Inorganic Chemistry

Article

pubs.acs.org/IC

Detecting DNA Mismatches with Metallo-Insertors: A Molecular Simulation Study

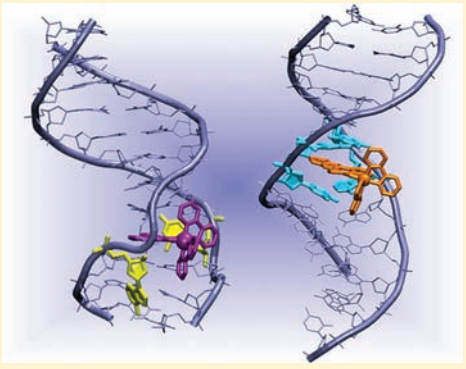
Attilio V. Vargiu[†] and Alessandra Magistrato*[,‡]

[†]CNR-IOM, Unità Operativa di Supporto SLACS, c/o Dipartimento di Fisica, Università di Cagliari, s.p. Monserrato-Sestu km 0.7, I-09042 Monserrato, Italy

[‡]CNR-IOM-Democritos, National Simulation Center, c/o International School for Advanced Studies (SISSA/ISAS), via Bonomea 265, 34136, Trieste, Italy

Supporting Information

ABSTRACT: Molecules that selectively recognize DNA mismatches (MMs) play a key role as nucleic acids probes and as chemotherapeutic agents. Metallo-insertors bind to the minor groove (mG) of double strand (ds) DNA, expelling the mismatched base pairs and acting as their π-stacking replacement. In contrast, metallo-intercalators bind to the major groove (MG) of ds DNA and π-stack to adjacent base pairs. In this study we focused on structural and energetic properties of Δ−[Rh(bpy)$_2$(chrysi)]$^{3+}$ (**1**), Δ−[Ru(bpy)$_2$(ddpz)]$^{2+}$ (**2**), and Δ-[Ru(bpy)$_2$(eilatin)]$^{2+}$ (**3**) as prototypical examples of metallo-insertors and intercalators. For all molecules we characterized both insertion and intercalation into a DNA dodecamer via force field based molecular dynamics (MD) and hybrid quantum-classical (QM/MM) MD simulations. A structural analysis of the **1**−**3**/DNA noncovalent adducts reveals that the insertion provokes an untwist of the DNA, an opening of the mG and of the phosphate backbone in proximity of the mismatch, while the intercalation induces smaller changes of these structural parameters. This behavior appears to be correlated with the size of the inserting/intercalating ligand in proximity of the metal coordination site. Moreover, our simulations show that the different selectivity of **1** toward distinct MM types may be correlated with the thermodynamic stability of the MMs in the free DNA and with that of the corresponding insertion adduct. Understanding the factors which tune a specific insertion is of crucial importance for designing specific luminescent probes that selectively recognize MMs, as well as for developing more effective anticancer drugs active in MM repair of deficient cells lines.



1. INTRODUCTION

Maintaining the fidelity of the genome is critical to a cell's survival.[1] DNA mismatches (MMs) occur frequently in the cell as a consequence of polymerase errors, chemicals, UV induced damages, genotoxic radicals, and so forth.[2] These errors, if left uncorrected, may lead to mutations upon DNA replication. In cells, a complex mismatch repair (MMR) machinery localizes and corrects these mutations. However, when deficiencies occur in the MMR machinery, the rate of mutation increases, along with the risk of developing cancers.[2,3] For example, mutations in repair genes are at the basis of a large amount of colon cancers.[4] Moreover, mutations may lead to a larger disposition to diseases' onset and to an altered response to pharmaceuticals. For this reason, there is a great interest in the development of molecules able to selectively target MMs in nucleic acids.[2]

Among the possible candidates, transition metal complexes, which bind and/or react at specific DNA sequences, have been largely studied for their potential applications as nucleic acid probes or chemotherapeutics.[3,5] Complexes, bearing Rh or Ru as metal centers and with one intercalating organic ligand, can bind between two DNA base pairs (bps) into the major groove (MG), provoking minimal distortions to the double helix.[6] These compounds are usually called metallo-intercalators (Figure 1).[2] A prototypical metallo-intercalator is [Ru(bpy)$_2$(dppz)]$^{2+}$ (bpy = 2,2′-bipyridine and ddpz = dipyrido[3,2-*a*:2′,3′-*c*]phenazine) (**2**, Figure 2).[6] This complex has solvatochromic luminescence when it intercalates into the DNA and presents a slight preference toward poly d(AT) over poly d(GC) sequences.[2,7] Although **2** preferentially intercalates into the MG, recent experimental findings suggest that a different binding mode, called insertion, may also be possible.[7] The insertion binding mode (Figure 1)[2] was first observed for Rh complexes, which are known as agents for photoactivated DNA strand cleavage. Among them, [Rh(bpy)$_2$(chrysi)]$^{3+}$ (**1**, Figure 2) was recently discovered to bind to single base MMs from the mG side, acting as π-stacking replacement of the mismatched bases, which are expelled toward the MG.[8,9] **1** binds enantiospecifically to double strand (ds) DNA, as only the Δ enantiomer fits into mG. Although NMR measurements showed that in solution **1** exclusively inserts, crystallographic studies revealed that both insertion and intercalation of this complex are possible.[8]

Received: August 1, 2011
Published: January 30, 2012

   dx.doi.org/10.1021/ic201659v | *Inorg. Chem.* 2012, 51, 2046−2057

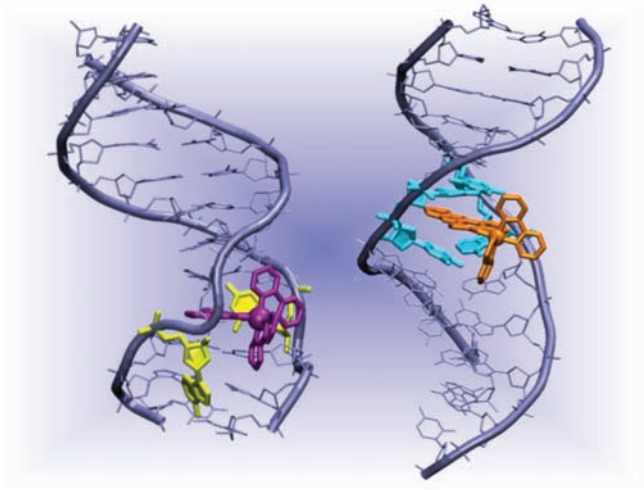


Figure 1. Simplified view of insertion (left) and intercalation (right) binding modes for **1** and **2**, respectively. Mismatched bases, expelled from ds DNA upon insertion, are yellow colored. The base pairs flanking the intercalation site are cyan colored.

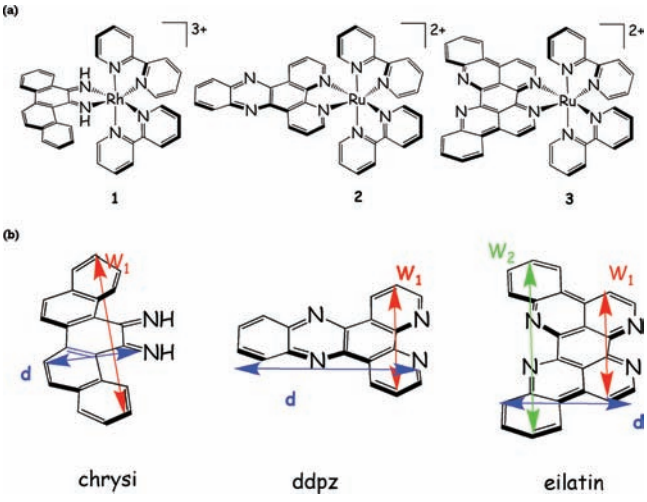


Figure 2. (a) Chemical structures of the rhodium and ruthenium complexes studied in this work. Δ-$[Rh(bpy)_2(chrysi)]^{3+}$, (**1**), Δ-$[Ru(bpy)_2(ddpz)]^{2+}$, (**2**), Δ-$[Ru(bpy)_2(eilatin)]^{2+}$, (**3**). (b) Definition of width (red) and length (blue) for the three inserting/intercalating ligands of **1**−**3**. In **3** a second width is defined (green).

An important property of **1** is its high specificity toward mismatched sites, with one MM targeted over 2700 DNA bps. In addition, **1** possesses a different binding affinity toward distinct MM types. Namely, the binding constants toward CC, AC and AA MMs are of 1.0×10^7 M^{-1}, 2.7×10^6 M^{-1} and 2.9×10^5 M^{-1}, respectively.[10] [These data refer to a MM flanked by GC bases. Although the binding affinity sensibly changes depending on the flanking sequence, it is believed that the relative order should not be affected.] Experimentally, this has been correlated with the different thermodynamic destabilization of the MMs provoked by their impaired capacity of H-bonding and π-stacking.[2] Furthermore, **1** is able to selectively recognize single base bulges and abasic sites, and it may be employed to recognize single nucleotide polymorfims.[2,11] Finally, and more significantly, **1** was demonstrated to have an antiproliferative effect in MMR deficient cell lines in vivo.[1,8] The important characteristics of compound **1** boosted the search for derivatives with enhanced selectivity properties. Among them is the second-generation metallo-insertor $[Rh(bpy)_2(phzi)]^{3+}$ (phzi = benzo[*a*]phenazine-5,6-quinone diimine) (**4**, Supporting Information, Figure S1).[2,12] While showing binding affinities and selectivity toward AC and CC mismaches[2] very similar to those of **1**, **4** is selective also toward CT. However, both **1** and **4** are unable to detect the most thermodynamically stable MMs, containing G nucleobases.

Many other attempts were done to improve the selectivity of inorganic MMs detectors, but the molecular reasons regulating the affinity toward MMs are not completely clear, and this limited the success of this approach. For example, the bulkiness of the intercalating/inserting ligand has been correlated to the preference of MM detectors toward the insertion.[13] However, recent experiments showed that this is not the only factor discriminating between intercalation and insertion.[9,14] For example $[Ru(bpy)_2(eilatin)]^{2+}$ (**3**, Figure 2), which bears the very bulky eilatin ligand, is not MM specific at all, but it can both intercalate or insert into matched and mismatched bps of ds DNA, respectively.[13]

In view of the importance of these molecules as potential chemotherapeutic agents in MMR deficient cells or fluorescent probes to detect DNA MMs, and, because of the lack of a molecular characterization of their adducts with DNA, we performed an extensive computational study of their binding to nucleic acids fragments. Namely, we characterized the structural and energetic properties for the insertion and the intercalation of **1**−**3** into a DNA dodecamer via classical and hybrid quantum/classical (QM/MM) molecular dynamics (MD) simulations.[5] Our results provide an atomistic picture of how these complexes bind to DNA, elucidating similarities and differences among the insertion/intercalations adducts. On the basis of our findings, we formulate hypotheses on which factors may allow to discriminate between insertion and intercalation, and what may tune their different MM affinity. Our study, therefore, could represent a source of useful information for the design of new highly specific diagnostic molecules and chemotherapeutic agents.

2. MATERIALS AND METHODS

2.1. Density Functional Theory (DFT) Calculations. Quantum chemical calculations were performed to geometry optimize the structures of compounds **1**−**4** (Figure 2 and Supporting Information, Figure S1). We started these simulations from the X-ray structure of compound **1**. For compounds **2** and **3**, we replaced Rh with Ru and we built the proper intercalating/inserting ligand, while for **4** we simply built the proper ligand. The CPMD package was employed, and a plane waves (PW) basis set was used to represent the electronic density with an energy cutoff of 70 Ry.[15] Core/valence interactions were described using norm-conserving pseudopotentials of the Martins−Troullier type.[16] Integration of the nonlocal parts of the pseudopotential was obtained via the Kleinman−Bylander[17] scheme for all atoms except rhodium and ruthenium, for which a Gauss−Hermite numerical integration scheme was used. For these latter elements the pseudopotentials were successfully used in other inorganic systems.[18−20] The gradient-corrected Becke exchange and the Perdew correlation functionals (BP)[21,22] were used. Isolated system conditions were applied.[23]

2.2. Classical MD Calculations. Classical all-atom MD simulations were performed on noncovalent adducts between compounds **1**−**3** and the oligonucleotide 5′-d[CGGAAATTCCCG]-3′. They were built using as a template the crystal structure of Δ-$[Rh(bpy)_2(chrysi)]^{3+}$ (**1**) in complex with the same oligonucleotide sequence (PDB Code: 2O1I).[8] This structure contained two terminal and symmetric $[Rh(bpy)_2(chrysi)]^{3+}$ inserted into the mG and one molecule intercalated in the MG.[24]

To build the insertion adducts, we kept in our model only the inorganic molecule **1** inserted in place of the C9-A16 mismatch. [Residues from 5′ to 3′ end on the first strand are numbered from 1 to 12, while 5′ to 3′ residues on the second strand are numbered from 13 to 24.] The other bases, A4-C21, which in the X-ray structure were

dx.doi.org/10.1021/ic201659v | Inorg. Chem. 2012, 51, 2046−2057

expelled from the double helix, were rotated and remodeled in the Watson−Crick configuration, and A4 was replaced by a G nucleobase (to form a matched GC base pair). To build the insertion adducts for **2** and **3** the bpy ligands, the metal and the coordination bonds of the inserting ligand were superimposed with those of **1** in the crystal structure.

To build a model for intercalation of **1** into the same crystallographic DNA dodecamer, instead, we replaced both A4 and A16 with a G nucleobase (and rotated the expelled bases to reestablish their Watson−Crick configurations also in this case). The inorganic complexes **2** and **3** were intercalated into the same DNA sequence with the same procedure followed for insertion.

Each model is labeled as *n*/DNA, where *n* is the number of the compound as it appears in Figure 2. No other label is present for intercalators, while in the case of insertors the MM type is also reported. In practice, **1/DNA_AC** and **1/DNA** stand for **1** inserting to an AC mismatch into mG and intercalating into MG, respectively.

To unveil which factors determine the different selectivity of **1** toward AC, CC, and AA MMs, we also constructed two additional models in which the AC mismatch, present in the X-ray structure, was replaced by CC and AA.

In addition, we considered models of the DNA dodecamers, free of any ligand and in canonical B-DNA conformation, as reference structures to evaluate the distortions induced to ds DNA by the binding of **1**−**3**. With **DNA** we indicate a dodecamer with the same sequence of 2O1I, but with the A bases of the MMs replaced by G. Instead, with the **DNA_AA**, **DNA_AC**, **DNA_CC** we indicate DNA dodecamers with an AA, AC, CC mismatch on nucleobases 16 and 9, respectively. Finally, we simulated the **1/DNA_AA**, **1/DNA_AC**, and **1/DNA_CC** adducts to evaluate the reasons of the different MM selectivity of compound **1**.

The PARMBSC0[25] refinement of the parm99 force field[26] was used for the oligonucleotide moieties, while for the inorganic molecules the parametrization was done following the standard AMBER procedure, as described in ref 27. RESP charges of the inorganic complex were calculated according to the Merz−Kollmann scheme by performing single point calculations on the previously optimized geometry of the complexes to calculate ESP charges. Then, the *resp* module of AMBER9 was used to derive atomic point charges.[28] Ru and Rh van der Waals (vdw) parameters were taken from literature data.[29] All systems were simulated in explicit water solution, and the water molecules were described with the TIP3P potential.[30] Na^+ ions were added to achieve neutrality and were modeled with the Aqvist potential.[31]

In total, we considered 13 different systems whose sizes ranged between 15,000 and 24,000 atoms (See Supporting Information, Table S1). The initial structures were relaxed by performing 1000 steps of steepest descent optimization. Then, 100 ps of MD at constant volume were performed during which the systems were heated up to 300 K. Finally, MD simulations at constant pressure were performed. The temperature was kept at 300 K by applying Langevin forces to all heavy atoms with the Langevin damping constant set to 5 ps^{-1}. The pressure was kept at 1.013 bar using the Nosé-Hoover Langevin piston pressure control. A time step of 2 fs was used, and periodic boundary conditions were applied. Electrostatic interactions were evaluated with the Particle Mesh Ewald (PME) method, with a real space cutoff of 10 Å and a grid spacing of 1 Å per grid point in each dimension. The van der Waals energies were calculated using a smooth cutoff (switching radius 8 Å, cutoff radius 10 Å).

Most systems were simulated for ∼30 ns, with the exception of **1/DNA_AC** and **1/DNA_CC** which were simulated for ∼70 ns, and **1/DNA_AA**, which was simulated for ∼50 ns, respectively, and the free DNA dodecamers, which were simulated for ∼16 ns (Supporting Information, Table S1). All MD simulations (for a total time of 460 ns) were carried out with the AMBER9 package.[28]

2.3. QM/MM MD Calculations. Hybrid simulations were carried out using the CPMD program[15] combined with a classical MD scheme, based on both GROMOS96[32] and AMBER force fields,[26] through the interface developed by Röthlisberger and co-workers.[33] The equilibrium conformations of complexes from classical MD simulations were taken as starting points for QM/MM MD calculations. The systems were divided into two regions: the first, treated at the QM level, comprised the inorganic molecule, while the rest of the system was modeled at the MM level (Supporting Information, Table S1).[34] The QM region was treated at the DFT level with the same computational setup described above in Section 2.1, while the MM region was described with the refined parm99 force field.[25,26] The electrostatic interactions between QM and MM atoms were calculated as described elsewhere.[33]

To evaluate the impact of point charges on the stability of the adducts, we reparametrized the RESP charges of **1** according to a force-matching scheme.[35−37] This was done in the simulations of **1** inserted in different MM types, namely, in **1/DNA_AA**, **1/DNA_AC**, and **1/DNA_CC**. Since, **1/DNA_AC** was simulated with both the RESP charges obtained with the standard AMBER procedure and those obtained from the force-matching scheme, the simulation carried out with the latter charges will be indicated as **1/DNA_AC′**.

2.4. Analysis. Structural parameters of DNA were defined according to the EMBO Workshop on DNA Curvature and Bending[38] and were calculated with the program Curves+.[39] Curves parameters of the insertion complexes were obtained either excluding the A16 nucleobase from the fitting to obtain the intra base pair (bp) parameters and groove parameters or by excluding both the A16 and C9 nucleobases, as they are completely expelled from ds DNA, for the inter-bps parameters. Instead, for the DNA parameters of the intercalation adducts all bases were considered in the fitting.

Dominant conformational clusters and root-mean-square fluctuations (RMSF) were extracted with the *ptraj* module of AMBER.[28] For the cluster analysis we used the average linkage algorithm with a cutoff root-mean-square-deviation (RMSD) of 1.5 Å. Visualization and energetic analysis of the trajectories were performed with the program VMD.[40]

The H-bond content (HBC) for each adduct was defined as a H-bond length per occupancy divided by 3.0 (3.0 Å is the average donor−acceptor distance of a H-bond).

To monitor the flexibility of bpy ligands we define the quantity:

$$d_d = [(C_x - C_z) + (C_y - C_w) + (C_x - C_w)]/3$$

as the time average of the sum of the distances between the *para* carbons (C_p) of the equatorial aromatic rings and the C_p of the axial aromatic rings of the other bpy ligand, plus the distance between C_p of the two equatorial aromatic rings (see Supporting Information, Figure S2).

To estimate the degree of penetration of the inorganic complex inside mG or MG through the quantity d_p, which measures the (time averaged) distance of the compound metal center M to the nitrogens of closest DNA bases and is defined as d_p = [(M-N1@G15) + (M-N3@C10) + (M-N1@A17) + (M-N3@T8)]/4 in case of insertion, and d_p = [(M-N4@T19) + (M-N1@A6) + (M-N1@T18) + (M-N3@T7)]/4, in case of intercalation. Free energies of binding ΔG_{bind} of each system is calculated as:

$$\Delta G_{bind} = G_{com} - (G_{rec} + G_{lig}) \quad (1)$$

where G_{com}, G_{rec}, and G_{lig} are the absolute free energies of complex, DNA, and ligand, respectively, averaged over the equilibrium trajectory. The Molecular Mechanics/Poisson−Boltzmann Surface Area (MM/PBSA) methodology[41−44] was used, according to which the free energy difference can be decomposed as $\Delta G = \Delta E_{MM} + \Delta G_{solv} - T\Delta S_{conf}$, where ΔE_{MM} is the difference in the molecular mechanics energy, ΔG_{solv} is the solvation free energy, and $T\Delta S_{conf}$ is the conformational entropy. The first two terms were calculated with the following equations:

$$\Delta E_{MM} = \Delta E_{bond} + \Delta E_{angle} + \Delta E_{torsion} + \Delta E_{vdw} + \Delta E_{ele} \quad (2)$$

$$\Delta G_{solv} = \Delta G_{solv,PB} + \Delta G_{solv,NP} \quad (3)$$

E_{MM} includes the molecular mechanics energy contributed by the bonded (E_{bond}, E_{angle}, and $E_{torsion}$) and nonbonded (E_{vdw} and E_{ele}) terms of the force field. ΔG_{solv} is the solvation free energy, which has an electrostatic contribution ($\Delta G_{solv,PB}$, evaluated using the Poisson−Boltzmann equation) and a nonpolar one ($\Delta G_{solv,NP} = \gamma\Delta_{SA} + b$,

dx.doi.org/10.1021/ic201659v | Inorg. Chem. 2012, 51, 2046−2057

proportional to the solvent-exposed surface area (Δ_{SA})). The electrostatic solvation free energy was calculated using the *pbsa* module of AMBER, with relative dielectric constants of 1 for solute and 78.3 for the solvent respectively. The electrostatic potential was calculated on a cubic lattice of length equal to 150% of the longest interatomic distance of the solute, using a grid spacing of 0.5 Å. Up to 1000 iteration steps were requested for the convergence of the energy (using the linear form of the PB equation). The surface area entering in the equation for $\Delta G_{solv,NP}$ was calculated using MOLSURF, with γ and b values of 0.00542 kcal/mol Å^2 and 0.92 kcal/mol, respectively, for use with PARSE atomic radii. The solvent probe radius was set to 1.4 Å.

The solute entropy contribution ($-T\Delta S_{conf}$) is composed by a rototranslational term, calculated with classical statistical mechanics, and by a vibrational term, estimated by normal-mode analysis using the NMODE module of AMBER 9.0.

The python script MMPBSA.py was used to perform MM/PBSA calculations on 50 selected snapshots of each system. The single trajectory approach was employed, in which geometries of the ligand and the receptors were also extracted from the trajectory of the complex.

3. RESULTS AND DISCUSSION

Extensive MD simulations were carried out to compare intercalation and insertion adducts of compounds **1**−**3** into ds DNA. In particular, we analyzed the structural or energetic characteristics dictating their preferential DNA binding mode.

3.1. Rhodium Complex. We initially investigated the insertion and the intercalation of **1** into ds DNA as only for this complex a crystal structure is available for both binding poses. Our simulations, in explicit water and at finite temperature, show that **1** can stably insert or intercalate into the DNA.[8] This is remarkable as **1** is not found to intercalate according to spectroscopic measurements.[14] We cannot exclude that on (very much) longer simulations time scales this complex may dissociate from the MG and find a more stable binding mode. Indeed, we found that this mode of binding is thermodynamically unstable compared to insertion (vide infra). To rationalize the differences between the two binding modes we compare in the following the structural and energetic features of **1/DNA_AC** and **1/DNA**.

3.1.1. Insertion. Interestingly, after ∼25 ns of equilibrium trajectory **1/DNA_AC** undergoes a rotation of A16 (already expelled from DNA), which gets fully solvated for about 2 ns, and then sticks back to the DNA forming H-bonds with the mG side (Supporting Information, Figure S3). Because of the slow convergence of the global DNA structural parameters of **1/DNA_AC**, we extended its simulation further by 30 ns. As the rotation of A16 comes along with a small enlargement of the mG and changes in other parameters (vide infra), we divided the trajectory in two parts, labeled as **1/DNA_ACI** and **1/DNA_ACII** for the first 25 and the last 30 ns of the equilibrated trajectory, respectively (Supporting Information, Figure S4).

A cluster analysis performed on each part of the equilibrium trajectory reveals the presence of a main representative cluster in both parts, including 86% and 79% of the sampled conformations, respectively (Table 1; Figure 3 and Supporting Information, Figure S3). Moreover, although the RMSD relative to A16 is 0.56 Å, between the two clusters' representative structures the total RMSD amounts to 3.10 Å, indicating that modest rearrangements of expelled bases can heavily affect the global duplex structure.

The RMSF profile per residue (Figure 4) indicates that, after the rearrangement of A16 (**1/DNA_ACII**), the flexibilities of both the MM region and **1** decrease, suggesting that the inorganic complex becomes more tightly packed inside the mG.

Table 1. Number and Population (%) of Configuration Clusters Extracted from the MD Trajectories of All Systems Studied

system/number of clusters	1	2	3	4	5
1/DNA_ACI	86	1	2	7	4
1/DNA_ACII	14	2	4	79	0
2/DNA_AC	42	1	47	1	10
3/DNA_AC	76	18	6	0	0
1/DNA	5	64	5	26	0
3/DNA	70	19	0	0	1
2/DNA	90	3	2	0	5
DNA	87	9	1	1	2
DNA_AA	95	1	2	0	2
DNA_AC	31	47	1	21	0
DNA_CC	88	6	3	2	1
1/DNA_AA	94	1	1	0	2
1/DNA_AC′	80	8	11	0	0
1/DNA_CC	12	75	9	3	2

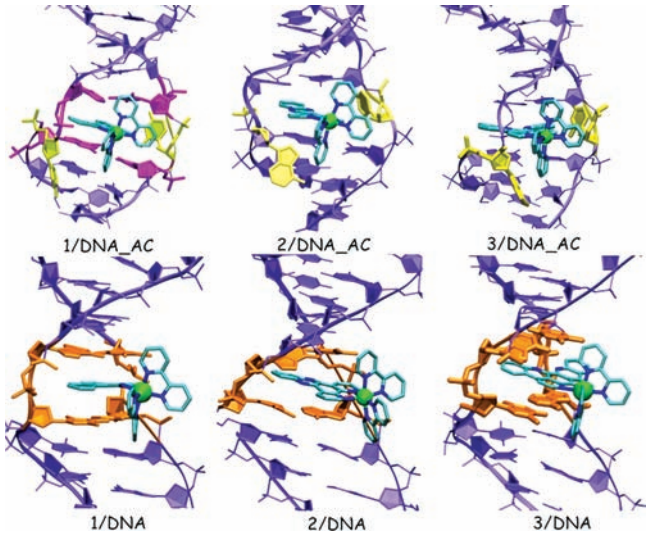


Figure 3. Close view of the most representative clusters for insertion (top) and intercalation (bottom) adducts of **1**−**3** (from left to right). Hydrogen atoms are omitted for clarity. Mismatched bases (C9-A16) expelled from DNA are colored in yellow. Bases flanking the intercalation site Flk1$_{int}$ (A6-T19) and Flk2$_{int}$ (T7-A18) are colored in orange. In **1/DNA_AC** also Flk1$_{ins}$ (T8-A17), top, and Flk2$_{ins}$ (C10-G15), bottom, are colored in magenta.

In both **1/DNA_ACI** and **II** the overall RMSF is slightly larger than in the free oligonucleotide (**DNA_AC**). Thus, the insertion of **1** does not stiffen the DNA duplex.

The analysis of **1/DNA_AC** structural parameters will be focused on global descriptors, such as helical axis bending and mG/MG widths and depths, as well as on the intra-bp (e.g., opening, propeller twist and buckle) and inter-bps local parameters (e.g., shift, slide, twist, roll and tilt). These latter will be analyzed, in particular, around the site of insertion, where, as expected, the largest differences were found with respect to canonical DNA values. As mentioned in the Materials and Methods, since upon insertion the mismatched bp C9-A16 (hereafter MM) is expelled outside the DNA, the inter-bps parameters will be evaluated between bps T8-A17 (hereafter Flk1$_{ins}$) and C10-G15 (hereafter Flk2$_{ins}$) to get useful information about unwinding, rolling, tilting of the DNA helix (Table 2). In addition, the inter-bps parameters will be reported

dx.doi.org/10.1021/ic201659v | *Inorg. Chem.* 2012, 51, 2046−2057

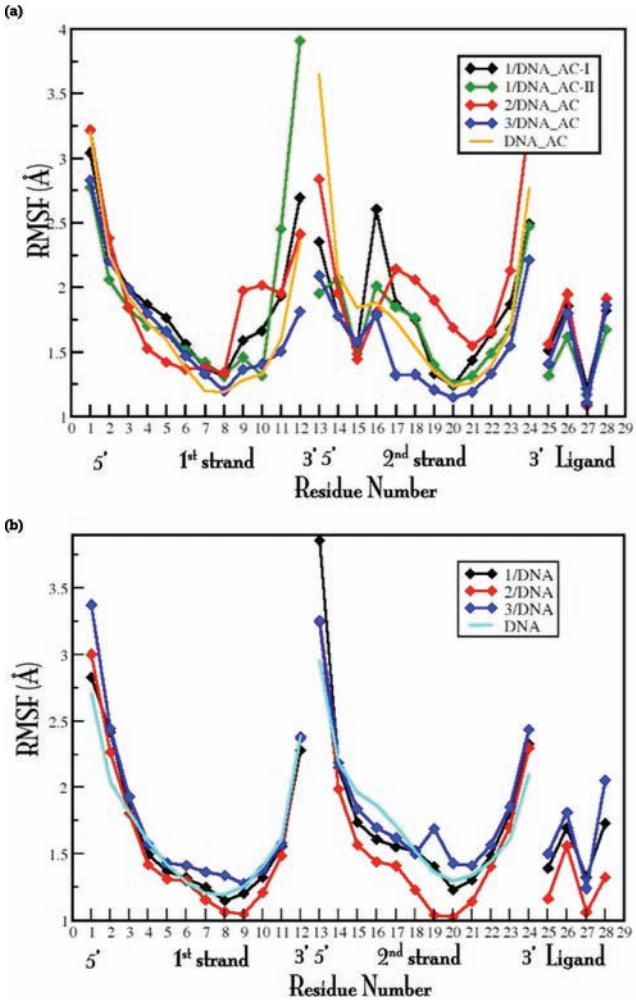


Figure 4. Root mean square fluctuations (RMSF) (Å) per residue for the insertion adducts of **1**−**3** into an AC mismatch and for intercalation adducts are shown in (a) and (b), respectively. The RMSF per residue of the corresponding DNA sequence free of any ligand is also reported for comparison. Residues from 1 to 12 and from 13 to 24 refer to the first and the second DNA strands, respectively; residues from 25 to 29 refer to the inorganic complex.

also for bps steps T7-A18/T8-A17 (Flk1-1_{ins}/$Flk1_{ins}$) and C10-G15/C11-G14 ($Flk2_{ins}$/Flk2+1_{ins}) to verify if and to what level the DNA duplex is able to recover the conformational stress induced by the insertion of the inorganic compounds (Supporting Information, Table S3). These parameters are compared to that of **DNA_AC**.

The intra-bp DNA parameters of the bps flanking the MM show only small differences from canonical DNA, with the propeller of $Flk1_{ins}$ and $Flk2_{ins}$ being mostly affected (Table 2 and Supporting Information, Table S2). Obviously, the differences in the intra-bp parameters become extremely large at the MM site, since the mismatched bases are expelled outside the DNA. Concerning, instead, the inter-bps parameters of $Flk1_{ins}$/$Flk2_{ins}$, the insertion of **1** determines an increase of the tilt. The rise and the roll are, instead, not significantly affected with respect to free DNA (Table 2). A marked unwinding of the DNA is also observed, with the twist angle decreasing by −15 and −19 deg before and after the rotation of A16, respectively. The analysis of the inter-bps parameters of the bps steps surrounding the insertion site shows only small variations in the DNA parameters of Flk1−1_{ins}/$Flk1_{ins}$, while at $Flk2_{ins}$/Flk2+1_{ins} an unwinding of the DNA occurs (Supporting Information, Table S3). This latter parameter increases after the rotation of A16 in **1**/**DNA_AC**.

Furthermore, the insertion is also characterized by a widening of the mG and by a decrease of its depth (Table 3). Finally, the overall axis bending (Supporting Information, Table S4 and Figure S5) presents a marked increase in **1**/**DNA_ACII**.

The unwinding of the DNA duplex at the insertion site corresponds to a large P@C10−P@A17 distance, which has an average value of 18.8 (1.1) Å and of 20.2 (0.9) Å in **1**/**DNA_ACI** and **1**/**DNA_ACII**, respectively, and fluctuates significantly along the simulation (Supporting Information, Figure S6). The opening of mG and of the phosphate backbone depends on the average width and length of chrysi (Figure 2b). These are of 11.43 (0.07) Å and 5.81 (0.06) Å, respectively.

The analysis of the H-bonds shows an interesting behavior of the expelled bases (Supporting Information, Table S5), which is consistent with spectroscopic measurements.[14] Indeed, while C9 is anchored to the phosphate group of T8 for all the simulation, A16 is remarkably mobile and H-bonds with different bases in the mG. A16 H-bonds, initially, with G14 and

Table 2. DNA Intra-bp Parameters[a] Relative to Insertion for T8-A17 ($Flk1_{ins}$) and Inter-bps Parameters[b] Relative to T8-A17, C10-G15 ($Flk1_{ins}$/$Flk2_{ins}$)[c]

	$Flk1_{ins}$			$Flk1_{ins}$/$Flk2_{ins}$					
system	buckle	propeller	opening	shift	slide	rise	tilt	roll	twist
2/DNA_AC	−13 (8)	−1 (7)	6 (4)	0.4 (0.6)	4.9 (0.7)	6.9 (0.4)	0.7 (5.2)	3.2 (6.0)	42.8 (4.7)
3/DNA_AC	−12 (8)	−1 (8)	8 (6)	0.1 (1.0)	1.7 (1.4)	6.7 (0.4)	−0.1 (4.9)	12.7 (6.9)	40.7 (8.8)
1/DNA_ACI	−4 (10)	−7 (10)	6 (6)	1.4 (0.7)	2.5 (1.2)	6.7 (0.4)	7.2(5.4)	8.2 (6.5)	47.8 (8.1)
1/DNA_ACII	−5 (9)	−3 (10)	6 (6)	1.5 (0.7)	3.1 (1.2)	6.5 (0.4)	8.9 (5.4)	7.6 (6.3)	44.4 (7.9)
1/DNA_AA	−11 (10)	−12 (10)	5 (5)	0.6 (0.8)	−1.5 (1.9)	6.4 (0.4)	7.3 (6.0)	24.1 (5.4)	26.1 (7.1)
1/DNA_AC′	−1 (9)	0 (8)	7 (5)	−0.7 (0.7)	−0.4 (0.9)	6.6 (0.3)	1.3 (5.1)	23.4 (6.5)	23.3 (8.4)
1/DNA_CC	1 (9)	2 (8)	7 (5)	−0.7 (0.7)	−0.3 (0.9)	6.5 (0.3)	1.5 (4.7)	25.2 (5.8)	24.1 (8.4)
DNA	1 (12)	−16 (8)	3 (5)	0.1 (0.8)	−1.3 (1.4)	6.9 (0.4)	−1.7 (5.6)	8.4 (8.1)	64.5 (6.7)
DNA_AA	−1 (12)	−20 (9)	3 (7)	−0.0 (0.9)	−0.4 (1.3)	6.7 (0.5)	−0.1 (6.0)	13.8 (8.8)	60.1 (7.2)
DNA_AC	2 (12)	−21 (8)	4 (6)	0.2 (0.8)	−0.7 (1.1)	6.6 (0.4)	−1.5 (6.1)	9.9 (9.3)	62.9 (6.8)
DNA_CC	1 (11)	−21 (8)	5 (6)	0.3 (0.9)	0.3 (0.9)	6.5 (0.5)	0.8 (6.9)	10.7 (10.2)	69.1 (8.3)
X-ray	−8.2	−8.6	−0.8	1.1	2.4	3.2	8.4	3.7	39.7

[a]Buckle (deg), propeller (deg), opening (Å). [b]Shift (Å), slide (Å), rise (Å), tilt (deg), roll (deg) and twist (deg). [c]Standard deviations are reported in parentheses.

dx.doi.org/10.1021/ic201659v | Inorg. Chem. 2012, 51, 2046−2057

Table 3. mG and MG Depths and Widths (Å) for the MM Site (Nucleobase 9 and 16) and the bps Flanking the Insertion Site (T8-17, $Flk1_{ins}$ and C10-G15, $Flk2_{ins}$) (a) and the Intercalation Site (A6-T19, $Flk1_{int}$ and T7-A18, $Flk2_{int}$) (b)[a]

(a)	$Flk1_{ins}$				MM				$Flk2_{ins}$			
system	mW	md	MW	Md	mW	Md	MW	Md	mW	Md	MW	Md
2/DNA_AC	9.3 (1.4)	5.6 (3.1)	10.8 (0.8)	0.7 (1.0)	10.0 (1.0)	1.78 (2.5)	9.9 (1.0)	0.9 (1.7)	9.8 (1.9)	3.5 (2.9)	12.5 (0.8)	−1.9 (0.3)
3/DNA_AC	10.2 (0.8)	2.4 (1.0)	12.9 (1.6)	4.5 (2.4)	10.4 (1.1)	2.4 (2.1)	12.3 (1.3)	3.7 (1.9)	7.4 (0.8)	4.7 (0.8)	12.3 (0.9)	−1.7 (0.4)
1/DNA_ACI	10.4 (0.7)	2.6 (1.2)	12.6 (1.4)	3.1 (2.5)	11.0 (1.5)	2.3 (2.2)	10.3 (1.1)	2.7 (1.5)	8.7 (1.6)	4.3 (2.2)	12.7 (0.7)	−2.0 (0.4)
1/DNA_ACII	10.7 (0.7)	2.6 (1.3)	12.7 (1.3)	1.7 (2.4)	11.4 (1.5)	1.6 (2.2)	11.2 (1.3)	1.2 (1.9)	10.0 (1.5)	1.9 (2.4)	12.7 (0.7)	−2.0 (0.4)
1/DNA_AA	7.6 (1.1)	4.1 (0.7)	17.5 (1.5)	4.3 (1.6)	9.0 (1.0)	2.1 (0.8)	9.3 (1.8)	1.5 (0.9)	10.8 (0.4)	1.6 (0.2)	10.6 (0.20)	0.15 (0.20)
1/DNA_AC′	10.2 (0.5)	1.6 (0.4)	13.6 (2.2)	6.1 (1.4)	10.2 (0.9)	−0.4 (1.3)	11.6 (2.1)	0.3 (2.9)	6.5 (0.8)	3.8 (0.5)	13.1 (1.0)	−1.6 (0.5)
1/DNA_CC	10.3 (0.6)	1.6 (0.5)	13.0 (2.2)	6.9 (1.2)	10.4 (0.8)	−0.5 (1.2)	11.5 (2.0)	0.1 (3.1)	6.1 (1.1)	3.7 (0.5)	13.6 (1.5)	−1.5 (0.5)
DNA_AA	6.5 (1.6)	4.7 (0.8)	12.7 (1.7)	5.8 (1.7)	7.6 (1.8)	4.1 (0.9)	12.9 (1.7)	4.0 (1.5)	8.2 (2.0)	3.9 (1.4)	12.2 (1.2)	−0.3 (1.0)
DNA_AC	6.7 (1.5)	4.9 (0.8)	12.6 (1.8)	6.3 (1.7)	7.3 (1.6)	3.9 (0.7)	12.7 (1.5)	4.2 (1.7)	8.0 (1.5)	4.0 (0.7)	13.6 (1.0)	−0.6 (0.6)
DNA_CC	5.9 (1.8)	4.8 (0.7)	12.3 (1.9)	6.0 (1.9)	6.3 (1.8)	4.2 (0.8)	13.0 (1.5)	4.9 (1.1)	6.8 (1.9)	4.3 (0.6)	14.6 (0.6)	−0.4 (0.4)

(b)	$Flk1_{int}$				$Flk2_{int}$			
system	mW	Md	MW	Md	mW	Md	MW	Md
2/DNA	9.4 (0.9)	3.8 (0.9)	18.1 (2.1)	6.6 (2.3)	8.8 (0.9)	3.4 (0.7)	17.7 (2.0)	6.1 (2.4)
3/DNA	10.2 (1.0)	3.0 (1.6)	19.7 (2.6)	6.92(2.3)	9.8 (1.2)	3.1 (1.3)	19.5 (2.6)	6.1 (2.2)
1/DNA	9.1 (1.4)	3.5 (1.2)	20.6 (2.6)	7.3 (2.3)	9.4 (1.2)	3.1 (1.1)	20.4 (2.5)	7.4 (1.8)
DNA	5.5 (1.4)	4.8 (0.6)	13.1 (2.2)	6.4 (1.8)	5.5 (1.4)	4.7 (0.6)	13.2 (2.1)	6.2 (1.9)

[a]mW, md, MW, Md stand for mG width and depth and MG width and depth, respectively. Standard deviations are reported in parentheses.

dx.doi.org/10.1021/ic201659v | Inorg. Chem. 2012, 51, 2046−2057

Table 4. Intra-bp DNA Structural Parameters[a] Relative to Intercalation for A6-T19 ($Flk1_{int}$) and T7-A18 ($Flk2_{int}$) and Inter-bps DNA Parameters[b]) for $Flk1_{int}/Flk2_{int}$[c]

	$Flk1_{int}$			$Flk2_{int}$			$Flk1_{int}/Flk2_{int}$			
system	buckle	propeller	opening	buckle	propeller	opening	rise	tilt	roll	twist
2/DNA	29 (10)	−13 (8)	3 (5)	0 (1)	0 (1)	0 (1)	7.3 (0.3)	−11 (5)	−4 (6)	24 (4)
3/DNA	2 (9)	−8 (8)	2 (5)	−1 (8)	−1 (7)	4 (5)	6.9 (0.3)	−2 (5)	−5 (7)	4 (10)
1/DNA	0 (9)	0 (8)	1 (5)	−8 (9)	−6 (8)	1 (5)	7.1 (0.4)	0 (5)	−9 (7)	8 (9)
DNA	−1 (10)	−15 (8)	3 (5)	0 (9)	−15 (9)	3 (5)	3.3 (0.3)	0 (4)	−2 (4)	32 (3)
X-ray	10	−8	3	2	6	8	7.3	0	−12	31

[a]Buckle (deg), propeller (deg), opening (Å). [b]Rise (Å), tilt (deg), roll (deg) and twist (deg). [c]Standard deviations are reported in parentheses.

G12, while after A16 rotation, it anchors to G15, determining a slight enlargement of the mG (Supporting Information, Figure S3). Remarkably, no H-bond anchors **1** to DNA (Supporting Information, Table S5).

3.1.2. Intercalation. The main representative cluster of **1/DNA** accounts for 63% of the configurational space (Figure 3). The flexibility of **1/DNA** is similar to that of the "regular" duplex **DNA**: it slightly decreases in proximity of the intercalation site, increasing at the terminal regions of the DNA duplex.

An analysis of the DNA structural parameters of the bps flanking the intercalation site (A6-T19, hereafter $Flk1_{int}$ and T7-A18, hereafter $Flk2_{int}$) of **1/DNA** with those of **DNA** reveals no significant variations for the intra-bp parameters, apart from a decrease in the propeller at $Flk1_{int}$ and in the buckle at $Flk2_{int}$ (Table 4). Concerning, instead, the inter-bps parameters of $Flk1_{int}/Flk2_{int}$, an obvious increase in the rise (by 3.7 Å), along with a decrease of the roll and the twist (−7 and −23 deg, respectively, Table 4) are observed. Also an untwisting of the DNA occurs upon intercalation to a comparable extent than upon insertion (−24 deg).

We have also analyzed the intra-bp of A5-T2 ($Flk\text{-}1_{int}$) and T8-A17 ($Flk2{+}1_{int}$); the buckle features the largest deviations in both cases. In contrast, only small changes occur, with respect to **DNA**, in the inter-bps parameters of $Flk\text{-}1_{int}/Flk1_{int}$ and $Flk2_{int}/Flk2{+}1_{int}$ (Supporting Information, Table S6).

Interestingly, the intercalation provokes a marked increase of MG width (Table 3b), while only a small decrease occurs in the total bending of the DNA duplex (Supporting Information, Table S4).

Here, the P@A6-P@T19 distance is of 18.1 (0.5) Å and the average width and length (Figure 2b) of chrysi are almost identical to insertion (11.38 (0.07) Å and 5.82 (0.06) Å, respectively).

Finally, the H-bond pattern of the bps flanking the intercalation site (A6-T19 and T7-A18) is not altered: two H-bonds are present in each bp for the entire simulation time (Supporting Information, Table S5).

3.2. Ruthenium Complexes. To rationalize differences and similarities among Ru and Rh complexes noncovalently bonded to ds DNA, we carried out the same structural characterization presented for **1** also for the insertion and intercalation adducts of **2** and **3**.

3.2.1. Insertion. A cluster analysis on the trajectory of **2/DNA_AC** reveals two main configuration groups (with populations of 47% and 42% and differing only in the orientation of the terminal bases of the DNA), while **3/DNA_AC** features one main cluster (70%, Table 1). Consistently, the RMSF profile of **3/DNA_AC** is similar to that of **1/DNA_AC** (Figure 4), while **2/DNA_AC** presents RMSF values usually larger than those of the other adducts. A close view of the most representative structures of all adducts is shown in Figure 3.

Concerning the intra-bp structural deformations induced on the DNA by the insertion of **2** and **3**, changes are found in the buckle and in the opening of $Flk1_{ins}$ (Table 2). Small variations are also observable for the intra-bp parameters of $Flk2_{ins}$ (Supporting Information, Table S2).

More relevant are, however, the deformations of the inter-bps parameters of $Flk1_{ins}/Flk2_{ins}$ (Table 2). In all insertion complexes the roll shows a positive value, as the bases are opened toward the mG. At variance with **1/DNA_AC** the tilt is almost zero suggesting that the two bps flanking the MM remain coplanar. Similarly to compound **1**, there is a significant decrease of the twist angle upon insertion of **2** and **3**. In addition, the insertion provokes an increase of the mG width and a decrease of its depth at both $Flk1_{ins}$ and MM (Table 3). The changes of the inter-bps parameters of $Flk1_{ins}/Flk\text{-}1_{ins}$ and of $Flk2_{ins}/Flk2{+}1_{ins}$ (Supporting Information, Table S3) are small, suggesting that the DNA helix is most likely flexible enough to absorb the deformations provoked by the insertion already in the bps flanking the binding site. Here, the only significant difference is an increase of the twist by 8 deg for $Flk1_{ins}/Flk\text{-}1_{ins}$ in **2/DNA_AC**.

The mG has the smallest width in **2/DNA_AC**, while the largest widening occurs in **1/DNA_AC**. In all insertion complexes the DNA bending increases with respect to **DNA_AC**, assuming the largest value in **1/DNA_ACII** and the smallest in **2/DNA_AC** (Supporting Information, Table S4 and Figure S5).

We also examined the local opening of the DNA at the insertion site. The P@C9−P@A16 distances in **2/DNA_AC** and **3/DNA_AC** (19.0 (0.6) Å and 17.7 (0.8) Å, respectively) are both smaller than in **1/DNA_AC**. This aspect is surprising as compound **3** was designed with a very large ligand to increase its affinity for insertion (Figure 2). However, eilatin is characterized by two widths and that affecting the opening of the phosphate backbone is the smallest one, W1 (Figure 2b) (W1 = 9.11 (0.07) Å) (Figure 3). Instead, the largest width, W2 (W2 = 13.1 (0.1) Å), is deeply inserted into mG, almost emerging from the MG side (Supporting Information, Figure S7).[13]

In summary, our data highlight that **1**, which experimentally has the highest propensity to recognize the MMs by inserting into the DNA, provokes the maximal enlargement of the mG, the largest increase of the P@C10−P@A17 distance, and the largest increase of the DNA axis bending, a finding that does not follow simple intuition.

3.2.2. Intercalation. The cluster analysis (Table 1) shows that **2/DNA** and **3/DNA** are characterized by one main representative structural cluster, grouping ∼90% and ∼76% of the sampled conformations, respectively. Moreover, **2/DNA** has an RMSF per residue (Figure 4b) markedly lower than **DNA** (its flexibility is the lowest among the three intercalation

dx.doi.org/10.1021/ic201659v | Inorg. Chem. 2012, 51, 2046−2057

adducts), while **3/DNA** has slightly larger RMSF at the intercalation site.

Concerning the intra-bp DNA structural parameters of $Flk1_{int}$ and $Flk2_{int}$, no relevant variations are seen with respect to **DNA**. The only exception is a large increase in the buckle (28 deg, Table 4) at $Flk1_{int}$ of **2/DNA**.

Instead, for the inter-bps parameters of $Flk1_{int}/Flk2_{int}$, apart the obvious increase in the rise, a decrease of the twist of only 7 deg is observed in **2/DNA** (Table 4), compared with the larger deviations occurring in **3/DNA** and **1/DNA**. No significant changes can be observed in the intra-bp parameters of $Flk1-1_{int}$ and $Flk2+1_{int}$, as well as in the inter-bps parameters of $Flk1-1_{int}/Flk1_{int}$ and $Flk2_{int}/Flk2+1_{int}$ steps (Supporting Information, Table S6).

Compared to other intercalation adducts, **2** induces the smallest opening of the MG, while **1** provokes the largest one (Table 3b). Interestingly, while in **1/DNA** and **3/DNA** a small increase of the total axis bending occurs, due probably to intercalation of a large ligand, no variation is observed in **2/DNA** (Supporting Information, Table S4).

The P@A6-P@T19 distances in **2/DNA** and **3/DNA**, 17.6 (0.1) Å and 16.7 (0.1) Å, respectively, are smaller than that of **1/DNA**. This can be easily explained considering that ddpz has a smaller width than eilatin and chrysi. The widths are, indeed, of 9.11 (0.07) Å, 13.1 (0.1) Å, 9.11 (0.06) Å, and 11.38 (0.07) Å for ddpz, for W2 and W1 of eilatin and for chrysi, respectively. The large size of W2 in eilatin does not allow a complete intercalation. However, because of eilatin's shape, **3** can asymmetrically accommodate into MG, without deforming it substantially (Figure 3 and Supporting Information, Figure S8). In contrast, chrysi, being large and short, needs a larger opening of the MG to π-stack with the flanking bps (Figure 3 and Supporting Information, Figure S8).[13] Thus, the length of the intercalating/inserting ligand is another important parameter affecting the amount of π-stacking interactions with the flanking bps. These lengths are 10.20 (0.06) Å, 7.11 (0.06) Å, 5.82 (0.06) Å for ddpz, eilatin, and chrysi, respectively.

For all adducts no decrease of the HBC between the bps flanking the intercalation site is observed with respect to **DNA** (Supporting Information, Table S5).

3.3. QM/MM MD Simulations. Since the flexibility of the bpy ligands in **1**−**3** may be correlated with their capability of insertion/intercalation, we also performed 3 ps of hybrid QM/MM MD simulations for all investigated adducts. This allows having a more accurate description of the d_d parameter (see Materials and Methods, and Supporting Information, Figure S2).[27,45] In these simulations, the inorganic complex is fully described at the QM level, while the DNA, the solvent, and the counterions are treated at the force field level.[34] Dispersion interactions among the intercalating/inserting ligand and the DNA bps are at the basis of either the insertion or the intercalation. These interactions are not reproduced by most DFT functionals,[46] including that employed in our calculations. However, the partitioning scheme between the QM and the MM regions adopted here allows to fully account for these interactions at the force field level.[47,48]

Interestingly, **1/DNA_AC** and **1/DNA** have d_d of 7.29 (0.10) Å and 7.11 (0.17) Å, respectively. Thus, upon both insertion and intercalation the complex slightly shrinks to fit into the DNA with respect to the gas phase value (7.41 Å). This is particularly evident in case of intercalation.

Instead, d_d is of 7.18 (0.22) Å and 7.15 (0.22) Å in **2/DNA_AC** and **2/DNA**, respectively (vs 7.37 Å in the gas phase). Thus, **2** distorts upon both insertion and intercalation, although less than **1**. A larger difference is found between **3/DNA_AC** and **3/DNA**. Here d_d is of 7.41 (0.17) Å and 7.23 (0.18) Å, respectively. The asymmetric accommodation of eilatin inside the MG imposes a distortion of the bpy to intercalate; instead, when **3** inserts, the average d_d is the same as in the gas phase (7.40 Å).

To analyze the differences between insertion and intercalation we also monitored the average penetration distances of the inorganic complex d_p. This analysis, however, showed no significant differences among the adducts. In addition, the analysis of the RESP charges[35] generated along the QM/MM MD trajectory revealed very similar polarizations of **1**-**3**, independently of the binding mode.

Since the widths of the ligands appear to be important in discriminating between intercalation and insertion, we also measured them along the QM/MM trajectories. For the insertion the widths are of 10.90 (0.08) Å, 8.98 (0.04) Å and 9.06 (0.03) Å, 13.2 (0.1) Å for chrysi, ddpz, and eilatin (W1 and W2), respectively. Instead, the corresponding values for intercalation are 10.70 (0.10) Å, 9.05 (0.06) Å, 9.03 (0.05) Å, 13.30 (0.10) Å. These results indicate that in both binding modes the intercalating/inserting ligands do not distort significantly.

3.4. Selectivity Toward MM Types. **1** was demonstrated to have a different selectivity toward distinct MM types with binding constants of 1.04×10^7 M, 2.71×10^6 M, 2.90×10^5 for the CC, AC, and AA MMs, flanked by the GC bases.[10] In the oligonucleotide employed for our simulations the MM is flanked by TG and GA on the first and the second strands, respectively. It has been experimentally observed that the binding affinity of **1** toward different MMs can markedly change for the different flanking bases;[10] however, we attempted at rationalizing the origin of the observed differences by performing classical MD simulations of **1** inserted into DNA dodecamers, bearing either a AC, a CC, or an AA mismatch (the rest of the DNA, instead, preserved the same sequence). In these simulations, the RESP charges of the inorganic complexes were extracted from QM/MM MD simulations via a force-matching scheme.[35,37,50] In addition, we simulated free DNA duplexes containing of the three different MM types, as reference structures.

DNA_AA and **DNA_CC** are characterized by one main representative cluster (covering 94% and 88% of the sampled space, respectively), while **DNA_AC** has three main clusters (47, 31, and 21%, Table 1). Moreover, **DNA_AA** has in general a lower RMSF profile than the other sequences (Figure 5a), except for the MM region. **DNA_AC** is flexible mostly far from the MM site, while **DNA_CC** is highly flexible at the MM, and in the flanking residues.

A comparison of the intra-bp parameters at $Flk1_{ins}$, MM and $Flk2_{ins}$ (Table 2 and Supporting Information, Table S2) of the different mismatched DNA sequences reveals the largest differences in the propeller and in the opening of MM and $Flk2_{ins}$. The largest opening takes place at MM for **DNA_AA**, consistently with the fact that AA contains the bulkier mismatched bases.

Concerning the inter-bps parameters of steps $Flk1_{ins}$/MM and MM/$Flk2_{ins}$ (Supporting Information, Table S7), it is noticeable that **DNA_AA** is differing from other sequences mostly in the shift and the slide of the bases. Moreover, **DNA_AA** is partially untwisted and overtwisted at $Flk1_{ins}$/MM and at MM/$Flk2_{ins}$, respectively. As expected, **DNA_CC**, bearing the smallest mismatched bases among DNA sequences studied here, presents the smallest changes from canonical B-DNA.

dx.doi.org/10.1021/ic201659v | Inorg. Chem. 2012, 51, 2046−2057

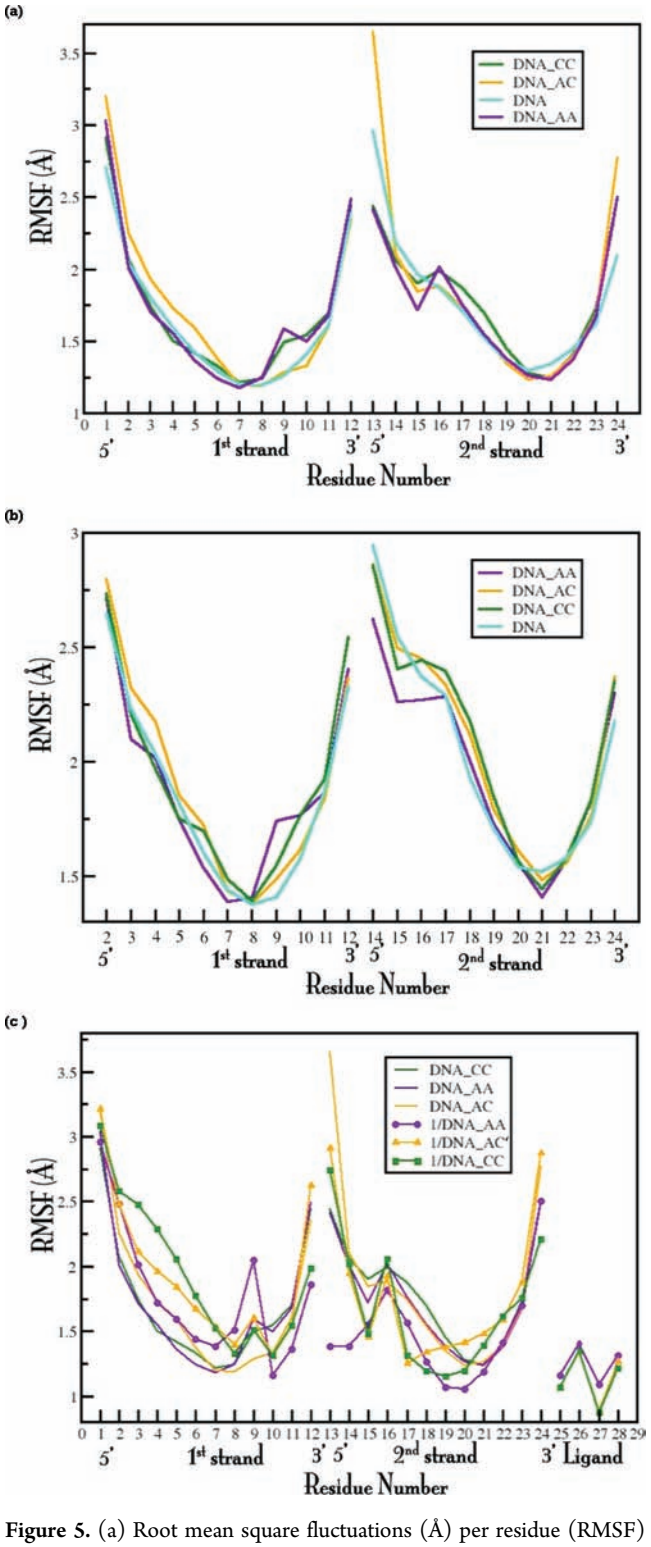


Figure 5. (a) Root mean square fluctuations (Å) per residue (RMSF) for free DNA sequences containing an AC, a AA, and a CC MMs versus a DNA sequence with no mismatch. (b) RMSF of P atoms of the free DNA sequences. (c) RMSF per residue of **1/DNA_AC′**, **1/DNA_AA**, and **1/DNA_CC** adducts. The RMSFs per residue of the corresponding DNA sequences free of any ligand are also reported for comparison. Residues from 1 to 12 and from 13 to 24 refer to the first and the second DNA strands, respectively; residues from 25 to 29 refer to the inorganic complex.

Since upon insertion a large opening of mG and of the phosphate backbone occurs at the MM site, if the selectivity depended on the structural features of the target DNA, one should expect the metallo-insertor to have the largest affinity toward the DNA sequence displaying the widest mG and the largest opening of the phosphate backbone. Instead, no correlation exists between these parameters and the binding affinity. Indeed, the average P@C10−P@A17 distances are quite similar, being 15.3(1.2) Å, 16.3 (0.8) Å, and 16.3 (1.5) Å for **DNA_CC**, **DNA_AA**, and **DNA_AC**, respectively (Supporting Information, Figure S9) and no trend is present in the mG widths (Table 3).

The RMSF of the P atoms (Figure 5b) shows the largest values at the MM of **DNA_CC**. **DNA_AC** and **DNA_AA** are both characterized by lower RMSFs on the strand containing the MM.

An analysis of the electrostatic and vdw interaction energies of the MM with the bps flanking the mismatch and with the whole DNA demonstrates that CC is the most destabilized MM (Supporting Information, Table S8). These data correlate with experimental findings.[10]

The same analyses were performed also for the insertion adducts of **1** on these DNA sequences. These adducts are all characterized by one most representative cluster, grouping 75%, 80%, and 94% of the sampled conformations for **1/DNA_CC**, **1/DNA_AC′**, and **1/DNA_AA**, respectively (Table 1). A comparison between the clusters of the adducts and those of the free DNAs indicates that the binding of the inorganic complexes strongly affects the flexibility the DNA duplexes. The RMSF of the adducts shows that **1/DNA_AA** and **1/DNA_AC′** become more flexible with respect to corresponding free DNAs (Figure 5c). Instead, **1/DNA_CC** is characterized by an overall stabilization with respect to **DNA_CC**, particularly on the second strand. Consistently with what observed for **1/DNA_AC** also **1/DNA_CC** presents a rotation of A16 at 45 ns after equilibration and this base remains solvent exposed for 10 ns. This comes along with a marked opening of the P backbone (Supporting Information, Figure S10).

Concerning the inter-bps DNA structural parameters of $Flk1_{ins}/Flk2_{ins}$, **1/DNA_CC** and **1/DNA_AA** show the largest and the smallest twist angles, respectively. Consistently with previous findings, all complexes are markedly untwisted with respect to their corresponding mismatched free DNA sequences. All other DNA parameters are similar with exception of the tilt and the shift, showing the largest deviations in **1/DNA_AA.** This is consistent with the fact the largest MM induces the largest perturbation to the double helix (Table 2). No relevant changes of the inter-bps parameters can be observed at $Flk\text{-}1_{ins}/Flk1_{ins}$ and $Flk2_{ins}/Flk2{+}1_{ins}$ (Supporting Information, Table S3). The insertion of the drugs induces a small increase of the overall axis bending with respect to the corresponding free DNA sequences. This is particularly evident in **1/DNA_AA**, which is characterized by a very large overall bending (Supporting Information, Table S4).

Once again **1/DNA_CC**, the experimentally favored insertion adduct, is characterized by the largest P@C10−P@A17 average distance (16.4 (1.2) Å, 16.0 (0.6) Å, and 15.7 (0.7) Å for **1/DNA_CC**, **1/DNA_AC′**, and **1/DNA_AA**, respectively), although these differences are very small.

The interactions energies between the complex, the MM, the flanking bps, and the whole DNA markedly favor the **1/DNA_CC** (Supporting Information, Table S8), mainly because of the electrostatic contribution. In fact, besides having a larger HBC than the other adducts, it is also the only one in which

dx.doi.org/10.1021/ic201659v | Inorg. Chem. 2012, 51, 2046−2057

1 H-bonds to the oligonucleotide (Supporting Information, Table S5).

In summary, our data confirm that the binding affinity toward different MM types depends on the thermodynamic destabilization of the MM and the flexibility of the P backbone in the free DNA sequences, as well as on the thermodynamic stability of the corresponding insertion adduct.

3.5. Binding Free Energies. We performed MM-PBSA calculations to evaluate the binding free energies of the insertion and intercalation adducts (Table 5). We have to remark that the calculated absolute values markedly differ from the experimental ones. In fact, experimentally the binding free energy of **1** oscillates between −7 and −10 kcal/mol, depending on the DNA sequence.[10] Instead, the calculated insertion binding free energy of **1** is of ∼−30 kcal/mol. In addition, standard deviations are larger than the free energy differences among the various adducts, limiting a semi-quantitative analysis. We remark that in all cases the conformational entropy contribution is similar, and the differences in free energies among the complexes can be ascribed to the enthalpic contribution. However, MM-PBSA calculations can lead to large errors in the evaluation of the binding free energies,[48] especially when transition metals are involved.[51] Thus, in the present case the results are discussed only to draw a qualitative trend, which supports some of the findings presented above.

Table 5. Binding Free Energies (kcal/mol) According to MM-PBSA Calculations[a]

system	ΔG	ΔH	$T\Delta S$
2/DNA	−23 (8)	−44 (2)	−22 (7)
3/DNA	−26 (8)	−49 (3)	−23 (8)
1/DNA	−25 (10)	−49 (2)	−23 (10)
1/DNA_ACI	−29 (9)	−53 (3)	−24 (9)
1/DNA_ACII	−31 (9)	−53 (3)	−22 (9)
3/DNA_AC	−36 (7)	−59 (4)	−22 (7)
2/DNA_AC	−38 (10)	−60 (4)	−23 (10)
1/DNA_AA	−28 (9)	−52 (3)	−24 (9)
1/DNA_AC′	−28 (9)	−52 (3)	−24 (9)
1/DNA_CC	−37 (9)	−60 (3)	−24 (9)

[a]Standard deviations are reported in parentheses.

According to the MM-PBSA calculations, the conformational changes induced by the rotation of A16 of **1/DNA_AC** correspond to an increase of 2 kcal/mol in the binding strength. In addition, the binding free energies calculated for all adducts of **1** with different MM types correlates with experimental findings as **1/DNA_CC** is the most favored adduct. In general, MM-PBSA calculations suggest that insertion is thermodynamically favored with respect to intercalation. However, we stress again that these energies are characterized by large standard deviations. Thus, the binding free energies are very close to each other and do not allow to discriminate quantitatively which is the most favored insertion or intercalation adduct among the different inorganic complexes. These values are most probably meaningful only when comparing the relative free energy differences of very similar systems (i.e., the same complex and the same binding mode). However, the qualitative picture that emerges from these calculations is that the size of the intercalating/inserting ligand seems to have a small impact on the thermodynamic stability of the adducts. Instead, it is very likely to affect their binding kinetics.[43]

4. SUMMARY AND CONCLUSIONS

We performed a detailed structural characterization of **1**-**3**/**DNA** insertion and intercalation adducts to rationalize the factors discriminating between the two binding modes. Our analysis shows that both insertion and intercalation do not markedly perturb the DNA duplex. The structural changes are mainly concentrated at the insertion/intercalation site.

For insertion the most significant variation involves an unwinding of the DNA between the bps flanking the insertion site. Counterintuitively, the unwinding is more pronounced for the binding of **1**, which, experimentally, was shown to have the highest affinity for insertion. The DNA untwisting comes along with an increase of the mG width and of the P@C10−P@A17 distance, which is the largest upon the insertion of **1.** This appears to be independent of the size of the intercalating/inserting ligand, which is the largest in **3**. However, a visual inspection of the adducts reveals that upon insertion of **3** only the smallest width, W1, affects the opening of the P backbone (Figures 2b, 3, and Supporting Information, Figure S7). Instead, the largest width W2, being farther from the Ru coordination site, does not affect the structure of the DNA.

Several attempts were made to increase the affinity of the inorganic complexes toward the destabilized regions of the DNA. Most of the proposed inorganic complexes had larger (and/or longer) ligands than chrysi far from the metal site and did not show any significant improvement of the selectivity.[13,52] Only complex **4** (Supporting Information, Figure S1), which has the phzi ligand, very similar in size to chrysi (11.6 Å and 6.9 Å and 11.6 Å and 5.9 Å for the gas phase width and depth of phzi and chrysi, respectively calculated by us), presented an increased affinity for CT MMs and a binding affinity similar to **1** for the other MM types.[2,12] Therefore, our results may help to rationalize why the other attempts to increase the selectivity of these inorganic molecules toward MMs failed so far.

Intercalation is characterized by a small increase of the MG width and by a modest unwinding of the DNA. These changes are the smallest for **2**, which experimentally was shown to be the best metallo-intercalator among those studied here, suggesting that a ligand with a small width (and a long depth) may increase the insertion binding affinity.

We attempted at rationalizing the different binding affinities of **1** toward distinct MMs. Consistently with the experiments, this aspect seems to be correlated with the instability of the MM types. Furthermore, this may also be correlated with a larger thermodynamic stability of the resulting insertion adduct. In fact, **1/DNA_CC**, which is experimentally the most favored adduct, presents the largest number of intra- and intermolecular H-bonds.

On the basis of our findings, we propose a possible explanation of the experimentally observed behavior. Since upon insertion the mismatched bp has to be expelled from DNA, the inorganic complex has to pay, most probably, a higher free energy cost to insert than to intercalate. In fact, upon both insertion and intercalation stacking interactions must be destroyed, but the intra-bp H-bonds of the mismatch have to be broken only upon insertion.

In addition to this higher free energy cost, a larger number of matched bps is present in the DNA filaments with respect to the mismatched ones. Thus, octahedral inorganic complexes may initially intercalate into ds DNA. However, intercalation may be effective only for complexes bearing narrow and/or long intercalating ligands. Thus, for complexes like **1**, having a

dx.doi.org/10.1021/ic201659v | Inorg. Chem. 2012, 51, 2046−2057

large and short intercalating/inserting ligand, intercalation will be impaired. Consistently with our hypothesis, spectroscopic measurements reveled that **1** does not intercalate into ds DNA in solution.[14]

In summary, the largest and the shortest is the intercalating ligand the most disfavored will be the intercalation, and the binding equilibrium of the inorganic complex will be shifted toward insertion. These data are not a definitive proof of this hypothesis, and a detailed calculation of the binding free energies and of the free energies barriers for insertion and intercalation, would be required to confirm that.[53]

In conclusion, our study provides for the first time a detailed atomistic picture of insertion and intercalation of octahedral inorganic complexes into ds DNA, providing a detailed comparison of their structural and energetic properties. According to our results we formulate some hypotheses on the factors that may finely tune the preferential binding of these inorganic complexes to DNA, providing useful information for the design of selective anticancer drugs active in MM deficient cell lines as well as of fluorescent probes able to selectively detect DNA defects.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Further details are given in Table S1−S10 and Figures S1−S10. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: alema@sissa.it.

■ ACKNOWLEDGMENTS

The authors thank the CASPUR and the CINECA supercomputing centers for computational resources. A.V.V. acknowledges financial support from "Regione Autonoma della Sardegna" through a Research Fellowship on fundings of the Project "PO Sardegna FSE 2007-2013, L.R.7/2007 Promozione della ricerca scientifica e dell'innovazione tecnologica in Sardegna".

■ REFERENCES

(1) Ernst, R. J.; Song, H.; Barton, J. K. *J. Am. Chem. Soc.* **2009**, *131*, 2359.
(2) Zeglis, B. M.; Pierre, V. C.; Barton, J. K. *Chem. Commun.* **2007**, 4565.
(3) Zeglis, B. M.; Barton, J. K. *Nat. Protoc.* **2007**, *2*, 357.
(4) Zeglis, B. M.; Boland, J. A.; Barton, J. K. *J. Am. Chem. Soc.* **2008**, *130*, 7530.
(5) Spiegel, K.; Magistrato, A. *Org. Biomol. Chem.* **2006**, *4*, 2507.
(6) Erkkila, K. E.; Odom, D. T.; Barton, J. K. *Chem. Rev.* **1999**, *99*, 2777.
(7) Lim, M. H.; Song, H.; Olmon, E. D.; Dervan, E. E.; Barton, J. K. *Inorg. Chem.* **2009**, *48*, 5392.
(8) Pierre, V. C.; Kaiser, J. T.; Barton, J. K. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 429.
(9) Zeglis, B. M.; Pierre, V. C.; Kaiser, J. T.; Barton, J. K. *Biochemistry* **2009**, *48*, 4247.
(10) Jackson, B. A.; Barton, J. K. *Biochemistry* **2000**, *39*, 6176.
(11) Zeglis, B. M.; Boland, J. A.; Barton, J. K. *Biochemistry* **2009**, *48*, 839.
(12) Junicke, H.; Hart, J. R.; Kisko, J.; Glebov, O.; Kirsch, I. R.; Barton, J. K. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3737.
(13) Zeglis, B. M.; Barton, J. K. *Inorg. Chem.* **2008**, *47*, 6452.
(14) Cordier, C.; Pierre, V. C.; Barton, J. K. *J. Am. Chem. Soc.* **2007**, *129*, 12287.
(15) http://www.cpmd.org/.
(16) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 8861.
(17) Kleinman, L.; Bylander, D. M. *Phys. Rev. Lett.* **1982**, *48*, 1425.
(18) Krumper, J. R.; Gerisch, M.; Magistrato, A.; Rothlisberger, U.; Bergman, R. G.; Tilley, T. D. *J. Am. Chem. Soc.* **2004**, *126*, 12492.
(19) Maurer, P.; Magistrato, A.; Rothlisberger, U. *J. Phys. Chem. A* **2004**, *108*, 11494.
(20) Corral, E.; Hotze, A. C. G.; Magistrato, A.; Reedijk, J. *Inorg. Chem.* **2007**, *46*, 6715.
(21) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(22) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.
(23) Barnett, R. N.; Landman, U. *Phys. Rev. B* **1993**, *48*, 2081.
(24) Hart, J. R.; Glebov, O.; Ernst, R. J.; Kirsch, I. R.; Barton, J. K. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 15359.
(25) Orozco, M.; Noy, A.; Perez, A. *Curr. Opin. Struct. Biol.* **2008**, *18*, 185.
(26) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668.
(27) Robertazzi, A.; Vargiu, A. V.; Magistrato, A.; Ruggerone, P.; Carloni, P.; de Hoog, P.; Reedijk, J. *J. Phys. Chem. B* **2009**, *113*, 10881.
(28) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; Debolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1.
(29) Gossens, C.; Tavernelli, I.; Rothlisberger, U. *J. Am. Chem. Soc.* **2008**, *130*, 10921.
(30) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(31) Aqvist, J. *J. Phys. Chem.* **1990**, *94*, 8021.
(32) van Gunsteren, W. F. *GROMOS96*; Hochschulverlag AG der ETH: Zürich, Switzerland, 1996
(33) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941.
(34) Magistrato, A.; Ruggerone, P.; Spiegel, K.; Carloni, P.; Reedijk, J. *J. Phys. Chem. B* **2006**, *110*, 3604.
(35) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Phys. Chem. B* **2002**, *106*, 7300.
(36) Maurer, P.; Laio, A.; Hugosson, H. W.; Colombo, M. C.; Rothlisberger, U. *J. Chem. Theory Comput.* **2007**, *3*, 628.
(37) Spiegel, K.; Magistrato, A.; Maurer, P.; Ruggerone, P.; Rothlisberger, U.; Carloni, P.; Reedijk, J.; Klein, M. L. *J. Comput. Chem.* **2008**, *29*, 38.
(38) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *6*, 655.
(39) Lavery, R.; Moakher, M.; Maddocks, J. H.; Petkeviciute, D.; Zakrzewska, K. *Nucleic Acids Res.* **2009**, *37*, 5917.
(40) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.
(41) Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. *J. Comput. Chem.* **2002**, *23*, 128.
(42) Guvench, O.; MacKerell, A. D. *Curr. Opin. Struct. Biol.* **2009**, *19*, 56.
(43) Gilson, M. K.; Zhou, H. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21.
(44) Vargiu, A. V.; Ruggerone, P.; Magistrato, A.; Carloni, P. *Biophys. J.* **2008**, *94*, 550.
(45) Robertazzi, A.; Magistrato, A.; de Hoog, P.; Carloni, P.; Reedijk, J. *Inorg. Chem.* **2007**, *46*, 5873.
(46) Sgrignani, J.; Franco, D.; Magistrato, A. *Molecules* **2011**, *16*, 442.
(47) Magistrato, A.; Pregosin, P. S.; Albinati, A.; Rothlisberger, U. *Organometallics* **2001**, *20*, 4178.
(48) Magistrato, A.; Togni, A.; Rothlisberger, U. *Organometallics* **2006**, *25*, 1151.
(49) Vargiu, A. V.; Ruggerone, P.; Magistrato, A.; Carloni, P. *J. Phys. Chem. B* **2006**, *110*, 24687.
(50) Spiegel, K.; Magistrato, A.; Carloni, P.; Reedijk, J.; Klein, M. L. *J. Phys. Chem. B* **2007**, *111*, 11873.

(51) Ciancetta, A.; G., S.; Ryde, U. *J. Comput.-Aided Mol. Des.* **2011**, *25*, 729.
(52) Ruba, E.; Hart, J. R.; Barton, J. K. *Inorg. Chem.* **2004**, *43*, 4570.
(53) Vargiu, A. V.; Ruggerone, P.; Magistrato, A.; Carloni, P. *Nucleic Acids Res.* **2008**, *36*, 5910.