Interactions between minor groove binders and DNA: insights from simulations on anthramycin-DNA adducts

Attilio V. Vargiu[a], Paolo Ruggerone[b]*, Alessandra Magistrato[a], and Paolo Carloni[a]*

[a] *SISSA/ISAS and CNR-INFM-DEMOCRITOS Modeling Center, Via Beirut 4, I-34014, Italy.*

[b] *CNR-INFM-SLACS and Physics Department of Cagliari, Cittadella Universitaria I-09042 Monserrato, Italy.*

Abstract

The anticancer drug anthramycin inhibits replication and transcription processes by covalently binding to DNA. Here we use molecular simulations to investigate the interaction between this ligand and the dodecanucleotide d(GCCAACGTTGGC). Starting from the X-ray structure of the covalent adduct anthramycin-d(CCAACGTTG*G) [1], we simulate the non-covalent complexes with the anhydro and hydroxy forms of the drug (i.e prior to the reaction event), along with the covalent complex. Classical molecular dynamics (MD) simulations show that only the hydroxy-form lies in front of the reactive center for the whole simulation (∼20 ns). Instead, the anhydro-form moves inside the minor groove to the nearest base pair after ∼10 ns. This sliding process is associated to an increase of the non-bonding interactions between the drug and the oligonucleotide. Hybrid quantum mechanics/molecular mechanics (QM/MM) calculations suggest that the hydroxy and, even more, the anhydro form, are polarized by the DNA, pointing up an electrostatic role of this latter in the alkylation reaction. We conclude that the binding may be characterized by a multi-step pathway, which is catalyzed by the electric field of the DNA. The implications of these findings for drug design are discussed.

Key words: anthramycin, minor groove binders, anticancer drugs, molecular simulations, QM/MM.

* Corresponding authors.

Introduction

Several anticancer drugs exert their primary action by covalently binding to the minor groove of DNA, originating a cascade of events that lead to cell death [2–4]. These drugs are small organic molecules, which bind preferentially to guanines and, to a minor extent, to adenines [5, 6]. In spite of their low molecular weight, these drugs are quite sequence-selective, recognizing codes $3 \div 7$ base pairs (bps) long [3, 7–9]. Drug binding is believed to proceed by formation of a non-covalent complex [10]:

$$DNA + drug \overset{K_B}{\rightleftharpoons} (DNA \bullet drug)_{noncov}$$

followed by a covalent linkage:

$$(DNA \bullet drug)_{noncov} \overset{k_r}{\rightarrow} (DNA - drug)_{cov}$$

The molecular recognition process is supposed to be driven by (non-specific) non-bonded interactions and sequence-specific structural features (minor groove width and depth, and inherent sequence flexibility), rather than by specific H-bond patterns [5]. However, the molecular details of this process are not yet fully understood, due also to the modest available experimental information on minor groove binders [5].

Anthramycin [1] (Chart I) is one of the very few organic minor groove binders for which the X-ray structure of the covalent complex with DNA is available [1, 11–16]. This natural antibiotic belongs to the family of pyrrol[1,4]benzodiazepines (**PBD**'s), cytotoxins [17, 18] featuring a powerful antitumoral activity [19–21]. All **PBD**'s share a condensed three ring moiety featuring a right-handed twist between the phenol and pyrrol rings (respectively A and C in Chart I), which allows them to snugly fit into the minor groove of B-DNA. They covalently bind to the exocyclic amino group of guanine through the $C11$ carbon [1], showing a modest sequence selectivity for guanines embedded in Pu-G-Pu sequences (namely AGA-5'>AGG-5'≈GGA-5'>GGG-5' for anthramycin) [1, 7, 8, 10, 22, 23].

Although cardiotoxic [19, 24], anthramycin is largely used as template in drug design, and recently a number of its derivatives [25–29] have shown improved antitumoral activity (in particular SJG-136 [30, 31] is currently in Phase I clinical trials in both UK and USA). **PBD**'s have also being used as components of a gene targeting strategy [3, 32] aiming at

designing molecules able to span more than 12 bps of DNA, in order to target a specific cancer-related gene. Thus, anthramycin is a prototype of minor groove binders ideal to elucidate both the factors affecting molecular recognition and binding. Moreover, anthramycin represents a good candidate for MD-based studies: the very same chemical identity of the drug in the X-ray, *in vitro* and *in vivo* adducts [1, 33], ensures the relevance of the results obtained from atomistic simulations.

Here we use molecular simulation tools to investigate the binding of anthramycin to the DNA. Classical MD simulations are carried out on two non-covalent complexes between a 12-mer oligonucleotide and the hydroxy (**ANT**) and anhydro (**IMI**) forms of the drug (Chart I) in aqueous solution, along with the covalent complex. Theoretical studies on PBD-oligonucleotide adducts have been performed [24, 34–36], without inclusion of explicit waters and counterions. However, an accurate theoretical treatment of solvent effects, along with long-range interactions, is necessary for proper MD modeling of nucleic acids (see [37] and references therein). In the non-covalent complexes the drug are initially positioned in front of the reactive guanine. Our results show that the binding of the drug does not induce DNA bending. Only covalent bonding causes significant distortions in both the drug and the oligonucleotide (at a local level). While hydroxy-anthramycin stays in front of the reactive site for the whole simulation, anhydro-anthramycin slides inside the minor groove towards another site. This is consistent with a stabilization of the non-covalent adducts mainly due to hydrophobic interactions. In addition, hybrid QM/MM simulations suggest that the polarization effects of the DNA may play an important catalytic role in the alkylation.

Systems and Methods

Classical MD simulations

The calculations were based on the X-ray structure of the anthramycin covalently linked to the oligonucleotide d(CCAACGTTG*G)d(CCAACGTTG*G) [1] (**ANT-DNA**). The hydroxy-anthramycin•DNA (**ANT•DNA**) and anhydro-anthramycin•DNA (**IMI•DNA**) non-covalent complexes have been obtained by: *i)* cutting the bond between the $C11$ of the drug and the $N2$ of deoxyguanine, and manually pulling out the drug until distance $N2$-$C11$ was $\sim 3.5\,$Å; *ii)* adding the OH group on $C11$ in the case of **ANT•DNA**, or removing an

H atom from *N*10 in the case of **IMI•DNA** (see Chart I); *iii)* optimizing *in vacuo* the geometries of the drugs at the B3LYP-6-31G* [38, 39] level of calculations with the program GAUSSIAN [40]. Two CG base pair (bp) were added at each end of the oligonucleotides, in order to stabilize its terminal parts. The following reference systems were also built: *i)* a 12-mer oligonucleotide in the B-DNA conformation with sequence GCCAACGTTGGC (**DNA**), which has been constructed with the *nucgen* module of the AMBER package [41]; *ii)* hydroxy-anthramycin (**ANT**) and anhydro-anthramycin (**IMI**) in water.

ANT-DNA, **ANT•DNA**, **IMI•DNA** and **DNA** were neutralized adding 22 sodium counter-ions, and immersed in boxes of $58 \times 72 \times 58\,\text{Å}^3$, $60 \times 62 \times 66\,\text{Å}^3$, $64 \times 59 \times 66\,\text{Å}^3$, $56 \times 55 \times 56\,\text{Å}^3$ containing about 6000, 7700, 8000, 5500 water molecules respectively. The resulting systems contained ∼18800, ∼23900, ∼24800, ∼17300 atoms respectively.

Neutral molecules **ANT** and **IMI** were solvated with 1900 and 2900 water molecules within boxes of $45 \times 36 \times 35\,\text{Å}^3$ and $47 \times 47 \times 46\,\text{Å}^3$ (∼5700 and ∼9000 atoms respectively).

The AMBER parm99 [42, 43], TIP3P [44] and Aaqvist [45] force field parameters were used for the oligonucleotide, the water and the sodium counter ions respectively. The structural parameters of the drugs were taken from the gaff [46] database, or constructed with the *parmcal* module of the AMBER package [41] (TabS. 1), while RESP [47] charges were obtained using the *resp* utility of AMBER, after minimization of the electronic structure with GAUSSIAN [40].

Periodic boundary conditions were applied, and the electrostatic interactions were calculated with the particle-mesh Ewald method (PME) [48], using a 12 Å cutoff, as for the van der Waals interactions. The minimum distance between the solute and its images was 24 Å. NPT simulations at $310K$ and $1\,atm$ were performed using the Nosé-Hoover [49, 50] thermostat and the Parrinello-Rahman [51] pressure-coupling scheme. A time step of $1.5\,fs$ was set for all the simulations. H-bond lengths were constrained using the lincs algorithm [52], and the translational and rotational motions of the center of mass of the solute were removed every 25 MD steps. We performed a simulation of $21\,ns$ for **ANT-DNA**, $19\,ns$ for **ANT•DNA** and **IMI•DNA**, $21\,ns$ for the reference oligonucleotide **DNA**, and $6\,ns$, $10\,ns$ respectively for **ANT** and **IMI**, using the GROMACS package [53, 54].

After structural optimization (500 and 24500 steps of steepest descent and conjugate gradients protocols respectively), all the systems were gradually heated up to $310K$ by NVT

MD, imposing a harmonic constraint (k=200 $kcal/nm$) to the solute. As last step preceding the extensive dynamics, 200 ps of NPT runs were carried out with a progressive removal of the constraints.

The structural parameters of DNA were calculated with the program Curves 5.2 [55]. The minor groove widths were defined as the distances between sugar $C4'$ atoms, subtracted by two carbon vdW radii [1]. A propeller twist and a buckle angle between the pyrrol and the phenol rings of the drug were also defined, according to Ref. [1]. The relative energies of non-bonded interactions were estimated using the terms in the AMBER force field [41–43]. The molecular Solvent Accessible Surface Area (SASA) [56] of various molecules was calculated with GROMACS [54], using a probe radius of 1.4 Å. The variation in the hydrophobic SASA was evaluated as $\Delta\text{SASA} = \text{SASA}_{complex} - (\text{SASA}_{DNA} + \text{SASA}_{drug})$, with the terms in brackets referring to the isolated subsystems. All contributions to the binding are calculated limitedly to the bps from $G7$-$C18$....$G11$-$C14$, directly involved in the binding and covering the sliding path of anhydro-anthramycin (the interaction of the drug with the rest of the oligonucleotide amounts to less than 5% of the total).

The images showed in the paper are produced with the programs VMD [57] and gOpenMol [58, 59].

QM/MM calculations

The QM/MM scheme developed by Rothlisberger and coworkers [60] and implemented in the program CPMD [61] was used, partitioning all the drug-DNA adducts in two regions. One (QM), comprised the drug and the nucleoside of the reactive guanine $G10$ ($56 \div 59$ atoms), and was treated at the DFT level [62, 63], using the BLYP [38, 39] exchange-correlation functional. The guanine was cut at the $C1'$ carbon, and the valence of this atom was saturated with two capping hydrogens [64]. The remaining part of the system (MM) was treated using the AMBER [41–43] effective potential.

Quantum calculations were performed using plane waves decomposition up to a kinetic energy of 70 Ry, along with pseudo-potentials of the Martins-Trouillers type [65]. The QM part is contained within a rectangular box of $6000 \div 7000$ Å^3, and was treated as an isolated system [66]. The electrostatic interactions between QM and MM atoms were calculated

using a hierarchical scheme [60] as described elsewhere [67]. Starting structures were taken from the last MD snapshots, and annealing was performed in order to relax the structures at a temperature close to $0\,K$. A time step of $5\,au$ ($\sim 0.12\,fs$) and a fictitious electronic mass of $600\,au$ were used. NVT simulations were carried out by coupling the systems to a Nosé-Hoover thermostat [49, 50].

Boys orbitals (BO's) [68] were calculated during the QM/MM simulations, and, based on their position, we have estimated the ionicity of selected bonds. To this aim, following Ref. [69], we defined a bond ionicity of the A-B bond as $\mathrm{BI}_{AB} = \frac{|\overrightarrow{d_A}|\cos\theta}{|\overrightarrow{d_{AB}}|}$, where $\overrightarrow{d_A}$ is the position vector of the BO along A-B with respect to A, $\overrightarrow{d_{AB}}$ is the position vector of B with respect to A, and θ is the angle between the two position vectors.

Results and Discussion

We initially analyze the results from classical MD, focusing on the structural and energetic features of the *non-covalent* complexes. Comparison is also made with separate simulations of only the DNA or the drugs in solution. The effect of the surrounding on the electronic properties of the drug-guanosine subsystem is investigated by QM/MM calculations. Finally the data from the simulation of the *covalent complex* are analyzed and compared with the X-ray structure.

Non-covalent complexes

Classical MD

In **ANT•DNA** the drug oscillates around the initial binding site for the whole simulation time (see inset in Fig. 1). Instead in **IMI•DNA**, after $10\,ns$, the drug moves along the minor groove, towards the central part of the DNA fragment (from site **1** - configuration **IMI•DNA 1** - to **2** - configuration **IMI•DNA 2** - in Fig. 2). This movement covers approximately a distance of one base pair (bp) step, so that $C11_{Ant}$ sits in the vicinity of $T9$ in the configuration **2**. This shift of the drug determines the sudden increase of its RMSD in Fig. 1.

The DNA structures reach their average conformations after $\sim 8\,ns$. No relevant structural differences are observed with respect to the initial solvated structures (Fig. 1) and to that of **DNA** (Fig. 3). In particular, neither the global helical axis (Fig. 3) nor the local DNA base/base and base-pair/base-pair (Fig. 4) parameters show relevant variations upon non-covalent binding. The only significant structural deviation is a narrowing of the minor groove backbone around the ligand region in **IMI•DNA 2** (also compared to the same system in **1** and to **ANT•DNA** [75], see Figs. 3 and $4a$).

The drugs exhibit a slight distortion upon binding, as shown by the average values of the propeller and twist angles reported in Tab. I. The large values of the standard deviation are indicative of a drug flexibility. Nevertheless, within these deviations, it seems that the twisting is anti-correlated with the degree of binding, with the only exception of **IMI•DNA 2** (*vide infra*). The buckle instead does not change significantly for both the form of the drugs.

Drug-DNA interactions. In order to qualitatively compare the binding to DNA of the two forms of the drug, we use the non-bonded terms of the AMBER force field [42, 43], that give an approximate estimate of the interaction energies. In both the systems the LJ interactions account for ∼70-80% of the (non-bonded) energy, while electrostatics plays a minor role (Tab. II). Note that the LJ (Coulomb) interaction is slightly stronger (weaker) in **IMI•DNA** than in **ANT•DNA**: this is due to the presence in the latter molecule of the $OH11$ group, that prevents the drug to be as well "packed" by the DNA backbone as in **IMI•DNA**. In addition, the relative decrease of the SASA [56] upon formation of the adduct is larger in **IMI•DNA** than in **ANT•DNA** (Tab. III), meaning a larger relative gain in the free energy upon formation of the adduct . Thus, **ANT** establishes a more favorable non-bonded interaction with the DNA frame, with respect to the hydroxy form of the drug.

The H-bond pattern appears to be quite poor: few H-bonds are present, involving the nitrogens of $G10$, $G11$, $A16$, and $O4'_{G11}$ on the oligonucleotide, $O9$, $O11$ and $N10$ on the drug (see Fig. 5 and Tab. V). No water molecules are present between the reactive guanine $G10$ and the drugs. However, the presence of the $OH11$ group in **ANT•DNA** allows the entrance of water(s) in the region between the drug and the DNA backbone (Tab. V). Interestingly, the H-bonds between the solvent and the systems **IMI**, **ANT**, **DNA**

(calculated only for the atoms at the drug-DNA interface) are only partially replaced in the non-covalent complexes (Tabs. IV and V).

All these results suggest that the formation of the adducts have an entropic contribution, and their stabilization is mainly due to non-bonded interactions.

Anhydro-anthramycin sliding. Here we analyze the structural and energetic rearrangements undergone by the **IMI•DNA** system upon sliding of the drug. In **2** the structural parameters of the oligonucleotide are similar to those of canonical B-DNA (Fig. 4*b, c* and *d*), and the minor groove, in particular in the central region of the oligonucleotide, becomes narrower (Fig. 4*a*). Although the number of drug-DNA H-bonds is reduced in going from **1** to **2** (two *vs* three in Tab. IV), the overall interaction increases by few $kcal/mol$, as can be seen in Fig. 6. This mirrors a better accommodation of the drug into the DNA minor groove after the sliding, as evidenced also by the decrease of its twist angle (29 *vs* 22 deg in Tab. I), and by an increase of 4% in the hydrophobic contact area (26% *vs* 30% in Tab. III). Our data indicates that the change of entropy associated to the sliding process is very small. In fact: *(i)* the first water shell around the drug and the "binding bases" do not change, as can be seen from Tab. IV and by a calculation of the radial distribution function of waters around the binding region (data not shown); *(ii)* the conformational entropy of the DNA (calculated with the Schlitter's formula [71]) is unchanged. All these findings suggest that the sliding of the drug is mainly due to a gain in non-specific interactions with the DNA frame.

Reactant Polarization

In this section, we investigate the electrostatic effect of the biomolecular frame on the electronic structure of the reactants (the drug and its guanosine target) by QM/MM calculations. In particular we focus on the atoms $N2_{G10}$ and $C11_{drug}$, directly involved in the covalent reaction. The QM/MM scheme is used here to dissect the source of polarization into various contributions, namely the electrostatic effects due to the DNA and the solvent. We also assess the amount of reciprocal polarization of the reactants by performing additional QM calculations *in vacuo*. To simplify the discussion we introduce the following notation: the term *with bioframe* refers to the complete QM/MM calculations, *without bioframe* to the calculations with all the MM charges switched off, and *without solvent* to the calculations

with the charges of the solvent switched off.

Comparison of the electrostatic potential *with bioframe* and *without bioframe*, calculated on the QM/MM minimized structures [76], shows that the biomolecular frame modifies the electric field in the region between the drug and the minor groove, in particular between $N2_{G10}$ and $C11_{drug}$ (see Fig. 7). However, while in **IMI•DNA** a field directed from $N2_{G10}$ to $C11_{drug}$ is created, which might assist the alkylation reaction by rendering the latter atom more electrophilic, in **ANT•DNA** the largest effect occurs in proximity of the amino group of $G10$. Inspection of several QM/MM snapshots confirms that these features are present also during the dynamics. These effects can be ascribed almost completely to the DNA electric field, as confirmed by the identical results obtained from calculations *without solvent* and *with bioframe* (data not shown).

To have a direct picture of the effect of the electrostatic field on the electronic structure of the reactants, we calculate the BI's, averaged along the QM/MM trajectory (Tab. VI). BI's indicate bond polarization, and are particularly useful to quantify the effect of an external electric field. Note that also little differences between these values are significant, because of the restricted range of values of the BO projection onto the bond [69]. The polarization of drug and guanine bonds in both the systems parallels the differences in the potential map. In particular, a correlation between *position* and *broadness* of the maximum ΔV and bond polarization is detectable from our calculations. In fact, in **IMI•DNA** the $N10$-$C11$ π bond is the most polarized upon inclusion of the DNA electric field [77], as shown by a 3% shift of the BI's, while in **ANT•DNA** the variation of BI's indicates an easier polarization of the $N2_{G10}$ lone pair. Note that the latter BI, along with the one associated with the $N2$-$C2$ bond, undergoes a noticeable shift under the electric field of the drugs.

Finally, in Tab. VII we collect the energy gaps between the orbitals showing the largest projection on the reactive atoms of drug and guanine. The presence of the biomolecular frame reduces the gaps, and provokes a shift of the orbitals localized on $N2$ and $C11$ (Tab. VI).

Clearly these results provide only a qualitative picture of the catalytic effect of DNA on drugs in the formation of the covalent complex. However, a comparison of the data obtained *with* and *without bioframe* shows that a single nucleobase may not be able to create the appropriate conditions for binding. This result agrees with experiments, which have shown

a net preference of anthramycin to bind double-stranded DNA (dsDNA), rather than RNA, single-stranded DNA, single nucleobases, or proteins [1, 19, 70].

Covalent Complex

Since the X-ray structure of the covalent complex is available, a comparison between our calculations and the experimental data allows us to validate the computational setup, and provides a reference for the non-covalent and free systems. The structural features and the B-factors (Fig. 8) of the complex agree well with the experimental ones, except around the region $C24$-$G1$....$T20$-$A5$, from where we removed a second drug (bonded to $G22$ in the crystal structure).

Four persistent H-bonds are present during the simulation: the OH group on the phenol ring is the most involved, forming strong bonds with both NH_{G11} and O_{C15}. The average number of direct H-bonds between drug and DNA was 1.7, with large fluctuations (from 1 to 4). Three H-bonds (the first in Tab. IV) are also present in the X-ray structure, that counts a total of six H-bonds (Tab. VIII). This discrepancy is limited to the tail of the drug (Tab. VIII), and might be due to the effect of packing forces in the crystallographic structure, which are particularly important at the ends of the oligonucleotide [5]. Additionally, in our simulation one or more waters bridge the amide group $HN15_{Ant}$ and $O2_{T9}$, $O4'_{G10}$, $N3_{A17}$ atoms of the DNA: these waters push the drug tail towards the strand containing $A17$, in such a way that $HN15_{Ant}$ links to N_{A17} only, and looses the contact with $O2_{T9}$ and $O4'_{G10}$. Note that these bridging waters are not present in the X-ray structure. According to these results the role of the H-bond pattern connecting the drug and the DNA have a minor role in stabilization, as compared to Ref. [1]. However, there is a "secondary" stabilization coming from the water mediated bonds, that are absent in the experimental complex.

Although **ANT-DNA** is mostly stabilized by means of the covalent bond, we also calculated electrostatic and LJ energies between the drug and the oligonucleotide, in order to have a comparison with the non-covalent complexes. The electrostatic interactions are comparable with the LJ ones (Tab. II): this is due in part to the changes in the RESP charges of the $G10$ base (see TabS. 2), in part to the shortening of the distances between the charged atoms upon formation of the covalent bond.

Summary and Conclusions

We have investigated by classical and hybrid QM/MM molecular simulations the interaction between anthramycin and DNA. Classical MD simulations show that both drug and DNA undergo significant structural changes only upon covalent binding, while only minor distortions occur in the non-covalent complexes. The DNA helical axis does not show appreciable bending upon non-covalent or covalent binding, in agreement with experiments [1, 20]. Instead, the buckle and propeller twist of the drug, along with the *local* structural parameters of DNA, progressively increase in going from the bulk, to the non-covalent adducts, and to the covalent complex. These findings are consistent with an "induced fit mechanism" of binding [23], in which the transition state for the covalent bond is stabilized through conformational changes of both the DNA and the drug.

The calculation of the energetics of the systems reveals that the non-covalent complexes are mainly stabilized by dispersion interactions. This can be a reason for the modest sequence selectivity of these drugs [3, 7–9].

A remarkable difference is observed in dynamical properties of the two forms of the drug. After $10\,ns$ the anhydro form slides within the minor groove to the adjacent bp (from **1** to **2** in Fig. 2). Anhydro-anthramycin is therefore able to sit in two different sites of the oligonucleotide, consistently with a picture of non-specific drug-DNA interactions. In contrast, hydroxy-anthramycin remains in its the binding site during the whole length of our simulation. This may suggest that: *i)* the hydroxy form is the only reactive species [19, 72]; *ii)* the anhydro form is also reactive, but the alkylation may occur after the shuffling of the drug between different binding sites. Free energy calculations are required to further clarify this issue. It is worthy to note that shuffling or translocation of minor groove binders have been already detected experimentally [73, 74]. However, to the best of our knowledge, this is the first time that such a behavior is observed for minor groove alkylating drugs in a simulation.

QM/MM calculations suggest that the dsDNA may assist the alkylation reaction by altering the electrostatic potential between the reactants. In particular, in **IMI•DNA** the electric field due to DNA charges renders the drug reactive carbon more electrophilic. In **ANT•DNA** instead the lone pair on the exocyclic amino-nitrogen of guanine is the most sensitive to the surrounding. These polarization effect disappears upon removal of the elec-

tric field of dsDNA. On the other hand the drugs, in particular **ANT**, polarize significantly the C-N bond and the nitrogen lone pair of the guanosine, even if the standard deviations are much higher in this case. Therefore, our results point to a major catalytic role of dsDNA, consistently with experimental suggestions [1, 19, 70].

In conclusion, we found that entropic effects and electrostatic interactions are key factors for both the molecular recognition and the binding of anthramycin to DNA. All these results provide a valid complement to the available experimental data on molecular recognition between anthramycin and DNA, and more in general between covalent minor groove binders and DNA. In particular, this work points up that the formation of the reactants complex might require a significant activation free energy. This could be an important issue to be taken into account for the design of drugs similar to the one investigated here.

Acknowledgments

The authors thank Katrin Spiegel for her precious help. We are also grateful to Stefano Piana and Simone Raugei for the useful suggestions. Computational resources have been granted by CINECA (INFM grants) and CASPUR. This project represents a scientific collaboration between Trieste and Cagliari units of the INFM-Democritos.

---

[1] Kopka, M. L., Goodsell, D. S., Bailakov, I., Grzeskowiak, K.,Cascio, D., and Dickerson, R.E. (1994) Crystal structure of a covalent DNA-drug adduct: anthramycin bound to C-C-A-A-C-G-T-T-G-G and a molecular explanation of specificity. *Biochemistry*, **33**, 13593-13610 .

[2] Arimondo, P. B., and Hélène, C. (2001) Design of New Anti-Cancer Agents Based on Topoisomerase Poisons Targeted to Specific DNA Sequences. *Curr. Med. Chem.*, **1**, 219-235.

[3] Baraldi, P. G., Bovero A., Fruttarolo F., Preti D., Tabrizi M.A., Pavani M.G., and Romagnoli R. (2004) DNA minor groove binders as potential antitumor and antimicrobial agents. *Med. Res. Rev.*, **24**, 475-528.

[4] Hurley, L. H. (2002) DNA and its associated processes as targets for cancer therapy. *Nat. Rev. Cancer*, **2**, 188-200.

[5] Neidle, S. (1997) Crystallographic insights into DNA minor groove recognition by drug. *Biopolymers*, **44**, 105-121.

[6] Neidle, S. (2002) Nucleic acid structure and recognition. Oxford University Press, Oxford, NY.

[7] Hertzberg, R. P., Hecht, S. M., Reynolds, V. L., Molineux, I. J., and Hurley, L. H. (1986) DNA sequence specificity of the pyrrolo[1,4]benzodiazepine antitumor antibiotics. Methidiumpropyl-EDTA-iron(II) footprinting analysis of DNA binding sites for anthramycin and related drugs. *Biochemistry*, **25**, 1249-1258.

[8] Hurley, L. H., Reck, T., Thurston, D. E., and Langley, D. R. (1988) Pyrrolo[1,4]benzodiazepine antitumor antibiotics: relationship of DNA alkylation and sequence specificity to the biological activity of natural and synthetic compounds. *Chem. Res. Toxicol.*, **1**, 258-268.

[9] Denny, W. A. (2001) DNA minor groove alkylating agents. *Curr. Med. Chem.,* **8**, 533-544.

[10] Warpehoski, M. A., and Hurley, L. H. (1988) Sequence selectivity of DNA covalent modification. *Chem. Res. Toxicol.,* **1**, 315-333.

[11] Gniazdowski, M., Denny, W. A., Nelson, S. M., and Czyz, M. (2005) Effects of anticancer drugs on transcription factor-DNA interactions. *Expert. Opin. Ther. Tar.*, **9**, 471-489.

[12] Zegar, I. S., Setayesh, F. R., DeCorte, B. L., Harris, C. M., Harris, T. M., and Stone, M. P. (1996) Styrene oxide adducts in an oligodeoxynucleotide containing the human N-ras codon 12 sequence: structural refinement of the minor groove R(12,2)- and S(12,2)-alpha-(N2-guanyl) stereoisomers from 1H NMR. *Biochemistry*, **35**, 4334-4348.

[13] Schnell, J. R., Ketchem, R. R., Boger, D. L., and Chazin, W. J. (1999) Binding-Induced Activation of DNA Alkylation by Duocarmycin SA: Insights from the Structure of an Indole Derivative-DNA Adduct. *J.Am.Chem.Soc.,* **121**, 5645-5652.

[14] Shimotakahara, S., Gorin, A., Kolbanovskiy, A., Kettani, A., Hingerty, B. E., Amin, S., Broyde, S., Geacintov, N., and Patel, D. J. (2000) Accomodation of S-cis-tamoxifen-N(2)-guanine adduct within a bent and widened DNA minor groove. *J.Mol.Biol.*, **302**, 377-393.

[15] Sastry, M., Fiala, R., Lipman, R., Tomasz, M., and Patel, D. J. (1995) Solution structure of the monoalkylated mitomycin C-DNA complex. *J.Mol.Biol.*, **247**, 338-359.

[16] Zegar, I. S., Kim, S. J., Johansen, T. N., Horton, P. J., Harris, C. M., Harris, T. M., and Stone, M. P. (1996) Adduction of the human N-ras codon 61 sequence with (-)-(7S,8R,9R,10S)-7,8-dihydroxy-9,10-epoxy-7,8,9,10-tetrahydrobenzo[a] pyrene: structural refinement of the intercalated SRSR(61,2) (-)-(7S,8R,9S,1 0R)-N6-[10-(7,8,9,10- tetrahydrobenzo[a]pyrenyl)]-2'-

deoxyadenosyl adduct from 1H NMR. *Biochemistry*, **35**, 6212-6224.

[17] Horwitz, S. B, Chang, S. C., and Grollman, A. P. (1971) Chemosterilant action of anthramycin: a proposed mechanism. *Science*, **174**, 159-161.

[18] Hurley, L. H., Chandler, C., Garner, T. F., Petrusek, R. and Zimmer, S. G. (1979) DNA binding, induction of unscheduled DNA synthesis, and excision of anthramycin from DNA in normal and repair-deficient human fibroblasts. *J. Biol. Chem.*, **254**, 605-608.

[19] Hurley, L. H. (1977) Pyrrolo(1,4)benzodiazepine antitumor antibiotics. Comparative aspects of anthramycin, tomaycin and sibiromycin. *J. Antibiot.*, **30**, 349-370.

[20] Petrusek, R. L., Anderson, G. L., Garner, T. F., Fannin, Q. L., Kaplan, D. J., Zimmer, S. G., and Hurley, L. H. (1981) Pyrrolo[1,4]benzodiazepine antibiotics. Proposed structures and characteristics of the in vitro deoxyribonucleic acid adducts of anthramycin, tomaymycin, sibiromycin, and neothramycins A and B. *Biochemistry*, **20**, 1111-1119.

[21] Petrusek, R. L., Uhlenhopp, E. L., Duteau, N., and Hurley, L. H. (1982) Reaction of anthramycin with DNA. Biological consequences of DNA damage in normal and xeroderma pigmentosum cell. *J. Biol. Chem.*, **257**, 6207-6216.

[22] Hurley, L. H., and Needham-VanDevanter, D. R. (1986) Covalent binding of antitumor antibiotics in the minor groove of DNA. Mechanism of action of CC-1065 and the pyrrolo(1,4)benzodiazepines. *Acc. Chem. Res.*, **19**, 203-210.

[23] Kizu, R., Draves, P. H., and Hurley, L. H. (1993) Correlation of DNA sequence specificity of anthramycin and tomaymycin with reaction kinetics and bending of DNA. *Biochemistry*, **32**, 8712-8722.

[24] Remers, W. A., Mabilia, M., and Hopfinger, A. J. (1986) Conformations of complexes between pyrrolo[1,4]benzodiazepines and DNA segments. *J. Med. Chem.*, **29**, 2492-2503.

[25] Thurston, D. E. , Bose, D. S., Howard, P. W., Jenkins, T. C., Leoni, A., Baraldi, P. G., Guiotto, A., Cacciari, B., Kelland, L. R., Foloppe, and M-P.,Rault, S. (1999) Effect of A-Ring Modifications on the DNA-Binding Behavior and Cytotoxicity of Pyrrolo[2,1-c][1,4]benzodiazepines. *J. Med. Chem.*, **42**, 1951-1964.

[26] Langlois, N., Rojas-Rousseau, A., Gaspard, C., Werner, G. H., Darro, F., and Kiss, R. (2001) Synthesis and Cytotoxicity on Sensitive and Doxorubicin-Resistant Cell Lines of New Pyrrolo[2,1-c][1,4]benzodiazepines Related to Anthramycin. *J. Med. Chem.*, **44**, 3754-3757.

[27] Chen, Z., Gregson, S.J., Howard, P.W., and Thurston, D.E. (2004) A novel approach to

the synthesis of cytotoxic C2-C3 unsaturated pyrrolo[2,1-c][1-4]benzodiazepines (PBDs) with conjugated acrylyl C2-substituents. *Bioorg. Med. Chem. Lett.*, **14**, 1547-1549.

[28] Kamal, A., Ramesh, G., Laxman, N., Ramulu, P., Srinivas, O., Neelima, K., Kondapi, A. K., Sreenu, V. B., and Nagarajaram, H. A. (2002) Design, Synthesis, and Evaluation of New Noncross-Linking Pyrrolobenzodiazepine Dimers with Efficient DNA Binding Ability and Potent Antitumor Activity. *J. Med. Chem.*, **45**, 4679-4688.

[29] Masterson, L. A., Croker, S. K., Jenkins, T. C., Howard, P. W., and Thurston, D. E. (2004) Synthesis and Biological Evaluation of Pyrrolo[2,1-c][1,4]Benzodiazepine (PBD) C8 Cyclic Amine Conjugates. *Bioorg. Med. Chem. Lett.*, **14**, 901-904.

[30] Alley, M. C., Hollingshead, M. G., Pacula-Cox, C. M., Waud, W. R., Hartley, J. A., Howard, P. W., Gregson, S. J., Thurston, D. E., and Sausville E. A. (2004) SJG-136 (NSC 694501), A Novel Rationally Designed DNA Minor Groove Interstrand Cross-Linking Agent With Potent and Broad Spectrum Antitumour Activity: Part 2: Efficacy Evaluations. *Cancer Res.*, **64**, 6700-6706.

[31] Martin, C., Ellis, T., McGurk, C. J., Jenkins, T. C., Hartley, J. A., Waring, M. J., and Thurston, D. E. (2005) Sequence-Selective Interaction of the Minor-Groove Interstrand Cross-Linking Agent SJG-136 with Naked and Cellular DNA: Footprinting and Enzyme Inhibition Studies, *Biochemistry*, **44**, 4135-4147.

[32] James, A. M., Blaikie, F. H., Smith, R. A. J., Lightowlers, R. N., Smith, P. M., and Murphy, M. P. (2003) Specific targeting of a DNA-alkylating reagent to mitochondria. *Eur. J. Biochem.*, **270**, 2827-2836.

[33] Reynolds, V. L., and Hurley, L. H. (1982) Comparison of properties of the in vitro and cellular anthramycin-DNA adducts and characterization of the reaction of anthramycin with chromatin. *Chem-Biol. Interact.*, **42**, 141-151.

[34] Rao, S. N., Singh, U. C., and Kollman, P. A. (1986) Molecular mechanics simulations on covalent complexes between anthramycin and B DNA. *J. Med. Chem.*, **29**, 2484-2492.

[35] Zakrewska, K., and Pullman, B. (1986) A theoretical investigation of the sequence specificity in the binding of the antitumor drug anthramycin to DNA. *J. Biom. Struct. Dyn.*, **4**, 127-136.

[36] Adams, L. J., Jenkins, T. C., Banting, L., and Thurston, D. E. (1999) Molecular modelling of a sequence-specific DNA-binding agent based on the pyrrolo[2,1-c][1,4]benzodiazepines. *ad. Pharmacol. Commun.*, **5**, 555-560.

[37] Beveridge, D. L., and McConnell, K. J. (2000) Nucleic acids: theory and computer simulation, Y2K. *Curr. Opin. Struct. Biol.*, **10**, 182-196.

[38] Becke, A. D. (1988) Density-functional exchange-energy approximation with corrct asymptotic behavior. *Phys. Rev. A*, **38**, 3098-3100.

[39] Lee C., Yang W., and Parr R. G. (1988) Development of the Colle-Salvetti correlation-energy formulainto a functional of the electron density. *Phys. Rev. B*, **37**, 785-789.

[40] Frisch, M. J., Trucks, G. W., Schiegel, H. B., Scueseria, G. E., Robb, M. A., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Octerski, J., Petersson, G.A., Ayala, P.Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Baboul, A. G., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Allaham, M. A., Peng, C. Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P. M. W., Johnson, B. G., Chen, W., Wong, M. W., Andres, J. L., Head-Gordon, M., Replogle, E. S., and Pople, J. A. *Gaussian 98 (Revision A. 3)*. Gaussian, Inc., Pittsburgh, 1998.

[41] Pearlman, D. A., Case, D. A., Caldwell, J. W. Ross, W. R., Cheatham III, T. E., DeBolt, S., Ferguson, D., Seibel, G., and Kollman, P. (1995) AMBER, a computer program for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to elucidate the structures and energies of molecules. *Comp. Phys. Commun.*, **91**, 1-41.

[42] Cheatham III, T. E., Cieplak, P., and Kollman, P. A. (1999) A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.*, **16**, 845-862.

[43] Wang, J., Cieplak, P., and Kollmann, P. A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?. *J. Comp. Chem.,* **21**, 1049-1074.

[44] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.*, **79**, 926-935.

[45] Aaqvist, J. (1990) Ion-Water Interaction Potentials Derived from Free Energy Perturbation Simulations. *J. Phys. Chem.*, **94**, 8021-8024.

[46] Wang, J., Wolf, R. M., Caldwell, J. W., and Kollman, P. A. (2004) Development and testing

of a general Amber force field. *J. Comp. Chem.*, **25**, 1157-1174.

[47] Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. (1993) A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints For Determining Atom-Centered Charges: The RESP Model. *J. Phys. Chem.*, **97**, 10269-10280.

[48] Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089-10092.

[49] Nosé, S. (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.*, **52**, 255-268.

[50] Hoover, W. G. (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, **31**, 1695-1697.

[51] Parrinello, M., and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.*, **52**, 7182-7190.

[52] Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.*, **18**, 1463-1472.

[53] Berendsen, H. J. C., van der Spoel, D., and van Drunen, R. (1995) GROMACS: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.*, **91**, 43-56.

[54] Lindahl, E., Hess, B., and van der Spoel, D. (2001) GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod.* **7**, 306-317.

[55] Swaminatha, S., Ravishanker, G., Beveridge, D. L., Lavery, R., Etchebest, C., and Sklenar, H. (1990) Conformational and helicoidal analysis of the molecular dynamics of proteins: "curves," dials and windows for a 50 psec dynamic trajectory of BPTI. *Proteins* **8**, 179-193.

[56] Lee, B., and Richards, F. (1971) The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.*, **55**, 379-400.

[57] Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD - Visual Molecular Dynamics. *J. Mol. Graphics*, **14**, 33-38.

[58] Laaksonen, L., (1992) A graphics program for the analysis and display of molecular dynamics trajectories, *J. Mol. Graphics*, **10**, 33-34.

[59] Bergman, D. L., Laaksonen, L. and Laaksonen, A., (1997) Visualization of solvation structures in liquid mixtures, *J. Mol. Graph. Model.*, **15**, 301-306.

[60] Laio, A., VandeVondele, J., and Rothlisberger, U. (2002) A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations. *J. Chem. Phys.*, **116**, 6941-

6948.

[61] Hutter, J., Alavi, A. Deutsch, T., Ballone, P., Bernasconi, M., Focher, P., Goedecker, S., Tuckerman, M., and Parrinello, M. (1995-1999) CPMD, *Max-Planck-Institut für Festkörperforschung, Stuttgart and IBM Research Laboratory Zürich.*

[62] Hohenberg, P., and Kohn, W. (1964) Inhomogeneous electron gas. *Phys. Rev.*, **136**, B864-B871.

[63] Kohn, W., and Sham, L. J. (1965) Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, **140**, A1133-A1138.

[64] Sebastiani, D., and Rothlisberger, U. (2003). In Carloni, P., and Alber, F. (Eds.) *Quantum Medicinal Chemistry.* Wiley-VCH, Weinheim, pp. 5-36.

[65] Troullier, N., and Martins, J. L. (1991) Efficient pseudopotentials for plane-wave calculation. *Phys. Rev. B*, **43**, 1943-2006.

[66] Martyna, G. J., and Tuckerman, M. E., (1999) A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. *J. Chem. Phys.*, **110**, 2810-2821.

[67] Spiegel, K., Rothlisberger, U., and Carloni, P. (2004) Cisplatin Binding to DNA Oligomers from Hybrid Car-Parrinello/Molecular Dynamics Simulations. *J. Phys. Chem. B*, **108**, 2699-2707.

[68] Marzari, N., and Vanderbilt, D. (1997) *Phys. Rev. B*, **56**, 12847-12865.

[69] Alber, F., Folkers, G., and Carloni, P. (1999) Dimethyl Phosphate: Stereoelectronic versusEnvironmental Effects. *J. Phys. Chem. B*, **103**, 6121-6126.

[70] Farmer Jr, J. D., Gustafson, G. R., Conti, A., Zimmt, M. B., and Suggs, J. W. (1991) DNA binding properties of a new class of linked anthramycin analogs. *Nucl. Acid. Res.*, **19**, 899 (1991).

[71] Schlitter, J. (1993) Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chem. Phys. Lett.*, **256**, 617-622.

[72] Malhotra, R. K., Ostrander, J. M., and Hurley, L. H. (1982) Chemical conversion of antrhamycin 11-methyl ether to didehydroanhydro-anthramycin and its utilization in studies of the biosynthesis and mechanism of action of anthramycin. *J. Nat. Prod.*, **44**, 38-44.

[73] Bailly, C., Graves, D. E., Ridge G., and Waring, M. J. (1994) Use of a photoactive derivative of actinomycin to investigate shuffling between binding sites on DNA. *Biochemistry*, **33**, 8736-8745.

[74] Gunz, D., and Naegeli, H. (1996) A noncovalent binding-translocation mechanism for site-specific CC-1065-DNA recognition. *Biochem. Pharmacol.* **52**, 447-453 (1996).

[75] The presence of the OH group and of the H atom on the reactive carbon $C_{11}$of hydroxy-anthramycin does not allow for an equivalent good packing of drug into groove of DNA.

[76] The minimization has been carried out with a simulated annealing procedure, to ensure that the system was not trapped in a local energy minimum.

[77] Note that, due to the delocalization of π bonds [69], these results should be taken as an indication.

Figure Legends

Figure 1: Mass weighted RMSDs of the oligonucleotides in **ANT•DNA**, **IMI•DNA**, **ANT-DNA**, and **DNA**, obtained after least square (LSQ) fit to the minimized structures in solvent: the averages values are 4.0 Å, 2.9 Å, 1.8 Å and 2.7 Å, respectively. The RMSD of **DNA** is almost constant and is reported as dott-dashed line. The inset reports the RMSD's relative only to the drugs in the non-covalent complexes.

Figure 2: Superimposition of the average **IMI•DNA** structures before (green, **1**) and after (orange, **2**) $10\,ns$. The drug moves by 1bp step. The mean distances between reactive atoms $N2_{G11}$ and $C11_{Ant}$ are respectively 3.3 ± 0.2 Å in **1** and 6.5 ± 0.5 Å in **2**.

Figure 3: Superimposition between the average structure of **DNA** (red) and those of **ANT•DNA** (blue), **IMI•DNA 2** (orange) , **IMI•DNA 1** (green) and **ANT-DNA** (purple).

Figure 4: Mean values of selected structural parameters of the DNA. *(a) Minor groove widths:* Binding by hydroxy-anthramycin affects the groove width only in the central part of the DNA, whilst the anhydro form fits better into the groove and causes a narrowing similar (and also more extended) that in **ANT-DNA**. *(b) Twist:* the values for **DNA** are similat to those of **ANT•DNA** and **IMI•DNA 2**; the data for **IMI•DNA 1** are instead dissimilar from the others. The data for the covalent adduct in graphs *a* and *b* agree with experiment [1] except that in the region between $C2$-$G23$ and $C6$-$G19$, due to the presence of a second drug bonded at $G22$ in the X-ray structure. *(c) Buckle and (d) propeller:* values for **IMI•DNA 2** are in general closer to those of **DNA** with respect to the configuration **1**. The largest differences are localized in the bps $T9-A16$ and $G10^{*}-C15$.

Figure 5: H-bonds formed by anthramycin in **ANT-DNA**.

Figure 6: Variations of the non-bonded (electrostatic+LJ) energy of the drug-$d(G_7T_8T_9G_{10}G_{11})d(C_{18}A_{17}A_{16}C_{15}C_{14})$, drug-solvent, and solvent-$d(G_7T_8T_9G_{10}G_{11})d(C_{18}A_{17}A_{16}C_{15}C_{14})$ moieties, upon sliding of anhydro-anthramycin (dashed lines correspond to running averages performed on blocks of 100 steps). As can be seen there is a significant and permanent change of about -10 $kcal/mol$ in the energy between the drug and the bps directly involved in the sliding process, while no clear trend is evident for the others curves.

Figure 7: Contour plots of the electrostatic potential (Volts) within the QM region, through the plane defined by atoms $N2$, $N10$, $C11$. The plots report (*a:* **IMI•DNA**, *b:* **ANT•DNA**) the ΔV between the potential calculated *with* and *without bioframe*. Note that only in the first system (*a*) the largest difference is *localized* near to the electrophilic carbon, while in *b* it is *spread* in front of the reactive amino group of the guanosine. The $N10$-$C11$ arrow indicates the polarization of the bond, while the arrow between $N2$ and$C11$ indicates the gradient in ΔV.

Figure 8: Comparison between calculated and experimental B-factors ($T_{exp} \simeq 40K$) of the DNA and the drug. Note the good agreement with experiment for the data concerning the drug and the first strand (binding strand - red line) of the oligonucleotide at the binding region. The rest of this strand is instead much more mobile, due to the removal the second drug from the X-ray structure. Also the B-factors for the non-bonded strand (blue line) are well reproduced near the binding zone, even if the C5' base has very high fluctuations.

Tables

$angle(^\circ)\backslash system$	**ANT-DNA**	**ANT•DNA**	**ANT**	**IMI • DNA 2** **IMI • DNA 1**	**IMI**
PROPELLER	31 ± 7 (37.1)	26 ± 8	22 ± 8 (34.4*)	22 ± 7 29 ± 7	26 ± 9
BUCKLE	13 ± 6 (17.6)	10 ± 5	12 ± 6 (8.2*)	8 ± 4 8 ± 4	9 ± 5

Table I: Propeller twist and buckle of the drugs. In parenthesis are reported available experimental results. Data labelled with * refers to the $O11$-methyl ether of the drug.

$interaction(kcal/mol)\backslash system$	**ANT•DNA**	**IMI • DNA 2** **IMI • DNA 1**	**ANT-DNA**
LJ	-30 ± 2	-40 ± 3 -33 ± 3	-34 ± 3
COULOMB	-12 ± 3	-11 ± 3 -8 ± 3	-20 ± 4

Table II: LJ and coulomb interaction energies extracted from the non-bonded terms of the AMBER force field [42, 43].

	IMI	**ANT**	**GTTGG**	**IMI•CTTGG**	**ANT•GTTGG**	**ANT-GTTGG**
SASA (Å)	191 ± 5	161 ± 5	764 ± 17	667 ± 18 (**2**) 709 ± 17 (**1**)	746 ± 25	739 ± 21
ΔSASA (Å)				-288 (30%, **2**) -246 (26%, **1**)	-179 (19%)	-216 (26%, **IMI**) -186 (20%, **ANT**)

Table III: Calculated hydrophobic SASA's. The hydrophobic effect plays a role in the formation of the complex: a negative value of ΔSASA upon solvation in water corresponds to a gain in the free energy of the system.

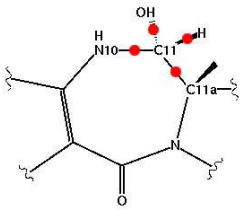
$H-bond \backslash system$	IMI•DNA 1	IMI•DNA 2	ANT•DNA	ANT-DNA	DNA	IMI	ANT
$\mathbf{NH_{G11}...O9_{Ant}}$	2.2 ± 0.3Å 58%	–	1.9 ± 0.3Å 73%	2.2 ± 0.3Å 59%	–	–	–
$\mathbf{NH_{G10}...O9_{Ant}}$	–	2.2 ± 0.3Å 60%	–	–	–	–	–
$\mathbf{N_{G11}...HO9_{Ant}}$	2.4 ± 0.4Å 50%	–	–	–	–	–	–
$\mathbf{O_{C15}...HO9_{Ant}}$	–	–	–	1.9 ± 0.2Å 81%	–	–	–
$\mathbf{N_{A16}...HO9_{Ant}}$	–	2.3 ± 0.3Å 39%	–	–	–	–	–
$\mathbf{N_{G11}...HN10_{Ant}}$	–	–	2.7 ± 0.2Å 16%	–	–	–	–
$\mathbf{O_{C15}...HN10_{Ant}}$	–	–	–	2.1 ± 0.2Å 65%	–	–	–
$\mathbf{NH_{G10}...N10_{Ant}}$	2.4 ± 0.3Å 69%	–	–	–	–	–	–
$\mathbf{O4'_{G11}...HO11_{Ant}}$	–	–	3.3 ± 0.9Å 14%	–	–	–	–
$\mathbf{N_{A17}...HN15_{Ant}}$	–	–	–	1.9 ± 0.3Å 61%	–	–	–
$\mathbf{HO9_{Ant}...wat}$	s : 75%	s : 17%	s : 79% d : 11%	s : 50%	–	s : 36% d : 42% t : 14%	s : 22% d : 61% t : 12%
$\mathbf{HO11_{Ant}...wat}$	–	–	s : 90% d : 17%	–	–	–	s : 17% d : 62% t : 19%
$\mathbf{HN10_{Ant}...wat}$	–	–	–	–	–	–	s : 37%
$\mathbf{N10_{Ant}...wat}$	–	–	-	–	–	s : 10%	–
$\mathbf{HN15_{Ant}...wat}$	s : 96%	s : 99%	s : 84% d : 6%	s : 33%	–	s : 68% d : 21%	s : 66% d : 21%
$\mathbf{HN_{G11}...wat}$	–	s : 74%	s : 17%	–	s : 71%	–	–
$\mathbf{N3_{G11}...wat}$	-	s : 69%	s : 14%	–	s : 74%		
$\mathbf{O4'_{G11}...wat}$	-	-	s : 27%	-	s : 39%	–	–
$\mathbf{O2_{C14}...wat}$	s : 84%	s : 86%	s : 87%	s : 83%	s : 85%	–	–
$\mathbf{HN_{G10}...wat}$	-	-	s : 17%	–	s : 71%	–	–
$\mathbf{N3_{G10}...wat}$	–	-	s : 63%	–	s : 76%	–	–
$\mathbf{O2_{C15}...wat}$	s : 77%	s : 78%	s : 8%	–	s : 83%	–	–
$\mathbf{O2_{T9}...wat}$	s : 95%	–	s : 91% d : 14%	s : 91%	s : 70% d : 14%	–	–
$\mathbf{N3_{A16}...wat}$	–	–	–	–	s : 85%	–	–
$\mathbf{O2_{T8}...wat}$	s : 91%	s : 95%	s : 85%	s : 70% d : 13%	s : 85% d : 6%	–	–
$\mathbf{N3_{A17}...wat}$	s : 73%	–	s : 82%	s : 19%	s : 81%	–	–
$\mathbf{HN_{G7}...wat}$	s : 71%	s : 68%	s : 72%	s : 67%	s : 70%	–	–
$\mathbf{N3_{G7}...wat}$	s : 55%	s : 65%	s : 59%	s : 65%	s : 64%	–	–
$\mathbf{O2_{C18}...wat}$	s : 90% d : 11%	s : 98%	s : 85% d : 10%	s : 91%	s : 85% d : 7%	–	–

Table IV: H-bonds among the drug, the oligonucleotide and the solvent. The table shows only bonds formed at the drug-DNA interface. We report hydrogen-acceptor distances (only for the drug-DNA network) and life times, expressed as percentage over the simulation time. The symbols *s*, *d* and *t* indicate the formation of 1, 2 and 3 bonds respectively, and the labeling of the atoms refers to Fig. 5.

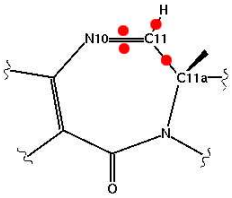
	IMI•DNA 1	IMI•DNA 2	ANT•DNA	ANT-DNA	DNA	IMI	ANT
drug	3 3.6 ± 1.0	2 2.2 ± 0.8	3 4.3 ± 1.1	4 3.6 ± 0.9	–	6 2.3 ± 1.1	9 5.4 ± 1.3
DNA	9 8.6 ± 1.5	8 7.5 ± 1.4	13 8.8 ± 1.7	8 5.6 ± 1.1	14 11.4 ± 1.6	–	–

Table V: Total number of persistent H-bonds (present for more than 10% of the simulation time) formed by the drug and the five bps $d(G_7T_8T_9G_{10}G_{11})d(C_{18}A_{17}A_{16}C_{15}C_{14})$ involved in the binding. Averages and standard deviations of the total number of H-bonds are also reported in the second row of each cell.

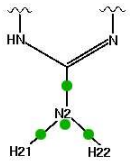
		with bioframe	without bioframe	without solvent	vacuo
IMI•DNA	**C11 – H11**	0.66 ± 0.01	0.68 ± 0.04	0.66 ± 0.01	0.68
	C11 – N10	0.57 ± 0.01	0.54 ± 0.01	0.58 ± 0.01	0.55
	C11 – N10	0.58 ± 0.01	0.55 ± 0.01	0.59 ± 0.01	0.56
	C11 – C11a	0.51 ± 0.01	0.51 ± 0.01	0.50 ± 0.01	0.51
	N2 – H21	0.57 ± 0.01	0.59 ± 0.01	0.57 ± 0.01	0.59
	N2 – H22	0.59 ± 0.01	0.58 ± 0.01	0.57 ± 0.01	0.58
	N2 – C2	0.36 ± 0.04	0.37 ± 0.04	0.38 ± 0.06	0.33
	N2 lone pair	0.21 ± 0.05	0.19 ± 0.05	0.21 ± 0.07	0.24
ANT•DNA	**C11 – H1**	0.66 ± 0.01	0.68 ± 0.04	0.66 ± 0.01	0.67
	C11 – N10	0.57 ± 0.01	0.57 ± 0.01	0.58 ± 0.01	0.57
	C11 – C11a	0.50 ± 0.01	0.50 ± 0.01	0.50 ± 0.01	0.51
	C11 – O11	0.63 ± 0.01	0.62 ± 0.01	0.62 ± 0.01	0.62
	N2 – H21	0.57 ± 0.01	0.59 ± 0.01	0.57 ± 0.01	0.59
	N2 – H22	0.58 ± 0.01	0.58 ± 0.01	0.59 ± 0.01	0.58
	N2 – C2	0.37 ± 0.04	0.38 ± 0.04	0.38 ± 0.04	0.33
	N2 lone pair	0.20 ± 0.05	0.15 ± 0.06	0.15 ± 0.06	0.24



ANT



IMI



GUA

Table VI: Bond ionicity of X-$N2_{G10}$ and X-$C11_{drug}$ bonds. Last column refers to calculations *in vacuo* on the drug or the guanosine. Two rows with the same label indicate a double bond, while the last row in each system represents the BI associated to the lone pair of the guanine nitrogen. The figures on the right side depict the BO's considered for the BI calculations (see Methods).

	IMI•DNA	**ANT•DNA**
with bioframe	$\Delta(lumo/homo-2) \approx 1.4\,eV$	$\Delta(lumo/homo) \approx 1.2\,eV$
without solvent	$\Delta(lumo/homo-1) \approx 1.3\,eV$	$\Delta(lumo+1/homo) \approx 2.1\,eV$
without bioframe	$\Delta(lumo/homo-1) \approx 1.9\,eV$	$\Delta(lumo+1/homo) \approx 3.2\,eV$

Table VII: Influence of the environment on the gaps between Kohn-Sham energy levels. The orbitals considered are those with the maximum projection on the reactive atoms.

	$N3_{G11}...O9_{Ant}$	$O2_{C15}...O9_{Ant}$	$O2_{C15}...N10_{Ant}$	$N3_{G11}...N10_{Ant}$	$O2_{T9}...N15_{Ant}$	$O4'_{G10}...N15_{Ant}$
$\mathbf{X-ray}$	3.9Å	3.1Å	3.4Å	3.7Å	2.9Å	3.3Å
MD	4.0 ± 0.3Å	2.9 ± 0.3Å	2.9 ± 0.2Å	3.3 ± 0.2Å	5.3 ± 0.4Å	4.9 ± 0.9Å

Table VIII: Comparison between X-ray and calculated acceptor-donor distances involving drug and DNA.

Figures

Chart I

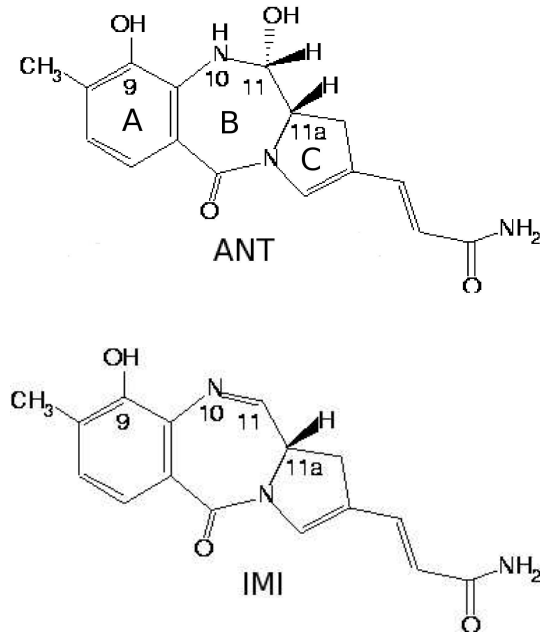


Fig. 1

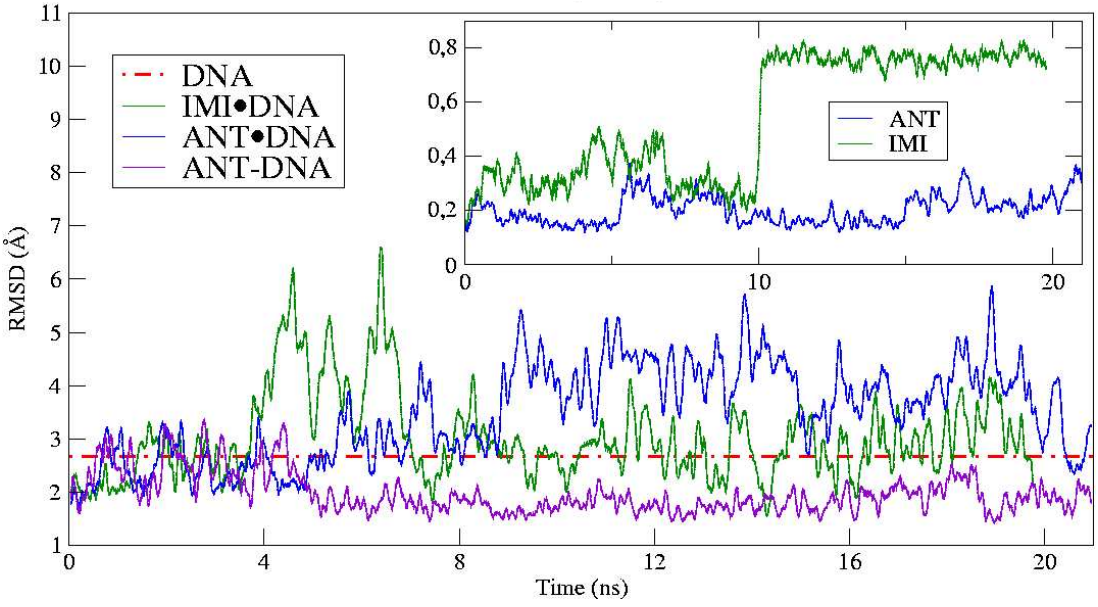


Fig. 2

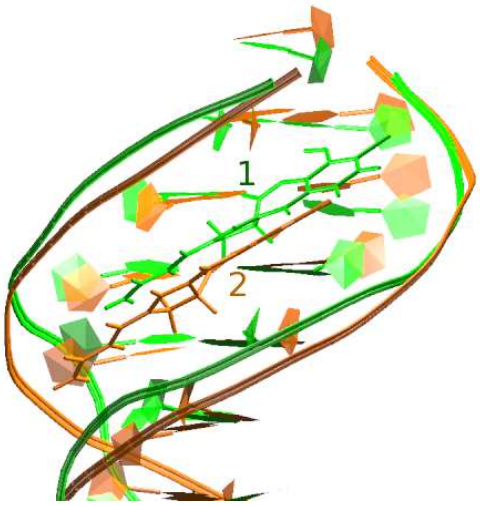


Fig. 3

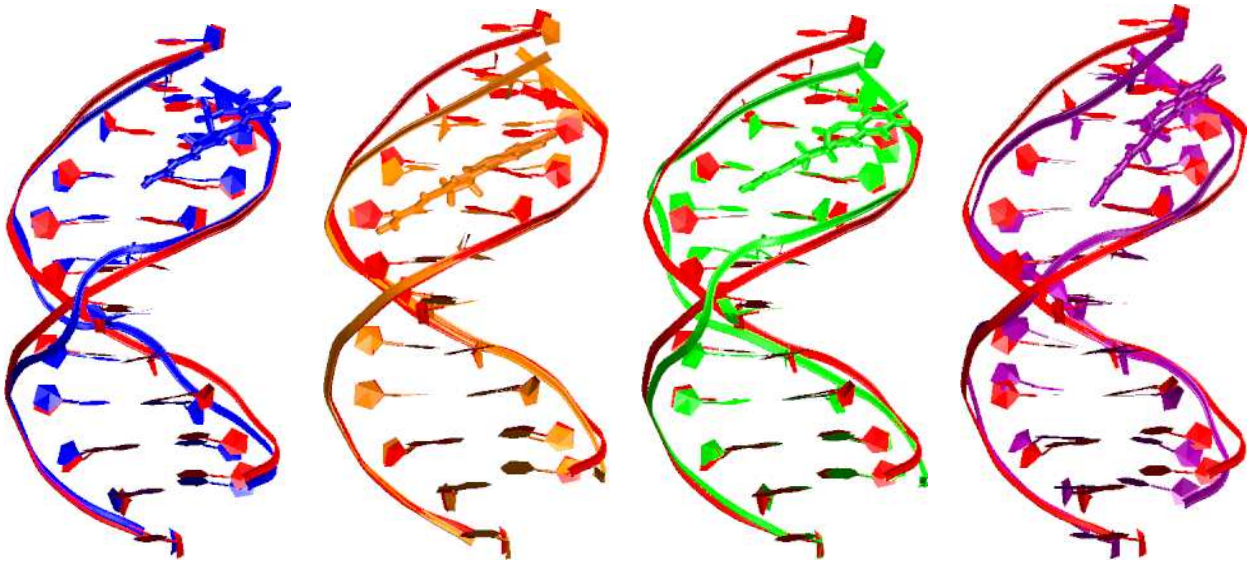


Fig. 4

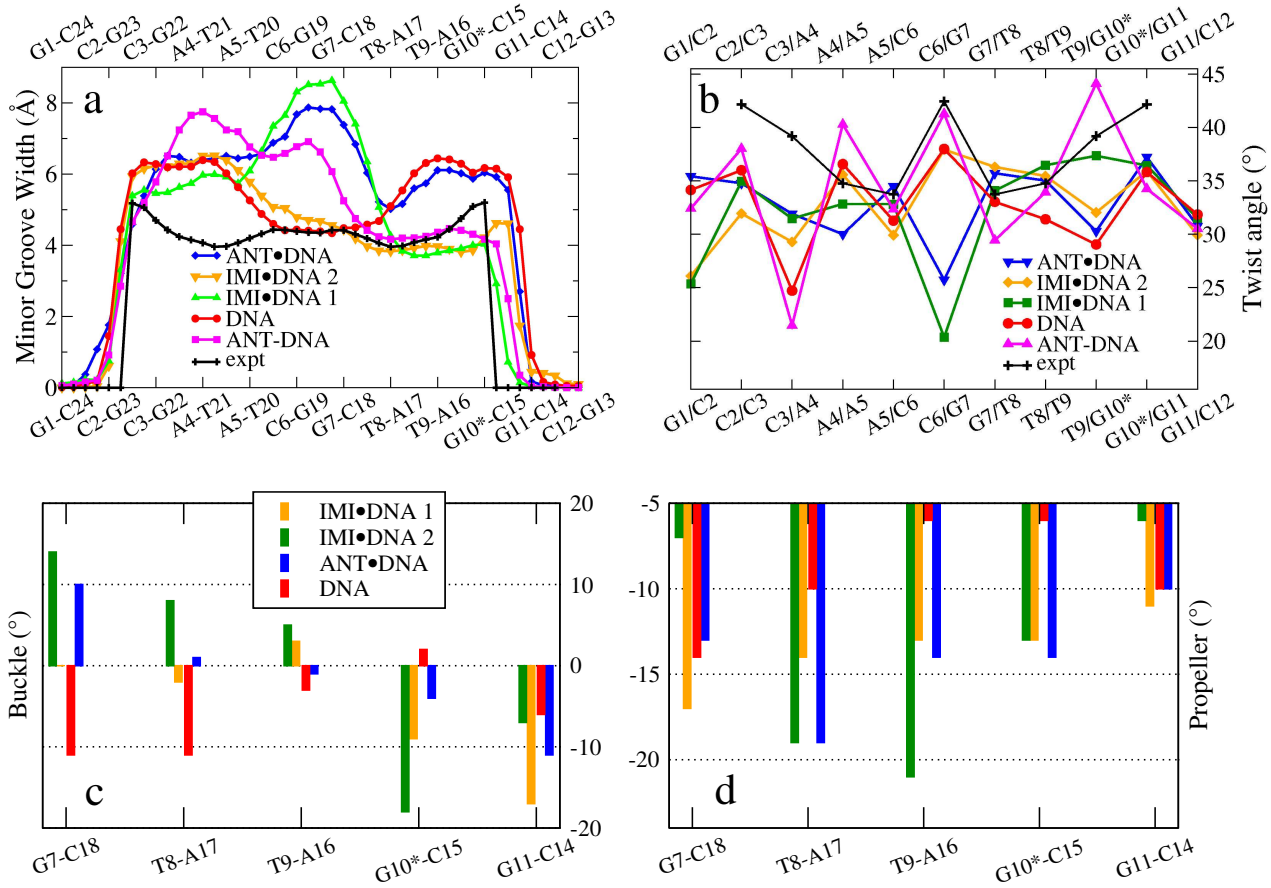

a
Minor Groove Width (Å)
G1-C24
C2-G23
C3-G22
A4-T21
A5-T20
C6-G19
G7-C18
T8-A17
T9-A16
G10*-C15
G11-C14
C12-G13
ANT•DNA
IMI•DNA 2
IMI•DNA 1
DNA
ANT-DNA
expt
b
Twist angle (°)
G1/C2
C2/C3
C3/A4
A4/A5
A5/C6
C6/G7
G7/T8
T8/T9
T9/G10*
G10*/G11
G11/C12
c
Buckle (°)
IMI•DNA 1
IMI•DNA 2
ANT•DNA
DNA
d
Propeller (°)

Fig. 5

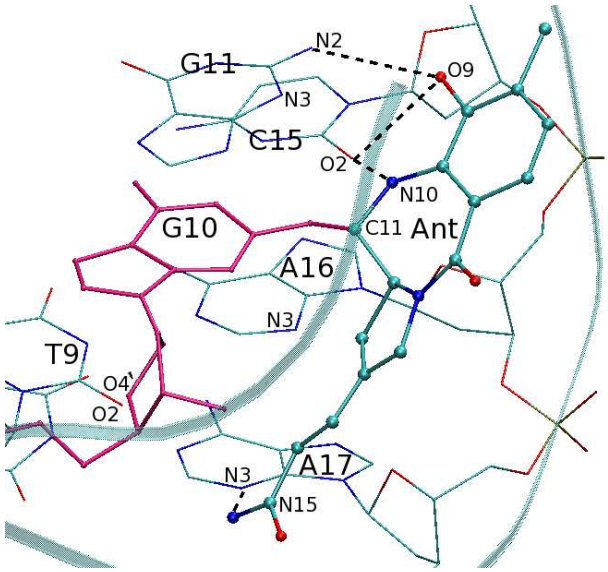

G11
N2
O9
N3
C15
O2
N10
C11
Ant
G10
A16
N3
T9
O4'
O2
N3
A17
N15

Fig. 6

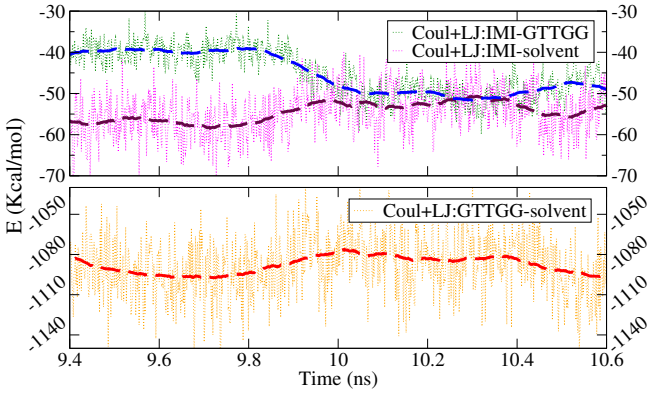


Fig. 7

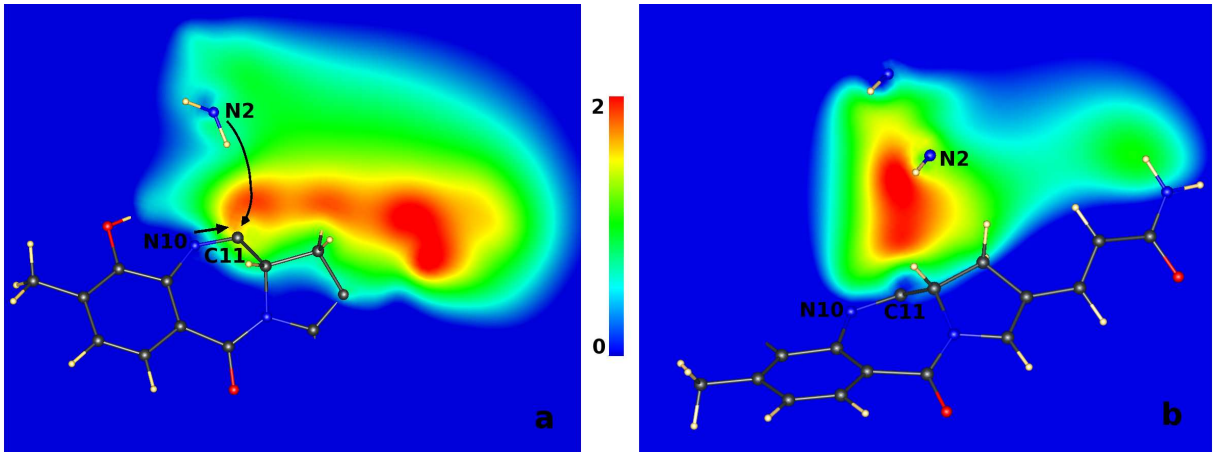


Fig. 8