DNA Sequence-Specific Reading by Echinomycin: Role of Hydrogen Bonding and Stacking Interactions

José Gallego,† F. J. Luque,‡ Modesto Orozco,§ Carolina Burgos,‖ Julio Alvarez-Builla,‖ M. Melia Rodrigo,⊥ and Federico Gago*,†

Departamentos de Fisiología y Farmacología, Química Orgánica, and Química Física, Universidad de Alcalá de Henares, 28871 Madrid, Spain, and Departament de Farmàcia (Unitat Físico-Química), Facultat de Farmàcia, and Departament de Bioquímica i Fisiología, Facultat de Química, Universitat de Barcelona, 08028 Barcelona, Spain

Received December 10, 1993®

The binding of echinomycin to DNA hexamers of the form GpApXpZpTpC, where the central XpZ step can be CpG, TpA, GpC, or ApT, has been studied by molecular modeling and molecular mechanics techniques. Interaction energies have also been calculated for the complexation of echinomycin with sequences containing the preferred central CpG step and different flanking base pairs. Besides, two more sets of sequences incorporating either 2,6-diaminopurine (DAP) or hypoxanthine in place of adenine or guanine, respectively, have been examined. The aim of this work was to evaluate the relative importance of hydrogen-bonding and stacking interactions in the association of echinomycin with DNA and further rationalize the experimental evidence. The results of these calculations are in consonance with available data from footprinting experiments and appear to support our previous hypothesis that, in addition to the crucial intermolecular hydrogen bonds in the central region, the stacking interactions involving the quinoxaline-2-carboxamide chromophores of the drug and the DNA base pairs play an important role in modulating the binding specificity of this bisintercalating antitumor antibiotic. This is most clearly seen when sequences with similar minor-groove environments are compared (*e.g.* CpI vs TpA or CpG vs TpDAP). The dipole moment of *N*-methylquinoxaline-2-carboxamide has been measured (μ = 4.15 ± 0.03 D) and compares very well with the calculated value (μ = 4.14 D). The fact that G:C, I:C, A:T, and DAP:T base pairs are shown to be endowed with distinct van der Waals and electrostatic stacking properties with respect to this heteroaromatic ring system could have important implications for the design of novel DNA mono- and bis-intercalating agents.

Introduction

Echinomycin is a staple-shaped antitumor antibiotic from *Streptomyces echinatus* which binds to DNA as a bifunctional intercalator.[1] Its two quinoxaline rings bisintercalate into double helical DNA whereas the inner part of the bicyclic depsipeptidic linker faces the minor-groove region of the two base pairs comprised between the chromophores where it establishes a number of hydrogen bonds with the DNA bases (Figure 1). The resulting interaction has been compared to that of a vice clamping the inner bases.[2]

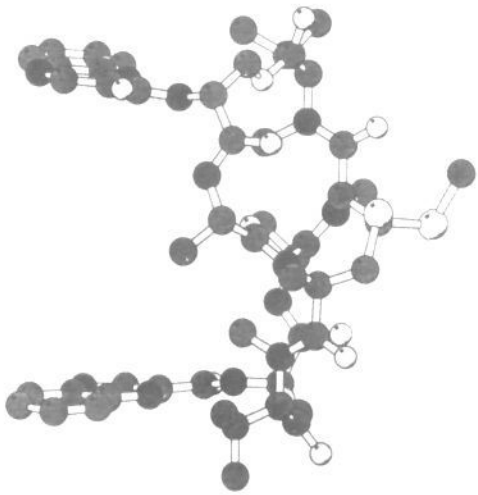
In common with its close relative triostin A, echinomycin shows a marked preference for binding to CpG steps.[3,4] In this selectivity it is commonly accepted that the hydrogen bonds established between the NH and carbonyl groups of the antibiotic's alanines and the N3 and 2-amino groups of the guanines play a predominant role. Some preference for A:T as the flanking base pairs has also been reported for echinomycin[3,4] although no definitive rules have been established in this respect.[5] Interestingly, when a purine nucleoside is on the 5′ side of a CpG binding site, its base ring can rotate 180° about the glycosidic bond, giving rise to a Hoogsteen base pairing scheme,[6] as shown experimentally by X-ray diffraction[7-9] and NMR analyses on several oligonucleotides.[10-12] For longer DNA sequences, however, no evidence exists that Hoogsteen pairing occurs at internal base pairs surrounding an isolated binding site.[5]

In a recent theoretical study,[13] we have suggested that the origin of the different conformational behavior of the bases flanking the echinomycin binding site in $d(ACGT)_2$ and $d(TCGA)_2$[10] could lie in the drastic change in dipole moment taking place in an A:T base pair when going from Watson–Crick to Hoogsteen pairing, which results in better stacking interactions with the drug's chromophores in the former complex but not in the latter. This same study highlighted an unfavorable electrostatic interaction between the quinoxaline-2-carboxamide system of the antibiotic and the sandwiched G:C base pairs, which led us to hypothesize that modulation of the dipole moments of the intercalating chromophores could be an additional element to be taken into account in the recognition process.

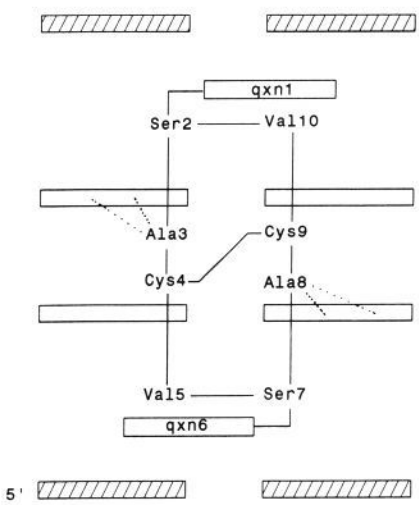
Traditional work in structure–affinity relationships (SAR) for the quinoxaline family of antibiotics has dealt with the effects that introduction of new substituents or removal of existing ones have on the binding properties of a given drug, as assessed mainly by DNA footprinting experiments.[14-17] The best known examples are probably those provided by des-*N*-tetramethyltriostin A (TANDEM) and [*N*-MeCys[3],*N*-MeCys[7]]TANDEM (CysMeTANDEM), triostin A analogues lacking either all or half of the *N*-methyl groups of the cysteines and valines, respectively, which bind better to TpA.[18,19] More recently, a series of elegant experiments utilizing molecular biology techniques have gone one step further: modification of the DNA itself. This has turned out to be another very important aspect of SAR studies. Thus, it has been shown that removal of the exocyclic amino group of guanine (by

† Departamento de Fisiología y Farmacología, Universidad de Alcalá de Henares.
‡ Departament de Farmàcia, Universitat de Barcelona.
§ Departament de Bioquímica i Fisiología, Universitat de Barcelona.
‖ Departamento de Química Orgánica, Universidad de Alcalá de Henares.
⊥ Departamento de Química Física, Universidad de Alcalá de Henares.
® Abstract published in *Advance ACS Abstracts*, April 15, 1994.

0022-2623/94/1837-1602$04.50/0



a



b

Figure 1. (a) Side view[43] of the echinomycin molecule showing the two quinoxaline chromophores and the bridging bicyclic depsipeptide. Hydrogen atoms are not displayed for clarity. (b) Schematic representation of the interaction of echinomycin with a DNA tetranucleotide as seen from the minor groove. The DNA base pairs flanking the bisintercalation site are hatched. Each alanine residue can engage in a maximum of two intermolecular hydrogen bonds; the actual number depends on base composition and sequence.

replacement of this base with hypoxanthine) leads to the loss of echinomycin's specificity for CpG steps,[20] lending further credence to the crucial role played by this group in the interaction of this ligand with the DNA minor groove. On the contrary, introduction of an extra amino group in the minor groove of AT regions (by replacing adenine with 2,6-diaminopurine [DAP]) favors echinomycin binding to any pyrimidine–purine (YpR) combination other than the usual CpG step. Furthermore, not only is the selectivity drastically changed but also the affinity of echinomycin for these new binding sites in the modified DNA appears to be at least 1 order of magnitude greater than that for normal DNA.[21]

All of these findings reveal a complex interaction behavior that is still far from being completely understood even though X-ray[7-9] and NMR techniques[10-12] have provided a wealth of structural data on some of these complexes. Computational chemistry tools, on the other hand, have also been used in an attempt to assess the relative importance of the considerable number of forces and conformational variables involved in the binding of this family of compounds to DNA.[13,22] In order to explore to what extent hydrogen-bonding and stacking interactions play a role in determining the binding preferences of echinomycin, we have generated a series of molecular models for the 1:1 complexes of echinomycin with standard and modified DNA sequences in order to evaluate the different contributions to their relative stability. This knowledge should aid in our understanding of drug–DNA interactions at the molecular level and may be useful for the design of new agents which will bind selectively to predetermined sequences.

In the present study[23] we focus on 13 double helical DNA hexamers in which echinomycin sandwiches the central dinucleotide step. These oligomers are intended to represent a section of a longer DNA fragment; for this reason a common terminal G:C base pair on both 5′ and 3′ ends was added to the central tetranucleotide where the most intimate interactions with the drug take place. One advantage of this procedure is that any possible end effects will not affect the central region where the drug is bound. These hexanucleotides can be grouped in four different categories: (i) a first family of four sequences containing GpApXpZpTpC as the basic unit where the central XpZ step can be CpG, TpA, GpC, or ApT, (ii) a second subset of GpXpCpGpZpC sequences, where the canonical CpG central step is conserved and the X–Z combinations are T–A, G–C, and C–G, (iii) three modified DNA sequences of general formula IpApCpIpTpC in which inosine substitutes for guanosine in either one or both strands, and (iv) three modified DNA sequences containing DAP in place of adenine, GpDpTpDpTpC, GpDpTpGpTpC, and GpDpCpDpTpC. The whole set of sequences (hereafter referred to by their central tetranucleotide) thus contains every combination of YpR and RpY binding sites experimentally probed by echinomycin and provides information both on the stacking interactions between the quinoxaline-2-carboxamide system and any DNA base pair, and on all the possible hydrogen-bonding arrangements between the depsipeptide and the DNA atoms in the minor groove.

Results and Discussion

The refined models presented here provide structural data on the complexes of echinomycin with 13 different DNA sequences and, more importantly, allow for a detailed computational description of the molecular interactions in terms of distinct energy contributions. When the interaction energies between echinomycin and each hexanucleotide are compared (Figure 2), it can be clearly seen that the sequences containing the central CpG step are favored with respect to any other natural YpR or RpY combination, the differences between them being a reflection of the different nature of the flanking bases. This result is in consonance with the large body of experimental evidence showing that this is the consensus binding site for echinomycin in natural DNA, and is usually attributed to the hydrogen bonds formed between the alanine residues of echinomycin and both the N3 atoms and 2-NH_2 groups of guanines in the DNA minor groove.[1] In fact, substitution

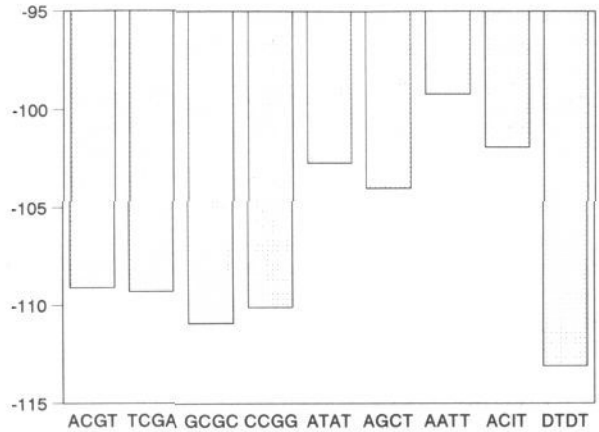


Figure 2. Total interaction energies (kcal mol^{-1}, calculated with no distance cutoff for the nonbonded pair list and with a relative permitivitty of $\epsilon = 4r_{ij}$) between echinomycin and the DNA hexamers. (The different oligonucleotides are identified by their central four base pair sequences. The results for the hybrid molecules studied are not shown for clarity but are included in Figure 4.)

Figure 3. Contributions (kcal mol^{-1}) of distinct echinomycin fragments to the interaction energies shown in Figure 2 (for each complex, from left to right): (1) the *N*-methylquinoxaline-2-carboxamide aromatic systems, made up by the quinoxaline rings and the $NHC_\alpha H_\alpha$ atoms of the serine residues; (2) the alanine residues; (3) the valine residues and the remaining serine atoms; and (4) the cysteine residues.

of inosine for guanosine in the central binding step leads to a substantial reduction in interaction energy, in accord with footprinting results for CpI steps.[20] Likewise, the RpY combinations studied here also appear as poor binding sites for echinomycin, in agreement with the experimental evidence,[3,4] despite the fact that the O2 of the pyrimidine base occupies an equivalent position to that of the N3 of purine in YpR steps and is therefore available for hydrogen bonding with the alanines' NH groups. The most favorable binding sites, however, are those in which the adenines have been replaced with DAP, and most notably the sequence DpTpDpT. This result also agrees with recent findings showing that in a *tyrT* DNA fragment thus modified these new sites are protected from DNAse cleavage very efficiently. What was not clear in this experimental work[21] is why, in this modified DNA, echinomycin no longer affords protection at the canonical CpG sites and why the affinity of the antibiotic for DAP-containing DNA is at least 1 order of magnitude larger than for natural DNA.

Decomposition of the interaction energies between echinomycin and the hexanucleotides provides a deeper understanding of the principal interactions involved in each complex. For this purpose, the echinomycin molecule was divided into four fragments: the *N*-methylquinoxaline-2-carboxamide systems, the alanine residues, the "terminal" residues (valines and serines), and the cysteines, whose relative contributions to the interaction energies are graphically shown in Figure 3. By looking at the differences between the complexes it is readily apparent that the binding selectivity mostly arises from the interactions involving both the alanine residues and the quinoxaline-2-carboxamide chromophores of echinomycin. Furthermore, a good linear correlation exists between these two energy components and the total interaction energy (Figure 4). The different hexamers appear clustered in two distinct families: the first one (lower left hand corner) represents a subset of good binding sites for echinomycin, presenting both a central dinucleotide step endowed with full hydrogen-bonding capabilities in the minor groove and an arrangement of base pairs giving rise to an overall favorable stacking interaction with the antibiotic's chromophores; the second one (upper right hand corner) shares

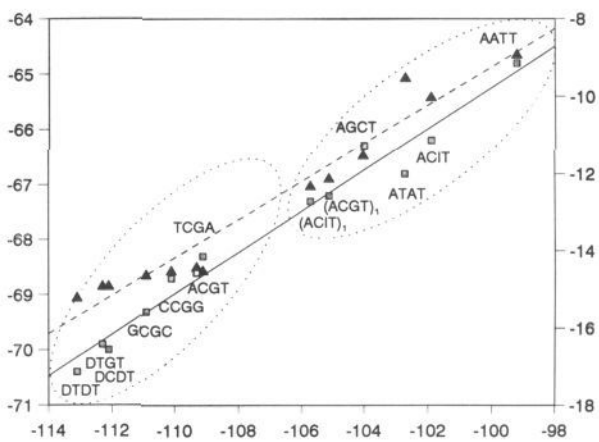


Figure 4. Correlation between the total interaction energies and the contributions of the alanine (▲, $r^2 = 0.976$) and *N*-methylquinoxaline-2-carboxamide (□, $r^2 = 0.986$) fragments of echinomycin for the 13 complexes studied. [$(ACIT)_1$ and $(ACGT)_1$ stand for the "Watson" and "Crick" inosine-substituted strands, respectively. D stands for 2,6-diaminopurine.] All units are kcal mol^{-1}.

poorer hydrogen-bonding possibilities and overall weaker interactions with the quinoxaline-2-carboxamide systems.

The differences in interaction energies involving the alanine residues are mainly due to the electrostatic and hydrogen bonding terms since the van der Waals contribution is rather similar in all the complexes (data not shown). The suitable geometrical disposition of hydrogen-bond donor and acceptor atoms in the minor groove is the main factor determining the binding selectivity of echinomycin for the central dinucleotides. This can be clearly construed from Figures 2 and 3, where the most favored complexes are shown to be those in which the antibiotic can bind to both the N3 atoms and the NH_2 groups of the purine bases (CpG and TpD) through the HN and CO of alanines. Other central YpR steps with only the N3 atom of purines (TpA and CpI), or putative RpY sites with only the O2 atom of pyrimidines (ApT and GpC), available for hydrogen bonding are less favored because only two

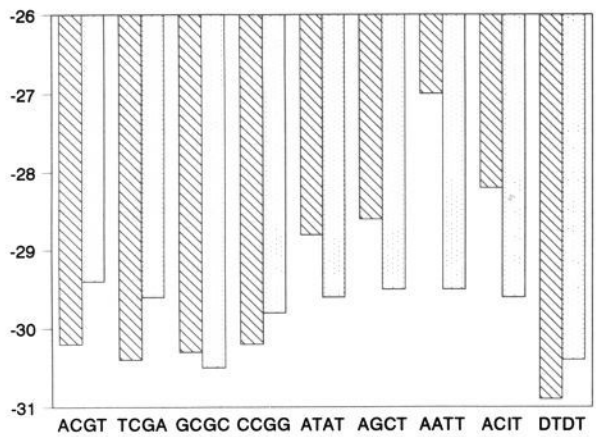


Figure 5. van der Waals contributions (kcal mol^{-1}) to the stacking interaction energies between the *N*-C_α-quinoxaline-2-carboxamide chromophores of echinomycin and either the central (hatched bars) or the flanking (dotted bars) base pairs in the different complexes studied. [Note that a general trend can be established: for the central base pairs, YpR steps are preferred over RpY steps, with base pairs containing exocyclic amino groups favored over those lacking them; for the flanking base pairs the differences are much smaller, with R bases preferred over Y bases on the 5′ side of the central step, and again base pairs containing exocyclic amino groups more favored than those not having them. These energies correlate with the *N*-methylquinoxaline-2-carboxamide–DNA interaction energies of Figures 3 and 4, indicating that differences in the total interaction energies of Figure 2 are mainly determined by the van der Waals term of the stacking interactions, in addition to the hydrogen bonds involving the alanines. The electrostatic term is comparatively much smaller, but its analysis is deemed necessary for a full understanding of the forces involved in these molecular associations (*cf.* Figure 6).]

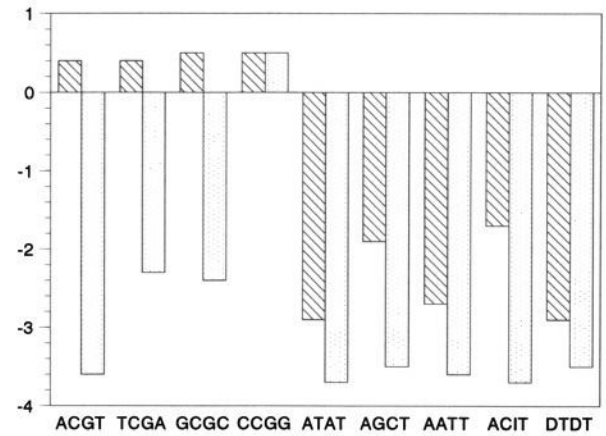


Figure 6. Electrostatic contributions (kcal mol^{-1}) to the stacking interaction energies between the *N*-C_α-quinoxaline-2-carboxamide chromophores of echinomycin and either the central (hatched bars) or the flanking (dotted bars) base pairs in the different complexes studied. [These values were calculated using a point charge interaction model and the C1′ atoms of deoxyribose and C_α atoms of serines as buffers in order to achieve electrical neutrality.[13] Calculations not including these atoms yielded the same relative results (data not shown). A dielectric constant of 1 was utilized since the chromophores are closely facing each other with no other atoms between them.]

Table 1. Angles (in degrees) Formed between the Dipole Moment Vector of the *N*-C_α-Quinoxaline-2-carboxamide Systems and That of Either the Central or the Flanking Base Pairs[a] in the Different Complexes

	central	flanking		central	flanking
ACGT	15.9	154.6	AGCT	160.5	155.1
TCGA	15.7	79.9	AATT	141.4	154.6
GCGC	15.9	132.8	ACIT	34.2	153.3
CCGG	15.4	28.1	DTDT	28.5	137.3
ATAT	65.2	154.2			

[a] Averaged over both sides of the complexes. The point charge interaction energy values of Figure 6 are in good agreement with the *magnitude* (*cf.* Figure 7) and *relative orientation* of the dipole moments of the stacked systems, which grossly represent the polarity of the charge distributions. Values close to 180° represent antiparallel arrangements of vectors giving rise to favorable electrostatic interactions; values closer to 0° denote parallel arrangements and therefore unfavorable electrostatic interactions.

hydrogen bonds can be formed. This factor contributes to explaining why echinomycin prefers the consensus CpG binding site over TpA or RpY sites in natural DNA and why removal of the exocyclic amino group from the minor groove by replacing guanosine with inosine abolishes specific binding of echinomycin to DNA.[20]

Compared to hydrogen bonding, stacking interactions have usually received little attention in echinomycin–DNA complexes, but a more detailed analysis is necessary for a full understanding of the experimental data. The binding enthalpy due to stacking interactions can be decomposed into van der Waals (Figure 5) and electrostatic components (Figure 6). Each of these have, in turn, been calculated for the sandwiched base pairs and for the base pairs flanking the bisintercalation site. It is illustrative to view the footprinting results in the light of these calculations, especially when comparing sequences presenting a very similar minor-groove environment, so as to understand, for example, why echinomycin prefers DAP-containing sites over guanine-containing sites.[21] By first examining those complexes with a common central CpG step, it can be seen that the minor differences between them are, as expected, restricted to interactions with the flanking bases (Figures 5 and 6). The electrostatic contribution, however, appears to be more discriminating than the corresponding van der Waals component, as most clearly illustrated by CpCpGpG (apparently a weaker binding site for echinomycin than ApCpGpT) for which this electrostatic term is slightly repulsive. This observation is in consonance with *ab initio*[24] and semiempirical[25] calculations on stacked DNA bases showing that whereas the dispersion term contributes to the overall stability, the electrostatic term is responsible for the sequence and orientation dependence of the stacking interaction, in agreement with early observations on the stacking patterns of nucleic acid constituents.[26] Moreover, the conformational preferences of DNA base-pair steps have been recently rationalized on the basis of the shapes and charge distributions of the stacked bases.[27]

As regards the central bases, the *unfavorable* electrostatic interaction already detected between the drugs' chromophores and the central CpG step, which we have depicted in a simplified fashion in terms of the magnitudes and relative orientations of the dipole moments of both the chromophores and the G:C base pairs,[13] is further confirmed in the present calculations (Figure 6 and Table 1). Since echinomycin is found to bind rather strongly to this central dinucleotide, this negative effect must be outweighed by the very favorable electrostatic and hydrogen-bonding interactions with the minor groove reported above. On the other hand, a DAP:T base pair, while also presenting the 2-amino group in the minor groove, is endowed with a significantly lower dipole moment (Figure 7), and the different charge distribution gives rise to an *attractive* electrostatic stacking interaction

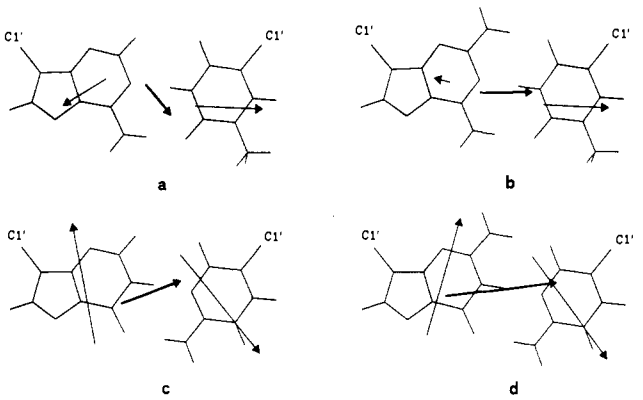


Figure 7. Dipole moments (μ/D) of the base pairs (thick arrows) found in the DNA–echinomycin complexes studied: (a) A:T, μ = 1.74; (b) DAP:T, μ = 2.31; (c) I:C, μ = 2.77; (d) G:C, μ = 5.08. (Dipole moments were calculated using the point charges employed in the force field and including the C1′ atoms of the sugars as buffers in order to achieve electrical neutrality.[13] The dipole moments of individual bases are drawn as thin arrows. The midpoint of each vector is centered on the geometrical center of the system considered.)

with the quinoxaline-2-carboxamide system (Figure 6), which leads to an improved calculated binding energy and to a larger experimental association constant. The results accounting for the enhanced affinity of echinomycin for DpTpDpT over ApCpGpT are reinforced by the interaction energies obtained for the "mixed" sequences also present in the modified DNA employed in the experiments,[21] *i.e.* DpCpDpT and DpTpGpT, for which values half way between those in the ApCpGpT and DpTpDpT complexes are found, as expected (Figure 4). In this modified DNA, the originally protected CpG sites are now adjacent to these DAP-containing *high-affinity* sites, and this fact may explain why binding to the canonical CpG steps is precluded.

The dipole moment of an I:C base pair, although conserving virtually the same direction as the previous two, is also reduced relative to that of G:C (Figure 7), which translates into a better electrostatic contribution to the stacking interaction with the quinoxaline-2-carboxamide (Figure 6). Nevertheless, due to the absence of the exocyclic amino group on the purine ring, the weakened hydrogen-bonding interactions with the minor groove, together with the decreased van der Waals component of the stacking interactions, make a CpI step a poor binding site for echinomycin. When the guanines of only one DNA strand are replaced with inosine, the interaction energies are also intermediate between those of a completely substituted DNA and the reference ApCpGpT complex (Figure 4). In agreement with these results, echinomycin is found to afford no protection against nuclease cleavage on these hybrid molecules either.[20]

The present work highlights the well-established unique role played by the 2-amino group in the minor groove for ligand binding,[1,20,28] but also unravels the distinct stacking properties of G:C, I:C, A:T, and DAP:T base pairs with respect to the quinoxaline-2-carboxamide chromophore of echinomycin. Both the experimental determination and *ab initio* calculation of the dipole moment of *N*-methylquinoxaline-2-carboxamide (μ = 4.1 D) have confirmed the reported dipolar nature of this chromophore[13] (see Methodology). Thus, the differential electrostatic characteristics of the bases (Figure 7) must have a profound influence on the interaction of echinomycin with DNA. In relation to this, it is worth remembering that the ascendancy of the 2-NH_2 group of guanines disappears when the antibiotic engages its CO of alanines in two intramolecular hydrogen bonds with the NH of valines, as is the case in TANDEM[29] and CysMeTANDEM.[30] According to our calculations (unpublished results), the observed preference of these ligands for TpA steps over CpG steps may partly originate in the fact that the electrostatic term of the stacking interactions is more favorable for the former dinucleotide step (Figure 6). In this regard, it is noteworthy that CysMeTANDEM does not bind to CpG sites, and binds more tightly to TpA than to CpI,[31] which is also in accord with the results shown in Figures 5 and 6.

Two relevant issues, which have not been addressed in the present investigation and are likely to have some bearing on the sequence preference of echinomycin, are the energetic cost associated with DNA unstacking on both sides of the bisintercalation site and the conformational state of the DNA prior to binding of the drug. Replacing all the adenine bases in a natural DNA with DAP has the effect of providing the minor groove with an exocyclic amino group in every step. The concomitant substitution of guanine with hypoxanthine would effectively "relocate" this NH_2 group from one type of purine ring to another, making this DNA more akin to natural DNA. On the other hand, it is well-known that thermal stabilities of DNA polynucleotides depend significantly not only on their base composition but also on their nucleotide sequences.[32] Theoretical calculations on stacked base pairs have shown the influence of these two factors on stacking energy heterogeneity,[25] with GpC consistently yielding the strongest stacking interaction, which implies that unstacking a guanine base on the 5′ side of the central CpG step should be more difficult than unstacking any other base in the same position. This reasoning is complementary to the above discussion on stacking interaction energies (Figures 5 and 6): whereas CpCpGpG cannot be regarded as a good binding site on electrostatic grounds, these interactions should make GpCpGpC, destacking energies notwithstanding, a good target for echinomycin binding. In this respect, it is usually agreed that echinomycin shows a preference for A:T pairs flanking the central CpG step[3,4] despite the fact that some NMR data exist on the binding of this antibiotic to $d(GCGC)_2$ and $d(CCGG)_2$,[10] and a crystal structure is available in which a guanine is found on the 5′ side of the CpG step sandwiched by triostin A, albeit in a Hoogsteen-like conformation.[9] Moreover, recent DNase I footprinting titration experiments analyzing the *tyrT* DNA restriction fragment at single-bond resolution have shown that the CpG bond at position 76 (*i.e.* the center of a GpCpGpC stretch) affords the most strongly protected site in a curve representing typical protection behavior.[5] On the contrary, this same tetranucleotide sequence centered around position 95 gives rise to a plot characteristic of weak binding sites, and this particular CpG bond appears as the least protected of all. These findings highlight the difficulties inherent to the characterization of "the best binding site" and the influence of context in this sort of experiments. Our calculations for sequences containing a CpG central step show no great differences in the overall interaction energies (Figure 2), which is in general agreement with the estimation that the range of binding strength of the various central CpG steps is probably no greater than 1 order of magnitude.[5] Consequently, any CpG could serve

in principle as a good binding site for echinomycin in natural DNA although we show that the absolute magnitude can be modulated by the stacking interactions with the flanking bases, which on the other hand have also been suggested to contribute to the sequence-dependent Hoogsteen pairing of these bases.[10,11,13]

The important role played by the stacking interactions, which is partially masked by the more prominent hydrogen-bonding interactions, manifests itself most dramatically at TpD and CpG central steps in a modified DNA containing DAP:T pairs in place of standard A:T base pairs. These two sequences present a similar minor-groove environment, and the differences in their binding affinities for echinomycin are a reflection of the distinct stacking properties of DAP:T and G:C base pairs, which favor binding of this antibiotic to a TpD step over the canonical CpG step, as shown recently by footprinting experiments.[21] In line with this observation about the contribution of the electronic properties of chromophores and base pairs to the binding selectivity of echinomycin is an early report that demonstrates a correlation between specificity of binding to G:C base pairs and polarizability of the chromophore in a series of monointercalating heteroaromatic ligands lacking special hydrogen-bonding functions.[33]

Conclusions

The overall agreement between the calculated interaction energies and the footprinting results is rewarding. The prevailing role of the 2-amino group in the minor groove for hydrogen-bonding interactions has been confirmed, but other energy contributions have emerged as additional selectivity factors. The preference of echinomycin for both central and flanking sequences has been shown to be the result of a balance of forces in which interactions emanating from the stacked systems play an important part. This is most clearly seen when sequences presenting a very similar minor-groove environment with full hydrogen-bonding capabilities but different stacking characteristics are compared (*i.e.* CpG and TpD) and the large dipole moment of *N*-methylquinoxaline-2-carboxamide is considered. Decomposition of the stacking interactions into van der Waals and electrostatic components highlights their relative importance in accounting for the observed footprinting patterns. G:C, I:C, A:T, and DAP:T base pairs are shown to be endowed with distinct stacking properties with respect to the quinoxaline chromophore regarding both van der Waals surface area and electron distribution. In the present calculations, of all the sequences explored, DpTpDpT appears as the most favorable binding site for echinomycin. On this basis, a prediction that we can make with a substantial degree of confidence is that binding of echinomycin to poly[d(DAP·T)]·poly[d(DAP·T)] should be stronger than to poly[d(G·C)]·poly[d(G·C)].

An accurate definition of the *optimal binding site* for echinomycin by footprinting experiments alone has been hampered by the fact that most of the DNA fragments used in these assays contain a limited, and often overlapping, repertoire of binding sites. From our results, we would support the notion[1] that a step comprising two adjacent base pairs falls somewhat short, and a longer sequence of at least *four base pairs* is more appropriate for this class of compounds. Nevertheless, the influence of local DNA conformation on ligand binding cannot be neglected and this depends on the surrounding sequences.

It is obvious that the interaction model presented here must be validated by accurate physical measurements. Complete thermodynamic profiles for echinomycin complexation with the different oligonucleotides studied here should be extremely valuable for characterizing the enthalpic and entropic contributions to the binding free energy in each complex but these studies have been hampered so far by the low solubility of the antibiotic in aqueous solution. From our data it would seem that, at least for sequences with very similar minor-groove environments, the entropy change must be of the same order of magnitude since the calculations satisfactorily reproduce the observed binding preferences. In any case, these results help to improve our understanding of the stacking interactions involved in the recognition process and are likely to provide new ground for experimental verification. These ideas may also be useful for the structure-based design of novel DNA-binding drugs.

Methodology

(1) Model Building. The solution structure of a complex between CysMeTANDEM and $d(GATATC)_2$[30] retrieved from the Brookhaven Data Bank[34] was used as a template in the construction of the complexes studied.[35] This is the only experimentally determined structure formed between a quinoxaline antibiotic and an oligonucleotide in which all the base pairs are in the Watson–Crick conformation. The rationale behind this modeling approach is that there is no evidence of Hoogsteen pairs taking place at internal positions surrounding CpG binding sites in longer DNA molecules.[5] Besides, little or no difference has been reported to exist between the conformations of a TpA step and a CpG step when both are bound by either CysMeTANDEM or triostin A, respectively, as judged by a root mean square difference of just 0.55 Å for all common atoms between these two DNA steps.[30] Echinomycin and triostin A can be compared with confidence as the complexes of these two drugs with $d(CGTACG)_2$ are virtually identical.[8]

Of the eight solution structures for the CysMeTANDEM-$d(GATATC)_2$ complex, that yielding the lowest potential energy and the best intermolecular hydrogen-bonding scheme was selected. Echinomycin was modeled as recently reported[13] and its two chromophores were then superimposed on the chromophores of CysMeTANDEM in the above experimental complex in order to replace one drug with the other. For the DNA hexamers, the appropriate modifications of base composition were introduced by using standard geometries[36] and replacing the respective purine and pyrimidine bases where necessary. 10 counterions resembling hexahydrated sodium ions were then placed in the bisector of each O–P–O group in order to achieve electroneutrality.[22,37]

(2) Molecular Mechanics Force Field. The AMBER all-atom force field parameters[36] were used for the standard DNA bases and the standard amino acid residues of echinomycin. Additional parameters describing bonded interactions for echinomycin have already been reported,[13] and those needed for DAP and inosine were derived by analogy with those already present in the AMBER database.[38] Given the importance attached to the electrostatic term in both hydrogen-bonding and stacking interactions in these systems,[13] it was essential to assign point charges to the nonstandard DNA bases and to echinomycin that could reliably represent the electron-

density distribution in these systems. Therefore the molecular electrostatic potentials for the N^9-methylated derivatives of hypoxanthine and DAP, and for suitable fragments[39] of the echinomycin molecule were calculated from the corresponding *ab initio* wave functions using a 6-31G* basis set. This level of quality has been demonstrated to provide a very accurate representation of the real electrostatic charge distribution.[40] Point charges were then derived (supplementary material) by fitting the rigorous quantum mechanical molecular electrostatic potential to a monopole–monopole expression, as reported elsewhere.[39] Strong support for the validity of this procedure is provided by the excellent agreement found between the calculated electrostatic dipole moment (μ = 4.14 D) for *N*-methylquinoxaline-2-carboxamide and the experimental value (μ = 4.15 ± 0.03 D) (see below).

(3) Energy Minimization. The initial models were refined by progressively minimizing their potential energy using the MINMD module of AMBER 4.0:[36] firstly the hydrogen atoms only, then both the peptidic part of the antibiotic and the sugar–phosphate backbone of the DNA hexamers, and finally the whole systems. Before each minimization stage, a short optimization run constraining the atoms to their initial coordinates allowed readjustment of covalent bonds and van der Waals contacts without changning the overall conformation of the complexes. All atom pairs were included in the calculation of the nonbonded interactions. The optimizations were carried out in a continuum medium of relative permitivitty $\epsilon = 4r_{ij}$ for simulating the solvent environment.[22,37,41] For the first 1000 steps of the minimization procedure, all possible hydrogen bonds between the alanine residues and suitable donor and acceptor atoms in the floor of the minor groove were reinforced with additional distance and angle constraining functions with force constants of 50 kcal mol^{-1} Å^{-2} and 50 kcal mol^{-1} rad^{-2}, respectively. All in all, the optimizations covered a total of 4000 steps of steepest descent energy minimization after which the final root mean square difference of the derivatives of the potential energy with respect to the atomic coordinates was within 0.15 ± 0.03 kcal mol^{-1} Å^{-1} in all the complexes.

Experimental Section

All chemicals were reagent grade and were used as received. The melting point was determined on an Electrothermal IA6304 in an open capillary tube and is uncorrected. The IR spectrum was recorded on a Perkin-Elmer 883 spectrophotometer using a KBr pellet. The ^{1}H NMR spectrum was recorded on a Varian UNITY spectrometer (300 MHz). Parameters were deduced from the analysis with the LAOCOON III program. The mass spectrum was determined on a Hewlett-Packard 5988A (70 eV) spectrometer by electronic impact. The elemental analysis was performed on a Perkin-Elmer elemental analyzer Model 240E. Flash column chromatography was performed with the indicated solvents on Merck 60 F_{254} silica gel.

(1) Synthesis of *N*-Methylquinoxaline-2-carboxamide. Methylamine hydrochloride was treated with 2-quinoxalinecarbonyl chloride and potassium carbonate suspended in dry dichloromethane. The mixture was refluxed for 6 h and then washed with water and dried with sodium sulfate. The solvent was evaporated under reduced pressure, and the residual oil was purified on a silica gel column, using ethyl acetate–hexane (1:1) as eluent. The residue obtained was crystallized from dichloromethane/hexane. The product appeared as white needles: mp 153–155 °C, IR 3338, 1683, 1530, 1490, 1401, 1127, 968, 800, 777 cm^{-1}; ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.68 (s, 1H, H2), 8.20 (m, 1H, H6, J_{H6-H5} = 7.54 Hz, J_{H6-H4} = 1.29 Hz, J_{H6-H3} = 0.18 Hz), 8.09 (m, 1H, H3, J_{H3-H4} = 5.89 Hz, J_{H3-H5} = 2.56 Hz, J_{H3-H6} = 0.18 Hz), 7.89 (bs, 1H, NH), 7.86 (m, 1H, H5, J_{H5-H4} = 6.80 Hz, J_{H5-H6} = 7.54 Hz, J_{H5-H3} = 2.56 Hz), 7.84 (m, 1H, H4 J_{H4-H3} = 5.89 Hz, J_{H4-H5} = 6.80 Hz, J_{H4-H6} = 1.29 Hz), 3.13 (d, 3H, CH_3) J = 5.10 Hz); MS *m/e* (rel. intensity) 187 (M^+, 49), 158 (21), 130 (100), 103 (3), 76 (15). Anal. Calcd for $C_{10}H_9N_3O$: C, 64.16; H, 4.85; N, 22.45. Found: C, 63.93; H, 4.98; N, 22.15.

(2) Dipole Moment Determination. Dielectric measurements of solutions of *N*-methylquinoxaline-2-carboxamide in benzene were performed at 25 °C with a WTW dipolmeter DM 01 at a fixed frequency of 2.0 MHz. The DFL1 cell was calibrated at 25 °C using toluene, carbon tetrachloride, and cyclohexane, all of known dielectric constant. Increments of the refractive indices of the solutions with respect to the solvent were determined at 25 °C in a Brice-Phoenix 2000-V differential refractometer. Values of the dipole moment were calculated from the equation of Guggenheim and Smith:[42]

$$\mu^2 = (27kTM)(4\pi\rho N)^{-1}(\epsilon_1 + 2)^{-2}((d\epsilon/d\omega) - 2n_1(dn/d\omega))$$

where k is the Boltzmann constant, T is the absolute temperature, M is the molecular weight of the solute, ρ is the density of the solvent, N is Avogadro's number, and ω is the weight fraction of the solute. ϵ and n represent the dielectric constant and index of refraction of the solutions, respectively; ϵ_1 and n_1 represent the same quantities for the solvent. Values of $d\epsilon/d\omega$ and $dn/d\omega$ were obtained as the slope from plots of the increments of the dielectric constant ($\Delta\epsilon = \epsilon - \epsilon_1$) and the index of refraction ($\Delta n = n - n_1$) against ω, in the vicinity of $\omega \rightarrow 0$.

Acknowledgment. J.G. is recipient of a predoctoral grant from the Spanish Ministerio de Educación y Ciencia (FPI, PN89-25126387). Biosym Technologies, Inc. (San Diego, California) contributed a software license. This research has been financed in part by the Spanish Comisión Interministerial de Ciencia y Tecnología (FAR91-0277) and the University of Alcalá de Henares (Madrid, Spain).

Supplementary Material Available: Two tables containing partial atomic charges for the echinomycin residues and the nonstandard DNA bases (2 pages). Ordering information is given on any current masthead page. Cartesian coordinates for the 13 energy-minimized complexes in PDB format are available from the authors on request (e-mail: ffgago@alcala.es).

References

(1) Waring, M. J. The Molecular Basis of Specific Recognition Between Echinomycin and DNA. In *Molecular Basis of Specificity in Nucleic Acid-Drug Interactions*. Pullman, B.; Jortner, J., Eds.; Kluwer Academic Publishers: Dordrecht, 1990, pp 225–245.
(2) Leroy, J. L.; Gao, X.; Misra, V.; Guéron, M.; Patel, D. J. Proton exchange in DNA-luzopeptin and DNA-echinomycin bisintercalation complexes: rates and processes of base-pair opening. *Biochemistry* **1992**, *31*, 1407–1415.
(3) Low, C. M. L.; Drew, H. R.; Waring, M. J. Sequence-specific Binding of Echinomycin to DNA: Evidence for Conformational Changes Affecting Flanking Sequences. *Nucleic Acids Res.* **1984**, *12*, 4865–4879.
(4) Van Dyke, M. M.; Dervan, P. B. Echinomycin Binding sites in DNA. *Science* **1984**, *225*, 1122–1127.
(5) Sayers, E. W.; Waring, M. J. Footprinting Titration Studies on the Binding of Echinomycin to DNA Incapable of Forming Hoogsteen Base Pairs. *Biochemistry* **1993**, *32*, 9094–9107.
(6) Hoogsteen, K. The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **1959**, *2*, 822–823.
(7) Wang, A. H.-J.; Ughetto, G.; Quigley, G. J.; Hakoshima, T.; van der Marel, G. A.; van Boom, J. H.; Rich, A. The Molecular Structure of a DNA-triostin A complex. *Science* **1984**, *225*, 1115–1121.
(8) Ughetto, G.; Wang, A. H.-J.; Quigley, G. J.; van der Marel, G. A.; van Boom, J. H.; Rich, A. A Comparison of the Structure of Echinomycin and Triostin A Complexed to a DNA Fragment. *Nucleic Acids Res.* **1985**, *13*, 2305–2323.
(9) Wang, A. H.-J.; Ughetto, G.; Quigley, G. J.; Rich, A. Interactions of Quinoxaline Antibiotic and DNA: The Molecular Structure of a Triostin A-d(GCGTACGC) Complex. *J. Biomol. Struct. Dyn.* **1986**, *4*, 319–342.
(10) Gao, X.; Patel, D. J. Antitumor Drug-DNA Interactions: NMR Studies of Echinomycin and Chormomycin Complexes. *Q. Rev. Biophys.* **1989**, *22*, 93–138.

(11) Gilbert, D. E.; Feigon, J. The DNA Sequence at Echinomycin Binding Sites Determines the Structural Changes Induced by Drug Binding: NMR Studies of Echinomycin Binding to [d(ACGTACGT)]$_2$ and [d(TCGATCGA)]$_2$. *Biochemistry* **1991**, *30*, 2483–2494.
(12) Gilbert, D. E.; Feigon, J. Proton NMR Studies of the [d(ACGTATACGT)]$_2$–2echinomycin Complex: Conformational Changes between Echinomycin Binding Sites. *Nucleic Acids Res.* **1992**, *20*, 2411–2420.
(13) Gallego, J.; Oritz, A. R.; Gago, F. A molecular dynamics study of the bis-intercalation complexes of echinomycin with d(ACGT)$_2$ and d(TCGA)$_2$: Rationale for sequence-specific Hoogsteen base-pairing. *J. Med. Chem.* **1993**, *36*, 1548–1561.
(14) Lee, J. S.; Waring, M. J. Interaction between Synthetic Analogues of Quinoxaline Antibiotics and Nucleic Acids. *Biochem. J.* **1978**, *173*, 129–144.
(15) Fox, K. R.; Gauvreau, D.; Goodwin, D. C.; Waring, M. J. Binding of Quinoline Analogues of Echinomycin to Deoxyribonucleic Acid. Role of the Chromophores. *Biochem. J.* **1980**, *191*, 729–742.
(16) Low, C. M. L.; Fox, K. R.; Waring, M. J. DNA Sequence-selectivity of Three Biosynthetic Analogues of the Quinoxaline Antibiotics. *Anti-Cancer Drug Des.* **1986**, *1*, 149–160.
(17) Low, L. C. M.; Olsen, R. K.; Waring, M. J. Sequence Preferences in the Binding to DNA of Triostin A and TANDEM as Reported by DNase I Footprinting. *FEBS Lett.* **1984**, *176*, 414–420.
(18) Addess, K. J.; Gilbert, D. E.; Olsen, R. K.; Feigon, J. Proton NMR studies of [N-MeCys^3,N-MeCys^7]TANDEM binding to DNA oligonucleotides: Sequence-specific binding at the TpA site. *Biochemistry* **1992**, *31*, 339–350.
(19) Waterloh, K.; Olsen, R. K.; Fox, K. R. Bifunctional intercalator [N-MeCys^3,N-MeCys^7]TANDEM binds to the dinucleotide TpA. *Biochemistry* **1992**, *31*, 6246–6253.
(20) Marchand, C.; Bailly, C.; McLean, M.; Moroney, S. E.; Waring, M. J. The 2-amino group of guanine is absolutely required for specific binding of the anti-cancer antibiotic echinomycin to DNA. *Nucleic Acids Res.* **1992**, *20*, 5601–5606.
(21) Bailly, C.; Marchand, C.; Waring, M. J. New binding sites for antitumor antibiotics created by relocating the purine 2-amino group in DNA. *J. Am. Chem. Soc.* **1993**, *115*, 3784–3785.
(22) Singh, U. C.; Pattabiraman, N.; Langridge, R.; Kollman, P. A. Molecular mechanical studies of d(CGTACG)$_2$:complex of triostin A with the middle AT base pairs in either a Hoogsteen or Watson-Crick pairing. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 6402–6406.
(23) A preliminary account of this work was presented at a workshop on *DNA-Drug Interactions* at the Center for International Meetings on Biology, Fundación Juan March: Madrid, November 1993.
(24) Aida, M.; Nagata, C. An *ab initio* molecular orbital study on the stacking interaction between nucleic acid bases: Dependence on the sequence and relation to the conformation. *Int. J. Quantum Chem.* **1986**, *29*, 1253–1261.
(25) Ornstein, R. L.; Rein, R.; Breen, D. L.; MacElroy, R. D. An optimized potential function for the calculation of nucleic acid interaction energies. I. Base stacking. *Biopolymers* **1978**, *17*, 2341–2360.
(26) Bugg, C. E.; Thomas, J. M.; Sundaralingam, M.; Rao, S. T. Stereochemistry of nucleic acids and their constituents. X. Solid-state base-stacking patterns in nucleic acid constituents and polynucleotides. *Biopolymers* **1971**, *10*, 175–219.
(27) Hunter, C. A. Sequence-dependence DNA structure. The role of base stacking interactions. *J. Mol. Biol.* **1993**, *230*, 1025–1054.
(28) Gago, F.; Richards, W. G. Netropsin binding to poly[d(IC)]·poly[d(IC)] and poly[d(GC)]·poly[d(GC)]: A computer simulation. *Mol. Pharmacol.* **1990**, *37*, 341–346.
(29) Viswamitra, M. A.; Kennard, O.; Cruse, W. B. T.; Egert, E.; Sheldrick, G. M.; Jones, P. G.; Waring, M. J.; Wakelin, L. P. G.; Olsen, R. K. Structure of TANDEM and its implication for bifunctional intercalation into DNA. *Nature* **1981**, *289*, 817–819.
(30) Addess, K. J.; Sinsheimer, J. S.; Feigon, J. Solution structure of a complex between [N-MeCys^3,N-MeCys^7]TANDEM and [d(GATATC)]$_2$. *Biochemistry* **1993**, *32*, 2498–2508.
(31) Addess, K. Personal communication.
(32) Gotoh, O.; Tagashira, Y. Stabilities of nearest-neighbor doublets in double-helical DNA determined by fitting calculated melting profiles to observed profiles. *Biopolymers* **1981**, *20*, 1033–1042.
(33) Müller, W.; Crothers, D. M. Interactions of heteroaromatic compounds with nucleic acids. 1. Influence of heteroatoms and polarizability on the base specificty of intercalating ligands. *Eur. J. Biochem.* **1975**, *54*, 267–277.
(34) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank: A Computer-based Archival File for Macromolecular Structures. *J. Mol. Biol.* **1977**, *112*, 535–542.
(35) INSIGHT-II, version 2.1.0; Biosym Technologies. 9685 Scranton Road, San Diego, CA 92121-2777.
(36) Pearlman, D. A.; Case, D. A.; Caldwell, J.; Seibel, G.; Singh, U. C.; Weiner, P.; Kollman, P. A. *AMBER* version 4.0 **1991**. Department of Pharmaceutical Chemistry, University of California, San Francisco.
(37) Gago, F.; Reynolds, C. A.; Richards, W. G. The Binding of Nonintercalative Drugs to Alternating DNA Sequences. *Mol. Pharmacol.* **1989**, *35*, 232–241.
(38) Weiner, S. J.; Kollman,P. A.; Nguyen, D. T.; Case, D. A. An All Atom Force Field for Simulations of Proteins and Nucleic Acids. *J. Comput. Chem.* **1986**, *7*, 230–252.
(39) Orozco, M.; Luque, F. J. A practical procedure for the determination of electrostatic charges for molecules. *J. Comput.-Aided Mol. Design* **1990**, *4*, 411–426.
(40) Luque, F. J.; Orozco, M.; Illas, F.; Rubio, J. Effect of electron correlation on the electrostatic potential distribution of molecules. *J. Am. Chem. Soc.* **1991**, *113*, 5203–5211.
(41) Orozco, M.; Laughton, C. A.; Herzyk, P.; Neidle, S. Molecular-mechanics modelling of drug-DNA structures: the effects of differing dielectric treatment on helix parameters and comparison with a fully solvated structural model. *J. Biomol. Struct. Dyn.* **1990**, *8*, 359–373.
(42) Guggenheim, E. A. The computation of electric dipole moments. *Trans. Faraday Soc.* **1951**, *47*, 573–576.
(43) Kraulis, P. J. Molscript: A Program to Produce both Detailed and Schematic Plots of Protein Structures. *J. Appl. Crystallogr.* **1991**, *24*, 946–950.