Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 7673–7677, August 1994
Biophysics

Relative binding affinities of distamycin and its analog to d(CGCAAGTTGGC)·d(GCCAACTTGCG): Comparison of simulation results with experiment

SURESH B. SINGH*†, AJAY*, DAVID E. WEMMER‡, AND PETER A. KOLLMAN*

*Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143; and ‡Department of Chemistry, University of California, Berkeley, CA 94720

Communicated by I. Tinoco, Jr., March 7, 1994

ABSTRACT **We report here an effort to use molecular dynamics/free energy perturbation methodology to calculate relative binding affinities of two related drugs to DNA. Specifically, we focus on the relative binding free energies of distamycin (Dst) and its analog, 2-imidazoledistamycin (2-ImD), to d(CGCAAGTTGGC)·d(GCCAACTTGCG). The pyrrole (Dst) and the imidazole variant (2-ImD) differ only in that the C-H is substituted by an N in the central ring. The starting conformation for these calculations was the previously determined solution structure of two 2-ImD molecules in the minor groove of the above 11-residue DNA. In this complex both the ligands have the imidazole nitrogen (N3) oriented toward the amino group of G^6. However only ligand 1 (site I) has N3 within the hydrogen bonding distance from N2 amino proton of G^6. We have calculated the difference in free energy of binding of 2-ImD versus Dst in three different cases by mutating 2-ImD → Dst reversibly. In the first case ligand 1 (site I) is mutated, in the second case ligand 2 (site II) is mutated, and in the third case both the ligands are mutated. These calculations show that at site I Dst has weaker binding affinity than 2-ImD by 0.7 kcal/mol, at site II Dst has stronger binding affinity than 2-ImD by 2.9 kcal/mol, and when occupying both site I and site II, Dst binds with greater affinity than 2-ImD by 1.8 kcal/mol. Recent experimental findings agree semiquantitatively (within 1 kcal/mol) with the calculations presented here. Hence the methodology presented here can be used to predict relative binding energies of two or more closely related molecules to DNA.**

One of the goals of modeling drug–receptor interactions is to understand the molecular basis of action of ligands on proteins and DNA. Thus our hope is that we will be able to utilize computational methods and graphics to identify key determinants of the receptor–drug interactions and ultimately be able to develop better drugs that will exhibit greater affinity toward the receptor. Hence it is important to develop methods to predict binding affinities of new drug molecules to receptors. In the case of DNA, one is interested in developing drugs that target a specific sequence. So far molecular dynamics (MD)/free energy perturbation (FEP) methodology has been successfully applied to predict binding affinities of drugs to proteins (1–3). There have been relatively few applications to DNA–drug interactions. The reason for this is that DNA is inherently more flexible than proteins, and also since DNA is a polyanion the treatment of electrostatic interactions is a challenging task. The large size of the system (DNA, ions, and water molecules) makes these simulations computationally intensive. Therefore, we present here a feasible methodology that can be effectively utilized to calculate relative association energies between any two or more similar ligands that bind to the same site on DNA.

Some recent studies of DNA–ligand interaction using MD include the drugs actinomycin D (4), spermine (5), and distamycin (Dst; ref. 6). Free energy calculations on DNA–drug complexes were presented by Cieplak *et al.* (7), who studied the base specificity of the interactions of daunomycin and acridine to DNA. Gago and Richards (8) and Hard and Nilsson (9) have also studied the base specificity of the interactions of netropsin with DNA. To our knowledge, this is the first reported effort to calculate the relative binding free energies of ligands to DNA.

One of the difficulties in applying free energy simulations on macromolecular systems is the long simulation times required to sample the possible conformations while maintaining equilibrium at each stage during the perturbation process. At present there is no way to assure that a simulation has been run long enough so that all possible conformations accessible to the system on the path of transformation have been sampled. In addition to sampling, there is the issue of the accuracy of force field parameters employed in these calculations and also several approximations involved in the handling of the long-range nonbonded interactions. Therefore, it is important to gain as much empirical experience with running FEP calculations under different conditions to study and understand the validity of the approximations involved in these calculations.

The binding of Dst and its variant 2-imidazole distamycin (2-ImD) to DNA represents not only an interesting system to understand DNA–drug interaction but also provides the opportunity to test the FEP methodology. This is because the mutations of the structure are small, involving only the mutation of the van der Waals parameters of N-D → C-H (where D is the dummy atom) and the partial electrostatic charges of all the atoms on the ligand.

Dst, like netropsin, is known to bind specifically to the minor groove of A+T-rich regions of B-DNA. Lavery and Pullman (10) have carried out electrostatic potential calculations and have shown that the minor groove of A+T-rich B-DNA sequence has the lowest negative potential. From this they inferred that the cationic drugs will exhibit binding specificity to the minor groove regions of A+T-rich sequences. The crystal structure of a DNA·Dst complex showed that there is only one molecule bound to the minor groove of an AATT DNA site (11). However, recent experimental studies carried out by Dwyer *et al.* (12) have shown that the introduction of a G·C base pair in an otherwise A+T-rich region widens the minor groove and allows the accommodation of two Dst molecules. According to the recently reported experimental studies by Dwyer *et al.* (12)

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Dst, distamycin; 2-ImD, 2-imidazoledistamycin; MD, molecular dynamics; FEP, free energy perturbation.

†Present address: Wyeth–Ayerst Research, CN 8000, Princeton, NJ 08543.

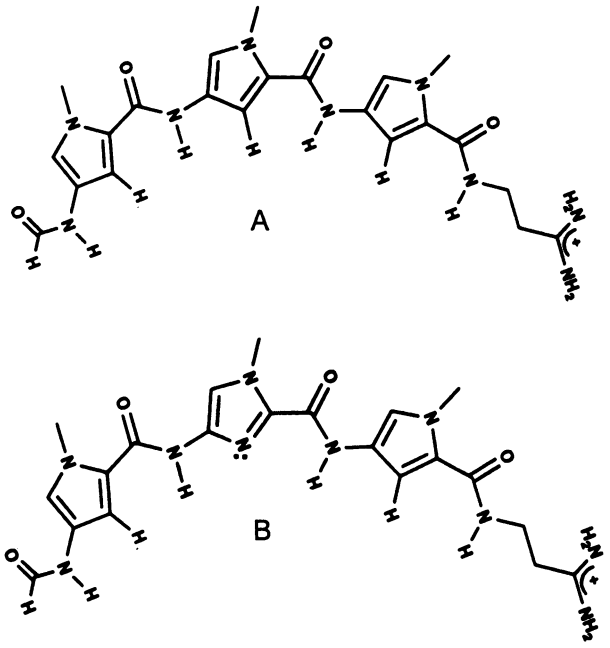

FIG. 1. (*A*) Structure of Dst. (*B*) Structure of 2-ImD.

and Geierstanger *et al.* (13), both Dst and its imidazole variant are found to bind tightly to the 5-bp AAGTT·AACTT site of an 11-mer duplex.

In our present studies we have employed the FEP methodology to predict relative binding affinities of Dst and its analog 2-ImD to d(CGCAAGTTGGC)·d(GCCAACTTGCG). This study was initiated when one of us (D.E.W.) offered to provide the University of California (San Francisco) group with NMR-determined coordinates for the 2:1 2-ImD·2-ImD·DNA complex without any information about the experimental relative free energies of binding of Dst·2-ImD·DNA, 2-ImD·Dst·DNA, Dst·Dst·DNA, and 2-ImD·2-ImD·DNA. Thus, the theoretical calculations were carried out without such knowledge.

Dst has three pyrrole rings, whereas its analog 2-ImD has an imidazole ring substituted for the central pyrrole ring (see Fig. 1). The pyrrole and the imidazole variant differ only in that the C-H is substituted by an N in the central ring. The starting conformation for these calculations is the solution structure of 2-ImD·2-ImD·DNA minor groove complex (see Fig. 2) determined by Dwyer *et al.* (12). In this complex both ligands point the imidazole N3 toward the N2 of G^6. However only ligand 1 has N3 within the hydrogen-bonding distance from N2 of G^6 (see Fig. 3). Ligand 1 lies along the strand with the AAGTT sequence, and ligand 2 lies along the strand with AACTT sequence. For the sake of clarity, we label the site where ligand 1 is bound as site I and the site of ligand 2 as site

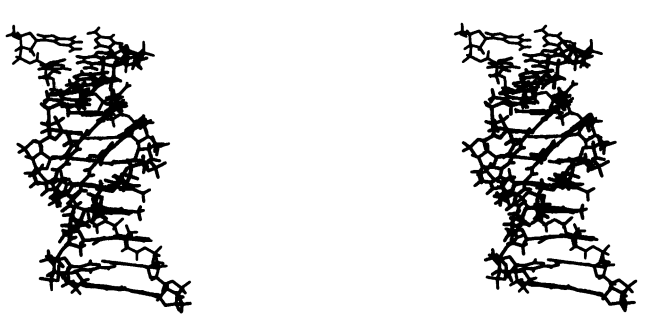

FIG. 2. Starting structure of the 2:1 complex of 2-ImD bound to d(CGCAAGTTGGC)·d(GCCAACTTGCG) shown in a stereoview.

II. We have carried out the following set of simulations to calculate free energy differences for the drug:

(*i*) 2-ImD $\rightleftharpoons$ Dst isolated in gas phase and

(*ii*) 2-ImD $\rightleftharpoons$ Dst isolated in solvent.

The following calculations were carried out on the complex in solution:

(*i*) 2-ImD·2-ImD·DNA $\rightleftharpoons$ Dst·2-ImD·DNA

(*ii*) 2-ImD·2-ImD·DNA $\rightleftharpoons$ 2-ImD·Dst·DNA

(*iii*) 2-ImD·2-ImD·DNA $\rightleftharpoons$ Dst·Dst·DNA.

The thermodynamic cycles employed in these simulations are given in Fig. 4. The experimental binding affinities of Dst and 2-ImD to DNA are given by ΔG_1, ΔG_2, ΔG_5, and ΔG_7. These values can be related to ΔG_3, ΔG_4, ΔG_6, and ΔG_8 computed by FEP simulations. The relative binding affinities of these various complexes can be computed from

$$\Delta\Delta G_1 = \Delta G_4 - \Delta G_3 = \Delta G_1 - \Delta G_2 \quad [1]$$

$$\Delta\Delta G_2 = \Delta G_6 - \Delta G_3 = \Delta G_5 - \Delta G_2 \quad [2]$$

$$\Delta\Delta G_3 = \Delta G_8 - 2\Delta G_3 = \Delta G_7 - \Delta G_2. \quad [3]$$

METHODS

The starting conformation for the FEP calculations was the NMR structure of the 2:1 complex of 2-ImD bound in the minor groove of d($C^1G^2C^3A^4A^5G^6T^7T^8G^9G^{10}C^{11}$)·d($G^{12}C^{13}$·$C^{14}A^{15}A^{16}C^{17}T^{18}T^{19}G^{20}C^{21}G^{22}$) (12). The starting geometry for Dst was obtained by taking 2-ImD from the NMR structure and then replacing N3 with the C-H group. The resultant geometry of Dst was then optimized using the STO-3G basis set (14). The electrostatic charges for 2-ImD and Dst were then calculated by employing the STO-3G basis set and fitting the point charges that reproduce the quantum mechanical electrostatic potential. All *ab initio* calculations were carried out with Gaussian 92 (15). The other force field parameters for Dst and 2-ImD not present in the force field of Weiner *et al.* (16) were assigned by analogy with chemically related groups (these and the charges are available on request from the authors).

The setup for MD/FEP calculations involved neutralizing the phosphate charges with Na^+ ions placed at 3.6 Å from the phosphate bisector. The drug complex was solvated by placing three site TIP3P water molecules (17) around the solute such that the solvent molecules were placed up to 10 Å away along each of the rectangular coordinate axes and as close as 2.5 Å from any given solute atom. We chose not to neutralize the positive charge on the Dst or 2-ImD to avoid computational complexities associated with highly mobile positive and negative ions in solution. Since the positive charge on the drug would be present in both the initial and final states of our calculation, we assume that the effects due to the unneutralized charges will cancel out and the perturbation on the system will be minimal. For isolated drug calculations, we solvated 2-ImD molecules by placing solvent molecules up to 12 Å away and as close as 2.5 Å from the solute.

We carried out the following equilibration protocol for relieving bad contacts between the solute and solvent and allowing the reordering of water with minimal distortion to the starting structure, before calculation of the free energy differences. We restrained the starting structure with a harmonic restraint of 25 kcal per mol per residue and carried out potential energy minimization for about 1000 cycles followed by MD at 300 K for about 3 ps. This was done to allow the water molecules to reorient and make favorable contacts with the solute. Then we released the restraints on the solute in five steps by reducing the restraint by 5 kcal per mol per residue per step followed by 600 steps of conjugate gradient minimization. Thus in the fifth cycle the whole system was

Ligand 1:	N3···H21	2.34
Ligand 1:	D3···H21	1.68
Ligand 2:	N3···H21	3.58
Ligand 2:	D3···H21	2.40

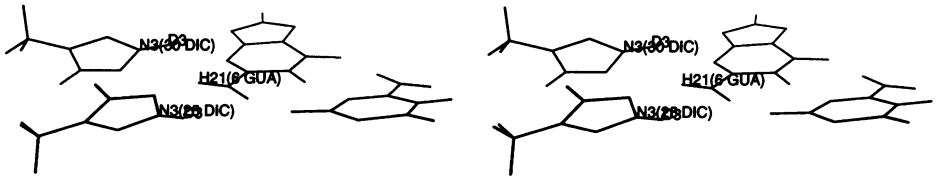

FIG. 3. The G^6-C^{17} base pair and the central imidazole rings of the two ligands in the starting structure are shown in stereo. The distances between N3 and D3 of the central imidazole rings of ligands 1 and 2 and H21 of G^6 are given above the stereoview. Ligand I (30 DIC) is on the top and ligand 2 (25 DIC) is on the bottom.

minimized without harmonic restraints. After this, MD was started by heating the system up to 300 K with a temperature coupling constant of 0.2 ps. The SHAKE routine (18) was applied to all the covalent bonds with hydrogens to a tolerance of 0.0005 Å. We have used a 1-fs time step during all our calculations. The temperature of the system was allowed to fluctuate around 300 K with a temperature coupling constant of 0.2 ps, and the pressure was allowed to fluctuate around 1 bar (1 bar = 100 kPa) with a pressure coupling constant of 0.6 ps. We implemented cubic periodic boundary conditions with a 10-Å cutoff for the electrostatic and Lennard–Jones potential evaluations for solute–water and water–water interactions. All solute–solute interactions (including sodium ions) were evaluated under the assumption that this will avoid cutoff artifacts on electrostatic interactions of sodium ions with phosphates without serious consequences on the solute–solvent interactions. The evaluation of all solute–solute interactions in our periodic simulations is possible since we do not generate images for the solute atoms.

The FEP calculations on these equilibrated systems was carried out by using the thermodynamic windows method as described elsewhere (19). The free energy perturbation was carried out by mutating 2-ImD → Dst in the complexed and in the isolated forms, in both the forward ($\lambda = 1 \rightarrow 0$) and backward ($\lambda = 0 \rightarrow 1$) directions.

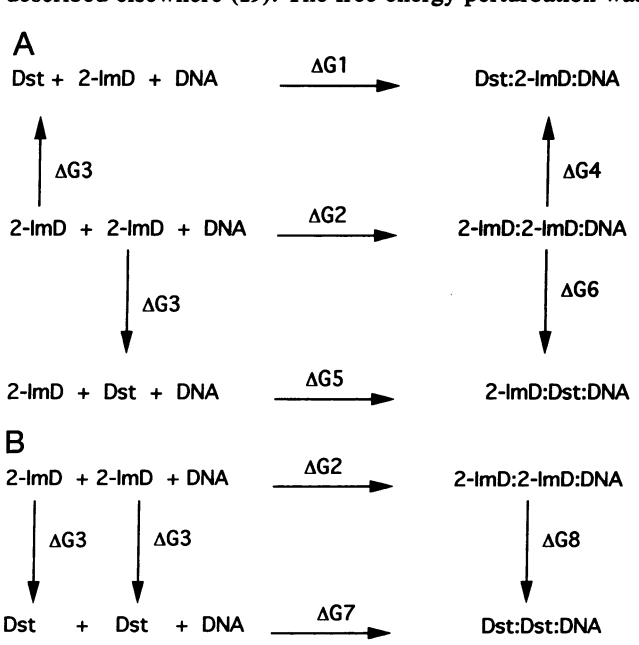

FIG. 4. Thermodynamic cycles for the calculation of the free energy difference of Dst and 2-ImD to d(CGCAA GTTGGC)·d(GCC-AACTTGCG). (*A*) The scheme for mutating 2-ImD at each of the two sites. (*B*) The scheme for mutating both the drug molecules at both the sites.

All our calculations were carried out using AMBER 4.0 (20). The potential energy function that describes the interactions between atoms is

$$V_{\text{total}} = \sum_{\text{bonds}} K_r(r - r_{\text{eq}})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{\text{eq}})^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j} \left\{ \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} \right] + \frac{q_i q_j}{\varepsilon R_{ij}} \right\} + \sum_{\text{H bonds}} \left[\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}} \right] + \sum_{\text{restraints}} K_r(r - r_0)^2. \quad [4]$$

In the FEP calculations, we have used 201 windows with a $\Delta\lambda$ of 0.005 per window. At each window we carried out 0.125 ps of equilibration and 0.125 ps of data collection. Thus the total length of the simulation in each direction was 50.250 ps (201 × 0.250 ps). All the calculations were carried out on the Cray YMP-C90 at the National Science Foundation's Pittsburgh Supercomputer Center.

To avoid the unraveling of the ends, we imposed a weak harmonic restraint of 5.0 kcal per mol per residue on the terminal base pairs. To achieve structural reversibility in the FEP calculations we imposed a harmonic restraint of 5.0 kcal per mol per residue on the terminal atoms of the drug molecules. This was done to restrain the drug from sliding away from its initial structure but nonetheless allows enough flexibility around the site of mutation on the drug. In the case of the mutation of ligand 1 in the complex, we found that the free energy difference for the forward mutation was very similar to that computed without restraints, indicating that the restraints have not severely affected the results.

RESULTS AND DISCUSSION

The distances between N3 and D3 (the atomic label for the dummy atom in 2-ImD) in 2-ImD or C3 and H3 in Dst of ligands 1 and 2 and the amino proton of G^6 are given in Table 1. There is a weak hydrogen bond between N3 of ligand 1 and the amino proton (H21) of G^6 (2.34 Å) in the NMR structure, whereas for ligand 2 the corresponding distance is 3.58 Å (see Fig. 3). These distances are shorter in the equilibrated structure (1.84 and 3.16 Å, respectively), indicative of stronger hydrogen bonds between DNA and the ligands. In the

Table 1. Distances in angstroms between N3/C3 and H3/D3 atoms of the central imidazole/pyrrole ring and the second proton (H21) on N2 of G^6

Conformer/ mutation site	Site I		Site II	
	N3/C3···H21	D3*/H3···H21	N3/C3···H21	D3*/H3···H21
NMR	2.34	1.68	3.58	2.40
Eqb†	1.84	1.21	3.16	1.93
Ligand 1‡	2.58	2.05	2.60	1.96
Ligand 2§	2.08	1.23	3.10	2.32
Ligands 1 and 2¶	2.45	2.04	3.31	2.57

*D3 is the atomic label for the dummy atom in 2-ImD.
†Structure after 10 ps of equilibration.
‡Distances after 50 ps of FEP of 2-ImD·2-ImD·DNA → Dst·2-ImD·DNA mutation. Site I, C3 and H3; site II, N3 and D3.
§Distances after 50 ps of FEP of 2-ImD·2-ImD·DNA → 2-ImD·Dst·DNA mutation. Site I, N3 and D3; site II, C3 and H3.
¶Distances after 50 ps of FEP of 2-ImD·2-ImD·DNA → Dst·Dst·DNA mutation. Sites I and II, C3 and H3.

2-ImD·2-ImD·DNA → Dst·2-ImD·DNA (site I) perturbation, the mutation of N3-D3 → C3-H3 leads to an increased separation of C3 of ligand 1 (2.58 Å) from the amino proton compared to the equilibrated structure (1.84 Å). This occurs in order to accommodate the hydrogen atom in the vicinity of the amino proton of G^6, which most likely leads to the observed rearrangement of the G^6·C^{17} base pair. In 2-ImD·2-ImD·DNA → 2-ImD·Dst·DNA (site II) perturbation, the distances between the C3 and H3 atoms of ligand 2 and the amino proton of G^6 are similar to those in the NMR structure, which indicates that the replacement of 2-ImD with Dst at site II leads to no observable disruption at the binding site. In the simulation with simultaneous perturbations at sites I and II, the structural changes are the cumulative effects of the independent mutations at sites I and II as described above. These changes are directly reflected in the changes in free energies for these simulations.

The free energy changes for the 2-ImD → Dst mutation in gas and solution phases are given in Table 2. The free energy of solvation ($\Delta\Delta G_{sol} = \Delta G_3 - \Delta G_{gas}$) is 1.82 kcal/mol. The desolvation free energy of 2-ImD is measurably greater than that of Dst. This is to be expected considering the fact that 2-ImD has one hydrogen bond donor (N3) more than Dst. The difference in desolvation energies of these two ligands should be reflected in their relative binding affinities to different binding sites on DNA.

The total free energy changes and the relative binding affinities of each of the three mutations are given in Table 2. In the case of the 2-ImD·2-ImD·DNA → Dst·2-ImD·DNA mutation, we observe that Dst has a less favorable free energy of association than 2-ImD at site 1 by 0.65 kcal/mol (Table 2). The lower affinity of Dst at this site is most likely due to the loss of the specific hydrogen bond interaction between the N3 and the amino proton of G^6 and the steric repulsion between H3 and amino proton of G^6. However, in the case of the 2-ImD·2-ImD·DNA → 2-ImD·Dst·DNA mutation, Dst has more favorable free energy of association than 2-ImD by 2.88 kcal/mol (Table 2). The stronger binding of Dst at site II is probably due to the introduction of favorable van der Waals interactions between the pyrrole and the minor groove groups on DNA since 2-ImD at site II has more space between the N3 and H21 of G^6 (see Table 1) than Dst. The more favorable desolvation energy of Dst will also contribute. Finally, in the third case, where 2-ImD·2-ImD·DNA → Dst·Dst·DNA mutation is carried out, the two Dst molecules bind with lower free energy than two 2-ImD molecules by 1.79 kcal/mol (Table 2). This, not surprisingly, is approximately the sum of the free energies for the two individual mutations. The first question addressed by the theory was the relative stability of 2-ImD·Dst·DNA versus Dst·2-ImD·DNA. Only the former is observed in equilibrium with Dst·Dst·DNA when both ligands are present. This is consistent with our calculated free energy difference between 2-ImD·Dst·DNA and Dst·2-ImD·DNA ($\Delta G_4 - \Delta G_6$) of 3.5 kcal/mol.

Experimental values for the relative binding affinities of Dst and 2-ImD have been determined using NMR spectroscopy, UV detected melting, and calorimetry. When the ligands Dst and 2-ImD are present in a 2-fold excess with the AAGTT oligomer, resonances from both the 2-ImD·Dst·DNA complex and the Dst·Dst·DNA complex can be seen in a ratio of about 15:1. This population difference corresponds to a ΔG of 1.6 kcal/mol difference (see Table 2) in binding free energy for 2-ImD relative to Dst at site I. This correlates reasonably well with the calculated value of 0.7 kcal/mol (Table 2). The absence of resonances from the Dst·2-ImD·DNA and 2-ImD·2-ImD·DNA complexes indicates that these complexes are significantly higher in free energy, as also found by our calculations.

A comparison of the 2:1 complexes of Dst and 2-ImD can be made by using data from UV melting and titration calorimetry (D. Rentzeperis, L. Marky, B. Geierstanger, and D.E.W., unpublished results). In this case the accuracy is

Table 2. Free energy in kcal/mol of the 2-ImD → Dst mutations in the isolated drug and the DNA complexes

Isolated drug			DNA–drug complex*			Free energy differences					
						Compared†			Experimental‡		
ΔG_{gas}	ΔG_3 (solution)	$\Delta\Delta G_{sol}$ ($\Delta G_3 - \Delta G_{gas}$)	ΔG_4 (site I)	ΔG_6 (site II)	ΔG_8 (sites I and II)	$\Delta\Delta G_1$	$\Delta\Delta G_2$	$\Delta\Delta G_3$	$\Delta\Delta G_1$	$\Delta\Delta G_2$	$\Delta\Delta G_3$
5.23 ± 0.01	7.05 ± 0.04	1.82	7.70 ± 0.24	4.17 ± 0.23	12.31 ± 0.12	0.65	−2.88	−1.79	1.6§	NA¶	−2.3‖

The free energy values presented here are an average of the forward and backward simulations (50 ps in each direction) with the error estimate (average free energy minus forward or backward value) for each mutation.
*ΔG_4 = 2-ImD·2-ImD·DNA → Dst·2-Imd·DNA, ΔG_6 = 2-ImD·2-ImD·DNA → 2-ImD·Dst·DNA, ΔG_8 = 2-ImD·2-ImD·DNA → Dst·Dst·DNA.
†$\Delta\Delta G_1 = (\Delta G_4 - \Delta G_3)$, $\Delta\Delta G_2 = (\Delta G_6 - \Delta G_3)$, $\Delta\Delta G_3 = (\Delta G_8 - 2\Delta G_3)$.
‡$\Delta\Delta G_1 = (\Delta G_1 - \Delta G_2)$, $\Delta\Delta G_2 = (\Delta G_5 - \Delta G_2)$, $\Delta\Delta G_3 = (\Delta G_7 - 2\Delta G_2)$ (see Eqs. 1–3).
§Calculated from the relative populations of Dst·Dst·DNA (1) and 2-ImD·Dst·DNA (15). $\Delta G = -RT \ln(1/15) = 1.6$ kcal/mol.
¶Not available. This value cannot be computed from the experimental data since the Dst·2-ImD·DNA complex is not detected in solution.
‖Calculated from the relative populations of Dst·Dst·DNA (50) and 2-ImD·2-ImD·DNA (1). $\Delta G = -RT \ln(50/1) = -2.3$ kcal/mol.

somewhat lower, but the affinity of 2-ImD is lower than that of Dst by about a factor of 50. This corresponds to a free energy difference ΔG of 2.3 kcal/mol (see Table 2), which is in good agreement with the calculated value of 1.8 kcal/mol (Table 2).

The calculations presented here reproduced the correct rank order of stability of the different complexes, including that of Dst at the AAGTT site. The traditional ideas about Dst binding would have considered this to be an unfavorable site, since the amino group of G^6 protrudes into the minor groove. However, in the 2:1 binding motif, this can easily be accommodated, yielding a complex with fairly high affinity. The affinity can be increased, again as reproduced well in the calculations, through the presence of specific ligand–DNA hydrogen bonds. The calculations also indicate that the desolvation of 2-ImD is a dominant unfavorable factor in its poor binding at site II, since there is no compensating hydrogen bond formed to the DNA.

The structure of the complex during the course of the FEP calculations remained fairly stable and in the vicinity of the NMR structure. This is reflected in the total rms deviations for all the atoms of the structure after 50 ps from the NMR structure (2.51 Å). The major contribution to this deviation is from the structural changes in the base pair separations while very small changes take place in the DNA–drug interactions. All the hydrogen bonds between the peptide groups on the drug and the minor groove groups (O2 of T^7, T^8, C^{17}, T^{18}, and T^{19} and N3 of G^6 and A^{16}) are preserved in the equilibrated structure and also in all the perturbed structures. Thus we note that the change in hydrogen-bonding interactions between the drug and the DNA takes place only between H21, the amino proton, of G^6 and N3 of the ligand. The rms deviations of the structures after 50 ps of FEP at site I, site II, and sites I and II from the NMR structure are 2.53, 2.67, and 2.51 Å, respectively.

CONCLUSIONS

We present here the calculations of the relative binding affinities of two different ligands to DNA. Our calculations show the following order of binding affinities of Dst and 2-ImD to d(CGCAAGTTGGC)·d(GCCAACTTGCG): 2-ImD·Dst·DNA > Dst·Dst·DNA > 2-ImD·2-ImD·DNA > Dst·2-ImD·DNA, a hierarchy qualitatively consistent with the experimental data; we also find the relative free energies of association of 2-ImD·Dst·DNA, Dst·Dst·DNA, and 2-ImD·2-ImD·DNA in semiquantitative (within 1 kcal/mol) agreement with experiment. The particular order of binding strength is due to the fact that 2-ImD has greater affinity for site I than Dst and Dst has greater affinity for site II than 2-ImD, with desolvation effects modulating these differences. All our calculations were carried out without the knowledge of the experimental thermodynamic data and structural data other than the use of the 2-ImD·2-ImD·DNA structure from Dwyer *et al.* (12) as a starting geometry. Thus, it was encouraging that the calculated free energies agreed reasonably well with experiment. Hence the approach presented here can be used to predict relative binding affinities of two or more similar ligands that bind to a specific site on DNA.

D.E.W. thanks Bernard H. Geierstanger for his useful comments and suggestions on this manuscript. We acknowledge the support for the current research from the National Institutes of Health (Grant CA 25644 to P.A.K. and Grant GM 43129 to D.E.W.) and the use of the University of California (San Francisco) Computer Graphics Laboratory (supported by Grant RR-1081; R. Langridge, principal investigator). We also acknowledge the computing support provided by the National Science Foundation's Pittsburgh Supercomputer Center through the Grand Challenge Award MCA93S017P to P.A.K.

1. Bash, P. A., Singh, U. C., Brown, F. K., Langridge, R. & Kollman, P. (1987) *Science* **235**, 574–576.
2. Merz, K. M. & Kollman, P. A. (1989) *J. Am. Chem. Soc.* **111**, 5649–5658.
3. Ferguson, D. M., Radmer, R. J. & Kollman, P. A. (1991) *J. Med. Chem.* **34**, 2654–2659.
4. Creighton, S., Rudolph, B., Lybrand, T., Singh, U. C., Shafer, R., Brown, S., Kollman, P., Case, D. & Andrea, T. (1989) *J. Biomol. Struct. Dyn.* **6**, 929–969.
5. Feuerstein, B. G., Pattabiraman, N. & Marton, L. J. (1989) *Nucleic Acids Res.* **17**, 6883–6892.
6. Boehncke, K., Nonella, M., Schulten, K. & Wang, A. H.-J. (1991) *Biochemistry* **30**, 5465–5475.
7. Cieplak, P., Rao, S. N., Grootenhuis, P. D. J. & Kollman, P. A. (1990) *Biopolymers* **29**, 717–727.
8. Gago, F. & Richards, W. G. (1990) *Mol. Pharmacol.* **37**, 341–346.
9. Hard, T. & Nilsson, L. (1991) *Nucleosides Nucleotides* **10**, 701–709.
10. Lavery, R. & Pullman, B. (1981) *Int. J. Quantum Chem.* **20**, 259–272.
11. Coll, M., Frederick, C. A., Wang, A. H.-J. & Rich, A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8385–8389.
12. Dwyer, T. J., Geierstanger, B. H., Bathini, Y., Lown, J. W. & Wemmer, D. E. (1992) *J. Am. Chem. Soc.* **114**, 5911–5919.
13. Geierstanger, B. H., Dwyer, T. J., Bathini, Y., Lown, J. W. & Wemmer, D. E. (1993) *J. Am. Chem. Soc.* **115**, 4474–4482.
14. Hehre, W. J., Stewart, R. F. & Pople, J. A. (1969) *J. Chem. Phys.* **51**, 2657–2664.
15. Frisch, M. J., Trucks, G. W., Head-Gordon, M., Gill, P. M. W., Wong, M. W., Foresman, J. B., Johnson, B. G., Schlegel, H. B., Robb, M. A., Replogle, E. S., Gomperts, R., Andres, J. L., Raghavachari, K., Binkley, J. S., Gonzalez, C., Martin, R. L., Fox, D. J., Defrees, D. J., Baker, J., Stewart, J. J. P. & Pople, J. A. (1992) *Gaussian 92, Revision C.4* (Gaussian, Inc., Pittsburgh).
16. Weiner, S. J., Kollman, P. A., Nguyen, D. T. & Case, D. A. (1986) *J. Comp. Chem.* **7**, 230–252.
17. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983) *J. Chem. Phys.* **79**, 926–935.
18. van Gunsteren, W. F. & Berendsen, H. J. C. (1977) *Mol. Phys.* **34**, 1311–1327.
19. Singh, U. C., Brown, F. K., Bash, P. & Kollman, P. A. (1987) *J. Am. Chem. Soc.* **109**, 1607–1609.
20. Pearlman, D. A., Case, D. A., Caldwell, J. C., Seibel, G. L., Singh, U. C., Weiner, P. & Kollman, P. A. (1991) *AMBER (4.0)* (Univ. of California, San Francisco).