JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Molecular Recognition of Platinated DNA from Chromosomal HMGB1

Trung Hai Nguyen,[†,‡,§] Giulia Rossetti,*[,†,‡,§,∥,⊥,▽] Fabio Arnesano,[#] Emiliano Ippoliti,[†,‡,§] Giovanni Natile,*[,#] and Paolo Carloni[†,‡,§]

[†]Computational Biophysics, German Research School for Simulation Sciences (joint venture of RWTH Aachen University and Forschungszentrum Jülich, Germany), D-52425 Jülich, Germany

[‡]Institute for Advanced Simulation IAS-5, Computational Biomedicine, Forschungszentrum Jülich, D-52425 Jülich, Germany

[§]Computational Biomedicine section (INM-9), Institute for Neuroscience and Medicine (INM), 52425 Jülich, Germany

[∥]Institute for Research in Biomedicine (IRB Barcelona), Baldiri Reixac 10, Barcelona 08028, Spain

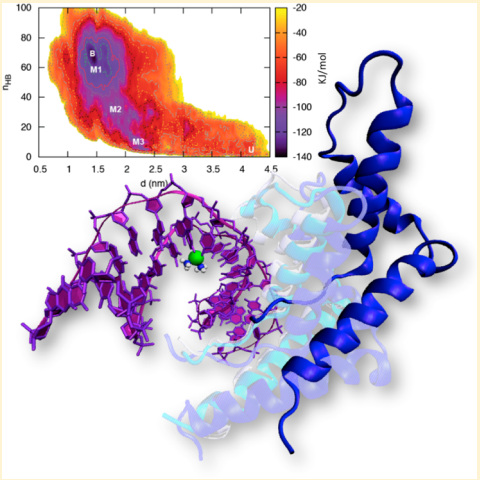
[⊥]Joint IRB-BSC Program in Computational Biology, Barcelona, Spain

[▽]Jülich Supercomputing Centre (JSC), Forschungszentrum Jülich, D-52425 Jülich, Germany

[#]Department of Chemistry, University of Bari "A. Moro", via Edoardo Orabona 4, I-70125 Bari, Italy

Supporting Information

ABSTRACT: Cisplatin cures testicular and ovarian cancers with unprecedented potency. It induces its beneficial activity by covalently binding to DNA. Repair enzymes, which remove the platinated lesions from DNA, cause drug resistance. Chromosomal High Mobility Group Box proteins (HMGB) may interfere with this process by binding to platinated DNA. Using 8 μs multiple-walker well-tempered metadynamics simulations, here, we investigated the structural and the energetic determinants of one of the HMGB proteins (HMGB1A) in complex with the platinated oligonucleotide $[Pt(NH_3)_2]^{2+}$-d(CCUCTCTG*G*ACCTTCC)-d(GGAGAGACCTGGAAGG) (*G are platinated guanines), for which experimental structural information is available. The calculated affinity is in good agreement with experiment. The process is predicted to be enthalpy-driven, as found for other protein/DNA complexes. The Lys7 residue, whose side-chain was not resolved in the X-ray structure, is found to interact with the C4 5′-phosphate and this interaction emerges as a key facet for the molecular recognition process. In addition, our calculations provide a molecular basis for the experimentally measured decreased affinity of HMGB1A for platinated DNA, as a consequence of Cys22-Cys44 S−S bridge formation (such an oxidation cannot take place in some members of this protein family present in the testis, where the drug is particularly effective). This decrease is likely to be caused by a small yet significant rearrangement of helices H1 and H2 with consequent alteration of the Phe37 juxtaposition.



INTRODUCTION

Cisplatin [cis-diamminedichloridoplatinum(II)] is one of the most effective drugs used in cancer therapy. It exerts the cytotoxic activity by covalently binding DNA[1,2] forming preferentially (>80%) 1,2-intrastrand cross-links with guanines[1−3] that bend DNA. However, nucleotide excision repair (NER) enzymes are able to remove the platinum lesions from DNA, leading to drug resistance.[1−3] Such a repair may be inhibited by chromosomal high-mobility group proteins (HMGB), HMGB1, HMGB2 and HMGB4[4] as shown by in vitro[5−7] and in cell experiments.[8] Notably, the binding affinity of the reduced state of HMGB1 protein for platinated DNA (Pt-DNA) is larger than that of the oxidized state, where two fully conserved cysteine residues (Cys22[9] and Cys44[10]) form a S−S bridge.[11] Both the reduced state and the Cys−Cys bridged protein could be present in disease conditions. This has led to the hypothesis that the high activity of cisplatin toward testicular tumors could be due to the unique expression in the testis[10] of HMGB4, which lacks one of the two cysteines and hence cannot form the S−S bond.[11]

Seminal X-ray studies show that the 3-helix bundle domain 'A' of the HMGB1 protein in the reduced state is L-shaped when in complex with the kinked $[Pt(NH_3)_2]^{2+}$-d(CCUCTCTG*G*ACCTTCC)-d(GGAGAGACCTGGAAGG) (*G are the platinated guanines) oligonucleotide (Pt-DNA•HMGB1A hereafter).[12] Most importantly, the aromatic residue of Phe37 intercalates into the hydrophobic notch formed by the two

Received: May 8, 2014
Published: June 30, 2014

   dx.doi.org/10.1021/ct500402e | J. Chem. Theory Comput. 2014, 10, 3578−3584

platinated guanine bases.[12] 50 ns-long molecular dynamics (MD) studies[13,14] based on this X-ray structure further suggested that the number of contacts resulting from this interaction is larger than in the case of oxaliplatin, a second-generation platinum-based drug, thus supporting the larger affinity of cisplatin-DNA•HMGB1A relative to the oxaliplatin-DNA•HMGB1A complex: the experimental free energies of binding (ΔG) were −52.5 kJ/mol[11] and −43 kJ/mol,[13,15,16] respectively. In these early studies, ΔG was not calculated, nor the molecular recognition process and the structural determinants of the oxidized form investigated.

Here we have used multiscale, multimicrosecond-long molecular simulation methods to gain insights into the recognition process of HMGB1A in the reduced state, as well as on the free energy landscape associated with it. This study is complemented by a structural prediction of the cisplatin-DNA•HMGB1A complex in the oxidized state, in which a S−S bond is formed.

We first developed accurate, AMBER-type[17] force field parameters for the platinum moieties based on the QM/MM simulations via the so-called "force matching" procedure.[18,19] In this approach, the parameters are derived so as to reproduce the QM/MM forces, the electrostatic potential, and the electric field obtained by QM/MM simulations. Then, we equilibrated the structure of Pt-DNA•HMGB1A complex in the reduced and oxidized state in aqueous solution by 0.2 μs-long MD simulations based on the obtained force field parameters for the platinated moiety, along with the AMBER Amber99SB-ILDN force field[17,20,21] with the Barcelona modification.[22] The latter, not used so far for Pt-DNA containing systems, is required for an accurate simulation of DNA based on the AMBER force field.[22,23] We finally used an enhanced sampling technique to calculate the free energy of unbinding of the Pt-DNA•HMGB1A complex as a function of appropriate reaction coordinates, already used for a variety of ligand/DNA complexes.[24−26] Specifically, we performed multiple-walker well-tempered metadynamics simulation[27] (overall trajectory collected for 8 μs) as a function of the distance (d) between HMGB1A and the Pt-DNA and the number of hydrogen bonds (n_{HB}) and hydrophobic contacts (n_{HC}) established at the interface between the two moieties (Pt-DNA and HMGB1A). The calculated free energy is in fair accord with the experiment.[11] Our calculations point to a key role of Lys7, a fully conserved residue across mammalian HMGB proteins, for the molecular recognition process. Moreover, the displacement of Phe37 from the hydrophobic notch formed by the cross-linked G8 and G9 bases, as a consequence of S−S bridge formation between Cys22 and Cys44, appears to be responsible for the decreased affinity upon protein oxidation.

MATERIALS AND METHODS

Our calculations were based on the X-ray structure of Pt-DNA•HMGB1A (PDB code 1CKT)[12] (Figure 1). The complex was immersed in a box of water. Twenty Na^+ ions were added to neutralize the total charge.

Parameterization of the Platinated Site. AMBER-like parameters for the platinated moiety were obtained using the force matching approach.[18,19] The whole system underwent first 10 ns-long classical MD simulation, with an approximate force field for the platinated site and the AMBER force field.[17,20,21] This simulation was used only to provide an initial structural model for the subsequent QM/MM simulation. The last snapshot was selected for the subsequent 10 ps-long QM/MM simulation (see Supporting Information, section S1). The CPMD program,[29,30] combined with the classical MD Gromos96 code[31] through the interface developed by Rothlisberger et al.,[32,33] was employed for the QM/MM simulation. The QM part includes two guanine bases, G8 and G9, the Pt^{2+} ion, and two NH_3 ligands. The MM part consists of the rest of the system. It was treated with the AMBER99SB-ILDN force field[17,20,21] (see Supporting Information, Table S1).

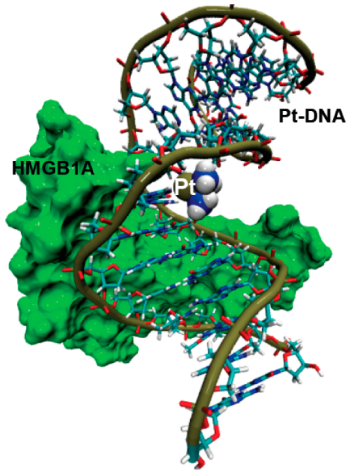


Figure 1. X-ray structure of Pt-DNA•HMGB1A (PDB code 1CKT).[12] Hydrogen atoms and missing side chains were added using the Leap program.[28] DNA is shown with bond representation, HMGB1A with Connolly surface representation and the cisplatin moiety with van der Waals spheres.

The force matching procedure[18,19] was applied over 200 conformations selected from the QM/MM trajectory. The atomic partial charges (Supporting Information, Table S2 and Figure S1) were obtained through a fit to the electrostatic potential and the electric field. The electrostatic forces derived from those charges, along with the ones coming from the van der Waals parameters taken from AMBER parm99SB force field,[17,20] were subtracted from the total forces acting on the QM atoms to yield the bonded forces. A least-squares fit of the latter allowed obtaining the AMBER-type bond, angle, and dihedral angle force field parameters (Supporting Information, Tables S3, S4 and Figure S2). Equilibrium bond lengths and bond angles were taken from the QM/MM averaged values. For the remaining calculations presented here, we used this force field, along with the Amber99SB-ILDN force field[17,20,21] with the Barcelona modification[22] for the DNA frame and the protein, as well as the TIP3P[34] and the Joung and Cheatham[35] force fields for water and the counterions, respectively.

MD and Free Energy Calculations. The X-ray structure of Pt-DNA•HMGB1A (PDB code 1CKT) was solvated in a truncated octahedral box containing water and 86 Na^+ and 66 Cl^- ions (see Supporting Information, section 2 and Table S5). The ionic strength corresponds to the experimental one (100 nM).[11] Long-range electrostatic interactions were evaluated using the Particle Mesh Ewald (PME) method.[36] The cutoff for the real part of the PME and for the van der Waals interaction was set to 10 Å. All bonds involving hydrogen atoms were constrained using Lincs algorithm.[37] An integration step of 2 fs was employed. NPT conditions were achieved by coupling the system to the modified Berendsen thermostat[38] and the Parrinello−Rahman barostat.[39]

dx.doi.org/10.1021/ct500402e | J. Chem. Theory Comput. 2014, 10, 3578−3584

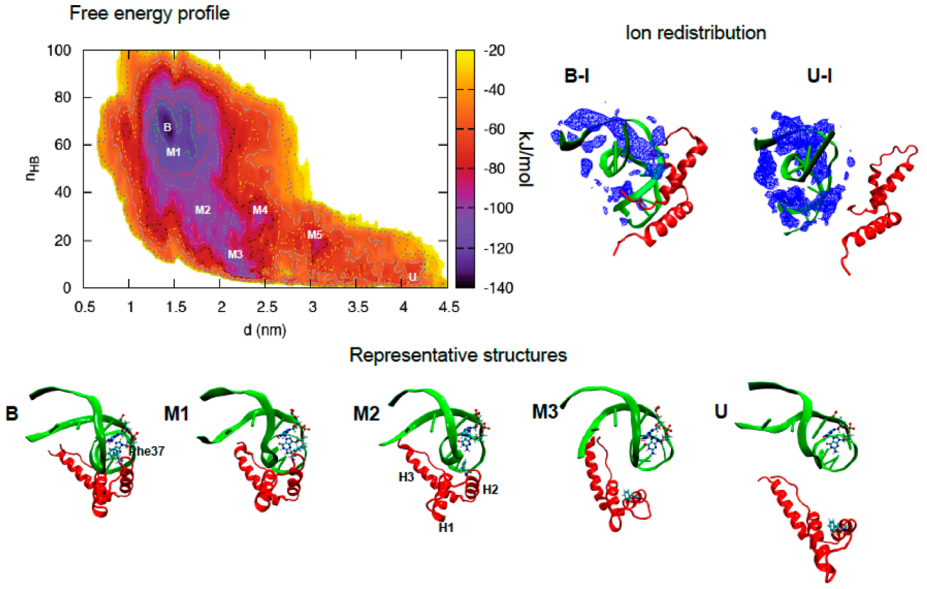


Figure 2. Free energy profile (top left) projected on d and n_{HB} collective variables [d is the distance between the centers of mass of Pt-DNA and HMGB1A and n_{HB} is the number of hydrogen bonds at the interface between the two moieties]. B denotes the bound while U represents the unbound minimum [Pt-DNA and HMGB1A protein ∼4 nm apart]. M1−M5 are the local minima. Figure S5 in Supporting Information shows the plot of the free energy profile projected on d and n_{HP} [n_{HP} is the number of hydrophobic contacts at the interface between the two moieties]. The counterion distribution for B (B-I) and U (U-I) is shown on the top right. The most representative (>55%) structures corresponding to the B, M1−M3, and U are shown on the bottom row. The Pt-DNA is represented by green ribbon, the HMGB1A protein by red ribbon, and the platinated guanines and the Phe37 residue are shown in balls and sticks.

200-ns MD simulations were carried out for Pt-DNA•HMGB1A, its oxidized form, and the naked d-(CCUCTCTGGACCTTCC)-d(GGAGAGACCTGGAAGG) unplatinated DNA (see Supporting Information, section S2 and Tables S5−S8 for details).

Finally, Pt-DNA•HMGB1A underwent 8 μs-long multiple-walker, well-tempered (WT) metadynamics[27,40−42] free energy calculations, with 8 walkers. The free energy G was calculated as a function of the following collective variables (CVs): (i) the distance d between the centers of mass of Pt-DNA and HMGB1A; (ii) the number of hydrogen bonds n_{HB} between the Pt-DNA and HMGB1A [n_{HB} was calculated by eq 1

$$n_{HB} = \sum_{i,j} \frac{1 - (r_{ij}/r_0)^6}{1 - (r_{ij}/r_0)^{12}} \quad (1)$$

where i and j runs over H-bond donors and acceptors of Pt-DNA and HMGB1A, respectively, while r_0 = 2.5 Å]; (iii) the number of hydrophobic contacts n_{HP} between nonpolar carbons on Pt-DNA and HMGB1A [n_{HP} was also calculated by eq 1 in which r_0 = 6 Å: the latter value accounts for the typical carbon−carbon distance (4/4.5 Å) and the thermal fluctuation amplitude (1.5/2 Å)]. Similar CVs have been employed to study DNA/ligand and protein/ligand complexes.[24−26,43−45] The initial height of the Gaussian hills was set to 1 kJ/mol. The widths of Gaussians were chosen to be $\sim^1/_3$ of the typical fluctuations of the CVs during the MD simulation, following ref 46. The deposition time of the biasing Gaussians was fixed to 1 ps. All calculations used the Gromacs 4.5.3 program[47] with the Plumed plugin.[48] The contribution to the free energy of unbinding from the metadynamics ΔG^{MetaD} was calculated as the free energy difference between the minimum at d = 4.2 nm and the bound state at d = 1.5 nm (U and B in Figure 2, respectively). The contribution for d > 4.2 nm (ΔG^{Elec}) was estimated through the nonlinear Poisson−Boltzmann equation by using Delphi v.5.1 program.[49] The total calculated value for the free energy of binding was obtained as $\Delta G = \Delta G^{MetaD} + \Delta G^{Elec}$. The standard-state free energy of binding was calculated by eq 2 in which

$$\Delta G^{\circ} = \Delta G - RT \ln\left(\frac{[P]}{[P]^0}\right) \quad (2)$$

R is the molar constant, $[P]$ is the concentration of the protein in our simulation box, and $[P]^0$ = 1 M is the standard-state concentration.[50] ΔG° was compared with the experimental binding energy through the relationship $\Delta G^{\circ} = -RT \ln(k_{eq})$, where k_{eq} is the experimental equilibrium constant. The enthalpy contribution ΔH to the ΔG^{MetaD} is estimated as the difference in electrostatic and van der Waals energies between bound and unbound states. The entropy contribution, $-T\Delta S$, is then determined from the relationship given in eq 3.

$$\Delta G^{MetaD} = \Delta H - T\Delta S \quad (3)$$

Structural Model of the Complete Protein in Complex with Pt DNA. The N- and C-terminal tails (res 1−6 and res 79−89) of HMGB1A are missing in the X-ray structure of the complex with the Pt-DNA.[12] We added the missing tails using the Robetta server.[51,52] Several models were generated. We selected the two models (Pt-DNA•HMGB1A-Full-1 and Pt-DNA•HMGB1A-Full-2) featuring the lowest discrepancy, in terms of secondary structure and of RMSD, from the X-ray structure.[12] These underwent 200 ns-long MD simulations with the same protocol as described above.

Calculated Structural Properties. The helical parameters of the base pair steps of DNA, such as twist, roll, tilt, rise, shift, and slide were obtained using the Curves+ program.[53] Cluster analysis from ref 54 was performed to find the most representative structures for the minimum basins.

dx.doi.org/10.1021/ct500402e | J. Chem. Theory Comput. 2014, 10, 3578−3584

RESULTS AND DISUSSIONS

To carry out this multiscale modeling study, we first determined force field parameters compatible with those of the AMBER force field for the platinated moiety in the Pt-DNA•HMGB1A complex by QM/MM simulations and force matching procedure. Based on this force field and on the X-ray structure,[12] 0.2 μs-long MD simulation of the Pt-DNA•HMGB1A complex (in the reduced form) results to be rather similar to the crystal structure[12] (RMSD ∼ 2.4 Å, Supporting Information, Figure S3 and Table S8). In particular, the contacts between the two moieties are maintained (Supporting Information, Figure S4). Eight μs-long multiple-walker WT metadynamics calculations, based on the last MD snapshot of the complex, provide the free energy as a function of three collective variables. These are d, the distance between the centers of mass of the two interacting moieties (Pt-DNA and HMGB1A), n_{HB}, the number of hydrogen bond contacts (Figure 2), and n_{HP}, the number of hydrophobic contacts at the interface. A cluster analysis was performed to provide a structural ensemble of all minima.

Free Energy Landscape. The free energy displays a well-defined absolute minimum, which can be seen in the figures of the free energy as a function of two CVs (B in Figure 2 and Suporting Information, Figure S5). Expectedly, this corresponds to the bound state. The most representative structure corresponding to B is very similar to the MD structure of Pt-DNA•HMGB1A discussed above. It is also rather similar to the X-ray structure (Figure 3a, RMSD ∼ 2.6 Å) despite the total

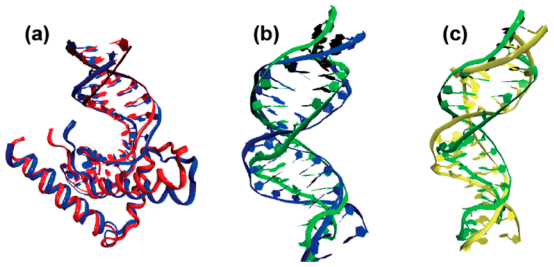


Figure 3. (a) Superimposition of the Pt-DNA•HMGB1A backbones derived from the X-ray structure (red)[12] and from metadynamics simulations (minimum B, blue). (b) Superimposition of unplatinated DNA from MD simulations (green) and Pt-DNA as it occurs in B (blue) of metadynamics simulations (the protein was removed for clearness). (c) Superimposition of unplatinated DNA from MD simulations (green) and Pt-DNA in U (yellow) of metadynamics simulations. Only the representative structures from cluster analysis (see methods for details) are shown in the figure.

helix axis bends more in B (Figure 4). The helical structural parameters of structure B (bending, twist, roll, tilt, rise, shift, and slide) are not too dissimilar from those of the X-ray structure, except for those near the platinated site (Figure 4 and Supporting Information, Figure S6), which deformed more dramatically in B. Accordingly, the DNA axis bends near the platinated site more in B than in the X-ray structure (Figure 4). HMGB1A binds to the minor groove of Pt-DNA, opposite to the platinated site, while the phenyl ring of the Phe37 residue inserts into the hydrophobic cleft formed by the cross-linked guanine bases G8 and G9. Furthermore, the Lys7 residue makes a very strong hydrogen bond with the 5′-phosphate of the C4 base. This hydrogen bond could not be observed in the X-ray structure due to the missing side chain of Lys7.[12]

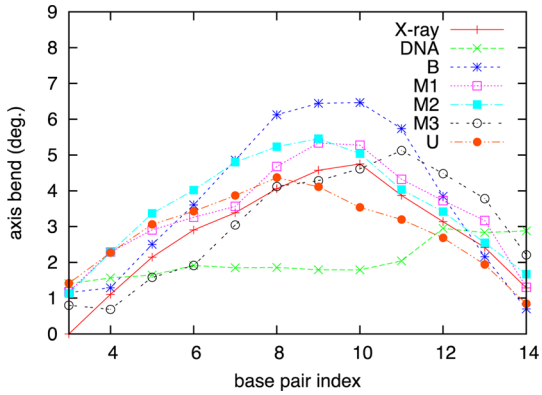


Figure 4. Bending of the DNA's axis, as a function of base-pair sequence, for the X-ray structure, the MD simulation of unplatinated DNA (labeled as "DNA" in the figure), and the metadynamics simulations of Pt-DNA at different stages of interaction with HMGB1A (B, M1−M3, and U). Base pair indices 8 and 9 correspond to the platinated bases. Only the representative structures from cluster analysis (see methods for details) were used for the figure.

U in Figure 2 is the local minimum at the highest free energy. In the representative structure of this minimum, HMGB1A is detached from Pt-DNA. The complex is characterized by a lack of intermolecular hydrogen bonds and of hydrophobic interactions, which instead are maximized in B (Figure 2 and Supporting Information Figure S5). The bending of the Pt-DNA axis in U decreases with respect to that in B by ∼2° near the platinated site (Figure 4). The Na^+ counterions condense at the Pt-DNA/protein interface (U-I in Figure 2). These ions are absent in B (B-I in Figure 2).

It is well-known that platination causes a considerable bending of the DNA axis; however, there are some significant differences between the X-ray structure and the Pt-DNA solution structures in B (the protein-bound state), and in U (the protein-unbound state). The bending of the Pt-DNA axis is highest in B (Figure 4), clearly indicating that the binding of the protein further increases the bending of the Pt-DNA axis and this could be in solution even greater than that observed in the solid state.

The investigation of the unbinding process by metadynamics allows to estimate the binding free energy. The free energy difference between B and U is about −58 kJ/mol (Figure 2 and Supporting Information, Figure S7). The residual binding free energy on passing from U to a complete separation of the two molecules is expected to be significant because of the long-range electrostatic interactions between two highly charged molecules. An estimate of this free energy, calculated by solving the nonlinear Poisson−Boltzmann eq (Supporting Information, Figure S8), is −20 ± 4 kJ/mol. Therefore, the total binding free energy emerging from metadynamics and electrostatics calculations is −78 ± 4 kJ/mol. The standard state free energy (ΔG°), which takes into account the concentration of the protein in our simulation [eq 2] gives a value of −63 ± 4 kJ/mol, which is not too dissimilar from the experimental value of −52.5 ± 0.3 kJ/mol.[11]

The calculated enthalpy (ΔH ∼ −107 kJ/mol) and entropy ($-T\Delta S$ ∼ 49 kJ/mol) changes between U and B govern the binding process. The decrease in enthalpy favors the bound state, while the loss of conformational entropy disfavors the binding. However, the entropic penalty is smaller than the enthalpic gain, which results in a net decrease in free energy

dx.doi.org/10.1021/ct500402e | J. Chem. Theory Comput. 2014, 10, 3578−3584

upon binding. We expect that the relatively small contribution from the residual binding free energy does not change this picture. Hence, overall, our calculations suggest that the Pt-DNA•HMGB1A adduct formation is enthalpy driven. Such a compensation phenomenon has been observed in the formation of other protein−DNA adducts.[55−57]

Molecular Recognition Process. Our study allows identifying local minima during the unbinding process. This information might be used to gain insights on the molecular recognition process. The three local free energy minima M1−M3 (Figure 2), whose depth relative to B decreases progressively, appear to lie along the unbinding pathway of the complex.[58] Hence, the structural features of the complex in these local minima might give insights into the binding and unbinding mechanism of the complex. In M1 the relative position of Pt-DNA and HMGB1A is essentially unchanged with respect to B, except that the phenyl ring of the Phe37 residue has moved away from the hydrophobic notch formed by the cross-linked G8 and G9 bases (M1 in Figure 2). A few water molecules enter the resulting cavity (Supporting Information, Table S10).

In M2, the minor groove close to the cisplatin moiety detaches from the edge of the L-shaped H1 and H2 helices of HMGB1A (Figure 2). This allows water to fully penetrate into the interstitial space. The edge of the L-shaped H2 and H3 helices of HMGB1A remains still in contact with the minor groove few base steps away from the platinated site. The electrostatic contacts between Lys11 and C15, as well as between Lys7 and the 5′-phosphate of the C4 base, still remain.

In M3, the HMGB1A protein is completely detached from the minor groove of the Pt-DNA (Figure 2). The only remaining electrostatic contact between the Pt-DNA and the HMGB1A protein is the hydrogen bond formed between the Lys7 residue and the 5′-phosphate of the C4 base.

To investigate the structural properties of Pt-DNA in M1−M3, we also calculated the bending of the DNA axis (Figure 4) and other helical parameters such as roll, tilt, twist, rise, shift, and slide (Supporting Information, Figure S6).[53] The bending of the DNA axis in M1−M3 decreases slightly with respect to that in B. This suggests that the unbinding of HMGB1A is accompanied by a decrease of the Pt-DNA bending. This is consistent with the X-ray studies, which have shown that the bending of Pt-DNA alone[59] is smaller than that of the Pt-DNA•HMGB1A complex.[12] The other helical parameters are quite similar to those of structure B (Supporting Information, Figure S6).

Effect of Oxidation. We compared structure B with that of the complex in which Cys22 and Cys44 form a S−S bridge (oxidized state) as obtained by a 0.2 μs-long MD simulation in aqueous solution. Although most intermolecular contacts between the two biomolecules remain similar to those of the X-ray[12] and MD structures of the reduced Pt-DNA•HMGB1A complex (Supporting Information, Table S9 and Figure S4), helix H2 moves slightly toward helix H1 (Supporting Information, Figure S9a) and, as a result, Phe37, located on the top of helix H2, is slightly displaced from the hydrophobic notch formed by the platinated bases G8 and G9 (Supporting Information, Figure S9b). The orientation of this residue in the oxidized state is very different from that in the reduced state, as shown by a recent NMR study of free HMGB1A.[60] Hence, the conformational change of the Phe37 residue is expected to be an important structural determinant of the experimentally observed decrease in binding free energy (~5.5 kJ/mol), upon Pt-DNA•HMGB1A complex formation, on passing from the reduced to the oxidized protein state.[11]

Role of the N- and C-Terminal Tails. Residues 1−6 and 79−89 are missing in the X-ray structure.[12] Here, we have predicted their structural determinants using the Robetta approach.[52] Two models (Pt-DNA•HMGB1A-Full-1 and Pt-DNA•HMGB1A-Full-2), differing for the orientations of the N-tail, were generated and these were subjected to 200 ns MD simulation. In the first system, the N-terminal tail extends to the solvent (Supporting Information, Figure S10a). Lys7 and Lys6 are in contact with the phosphate groups of C4 and C2 bases, respectively, while Lys2 is mostly solvent-exposed. The other residues are not in contact with the Pt-DNA. The position of the side arm of Lys7 is very similar to that observed when residues 1−6, not localized in the X-ray structure determination, were not modeled. In the second system, the N-terminal tail curves back into the major groove of the DNA (Supporting Information, Figure S10b). The terminal NH_3 group makes electrostatic contact with the phosphate group of T7, while Lys7 is in contact with G5′s phosphate group. Lys2 and Lys6 are mostly exposed to water. The other residues are not in contact with Pt-DNA. Notably, in both systems, the C-terminal tail is very disordered and does not have contacts with the Pt-DNA.

This investigation provides two interesting results: first, the N-terminus can curve back in the major groove where the platinum moiety is located; second, it can have very different localizations spanning from interaction with the phosphates of the platinated strand to phosphates of the complementary strand of DNA. Such an undetermined localization can fully explain the failure to detect residues 1−6 in the X-ray structure. Moreover, Lys7 is found in contact with a phosphate group of the DNA not only in our metadynamics free energy minimum B but also in the two models described above; in contrast, Lys6 and Lys2 are mostly solvent exposed. Therefore, it is possible to conclude that the role of Lys7 for the molecular recognition may be prominent on that of other lysine residues present at the N-terminus.

Limitations of the Present Work. As in any molecular modeling study, several limitations may affect our predictions. First, the accuracy of the calculated free energy depends on the choice of the collective variables (CVs). Our choice of CVs [including distance (d), number of hydrogen bonds (n_{HB}), and hydrophobic contacts (n_{HP})] takes into account the most important interactions at short distances. These CVs have already been used successfully in studying biomolecular recognition processes.[24−26,43−45] Nevertheless, the recognition process studied here is very complex, involving two large and highly charged molecules. In particular, at long distances the short-range CVs n_{HB} and n_{HP} become void and ineffective to describe the relevant motions of the process. Other CVs, might be required to account for the favorable orientation of the two molecules upon approaching each other. In the present study, these long-range CVs were not included, which might affect the accuracy of the calculated free energy profile at long distances. A second source of error can be the use of a force field with fixed point-charges. To reduce the latter error, we have chosen an accurate, most-updated, force field for protein and DNA.[17,20−22] Furthermore, we developed the force parameters for the platinated moiety using a sophisticated force-matching technique,[18,19] which takes into account effects of solvent and finite temperature. In addition, the calculated overall free energy of the process contains also a significant contribution

from highly approximate nonlinear Poisson−Boltzmann calculations. This is an additional source of inaccuracy in our estimate of the free energy.

CONCLUSIONS

The present investigation concerns the molecular recognition process between cisplatin-cross-linked DNA and the chromosomal HMGB1A protein.

Our calculated free energy is fairly consistent with the experimental data.[11] The molecular recognition process appears to be enthalpy driven, as found in other DNA/protein complexes.[55−57]

The work by Lippard et al.[12] showed how the phenyl group of Phe37 inserts in between the cross-linked guanines and the resulting stacking interaction contributes to the stabilization of the protein−DNA complex. This computational investigation has shown that, indeed, this interaction locks the protein−DNA complex in the absolute free energy minimum (B in Figure 2); moreover, the disruption of this interaction is the first step in the dissociation process, further supporting its key role in the stabilization of the complex. We also notice that in B, as in the X-ray structure of the complex,[12] HMGB1A establishes only one hydrophobic interaction with the cross-linked G8 and G9 bases and several hydrophobic and hydrogen bonding interactions with adjacent nucleotides. This feature is similar to what found in a recent structural prediction on the interaction of Pt-DNA with the DNA repair protein MutSα.[61]

Our calculations point to a prominent role of the Lys7 residue in the molecular recognition process. Lys7 is fully conserved across mammals (using clustal W[62]) and forms a hydrogen bond with the 5′-phosphate of cytosine 4 (C4). Significantly, the amide $^{1}H/^{15}N$ NMR resonances of Lys7 were shown to undergo a large shift upon formation of a HMBGB1A/14-mer complex,[63,64] and, in the Pt-DNA•HMGB1A X-ray structure, Lys7 was the first residue of the N-terminus to be seen in the electron density map[12] although its side chain was not solved leaving undefined its role. From our simulations, this residue seems to be the last 'anchor' between the HMGB1A protein and the [Pt(NH$_3$)$_2$]$^{2+}$-d(CCUCTCTG*G*ACCTTCC)-d(GGAGAGACCTGGAAGG) DNA. Most likely, the Lys7-C4 5′-phosphate interaction is also the first step in the protein−DNA recognition process. Thus, this interaction could influence both the formation as well as the dissociation process, by favoring the first and disfavoring the latter, and contribute significantly to the stability of the DNA−protein complex.

Based on the molecular contacts, which are progressively lost on passing from B to U, it can be predicted that mutation of Lys11, Arg23, Lys42, and Ser45 to Gly or Ala could significantly affect the binding affinity of HMGB1A for Pt-DNA (see Supporting Information, Figure S12).

Our calculations also provide a structural basis for the experimentally observed decrease in affinity upon protein oxidation and consequent formation of a Cys22−Cys44 disulfide bridge (5.5 kJ/mol).[11] This has led to the hypothesis that the high efficacy of cisplatin toward testicular tumors could be due to the presence in the testis of a member of the HMGB superfamily (the HMGB4 protein)[10,11] which lacks one of the two cysteine residues. The latter protein cannot form the S−S bridge upon oxidation. Furthermore, our calculations suggest that the decrease in binding free energy upon formation of the Cys22−Cys44 disulfide bridge is likely due to a structural rearrangement of helices H1 and H2, which causes Phe37, located on top of helix H2, to slightly withdraw from the hydrophobic notch formed by the cross-linked G8 and G9 bases. The latter result is rather remarkable since it shows that computational methods can correctly evaluate loss of stability consequent to small conformational changes.

A further step in the modeling of this system should take into account the possibility that, in vivo, some lysine residues can be acetylated.[65,66] Indeed, this post-translational modification is known to enhance the affinity of HMGB1A for Pt-DNA.[64,67,68] Hence, it would be very important to determine how lysine acetylation impacts on the molecular recognition of Pt-DNA by HMGB1. Once this problem has been settled, we plan to investigate enantiomers of oxaliplatin to sort out how the different stereochemistry of the diamine ligand influences the strength of the interaction of HMGB1 with Pt-DNA. In the latter case there are already experimental data,[69−73] which can be used to test the predictivity of our computational methods.

ASSOCIATED CONTENT

Supporting Information

Methods, Tables S1−S10, and Figures S1−S12. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*Email: giovanni.natile@uniba.it.
*Email: g.rossetti@grs-sim.de.

Funding

The University of Bari; the Consorzio Interuniversitario di Ricerca in Chimica dei Metalli nei Sistemi Biologici (CIRCMSB); the Italian Ministero dell'Università e della Ricerca [PON 01078]; and the European Commission (COST Action CM1105) to [F.A. and G.N.]. T.H.N. acknowledges the Deutsche Forschungsgemeinschaft (DFG) for financial support (Grant No. CA 973/8-1). We gratefully acknowledge that this computational work has been accomplished using the PRACE (http://www.prace-ri.eu/) research infrastructure resource CURIE hosted by CEA, France.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Jung, Y.; Lippard, S. J. *Chem. Rev.* **2007**, *107*, 1387−1407.
(2) Wang, D.; Lippard, S. J. *Nat. Rev. Drug. Discovery* **2005**, *4*, 307−320.
(3) Jamieson, E. R.; Lippard, S. J. *Chem. Rev.* **1999**, *99*, 2467−2498.
(4) Both HMGB1 and HMGB2 are ubiquitously expressed, whereas HMGB4 is exclusively expressed in testis. HMGB4 lacks the acidic C-terminal tail and one of the conserved cysteines, Cys22, which can form a disulfide bridge with Cys44 under oxidizing conditions.
(5) Ugrinova, I.; Zlateva, S.; Pashev, I. G.; Pasheva, E. A. *Int. J. Biochem. Cell Biol.* **2009**, *41*, 1556−1562.
(6) Huang, J. C.; Zamble, D. B.; Reardon, J. T.; Lippard, S. J.; Sancar, A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 10394−10398.
(7) Malina, J.; Kasparkova, J.; Natile, G.; Brabec, V. *Chem. Biol.* **2002**, *9*, 629−638.
(8) Yusein-Myashkova, S.; Ugrinova, I.; Pasheva, E. *BMB Rep.* **2013**, 2555.
(9) *We use the numeration adopted in the deposited X-Ray structure 1CKT (12).*
(10) Catena, R.; Escoffier, E.; Caron, C.; Khochbin, S.; Martianov, I.; Davidson, I. *Biol. Reprod.* **2009**, *80*, 358−366.
(11) Park, S.; Lippard, S. J. *Biochemistry* **2011**, *50*, 2567−2574.

(12) Ohndorf, U. M.; Rould, M. A.; He, Q.; Pabo, C. O.; Lippard, S. J. *Nature* **1999**, *399*, 708−712.
(13) Ramachandran, S.; Temple, B. R.; Chaney, S. G.; Dokholyan, N. V. *Nucleic Acids Res.* **2009**, *37*, 2434−2448.
(14) Ramachandran, S.; Temple, B.; Alexandrova, A. N.; Chaney, S. G.; Dokholyan, N. V. *Biochemistry* **2012**, *51*, 7608−7617.
(15) Wei, M.; Cohen, S. M.; Silverman, A. P.; Lippard, S. J. *J. Biol. Chem.* **2001**, *276*, 38774−38780.
(16) Malina, J.; Novakova, O.; Vojtiskova, M.; Natile, G.; Brabec, V. *Biophys. J.* **2007**, *93*, 3950−3962.
(17) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(18) Maurer, P.; Laio, A.; Hugosson, H. W.; Colombo, M. C.; Rothlisberger, U. *J. Chem. Theory Comput.* **2007**, *3*, 628−639.
(19) Spiegel, K.; Magistrato, A.; Maurer, P.; Ruggerone, P.; Rothlisberger, U.; Carloni, P.; Reedijk, J.; Klein, M. L. *J. Comput. Chem.* **2008**, *29*, 38−49.
(20) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712−725.
(21) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins* **2010**, *78*, 1950−1958.
(22) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., 3rd; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817−3829.
(23) Perez, A.; Lankas, F.; Luque, F. J.; Orozco, M. *Nucleic Acids Res.* **2008**, *36*, 2379−2394.
(24) Vargiu, A. V.; Ruggerone, P.; Magistrato, A.; Carloni, P. *Nucleic Acids Res.* **2008**, *36*, 5910−5921.
(25) Biarnes, X.; Bongarzone, S.; Vargiu, A. V.; Carloni, P.; Ruggerone, P. *J. Comput.-Aided Mol. Des.* **2011**, *25*, 395−402.
(26) Coletta, A.; Desideri, A. *Nucleic Acids Res.* **2013**, *41*, 9977−9986.
(27) Barducci, A.; Bussi, G.; Parrinello, M. *Phys. Rev. Lett.* **2008**, *100*, 020603.
(28) Case, D.; Darden, T. A.; Cheatham, T. E.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Crowley, M.; Walker, R.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. *AMBER12*; University of California: San Francisco, 2012.
(29) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471−2474.
(30) *CMPD*; IBM Corp, 1990−2008; MPI für Festkörperforschung Stuttgart, 1997−2001; http://www.cpmd.org(accessed June, 14, 2014).
(31) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and Userguide*; Hochschuleverlag AG an der ETH: Zürich, 1996.
(32) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Phys. Chem. B* **2002**, *106*, 7300−7307.
(33) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941−6947.
(34) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(35) Joung, I. S.; Cheatham, T. E., 3rd *J. Phys. Chem. B* **2008**, *112*, 9020−9041.
(36) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(37) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(38) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101−014107.
(39) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182−7190.
(40) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.
(41) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. *J. Phys. Chem. B* **2006**, *110*, 3533−3539.
(42) Leone, V.; Marinelli, F.; Carloni, P.; Parrinello, M. *Curr. Opin. Struct. Biol.* **2010**, *20*, 148−154.
(43) Barducci, A.; Chelli, R.; Procacci, P.; Schettino, V.; Gervasio, F. L.; Parrinello, M. *J. Am. Chem. Soc.* **2006**, *128*, 2705−2710.
(44) Piana, S.; Laio, A. *J. Phys. Chem. B* **2007**, *111*, 4553−4559.
(45) Pietrucci, F.; Marinelli, F.; Carloni, P.; Laio, A. *J. Am. Chem. Soc.* **2009**, *131*, 11811−11818.
(46) Gervasio, F. L.; Laio, A.; Parrinello, M. *J. Am. Chem. Soc.* **2005**, *127*, 2600−2607.
(47) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(48) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. *Comput. Phys. Commun.* **2009**, *180*, 1961−1972.
(49) Li, L.; Li, C.; Sarkar, S.; Zhang, J.; Witham, S.; Zhang, Z.; Wang, L.; Smith, N.; Petukh, M.; Alexov, E. *BMC Biophys.* **2012**, *5*, 9.
(50) Kranjc, A.; Bongarzone, S.; Rossetti, G.; Biarnés, X.; Cavalli, A.; Bolognesi, M. L.; Roberti, M.; Legname, G.; Carloni, P. *J. Chem. Theory Comput.* **2009**, *5*, 2565−2573.
(51) Moult, J.; Fidelis, K.; Kryshtafovych, A.; Tramontano, A. *Proteins* **2011**, *79* (Suppl 10), 1−5.
(52) Kim, D. E.; Chivian, D.; Baker, D. *Nucleic Acids Res.* **2004**, *32*, W526−531.
(53) Lavery, R.; Moakher, M.; Maddocks, J. H.; Petkeviciute, D.; Zakrzewska, K. *Nucleic Acids Res.* **2009**, *37*, 5917−5929.
(54) Micheletti, C.; Seno, F.; Maritan, A. *Proteins* **2000**, *40*, 662−674.
(55) Joynt, S.; Morillo, V.; Leng, F. *Biophys. J.* **2009**, *96*, 4144−4152.
(56) Vaz de Andrade, E.; Freitas, S. M.; Ventura, M. M.; Maranhao, A. Q.; Brigido, M. M. *Biochim. Biophys. Acta* **2005**, *1726*, 293−301.
(57) Cui, T.; Wei, S.; Brew, K.; Leng, F. *J. Mol. Biol.* **2005**, *352*, 629−645.
(58) M4 and M5 are characterized by more than one representative structure (Supporting Information Figure S10). They are sampled during the later part of the metadynamics trajectory when Pt-DNA and HMGB1A protein approach each other from the U state. These structures correspond to metastable binding modes between the Pt-DNA and the HMGB1A protein and they are unlikely to contribute to the molecular recognition process.
(59) Takahara, P. M.; Rosenzweig, A. C.; Frederick, C. A.; Lippard, S. J. *Nature* **1995**, *377*, 649−652.
(60) Wang, J.; Tochio, N.; Takeuchi, A.; Uewaki, J.; Kobayashi, N.; Tate, S. *Biochem. Biophys. Res. Commun.* **2013**, *441*, 701−706.
(61) Negureanu, L.; Salsbury, F. R., Jr. *J. Biomol. Struct. Dyn.* **2014**, *32*, 969−992.
(62) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. *Bioinformatics* **2007**, *23*, 2947−2948.
(63) Hardman, C. H.; Broadhurst, R. W.; Raine, A. R.; Grasser, K. D.; Thomas, J. O.; Laue, E. D. *Biochemistry* **1995**, *34*, 16596−16607.
(64) Assenberg, R.; Webb, M.; Connolly, E.; Stott, K.; Watson, M.; Hobbs, J.; Thomas, J. O. *Biochem. J.* **2008**, *411*, 553−561.
(65) Yang, H.; Antoine, D. J.; Andersson, U.; Tracey, K. J. *J. Leukocyte Biol.* **2013**, *93*, 865−873.
(66) Elenkov, I.; Pelovsky, P.; Ugrinova, I.; Takahashi, M.; Pasheva, E. *Int. J. Biol. Sci.* **2011**, *7*, 691−699.
(67) Ugrinova, I.; Pasheva, E. A.; Armengaud, J.; Pashev, I. G. *Biochemistry* **2001**, *40*, 14655−14660.
(68) Pasheva, E.; Sarov, M.; Bidjekov, K.; Ugrinova, I.; Sarg, B.; Lindner, H.; Pashev, I. G. *Biochemistry* **2004**, *43*, 2935−2940.
(69) Malina, J.; Kasparkova, J.; Natile, G.; Brabec, V. *Chem. Biol.* **2002**, *9*, 629−638.
(70) Delalande, O.; Malina, J.; Brabec, V.; Kozelka, J. *Biophys. J.* **2005**, *88*, 4159−4169.
(71) Malina, J.; Novakova, O.; Vojtiskova, M.; Natile, G.; Brabec, V. *Biophys. J.* **2007**, *93*, 3950−3962.
(72) Kasparkova, J.; Vojtiskova, M.; Natile, G.; Brabec, V. *Chemistry* **2008**, *14*, 1330−1341.
(73) Malina, J.; Novakova, O.; Natile, G.; Brabec, V. *Chem.—Asian J.* **2012**, *7*, 1026−1031.