BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Recognition of Platinum–DNA Adducts by HMGB1a

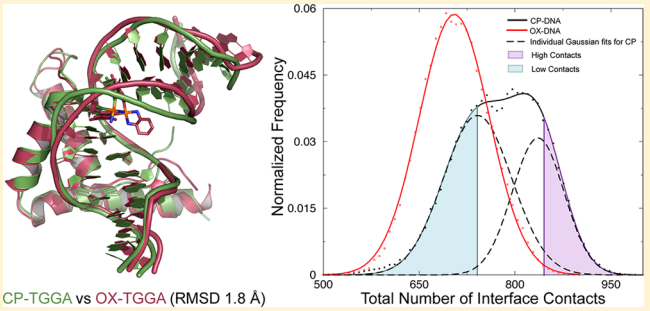
Srinivas Ramachandran,[†] Brenda Temple,[‡] Anastassia N. Alexandrova,[§] Stephen G. Chaney,*[,†] and Nikolay V. Dokholyan*[,†]

[†]Department of Biochemistry and Biophysics and [‡]R. L. Juliano Structural Bioinformatics Core Facility, University of North Carolina, Chapel Hill, North Carolina 27599, United States

[§]Department Chemistry and Biochemistry, University of California at Los Angeles, Los Angeles, California 90095-1569, United States

Supporting Information

ABSTRACT: Cisplatin (CP) and oxaliplatin (OX), platinum-based drugs used widely in chemotherapy, form adducts on intrastrand guanines (5′GG) in genomic DNA. DNA damage recognition proteins, transcription factors, mismatch repair proteins, and DNA polymerases discriminate between CP- and OX-GG DNA adducts, which could partly account for differences in the efficacy, toxicity, and mutagenicity of CP and OX. In addition, differential recognition of CP- and OX-GG adducts is highly dependent on the sequence context of the Pt-GG adduct. In particular, DNA binding protein domain HMGB1a binds to CP-GG DNA adducts with up to 53-fold greater affinity than to OX-GG adducts in the TGGA sequence context but shows much smaller differences in binding in the AGGC or TGGT sequence contexts. Here, simulations of the HMGB1a-Pt-DNA complex in the three sequence contexts revealed a higher number of interface contacts for the CP-DNA complex in the TGGA sequence context than in the OX-DNA complex. However, the number of interface contacts was similar in the TGGT and AGGC sequence contexts. The higher number of interface contacts in the CP-TGGA sequence context corresponded to a larger roll of the Pt-GG base pair step. Furthermore, geometric analysis of stacking of phenylalanine 37 in HMGB1a (Phe37) with the platinated guanines revealed more favorable stacking modes correlated with a larger roll of the Pt-GG base pair step in the TGGA sequence context. These data are consistent with our previous molecular dynamics simulations showing that the CP-TGGA complex was able to sample larger roll angles than the OX-TGGA complex or either CP- or OX-DNA complexes in the AGGC or TGGT sequences. We infer that the high binding affinity of HMGB1a for CP-TGGA is due to the greater flexibility of CP-TGGA compared to OX-TGGA and other Pt-DNA adducts. This increased flexibility is reflected in the ability of CP-TGGA to sample larger roll angles, which allows for a higher number of interface contacts between the Pt-DNA adduct and HMGB1a.



Cisplatin (CP) and oxaliplatin (OX) are platinum-based drugs used widely in the treatment of various cancers. The main mode of action of Pt drugs is the formation of covalent adducts on genomic DNA. The most abundant such adduct is formed on adjacent guanines. CP and OX form similar adducts on similar sites of genomic DNA, with the only difference being their carrier ligand: CP has diammine as the carrier ligand, while OX has diamminocyclohexane. However, CP and OX have major differences in both cellular efficacies and molecular recognition by different proteins.

Many proteins discriminate between CP- and OX-DNA adducts, but the most significant differences are observed for HMG domain proteins like HMGB1, LEF-1, TBP, and hUBF.[1−3] Consequently, the differential binding affinities of HMG domain proteins for CP- and OX-DNA adducts have been proposed to contribute to the distinctive cellular effects of CP- and OX-DNA adducts.[4] For example, HMGB1 has been shown to inhibit translesion synthesis past CP adducts to a greater extent than past OX adducts.[5] Furthermore, the level of HMGB1 expression is elevated in many cancer types, including colorectal cancers,[6] which frequently display resistance to CP but not OX.[7,8] Hence, understanding the basis of the differential binding of HMG domain proteins to CP- and OX-DNA could lead to a better understanding of the differential efficacies of CP and OX.

HMGB1 is the best-characterized HMG domain protein in terms of binding Pt-DNA adducts. Even though HMGB1 contains two domains (HMGB1a and HMGB1b), footprinting studies of full-length HMGB1 have indicated that only HMGB1a binds to Pt-GG adducts.[9] Hence, studies conducted with isolated HMGB1a correspond well with those with full-length HMGB1. The differential binding affinity of HMGB1a for CP- and OX-DNA is also dependent on the identity of bases flanking the central intrastrand GG.[1,2] The binding affinity of HMGB1a for CP-DNA adducts has been shown to

Received: June 26, 2012
Revised: August 29, 2012
Published: September 5, 2012

be ~53 times greater than that for OX-DNA adducts in the TGGA sequence context,[1] while in the AGGC[1] and the CGGA[2] sequence contexts, the binding affinity for CP-DNA is only 3 times greater than that for OX-DNA adducts. In the TGGT sequence context, the binding affinity of HMGB1a for CP-GG and OX-GG adducts is almost the same.[2] This sequence-dependent effect has also been shown for HMGB1b and TBP, but the trends are slightly different.[1]

The structure of HMGB1a bound to CP-DNA in the TGGA sequence context has been determined and suggests the basis for specific recognition of Pt-DNA adducts by HMGB1a. The structures of free Pt-DNA adducts show extensive distortion of the minor groove and significant bending of DNA due to Pt adduct formation. The binding of HMGB1a distorts the Pt-DNA adduct further, forming a highly complementary binding interface on the minor groove with extensive hydrophobic and electrostatic interactions. Phe37 of HMGB1a specifically recognizes the GG base pair step distorted by the formation of the Pt adduct and forms energetically favorable $\pi-\pi$ stacking interactions with them. Mutation of Phe37 to alanine results in a >667-fold decrease in binding affinity.[10] Several other structures of HMG domain proteins bound to undamaged DNA reveal similar intercalation, but involving different residues (methionine in LEF-1[11] and isoleucine in SRY,[12] for example). The crystal structure of HMGB1a bound to the Pt-DNA adduct in the TGGA sequence context[13] also shows a hydrogen bond between Ser41 and the N3 atom of adenine 3′ of the Pt-GG adduct. An S41A mutation resulted in an only 4-fold decrease in binding affinity, an effect much weaker than that of the Phe37Ala mutation.[10]

Previously, we postulated that the differences in conformational dynamics of CP- and OX-DNA adducts could lead to the differential recognition of these adducts by HMG domain proteins. We tested this hypothesis by performing all-atom molecular dynamic simulations of CP- and OX-DNA adducts in the AGGC,[14] TGGA,[15] and TGGT[16] sequence contexts. These data showed that in the TGGA sequence context, the CP-GG adduct was more flexible than the OX-GG adduct and was, therefore, able to explore greater bend angles and greater rolls for the central Pt-GG base pair step, which would favor Phe37 $\pi-\pi$ stacking.[15] While the CP-GG adduct was more flexible than the OX-GG adduct in the AGGC sequence context also, only subtle differences were observed in the minor conformations explored by the CP-GG adduct that might favor HMGB1a binding.[14] As might be expected, no significant differences in either conformation or conformational dynamics were observed between the CP-GG and OX-GG adducts in the TGGT sequence context.[16] These studies explored only whether the conformations sampled by CP- and OX-GG DNA adducts in solution might favor binding to HMGB1a. In this study, we directly address the basis of the high differential binding affinity of HMGB1a for CP- and OX-DNA in the TGGA sequence context compared to TGGT and AGGC sequence contexts by performing all-atom molecular dynamic simulations of the Pt-DNA-HMGB1a complex in all three sequence contexts. These data show that the CP-TGGA-HMGB1a complex is uniquely able to sample larger roll angles that favor high-affinity binding between the Pt-DNA adduct and HMGB1a.

MATERIALS AND METHODS

Pt-DNA-HMGB1a Simulations. We used the crystal structure of rat HMGB1a bound to CP-DNA in the TGGA sequence context as the starting structure for the CP-TGGA simulations [Protein Data Bank (PDB) entry 1CKT[13]]. With this structure, Insight II (Accelrys Inc.) was used to modify the identity of the flanking bases (T and A on the 5′ side and A, C, and T on the 3′ side of the adduct) and the type of adduct (cisplatin or oxaliplatin) to generate starting structures for OX-TGGA-, CP-AGGC-, OX-AGGC-, CP-TGGT-, and OX-TGGT-HMGB1a simulations. The complete sequences of the DNA parts of these complexes are shown in Figure S1 of the Supporting Information. The force field parameters used for DNA and protein were from the force field of Duan et al.,[17,18] while the ff03 parameter set and the partial charges for the Pt-GG adducts were taken from Sharma et al.[14] Besides the atomic partial charges, other force field parameters of the Pt-GG adducts were referenced from AMBER parm99 force field parameters or from previous work by Yao et al.[19] and Scheeff et al.[20] For each of the Pt-DNA-HMGB1a complexes, we performed three sets of 50 ns simulations. The starting structures were similar for each of the three simulations, but the initial velocities were randomized. The simulation system was prepared by first removing the hydrogen atoms of the starting structures in Insight II. The structures were then prepared with the LEaP module of AMBER version 9.0 as follows. First, hydrogen atoms were added back according to the nucleotide templates in the AMBER force field library. Second, the structures were neutralized with Na^+ ions. Third, the neutralized system was fully solvated in an octahedral water box using TIP3P model water molecules. The distance between the wall of the periodic box and the closest atom in the solute was set to 12.5 Å.

In the molecular dynamics (MD) simulations, the nonbonded cutoff was set to be 9.0 Å and the nonbonded list was updated every 10 steps. The MD simulations were always conducted under *NPT* conditions [constant pressure, using isotropic position scaling, at 1 atm (1 atm = 6.9 kPa), pressure relaxation time constants of 0.2 ps in relaxation MD and 2 ps in production MD; constant temperature, using weak-coupling algorithm, at 300 K, heat bath coupling time constants of 0.2 ps in relaxation MD and 1 ps in production MD] with a 1 fs time step. The SHAKE algorithm was applied to all bonds involving hydrogen atoms and in turn the regular force evaluation omitted for those bonds. The translational center-of-mass motion was removed every 1 ps. Every time the system was heated, no velocity information was inherited from the previous stage and a Maxwell distribution of velocities was re-established. In minimization and relaxation, a harmonic potential was applied to Pt-DNA adduct or DNA to restrain its motion. The restraint was gradually weakened and became zero at the final stage of relaxation. The following 50 ns production MD was conducted unrestrained.

Simulation Protocol. The whole system underwent a 140 ps minimization and relaxation before a 50 ns production MD conducted by the PMEMD module of AMBER version 9.0.[21] First, the whole system, including the Pt-DNA adduct, protein, water molecules, and counterions, was minimized for 5000 steps using the steepest descent method in a constant volume. Second, the system was heated from 0 to 300 K under *NVT* (constant volume and constant temperature) conditions in a 20 ps relaxation MD simulation. Third, the system was further relaxed in a 100 ps MD simulation under *NPT* conditions. Finally, the system was heated from 200 to 300 K under *NVT* conditions for a 20 ps run before beginning the 50 ns production MD simulation.

dx.doi.org/10.1021/bi3008577 | *Biochemistry* 2012, 51, 7608−7617

Analysis Protocols: Ser41 Hydrogen Bond. All the trajectories were analyzed for the presence of a hydrogen bond between the hydroxyl group of Ser41 in HMGB1a and all possible donors and acceptors based on a distance cutoff of 3.5 Å between the donor and acceptor and an angular cutoff of 135° among the donor, hydrogen, and acceptor. For selected hydrogen bonds, we computed the hydrogen bond energy using an orientation-dependent potential.[22]

Centroid Structures. To identify the centroid structures from the complete ensembles of CP- and OX-DNA-HMGB1a complexes in all three sequence contexts, we first selected 1000 snapshots equally spaced from the 5−50 ns stretch of the three independent simulations for each Pt-DNA-HMGB1a complex. Pairwise root-mean-square deviations (rmsds) (see below) among all 1000 structures of a given Pt-DNA-HMGB1a complex were calculated, and the centroid structure for a given ensemble was identified as the one with the lowest mean rmsd from all other structures.

Comparisons of Structures Using the Root-Mean-Square Deviation. We used the rmsd to compare structures. In these comparisons, we utilized only the Cα atoms of HMGB1a and phosphorus atoms from the DNA part of the complex to calculate the rmsd.

Analysis of the Protein−DNA Interface. We define an interface contact as a heavy atom from the protein and a heavy atom from the DNA within 5 Å of each other. Using this definition, we determine the total number of interface contacts in a given snapshot and build histograms of this measure across simulations. We also determine the contact map for a given ensemble. A contact map is a matrix in which each row is a nucleotide and each column is a protein residue. The elements of the matrix are the total number of contacts between each nucleotide (row) and each residue (column). These contact maps are represented as colors in a diagram (heat map). The range of the number of contacts and their corresponding color are shown as a scale at the bottom of the heat maps. Contact maps from two different ensembles are further compared by subtracting one from another, generating a difference contact map, which is also represented as a heat map.

Stacking Analysis. We define the mode of stacking through three geometric measures: the horizontal (r_1) and vertical (r_2) distances between the center of the stacking aromatic rings and the angle (θ) between the planes of the stacking aromatic rings (Figure S2 of the Supporting Information).

RESULTS

MD Simulations Feature Similar Protein−DNA Binding Modes. We performed three sets of simulations of 50 ns each of the complex between HMGB1a and the Pt-DNA adduct in three sequence contexts (TGGA, TGGT, and AGGC) with two drugs (CP and OX). With these simulations, our goal was to uncover the molecular basis of the ∼53-fold difference in binding affinity between CP- and OX-DNA adducts in the TGGA sequence context and also the sequence-dependent interactions of HMGB1a with the Pt-DNA adduct. We observed that these simulations showed uniform rmsd fluctuations after ∼5 ns (Figure S3 of the Supporting Information); hence, we utilized the final 45 ns of all the trajectories for various structural analyses. When we compare the centroid structure of the CP-DNA-HMGB1a complex in the TGGA sequence context to the published crystal structure of the HMGB1a-Pt-TGGA complex[13] (Figure 1), we observe

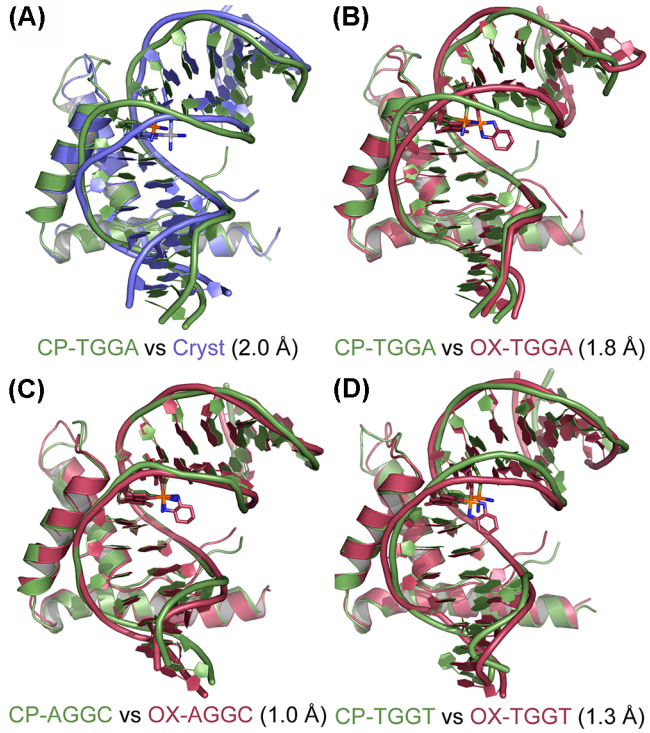


Figure 1. Comparison of the centroid structures. (A) Centroid structure from the CP-TGGA simulation aligned with the crystal structure of CP-TGGA using Cα atoms of the protein and phosphorus atoms of the DNA. (B−D) Similar structural alignments of centroid structures from the simulations of CP-TGGA and OX-TGGA (B), CP-AGGC and OX-AGGC (C), and CP-TGGT and OX-TGGT (D). The root-mean-square deviation for each pairwise comparison is shown in parentheses.

an rmsd of 2 Å, which illustrates an excellent congruence between the MD simulation and the crystal structure. Similarly, comparison of centroid structures of CP-DNA and OX-DNA adducts from the three sequence contexts reveals rmsds of <2 Å (Figure 1), indicating that our equilibrium simulations preserve the overall fold of the protein and the protein-Pt-DNA binding mode. The rmsds also indicate that the overall structures of the CP-DNA-HMGB1a and OX-DNA-HMGB1a complexes are similar in all three sequence contexts.

Analysis of Interface Contacts between the Pt-DNA Adduct and HMGB1a for CP- and OX-DNA Adducts in TGGA, TGGT, and AGGC Sequence Contexts. Previous structural[13] and site-directed mutagenesis[9,10] studies of the CP-TGGA-HMGB1a complex have implicated the roll of the central Pt-GG base pair step, Phe37 intercalation between central G residues, and the formation of hydrogen bonds between Ser41 and the base 3′ to the central Pt-GG adduct as being important for binding. To ensure an unbiased evaluation of the Pt-GG-HMGB1a interaction, we analyzed the total number of interface contacts between the Pt-DNA adduct and HMGB1a in all three sequence contexts (Figure 2). For each snapshot of the simulations, we calculated the number of atomic contacts formed between the DNA and protein, where an atomic contact is defined as two non-hydrogen atoms within 5 Å of each other. In structurally similar interfaces, the number of interface contacts can be used to compare the strength of the interfaces. Comparing CP- and OX-DNA adducts in different sequence contexts, we find that the TGGT and AGGC sequence contexts feature similar distributions of the number of

dx.doi.org/10.1021/bi3008577 | Biochemistry 2012, 51, 7608−7617

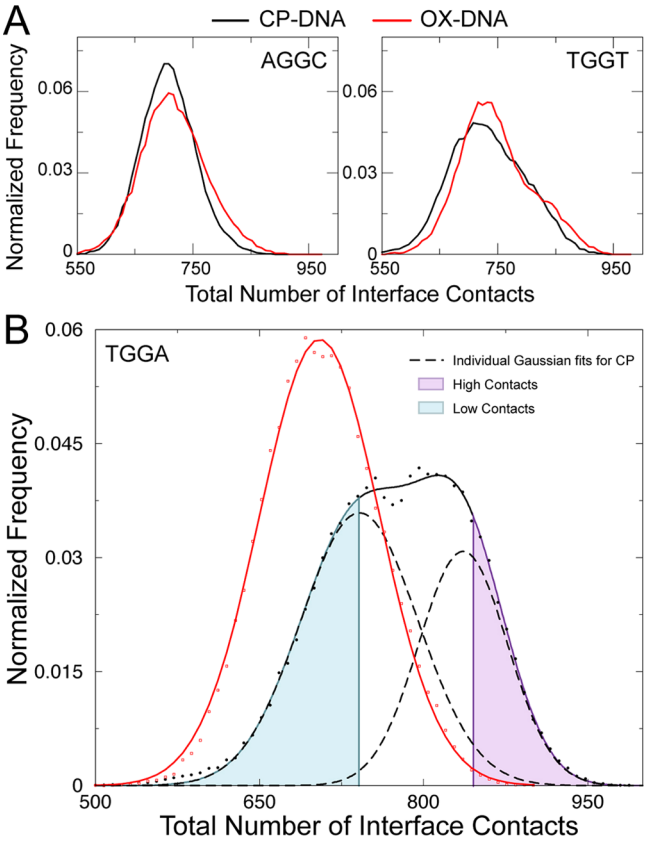


Figure 2. Distribution of the number of interface contacts between DNA and protein. (A) Distribution of the total number of interface contacts between HMGB1a and the Pt-DNA adduct plotted for CP- and OX-DNA in the AGGC and TGGT sequence contexts. (B) Distribution of the total number of interface contacts plotted for the HMGB1a-Pt-DNA complex for CP- and OX-DNA in the TGGA sequence context. We observe a bimodal distribution of interface contacts for the HMGB1a-CP-TGGA complex, and the two underlying Gaussian distributions are plotted with dashed lines. The shaded regions represent the ensembles with a high number of contacts and a low number of contacts.

protein−DNA interface contacts in the CP- and OX-DNA ensembles (Figure 2A). The TGGA sequence context, on the other hand, features a significantly higher number of interface contacts with HMGB1a in the CP-DNA ensemble than in the OX-DNA ensemble (Figure 2B). To examine the differences in the distributions of the number of interface contacts in the TGGA sequence context further, we deconvoluted the distributions assuming they form normal distributions. The distribution of the number of contacts with HMGB1a in OX-TGGA could be represented well with a single Gaussian distribution centered at 705 ± 77. However, the number of contacts with HMGB1a in CP-TGGA was best represented by two Gaussian distributions (χ^2 = 0.03 for a single Gaussian distribution, and χ^2 = 0.02 for two Gaussian distributions combined). These distributions were centered at 741 ± 75 and 836 ± 57 (Figure 2B). Thus, from the distribution of the number of contacts in the CP-TGGA-HMGB1a complex, we can discern two populations: one population with a distribution similar to that of the OX-TGGA-HMGB1a complex and the other with a distribution significantly right-shifted compared to that of the OX-TGGA-HMGB1a complex. The higher mean number of contacts in the CP-DNA adduct indicates a stronger complex with HMGB1a compared to that of the OX-DNA adduct. Thus, from the analysis of the distributions, we conclude that the ensembles of CP- and OX-DNA-HMGB1a complexes are very similar in the TGGT and AGGC sequence contexts, while in the TGGA sequence context, the CP-DNA-HMGB1a complex has a subpopulation that has much higher number of interface contacts than the OX-DNA-HMGB1a complex.

Buried Interface Surface Area. Parallel to the number of interface contacts, the solvent accessible surface area in the protein and DNA that is buried because of interface formation is also an indicator of interface strength upon comparison of similar protein−DNA complexes. The distribution of the buried surface area of the HMGB1a-Pt-DNA complexes follows a trend similar to that for interface contacts; CP-TGGA has larger buried interface area than OX-TGGA, while the distribution of buried interface area is similar between CP- and OX-DNA in other sequence contexts (Figure S4 of the Supporting Information).

Interface Contact Maps Reveal Regions in HMGB1a That Interact with Pt-DNA. To define the regions of the HMGB1a-Pt-DNA interface that feature differences in contacts when comparing different ensembles, we generated contact maps for each ensemble (Figure S5 of the Supporting Information). The contact map displays the mean number of heavy atom contacts formed between protein residues (represented on the *X*-axis) and the DNA bases (represented in the *Y*-axis) as a heat map. The higher the number of contacts between a given residue and a given base, the more the color tends toward red, while a smaller number of contacts tends to be toward blue. For the purpose of analyzing the DNA−protein interface, we divide the protein into its secondary structural segments [loop 1, L1; helix 1, H1; etc. (Figure S6 of the Supporting Information)]. From the contact map derived from the crystal structure, we observe that contacts are formed only between DNA and L1, H1, and H2, as described previously.[13] However, from the contact maps of the MD ensembles, we observe additional contacts formed between DNA and the N-terminal L1 and C-terminal H3 regions. These regions of HMGB1a were stabilized by crystal contacts in the experimental structure and, hence, were rigid and did not form contacts with DNA, even though they consist of several basic residues that would be predicted to form contacts with DNA. In the solution environment of the MD simulations, these regions are not constrained by adjacent crystal contacts and are free to interact with DNA, which results in a higher number of contacts with DNA compared to the number in the crystal structure.

Difference Contact Maps Reveal Regions of Ligand-Dependent Differences in Binding. Because all the ensembles featured similar contact maps, we calculated the difference in the mean frequency of contacts between each amino acid residue of HMGB1a and the DNA base for CP- and OX-DNA adducts in each sequence context and plotted those differences as difference contact maps (Figure 3). The locations of these protein−DNA interactions are shown in Figure S7 of the Supporting Information. When we compare the frequency of contacts between the MD simulation and the crystal structure of the HMGB1a CP-TGGA complex, the enhanced interactions between the L1 and H3 regions of HMGB1a and the Pt-DNA adduct are even more apparent. We also observe an enhanced interaction between Phe37 and the 3′-G residue of the central Pt-GG in the MD ensemble compared to the crystal structure. When we compare MD ensembles of CP- and OX-

dx.doi.org/10.1021/bi3008577 | *Biochemistry* 2012, 51, 7608−7617

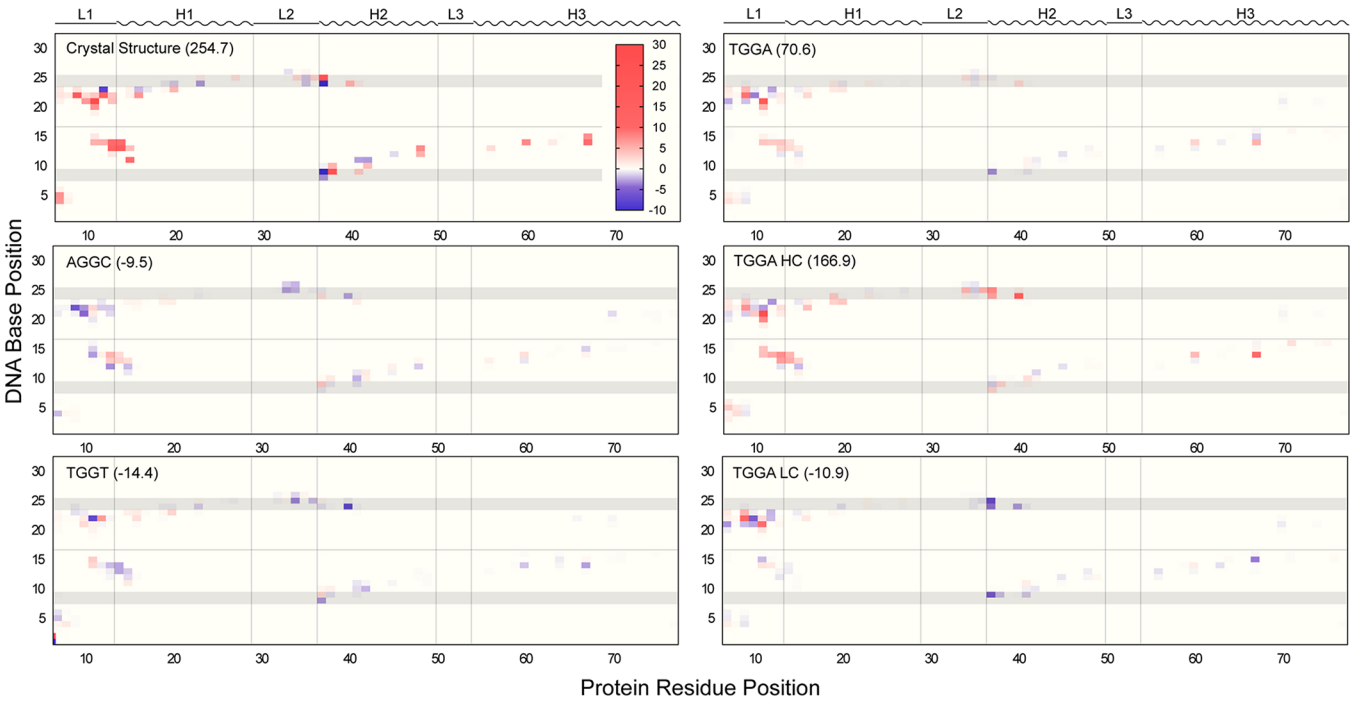


Figure 3. Difference contact maps of protein−DNA interactions. Difference contact maps (see Materials and Methods) are represented as heat maps for comparing different ensembles. The difference contact map obtained by subtracting the crystal structure contact map from the CP-DNA contact map is labeled crystal structure. Difference contact maps obtained by subtracting the OX-DNA contact map from the CP-DNA contact map for different sequence contexts are labeled AGGC, TGGT, and TGGA. Difference contact maps obtained by subtracting the OX-DNA contact map from high-contact and low-contact ensembles of CP-DNA in the TGGA sequence contexts are labeled TGGA HC and TGGA LC, respectively. The numbers in parentheses represent the differences in the total number of interface contacts between the ensembles being compared.

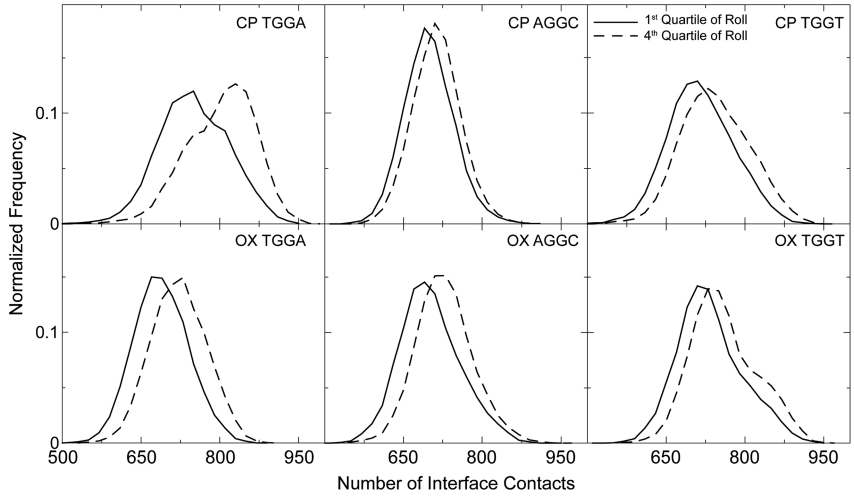


Figure 4. Effect of Pt-GG roll on the number of interface contacts. Distribution of total interface contacts of the structures corresponding to first and fourth quartiles of the Pt-GG roll distribution for CP- and OX-DNA in different sequence contexts.

DNA-HMGB1a complexes, we observe that the AGGC and TGGT sequence contexts feature similar frequencies of contacts. In contrast, significant differences are observed in the TGGA sequence context, reflecting the higher mean number of interface contacts in CP-DNA compared to OX-DNA in that sequence context. To distinguish the two populations in CP-TGGA (with respect to the number of interface contacts), we divided the CP-TGGA-HMGB1a ensemble into two extreme regions (shown shaded in Figure 2B), the "high-contact" structures (HC) that form the right half of the second Gaussian distribution and the "low-contact" structures (LC) that form the left half of the first Gaussian distribution, and then calculated the difference contact maps for both HC and LC CP-TGGA-HMGB1a versus OX-TGGA-HMGB1a. As expected, the differences between the binding of HMGB1a to CP-TGGA and OX-TGGA adducts were most pronounced for the HC CP-TGGA-HMGB1a complex. These differences were observed at L1 and H3 and around Phe37. An increased number of contacts with Phe37 and surrounding residues implies a higher extent of stacking of Phe37 with Pt-GG in the HC-CP-DNA-HMGB1a compared to the OX-DNA-HMGB1a complex. L1 and H3 represent the 3′ end of HMGB1A's binding interface with the Pt-DNA adduct and form the other end of the interface compared to the anchoring

dx.doi.org/10.1021/bi3008577 | *Biochemistry* 2012, 51, 7608−7617

Phe37 stacking. Thus, both ends of the protein−DNA interface feature an increased number of contacts, implying better complementarity and a tighter interface formed by the HC-CP-TGGA-HMGB1a complex compared to the OX-TGGA-HMGB1a complex. Because the only difference between the HC-CP-TGGA-HMGB1a complex and the OX-TGGA-HMGB1a complex is the Pt drug, our interface analysis points to the stronger ability of CP in the TGGA sequence context to mold to the HMGB1A interface compared to that of the OX-DNA adduct. This conclusion agrees well with our earlier observation that CP-DNA featured greater flexibility than OX-DNA in the TGGA sequence context.[15]

Pt-GG Roll Correlates with the Number of Interface Contacts. The roll of the Pt-GG base pair step was calculated with CURVES, version 5.3, for each of the MD simulations (Figure S8 of the Supporting Information). The roll was best described as a unimodal Gaussian distribution for each of the simulations. For both the CP-DNA-HMGB1a and the OX-DNA-HMGB1a complexes in the AGGC and TGGT sequences, the median roll angle for the P-GG base pair step was similar, and only slightly smaller than that observed in the crystal structure of the CP-TGGA-HMGB1a complex. In contrast, the roll of the OX-TGGA-HMGB1a complex was significantly smaller than that for the CP-TGGA-HMGB1a complex. We next asked if the higher number of interface contacts in CP-DNA compared to OX-DNA in the TGGA sequence context correlated with the increased flexibility of CP-DNA as evidenced by the increased roll of the Pt-GG base pair step. To test this hypothesis, we divided the distribution of the Pt-GG roll into quartiles and plotted the distribution of the number of protein−DNA interface contacts corresponding to structures in each of the quartiles of the Pt-GG roll distributions (Figure 4). We observe that for both the CP- and the OX-DNA-HMGB1a complexes, the mean number of interface contacts increases as we proceed from the structures in the lowest to the highest quartile of the Pt-GG roll. However, this effect is much more pronounced for CP-TGGA because of the presence of a much higher number of interface contacts when compared to the number for OX-TGGA. In the TGGT and AGGC sequence contexts, structures belonging to larger roll quartiles featured higher numbers of interface contacts, but this difference was not significant for the CP- and OX-DNA adducts. Thus, we conclude that the DNA flexibility as measured by the Pt-GG roll correlates with the number of interface contacts and that this effect is particularly strong for the CP-TGGA-HMGB1a complex. To further characterize this correlation, we performed the reverse analysis, i.e., if the HC ensemble features a much larger roll than the low-contact ensemble. We find that indeed the highest quartile of HMGB1a-Pt-DNA contacts corresponds to a larger Pt-GG roll angle than the lowest quartile of contacts in every sequence context, and that this effect is strongest for the HMGB1a-CP-TGGA complex because of the bimodal distribution of contacts (Figure S9 of the Supporting Information). While the overall distribution of the Pt-GG roll angle is similar for the HMGB1a-CP-DNA complex in all three sequence contexts and for the HMGB1a-OX-DNA complex in the AGGC and TGGT sequence contexts, the HMGB1a-CP-TGGA complex is unique in that a larger roll correlates with a quantitatively greater number of interface contacts than the number observed for any of the other HMGB1a-Pt-DNA complexes.

The Effect of Pt-GG Roll on Interface Contacts Is Not Driven Solely by Phe37 Stacking for the HMGB1a-CP-TGGA Complex. The increased number of protein−DNA contacts at larger roll angles could arise for two reasons; the overall interface could form more contacts, or the increased roll might facilitate an increased number of contacts between Phe37 and the Pt-GG base pair step. To differentiate these two effects, we plotted the distribution of the number of contacts of structures from different roll quartiles, after removing contacts formed by Phe37 (Figure S10 of the Supporting Information). Indeed, we observe that the distributions of the number of contacts for structures with small and large rolls are very similar in most cases, indicating that a larger roll mainly facilitates better contacts between Phe37 and the Pt-GG base pair step. The only exception is the CP-DNA-HMGB1a complex in the TGGA sequence context, which still features a significantly higher number of contacts in structures with a larger Pt-GG roll. These results again indicate that the cooperative effect of DNA flexibility on the overall protein−DNA interface is particularly strong in the case of CP-DNA in the TGGA sequence context.

Effect of Sequence Context on the Stacking of Phe37 with the Pt-GG Base Pair Step. An important interaction that drives the formation of a complex between HMGB1a and Pt-DNA is the stacking of the aromatic side chain of Phe37 with the aromatic bases of the Pt-GG base pair step.[10,13] Hence, we sought to analyze sequence- and carrier ligand-dependent stacking interactions. Energetic analysis of model aromatic molecules using quantum mechanical calculations[23,24] and analysis of stacking amino acids with each other in high-resolution protein crystal structures[25,26] point to two major modes of stacking: (i) the two aromatic rings parallel to each other, with their center slightly offset (parallel offset), and (ii) the two aromatic rings perpendicular to each other (perpendicular). However, in crystal structures of protein-DNA complexes, only parallel stacking of aromatic amino acids with DNA bases has been observed to date.[27] The modes of stacking can be geometrically described using three parameters: the horizontal (r_1) and vertical (r_2) distances between the center of the stacking aromatic rings and the angle (θ) between the planes of the stacking aromatic rings. Roughly, we define parallel stacking to have an r_1 of 1−2 Å, an r_2 of 3.5 Å, and a θ of <35°; for perpendicular stacking, r_1 = 1−2 Å, r_2 = 4.5−5.5 Å, and θ > 55°. Values of r_1, r_2, and θ outside these two ranges denote energetically less favorable packing. Thus, we have determined r_1, r_2, and θ values for Phe37 and both the 5′- and 3′-G residues of the central Pt-GG base pair step for the HMGB1a-CP-TGGA crystal structure in each of our MD simulations. The distribution of r_1 and r_2 at θ values of <35°, 35−55°, and >55° for the MD simulations of HMGB1a-CP-DNA and HMGB1a-OX-DNA complexes in all three sequence contexts is shown in Figure S11 of the Supporting Information for stacking with the 3′-G and Figure S12 of the Supporting Information for stacking with the 5′-G. From this analysis, it was clear that most of the structures with θ values of <35° represented parallel stacking and most of the structures with θ values of >55° represented perpendicular stacking. However, because of the distortion of the GG base pair step in the Pt-GG adduct, it is likely that perpendicular stacking of Phe37 with the 3′-G might correlate with parallel stacking of Phe37 with the 5′-G. Thus, we correlated the θ values for the 3′-G with the θ values for the 5′-G (Figure S13 of the Supporting Information). These data showed that, indeed, perpendicular stacking with one of the G residues corresponds to parallel stacking with the other G. Thus, 90% of the conformations explored during these

dx.doi.org/10.1021/bi3008577 | Biochemistry 2012, 51, 7608−7617

simulations corresponded to parallel stacking of Phe37 with one of the G residues in all three sequence contexts, and overall, there is no significant difference between Phe37 stacking for the CP- and OX-GG adducts in any of the sequence contexts (Figure 5A).

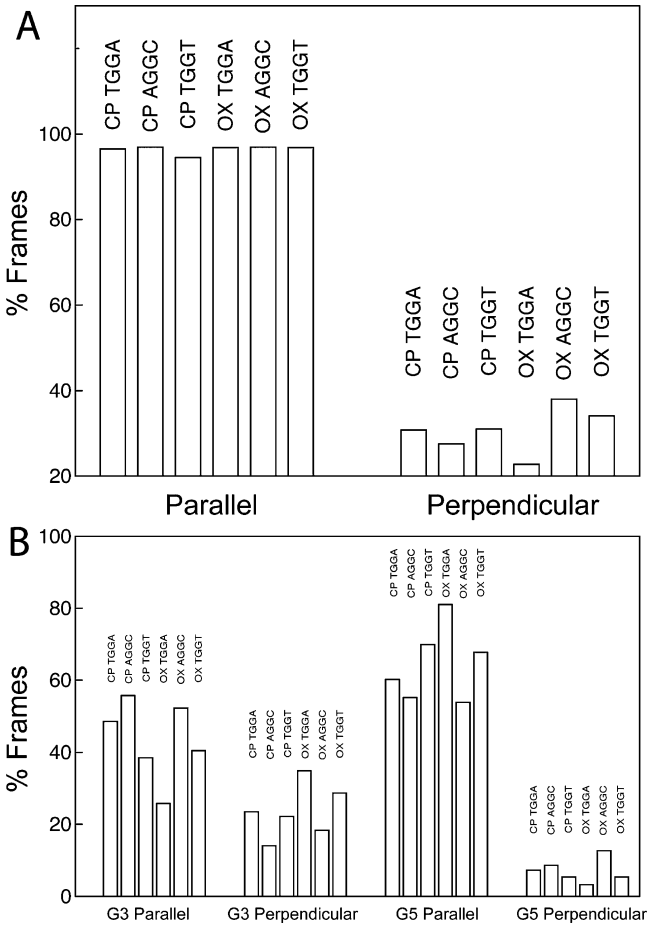


Figure 5. Prevalence of parallel and perpendicular stacking modes. (A) For CP- and OX-DNA ensembles in each of the sequence contexts, the percentage of structures that feature parallel and perpendicular stacking of Phe37 with either the 3′-G or 5′-G is plotted. (B) Similar to panel A, but broken down for the 5′-G and 3′-G.

To determine if different roll values of the Pt-GG base pair step are associated with different modes of stacking of Phe37 with Pt-GG, we plotted the distributions of r_1, r_2, and θ of Phe37 stacking with the 3′-guanine for structures corresponding to different quartiles of the roll distribution (Figure 6 and Figure S14 of the Supporting Information). In the TGGA sequence context (Figure 6), the Pt-GG roll correlates strongly with the mode of stacking of Phe37 with the 3′-G. In the first quartile (lowest roll), the stacking parameters peak at an r_1 of 2 Å, an r_2 of 4 Å, and a θ of ∼40°. These values correspond to a mode of stacking that is intermediate between parallel and perpendicular, which is not the most favorable mode. In the fourth quartile (largest roll), the stacking parameters peak at the following values: r_1 = 1.5 Å, r_2 = 3.5 and 4.5 Å, and $\theta \sim 20°$ and 65°. These values correspond to parallel and perpendicular modes of stacking that are energetically favorable. Furthermore, perpendicular stacking with the 3′-G corresponds to parallel stacking with the 5′-G, indicating energetically favorable stacking observed in structures in the fourth quartile of the roll distribution. We observe a similar effect in the AGGC sequence context, but to a lesser extent, while in the TGGT sequence context, we find the stacking modes to be similar in different quartiles of the Pt-GG roll distribution (Figure S14 of the Supporting Information).

Effect of Sequence Context on Ser41 Hydrogen Bonding. Because Ser41 of HMGB1a forms a hydrogen bond with the N3 atom of the adenine 3′ to Pt-GG adduct in the HMGB1a-CP-TGGA crystal structure[13] and mutagenesis experiments[9,10] suggest that this hydrogen bond has a minor (∼4-fold) effect on the affinity of HMGB1a for CP-DNA, we examined the ability of Ser41 to form hydrogen bonds with the DNA in each of our simulations. In our HMGB1a-Pt-DNA simulations, we observe hydrogen bonds between Ser41 and the base 3′ to the Pt-GG adduct regardless of the identity of the 3′ base or the Pt drug. In the AGGC and TGGT sequence contexts, the hydrogen bond is formed between Ser41 and the O2 atom of the 3′-C and 3′-T, respectively (Figure 7). When we calculated the energetic contribution of this hydrogen bond seen in our MD simulations using an orientation-dependent hydrogen bonding potential, we observe a mean hydrogen bond energy due to the Ser41−A hydrogen bond in the CP-TGGA-HMGB1a complex of 0.78 ± 0.003 kcal/mol (standard error of the mean), which compares well with the value of

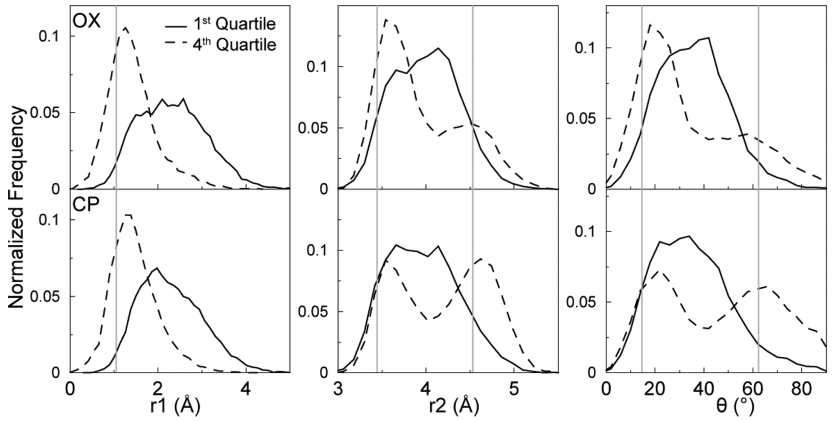


Figure 6. Effect of roll on stacking. The distributions of parameters used to describe stacking of Phe37 with the 3′-G of the platinated base pair of CP- and OX-DNA in the TGGA sequence context (r_1, r_2, and θ) are plotted for the structures belonging to the first and fourth quartiles of the Pt-GG roll distribution.

dx.doi.org/10.1021/bi3008577 | Biochemistry 2012, 51, 7608−7617

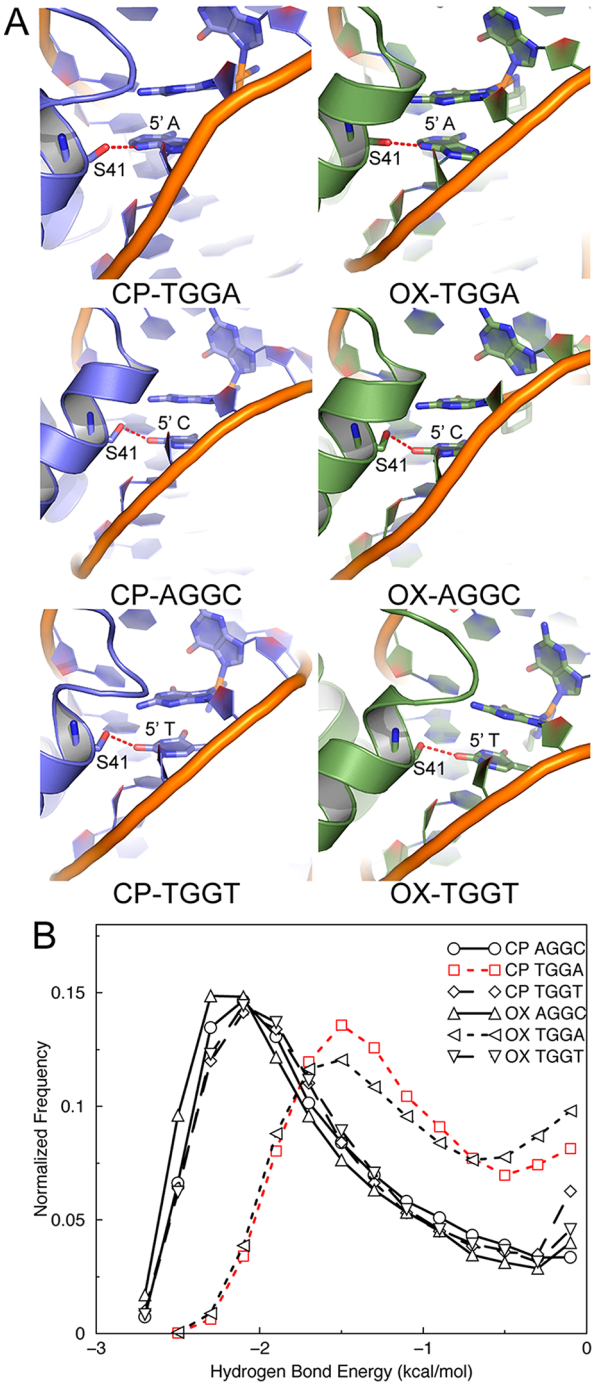


Figure 7. Hydrogen bond between Ser41 and the base 3′ to the Pt-GG base pair step. (A) Representative structures featuring the Ser41 hydrogen bond in each of the sequence context depict the hydroxyl oxygen of Ser41 and the electron donor connected by a dashed line. The structures are depicted using a cartoon representation with Ser41 and DNA bases also shown as sticks. (B) Distribution of hydrogen bond energies calculated for all the structures in the CP- and OX-DNA ensemble in the TGGA, AGGC, and TGGT sequence contexts.

0.832 kcal/mol observed experimentally: The S41A mutation causes a 4-fold increase in the binding dissociation constant. The OX-TGGA-HMGB1a complex features the same hydrogen bond with a similar hydrogen bond energy. However, the TGGT and AGGC sequence contexts feature pyrimidines on the 3′ side of the adduct, and hence an SP2-hybridized oxygen (O2). These oxygens form hydrogen bonds with Ser41 in our simulations, and further, they are almost twice as stabilizing compared to the TGGA sequence context. Thus, we observe Ser41 forming sequence-specific hydrogen bonds with the Pt-DNA adduct in all three sequence contexts. The strength of the hydrogen bonds formed appears to be greater in the TGGT and AGGC sequence contexts than in the TGGA sequence context. However, in all three sequence contexts, there are only small differences in hydrogen bond strength between CP- and OX-DNA adducts, indicating that Ser41 does not play a significant role in the ability of HMGB1a to discriminate between CP- and OX-DNA adducts.

DISCUSSION

Comparing the simulations of Pt-DNA-HMGB1a complexes and the free Pt-DNA simulations offers insight into how conformational dynamics of free DNA correlates with that of the protein-bound state. In the MD simulations of Pt-DNA-HMGB1a complexes, we observe all of the interactions between HMGB1a and the Pt-GG DNA adducts that were seen in the crystal structure.[13] In addition, we see interactions between the N-terminal tail and C-terminal helix of HMGB1a and the DNA minor groove that were not seen in the crystal structure. Those interactions are expected given the positively charged side chains in both regions of the protein but may not have been observed in the crystal structure because they were involved in crystal contacts.[13] In the full-length HMGB1 protein,[28] L1 and H3 are not constrained (L1 is flexible, and H3 is followed by a flexible linker). Hence, these regions could form contacts similar to what we observe for the isolated domain.

In our free Pt-DNA simulations in the TGGA sequence context, we observed increased flexibility of the CP-DNA adduct compared to that of the OX-DNA adduct as reflected in larger Pt-GG rolls and overall bend angles sampled by CP-DNA.[15] This increased flexibility in the free state translated to tighter complex formation in our simulations of the CP-TGGA-HMGB1a complex. We infer tighter complex formation based on the increased number of interface contacts and buried interface surface area between HMGB1a and CP-TGGA. As proposed previously,[13] the larger roll of the Pt-GG base pair step also correlated with more energetically favorable modes of stacking of Phe37 with Pt-GG to a great extent in the TGGA sequence context and to a lesser extent in the AGGC sequence context for both CP- and OX-DNA. However, a larger roll correlated with a higher number of interface contacts with residues other than Phe37 only for CP-DNA in the TGGA sequence context. These increased numbers of contacts are seen mainly in L1 and H3, which form contacts with the bases farthest from the Pt-GG adducts on the 3′ side (Figure S6 of the Supporting Information). Thus, the increased flexibility of CP-TGGA adduct correlates with strengthened interactions at both ends of the Pt-DNA-HMGB1a complex (stacking of Phe37 with Pt-GG at the 5′ end of the binding interface and interactions with L1 and H3 at the 3′ end of the interface). Of the three sequence contexts studied here, we observe significant differences between the CP- and OX-DNA-HMGB1a complexes in the TGGA sequence context and no major differences in the TGGT and AGGC sequence contexts, which agrees well with the reported ∼53-fold differences in the binding affinity in the TGGA sequence context,[1] the only ∼3-fold difference in the AGGC sequence context,[1] and no major difference in the TGGT sequence context.[2]

dx.doi.org/10.1021/bi3008577 | *Biochemistry* 2012, 51, 7608−7617

What is the basis of the increased flexibility of the CP-TGGA adduct? We speculate that the larger roll and higher flexibility observed in the CP-DNA adduct in the TGGA occurs due to a combination of sequence context and the identity of the carrier ligand. Because of DNA's right-handed helicity, the 3′ base is closer to the adduct than the 5′ base. Hence, the identity of the 3′ base has a greater bearing on the Pt-DNA adduct's flexibility. In terms of the carrier ligand, cisplatin is more flexible than oxaliplatin,[15] because in cisplatin, the Pt-ammines are not constricted by the cyclohexane moiety. Thus, a combination of sequence-dependent DNA flexibility on the 3′ side of the adduct and the flexibility of the carrier ligand results overall in an increased flexibility of the Pt-DNA adduct, which is also reflected in the increased level of binding by HMGB1a.

Phe37 stacks with Pt-GG in the crystal structure,[13] and the F37A mutation results in a >667-fold decrease in the affinity of HMGB1a for the Pt-DNA adduct. In the MD simulations, we observe stacking of Phe37 with both the 5′-G and the 3′-G. The mode of stacking observed in the crystal structure is parallel stacking of Phe37 with the 3′-G.[13] In the TGGA sequence context, we observe a significantly higher extent of 3′ parallel stacking in CP-DNA compared to OX-DNA, while in TGGT and AGGC sequence contexts, the percentage of frames in which Phe37 is stacking parallel to the 3′-G is similar in CP- and OX-DNA (Figure 5B). However, there is no theoretical reason to assume the Phe37 stacking is restricted to the 3′-G or that stacking with the 3′-G is energetically more favorable than stacking with the 5′-G. In the crystal structures of TATA-box binding protein complexed with DNA, phenylalanine stacking with both 5′ and 3′ bases is observed,[29] so we consider it likely that in solution, the HMGB1a-Pt-GG DNA complex will be able to explore Phe37 stacking with both G residues and that they would likely contribute equally to the binding affinity. Thus, via comparison of the stacking interactions involving both the 5′-G and the 3′-G in the Pt-DNA-HMB1a simulations, 90% of the conformations explored during the simulations stack with either the 3′- or 5′-G in all three sequence contexts, and there was no significant difference between stacking for the CP- and OX-GG adducts in any of the sequence contexts. Thus, from our simulations, we conclude that Phe37 stacking with the Pt-GG base pair does not contribute significantly to the ability of HMGB1a to discriminate between CP- and OX-DNA adducts.

The hydrogen bond between Ser41 and the base 3′ to the Pt-GG adduct is observed in the crystal structure, and the S41A mutation results in an ∼4-fold decrease in the affinity of HMGB1a for the Pt-DNA adduct. In the MD simulations, we observed the formation of a hydrogen bond from Ser41 with the base 3′ to the Pt-GG adduct in all three sequence contexts. In the AGGC and TGGT sequence contexts, the hydrogen bond is formed between Ser41 and the O2 atom of the 3′-C and 3′-T, respectively (Figure 7). However, no significant differences in the energy of hydrogen bond formation are observed between CP-DNA and OX-DNA adducts in any of the sequence contexts examined.

In conclusion, we infer that the high binding affinity of HMGB1a for CP-TGGA is due to the greater flexibility of CP-TGGA compared to those of OX-TGGA and other Pt-DNA adducts. This increased flexibility is reflected in the ability of CP-TGGA to sample larger roll angles than OX-TGGA, which allows for a higher number of interface contacts between the CP-TGGA adduct and HMGB1a. With two different DNA adducts and three different sequence contexts, we clearly demonstrate the effect of DNA sequence and carrier ligand on protein−DNA interactions. Most significantly, we show that for HMGB1a, it is DNA flexibility and not just specific interactions with bases on the 5′ and 3′ sides of the adduct that influences sequence-specific differences in binding affinity for CP- and OX-DNA adducts.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Fourteen figures supporting the results. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*N.V.D.: 3097 Genetic Medicine Building, Campus Box 7260, Chapel Hill, NC 27599; e-mail, dokh@unc.edu; phone, (919) 843-2513. S.G.C.: 3091 Genetic Medicine Building, Campus Box 7260, Chapel Hill, NC 27599; e-mail, sgc@med.unc.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the University of North Carolina Research Computing Group for providing and maintaining the computing resources used in this work.

■ REFERENCES

(1) Wei, M., Cohen, S. M., Silverman, A. P., and Lippard, S. J. (2001) Effects of spectator ligands on the specific recognition of intrastrand platinum-DNA cross-links by high mobility group box and TATA-binding proteins. *J. Biol. Chem. 276*, 38774−38780.

(2) Malina, J., Novakova, O., Vojtiskova, M., Natile, G., and Brabec, V. (2007) Conformation of DNA GG intrastrand cross-link of antitumor oxaliplatin and its enantiomeric analog. *Biophys. J. 93*, 3950−3962.

(3) Chvalova, K., Sari, M. A., Bombard, S., and Kozelka, J. (2008) LEF-1 recognition of platinated GG sequences within double-stranded DNA. Influence of flanking bases. *J. Inorg. Biochem. 102*, 242−250.

(4) Chaney, S. G., Campbell, S. L., Bassett, E., and Wu, Y. (2005) Recognition and processing of cisplatin- and oxaliplatin-DNA adducts. *Crit. Rev. Oncol. Hematol. 53*, L3−L11.

(5) Vaisman, A., Lim, S. E., Patrick, S. M., Copeland, W. C., Hinkle, D. C., Turchi, J. J., and Chaney, S. G. (1999) Effect of DNA polymerases and high mobility group protein 1 on the carrier ligand specificity for translesion synthesis past platinum-DNA adducts. *Biochemistry 38*, 11026−11039.

(6) Volp, K., Brezniceanu, M.-L., Bosser, S., Brabletz, T., Kirchner, T., Gottel, D., Joos, S., and Zornig, M. (2006) Increased expression of high mobility group box 1 (HMGB1) is associated with an elevated level of the antiapoptotic c-IAP2 protein in human colon carcinomas. *Gut 55*, 234−242.

(7) Raymond, E., Faivre, S., Woynarowski, J. M., and Chaney, S. G. (1998) Oxaliplatin-mechanism of action and antineoplastic activity. *Semin. Oncol. 25*, 4−12.

(8) Rixe, O., Ortuzar, W., Alvarez, M., Parker, R., Reed, E., Paull, K., and Fojo, T. (1996) Oxaliplatin, tetraplatin, cisplatin, and carboplatin: Spectrum of activity in drug-resistant cell lines and in the cell lines of the National Cancer Institute's Anticancer Drug Screen panel. *Biochem. Pharmacol. 52*, 1855−1865.

(9) Jung, Y., and Lippard, S. J. (2003) Nature of Full-Length HMGB1 Binding to Cisplatin-Modified DNA. *Biochemistry 42*, 2664−2671.

(10) He, Q., Ohndorf, U. M., and Lippard, S. J. (2000) Intercalating residues determine the mode of HMG1 domains A and B binding to cisplatin-modified DNA. *Biochemistry 39*, 14426−14435.

(11) Love, J. J., Li, X., Case, D. A., Giese, K., Grosschedl, R., and Wright, P. E. (1995) Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature 376*, 791−795.
(12) Werner, M. H., Huth, J. R., Gronenborn, A. M., and Clore, G. M. (1995) Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. *Cell 81*, 705−714.
(13) Ohndorf, U.-M., Rould, M. A., He, Q., Pabo, C. O., and Lippard, S. J. (1999) Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. *Nature 399*, 708−712.
(14) Sharma, S., Gong, P., Temple, B., Bhattacharyya, D., Dokholyan, N. V., and Chaney, S. G. (2007) Molecular dynamic simulations of cisplatin- and oxaliplatin-d(GG) intrastand cross-links reveal differences in their conformational dynamics. *J. Mol. Biol. 373*, 1123−1140.
(15) Ramachandran, S., Temple, B. R., Chaney, S. G., and Dokholyan, N. V. (2009) Structural basis for the sequence-dependent effects of platinum-DNA adducts. *Nucleic Acids Res. 37*, 2434−2448.
(16) Bhattacharyya, D., Ramachandran, S., Sharma, S., Pathmasiri, W., King, C. L., Baskerville-Abraham, I., Boysen, G., Swenberg, J. A., Campbell, S. L., Dokholyan, N. V., and Chaney, S. G. (2011) Flanking bases influence the nature of DNA distortion by platinum 1,2-intrastrand (GG) cross-links. *PLoS One 6*, e23582.
(17) Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem. 24*, 1999−2012.
(18) Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc. 117*, 5179−5197.
(19) Yao, S., Plastaras, J. P., and Marzilli, L. G. (1994) A molecular mechanics AMBER-type force field for modeling platinum complexes of guanine derivatives. *Inorg. Chem. 33*, 6061−6077.
(20) Scheeff, E. D., Briggs, J. M., and Howell, S. B. (1999) Molecular modeling of the intrastrand guanine-guanine DNA adducts produced by cisplatin and oxaliplatin. *Mol. Pharmacol. 56*, 633−643.
(21) Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Jr., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A. E., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C. E. A. F., Ross, W. S., and Kollman, P. A. (2006) *AMBER 9*, University of California, San Francisco.
(22) Ding, F., Tsao, D., Nie, H., and Dokholyan, N. V. (2008) Ab initio folding of proteins with all-atom discrete molecular dynamics. *Structure 16*, 1010−1018.
(23) Tsuzuki, S., Honda, K., Uchimaru, T., Mikami, M., and Tanabe, K. (2002) Origin of attraction and directionality of the π/π interaction: Model chemistry calculations of benzene dimer interaction. *J. Am. Chem. Soc. 124*, 104−112.
(24) Rutledge, L. R., Durst, H. F., and Wetmore, S. D. (2009) Evidence for Stabilization of DNA/RNA-Protein Complexes Arising from Nucleobase-Amino Acid Stacking and T-Shaped Interactions. *J. Chem. Theory Comput. 5*, 1400−1410.
(25) Misura, K. M., Morozov, A. V., and Baker, D. (2004) Analysis of anisotropic side-chain packing in proteins and application to high-resolution structure prediction. *J. Mol. Biol. 342*, 651−664.
(26) Hunter, C. A., Singh, J., and Thornton, J. M. (1991) π-π interactions: The geometry and energetics of phenylalanine-phenylalanine interactions in proteins. *J. Mol. Biol. 218*, 837−846.
(27) Baker, C. M., and Grant, G. H. (2007) Role of aromatic amino acids in protein-nucleic acid recognition. *Biopolymers 85*, 456−470.
(28) Stott, K., Tang, G. S., Lee, K. B., and Thomas, J. O. (2006) Structure of a complex of tandem HMG boxes and DNA. *J. Mol. Biol. 360*, 90−104.
(29) Kim, Y., Geiger, J. H., Hahn, S., and Sigler, P. B. (1993) Crystal structure of a yeast TBP/TATA-box complex. *Nature 365*, 512−520.