DNA Cross-Linking by Azinomycin B: Monte Carlo Simulations in the Evaluation of Sequence Selectivity

Stefano Alcaro,*,† Francesco Ortuso,† and Robert S. Coleman‡

Dipartimento di Scienze Farmaco-Biologiche, Università di Catanzaro "Magna Græcia", 88021 Roccelletta di Borgia, Catanzaro, Italy, and Department of Chemistry, The Ohio State University, 100 West 18th Avenue, Columbus, Ohio 43210

Received September 11, 2001

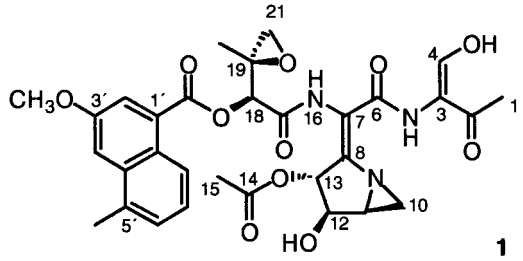
A new set of charges specifically developed for biologically relevant N7-alkylated purine adducts have been implemented in the AMBER* force field of the MacroModel package and applied to the conformational search of azinomycin B−DNA interactions. To perform a sequence dependent reactivity relationship study, four DNA triplets known to interact differently with the drug, 5′-GCT-3′, 5′-GCC-3′, 5′-GTC-3′, and 5′-GTT-3′, have been modeled in B-form and intercalative conformations. Monte Carlo simulations of all possible monoadducts and intercalative complexes have been carried out and analyzed using a filtering criterion that estimates the probability of covalent bond formation and covalent cross-linking. We observed a good correlation between existing experimental data and our computational estimations that validate the approach. The comparison of the conformational properties of the drug−DNA monoadducts and complexes confirms the most probable mechanism of action involving an initial aziridine and subsequent epoxide alkylation. The different hydrogen bond network in the monoadducts and in the intercalative complexes between the drug and the three base-pair receptor is the primary reason for the different cross-linking reactivity. In addition, steric hindrance of the major groove exposed methyl group of central thymine-based triplets plays an important role in the lack of the reactivity of these sequences. Synthetic work on the azinomycins and the information coming from this computational study will be important for the design of more potent or DNA sequence-selective agents based on the azinomycin skeleton.

Introduction

Azinomycin B (Figure 1) and the structurally related agent azinomycin A are naturally occurring antitumor agents[1] that exhibit both in vitro cytotoxic activity and in vivo antitumor activity.[2] Azinomycin B appears to exert its biological activity by the formation of covalent interstrand cross-links[3] in duplex DNA, presumably via the electrophilic epoxide and aziridine rings, but a detailed biological evaluation of these agents has been hampered by chemical instability and poor availability from natural sources. Cross-linking occurs within the major groove via the N7 position of two purine bases.[4,5] This unprecedented molecular mechanism of action and effective antitumor activity make the azinomycins attractive targets for synthetic[6,7] and mechanistic efforts;[4,5] biological evaluation has been limited to the original report.[2]

We have developed force-field parameters and described Monte Carlo conformational searching of the natural agents.[8] Because of the large number of populated conformers, we developed filtering protocols based on a distance and vector analysis of the cross-linking site in duplex DNA in order to analyze low energy conformers. In our earlier studies, we concluded that azinomycin A and B were highly preorganized for DNA cross-linking.

The interactions of azinomycin B with duplex DNA are challenging to study using computational methods



1

Figure 1. Azinomycin B structure and numbering scheme.

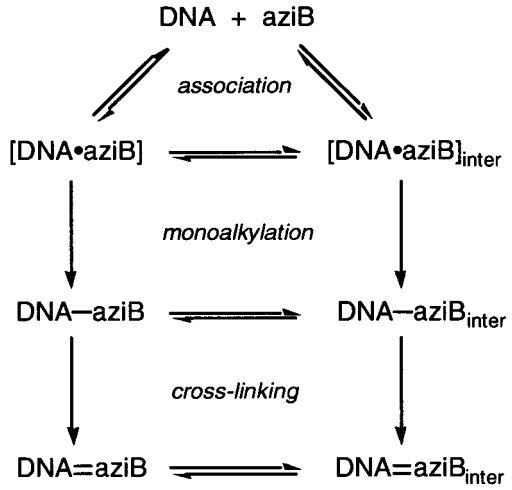
because of the conformationally mobile nature of the molecule. Thus, it is critical to have effective parameters and charges. In the present work, we have examined in detail the noncovalent association, monoalkylation, and interstrand DNA cross-linking by azinomycin B using Monte Carlo simulations of the agents bound to the putative DNA target sequence 5′-d(PuNPy)-3′ (Pu = purine, Py = pyrimidine, N = any nucleotide). We examined the interaction of the agent with a native B-DNA receptor and with a B-DNA receptor containing a preformed intercalation site. Our results provide a detailed model for the interaction of this agent with duplex DNA that is consistent with experimentally observed sequence selectivity and with computationally predicted nucleobase reactivity.[5,6]

In our preliminary communication,[9] we have considered the general paradigm for the interaction between **1** and the DNA shown in Scheme 1. Initial noncovalent association of the agent with DNA can occur by two alternative pathways: (1) association of the agent with the surface of the major groove; (2) association of the

* To whom correspondence should be addressed. Tel: +39-0961-391157. Fax: +39-0961-391490. E-mail: alcaro@unicz.it.
† Università di Catanzaro "Magna Græcia".
‡ The Ohio State University.

10.1021/jm011040w CCC: $22.00
Published on Web 01/10/2002

Scheme 1. Theoretical Reactions between Azinomycin B (aziB) and DNA in B-Form in Equilibrium with Intercalation



agent with the major groove accompanied by intercalation. Two covalent bonds are then formed in a sequential fashion between the agent and DNA: monoalkylation followed by interstrand cross-link formation. Each of these steps may or may not be accompanied by intercalation. We have now examined both pathways in detail, focusing on the conformational preferences of the agent when bound to DNA with respect to distance and geometric requirements for covalent bond formation between the electrophilic epoxide C21 and aziridine C10 carbons and the nucleophilic purine N7 atoms of the DNA. We have found a strong correlation of our modeling results with experimentally determined sequence selectivity patterns and with computationally based models of sequence-dependent nucleobase reactivity.

AMBER* Charge Improvement of N7-Alkylated Nucleotides

The DNA alkylation effected by many agents, including azinomycin B, generates adducts that increase the net charge by +1. In conformational searching of such agent−DNA adducts, especially with very flexible systems, the effect of this charge is nonnegligible and can significantly affect the energetic ordering of conformers. To improve our preliminary model,[9] we have performed a comparison between the AMBER* and ab initio charge distribution of dG and dA bases and their corresponding N7-methyl adducts. Other authors have reported a similar study considering the methylation at the N2 atom of guanine occurring by minor groove cross-linking agents.[10]

In the original AMBER* force field[11] as used in the MacroModel package,[12] the nucleic acid parameters and point charges are implemented using sugar and base substructures. The total charge of each nucleotide is −1.000 because of the negative charge contribution of the phosphate diester. A deeper analysis of the original AMBER* united atom (UA) notation implementation of DNA and RNA substructures reveals that the sugar and phosphate charge contribution is constant at −0.764 charge units.

When a methyl group is introduced onto the purine N7 position of either a dA or dG base, since no substructures are available for the N-methylated forms

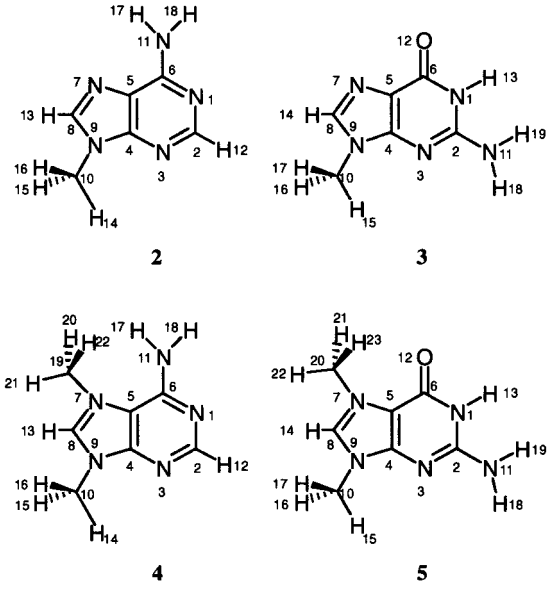


Figure 2. Numbering of compounds **2**, **3**, **4**, and **5**.

of the bases, the force field parameter assignment of AMBER* occurs using substructures not directly related to nucleic acid parameters (e.g., imine and amide parameters). The macroscopic effect of this operation results in an overall charge assigned to each methylated nucleotide that differs from expected values and differs between the two purines (dAMP = −0.470, dGMP = +0.236). In the case of azinomycin B alkylation of DNA in the self-complementary duplex oligonucleotide $d(TAGCTA)_2$, we surprisingly obtained different total charges when we built adducts with an N7-alkylated guanine versus adenine.[9] This electrostatic perturbation changed the outcome of conformational searches, especially when in the comparison of monoadducts that result in the same cross-linking pattern.

The first step of the present study has been devoted to atomic charge calculations of N9-methyladenine and N9-methylguanine (**2** and **3**, Figure 2). We have computed atomic charge distributions of the two bases using the following ab initio methods STO3G, HF 3-21G*, HF 6-31G*, HF 6-31G**, and HF 6−311 G**.[13] In Tables 1 and 2 we report the comparison between the original AMBER* force field set of charges and the computed ab initio ones along with correlation coefficients.

The best correlation between the AMBER* original charge distribution and the five ab initio methods is represented by HF 3-21G*. Therefore, this level of theory and basis set were adopted for ab initio charge calculations of the dimethylated purines N7,N9-dimethyladenine and N7,N9-dimethylguanine (**4** and **5**, Figure 2), setting the total charge equal to +1 (data not shown).

The difference in the computed charge distribution of the nominally equivalent hydrogen atoms is due to high conformational sensitivity of the ab initio calculations. Such difference mainly regards the aliphatic hydrogen atoms but is automatically solved in the UA implementation by the direct inclusion of their electrostatic contributions into the attached carbon charge.

The second step of this study involved the incorporation of a new set of charges for N-methylated bases **4** and **5** into the AMBER* force field. To follow the same

Table 1. AMBER* vs ab Initio Charge Distribution and Correlation for Compound **2**

	atom	AMBER*	STO 3G	3-21 G*	6-31 G*	6-31 G**	6-311 G**
1	N	−0.760	−0.716	−0.819	−0.775	−0.779	−0.782
2	C	0.603	0.608	0.598	0.547	0.549	0.555
3	N	−0.717	−0.690	−0.791	−0.737	−0.740	−0.741
4	C	0.695	0.511	0.585	0.474	0.479	0.466
5	C	−0.151	−0.058	−0.129	−0.084	−0.088	−0.099
6	C	0.813	0.738	0.905	0.872	0.874	0.873
7	N	−0.599	−0.545	−0.584	−0.565	−0.566	−0.558
8	C	0.426	0.227	0.154	0.097	0.099	0.081
9	N	−0.729	−0.047	−0.107	−0.001	−0.001	0.038
10	C	0.157	−0.308	−0.260	−0.184	−0.197	−0.211
11	N	−0.793	−0.793	−1.047	−1.058	−1.048	−1.019
12	H	−0.032	−0.025	0.059	0.063	0.064	0.062
13	H	0.062	0.077	0.170	0.176	0.176	0.176
14	H	0.038	0.120	0.118	0.085	0.088	0.089
15	H	0.038	0.122	0.134	0.103	0.107	0.108
16	H	0.038	0.129	0.140	0.109	0.113	0.114
17	H	0.337	0.330	0.439	0.442	0.438	0.427
18	H	0.337	0.320	0.434	0.435	0.431	0.420
	AMBER* r^2		0.828	**0.839**	0.802	0.802	0.787

Table 2. AMBER* vs ab Initio Charge Distribution and Correlation for Compound **3**

	atom	AMBER*	STO 3G	3-21 G*	6-31 G*	6-31 G**	6-311 G**
1	N	−0.746	−0.707	−0.940	−0.853	−0.851	−0.830
2	C	0.842	0.843	1.075	1.012	1.011	1.001
3	N	−0.702	−0.684	−0.810	−0.751	−0.755	−0.751
4	C	0.490	0.383	0.451	0.366	0.370	0.342
5	C	−0.088	−0.070	−0.184	−0.098	−0.101	−0.103
6	C	0.714	0.648	0.893	0.800	0.802	0.801
7	N	−0.575	−0.497	−0.477	−0.477	−0.477	−0.461
8	C	0.382	0.199	0.073	0.039	0.039	0.007
9	N	−0.651	−0.007	−0.041	0.021	0.024	0.076
10	C	0.157	−0.312	−0.233	−0.141	−0.157	−0.176
11	N	−0.758	−0.780	−1.037	−1.053	−1.042	−1.012
12	O	−0.459	−0.449	−0.629	−0.625	−0.625	−0.631
13	H	0.340	0.341	0.458	0.432	0.430	0.421
14	H	0.046	0.060	0.163	0.168	0.169	0.171
15	H	0.038	0.120	0.112	0.079	0.082	0.084
16	H	0.038	0.121	0.122	0.088	0.093	0.094
17	H	0.038	0.121	0.122	0.087	0.092	0.094
18	H	0.329	0.340	0.450	0.459	0.456	0.444
19	H	0.329	0.330	0.433	0.445	0.442	0.430
	AMBER* r^2		0.844	**0.849**	0.835	0.834	0.812

philosophy as our original AMBER* nucleotide parameter development,[8] we performed a normalization starting from the HF 3-21G* data that most effectively reproduced the N7-methylated purine +0.764 partial charge, as expected. This value is able to give a net positive charge to any aliphatic N7-adduct of the nucleobase. Technically, it was carried out following our previously reported approach[8] by adding two new substructures corresponding to the methylated purine bases **4** and **5** to the AMBER* force field (see Supporting Information).

We have tested the proper functioning of these new substructures directly on our DNA−azinomycin B models.[9] The presence of sodium counterions in our DNA duplex makes the nucleic acid neutral per se. As expected for the N7-monoalkylated adducts, we always obtained a total overall charge of +1. In addition, the two cross-linked models have been successfully tested, obtaining an overall charge of +2. The effect of the charge improvement in the conformational search of monoadduct is discussed in the Results section.

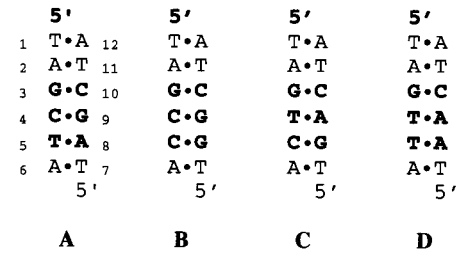


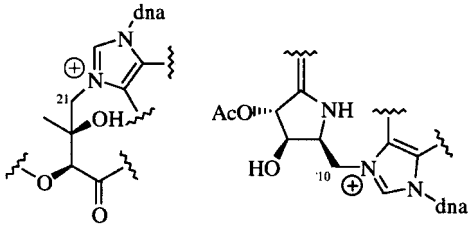
Figure 3. Structure of the duplex DNA sequences **A**, **B**, **C**, **D** modeled in B-form conformation with nucleotide numbering scheme.

Conformational Search of Azinomycin B−DNA Interactions

The charge modifications introduced in the AMBER* force field, together with the other parameters specifically developed for the azinomycin molecule,[8] have been applied in a detailed conformational evaluation of the drug−DNA interactions. To create a model closest to existing experimental data, we considered four DNA duplexes possessing a broad spectrum of the cross-linking reactivity.[5] The length of the DNA fragments was standardized to six nucleotides and the first, second, and sixth nucleotides were chosen adjacent to the receptor triplets according to a literature reference.[4] The four duplexes with their complementary strands are shown in Figure 3.

The only published experimental cross-linking reactivity comparison follows the trend: **A** = **B** > **C** ≫ **D**.[5] The conformational search of azinomycin B/DNA interactions has been carried out considering the relevant conformations of the molecular systems responsible for the cross-link and the first alkylation event. Extending what we initially reported for one DNA sequence,[9] we now have considered both monoadducts and intercalative docking.

Monoadducts. Our previous results indicate that the cross-linking mechanism is strongly related to the conformational properties of monoadducts formed between the drug and DNA.[9] In that case, working only with the duplex **A** containing the triplet 5′-GCT-3′ that is experimentally known to be effective in cross-linking,[4] we used the original AMBER* charge distribution corrected with our force field modifications specifically developed for the drug.[8]



The monoadducts were generated considering four possible reaction paths involving both the epoxide (C21) and the aziridine rings (C10), respectively, and the nucleophilic N7-purine atoms of the third and the eighth nucleotides. The four possible combinations have been classified using the acronyms MAB, MAT, MEB, and MET[14] as described in our previous communication[9] and repeated for each of the duplexes considered (**A**, **B**, **C**, and **D**). Starting with the global minimum energy

conformer of azinomycin B obtained from detailed Monte Carlo studies,[8] the four starting monoalkylation intermediates were generated by the following method: (1) energy minimization of the ring-opened epoxide and aziridine compounds; (2) attachment of the agent to the purine N7 by superimposition of the C10 or C21 atoms of the agents onto the methyl carbon of the appropriate N7-methylpurine; (3) deletion of the methyl group; and (4) connection of the agent to the DNA.

The Monte Carlo conformational search of the adducts was carried out using a computational protocol slightly modified from that adopted previously,[9] omitting the initial in vacuo search and performing the entire simulation using the GB/SA water model. Despite the increased CPU time required for this protocol, it was judged to be superior, based on the more accurate conformational space description of the solvated model.

Duplex DNA was modeled in the B-form conformation and was coordinated with fixed Na^+ counterions, positioned 2.5 Å from each anionic oxygen of the phosphates, thereby neutralizing the charge of such moieties, and minimizing unrealistic electrostatic interactions between the agent and phosphates. The DNA and the counterions in the monoadduct simulation were fixed using a force constant of 24 kcal/mol.[15] Each of the starting structures was subjected to a Monte Carlo simulation with all rotatable torsional angles on the azinomycin fragment including the purine N7-agent bond. Five thousand conformations were generated and minimized with our modified AMBER* force field,[8] now containing the charge improvement, directly in GB/SA water model (≤10000 iterations). Convergence in the Monte Carlo search was judged complete if the global minimum was found more than one time and, if in the final portion of the calculation, the random conformer generation was not able to reach deeper energy wells in the conformational space. Conformations that were >12 kcal/mol above the minimum were discarded from the Monte Carlo search.

Intercalative Docking. In our previous work,[9] we observed that the noncovalent association between azinomycin B and DNA is more productive for the occurrence of the first alkylation when the naphthalene intercalates into the duplex. To evaluate differences in the recognition between **1** and the four base-pair triplets, we modeled four duplex sequences containing an intercalation site compatible with the size of the naphthoate moiety. To take into account the intercalative recognition of the azinomycin with both sides of each triplet, we have generated four models **E**, **F**, **G**, and **H** containing the intercalation site preceded (upper) and followed (lower) by the same nucleotide triplet (Figure 4). The advantage of this computational expedient is that we can explore two intercalative recognition events in a single Monte Carlo simulation.

We adopted as a structural template for the nucleic acid receptor the experimentally determined crystallographic coordinates of a planar aromatic compound intercalated in the DNA major groove.[16] The intercalated duplex models were obtained from the PDB entry code 366d using the following steps: (1) We have removed the ligand, *N*-(2-(dimethylamino)ethyl)acridine-4-carboxamide, from the crystallographic model

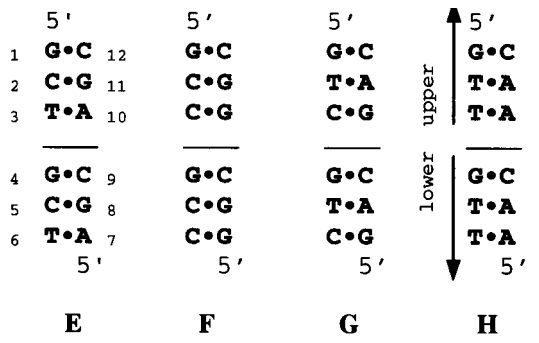
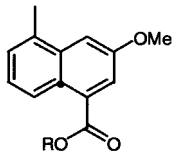


Figure 4. Structures and numbering schemes of the intercalation duplex models **E**, **F**, **G**, **H**. The black dash indicates the location of each intercalation site.

that contains only one DNA strand. (2) We have created the complementary strand while still maintaining the intercalation site. (3) We have used such a duplex model as a template for building our intercalated sequences, superimposing the upper and the lower triplets in B-conformation onto the bases directly adjacent to the intercalation site (i.e, 3rd−10th and 4th−9th nucleotides, see Figure 4). (4) The upper and lower triplet were linked using the backbone conformation of the crystallographic model. (5) The counterions were added with the same criterion adopted for the monoadduct models.



The specified naphthalene carbon (above) was constrained at a distance of 3.5 Å from the adjacent dG base of the 4th nucleotide with a force constant of 20 kJ/mol.[8] This constraint, defined between the C9′ of **1** and N1 dG of the 4th nucleotide, was considered weak enough to allow the naphthalene to make modest movements but strong enough to keep it intercalated. Constraining only one carbon of the naphthalene allowed the aromatic ring to flip by 180°, thereby obviating the need to perform two simulations to examine both orientations of the intercalated aromatic system.

We performed four Monte Carlo (MC) in situ docking simulations. Five thousand conformations were generated in GB/SA water using the naphthalene distance constraint, and then the energy of each was minimized without constraint, under the same solvation conditions.

Results and Discussion

To evaluate the alkylation probability in any complex or adduct that was modeled, we adopted a common set of criteria for the probability calculation that is based on the contemporaneous satisfaction of the distance and attack angle between the azinomycin B C10/C21 and the purine N7 nitrogens. We have fixed the distance threshold ≤3.8 Å and the angle acceptance ≥120°. These values are compatible with a reasonable attack of both epoxide and aziridine moieties by the purine N7 nitrogens.

The probability of cross-linking and/or alkylation for each duplex was evaluated by summation of the Boltz-

Table 3. Number of Energy Minimum Conformations Representative of ≥90% of the Boltzmann Population

1−DNA adduct	duplex sequence			
	A	**B**	**C**	**D**
MAB	1	1	1	5
MET	3	2	1	2
MAT	1	2	3	2
MEB	1	5	3	1

1·DNA intercalative complex	duplex sequence			
	E	**F**	**G**	**H**
	4	3	2	1

Table 4. Boltzmann Cross-Linking Probability in Percentage Computed at Room Temperature for Each Monoadduct

azinomycin−DNA monoadduct	duplex sequences			
	A	B	C	D
MAB	99.56	99.25	86.45	54.60
MET	100.00	97.17	0.00	0.00
MAT	0.00	4.03	0.00	0.00
MEB	0.00	0.00	0.00	0.00

mann probability of those minimum energy conformations that satisified the above filtering criteria.

An additional intra- and intermolecular hydrogen bond analysis of the most populated monoadduct and docked conformers has been carried out in order to understand the driving forces for interactions between the four DNA sequences and the drug in the cross-linking mechanism.

The most easily apparent result obtained from the conformational search is the low number of minimum energy conformations making up the 90% Boltzmann population at 298 K (Table 3).

It is illustrative to highlight that in several cases only one conformer dictates the entire set of properties of the DNA/drug system. In fact, the isolated molecule as already reported in our previous communication[8] was found to exist in several energy minimum conformations because the high degree of flexibility. We have now found that only a single (or very few) conformer seems to be responsible for the differences in the duplex selectivity of the azinomycin B, which significantly simplifies the interpretation of the results.

Monoadducts. The 16 monoadducts modeled have been analyzed applying the alkylation distance/angle criterion described above. In Table 4 we summarize the results obtained by the application of these criteria.

Comparison of the duplex **A** monoadduct conformational profile before[9] and after the charge improvement introduced into the AMBER* force field showed that the new simulations have dramatically changed the probability of cross-link formation from the MET adduct. This result is consistent with the recent observation[17] that found an epoxide-containing analogue of azinomycin alkylates guanosine residues in duplex DNA.

Comparison of the monoadduct cross-linking probabilities relative to the formation of two possible orientations of cross-link formation showed a dramatic difference in the MAB/MET versus the MAT/MEB probabilities. This confirmed that the reaction between azinomycin B and DNA must occur with the aziridine in bottom (nucleotide 8) and epoxide in top (nucleotide 3) positions, which is also in accordance with experimental observations.[4] The most interesting result of Table 4 is the differential probability of cross-link in the four DNA sequences **A**−**D**. To understand the relationships between the DNA and the drug in such MAB and MET adducts, we have performed a detailed analysis of the hydrogen bonding interactions (Table 5).

Table 5. Global Minimum Conformer Analysis MAB **1**−DNA Monoadducts Depicted in Figure 5

	bond type	acceptor atom	donor atom
	1−**A** Adduct, Boltzmann % = 99		
HB_1	intermolecular	O-3′	dA-2
HB_2	intermolecular	O-4	dC-10
HB_3	intermolecular	dG-9	H5
HB_4	intermolecular	O-14	dC-4
HB_5	intramolecular	O-6	OH-4
HB_6	intramolecular	O-2	H-9
	1−**B** Adduct, Boltzmann % = 98		
HB_1	intermolecular	O-3′	dA-2
HB_2	intermolecular	O-4	dC-10
HB_3	intermolecular	dG-9	H5
HB_4	intermolecular	O-14	dC-4
HB_5	intramolecular	O-6	OH-4
HB_6	intramolecular	O-2	H-9
	1−**C** adduct, Boltzmann % = 57		
HB_1	Intermolecular	O-3′	dA-2
HB_2	intermolecular	O-4	dC-10
HB_3	intramolecular	O-6	OH-4
HB_4	intramolecular	O-2	H-9
	1−**D** adduct, Boltzmann % = 50 no hydrogen bonds		

The analysis of the MAB adducts revealed interesting information that explained the different selectivity between the four sequences. The number of hydrogen bonds appeared to predict the different stability of the four very similar global minimum conformations (Figure 5).[18] The first two duplexes **A** and **B** with the MAB adduct showed the same high degree of intermolecular and intramolecular hydrogen bond network (Figure 5a,b). Two of the interactions with the DNA involved the central base-pair of the triplet (fourth and ninth nucleotides). This feature seemed to be crucial for the high MAB cross-linking probability in such adducts. Actually, duplex **C**, and much more so **D**, did not show the same hydrogen bonding pattern. In the global minimum of the duplex **C** we observed only two of the four intermolecular hydrogen bond interactions. In duplex **D** none of the hydrogen bonding interactions were present.

It is noteworthy that in the **C** and **D** sequences, the methyl group and carbonyl oxygen of the thymine (fourth nucleotide) present in the DNA major groove causes a steric and electrostatic repulsive interaction with the acetoxy group at C13 in the drug (Figure 5c,d). In **D**, this effect causes a consistent change in the drug C6−N5−C3−C2 torsion angle because of the lack of two key intramolecular hydrogen bonding contributions that so effectively stabilize the MAB adduct in the DNA duplexes **A** and **B**.

The difference in the hydrogen bond pattern was clearly reflected in the Boltzmann probability of each global minimum. In the case of the duplex **D**, the absence of any inter- and intramolecular hydrogen bonds made the global minimum significantly less populated with respect to the other three duplex monoadducts. In other words, it appeared that the probability of cross-linking was highly correlated with the hydrogen

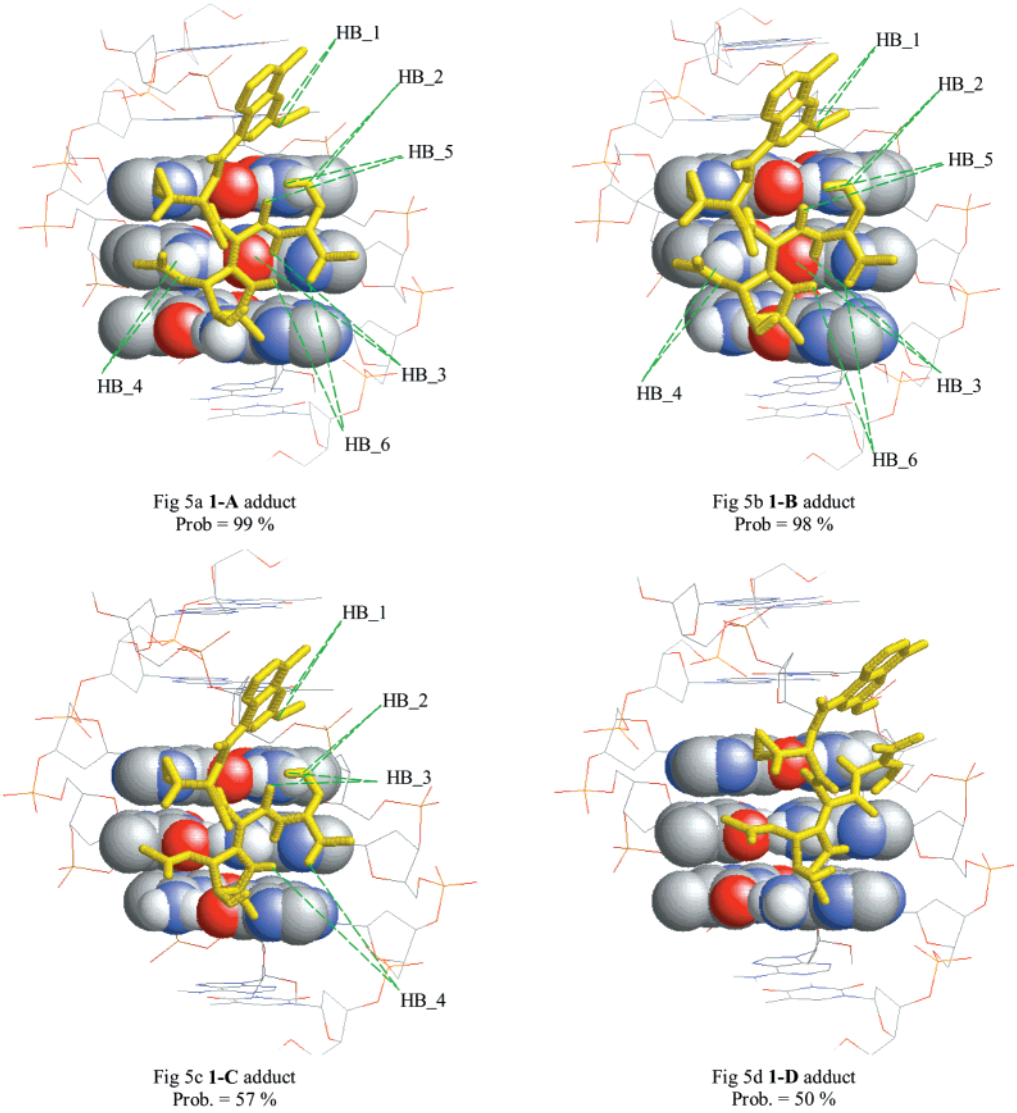


Figure 5. Hydrogen bond network in the energy minimum conformers of the MAB monoadducts.[9] Compound **1** covalently bound to the N7-purine of the eighth nucleotide of **A**, **B**, **C**, and **D** (Figure 3) is reported as yellow polytube; the reactive triplets are shown with the spacefill; sugars, backbone, and other bases are indicated with wireframe.[17] The dashed lines indicate the **1**−DNA hydrogen bond interactions (Table 5). For clarity, the Na^+ counterions and the interstrand hydrogen bonds have been omitted.

bonding network and stabilization of the global minimum conformers.

The MET structural analysis revealed that the minimum energy conformers are dissimilar with respect to DNA sequence. We observed two different clusters of conformers. In the first, we can consider the duplex **A** and **B** adducts to be similar even though the corresponding geometries are not exactly the same (Figure 6a,b). These conformers satisfy the cross-linking criterion, but show only four inter- and intramolecular hydrogen bonds (Table 6). In addition, these conformers do not contribute as much as the corresponding MAB adduct to the Boltzmann population. In other words, based solely on the number the hydrogen bonds observed, MET systems are not as effectively stabilized as the MAB systems. Conversely, there are other populated conformers higher in energy that help the MET system of **A** and **B** adducts to reach the high degree of cross-linking probability (Table 4).

The second cluster in the MET simulations includes the global minima with duplexes **C** and **D**. The common feature of these conformers is related to the location of the naphthalene moiety that interacts with the seventh and the eighth nucleotides preventing the satisfaction of the cross-linking criterion (Figure 6c,d). Actually, the aziridine moiety interacts mainly with the 10^{th} nucleotide and is always too far from the dG-N7 to achieve covalent bond formation. Consequently, in both MET cases we have zero cross-linking probability. A deep analysis of the hydrogen bond network at the fourth and ninth nucleotides has revealed one clear reason of such different behavior. The HB_1 in both **A** and **B** monoadducts involved the ammine hydrogen of the fourth nucleotide cytosine exposed into the major groove. In the other two monoadducts such a hydrogen bond donor is replaced by the adenine ammine hydrogen located on the same plane of the fourth and ninth nucleotides, but remarkably shifted in position and angle of productive hydrogen bonding to stabilize the same kind of energy minimum conformations of 1 with **A** and **B** sequences (Figure 6).

With respect to the MAT and MEB adducts, the very poor cross-linking probabilities make them uninteresting to investigate at the same level of detail as we have

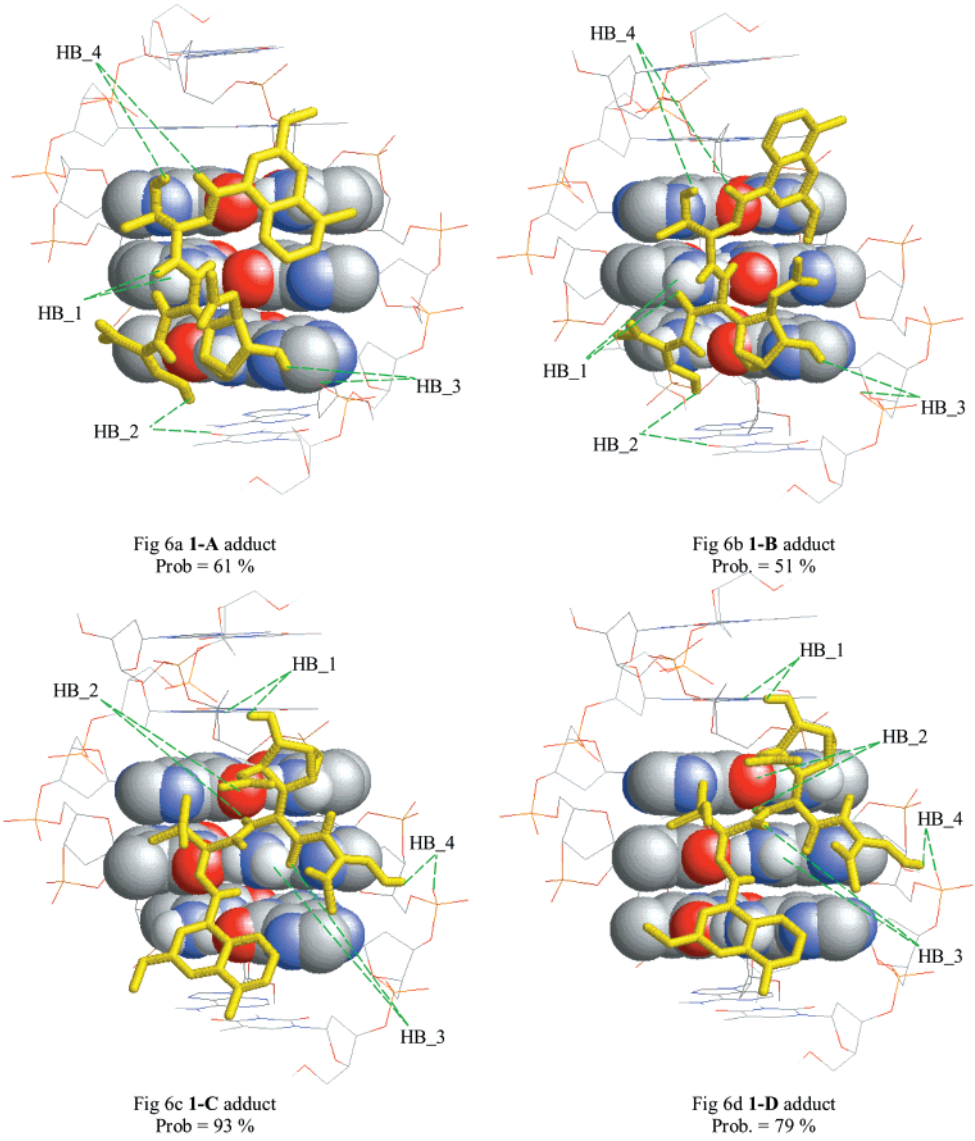


Figure 6. Hydrogen bond network in the energy minimum conformers of the MET monoadducts. Compound **1** covalently bound to the N7-dG of the third nucleotide of **A**, **B**, **C**, and **D** (Figure 3) is reported as yellow polytube; the reactive triplets are shown with the spacefill; sugars, backbone, and other bases are indicated with wireframe.[17] The dashed lines indicate the **1**−DNA hydrogen bond interactions (Table 6). For clarity, the Na^+ counterions and the interstrand hydrogen bonds have been omitted.

for the MAB and MET adducts. Geometrical analysis of the minimum energy structures for the MEB adducts revealed that the aziridine moiety was always too far from the purine N7 atom of interest. In the **A** and **B** duplexes, we observed a hydrogen bond network between the drug and the bottom nucleotides that forces contact of the aziridine with the fourth nucleotide, preventing a productive recognition of the C10 with the third nucleotide purine. In the **C** and **D** duplexes, the methyl group of dT at the fourth nucleotide sterically interacts with the naphthoyl group thereby preventing a sufficient level of "adhesion" of the drug with the DNA.

In addition, the MAT adducts cannot satisfy the distance criterion in the most populated conformers (Boltzmann pop >90%, see Table 3). Moreover, in adducts with **B**, **C**, and **D** duplexes, the purine N7 of the eighth nucleotide interacts with the C17-amide of **1** by hydrogen bonding, preventing a productive recognition of the epoxide. With duplex **A**, an intramolecular hydrogen bond between the C17-amide and the sp^2 carbonyl oxygen of the naphthoate moiety pushes the epoxide away from the DNA into the surrounding solvent.

Examining the higher energy conformers of MAT adducts, we observed the satisfaction of the cross-linking criterion only in the **B** system where, at the relative energy +2.15 kcal/mol, we found the first compatible conformer representing only 1.4% of population at 300 K.

DNA−Azinomycin Intercalative Recognition. The four docking simulations carried out for modeling the intercalative recognition between azinomycin B and DNA indicate the preponderance of **1** to interact with the lower triplet sequence (Figure 7).

Effective nucleophile−electrophile distances and angles expressed as alkylation probabilities are summarized in Table 7. Analysis of these numbers reveals a high degree of probability to form the MAB monoadduct for the duplexes **E** and **F**, whereas the two duplex sequences **G** and **H** are respectively evaluated to have

Table 6. Global Minimum Conformer Analysis MET **1**−DNA Monoadducts Depicted in Figure 6

	bond type	acceptor atom	donor atom
	1−**A** Adduct, Boltzmann % = 61		
HB_1	intermolecular	O-6	dC-4
HB_2	intermolecular	dT-7	OH-4
HB_3	intermolecular	phosphate-7	OH-12
HB_4	intramolecular	O-19	O-24
	1−**B** Adduct, Boltzmann % = 51		
HB_1	intermolecular	O-6	dC-4
HB_2	intermolecular	O-4	dC-5
HB_3	intermolecular	phosphate-7	OH-12
HB_4	intramolecular	O-19	O-24
	1−**C** Adduct, Boltzmann % = 93		
HB_1	intermolecular	dA-2	OH-12
HB_2	intermolecular	dG-3	H-16
HB_3	intermolecular	O-6	dA-9
HB_4	intermolecular	phosphate-8	OH-4
	1−**D** Adduct, Boltzmann % = 79		
HB_1	intermolecular	O-12	dA-2
HB_2	intermolecular	dG-3	H-16
HB_3	intermolecular	O-6	dA-9
HB_4	intermolecular	phosphate-8	OH-4

Table 7. Monoadduct Alkylation Probability versus Sequences **E**, **F**, **G**, and **H** at the Stage of Intercalative Recognition

	intercalative duplex			
alkylation description	**E**	**F**	**G**	**H**
C10/N7-dPu-8th aziridine bottom	98.61	98.00	43.12	0.00
C21/N7-dG-4th epoxide top	0.07	0.00	0.00	0.00
C10/N7-dG-4th aziridine top	0.00	0.00	0.00	0.00
C21/N7-dPu-8th epoxide bottom	0.00	0.00	0.00	0.00

medium and no probability of covalent bond formation. A low probability for the formation of MET adduct (<1%) was found in the case of the duplex **E** and is not related with the global minimum conformation shown in Figure 7.

A detailed analysis of the hydrogen bond network of the four **1**·DNA intercalated complexes (Table 8) revealed two possible origins of differences in the simulation results. Both are attributable to the fifth nucleotide. The mutation of dC with dT in going from sequences **E**/**F** to **G**/**H** is responsible for the lack of a fourth stabilizing intermolecular hydrogen bond between the

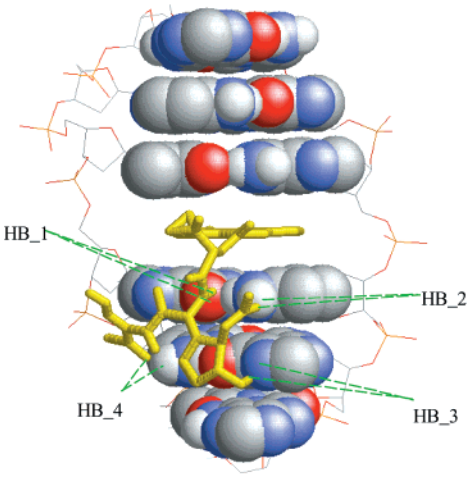


Fig 7a **1**·**A** complex
Prob. = 51 %

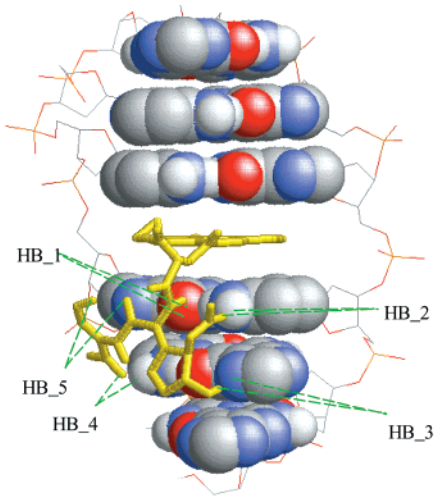


Fig 7b **1**·**B** complex
Prob. = 94 %

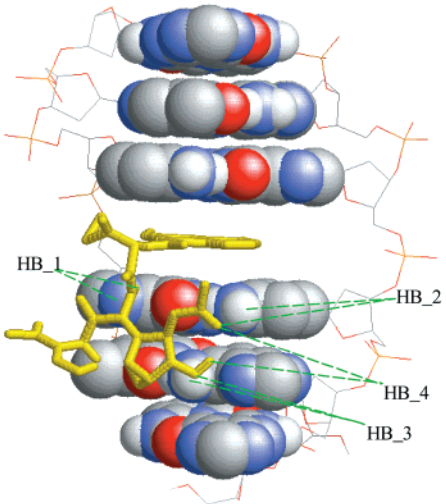


Fig 7c **1**·**C** complex
Prob. = 34 %

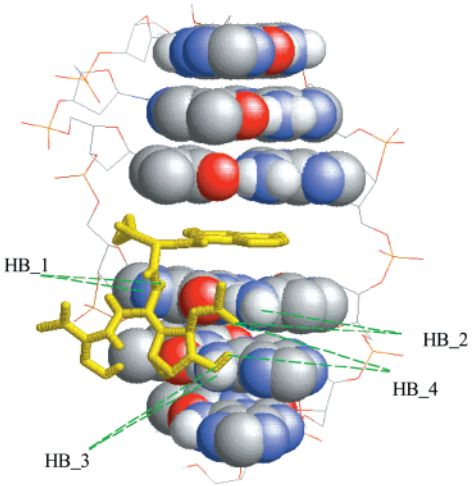


Fig 7d **1**·**D** complex
Prob. = 93 %

Figure 7. Hydrogen bond network in the energy minimum conformers of the intercalative complexes. Compound **1** complexed with the intercalative models of **A**, **B**, **C**, and **D** (Figure 4) is reported as yellow polytube; the reactive triplets are shown with the spacefill; sugars and backbone are indicated with wireframe.[17] The dashed lines indicate the **1**·DNA hydrogen bond interactions (Table 7). For clarity, the Na^+ counterions and the interstrand hydrogen bonds have been omitted.

Table 8. Global Minimum Conformer Analysis of the **1**·DNA Intercalative Complexes

	bond type	acceptor atom	donor atom
	1·E Complex, Boltzmann % = 50		
HB_1	intermolecular	dG-4	H-16
HB_2	intermolecular	O-15	dC-9
HB_3	intermolecular	dG-8	OH-12
HB_4	intermolecular	O-2	dC-5
	1·F Complex, Boltzmann % = 94		
HB_1	intermolecular	dG-4	H-16
HB_2	intermolecular	O-15	dC-9
HB_3	intermolecular	dG-8	OH-12
HB_4	intermolecular	O-2	dC-5
HB_5	intramolecular	O-6	OH-4
	1·G Complex, Boltzmann % = 34		
HB_1	intermolecular	dG-4	H-16
HB_2	intermolecular	O-15	dC-9
HB_3	intermolecular	O-12	dA-8
HB_4	intramolecular	O-15	OH-12
	1·H Complex, Boltzmann % = 92.99		
HB_1	intermolecular	dG-4	H-16
HB_2	intermolecular	O-15	dC-9
HB_3	intermolecular	O-12	dG-8
HB_4	intramolecular	O-15	OH-12

drug and the DNA (Figure 7). The C2 carbonyl oxygen of azinomycin B is the acceptor and the cytosine amino hydrogen exposed in the major groove is the donor in such an intermolecular hydrogen bonding interaction. Replacing cytosine with thymine at the fifth nucleotide (sequences **E**/**F** to **G**/**H**) changes the major groove amino group to a carbonyl oxygen, which is no longer a suitable hydrogen bond partner for the C2 carbonyl oxygen of the agent. In addition, the thymine C4 carbonyl interacts with the agent C2 carbonyl repulsively and significantly changes the agent conformation from one compatible with alkylation to one that is incompatible.

The second reason for the low probability for intercalative sequences **G** and **H** to undergo alkylation is that the thymine of the fifth nucleotide positions the C5-methyl group within the major groove so as to sterically interact with the enolic C1−C4 portion of the drug. In the minimum energy complexes (Figure 7), there is a consistent and clearly identifiable conformational effect in the enolic C1−C4 side chain of the drug. As a consequence, the aziridine moiety in the most populated **1**·DNA complex geometries is moved away from the N7-guanine of the eighth nucleotide, reducing the alkylation probability of **G** and **H** with respect to the **E** and **F** duplex sequences (Table 7).

In these docking simulations, we have taken into account one possible mechanism for the monoalkylation, which considers necessary the intercalation of the naphthalene moiety. However, we can also assume that the drug recognizes the DNA without intercalation. Experimental evidence on this issue seems to point to an extremely weak intercalation,[17] although the evidence is not unequivocal.

In a recent theoretical study,[19] the authors predicted the reactivity of purine bases in different chemical environments by HOMO/LUMO calculations. Along the 5′ to 3′ direction, the reactivity of the purine bases contained in the four triplets of the duplex sequences **E**, **F**, **G**, and **H** can be summarized as follows (reactive base underlined): 5′-G̲G-3′ > 5′-A̲G-3′ = 5′-G̲A-3′ > 5′-G̲C-3′ > 5′-G̲T-3′ > 5′-A̲A-3′. Considering this scale of reactivity, we observed that the duplexes **E**, **F**, and **G** show a higher probability to be alkylated initially on the N7 of the eighth nucleotide (bottom adduct). In contrast, duplex **H** has a greater likelihood to be alkylated on the third nucleotide (top adduct), but in this reported reactivity scale, 5′-GT-3′ and 5′-AA-3′ are judged poorly reactive. On the basis of this ab initio theoretical reactivity, the probability of the initial alkylation event in this intercalative model would be as follows: **F** > **E** = **G** ≫ **H**. These conclusions are in good agreement with the results reported in Table 7, but do not match perfectly with the qualitative experimental results.[5]

Correlation with Experimental Reactivity. The existing experimental data for sequence reactivity relationships between the four DNA triplets under examination and the agent **1** are only qualitative, so an attempt to perform a quantitative study will depend on the availability of more detailed experimental data. On the other hand, in our computational study we have clearly found the same trend between the experimental data and the reactivity predictions.

The two most reactive DNA triplets, 5′-GCT-3′ and 5′-GCC-3′, respectively modeled in duplex sequences **A**/**E** and **B**/**F**, show a high probability of alkylation or cross-linking probability. Even though our estimated cross-linking reactivity is identical for both the MAB and MET monoadducts in duplexes **A** and **B**, the analysis of the intercalative recognition with **E** and **F** indicates only one probable cross-linking mechanism via an initial alkylation at the bottom base to give the MAB adduct. The MAB adduct, in agreement with other observations,[4,19] is more compatible with the conformational properties of the most populated intercalative **1**·**E** and **1**·**F** complexes.

On the other hand, we observed a good correlation between the experimentally classified poor and zero cross-linking reactivity of the triplets 5′-GTC-3′ and 5′-GTT-3′, respectively modeled as duplex sequences **C**/**G** and **D**/**H**, and the low level of MAB/MET cross-linking and the alkylation probabilities obtained in these studies. The analysis of the MAB adducts (Figure 5 and Table 5) clearly indicates that the diminished (duplex **C**) or absent (duplex **D**) hydrogen bonding network is the primary reason for the lack of reactivity. In this context, the steric effect of the C5-methyl group of thymine at the fourth nucleotide position plays an important role in the destabilization of the reactive cross-linking conformation of the MAB monoadduct.

Conclusions

Using ab initio methods we have developed a new set of charges for the AMBER* force field implemented in the MacroModel package that is useful for accurate molecular mechanics simulation of biologically relevant N7-alkylated purine nucleotides. This set of charges can be applied to the study of the conformational properties of adducts between DNA and common alkylating agents and to the drug design of more potent analogues of major groove binding antitumor agents.

We have tested and applied this modified force field to the study of the sequence dependent reactivity of the DNA cross-linking agent azinomycin B by modeling duplex sequences containing four experimentally known DNA recognition triplets. We have found a good cor-

relation between the estimated and the experimental reactivity. The results presented herein represent a substantial and significant development in our goal to demonstrate the predictive potential of our protocol for Monte Carlo simulation of DNA/drug adducts, so that they can be confidently applied in the design of new analogues of **1**.

After the comparison of all possible alkylation attacks of the drug to the DNA, the mechanism that involves the aziridine "bottom" and the epoxide "top" is confirmed as the most probable orientation of the drug. The different hydrogen bond network of the monoadducts and the intercalative complexes between the drug and the duplex sequences is mostly responsible for the different cross-link reactivity. In addition, a steric hindrance by the major groove exposed methyl moiety of thymine-based triplets plays an important role in the lack of the reactivity.

Information coming from this study will be important for the design of more potent or DNA sequence-selective compounds based on the skeleton of the azinomycins.

Acknowledgment. This research was supported with funding from NATO (CRG 970160), the National Institutes of Health (CA 65875), and the Scientific-Technological Research P.O.P., Calabria 1994-1999. S.A. is grateful to the FIRC for a short-term fellowship (2001) and to the Young Researcher Project of the CNR. We thank Professor Domenicantonio Rotiroti for arranging a CNR Visiting Scientist Fellowship for R.S.C (1999).

Supporting Information Available: The substructures for compounds **4** and **5** in MacroModel force field format are included in this section. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Nagaoka, K.; Matsumoto, M.; Oono, J.; Yokoi, K.; Ishizeki, S.; Nakashima, T. Azinomycins A and B, New Antitumor Antibiotics. I. Producing Organism, Fermentation, Isolation and Characterization. *J. Antibiot.* **1986**, *39*, 1527−1532. Yokoi, K.; Nagaoka, K.; Nakashima, T. Azinomycins A and B, New Antitumor Antibiotics. II. Chemical Structures. *Chem. Pharm. Bull.* **1986**, *34*, 4554−4461.

(2) Ishizeki, S.; Ohtsuka, M.; Irinoda, K.; Kukita, K.-I.; Nagoka, K.; Nakashima, T. Azinomycins A and B, New Antitumor Antibiotics. III. Antitumor Activity. *J. Antibiot.* **1987**, *40*, 60−65.

(3) For a review on DNA cross-linking agents, see: Rajski, S. R.; Williams, R. M. DNA Cross-Linking Agents as Antitumor Drugs. *Chem. Rev.* **1998**, *98*, 2723−2795.

(4) Fujiwara, T.; Saito, I.; Sugiyama, H. Highly Efficient DNA Interstrand Cross-linking Induced by an Antitumor Antibiotic, Carzinophilin. *Tetrahedron Lett.* **1999**, *40*, 315−318.

(5) Armstrong, R. W.; Salvati, M. E.; Nguyen, M. Novel Interstrand Cross-Links Induced by the Antitumor Antibiotic Carzinophilin/Azinomycin B. *J. Am. Chem. Soc.* **1992**, *114*, 3144−3145.

(6) Coleman, R. S. Issues of Orthogonality and Stability: Synthesis of the Densely Functionalized Heterocyclic Ring System of the Antitumor Agents Azinomycins A and B. *Synlett* **1998**, 1031−1039.

(7) Coleman, R. S.; Li, J.; Navarro, A. Total Synthesis of Azinomycin A. *Angew. Chem., Int. Ed.* **2001**, *40*, 1736−1739. Coleman, R. S.; Kong, J.-S.; Richardson, T. E. Synthesis of Naturally Occurring Antitumor Agents: Stereocontrolled Synthesis of the Azabicyclic Ring System of the Azinomycins. *J. Am. Chem. Soc.* **1999**, *121*, 9088−9095. Coleman, R. S.; Richardson, T. E.; Carpenter, A. J. Synthesis of the Azabicyclic Core of the Azinomycins: Introduction of Differentiated *trans*-Diol by Crotylstannane Addition to Serinal. *J. Org. Chem.* **1998**, *63*, 5738−5739. Coleman, R. S.; Kong, J.-S. Stereocontrolled Synthesis of the Fully Elaborated Aziridine Core of the Azinomycins. *J. Am. Chem. Soc.* **1998**, *120*, 3538−3529. Coleman, R. S.; Carpenter, A. J. Coleman, R. S.; Carpenter, A. J. Synthesis of the Aziridino-[1,2-*a*]pyrrolidine Substructure of the Antitumor Agents Azinomycin A and B. *J. Org. Chem.* **1992**, *57*, 5813−5815.

(8) Alcaro, S.; Coleman, R. S. Molecular Modeling of the Antitumor Agents Azinomycins A and B: Force-Field Parametrization and DNA Cross-Linking Based Filtering. *J. Org. Chem.* **1998**, *63*, 4620−4625.

(9) Alcaro, S.; Coleman R. S. A Molecular Model For DNA Cross-Linking By The Antitumor Azinomycin B. *J. Med. Chem.* **2000**, *43*, 2783−2788.

(10) Jenkins, T. C.; Hurley, L. H.; Neidle, S.; Thurston, D. E. Structure of a covalent DNA minor groove adduct with a pyrrolobenzodiazepine dimer: evidence for sequence-specific interstrand cross-linking. *J. Med. Chem.* **1994**, *26*, 4529−4537.

(11) McDonald, D. Q.; Still, W. C. AMBER* Torsional Parameters for the Peptide Backbone. *Tetrahedron Lett.* **1992**, *33*, 7743.

(12) MacroModel version 5.5 for SGI. Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. MacroModel − An Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics. *J. Comput. Chem.* **1990**, *11*, 440−467.

(13) *Spartan Silicon Graphics*, Version 5.1.1; Wavefunction, Inc.: Irvine, CA.

(14) MAB = monoalkylation aziridine bottom at the purine N7 of eighth nucleotide; MAT = monoalkylation aziridine top at the N7-dG of the third nucleotide; MEB = monoalkylation epoxide bottom at the N7-dG of the third nucleotide; MET = monoalkylation epoxide top at the purine N7 of the eighth nucleotide.

(15) This force constant value is the default value suggested by the MacroModel Monte Carlo setup procedure.

(16) Todd, A. K.; Adams, A.; Thorpe, J. H.; Denny, W. A.; Wakelin, L. P.; Cardin, C. J. Major Groove Binding and "DNA-Induced" Fit in the Intercalation of a Derivative of the Mixed Topoisomerase I/II Poison N-(2-(Dimethylamino)ethyl)acridine-4-carboxamide (DACA) into DNA: X-ray Structure Complexed to d(CG(5-BrU)ACG)$_2$ at 1.3 Å Resolution. *J. Med. Chem.* **1999**, *42*, 536−540.

(17) Zang, H.; Gates K. S. DNA Binding and Alkylation by the "Left Half" of Azinomycin B. *Biochemistry* **2000**, *39*, 14968−14975.

(18) The graphic Figures 5, 6, 7 have been realized using the software RasMol 2.7.1 running on an 800 MHz IBM-compatible personal computer.

(19) Sugiyama, H.; Saito, I.; Sugiyama, H. Theoretical Studies of GG-Specific Photocleavage of DNA via Electron Transfer: Significant Lowering of Ionization Potential and 5′-Localization of HOMO of Stacked GG Bases in B−Form DNA. *J. Am. Chem. Soc.* **1996**, *118*, 7063−7068.

JM011040W