing 1,2-GG intrastrand CL is modulated by the nature of the bases that flank the platinum lesion (*21*, *22*).

The domain HMGB1a binds to the minor groove of the DNA double helix opposite the cisplatin 1,2-intrastrand CL located in the major groove (*19*, *23*). Distortions such as prebending, unwinding at the site of platination, and preformation of a hydrophobic notch as a consequence of the CL formation are important for the recognition and affinity of the proteins that specifically bind DNA damaged by cisplatin (*19*, *23*). Intercalating residues in HMG-domain proteins influence both the affinity and orientation of the protein binding to this CL (*19*). Moreover, the lysine-rich region of the A/B linker in the full-length HMGB1 protein attached to the N-terminus of the domain HMGB1b markedly enhances binding of this domain to DNA containing site-specific 1,2-GG intrastrand CL of cisplatin stressing an importance of this domain for specific binding of the HMGB1 protein to this platinum lesion (*24*). It has been also proposed (*25*) that binding of the A/B linker region within the major groove of DNA helps the two HMGB1 domains anchor to the minor DNA groove facilitating their DNA binding.

These studies have been extended to DNA interstrand CLs produced by cisplatin by using 110-base pairs (bp) oligodeoxyribonucleotide probe containing five equally spaced interstrand adducts (*26*). It has been found that calf thymus full-length HMGB1 protein binds to this interstrand cross-linked probe with an affinity similar to that of the similar DNA probe, but containing five equally spaced 1,2-GG intrastrand CLs. On the other hand, no binding of the HMGB1 protein to the probe containing transplatin-interstrand CL has been noticed.

A later study (*27*) has also proposed binding of the full-length HMGB1 protein to DNA containing interstrand CLs, whereas it has been reported quite recently that interstrand CLs of cisplatin are not recognized and bound by full-length rat HMGB1 and its domain A (*28*) and that HMGB1b does not exhibit a distinct gel shift in the presence of 15-bp duplex containing interstrand CL of cisplatin in a conventional electrophoretic mobility shift assay (*29*). In the present work, we have reexamined recognition and binding of the interstrand CL of cisplatin by full-length rat HMGB1 protein and its domains A and B by using various DNA probes containing single, site-specific interstrand CL of cisplatin as a substrate. We have found that the full-length HMGB1 protein and its domain B to which the lysine-rich region of the A/B linker is attached do bind to DNA interstrand cross-linked by cisplatin with an affinity similar to major 1,2-GG intrastrand CL of this drug. On the other hand, the domains HMGB1a or HMGB1b lacking this region of the A/B linker do not bind to the interstrand CL under conditions when these proteins readily bind to 1,2-GG intrastrand adduct. In addition, a structural model based on hydroxyl radical footprinting and molecular modeling of the complex formed between DNA interstrand cross-linked by cisplatin and the domain HMGB1b containing the A/B linker is proposed.

MATERIALS AND METHODS

Preparation of Oligonucleotide Probes Containing Single, Site-Specific Interstrand and 1,2-GG Intrastrand CLs. The synthetic oligodeoxyribonucleotides (Table 1) were pur-

Table 1: Oligodeoxyribonucleotide Sequences and Abbreviations[a]

abbreviation	sequence
TGGT	5′-CCTCTCTCT**GG**TCTTCTTCT-3′
ACCA(1)	5′-AGAAGAAGACCAGAGAGAGG-3′
TGCT	5′-TCTCCTTCT**G**CTCTCTTCTC-3′
AGCA(1)	5′-GAGAAGAGA**G**CAGAAGGAGA-3′
ACCA(2)	5′-GCTCAGAAGAAGACCAGAGAGAGGTCAG-3′
AGCA(2)	5′-GCTCGAGAAGAGA**G**CAGAAGGAGATCAG-3′
64-mer(1)	5′-GGGACCTGAACACGTACGGAATTCGATATCCTCGAGCCAG ATCTGCGCCAGCTGGCCACCCTGA-3′
64-mer(2)	5′-GAGCGCCAAGCTTGGGCTGCAGCAGGTCGACTCTAGAGG ATCCCGGGCGAGCTCGAATTCGCCC-3′
61-mer	5′-GGGGCGAATTCGAGCTCGCCCGGGATCCTCTAGAGTCGA CCTGCTGCAGCCCAAGCTTGGC-3′
59-mer	5′-GGTGGCCAGCTGGCGCAGATCTGGCTCGAGGATATCGAA TTCCGTACGTGTTCAGGTCC-3′

[a] Boldface indicates the sites of platination.

chased from IDT, Inc. (Coralville, IA), and purified as described previously (*30*, *31*); in the present work, their molar concentrations are related to the whole duplexes.

The 20-bp oligonucleotide duplexes uniquely and site-specifically interstrand and intrastrand cross-linked by cisplatin were prepared and characterized as described previously (*26*, *30*, *32*, *33*). The 148-bp duplex containing the single and central 1,2-GG intrastrand CL was prepared by using the 20-mer oligonucleotide TGGT containing the single 1,2-intrastrand CL of cisplatin at the GG site. The uniquely modified 20-mer was 5′ end-labeled, annealed with a set of five complementary and partially overlapping oligonucleotides [ACCA(2), 64-mer(1), 64-mer(2), 61-mer, and 59-mer], and ligated with T4 DNA ligase. The 148-bp duplex containing the single and central interstrand CL formed by cisplatin between guanine residues in the 5′-d(GC)/5′-d(GC) site was prepared by using the 20-mer oligonucleotide TGCT containing the single monofunctional adduct of cisplatin at the G site (*30*). This 20-mer monofunctionally modified by cisplatin was 5′ end-labeled and annealed first with the partially overlapping purine-rich bottom strand ACCA(2) (28-mer) to evolve the interstrand CL (*26*, *33*). The duplex was purified by FPLC (*26*, *33*) and further annealed with a set of four complementary and partially overlapping oligonucleotides [64-mer(1), 64-mer(2), 61-mer and 59-mer], and ligated with T4 DNA ligase. Full-length substrates containing the single intrastrand or interstrand CL were separated from unligated products in a 6% denaturing polyacrylamide gel, purified by electroelution, reannealed, and stored in annealing buffer [50 mM Tris-HCl (pH 7.9), 100 mM NaCl, 10 mM $MgCl_2$, and 1 mM DDT] at −20 °C. Other details of the purification of the 148-bp DNA substrates for recognition by full-length HMGB1 were the same as described previously (*34*, *35*).

Proteins. Expression and purification of recombinant rat full-length HMGB1 protein (Figure 1) and its domains (Figure 1) were carried out as described (*24*, *36*).

Electrophoretic Mobility-Shift Assays with HMGB1 Domain Proteins. Radioactively labeled probe DNAs (4 nM) were titrated with the domains HMGB1a, HMGB1a′, HMGB1b3, HMGB1b7, or HMGB1b′ in 10 μL sample volumes in buffer I composed of 10 mM HEPES, pH 7.5, 10 mM $MgCl_2$, 50 mM LiCl, 100 mM NaCl, 1 mM spermidine, 0.2 mg/mL bovine serum albumin, and 0.05% v/v 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 a running, precooled (4 °C), prerun (300 V, 1−2 h) 5% native polyacrylamide gels [29:1 acrylamide:bisacrylamide, 0.5× TBE buffer (45 mM Tris-HCl, 45 mM boric acid and 1 mM EDTA, pH 8.3)]. Gels were electrophoresed at 4 °C and 300 V for ∼1.5 h, dried, exposed to a molecular imaging plate, and analyzed on a Molecular dynamics phosphor imager (Storm 860 system), and the bands were quantitated with the ImageQuant software. The experiments with the full-length HMGB1 protein were performed in the same way, but with small modifications. The labeled DNA probes were titrated with the protein in the presence of 0.1 mg/mL sonicated calf thymus DNA in buffer II composed of 10 mM HEPES, pH 7.9, 10 mM $MgCl_2$, 150 mM NaCl, 0.2 mg/mL bovine serum albumin, 20% v/v glycerol, and 1 mM dithiothreitol.

Apparent dissociation constants, K_d, were estimated in the same way as described (*19*). Each K_d is the average of at least three measurements.

Hydroxyl Radical Footprinting Assays. These assays were performed in the same way as described (*19*) with small modifications. Labeled 20-bp duplexes with or without protein (20-fold excess protein over DNA) were incubated at 4 °C for 30 min in the buffer I. A fresh mixture of sodium ascorbate (20 mM), $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ (5 mM), EDTA (10 mM), and H_2O_2 (0.6%) was added to a 20 μL aliquot to initiate the reaction. After 5 min at 25 °C, 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/8 M urea gels. Gels were dried and quantitated on a Molecular dynamics phosphor imager (Storm 860 system). The peak areas corresponding to each band were compared, and when the difference between the control and protein-added lanes was more than 10%, the corresponding base was scored as being protected (*19*, *23*). Other details were the same as in previously published papers (*19*, *23*).

Molecular Modeling. The molecular modeling used the program AMBER Version 6 (*37*). The force-field file Parm98 (*38*) was supplemented by parameters defining the platinum coordination sphere as described previously (*39*). The Generalized Born (GB) solvation model (*40*) was used throughout. The molecular dynamics (MD) simulations were carried out at a constant temperature of 300 K and used the Berendsen algorithm for temperature control. The analysis of the MD trajectories used the CARNAL module of AMBER 4.1 and, for the DNA part, the program CURVES (*41*).

A model of the 20-bp duplex TGCT/AGCA(1) (Table 1) bearing an interstrand CL formed by cisplatin at the central

d(GC)/d(GC) sequence was constructed by fitting the backbone of the inner 10-bp part to the X-ray coordinates of a similarly cross-linked 10-bp duplex (entry 1A2E of the Protein Data Bank) (*14*). The coordinates of the inner dinucleotide bearing the CL were taken directly from the X-ray data. This approach is sustained by similarity of all three X-ray and NMR structures of DNA interstrand CL of cisplatin so far available (*12*−*14*). A similarly distorted structure was thus expected for the 20-bp interstrand cross-linked duplex tested in the present work. A starting structure for the protein was based on one of the NMR-derived models proposed for the 77 amino acids long (residues 88−164) part of the full-length HMGB1 protein (entry 1HMF, structure N°2 of the Protein Data Bank) (*42*) which was extended by attaching three additional amino acids to the N-terminus and 16 amino acids to the C-terminus in an unstructured linear form so that the final protein corresponded to the domain HMGB1b7 (amino acids 85−180 in Figure 1).

Both DNA and protein parts were first relaxed separately using energy minimization and MD techniques. The MD simulation of the protein part showed high mobility of the unstructured segment attached to the C-terminal; from the variety of positions, we chose one where this segment did not loop back onto the part where the DNA was supposed to bind. The duplex containing the interstrand CL was then manually docked onto the protein surface with the phenylalanine F103 from the helix I protruding into the cleft delimited by the two cross-linked guanines (see also the section Results). This procedure yielded two initial structures of the complex differing in the orientation of the DNA strands relative to the protein. Of the two resulting models, that in a qualitative agreement with the hydroxyl radical footprint was selected for further refinement.

The DNA−protein complex was subjected to a 1080 ps MD simulation. The last 480 ps of production were retained for analysis. Snapshots were stored every 1 ps. A structure averaged over the last 480 ps was subjected to 1000 cycles of GB energy minimization; this structure was used for calculations of the solvent-accessible-surface (SAS). The reference structure of free DNA (not bound to the protein) was obtained in a similar way.

SAS calculations were performed using the program msms (*43*) with probe radius value of 1.4 Å and a 3.0 density value. Contact surfaces were measured with the GBSA module of the Amber6 program. Hydrogen bonds were detected with the Carnal module of Amber6. The criterion used was a distance of ≤4.0 Å between the heavy atoms and a X−H···Y angle of more than 1 rad.

RESULTS

Full-Length HMGB1 Protein Binds to a 148-bp DNA Duplex Containing Single Interstrand CL of Cisplatin. To examine the affinity and specificity of the full-length HMGB1 protein for interstrand CL of cisplatin, a 148-bp probe (see the section Materials and Methods), nonmodified or containing single, site-specific interstrand CL formed between guanine residues in the central 5′-TGCT/5′-AGCA sequence was used in electrophoretic mobility-shift assay (EMSA) with recombinant rat HMGB1 protein (Figure 1). As indicated by the presence of a shifted band of the 148-bp probe containing the single interstrand CL of cisplatin

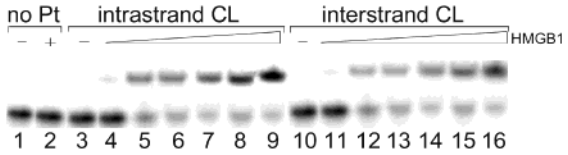


FIGURE 2: Gel mobility shift assay analysis of the titration of 148-bp DNA probe containing single, central site-specific 1,2-GG intrastrand CL of cisplatin (lanes 3−9) or single, central site-specific interstrand CL formed by cisplatin between guanine residues in the 5′-GC/5′-GC sequence (lanes 10−16) with full-length HMGB1 protein. Protein concentration was 0 nM (lanes 1, 3, 10), 97 nM (lanes 4, 11), 480 nM (lanes 5, 12), 970 nM (lanes 6, 13), 2170 nM (lanes 7, 14), 3500 nM (lanes 8, 15), and 6900 nM (lanes 2, 9, 16). The assay of the unplatinated 148-bp probe (containing the central sequence TGCT/AGCA) is shown in lanes 1 and 2.

with the protein [the intensity of which increased with growing protein concentration (Figure 2, lanes 11−16)], the protein exhibited affinity for uniquely interstrand cross-linked DNA. Not surprisingly, the HMGB1 protein also exhibited affinity for the same probe containing single 1,2-GG intrastrand CL of cisplatin in the central sequence TGGT/ACCA (Figure 2, lanes 4−9). The interaction between the protein and the interstrand cross-linked DNA was not a result of general DNA affinity for the protein because the protein failed to bind the unmodified 148-bp duplexes under identical conditions (Figure 2, lane 2). Thus, these results are consistent with our previous finding obtained with a different DNA probe, namely, a 110-bp oligodeoxyribonucleotide duplex containing five equally spaced interstrand adducts (*26*).

Since only a single shifted band formed following incubation of the 148-bp probe containing either interstrand or 1,2 intrastrand CL of cisplatin, detailed titration studies were carried out (Figure 2). Evaluation of the titration data for the probes containing either interstrand or intrastrand CL afforded the values of apparent dissociation constant, K_d, of 2 and 0.5 μM, respectively, indicating a slightly lower affinity of the full-length HMGB1 protein to the interstrand CL than to the 1,2-GG intrastrand adduct.

Isolated HMGB1 Domain A (HMGB1a) Does Not Bind to 20-bp DNA Duplex Containing a Single, Site-Specific Interstrand CL of Cisplatin. It has been demonstrated (*15*, *19*, *23*) that single domains HMGB1a and HMGB1b are sufficient for binding the 1,2-GG intrastrand CL of cisplatin and that these single domains bind to the duplexes containing the single, site-specific intrastrand CL as small as 15 bp. Therefore, in the present study we also investigated recognition and binding of single domains HMGB1a and HMGB1b (Figure 1) to a short 20-bp duplex TGCT/AGCA(1) (for its nucleotide sequence see Table 1) containing a single, site-specific interstrand CL formed by cisplatin between guanine residues at the central sequence 5′-GC/5′-GC. In these studies we employed in bandshift assay the domain HMGB1a (residues 1−84, Figure 1), the domain HMGB1a′ (residues 1−89, Figure 1), and various HMGB1 domains B containing the minimal HMGB1b (residues 92−162, Figure 1) to which several residues of the linker regions A/B or B/C were attached [HMGB1b′ (residues 85−162), HMGB1b3 (residues 89−180), and HMGB1b7 (residues 85−180), Figure 1]. For comparative purposes, the EMSAs under the same conditions were also performed with the 20-bp duplex TGGT/ACCA-(1) containing a single, site-specific 1,2-GG intrastrand CL of cisplatin (for its nucleotide sequence see Table 1).

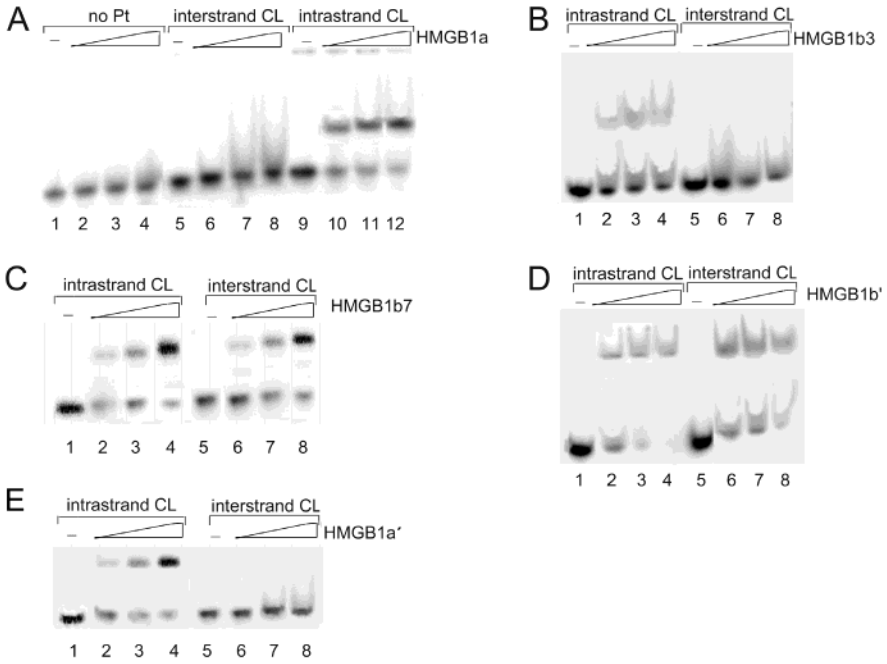


FIGURE 3: Gel mobility shift assay of 20-bp probes TGGT/ACCA(1) or TGCT/AGCA(1) containing single, site-specific 1,2-GG intrastrand CL of cisplatin or interstrand CL formed by this drug between guanine residues in the 5′-GC/5′-GC sequence, respectively, with the domains HMGB1a (A), HMGB1b3 (B), HMGB1b7(C), HMGB1b′ (D), and HMGB1a′ (E). The DNA probes were at the concentration of 4 nM. Lanes in panel A: 1−4, unplatinated duplex TGCT/AGCA(1); 5−8, the duplex containing interstrand CL; 9−12, the duplex containing intrastrand CL; protein concentration was 0 nM (lanes 1, 5, 9), 9.6 nM (lanes 2, 6, 10), 24 nm (lanes 3, 7, 11), and 48 nM (lanes 4, 8, 12). Lanes in panels B and D: 1−4, the duplex containing intrastrand CL; 5−8, the duplex containing interstrand CL; protein concentration was 0 μM (lanes 1, 5), 2.2 μM (lanes 2, 6), 5.4 μM (lanes 3, 7), and 10.8 μM (lanes 4 and 8). Lanes in panel C: 1−4, the duplex containing intrastrand CL; 5−8, the duplex containing interstrand CL; protein concentration was 0 μM (lanes 1, 5), 0.1 μM (lanes 2, 6), 0.5 μM (lanes 3, 7), and 1.5 μM (lanes 4 and 8). Lanes in panel E: 1−4, the duplex containing intrastrand CL; 5−8, the duplex containing interstrand CL; protein concentration was 0 nM (lanes 1, 5), 9.6 nM (lanes 2, 6), 24 nm (lanes 3, 7), and 48 nM (lanes 4, 8).

As expected the domain HMGB1a readily formed a complex with the 20-bp probe containing 1,2-GG intrastrand CL (Figure 3A, lanes 10−12). The protein−DNA complex was already apparent at the protein concentration as low as 1 nM (not shown), and evaluation of the titration data afforded K_d value of 10 nM. In contrast, the binding of the domain HMGB1a to the interstrand cross-linked 20-bp probe TGCT/AGCA(1) was not observed even at the protein concentration of 48 nM (Figure 3A, lanes 6−8), and it exhibited no binding to the unplatinated duplexes at the protein concentration as high as 1.6 μM (shown in Figure 3A for the protein concentrations in the range of 9.6−48 nM). Thus, we have confirmed the previous findings (*15*, *19*, *23*) that the domain HMGB1a binds to the 1,2-GG intrastrand CL of cisplatin, whereas it does not bind to the single interstrand CL of this drug contained in the short oligonucleotide duplexes (*28*).

Isolated HMGB1 Domain B (HMGB1b) to Which the Lysine-Rich Region of the A/B Linker is Attached at its N-Terminus Binds to the 20-bp DNA Duplex Containing the Single, Site-Specific Interstrand CL of Cisplatin. The 20-bp duplexes TGCT/AGCA(1) or TGGT/ACCA(1) containing either single interstrand or 1,2-GG intrastrand CL, respectively, were also screened for binding by the domain HMGB1b to which residues of the flanking linker regions A/B and/or B/C were attached (Figure 3B−D). Panels B and C of Figure 3 show the gel mobility shift assay of both duplexes with the domains HMGB1b3 and HMGB1b7, respectively, at various protein concentrations. Both proteins contain the minimal domain HMGB1b (residues 92−162), to which at its C terminus the fragment of 18 residues from the B/C linker (residues 163−180) is attached (Figure 1). The proteins HMGB1b3 and HMGB1b7 differ in the number of the residues contained in the fragment of the A/B linker attached to the N-terminus of the minimal domain HMGB1b. While HMGB1b7 contains seven residues ^{85}TKKKFKD^{91} of the A/B linker, the HMGB1b3 contains only three residues from this sequence (residues ^{89}FKD^{91}) (Figure 1). Both HMGB1b7 and HMGB1b3 bound the duplex containing the intrastrand CL (Figure 3B and C, lanes 2−4), whereas they exhibited no binding to the unplatinated duplexes at the protein concentration as high as 6.0 μM (shown in Figure 3C, lane 8 for HMGB1b7). Titration experiments revealed that the affinity of the domain HMGB1b7 to the 1,2-GG intrastrand CL was higher (K_d values were 0.4 or 3 μM for HMGB1b7 or HMGB1b3, respectively). This result is in accord with the previously published results (*24*) also indicating that attachment of seven amino acid residues of the lysine-rich region of the A/B linker to the minimal domain HMGB1b increases the binding affinity of HMGB1b to the 1,2-GG intrastrand CL of cisplatin. Interestingly, the attachment of this ^{85}TKKKFKD^{91} region of the A/B linker to the HMGB1b also markedly enhanced its binding to the interstrand CL of cisplatin (Figure 3C, lanes 5−7). While no binding of the domain HMGB1b3 to the duplex containing the interstrand CL of cisplatin was noticed even at the concentration of the protein of 10.8 μM (Figure 3B, lane 7), the concentration of the domain HMGB1b7 as low as 50 nM was sufficient to observe its binding to the interstrand cross-linked duplex (Figure 3C, lane 5) (K_d was 1.0 μM).

As the domain HMGB1b7 also contained 18 residues from the B/C linker attached to its C-terminus, we also examined

the affinity of the domain HMGB1b′ to the interstrand CL of cisplatin. Both domains HMGB1b′ and HMGB1b7 have attached to their N-terminus all seven residues ^{85}TKKKFKD^{91} of the A/B linker, but they differ in that the former contains no additional residues from the B/C linker (Figure 1). The domain HMGB1b′ also bound the 20-bp duplex TGCT/AGCA(1) containing the interstrand CL of cisplatin (Figure 3D, lanes 5−7) with an affinity (K_d was 1.4 μM) which was almost identical to that of the domain HMGB1b7. Thus, the enhanced affinity of the domain HMGB1b7 to the interstrand CL of cisplatin is mainly a consequence of the attachment of the lysine-rich region of the A/B linker to the N-terminus of the minimal domain B, whereas the attachment of the residues to its C-terminus from the B/C linker has a negligible effect. Hence, the results of the present work indicate that the attachment of the lysine-rich region of the A/B linker (residues ^{85}TKKKFKD^{91}) to the minimal domain HMGB1b not only augments its binding to both types of cisplatin CLs, but also plays a fundamental role in the binding of the domain B of HMGB1 to the interstrand CL of this platinum drug.

It could be argued that the lysine-rich part of the A/B linker could also increase affinity of the domain A to the interstrand CL of cisplatin. Therefore, we also examined the affinity of the domain HMGB1a′ (the residues 1−89, Table 1), i.e., the HMGB1 domain A to which the lysine-rich part of the A/B linker was attached to its C-terminus. The domain HMGB1a′ readily formed a complex with the 20-bp probe containing 1,2-GG intrastrand CL (Figure 3E, lanes 1−4). The complex formed between DNA and protein was already apparent at the protein concentration as low as 1 nM (not shown), and evaluation of the titration data afforded K_d value of 10 nM. Hence, it is evident that the attachment of the lysine-rich part of A/B linker to the C-terminus of the domain HMGB1a does not affect affinity of this domain to the 1,2-GG intrastrand CL of cisplatin. In contrast, the binding of the domain HMGB1a′ to the interstrand cross-linked 20-bp probe TGCT/AGCA(1) was not observed even at the protein concentration of 48 nM (Figure 3E, lanes 5−8). This result indicates that the attachment of the lysine-rich part of the A/B linker to the C-terminus of the HMGB1a domain does not considerably affect its affinity to the interstrand CL of cisplatin, in contrast to the case when this lysine-rich region is attached to the N-terminus of the HMGB1b domain. The values of apparent dissociation constants, K_d, determined for HMG-domain proteins interactions with oligonucleotide duplexes containing single, site-specific CLs of cisplatin used in the present work are summarized in the Table 2.

Table 2: K_d Values Determined for HMG-Domain Protein Interactions with Oligodeoxyribonucleotide Duplexes Cross-Linked by Cisplatin

protein	1,2-GG intrastrand CL	interstrand CL
HMGB1a	10 nM[a]	≫48 nM[a]
HMGB1a′	10 nM[a]	≫48 nM[a]
HMGB1b′	0.6 μM[a]	1.4 μM[a]
HMGB1b3	3 μM[a]	≫10.8 μM[a]
HMGB1b7	0.4 μM[a]	1.0 μM[a]
full-length HMGB1	0.5 μM[b]	2 μM[b]

[a] DNA probe was a 20-bp duplex containing central single, site-specific CL. K_d values were obtained in the absence of competitor DNA. [b] DNA probe was a 148-bp duplex containing central single, site-specific CL. K_d values were obtained in the presence of 0.1 mg/mL sonicated calf thymus DNA as competitor DNA.

Hydroxyl Radical Footprint of the Domain HMGB1b7 Bound to the 20-bp Duplex Containing the Single, Site-Specific Interstrand CL of Cisplatin Covers Five Base Pairs and Is Asymmetric. Figure 4 shows the detailed analysis of hydroxyl radical footprints of the domain HMGB1b7 on the 20-bp duplex TGCT/AGCA(1) containing the single, site-specific interstrand CL of cisplatin. In this duplex the top (pyrimidine-rich) or bottom (purine-rich) strands were labeled. The regions of protection from the hydroxyl radical cleavage reactions are shown schematically in Figure 4D. An asymmetric protection pattern was observed, in which the binding site of the HMGB1b7 protein extends toward the 3′ side of the platinated nucleotide in the top strand (TGCT) [or toward the 5′ side of the platinated nucleotide in the bottom strand (AGCA(1)].

Structural Model for the Complex Formed between DNA Containing Interstrand CL of Cisplatin and the Domain HMGB1b7. The structural studies of the interstrand CL of cisplatin have demonstrated that this adduct induces a bend in the direction of the minor groove, with extrusion of the two cytosines of the cross-linked d(GC)/d(GC) sequence (*12*−*14*). Thus, a reasonable assumption is that the HMGB1b domain, which has the L-shape characteristic of HMGB proteins (*42*), recognizes the CL by binding to the convex side of the kinked helix. This would mean that the protein makes the contacts mainly with the groups located in the major groove of the DNA near the CL. In contrast, in all cases examined to date, HMG-domain proteins bent DNA in the direction of the major groove, and the protein made contacts with the groups in the minor groove. We were interested to know whether such a recognition complex involving DNA bent toward the minor groove is structurally and energetically feasible without a major adaptation of either the DNA duplex or the protein. A recent X-ray crystal structure of the 10-bp DNA containing an interstrand CL of cisplatin (*14*) and an NMR structure of the HMGB1b domain (residues 88−164) (*42*) was used as a basis for the modeling of the recognition complex formed between the interstrand cross-linked duplex TGCT/AGCA(1) and the domain HMGB1b7.

Models for the two parts (constructed and relaxed as outlined in the sections Materials and Methods) were docked together, fitting the convex side of the kinked DNA into the concave side of the L-shaped protein. These docking experiments have revealed that phenylalanine F103 fits very well into the cavity delimited by the two cross-linked guanine residues. F103 protrudes from helix I of the HMGB1b domain and corresponds to the "primary intercalating residue" (*18*) used by HMG-domain proteins to stabilize kinked DNA in recognition complexes. It was therefore reasonable to assume that this classical intercalation mode could be used in the present complex as well. Both of the two possible orientations of the DNA duplex with respect to the protein were initially considered. However, relaxation of the structures revealed that only in one of the two orientations the protein binds to that side of the interstrand CL where the footprinting analysis indicated inhibition of cleavage. Hence, finally only this model was retained and further refined. The model was subjected to an unconstrained MD simulation at a simulated temperature of 300 K for 1080

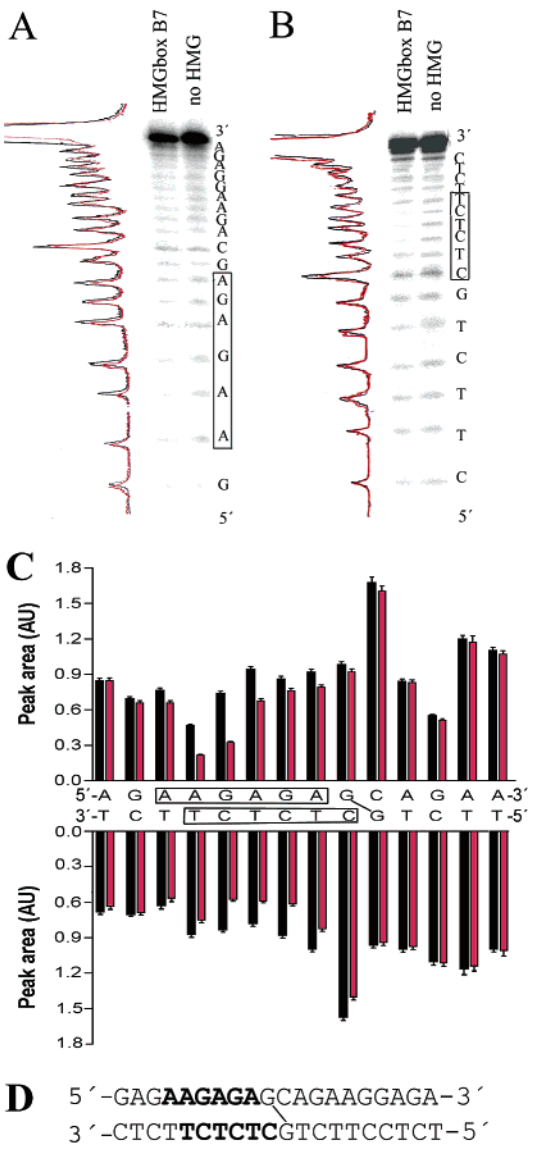


FIGURE 4: Hydroxyl radical footprint of the domain HMGB1b7 bound to the TGCT/AGCA(1) duplex containing the single, site-specific interstrand CL formed by cisplatin between guanine residues in the 5′-GC/5′-GC sequence in which the top (A) or bottom strand (B) was 5′-end radioactively labeled; black and red lines are the traces for a control sample without protein added and for the footprint with protein added, respectively. (C) The quantitative representation of the evaluation of the peak areas (in arbitrary units, AU) corresponding to each band in panels A and B; black and red bars represent the peak areas of the bands corresponding to the control samples without protein added and to the footprint with the protein added, respectively. Only when the difference between the control and protein-added lanes was more than 10%, was the corresponding base scored as being protected (19, 23). Boxed letters in panels 4A−C indicate the protection areas. (D) Summary of footprint results. Bold letters indicate the protected areas and a short line linking two G bases in the different strands of the duplex indicates the sites of platinum interstrand cross-linking.

ps. The fluctuations of atomic positions reached a plateau after 600 ps and during the last 480 ps the model showed stable behavior (rms ≤ 3.0 Å). No signs of dissociation were observed.

The final optimized model obtained by time-averaging over the 480 ps production period is shown in Figure 5. The calculated excluded SAS is 1760 Å^2, which is a typical value found for protein−DNA recognition modules (*44*). Figure 5 shows the perfect fit of the duplex bent in the direction of the minor groove into the L-shaped structure formed by the three helices of the domain HMGB1b7. Specifically, the helix I fits into the major groove of the branch of the kinked DNA extending from the interstrand CL toward the 5′-end of the top strand of the duplex TGCT/AGCA(1) and toward the 3′-end of its bottom strand (5′-branch). On the other hand, the helix III is inserted into the major groove of the other branch (3′-branch).

Consistent with the structures of other complexes formed between HMG-domains and DNA (*18*), in our model both the A/B and B/C linkers participate in the recognition of interstrand CL of cisplatin, contacting mainly the sugar−phosphate backbone of the 3′-branch (Figure 5). The B/C linker is wrapped around the sugar−phosphate backbone of the top strand of the duplex TGCT/AGCA(1), while the basic lysine-rich A/B linker binds very tightly to the sugar−phosphate backbone of both DNA strands. The tight binding of the lysine-rich region of the A/B linker gives a plausible rationale for the reduced affinity of the domain HMGB1b3 in which the three N-terminal lysines from the sequence ^{85}TKKKFKD^{91} of the A/B linker are missing (vide supra).

The Formation of the Complex between DNA Containing Interstrand CL of Cisplatin and the Domain HMGB1b7 Requires Relatively Moderate Adaptation of DNA and the Protein. A comparison of our model of the complex formed between the cross-linked duplex and HMGB1b7 (Figure 5) with the optimized structures of the cross-linked duplex and the protein not contained in the complex (not shown) has revealed that both have undergone relatively moderate structural changes as a consequence of the complex formation. The main alterations that occur in the protein moiety concern the two flexible tails which got fixed in the complex through electrostatic contacts with the sugar−phosphate backbone (vide supra). As for the DNA part, the most striking change is an extra unwinding of ∼40° at the CL site (increasing the total unwinding angle to ∼110°), apparently optimizing the positions of the phosphate residues interacting with the protein.

The Solvent-Accessible Surface (SAS) around Deoxyribose Protons and Amino Acid−Phosphate Hydrogen Bonding Contacts Correlate with the Hydroxyl Radical Footprint. Hydroxyl radical footprinting makes it possible to map areas of a DNA whose accessibility has been reduced by fixation of a recognition protein. Both major (*45*) and minor (*46*) groove binding can be detected with this method. Hydroxyl radicals induce strand breaks abstracting a deoxyribose proton. The contribution of the individual hydrogen atoms in deoxyribose residues to the overall DNA cleavage by hydroxyl radicals has been shown to be determined almost exclusively by the accessibility of these atoms (*47*, *48*). Hence, the experimental footprint (Figure 4) has been correlated with the sum of the SAS calculated over all deoxyribose protons. A similar approach was recently used to correlate footprinting experiments with MD simulations when the TATA-binding protein bound to its high-affinity DNA sequence (*49*). The difference between the sum of the SAS in the platinated DNA alone and that in the complex shows a good agreement with the relative cleavage observed

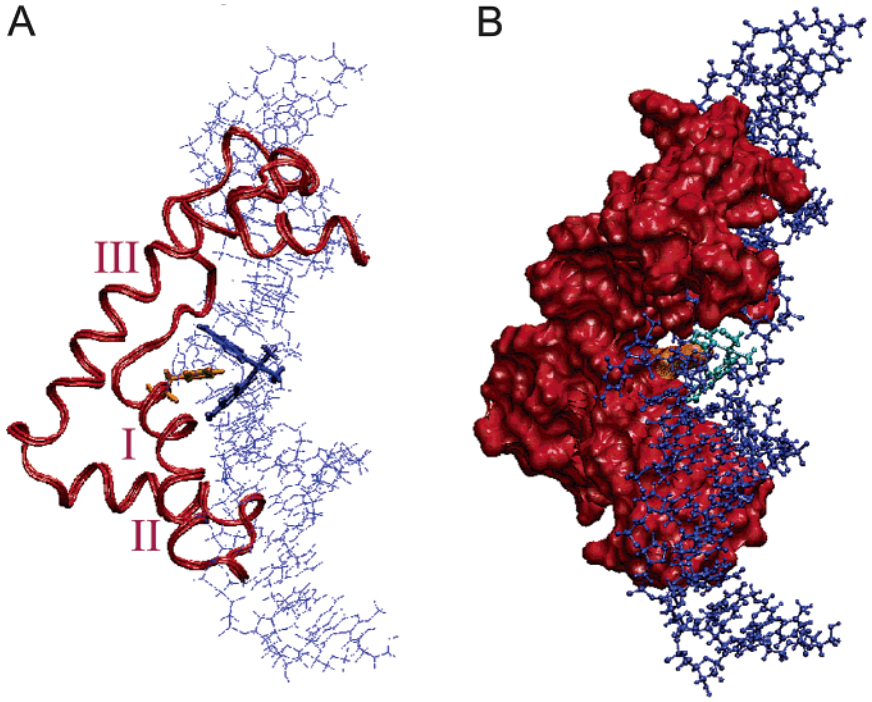


FIGURE 5: Two representations of the time-averaged structure of the complex formed between the domain HMGB1b7 and DNA containing the interstrand CL of cisplatin calculated over the last 480 ps of the MD simulation and energy-minimized. (A) The protein part is shown in a ribbon representation (in red), whereas the platinated DNA is displayed as a stick model (the top and bottom strands in pale and dark blue, respectively). (B) The protein part is shown as the Conolly surface (in red), whereas the platinated DNA is displayed as a stick-and-ball model (in blue). The DNA is oriented with the 3′-branch (comprising the 3′-end of the top strand and the 5′-end of the bottom strand) pointing upward in both figures. The helices I−III in the HMGB1b7 protein are indicated with Roman numbers.

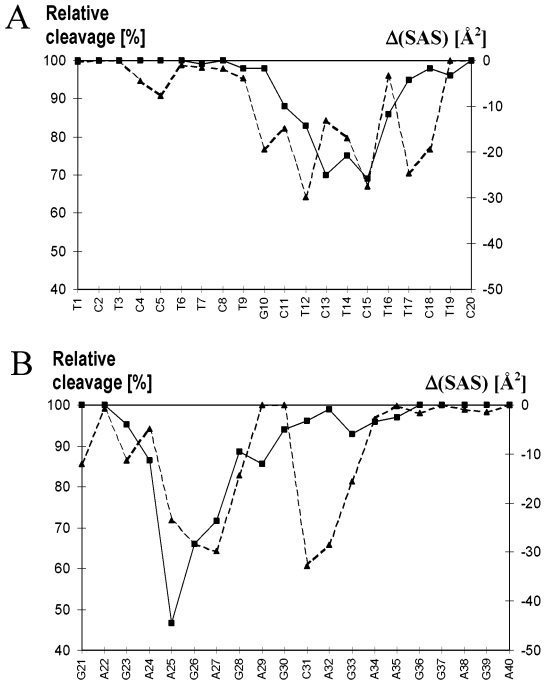


FIGURE 6: Comparison of the relative OH radical cleavage [% of the cleavage observed in the complex formed between the domain HMGB1b7 and DNA containing the interstrand CL of cisplatin (Figure 4) relative to free DNA (not bound in the complex with the protein) containing interstrand CL of cisplatin (squares)] and solvent-accessible surface (SAS) around all deoxyribose CH protons [difference between the complex and free DNA containing the interstrand CL (triangles)]. Top strand (A), bottom strand (B).

in hydroxyl radical footprinting experiments (Figure 6). The only site where the correlation is less satisfactory involves the extrahelical cytosine of the bottom strand of the duplex TGCT/AGCA(1) (C31) and its 3′-neighbor (A32) where the calculated SAS reduction seems overestimated. The small SAS calculated for these two nucleotides is due to the contacts of arginine R110 to the phosphate residues of the G30pC31pA32 fragment. It is possible that in the GB solvation model used in the present work, the individual hydration of the arginine and phosphate residues is underestimated and the direct contact exaggerated.

Another interesting correlation was observed between the reactivity of OH radicals and hydrogen bonding contacts. Figure 7, in which the H-bonding contacts of the time-averaged model (Figure 5) are summarized, indicates that these contacts are made mainly by the 3′-branch of the cross-linked duplex (see the text above for the definition of the 3′ and 5′ branches), where the nucleosides are protected from cleavage (Figure 4). A more sophisticated analysis consisted in counting the hydrogen bonding contacts that each nucleotide makes with the protein and averaged the counts over all snapshots of the MD trajectory. The average numbers of hydrogen bonds counted for a given nucleotide and its reactivity toward hydroxyl radicals are inversely correlated (Figure 8). This suggests that the hydrogen bonding contacts, although formed with phosphate and base residues (not with sugar residues), are the main forces keeping the amino acid side chains in the close proximity of sugar residues sufficient to protect these residues from being attacked by hydroxyl radicals. The van der Waals contacts between amino acid side chains and DNA bases in the major groove (which extend to both sides of the interstrand CL of cisplatin) do not apparently prevent hydroxyl radicals from reacting with the sugar residues. The inverse correlation between the number of hydrogen bonds and reactivity toward OH radicals is again less obvious at the pC31pA32 fragment, whose

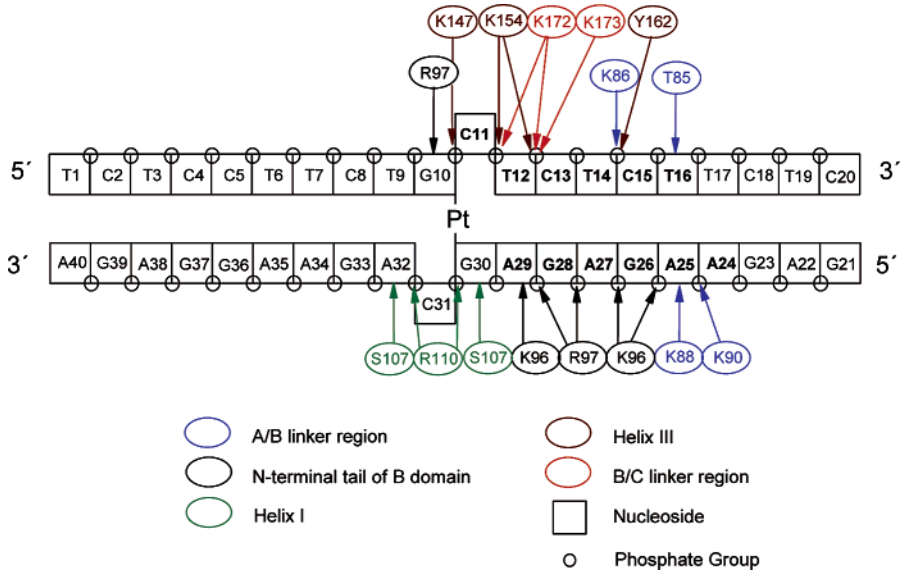


FIGURE 7: Schematic representation of the hydrogen bonding contacts between residues of the domain HMGB1b7 and DNA containing the interstrand CL of cisplatin in the structure shown in Figure 5. The nucleosides protected from OH radical cleavage (Figure 4) are labeled with bold letters.

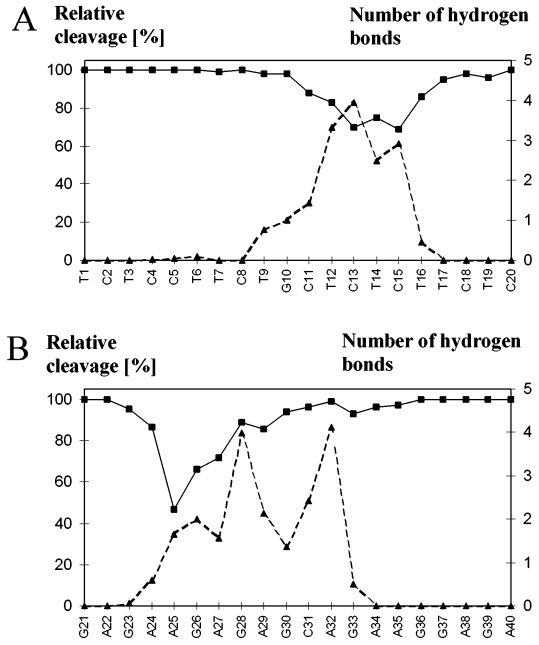


FIGURE 8: Comparison of the relative OH radical cleavage (squares) (for definition, see Figure 6) and the number of hydrogen bonds formed between the domain HMGB1b7 and DNA containing the interstrand CL of cisplatin (triangles) in their complex. Top strand (A), bottom strand (B).

interaction with arginine R110 is apparently overestimated by our simulation (vide supra).

Influence of the Base Sequence on Recognition. The top strand of the duplex containing the interstrand CL used in this study [TGCT/AGCA(1)] contains mainly pyrimidines, whereas the bottom strand is purine-rich. Thus, although the central interstrand CL-containing sequence 5′−d(GC)/5′−d(GC) is intrinsically centrosymmetric, the whole duplex is not. The observation that the hydroxyl radical footprint is highly asymmetrical (Figure 4) indicates that the protein is bound to DNA in one preferential orientation (relative to the site of the interstrand CL) (vide supra). Several sequence-specific hydrogen bonds between amino acid residues and DNA bases predicted by our model (Figure 7) may at least in part explain this selectivity in the orientation of the protein relative to the CL. These bonds include persistent hydrogen bonding between the ammonium group of lysine K88 and the N7 atom of adenine A25, which could also contribute to the strong inhibition of hydroxyl cleavage at this nucleotide.

DISCUSSION

We have shown in the present work that affinity of the full-length HMGB1 protein to the interstrand CL of cisplatin is only roughly 4 times lower than to the 1,2-GG intrastrand CL of this drug (Figure 2). Thus, we have confirmed our previous finding (*26*) that interstrand CLs of cisplatin are recognized by the full-length HMGB1 protein with an affinity which is not markedly different from the affinity of this protein to the major 1,2-GG intrastrand CL of this platinum drug. Importantly, the results of the present work and those reported previously (*26*) were obtained by using principally different DNA probes, which reinforces the validity of our finding. It has been reported recently (*28*) that the full-length HMGB1 protein does not bind the interstrand CL of cisplatin. This conclusion was deduced from the results of the experiments in which binding of the full-length HMGB1 protein to DNA interstrand cross-linked by cisplatin was examined using a short 18-bp oligodeoxyribonucleotide probe containing single, interstrand CL of cisplatin. It has been, however, demonstrated (*50*) that the minimum length of the cisplatin-modified DNA probe should be >60 bp. This probe size limitation very likely reflects a minimum requirement for a flanking DNA domain in order for the full-length HMGB1 protein binding to occur. Hence, it is not surprising that the full-length HMGB1 protein did not bind to the single interstrand CL of cisplatin contained in a duplex that was too short (*28*).

The most striking result of the present work is the observation that attachment of the lysine-rich region of the A/B linker to the domain HMGB1b results in its markedly enhanced affinity to the interstrand CL of cisplatin (Figures 3C,D and Table 2). Thus, the affinity to the interstrand CLs

of cisplatin of the domain HMGB1b to which residues ^{85}TKKKFKD^{91} of the A/B linker are attached (HMGB1b7 domain) is only slightly lower than that to 1,2-GG intrastrand CL of this drug (Table 2). This is in a deep contrast to the affinity of HMGB1a which binds with a high affinity to 1,2-GG intrastrand CL of cisplatin but does not bind to the interstrand CL (Table 2). Hence, the domain B of the HMGB1 protein with the A/B linker attached to its N-terminus should play a crucial role in the interaction between full-length HMGB1 protein and DNA interstrand cross-linked by cisplatin.

An eventuality raised by the reviewer that also the affinity of the domain A of HMGB1 (HMGB1a) to the interstrand CL of cisplatin could be enhanced by attaching the lysine-rich part of the A/B linker to the C-terminus of this domain deserves further discussion. The nonsequence specific HMG-boxes interact with DNA via the N-terminus (and consequently the binding affinity of the HMG-box is strengthen by the extension of this region with basic amino acid residues of the sequence TKKKFKD) and helices I and II, rather than via the helix III. For this reason, the attachment of a part of the linker region, TKKK sequence, to the C-terminus of helix III of HMGB1 domain A has only very little, if any, effect on binding of the domain A to 1,2-GG intrastrand CL (*24*) or supercoiled/linearized plasmid DNA (*24*, *25*). On the other hand, the attachment of the same linker sequence to the N-terminus of the HMGB1 domain B significantly increases affinity of this HMG-box for 1,2-GG intrastrand cross-link (*24*) and promotes the ability of the HMGB1 domain B to efficiently bend DNA (*36*, *51*). Thus, these data strongly support the view that affection of the affinity of HMGB1 domain A (HMGB1a) to DNA containing interstrand CL of cisplatin by attaching the TKKK sequence to the C-terminus of this domain is unlikely. Nevertheless, to further support the latter view, the affinity of the protein HMGB1a to which a short sequence of TKKKF was attached to its C-terminus (i.e., the protein HMGB1a′, the residues 1−89, Table 1) was examined (Figure 3E), and no binding of this protein to DNA probe containing the interstrand CL of cisplatin under conditions when this protein readily binds to 1,2-GG intrastrand CL of this drug was observed (Table 2).

The specific recognition of DNA duplex containing cisplatin interstrand CL is unprecedented since to date, HMG box proteins were found to recognize DNA duplexes bent in the direction of the major groove, making contacts with the groups in the minor groove. HMG-domain proteins are known to recognize kinked DNA through affinity to the convex side of the kink with the interior of the L-shaped structure formed by the three α helices. Hence, our observation that the full-length HMGB1 or HMGB1b7 proteins bind to DNA containing the interstrand CL of cisplatin is the first case when it seems reasonable to suggest that an HMG-domain protein recognizes DNA by making the contacts with its major groove. Using molecular modeling techniques, we have fitted the kinked, DNA duplex TGCT/AGCA(1) interstrand cross-linked by cisplatin with its convex side into the interior of the L-shaped HMGB1b7 and found that intercalation of phenylalanine F103 into the cleft formed by the two platinum-bound guanines markedly improves the fit. Only one of the two possible orientations of the DNA duplex with respect to the protein yielded, after relaxation, a model which was in agreement with the experimental hydroxyl radical footprint. This agreement was not only qualitative but also quantitative, as shown by the correlation between accessible areas around protected sugar protons and cleavage inhibition (Figure 6).

The fact that the molecular modeling yielded a model accounting for the experimental footprint does not prove, of course, that this is the only possible structure of the complex formed between DNA interstrand cross-linked by cisplatin and HMGB1b7 protein. However, the model does show that a stable complex could be formed without major structural changes of either partner. Moreover, the assumption used for its construction (fitting the convex side of the kinked DNA into the interior of the L-shaped protein) is plausible, and the resulting model shows all features of a physically reasonable structure: good surface complementarity, excluded SAS typical for complexes formed between HMG-domain proteins and DNA, and intercalation of a hydrophobic residue of the protein at the site of the kink in DNA, which is a classical recognition element used by HMG-domain proteins. Finally, the model suggests a rationale for the stabilizing role of the A/B linker in HMGB1 protein when the complex between this protein and DNA interstrand cross-linked by cisplatin is formed.

HMGB1 and HMGB2 proteins have been implicated in activation and repression of transcription (*18*). The capability of HMG-domain proteins to induce bends in DNA has been proposed to play a role in the assembly of transcription factor complexes: the bend could bring two transcription factors bound to initially remote sites of DNA to a close proximity and make their direct interaction possible. So far, bending toward the major groove has been considered as the unique DNA-structuring property of HMG-box proteins. Although the interstrand CL caused by cisplatin represents a very particular structure, our results suggest that the palette of DNA-binding motifs used by HMG-domain proteins may be broader than so far assumed and may include bending toward both the major and the minor grooves.

ACKNOWLEDGMENT

The authors acknowledge that their participation in the EC COST Chemistry Actions D20 and D21 enabled them to exchange regularly the most recent ideas in the field of platinum anticancer drugs with several European colleagues. J. Kasparkova is the international research scholar of the Howard Hughes Medical Institute. The research of J. Kasparkova and V. Brabec was also supported in part by the Wellcome Trust (U.K.). M.-A. Elionzo-Riojas was a recipient of the 1997 Gemini Award from the International Precions Metals Institute. Computer time from the IDRIS computer center of the Centre National de la Recherche Scientifique is gratefully acknowledged.

SUPPORTING INFORMATION AVAILABLE

Table S1: Detailed protocols of the MD simulations. Table S2: Atomic coordinates of the averaged structure presented in Figure 5 (pdb format). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Johnson, N. P., Butour, J.-L., Villani, G., Wimmer, F. L., Defais, M., Pierson, V., and Brabec, V. (1989) *Prog. Clin. Biochem. Med. 10*, 1−24.

2. Reedijk, J., and Teuben, J. M. (1999) Platinum−sulfur Interactions Involved in Antitumor Drugs, Rescue Agents, and Biomolecules, in *Cisplatin. Chemistry and Biochemistry of a Leading Anticancer Drug* (Lippert, B., Ed.) pp 339−362, Wiley-VCH, Zürich.
3. Brabec, V. (2002) *Prog. Nucleic Acid Res. Mol. Biol. 71*, 1−68.
4. 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.
5. Eastman, A. (1987) *Pharmacol. Ther. 34*, 155−166.
6. Lemaire, M. A., Schwartz, A., Rahmouni, A. R., and Leng, M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1982−1985.
7. Brabec, V., and Leng, M. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5345−5349.
8. Brabec, V. (2000) Chemistry and Structural Biology of 1,2-Interstrand Adducts of Cisplatin, in *Platinum-Based Drugs in Cancer Therapy* (Kelland, L. R., and Farrell, N. P., Eds.) pp 37−61, Humana Press Inc, Totowa, NJ.
9. van de Vaart, P. J. M., Belderbos, J., de Jong, D., Sneeuw, K. C. A., Majoor, D., Bartelink, H., and Begg, A. C. (2000) *Int. J. Cancer 89*, 160−166.
10. Leng, M., and Brabec, V. (1994) DNA Adducts of Cisplatin, Transplatin and Platinum-intercalating Drugs, in *DNA Adducts: Identification and Biological Significance* (Hemminki, K., Dipple, A., Shuker, D. E. G., Kadlubar, F. F., Segerbäck, D., and Bartsch, H., Eds.) pp 339−348, International Agency for Research on Cancer, Lyon, France.
11. Vrana, O., Boudny, V., and Brabec, V. (1996) *Nucleic Acids Res. 24*, 3918−3925.
12. Huang, H. F., Zhu, L. M., Reid, B. R., Drobny, G. P., and Hopkins, P. B. (1995) *Science 270*, 1842−1845.
13. Paquet, F., Perez, C., Leng, M., Lancelot, G., and Malinge, J. M. (1996) *J. Biomol. Struct. Dyn. 14*, 67−77.
14. Coste, F., Malinge, J. M., Serre, L., Shepard, W., Roth, M., Leng, M., and Zelwer, C. (1999) *Nucleic Acids Res. 27*, 1837−1846.
15. Zamble, D. B., and Lippard, S. J. (1999) Cisplatin and DNA Repair in Cancer Chemotherapy, in *Cisplatin. Chemistry and Biochemistry of a Leading Anticancer Drug* (Lippert, B., Ed.) pp 73−110, Wiley-VCH, Zürich.
16. Kartalou, M., and Essigmann, J. M. (2001) *Mutat. Res. 478*, 1−21.
17. Bustin, M., and Reeves, R. (1996) *Prog. Nucleic Acid Res. Mol. Biol. 54*, 35−100.
18. Thomas, J. O., and Travers, A. A. (2001) *TIBS 26*, 167−174.
19. He, Q., Ohndorf, U.-A., and Lippard, S. J. (2000) *Biochemistry 39*, 14426−14435.
20. Jamieson, E. R., and Lippard, S. J. (2000) *Biochemistry 39*, 8426−8438.
21. Cohen, S. M., Mikata, Y., He, Q., and Lippard, S. J. (2000) *Biochemistry 39*, 11771−11776.
22. Dunham, S. U., and Lippard, S. J. (1997) *Biochemistry 36*, 11428−11436.
23. Ohndorf, U. M., Rould, M. A., He, Q., Pabo, C. O., and Lippard, S. J. (1999) *Nature 399*, 708−712.
24. Stros, M. (2001) *Biochemistry 40*, 4769−4779.
25. Saito, K., Kikuchi, T., Shirakawa, H., and Yoshida, M. (1999) *J. Biochem. 125*, 399−405.
26. Kasparkova, J., and Brabec, V. (1995) *Biochemistry 34*, 12379−12387.
27. Turchi, J. J., Li, M., and Henkels, K. M. (1996) *Biochemistry 35*, 2992−3000.
28. Leng, M., Locker, D., Giraud-Panis, M. J., Schwartz, A., Intini, F. P., Natile, G., Pisano, C., Boccarelli, A., Giordano, D., and Coluccia, M. (2000) *Mol. Pharmacol. 58*, 1525−1535.
29. Mikata, Y., He, Q., and Lippard, S. J. (2001) *Biochemistry 40*, 7533−7541.
30. Brabec, V., Reedijk, J., and Leng, M. (1992) *Biochemistry 31*, 12397−12402.
31. Kasparkova, J., Farrell, N., and Brabec, V. (2000) *J. Biol. Chem. 275*, 15789−15798.
32. Bellon, S. F., Coleman, J. H., and Lippard, S. J. (1991) *Biochemistry 30*, 8026−8035.
33. Brabec, V., Neplechova, K., Kasparkova, J., and Farrell, N. (2000) *J. Biol. Inorg. Chem. 5*, 364−368.
34. Matsunaga, T., Mu, D., Park, C.-H., Reardon, J. T., and Sancar, A. (1995) *J. Biol. Chem. 270*, 20862−20869.
35. Buschta-Hedayat, N., Buterin, T., Hess, M. T., Missura, M., and Naegeli, H. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 6090−6095.
36. Stros, M. (1998) *J. Biol. Chem. 273*, 10355−10361.
37. Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., III, Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M. J., Stanton, R. V., Cheng, A. L. V., J. J., Crowley, M., Tsui, V., Radmer, R. J., Duan, Y., Pitera, J., Massova, I., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (1999) *AMBER 6*, University of California, San Francisco.
38. Cheatham, T. E., Cieplak, P., and Kollman, P. A. (1999) *J. Biomol. Struct. Dyn. 16*, 845−862.
39. Elizondo-Riojas, M.-A., Gonnet, F., Augé-Barrere-Mazouat, P., Allain, F., Berges, J., Attias, R., Chottard, J.-C., and Kozelka, J. (1997) Molecular Modeling of Platinum Complexes with Oligonucleotides: Methodological Lessons and Structural Insights, in *Molecular Modeling and Dynamics of Bioinorganic System* (Banci, L., and Comba, P., Eds.) pp 131−160, Kluwer Academic Publishers, Dordrecht, The Netherlands.
40. Tsui, V., and Case, D. (2000) *J. Am. Chem. Soc. 122*, 2489−2498.
41. Lavery, R., and Sklenar, H. (1989) *J. Biomol. Struct. Dyn. 6*, 655−667.
42. 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.
43. Sanner, M. F., Sphenner, J.-C., and Olson, A. J. (1996) *Biopolymers 38*, 305−320.
44. Nadassy, K., Wodak, S. J., and Janin, J. (1999) *Biochemistry 38*, 1999−2017.
45. Tullius, T. D., and Dombroski, B. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5469−5473.
46. Lee, S., and Hahn, S. (1995) *Nature 376*, 609−612.
47. Pogozelski, W. K., McNeese, T. J., and Tullius, T. D. (1995) *J. Am. Chem. Soc. 117*, 6428−6433.
48. Balasubramanian, B., Pogozelski, W. K., and Tullius, T. D. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 9738−9743.
49. Pastor, N., Weinstein, H., Jamison, E., and Brenowitz, M. (2000) *J. Mol. Biol. 304*, 55−68.
50. Chow, C. S., Barnes, C. M., and Lippard, S. J. (1995) *Biochemistry 34*, 2956−2964.
51. 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.

BI026695T