Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102
Volume 12, Issue Number 1 (1994), ©Adenine Press (1994).

Binding of Echinomycin to $d(GCGC)_2$ and $d(CCGG)_2$: Distinct Stacking Interactions Dictate the Sequence-Dependent Formation of Hoogsteen Base Pairs

José Gallego[1], F.J. Luque[2], Modesto Orozco[3], and Federico Gago[1]*

[1] Departamento de Fisiología y Farmacologia
Universidad de Alcalá de Henares, 28871 Madrid, Spain

[2] Departament de Farmàcia (Unitat Físico-Química)
Facultat de Farmácia, Universitat de Barcelona
08028 Barcelona, Spain

[3] Departament de Bioquímica i Fisiologia
Facultat de Química
Universitat de Barcelona
08028 Barcelona, Spain

Abstract

Molecular dynamics simulations have been used to explore the behavior of the complexes of echinomycin with the DNA tetramers $d(GCGC)_2$ and $d(CCGG)_2$ in which the terminal bases have been paired according to either a Hoogsteen or a Watson-Crick hydrogen bonding scheme. The energy of the four resulting complexes has been monitored along the dynamics trajectories and the interaction energy between echinomycin and DNA has been decomposed into contributions arising from the planar aromatic systems and the depsipeptide part of the antibiotic. Our calculations predict a large increase in overall stabilization upon protonation of the terminal cytosines and subsequent Hoogsteen pair formation in the complex of echinomycin with $d(GCGC)_2$ but not with $d(CCGG)_2$, in agreement with the experimental evidence [Gao and Patel, *Quart. Rev. Biophys. 22*, 93-138 (1989)]. The conformational preferences appear to arise mainly from differential stacking interactions in which the electrostatic component is shown to play a dominant role. Differences in hydrogen bonding patterns are also found among the complexes and these are compared in relation to available crystal structures. The binding of echinomycin to DNA appears as a complex process involving many interrelated variables.

Introduction

There has been some controversy in the past few years regarding the occurrence and significance of Hoogsteen base pairs in DNA restriction fragments and oligonucleotides (1,2), especially since it was unexpectedly found by x-ray crystallography that they could be induced upon binding of quinoxaline bis-intercalating antibiotics to certain

* Author to whom correspondence should be addressed.

DNA sequences (reviewed in 3). In the Hoogsteen conformation the purine ring adopts a *syn* orientation about the glycosidic bond giving rise to an alternative hydrogen bonding scheme with the pyrimidine on the opposite strand, as first reported for the 1:1 complex of 1-methylthymine with 9-methyladenine (4). The first Hoogsteen base pairs directly observed in an oligonucleotide structure were the central A:T base pairs separating the two CpG binding sites sandwiched by each drug molecule in the 2:1 crystal complex of triostin A with $d(CGTACG)_2$ (5). The complex of the same DNA duplex with echinomycin, another member of the quinoxaline family with similar binding preferences, highlighted the same unusual features (6). Shortly thereafter, the crystal structure of $d(GCGTACGC)_2$ with two triostin A molecules bound (7,8) revealed that Hoogsteen pairing was not restricted to A:T base pairs since the terminal G:C base pairs flanking the quinoxaline rings also displayed this alternative scheme. The fact that the crystals of this latter complex formed at acidic pH underscores the need for protonation at the N3 of the cytosine base for the Hoogsteen hydrogen bonding mode to be observed in G:C pairs (Figure 1).

Studies directed towards determining the existence of such drug-induced structural changes in solution included NMR experiments and treatment of drug-DNA complexes with several reagents. A number of NMR studies undertaken with echinomycin bound to different DNA oligomers showed that the rearrangement leading to Hoogsteen base-pairing is sequence-dependent since it only appears when the purine base is located on the 5′ side of the CpG binding steps, *v.g.* in $d(ACGT)_2$ (9), $d(GCGC)_2$ (10), and $d(ACGTACGT)_2$ (11), but not in $d(TCGA)_2$, $d(CCGG)_2$, or $d(TCGATCGA)_2$. In addition, the analysis of the $d(ACGTACGT)_2$:$(echinomycin)_2$ complex in solution revealed that whereas at 1 °C both *terminal* and *internal* A:T base pairs adopt a stable Hoogsteen scheme, at 45 °C this is maintained only for the former, with the latter alternating between Hoogsteen and either an open or a Watson-Crick paired state (11). Furthermore, if the CpG binding sites are separated by more than two base pairs, as in $d(ACGTATACGT)_2$:$(echinomycin)_2$, Hoogsteen base-pairing is not clearly detected in solution for the *internal* base pairs (12). Thus, the NMR results in solution have confirmed the original findings in the solid state but have additionally shown that this conformational change depends not only on the sequence but also on the temperature and on whether or not the base pairs are subjected to helical constraints within the DNA molecule. More recently, stable Hoogsteen base pairing with minimal base-base stacking distortions has been reported for the ApT step in the major product of the cross-linking reaction between $d(CGTAATTACG)_2$ and the non-intercalating cyclopropylpyrroloindole antitumor agent bizelesin (13).

For longer DNA stretches, hyperreactivity to diethylpyrocarbonate by purines both proximal and distal to echinomycin binding sites was initially suggested to be consistent with the formation of Hoogsteen pairs at these positions (14) although doubts were soon cast on this hypothesis (15). The similar reaction patterns toward the thymidine-specific reagent osmium tetroxide of a DNA fragment and the same fragment containing 7-deaza-2′-deoxy-adenosine in place of 2′-deoxy-adenosine in one of the strands further supported the notion that Hoogsteen base pair formation

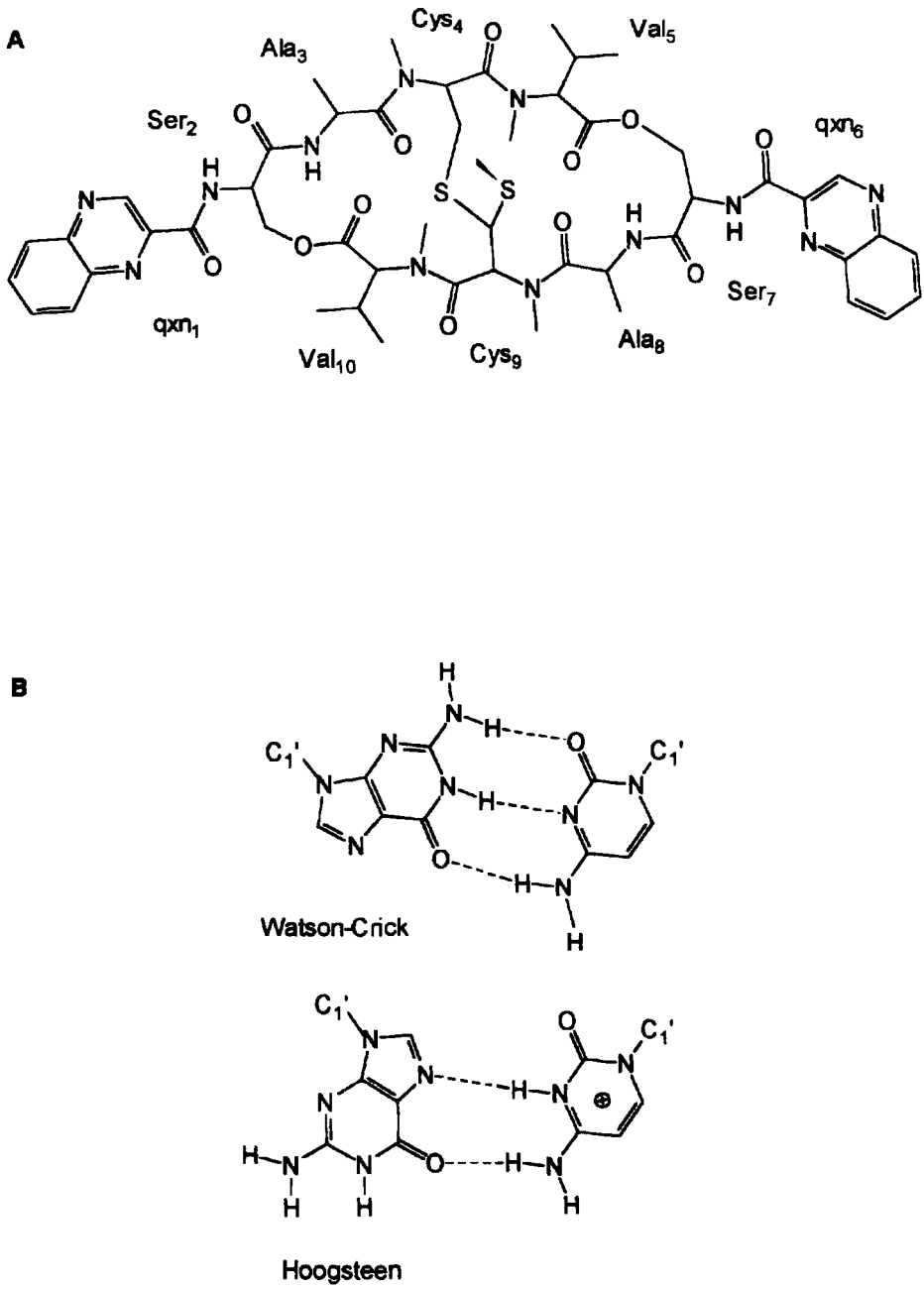

Figure 1: A. Chemical structure of echinomycin. **B.** Watson-Crick and Hoogsteen hydrogen-bonding schemes in a G:C base pair. Note that the Hoogsteen arrangement implies the loss of one hydrogen bond and the protonation of the cytosine base at the N3 atom.

was not a prerequisite for echinomycin binding since the replacement of N with C at position 7 of the purine ring, while preventing the formation of a hydrogen bond with N3-H moiety of thymine, did not preclude echinomycin binding (16). More

recently, Sayers and Waring (2) reported that incorporation of 7-deaza-purine analogues into *tyrT* DNA in place of the normal nucleotides did not result in decreased binding constants or in different footprinting patterns with respect to natural *tyrT* DNA. These findings were interpreted as rather conclusive proof that the *anti-syn* conformational transition of purine nucleosides is not essential for the binding of bis-intercalating quinoxaline antibiotics to DNA, in agreement with the NMR data for some sequences.

The question remains, however, of why Hoogsteen pairs are unambiguously detected in solution for some sequences and not for others. This alternating scheme was initially thought to be advantageous in the complexes of DNA with quinoxaline antibiotics because in this conformation the sugar-phosphate backbones from the two oligonucleotide strands are brought almost 2 Å closer together than they would be with Watson-Crick base pairs, thus allowing for improved van der Waals contacts between the two molecules (5,7,8). In an early molecular mechanics study of the $d(CGTACG)_2$:(triostin $A)_2$ complex, Singh *et al.* (17) agreed with this interpretation and concluded that the van der Waals' interaction between the valine residues and the DNA molecule was one of the main factors stabilizing the Hoogsteen conformation in the complex. Other factors which appeared to be involved were better base stacking of the A:T base pairs in Hoogsteen relative to Watson-Crick, more favorable drug-DNA stacking energies, and counterion effects. More recently, a theoretical study from our group pointed out that echinomycin-induced Hoogsteen formation for the flanking A:T base pairs in $d(ACGT)_2$, but not in $d(TCGA)_2$, could be stabilized by better electrostatic stacking interactions between the *N*-methyl-quinoxaline-2-carboxamide aromatic system of the drug and the A:T base pairs in this conformation. This effect was suggested to arise from the distinct electrostatic characteristics of an A:T base pair in the Hoogsteen conformation relative to its Watson-Crick counterpart (18).

An extension of this work to terminal G:C base pairs is presented here (19) by examining the complexes of echinomycin with the oligonucleotides $d(GCGC)_2$ and $d(CCGG)_2$, both containing the canonical CpG binding site for this antibiotic (3). NMR studies have revealed that the terminal G:C pairs in the $d(GCGC)_2$ complex, which are Watson-Crick paired at neutral pH, shift to Hoogsteen pairing as the pH is lowered, with a pK_a of 5.1 for the transition (10). Interestingly, no Hoogsteen base pairs were observed in the $d(CCGG)_2$ complex at any pH. In an attempt to rationalize these observations we have modelled and studied by means of molecular dynamics simulations in aqueous solution the complexes of echinomycin with $d(GCGC)_2$ and $d(CCGG)_2$ in which both terminal G:C base pairs adopt either a Hoogsteen (named G(h) and C(h), respectively) or a Watson-Crick conformation (G(w) and C(w)) (Figure 2). Knowledge about the factors favoring one type of arrangement over the other should improve our understanding of the stacking interactions involved in the recognition process and might be of value in the rational design of novel intercalating and bis-intercalating agents.

Methodology

The AMBER 4.0 suite of programs (20) was used throughout, implemented on a

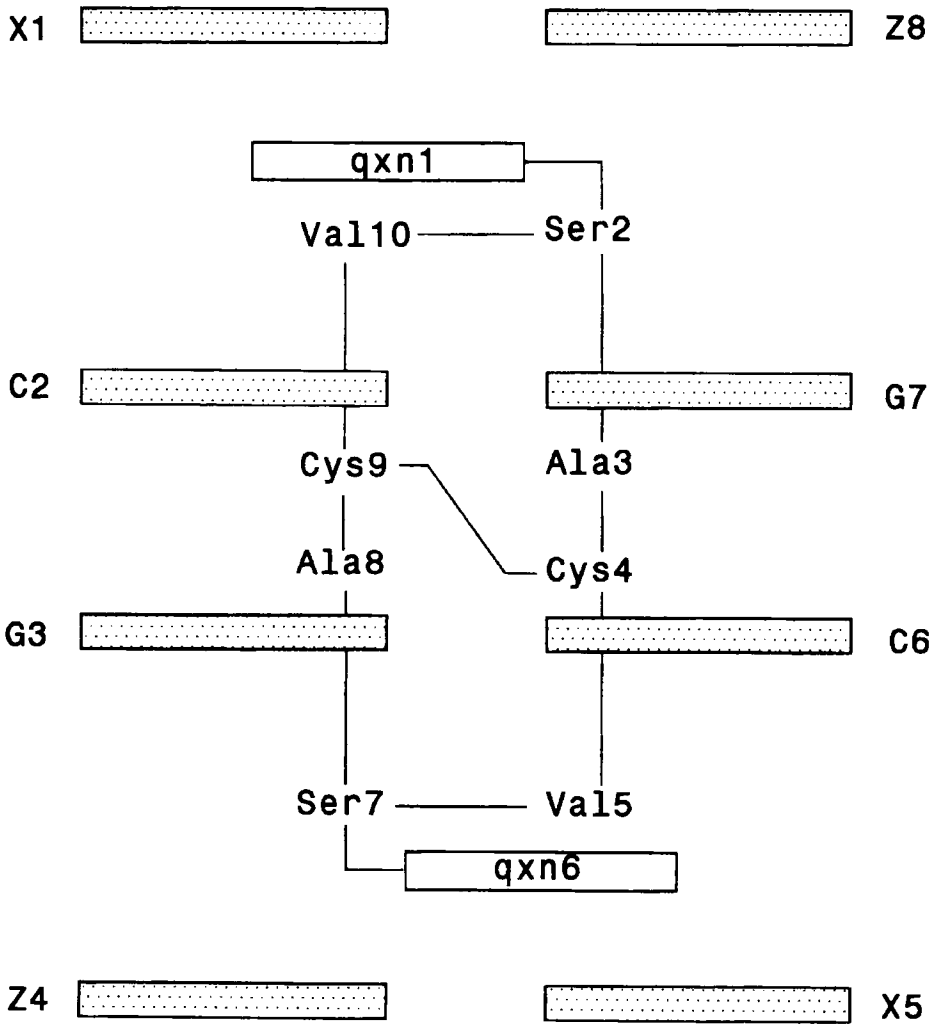

Figure 2: Schematic view from the major groove of a complex between echinomycin and a DNA tetramer containing the canonical CpG binding site. In the present work X:Z stands for either G:C or C:G base pairs. A number of hydrogen bonds established between the NH and carbonyl groups of the antibiotic's alanines and the N3 and 2-amino groups of guanines are crucial for the binding specificity.

cluster of Silicon Graphics Indigo and Control Data Cyber 910 workstations. The interactive molecular graphics software Insight II (21) was used to visualize and manipulate the structures.

Force Field Parameters

The AMBER all-atom force field parameters (22) were used for the DNA tetramers and the standard aminoacid residues of echinomycin. Additional parameters describing bonded interactions for echinomycin have already been reported (18) and those necessary for the deoxycytidinium residue were derived by analogy with those

present in the AMBER database (20). The water molecules were modeled by the rigid three-point charge TIP3P model (23), and the van der Waals parameters for the sodium counterions were taken from Aqvist (24).

The electrostatic term is critical for an adequate description of intermolecular interactions. Point charges for echinomycin and N^3-protonated N^1-methyl-cytosine were calculated *ab initio* at the 6-31G* level (25) using a modified version (26) of Momany's strategy (27). This level of quality has been demonstrated to provide a very accurate representation of the real charge distribution (28), and has been successfully used in molecular dynamics and free energy perturbation simulations (29). The charges were determined by fitting self-consistent field (SCF) and coulombic molecular electrostatic potentials (MEP) in 2 Connolly's layers (30) located at 1.4 and 1.8 times the van der Waals radii of the molecules. A density of 5 points $Å^{-2}$ was used to guarantee the statistical quality of the results (31). For echinomycin, charges were calculated for 4 suitable fragments, which together made up the whole molecule following the fractional model explained elsewhere (31). No dipole restrictions were included during any step of the charge parametrization. It is then remarkable that the largest difference found between the SCF and electrostatic dipole moments was 0.06 D. The validity of this procedure is also supported by the excellent agreement found between the calculated electrostatic dipole moment (μ = 4.14 D) for the *N*-methyl-quinoxaline-2-carboxamide fragment of echinomycin and the experimental value ($\mu = 4.15 \pm 0.03$ D) (32).

Model Building and Energy Minimization of the Complexes

The initial coordinates used for modelling G(h) and C(h) were those published for the $d(GCGTACGC)_2$ octamer complexed with two molecules of triostin A (8). This crystal is the only source of structural information about the stacking geometries of a quinoxaline antibiotic-DNA complex with terminal G:C pairs. For G(w) and C(w), the solution structure of a complex between CysMeTANDEM and $d(GATATC)_2$ (33), retrieved from the Brookhaven Data Bank (34), provided the template for the construction of these two complexes, since this is the most reliable complex of a quinoxaline antibiotic bound to a DNA molecule in which all the bases are Watson-Crick paired. Triostin A and CysMeTANDEM were replaced with echinomycin as previously described (18) based on the following rationale: (i) the complexes of echinomycin and triostin A with $d(CGTACG)_2$ are virtually identical (6), as are the complexes of triostin A with either $d(CGTACG)_2$ or $d(GCGTACGC)_2$; and (ii) only minor differences have 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 (33).

The starting models were initially refined in a continuum medium of relative permitivitty $\varepsilon = 4r_{ij}$ with an infinite cutoff for the nonbonded interactions. First the hydrogen atoms were optimized until the root-mean-square difference of the atomic gradients was less than 0.1 kcal mol^{-1} $Å^{-1}$. This was followed by 100 steps of steepest descent energy minimization during which the atoms were constrained to their initial coordinates in order to allow covalent bonds and van der Waals contacts to readjust without changing the overall conformation of the complexes. The systems were

then further relaxed using 1000 steps of steepest descent energy minimization and an upper-bound harmonic potential with a force constant of 10 Kcal mol^{-1} $Å^{-2}$ to restrain the alanine-guanine intermolecular hydrogen bonds and those between the terminal base pairs to idealized distances and angles.

The refined complexes were then placed in the center of rectangular boxes with dimensions such that the minimum distance between any atom in the complexes and the wall of the boxes was 6 Å. Water molecules were inserted in the box by immersing it into a Monte Carlo-equilibrated configuration of TIP3P water molecules and by subsequently removing all water molecules that were outside the box or whose oxygen or hydrogen atoms lay within 2 or 1 Å, respectively, of any DNA or echinomycin atom. This procedure yielded systems with 793, 786, 836 and 813 water molecules for the G(w), G(h), C(w) and C(h) models, on which periodic boundary conditions were applied. 1000 steps of steepest-descent energy minimization for the water molecules alone allowed optimization of the H_2O-H_2O and solute-H_2O contacts. During this and subsequent calculations a dielectric constant of 1 and a residue-based cutoff distance for nonbonded interactions of 8 Å were used. In order to achieve electrical neutrality, 4 (G(h) and C(h)) or 6 (G(w) and C(w)) sodium ions were then included in each of the systems as follows (35): the interaction energy was calculated at all water oxygen positions, and the water molecule with the most positive electrostatic interaction potential was replaced by a sodium ion. This procedure was iterated until all the sodium ions were placed in minimum energy configurations around the DNA:echinomycin complexes, after which both ions and water molecules were subjected to another 1000 steps of energy minimization.

The whole solvated neutral systems were then energy-minimized (3000 steps) and further subjected to molecular dynamics simulations in which both temperature and pressure were weakly coupled to thermal and pressure baths (36) with relaxation times of 0.1 and 0.5 ps, respectively. In a 5-ps heating phase, the temperature was linearly interpolated from 0.1 to 300 K. The reference pressure was 1 atm throughout the simulations, and all bonds were constrained to their equilibrium values by means of the SHAKE algorithm (37). The time step used was 2 fs, and the lists of nonbonded pairs were updated every 25 steps. System coordinates were saved every 0.1 ps for subsequent analysis, which were performed under the same conditions as before, *i.e.* using a residue-based cutoff of 8 Å and a dielectric of 1.

On the basis of the results from our previous simulations of echinomycin bound to $d(ACGT)_2$ and $d(TCGA)_2$ (18), we found it necessary to introduce a limited number of constraints. 1) Attempts to avoid the fraying effects observed in some unrestrained simulations by means of distance and angle constraints for the hydrogen bonds between the terminal bases were unsuccessful for some of the complexes. Therefore, the atoms of the terminal base pairs were restrained to their positions at 0 ps by means of a weak harmonic potential with a force constant of 2 Kcal mol^{-1} $Å^{-2}$. 2) An upper-bound constraint of 5 Kcal mol^{-1} $Å^{-2}$ was applied to the distance between the carbonyl oxygens of the two valine residues of echinomycin during the minimization and the first 10 ps of the dynamics simulations. This constraint, which was interpolated to 0 during the following 10-15-ps interval, was introduced

to prevent the carbonyl oxygens of valines from flipping and facing the DNA rather than the solvent, as already detected in some of our previous simulations and also found for one of the ester linkages of triostin A in a crystal structure (8). Thus, the antibiotic's conformation observed in most X-ray complexes was initially favored in all the complexes. 3) The NH(Ala)-N3(G) hydrogen bonds were reinforced by means of upper-bound harmonic restraining functions with force constants of 5 kcal mol^{-1} $Å^{-2}$ for distances and 5 kcal mol^{-1} rad^{-2} for angles, during the first 35 ps of the simulations. These restraints were later removed by linearly interpolating them to 0 during the ensuing 35-40-ps interval. Thereafter the simulations continued for 30 more ps, thus totalling 70 ps of dynamics simulation for each complex.

Results and Discussion

Monitoring the potential energies of the drug-DNA complexes along the trajectories provided an indication that the solvated systems were stabilized during the sampling period of the simulations (Figure 3A). The evolution of the root-mean-square deviation (RMSD) of each complex with respect to its corresponding initial structure allowed us to get a measure of the extent of the conformational changes undergone by the complexes as the simulation proceeded from the starting minimized coordinates (Figure 3B). It is noteworthy that the conformations of the G(h) and C(w) complexes were not significantly altered, as assessed by the low RMSD, even after removing the NH(Ala)-N3(G) hydrogen bonding constraints, whereas comparatively larger deviations from the initial structures were observed for G(w) and C(h). For this latter complex in particular, which was never detected in the NMR studies, the RMSD was found to increase substantially shortly after these restraints were deleted. These findings can be interpreted in terms of intrinsic greater stabilities for the complexes that were experimentally found at low pH, and are in agreement with additional calculations (see below), which predict a large increase in the interaction energy of echinomycin binding to $d(GCGC)_2$ upon protonation of the terminal cytosines and subsequent Hoogsteen pair formation.

General Structure of the Complexes

The four complexes retained their initial Hoogsteen (G(h) and C(h)) or Watson-Crick (G(w) and C(w)) pairing scheme during the molecular dynamics simulations in aqueous solution. The echinomycin molecule remained stably docked into the DNA tetramers until the end of the simulations. Only in the case of C(h), the complex whose existence has not been demonstrated experimentally, did one of the antibiotic's chromophores partially stick out after removing the intermolecular hydrogen bonding constraints.

The process of DNA intercalation is accompanied by helix unwinding and changes in sugar puckering. The unwinding angles in the four complexes were within the range found in X-ray crystallographic (6) and NMR (33) studies, or reported for dynamic simulations on similar systems (18). NMR analyses of the complexes of quinoxaline antibiotics with DNA oligonucleotides have reported an N-type conformation (near C3′ endo) for the sugars of the pyrimidine nucleotides sandwiched

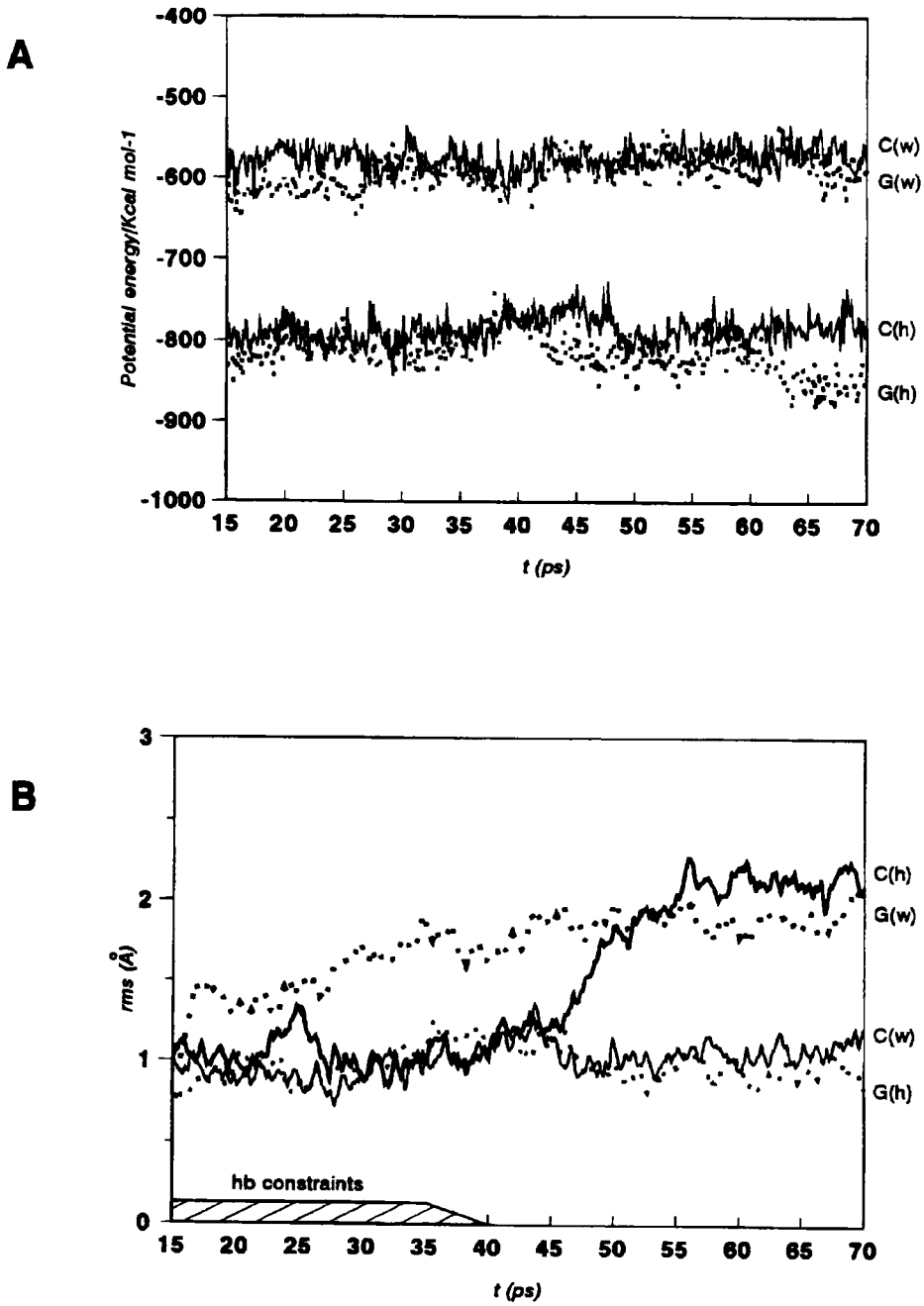

Figure 3: A. Potential energies of the complexes as a function of time. The differences between the Watson-Crick and Hoogsteen complexes are mainly due to the increased electrostatic term arising from the two protonated cytosines in G(h) and C(h). **B.** Root-mean-square deviations from the initial structures, calculated for all non-hydrogen atoms after least-square fitting of the structures using the same atoms.

Table I
Total interaction energies (kcal mol^{-1}) between echinomycin and each of the DNA tetramers averaged over the 25-35-ps and 40-70-ps intervals of the dynamics simulations. By comparing the two sets of values the effect of removing the intermolecular hydrogen bonding constraints in each complex can be assessed.

complex	25-35-ps interval	40-70-ps interval
G(w)	−129.8 ± 4.3	−108.9 ± 5.7
G(h)	−147.6 ± 4.3	−147.2 ± 4.7
C(w)	−131.3 ± 4.1	−125.7 ± 4.9
C(h)	−123.2 ± 4.9	−103.6 ± 7.4

between the two quinoxaline chromophores (10-12,33) as opposed to the C2′ endo (S-type) typical of B-DNA. In our systems, sugar puckering was monitored by measuring the pseudorotation phase angle of the deoxyriboses at the central dinucleotide step along the dynamics trajectory (data not shown). We found a general tendency for these sugars to fall into the N region of the pseudorotation cycle. C2 and C6 were consistently found to yield N-type sugar puckers in all the complexes, with the only exception of C6 in G(h), which smoothly fluctuated between C2′ endo and C4′ exo conformations.

The distribution of the sodium ions surrounding the complexes is thought to contribute significantly to the overall stabilization of the systems. The average Na^+ diffusion constants, calculated as described previously (18), are sufficiently low (G(h): $0.55\ 10^{-5}$, G(w): $0.75\ 10^{-5}$, C(h): $0.10\ 10^{-5}$, and C(w): $0.82\ 10^{-5}$ $cm^2\ s^{-1}$) to make us confident that the method used for their location effectively placed them in minimum energy configurations.

Echinomycin-DNA Interactions

The values shown in Table I are time-averaged echinomycin-DNA interaction energies. Echinomycin binding to $d(GCGC)_2$ with the terminal base pairs in Hoogsteen conformation is notably favored over the rest, followed by $d(CCGG)_2$ with all bases paired according to Watson-Crick. These are precisely the conformations found for these two echinomycin-bound tetranucleotides at low pH (10). In order to calculate the respective binding entalphies, the conformational energy change upon binding of the two molecules should be taken into account (38) but there are too many uncertainties regarding the conformational state of the DNA tetramers prior to binding of the drug, especially for the protonated species. Therefore, overall interaction energies were compared.

Of the total interaction energies about 50 per cent is contributed by the van der Waals′ interactions involving the quinoxaline ring systems alone (Table II), in agreement with previous calculations (17,18). In fact, the stacking interactions arising from these chromophores have been thought to enhance the stability of the complexes (39) and have also been suggested to contribute to the sequence requirements for Hoogsteen base pair formation (9,11,18). For these reasons, and in order to understand the differences among the four complexes studied and to facilitate the

Table II

van der Waals' contributions (kcal mol^{-1}) to the DNA interaction energy of individual residues of echinomycin averaged over the 25-35-ps and 40-70-ps intervals of the dynamics simulations.

	qxn		Ser		Ala		Val		Cys	
complex	25-35 ps	40-70 ps	25-35 ps	40-70 ps	25-35 ps	40-70 ps	25-35 ps	40-70 ps	25-35 ps	40-70 ps
G(w)	−55.8 ± 2.0	−52.7 ± 2.5	−8.0 ± 0.6	−6.2 ± 0.7	−7.6 ± 1.4	−6.5 ± 1.3	−5.6 ± 0.5	−6.1 ± 0.9	−2.4 ± 0.2	−2.5 ± 0.4
G(h)	−57.2 ± 2.3	−56.8 ± 2.4	−10.8 ± 0.8	−10.7 ± 0.9	−7.2 ± 1.5	−6.6 ± 1.6	−7.5 ± 0.8	−6.6 ± 0.7	−2.4 ± 0.1	−2.3 ± 0.2
C(w)	−57.4 ± 1.8	−56.8 ± 2.0	−9.9 ± 1.0	−9.1 ± 0.8	−7.5 ± 1.6	−7.8 ± 1.3	−7.1 ± 0.7	−8.0 ± 0.8	−2.7 ± 0.2	−2.7 ± 0.3
C(h)	−52.5 ± 2.4	−47.8 ± 3.3	−9.9 ± 0.8	−7.8 ± 0.9	−7.3 ± 1.6	−7.2 ± 1.3	−7.7 ± 1.0	−6.6 ± 1.0	−2.5 ± 0.2	−2.3 ± 0.2

Table III

40-70 ps average H-acceptor distances (in Å) between the NH and CO groups of the alanine residues of echinomycin and the N3 and 2-NH_2 of guanines (NH-N3 and O-HN2, respectively) determining intermolecular hydrogen bond formation in the four complexes studied.

	G(w)		G(h)		C(w)		C(h)	
	NH-N3	O-HN2	NH-N3	O-HN2	NH-N3	O-HN2	NH-N3	O-HN2
Ala3 - G7	2.0 ± 0.2	2.5 ± 0.3	1.9 ± 0.1	2.3 ± 0.2	2.2 ± 0.4	2.2 ± 0.2	3.6 ± 0.8	2.7 ± 0.5
Ala8 - G3	4.1 ± 0.5	2.3 ± 0.4	1.9 ± 0.1	2.6 ± 0.4	3.0 ± 0.4*	2.1 ± 0.3	2.0 ± 0.2	2.2 ± 0.3

* replaced with an intramolecular hydrogen bond (2.2 ± 0.3 Å) with the CO of the adjacent qxn-6 (see Figure 4).

interpretation of the results, we found it useful to dissect the overall interaction energies into the contributions originating from these chromophores and those involving the depsipeptide part of the antibiotic.

(1) Hydrogen Bonds and van der Waals' Interactions Involving the Depsipeptide Part of the Antibiotic

The echinomycin-DNA complexes are stabilized by a number of hydrogen bonds between the NH and CO groups of the alanine residues of the drug and the N3 and H-N2 atoms of guanines in the central CpG step (3,40). By comparing the four complexes studied (Table III), it is readily seen that only in G(h) is the pattern of hydrogen bonds analogous to that found in the crystal structures of the complexes of echinomycin or triostin A with $d(CGTACG)_2$ (6) and triostin A with $d(GCGTACGC)_2$ (8), *i.e.* two strong hydrogen bonds between the NH groups of the alanines and the N3 atoms of guanines, and a weaker one between the carbonyl group of one of the alanines and the facing 2-NH_2 group of guanine. The remaining O-N2 distance (ranging from 3.6 to 4.1 Å in the solid-state complexes) has been considered too long to be a hydrogen bond. In agreement with this, during the simulation of G(h) in water, both NH-N3 hydrogen bonds remained strong and stable, whereas the O-N2 distances were longer and gave rise to an additional weaker hydrogen bond for Ala-3-G7 and to a more transient and fluctuating hydrogen bond for the Ala-8-G3 interaction. As a matter of fact, G(h) is the only complex which, in common with the crystal structures, has an RpCpGpY sequence and the base pairs flanking the CpG site are Hoogsteen paired.

The crystallographic pattern is altered in one way or another in the other three complexes (Table III). In C(w) we detect an intramolecular hydrogen bond between the NH of Ala-8 and the CO of the neighboring quinoxaline residue (qxn-6) similar to that encountered in the simulations of echinomycin bound to $d(ACGT)_2$ and $d(TCGA)_2$ with all the base pairs in Watson-Crick conformation (18). In fact, N3 of G3 and CO of qxn-6 are competing for hydrogen bonding to this NH, and there also appears to be an exchange between the two NH-N3 hydrogen bonds on both sides of the complex (Figure 4). Interestingly, for the closely related triostin A, this sort of intramolecular hydrogen bond has been proposed to exist for one of the conformers in solution (41).

These findings might be significant for two main reasons: (i) the hydrogen bonds between the NH of alanines and the N3 of guanines were reinforced in all the complexes by means of distance and angle constraints during the first 35 ps of the simulations but they evolved differently thereafter; and (ii) a similar behavior was detected in the complexes of echinomycin with $d(ACGT)_2$ and $d(TCGA)_2$ during 40-ps unrestrained simulations in water (18). It thus appears that the distinct hydrogen bonding patterns observed depend to a certain extent on the nature and conformation of the terminal base pairs. The compression of the minor groove brought about by Hoogsteen pairing of the terminal bases in G(h) appears to bring the N3 atoms of the central guanines closer to the NH's of the alanines, which may account for the enhanced NH-N3 hydrogen bonding interactions in this complex.

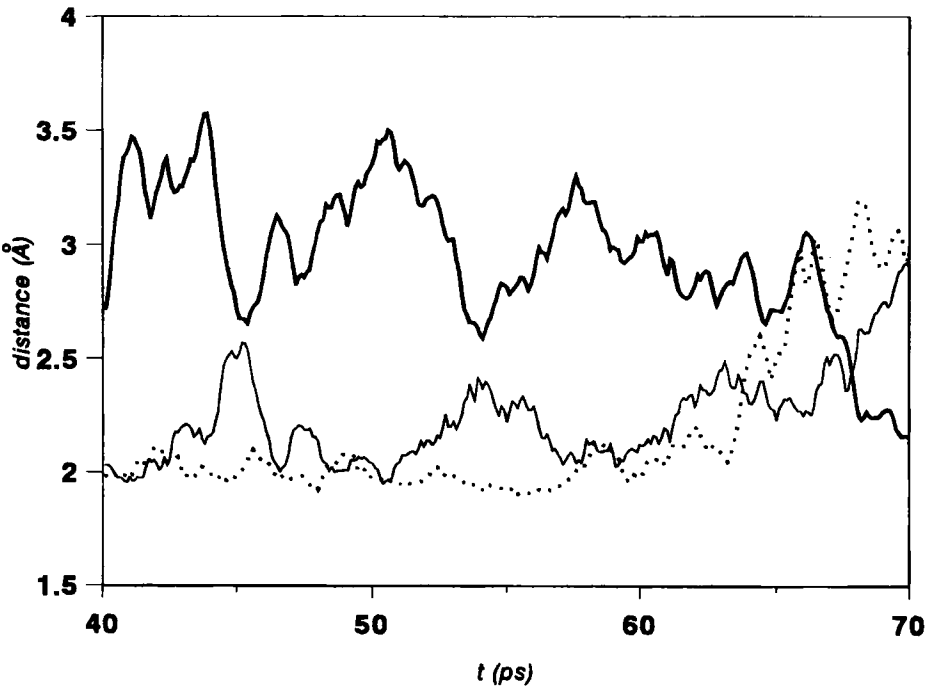

Figure 4: H-acceptor distances as a function of time for the NH(Ala3)-N3(G7) (dotted), NH(Ala8)-N3(G3) (thick), and NH(Ala8)-O(qxn6) (thin) hydrogen bonds in C(w). The data were smoothed after filtering high-frequency noise by "moving window" averaging *(boxcar smoothing)* using a window width of 1.1 ps (11 sets of distances).

One factor that was initially thought to be important for the stabilization of the Hoogsteen conformation in the $d(CGTACG)_2$-(triostin A)$_2$ crystal structure was the valine-DNA contacts, which would be presumably improved due to the narrowing of the DNA minor groove in this conformation (8,17). In our equilibrated Hoogsteen complexes, however, we do not detect any enhanced van der Waals' interaction between the valine residues and the DNA. Instead, it is the serine residues which in G(h) provide an extra dispersion energy of about 4.5 Kcal mol^{-1} with respect to G(w) (Table II), half of it contributed by enhanced interactions with the base pairs. Similar results were obtained in the simulations of $d(ACGT)_2$:echinomycin and $d(TCGA)_2$:echinomycin (18), which were supported by a number of NOE's detected in the NMR experiments for the former complex but not for the latter (9,10). In any case, as noted previously (9,11), these interactions cannot be the dominant force favoring one DNA conformation over the other because they are not sequence-dependent.

(2) Stacking Interactions Between the DNA Base Pairs and the Antibiotic's Chromophores

We have already stressed that, given the coplanarity and conjugation of the aromatic ring with the peptide bond linking it to the d-serine residue, the quinoxaline-2-carboxamide should be considered as the actual chromophore giving rise to stacking interactions with the neighboring DNA base pairs (18). These chromophores, as well as the G:C and $G:C^+$ base pairs they interact with, are highly polarized systems,

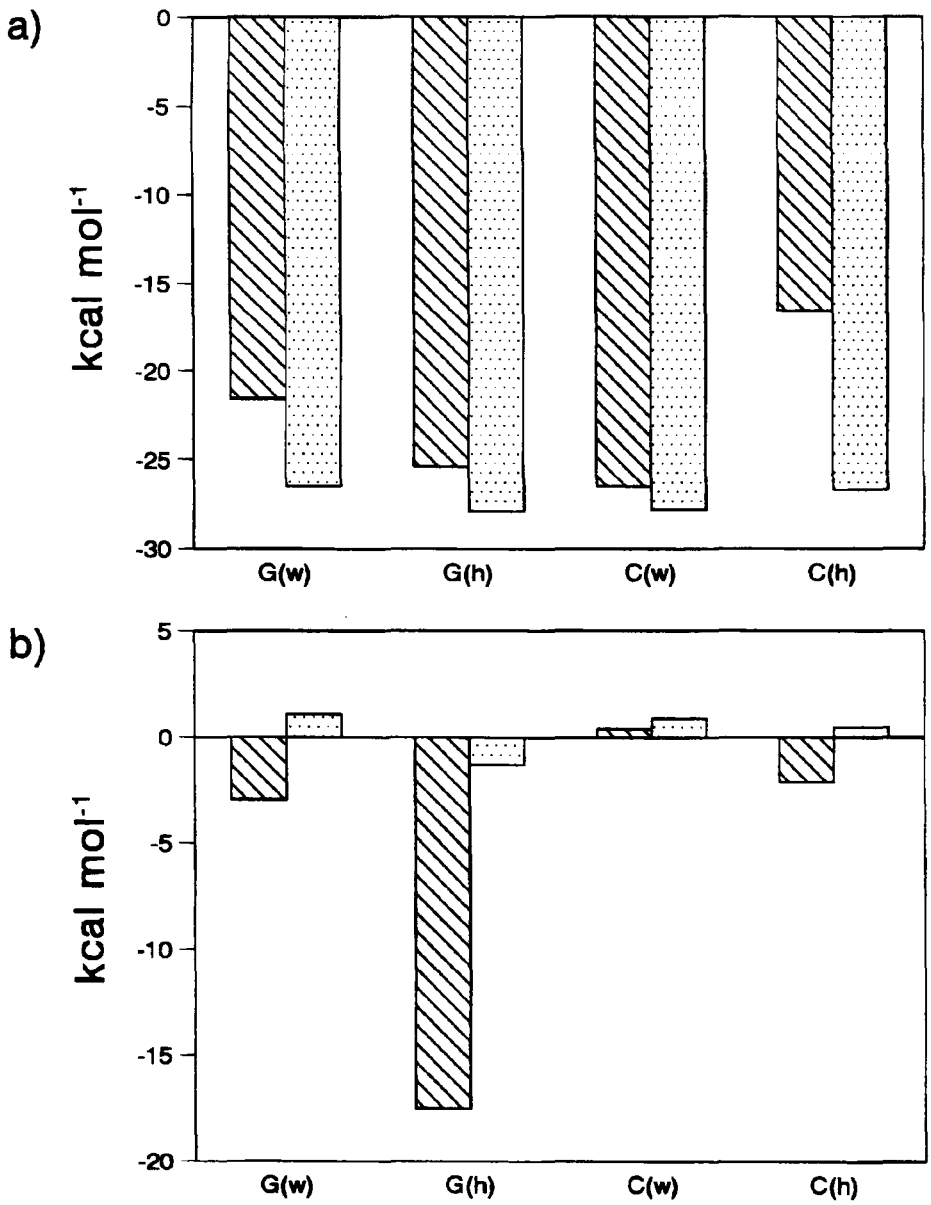

Figure 5: Van der Waals (a) and electrostatic (b) energy components (Kcal mol^{-1}) of the stacking interaction between the *N*-C_α-quinoxaline-2-carboxamide chromophores of echinomycin and either the central (dotted bars) or the flanking (hatched bars) G:C base pairs of the DNA tetramers, averaged over the 40-70 ps interval of the simulations. The electrostatic term was obtained by means of a point charge interaction model employing a dielectric constant of 1. The C1′ atoms of deoxyriboses and C_α atoms of serines were used as buffers in order to achieve electrical neutrality or a total charge of +1 for $G:C^+$ (18). All of the energy values result from the summation of the corresponding stacking interactions from both sides of the complexes. The calculated dipole moments for the C1′-buffered unprotonated G:C pairs and C_α-buffered *N*-C_α-quinoxaline-2-carboxamide systems, averaged over the last 30 ps of the simulations and over all the complexes, are 5.0 and 4.4 Debyes, respectively. The angles (in degrees) formed between the dipole moment vectors of the neutral G:C pairs and the *N*-C_α-quinoxaline-2-carboxamide systems, averaged over the 40-70-ps interval and both sides of the complexes, are the following: (i) with central G:C pairs: G(h), 21.3; G(w), 14.9; C(h), 16.1; C(w), 12.8. (ii) with flanking G:C pairs (*cf.* Figure 6): G(w), 139.4; C(w), 29.9.

and therefore their relative orientation has a profound influence on the magnitude of their interaction, especially regarding the electrostatic contribution.

The van der Waals' and electrostatic components of the stacking interactions have been calculated for the sandwiched base pairs and for the base pairs flanking the bisintercalation site, and are graphically displayed in Figure 5. Since all the complexes share a common central CpG step, the major differences among them are due to interactions involving the terminal base pairs. The most striking results are (i) the large gain in electrostatic stacking interaction energy for echinomycin binding to $d(GCGC)_2$ upon Hoogsteen pairing of the terminal bases, and (ii) the reduced van der Waals' stacking interaction if the terminal guanine bases in $d(CCGG)_2$ were to adopt a Hoogsteen conformation. The combination of these two factors can account for the experimental observation that at low pH Hoogsteen base pair formation is only observed for the echinomycin-$d(GCGC)_2$ complex (10).

The increase in stabilization energy for the G(h) complex can be explained by considering the orientation of the $G:C^+$ base pair with respect to the quinoxaline-2-carboxamide system of echinomycin (Figure 6). Note that the positively charged cytosine base lies directly above (or below) the most negatively charged region of the drug's chromophore (in the vicinity of the carbonyl oxygen), whereas the positively

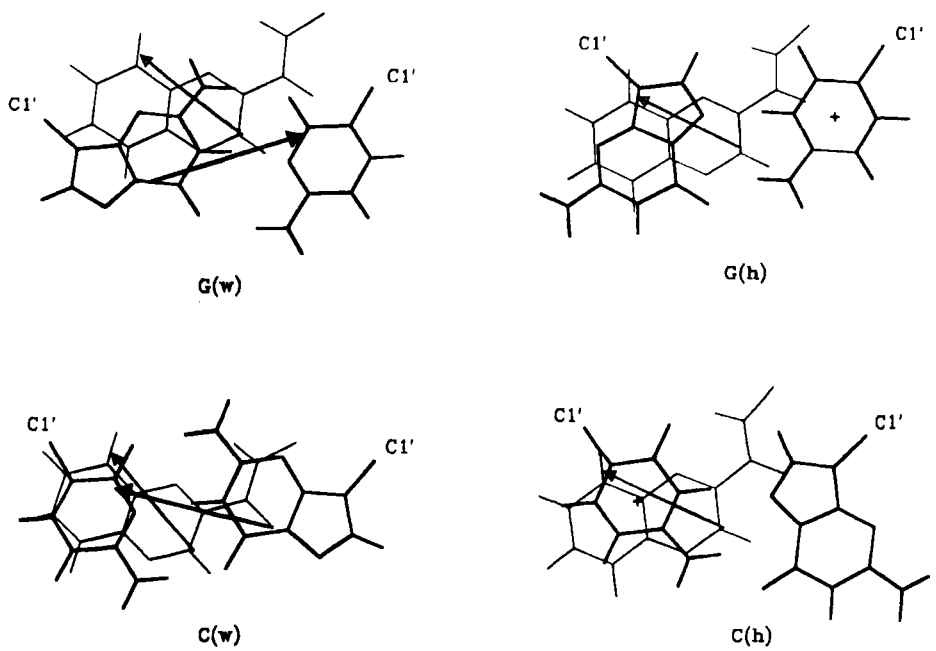

Figure 6: Stacking geometries between the quinoxaline-2-carboxamide chromophores of echinomycin (thin lines) and the terminal G:C base pairs (thick lines) in the four complexes studied. The structures were taken from the optimized 40-50-ps averaged coordinates of the simulations in aqueous solution. The dipole moments of the echinomycin chromophores and of the G:C base pairs in Watson-Crick conformation, depicted by thin and thick arrows, respectively, represent the polarity of the charge distributions. The midpoint of each vector is centered on the geometrical center of the system considered. In the Hoogsteen complexes, the terminal cytosines are protonated and marked with a "+" symbol.

charged edge of the benzene ring of quinoxaline is in close proximity to the electron-rich N3 of guanine. This optimized electrostatic complementarity between the two planar systems is also accompanied by an increased van der Waals' interaction with respect to G(w) (Figure 5a). Thus, this conformation is strongly stabilized at low pH. As regards the CpCpGpG sequence, in C(w) the carbonyl oxygen of quinoxaline-2-carboxamide is near the region in guanine associated with the most negative electrostatic potential (in the vicinity of N7 and O6), which results in unfavorable electrostatic interactions between these two stacked systems (Figure 5b), as can be seen in a simplified fashion by the parallel arrangement of dipole moment vectors (Figure 6). Upon cytosine protonation and Hoogsteen base pair formation, no better electrostatic complementarity is achieved for this complex even though the overall electrostatic interaction appears to improve slightly (presumably due to the long-range character of the Coulombic term employed in the force field, which varies as r^{-1}), but only at the cost of a larger loss in van der Waals' interaction (Figure 5).

On the other hand, the unfavorable electrostatic interactions already detected between the quinoxaline-2-carboxamide systems of the drug and the central CpG step (18) are further confirmed in the present complexes (Figure 5b), with the sole exception

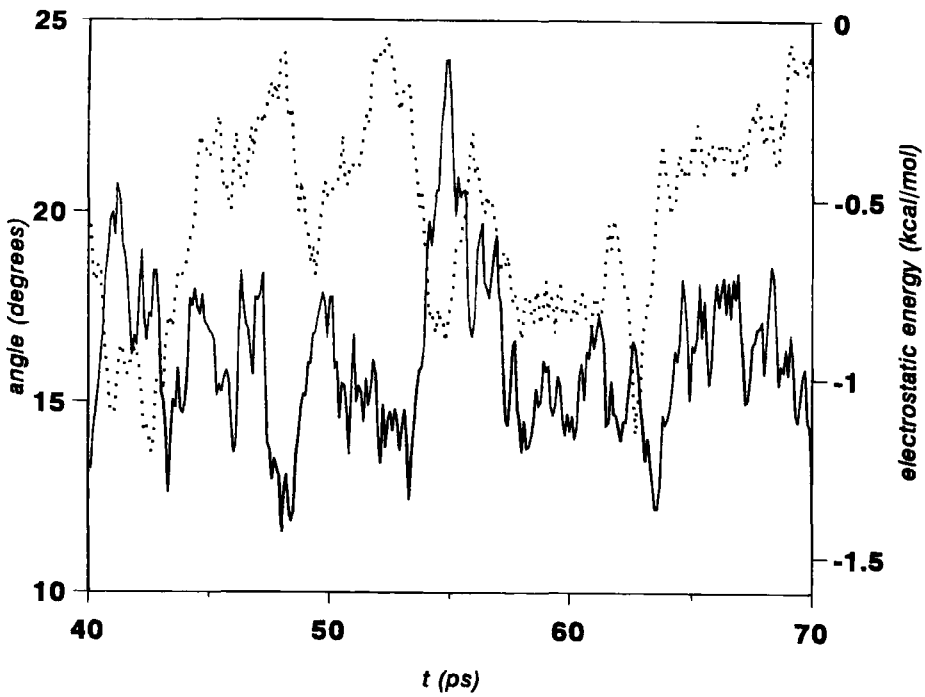

Figure 7: Dependence of the electrostatic contribution to the stacking interaction energy (thin line) on the orientation of the two stacked systems during the 40-70-ps sampling period of the dynamics simulation of G(h). The relative orientation was measured by the angle formed between the dipole moment vectors of one *N*-C_α-quinoxaline-2-carboxamide chromophore of echinomycin (qxn-1) and the neighboring central G:C base pair (dotted line). Angles close to 0° represent parallel arrangements of dipole moments (Figure 6), which translate into electrostatic charge repulsion, whereas angles close to 180° would denote antiparallel arrangements leading to favorable dipolar interactions. Smoothing of the data was accomplished as in Figure 4.

of G(h). It is precisely for this complex that the angle formed between the dipole moment vectors of the G:C base pairs and the *N*-C_α-quinoxaline-2-carboxamide systems is more open (*cf.* legend to Figure 5). These vectors, and the angle formed between them, provide a simple representation of the polarity of the charge distributions and a convenient way of measuring the relative orientation of the stacked systems. In fact, if this angle and the electrostatic interaction energy are monitored along the dynamics trajectories, a negative correlation can be found, clearly for some complexes, *e.g.* G(h) (Figure 7), and less so for others in which this picture is probably clouded by several overlapping effects (data not shown). This correlation supports the results found on simpler model systems (42) and highlights the importance of chromophore orientation in optimizing the electrostatic component of the stacking interactions. In the case of echinomycin, the chromophores are held rather rigidly by the depsipeptide ring, which allows very limited rotational freedom. Nevertheless, the generally unfavorable electrostatic stacking interaction between echinomycin and the central G:C base pairs is offset by the less discriminating van der Waals component plus the very favorable hydrogen bonding and van der Waals' interaction of the depsipeptide with the minor groove, and the antibiotic shows a marked specificity of binding to a central CpG step. This binding selectivity disappears, however, when 2,6-diamino-purine (DAP) is incorporated into DNA in place of adenine (43). We have found that a DAP:T base pair, while also presenting the 2-amino group in the minor groove, is endowed with a significantly lower dipole moment, and the different charge distribution gives rise to an attractive electrostatic stacking interaction with the quinoxaline-2-carboxamide system. This leads to an improved calculated interaction energy for echinomycin binding to TpDAP steps on this modified DNA (32), which is in consonance with the increased association constant determined experimentally (43).

Conclusions

Protonation of cytosine at low pH hampers the normal pairing with a guanine base due to the steric clash between the incoming proton and the hydrogen on N1 of guanine (Figure 1). Rotation of the guanine about the glycosidic bond to adopt a *syn* orientation with respect to the sugar allows an alternative pairing scheme to be formed in which one hydrogen bond has been lost with respect to a Watson-Crick G:C base pair. It was intriguing to us that this conformational change would take place in $d(GCGC)_2$ at low pH upon binding of echinomycin, but not in $d(CCGG)_2$ under the same conditions (10). The present results provide a plausible explanation for these findings and further suggest (18) that stacking forces may be playing an important and discriminatory role in the interaction of DNA bases with ligands containing dipolar planar chromophores.

We propose that, contrary to earlier hypotheses (8,44), the Watson-Crick to Hoogsteen transition in echinomycin-DNA complexes is triggered by the sequence-dependent electrostatic component of the stacking interactions between the quinoxaline chromophores of the drug and the base pairs flanking the central bis-intercalation step, as already reported for echinomycin binding to $d(ACGT)_2$ (18). Additional stabilization can then be brought about by improved van der Waals contacts between the DNA and both the depsipeptide and the planar ring systems of the antibiotic, in

agreement with previous suggestions (8,17). It is noteworthy that electrostatic interactions between the stacked bases have been also shown to be largely responsible for the conformational preferences of DNA base-pair steps (45), and a dominant factor in determining the stacking patterns of nucleic acid constituents (46) and highly polar heteroaromatic molecules (47).

The conformational changes brought about by the binding of echinomycin and other bisintercalating agents are not independent of helical constraints. Thus, demonstrating stable Hoogsteen pairing for internal base pairs in long DNA tracts has proved elusive so far. Among the limitations of current experimental procedures are the lack of suitable reagents and the random nature of most of the DNA sequences employed in the footprinting experiments, which do not contain the exact sequences for which the Hoogsteen rearrangement has been detected upon binding of *two* echinomycin or triostin A molecules by either NMR or x-ray crystallographic techniques (*e.g.* ACGTACGT or GCGTACGC). The shorter ACGT and GCGC sequences are indeed found in the *tyrT* fragment commonly used in these experiments (2). Bonds 76 and 95 are in fact a perfect match for GCGC but their context is completely different, which translates into very different protection profiles at both sites both on normal DNA and DNA containing 2'-deoxy-7-deazaguanosine in place of 2'-deoxyguanosine (2). This deaza-nucleoside could also in principle adopt a *syn* conformation but the resulting Hoogsteen-like pair with the facing cytosine would be stabilized by just one hydrogen bond as opposed to the three hydrogen bonds that can be formed when the pair adopts a Watson-Crick scheme. The footprinting results thus suggest that Hoogsteen pairs are not a prerequisite for echinomycin binding to DNA (2,16). Nevertheless, the fact that these changes do indeed take place in short oligonucleotides upon binding of echinomycin or triostin A provides an excellent test case for probing the nature and dependence of stacking interactions in DNA-drug complexes and hints at the rational modification of existing ligands.

Time-averaged (40-70 ps) cartesian coordinates for the four complexes in PDB format are available from the authors on request (e-mail: ffgago@alcala.es).

Acknowledgments

This research has been financed in part by the Spanish CICYT (Project FAR91-0277), and the University of Alcalá de Henares. We thank Dr. Manuel Pastor and Prof. Julio Alvarez-Builla for many helpful discussions and a generous allowance of computer time. J.G. is recipient of a pre-doctoral grant from the Spanish Ministerio de Educación y Ciencia (FPI, PN89-25126387). Biosym Technologies, Inc. (San Diego, California) contributed a software license.

References and Footnotes

1. Portugal, J., *Trends Pharmacol. Sci. 14*, 127-130 (1989).
2. E.W. Sayers and M.J. Waring, *Biochemistry 32*, 9094-9107 (1993).
3. M.J. Waring, in Molecular Basis of Specificitiy in Nucleic Acid-Drug Interactions, B. Pullman and J. Jortner, Eds., pp. 225-245, Kluwer Academic Publishers, Dordrecht (1990).
4. K. Hoogsteen, *Acta. Cryst. 2*, 822-823 (1959).

5. A.H-J. Wang, G. Ughetto, G.J. Quigley, T. Hakoshima, G.A. van der Marel, J.H. van Boom and A. Rich, *Science 225*, 1115-1121 (1984).
6. G. Ughetto, A.H-J. Wang, G.J. Quigley, G.A. van der Marel, J.H. van Boom and A. Rich, *Nucleic Acids Res. 13*, 2305-2323 (1985).
7. G.J. Quigley, G. Ughetto, G.A. van der Marel, J.H. van Boom, A.H-J. Wang and A. Rich, *Science 232*, 1255-1258 (1986).
8. A.H-J. Wang, G. Ughetto, G.J. Quigley and A. Rich, *J. Biomol. Struct. Dyn. 4*, 319-342 (1986).
9. X. Gao and D.J. Patel, *Biochemistry 27*, 1744-1751 (1988).
10. X. Gao and D.J. Patel, *Quart. Rev. Biophys. 22*, 93-138 (1989).
11. D.E. Gilbert and J. Feigon, *Biochemistry 30*, 2483-2494 (1991).
12. D.E. Gilbert and J. Feigon, *Nucleic Acids Res. 20*, 2411-2420 (1992).
13. F.C. Seaman and L. Hurley, *Biochemistry 32*, 12577-12585 (1993).
14. D. Mendel and P.B. Dervan, *Proc. Natl. Acad. Sci. USA 84*, 910-914 (1987).
15. J. Portugal, K.R. Fox, M.J. McLean, J.L. Richenberg and M.J. Waring, *Nucleic Acids Res. 16*, 3655-3670 (1988).
16. J.M. McLean, F. Seela and M.J. Waring, *Proc. Natl. Acad. Sci. USA 86*, 9687-9691 (1989).
17. U.C. Singh, N. Pattabiraman, R. Langridge and P.A. Kollman, *Proc. Natl. Acad. Sci. USA 83*, 6402-6406 (1986).
18. J. Gallego, A.R. Ortiz and F. Gago, *J. Med. Chem. 36*, 1548-1561 (1993).
19. A preliminary account of these findings was presented at the Workshop on *DNA-drug interactions* held at the Fundación Juan March, Madrid (15-17 November, 1993).
20. D.A. Pearlman, D.A. Case, J. Caldwell, G. Seibel, U.C. Singh, P. Weiner and P.A. Kollman, *AMBER* version 4.0 (1991), Department of Pharmaceutical Chemistry, University of California, San Francisco.
21. Insight-II, version 2.1.0 (1992), Biosym Technologies. 9685 Scranton Road, San Diego, CA 92121-2777.
22. S.J. Weiner, P.A. Kollman, D.T. Nguyen and D.A. Case, *J. Comp. Chem. 7*, 230-252 (1986).
23. W.L. Jorgensen, J. Chandrasekhar and J.D. Madura, *J. Chem. Phys. 79*, 926-935 (1983).
24. J. Aqvist, *J. Phys. Chem. 94*, 8021-8024 (1990).
25. P.C. Hariharan and J.A. Pople, *Chem. Phys. Lett. 16*, 217-219 (1972).
26. U.C. Singh and P.A. Kollman, *J. Comp. Chem. 5*, 129-145 (1984).
27. F.A. Momany, *J. Phys. Chem. 82*, 592-601 (1978).
28. F.J. Luque, M. Orozco, F. Illas and J. Rubio, *J. Am. Chem. Soc. 113*, 5203-5211 (1991).
29. M. Orozco, W.L. Jorgensen and F.J. Luque, *J. Comp. Chem. 14*, 1498-1503 (1993).
30. M.L. Connolly, *J. Appl. Crystal. 16*, 548-558 (1983).
31. M. Orozco and F.J. Luque, *J. Comp.-Aided Mol. Design 4*, 411-426 (1990).
32. J. Gallego, F.J. Luque, M. Orozco, C. Burgos, J. Alvarez-Builla, M.M. Rodrigo and F. Gago, *J. Med. Chem. 37*, 1602-1609 (1994).
33. K.J. Addess, J.S. Sinsheimer and J. Feigon, *Biochemistry 32*, 2498-2508 (1993).
34. F.C. Bernstein, T.F. Koetzle, G.J.B. Williams, E.F. Meyer Jr., M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi and M. Tasumi, *J. Mol. Biol. 112*, 535-542 (1977).
35. W.F. van Gunsteren, H.J.C. Berendsen, R.G. Geurtsen and H.R.J. Zwinderman, *Ann. N. Y. Acad. Sci. 482*, 287-303 (1986).
36. H.J.C., Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola and J.R. Haak, *J. Chem. Phys. 81*, 3684-3690 (1984).
37. J.P. Ryckaert, G. Ciccotti and H.J.C. Berendsen, *J. Comp. Phys. 23*, 327-341 (1977).
38. F. Gago, C.A. Reynolds and W.G. Richards, *Molec. Pharmacol. 35*, 232-241 (1989).
39. T.V. Alfredson and A.H. Maki, *Biochemistry 29*, 9052-9064 (1990).
40. C. Marchand, C. Bailly, M. McLean, S.E. Moroney and M.J. Waring, *Nucleic Acids Res. 20*, 5601-5606 (1992).
41. J.R. Kalman, T.J. Blake, D.H. Williams, J. Feeney and G.C.K. Roberts, *J. Chem. Soc. Perkin Trans. I*, 1313-1321 (1979).
42. R.L. Ornstein, R. Rein, D.L. Breen and R.D. MacElroy, *Biopolymers 17*, 2341-2360 (1978).
43. C. Bailly, C. Marchand and M.J. Waring, *J. Am. Chem. Soc. 115*, 3784-3785 (1993).
44. O. Kennard and W.N. Hunter, *Angew. Chem. Int. Ed. Engl. 30*, 1254-1277 (1991).
45. C.A. Hunter, *J. Mol. Biol. 230*, 1025-1054 (1993).
46. C.E., Bugg, J.M., Thomas, M. Sundaralingam and S.T. Rao, *Biopolymers 10*, 175-219 (1971).
47. J.M. Mínguez, T. Gandásegui, J.J. Vaquero, J. Alvarez-Builla, J.L. García-Navío, F. Gago, A.R. Ortiz, P. Gómez-Sal, R. Torres and M.M. Rodrigo, *J. Org. Chem. 58*, 6030-6037 (1993).

Date Received: May 5, 1994

Communicated by the Editor A.H.-J. Wang