0026-895X/90/030341-06$02.00/0

MOLECULAR PHARMACOLOGY, 37:341-346

ACCELERATED COMMUNICATION

Netropsin Binding to Poly[d(IC)]·poly[IC)] and Poly[d(GC]·poly[d(GC)]: A Computer Simulation

FEDERICO GAGO[1] and W. GRAHAM RICHARDS

Oxford Centre for Molecular Sciences and Physical Chemistry Department, Oxford University, Oxford OX1 3QZ, UK

Received September 14, 1989; Accepted December 15, 1989

SUMMARY

The thermodynamic cycle perturbation approach has been used to calculate the difference in the free energy of binding of netropsin to two different DNA molecules. In the computer simulations, all the inosine residues have been gradually 'mutated' into guanosine in a DNA dodecamer and in a complex of the same dodecamer with netropsin. The difference in binding free energy of about 4.3 kcal mol^{-1} agrees well with the experimentally determined value of 4.0 kcal mol^{-1}. One structural determinant of the specificity seems to be the width of the minor groove in the two complexes.

Netropsin (Fig. 1) has become one of the prototypes of drugs that bind in the minor groove of DNA (1) and a very useful model compound to investigate protein-DNA and drug-DNA recognition patterns. The strength of binding between netropsin and DNA molecules of different sequences is determined by the intermolecular interaction energies that are ultimately responsible for the specificity of the interaction. Different experimental (2–6) and theoretical (7–9) procedures have shown a preferential binding of netropsin to runs of A-T base pairs over G-C ones in double-stranded DNA. However, the relative importance of each contributing factor to the binding energy is still a matter of controversy (1–9). Although the calculated interaction energy can be broken down into electrostatic, dispersion-repulsion, and hydrogen bonding contributions, the structural and damping effects of water are usually neglected, which leads to the overestimation of the electrostatic term (7–9).

In pharmacology, the major problem in calculating theoretical values of binding energies of drugs is that they are differences in energy between drug molecules in solution, interacting with water, and drug molecules in a drug-receptor complex, interacting with each other (Fig. 2). Computer simulations are not as yet capable of calculating directly the difference $\Delta G_1 - \Delta G_2$ for large molecules. The difficulty in evaluating $\Delta G_1 - \Delta G_2$ is that each binding process must be simulated slowly enough to ensure thermodynamic equilibrium, which involves the generation of many representative configurations of the system (10), placing this sort of simulation beyond the scope of even the powerful supercomputers. However, the nonphysical process of perturbing one of the reactants into another, both in the free state and in the bound state, yields a relative free energy change, $\Delta G_3 - \Delta G_4$, that can be evaluated more easily and related to $\Delta G_1 - \Delta G_2$ (Fig. 2). Successful applications of this approach to a variety of molecular recognition and activity problems can be found in the literature (see Refs. 11–13 for reviews).

In this work, we have aimed at obtaining a quantitatively correct value for the difference in binding free energy of netropsin to two different alternating DNA sequences. The encouragement provided by the results obtained in the calculation of redox potentials (14), partition coefficients (15), and binding energies of carbonic anhydrase inhibitors (16) has led us to extend this method to a particularly interesting target for antitumour drug action, DNA. In this case, the perturbation consists of the hypothetical transformation of ICIC into GCGC in solution, both in the free state and with netropsin bound to it, as shown in Fig. 2. Experimental data available for netropsin binding to different polynucleotides (6) show that the binding affinities for ICIC and ATAT are similar and higher than that for GCGC. Therefore, it is the exocyclic amino group of guanine present in the minor groove that is decreasing the binding free energy. It is this group that forms the core of the simulation presented in this paper.

Methods

Molecular model. We have reported on the refined model of an ATAT-netropsin complex (9). A similar model considering all the

F. G. is grateful to the Consejo Superior de Investigaciones Cientificas (Spain) for a postdoctoral grant. The project is in part pursuant to a contract with the National Foundation for Cancer Research.

[1] Permanent address: Departamento de Farmacología, Universidad de Alcalá de Henares, Madrid (Spain).

ABBREVIATIONS: ICIC, poly[d(IC)]·poly[d(IC)]; GCGC, poly[d(GC)]·poly[d(GC)]; ATAT, poly[d(AT)]·poly[d(AT)].

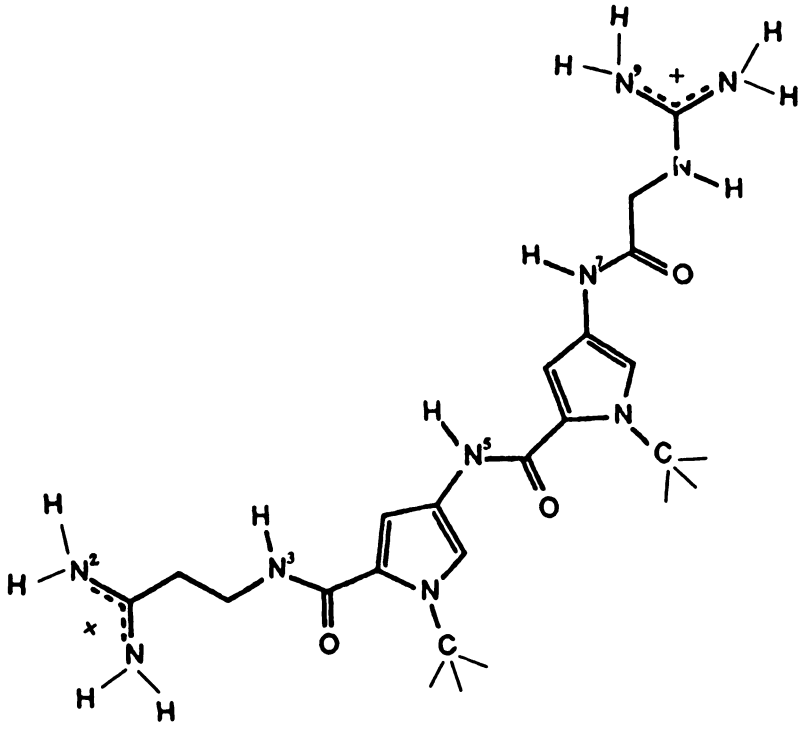

Fig. 1. Molecular structure of netropsin. The atom positions relevant to the text have been numbered.

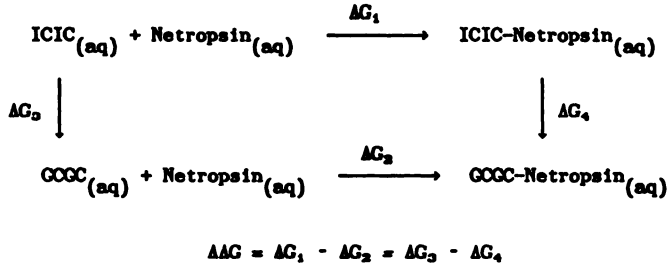

Fig. 2. Thermodynamic cycle for the association of netropsin with ICIC and GCGC in aqueous solution.

atoms explicitly was built for ICIC by replacing the respective purine and pyrimidine bases, followed by energy minimization techniques. The AMBER suite of programs (17) was used throughout. The resulting structure was immersed into a large water bath constructed of repeated cubes of transferable intermolecular potential water molecules. Each cube was a snapshot from a Monte Carlo simulation of liquid water, where water molecules are represented by a rigid three-point charge model (18). In order to achieve electroneutrality, an Na^+ counterion had been previously placed in the plane of each phosphate group at a distance of less than 3 Å from the phosphorus atom. After all the water molecules located less than 2.4 Å away from any solute atom were removed, the box consisted of 1570 water molecules and had a total volume of 50.9 × 35.8 × 34.1 nm^3, with the minimum distance between any solute atom and any of the edges being 0.5 nm.

The length of 12 base pairs was chosen so as to avoid end effects in the central region where the drug was bound and where the perturbation was going to be more relevant to the calculation of the free energy change. The residue numbers correspond to their sequence numbers along a given strand in the 5′→3′ direction.

For the free oligonucleotide, the same starting geometry as that in the complex was used but the drug was replaced by five water molecules, resembling a short 'spine' of hydration in the minor groove; it is known from crystallographic studies (4) that netropsin is able to displace some water molecules from the minor groove when it binds to the central region of the dodecamer CGCGAATT(Br)CGCG. After energy minimization and immersion of the resulting structure in a water bath as described above, the total number of water molecules considered was 1501 and the total volume of the box 50.9 × 35.6 × 34.2 nm^3. The dynamics and perturbation calculations were carried out under the same conditions as for the complex.

Molecular mechanics and dynamics. The resulting configurations were relaxed by minimizing the energy of the solvent molecules and, subsequently, by performing a conjugate gradient energy minimization of the whole system until the root mean square gradient was less than 0.1 kcal mol^{-1} $Å^{-1}$.

The AMBER force field does not include atomic partial charges for inosine and netropsin, but these were derived in a consistent manner (19) by fitting the atomic charges to the *ab initio* molecular electrostatic potentials, using the QUEST program (20). The atoms of netropsin were assigned the van der Waals and hydrogen bonding parameters of corresponding atom types, as described previously (8, 9); the additional parameters necessary were obtained in accordance with the interpolation method presented by Weiner *et al.* (21).

After energy minimization, a molecular dynamics run of 16 psec at 300°K and 1 atm (isothermal isobaric ensemble) was performed in order to equilibrate the system. The step length was 0.0005 psec during the first 4 psec and 0.002 psec thereafter. The SHAKE algorithm (22) was used to constrain bond lengths to their equilibrium values and improve the computational efficiency.

At this point the perturbation of the inosine C-2 hydrogen into an amino group was started with the aid of two dummy atoms (Fig. 3), which are characterized by having a point charge of zero and the nonbonded parameters set to zero. A value of 0.5 Å was assigned to the equilibrium bond distance H-dummy atom so that the appearance of the NH_2 group would disturb the structure more slowly. During the perturbation, the parameters for the dummy atoms were linearly increased so that at the end of the simulation they attained the values corresponding to H-bonding NH_2 hydrogens. In this respect, it is important to note that the dummy atoms must be included in the hydrogen-bonding list of the parameter data base, because failing to do this would result in the utilization by the program of the 6–12 van der Waals potential instead of the 10–12 hydrogen-bonding potential (21) and in over-long hydrogen-bonding distances. In all our simulations, periodic boundary conditions, a unit dielectric constant, and a cut-off for nonbonded interactions of 8 Å were used, and the list of nonbonded

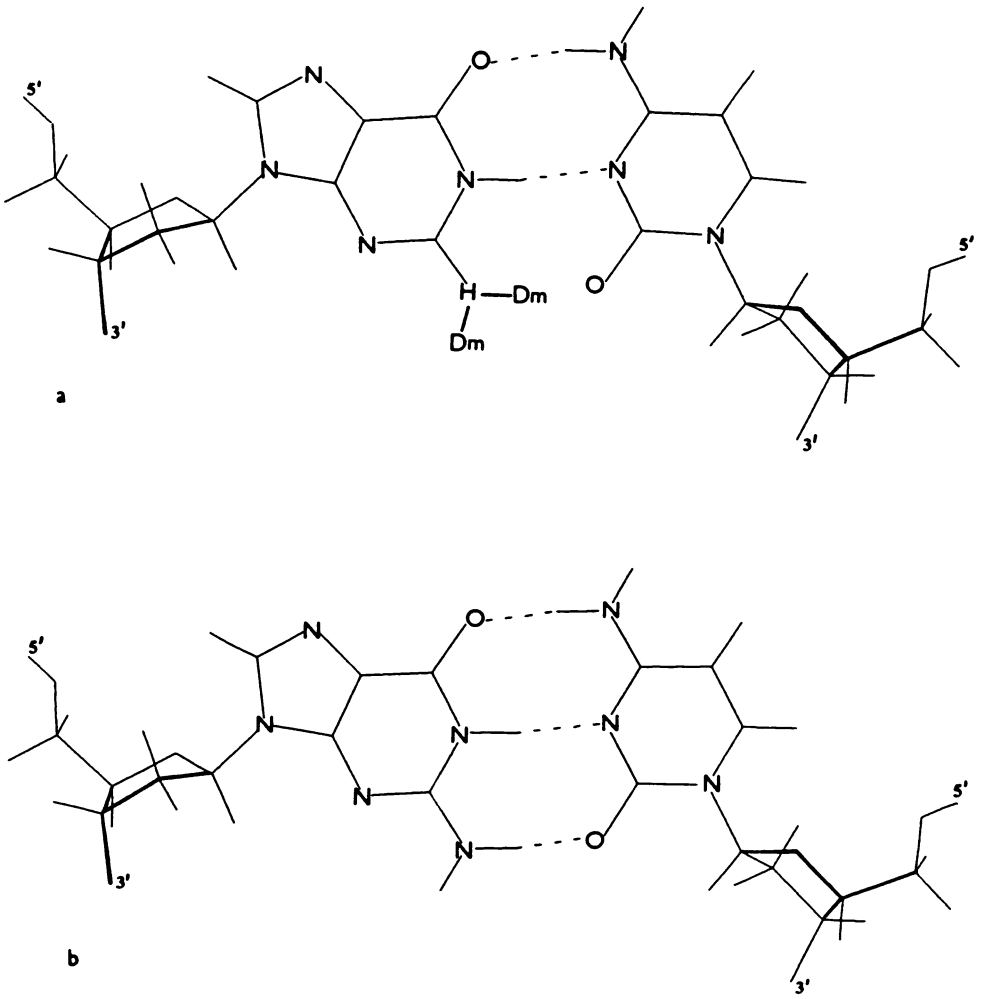

Fig. 3. Schematic representation of the perturbation of an I-C base pair (a) into a G-C base pair (b). *Dm*, dummy atom.

pairs was updated every 100 steps. A longer cut-off, although desirable, would have greatly increased the computational cost, and it is important to realize that no change in overall charge is going to take place in the perturbation.

Thermodynamic cycle perturbation. The perturbation method followed was that of 'windowing' (12), as implemented in the GIBBS module of AMBER, starting with a coupling parameter (λ) of 1.0 and increments ($\Delta\lambda$) of −0.025. Data were collected in the forward and backward directions (Fig. 4). Each thermodynamic 'window' was equilibrated for 250 steps before data collection during another 250 steps. Thus, the complete mutation took place in 80 psec. The duration of the perturbation and the selection of the small $\Delta\lambda$ were determined by the magnitude of the free energy change (Table 1), so that the energy change per window contributed less than $2RT$ (12) and sufficient representative configurations of the perturbed system were explored.

Results and Discussion

The refined model of the complex of netropsin with ICIC in water shows the drug molecule firmly anchored in the minor groove through four hydrogen bonds to the base ring acceptor atoms (Fig. 5) and close van der Waals contacts with the walls of the groove. Only the amide nitrogen atom N-5 of netropsin is not involved in hydrogen bonding; the inosine N-3 atoms of Ino-7 and Ino-19 facing it are about 3.8 Å away. At the end of the first 16 psec of dynamics, the hydrogen bonds of the amidimium and guanidinium ends help maintain the molecule essentially in the same orientation. However, as a result of growth of the exocyclic amino group in the purine base of each pair in the dodecamer during the perturbation, the minor groove becomes less deep, more polar, and less narrow, in a process that is accompanied by the extrusion of part of the netropsin molecule into the solvent with a less favorable interaction energy.

In this respect, it is of interest to compare the suggestion put forward by Marky and Breslauer (6) with the results of our simulation. These authors envisioned netropsin in its complex with GCGC as being bowed out but with the two charged ends still interacting with the double helix. This proposal was based on the fact that the electrostatic contribution to the binding energy is of the same order of magnitude for the association of netropsin with either ATAT or GCGC. At the end of our simulation, the amidinium end of netropsin is bent towards the aqueous medium and is found to be interacting with one of the oxygens of the phosphate group bridging Cyt-6 and Ino-7, whereas the amide nitrogen next to it hydrogen bonds to the neighboring 0-1′ sugar atom of Ino-7. The hydrogen bonding pattern of Fig. 5 is lost; only the amide nitrogen atom N-7 is still hydrogen bonding to a base ring acceptor atom (0-2 of Cyt-18) in a bifurcated hydrogen bond that also involves the 0-1′ sugar atom of Ino-19. The bulk of the drug molecule still is within the groove, but this has widened up considerably (see Table 2).

In relation to this, it is also interesting to analyze the behav-

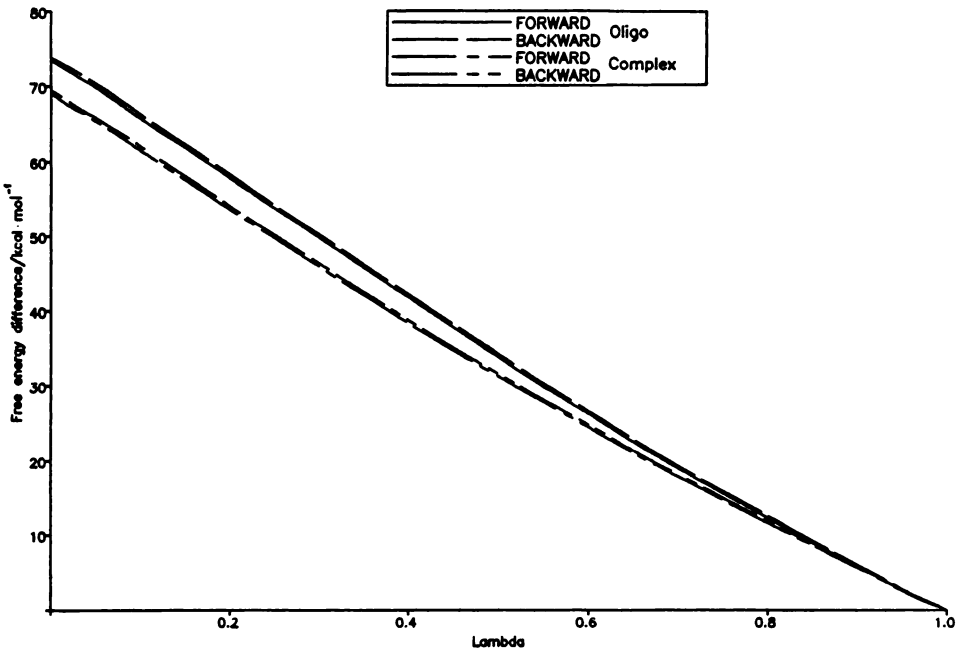

Fig. 4. Cumulative free energy change, in kcal mol^{-1}, during the perturbation (absolute value) of the free oligonucleotide (forward and backward) and the netropsin complex (forward and backward).

TABLE 1

Differences in free energy for the transformation of ICIC into GCGC in solution, both in the free state and in its complex with netropsin

	ΔG_3 (ICIC → GCGC)	ΔG_4 (ICIC-netropsin → GCGC-netropsin)
	kcal mol^{-1}	
Forward simulation	−73.45	−69.11
Backward simulation	73.85	69.49
Average	−73.65 ± 0.29	−69.30 ± 0.27

$\Delta\Delta G_{theory} = \Delta G_3 - \Delta G_4 = -4.35 \pm 0.28$
$\Delta\Delta G_{experiment} = \Delta G_1 - \Delta G_2 = -11.1 - (-7.1) = -4.0$

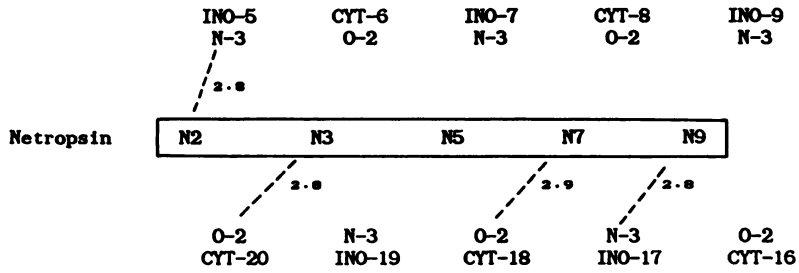

Fig. 5. Schematic representation of the hydrogen bonding between the electronegative atoms in the floor of the minor groove of ICIC and netropsin in solution. The interactions are indicated by *dashed lines*, and *numbers* indicate the length in Å of the hydrogen bonds.

ior of the five water molecules that were placed in the minor groove of the free oligonucleotide; only one remains in the groove after the perturbation, hydrogen bonding to Cyt-18 0-2. All the others have moved into the bulk solvent and the width of the groove has also increased (Table 2) (note that the larger standard deviation with respect to that during the equilibration is a consequence of the system still being perturbed and the sampling covering a longer time span). This probably reflects the dynamic nature of the hydration of the minor groove in alternating sequences and tempts us to speculate about differences in the strength of minor groove hydration in alternating and homopolymeric A-T runs. If the water molecules are more tightly bound in the minor groove of $d(A)_n$ tracts, the entropic contribution to the binding energy of drugs like netropsin, which are capable of displacing them, to this sequence should be greater. This hypothesis is supported by experimental measurements that show that netropsin binding to poly(dA)·poly(dT) is mostly entropy driven, whereas the magnitude of the binding entropy when ATAT is the acceptor site is much smaller, although the drug binds equally well to both sequences (6).

Fig. 6 shows the variation of potential energy during the dynamics simulation, first as the systems tend towards equilibrium and later during the perturbation. The energy change is mainly due to the electrostatic part of the potential; one new hydrogen bond between each base pair is now being formed and more interactions with the water molecules are also possible. Hydrogen bonds between base pairs in a double helix contribute 0.8 to 1.6 kcal mol^{-1}/hydrogen bond to the binding energy (23). On the other hand, the stacking interactions are expected to be similar for guanine and inosine, as shown experimentally for ribooligonucleotides (24).

The nucleotides at the ends of the double helix show a larger mobility than those in the central region, as found in previous simulations in water (25), highlighting the need for oligonucleotides longer than the span covered by the drug in the model

TABLE 2

Width of the minor groove in the free oligonucleotides and in the netropsin complexes: shortest distances across the groove between phosphate-phosphate (P-P distance minus 5.8 Å), and O1′ atoms

Only the distances in the central part of the dodecanucleotide spanning the stretch where the drug is bound in the complex are shown. The averages were taken at the end of the equilibration period (ICIC and ICIC-netropsin; 12–16 psec) and at the end of the perturbation (GCGC and GCGC-netropsin; 88–96 psec).

	Phosphate pair distance				Base pair distance			
	7-22	8-21	9-20	10-19	6-21	7-20	8-19	9-18
				Å				
Free oligonucleotide								
Start of dynamics	4.1	4.3	4.4	6.3	7.8	7.5	7.1	8.7
After equilibration (ICIC)	5.5 ± 0.2	7.9 ± 0.2	9.1 ± 0.2	7.9 ± 0.2	7.9 ± 0.2	8.9 ± 0.2	8.7 ± 0.4	10.4 ± 0.2
After perturbation (GCGC)	13.1 ± 0.5	13.0 ± 0.8	13.2 ± 1.0	11.7 ± 0.9	9.6 ± 0.4	8.9 ± 0.4	9.3 ± 0.4	11.1 ± 0.3
Netropsin complexes								
Start of dynamics	4.9	4.5	4.6	6.0	7.7	7.2	7.3	7.6
After equilibration (ICIC-netropsin)	6.0 ± 0.4	6.6 ± 0.6	7.5 ± 0.3	6.7 ± 0.4	8.9 ± 0.5	8.6 ± 0.5	8.8 ± 0.3	9.6 ± 0.3
After perturbation (GCGC-netropsin)	9.2 ± 0.4	7.0 ± 0.4	6.9 ± 0.3	11.6 ± 0.5	9.7 ± 0.4	8.9 ± 0.4	9.3 ± 0.4	11.1 ± 0.3

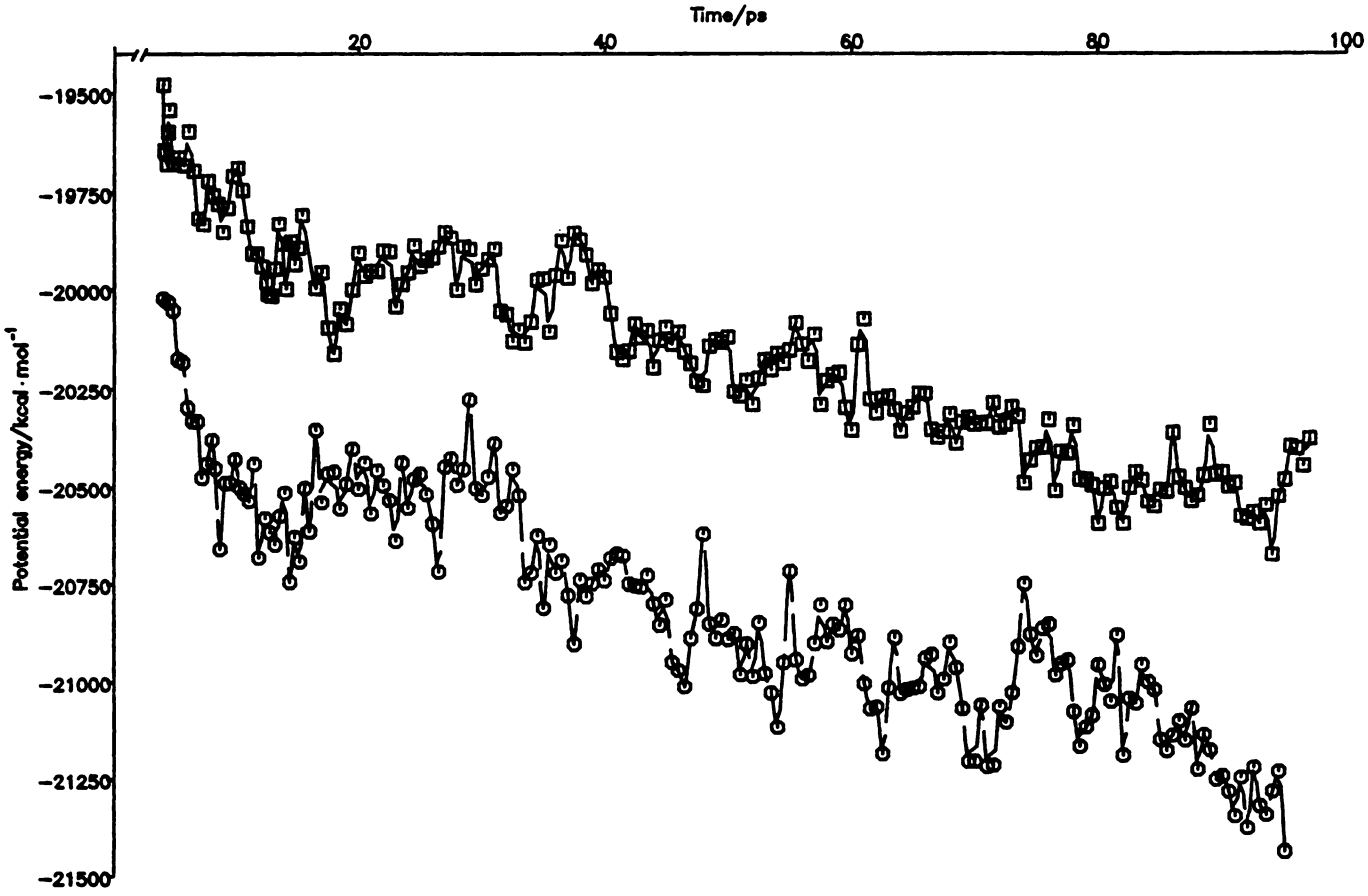

Fig. 6. Plot of the variation in potential energy, in kcal mol^{-1}, during the dynamics simulation. Each value corresponds to the average taken over 250 steps. □, Oligonucleotide; ○, netropsin complex.

system. Hydrogen bonds are more easily exchangeable with water at both the 5′ and 3′ ends, resulting in fraying effects.

Hydrogen bonds can be important determinants of specificity (23). In the case of nonintercalative DNA-binding drugs, recent results suggest an appreciable contribution of hydrogen bond interactions to the association reaction in the minor groove, based on the fact that this region is rather less polar than the bulk solvent, at least in the complex of Hoechst 33258 and ATAT (26). The minor groove of ICIC should present a very similar environment to that of ATAT, because it is equally accessible.

An interpretation of the difference in binding affinities could then be as follows. In the complexes of netropsin with ATAT or ICIC, the reduced minor groove width is stabilized by a number of electrostatic (including hydrogen bonding) and van der Waals interactions, making the interior of the groove a

rather low polarity medium. In the netropsin-GCGC complex, however, such close interactions cannot arise, because the minor groove width cannot be reduced to the same extent as in the ATAT- or ICIC-netropsin complexes, and that binding site probably exhibits a higher dielectric constant, so that exchange of netropsin-DNA hydrogen bonds for netropsin-water hydrogen bonds occurs more readily and the binding free energy is reduced.

As can be seen in Table 1, the calculated free energy change ($\Delta\Delta G$) is 4.35 ± 0.28 kcal mol^{-1} versus the experimental value of 4.0 kcal mol^{-1} (6). This good agreement is a consequence of the cancellation of errors that characterizes the free energy perturbation method (11–13) and is encouraging given the complexity of the calculations, applied for the first time to a nonintercalative DNA binding drug in solution. The standard deviation is an indication only of the statistical fluctuations inherent in the methodology. The statistical error could have been estimated from the standard deviation of results from independent simulations, and the determination of ΔG for the simulation running in the direction GCGC→ICIC would also have enabled us to ascertain the dependence of the free energy change upon the starting geometry. However, this was not carried out because of the serious computational demands of the calculation (126 hr of Convex-C2 central processing unit time for each complete run).

Conclusions

Despite the shortcomings of the relatively limited sampling of conformational space, we see the agreement with experiment more as a validation of the approach followed than as a coincidence. The averaged free energy change is of the same order of magnitude as that calculated for a protein-inhibitor complex (27). The calculations reproduce the observed preference of netropsin for binding to ICIC over GCGC in a quantitative manner and provide some reasons for the specificity of the association. They support our previous observation (9) that minor groove width can be an important determinant of complex stability in the interaction of nonintercalative agents with DNA.

More simulations will have to be done in the future to prove that this method is a useful tool for computer-aided antitumor drug design. Considerably simpler cases will be the perturbation of a drug in the acceptor site of a DNA molecule or the mutation of just one base or a base pair in a given sequence.

Acknowledgments

We thank C. A. Reynolds and P. M. King for many useful discussions.

References

1. Zimmer, C., and U. Wahnert. Nonintercalating DNA-binding ligands: specificity of the interaction and their use as tools in biophysical, biochemical and biological investigations of the genetic material. *Prog. Biophys. Mol. Biol.* **47**:31–112 (1986).
2. Patel, D. J. Antibiotic-DNA interactions: intermolecular nuclear Overhauser effects in the netropsin-d(CGCGAATTCGCG) complex in solution. *Proc. Natl. Acad. Sci. USA* **79**:6424–6428 (1982).
3. Van Dycke, M. W., R. P. Hertzberg, and P. B. Dervan. Map of distamycin, netropsin, and actinomycin binding sites on heterogeneous DNA: DNA cleavage-inhibition patterns with methidiumpropyl-EDTA·Fe(II). *Proc. Natl. Acad. Sci. USA* **79**:5470–5474 (1982).
4. Kopka, M. L., C. Yoon, D. Goodsell, P. Pjura, and R. E. Dickerson. The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc. Natl. Acad. Sci. USA* **82**:1376–1380 (1985).
5. Coll, M., J. Aymami, G. A. van der Marel, J. H. van Boom, A. Rich, and A. H.-J. Wang. Molecular structure of the netropsin-d(CGCGATATCGCG) complex: DNA conformation in an alternating sequence. *Biochemistry* **28**:310–320 (1989).
6. Marky, L. A., and K. J. Breslauer. Origins of netropsin binding affinity and specificity: correlations of thermodynamic and structural data. *Proc. Natl. Acad. Sci. USA* **84**:4359–4363 (1987).
7. Lavery, R., K. Zakrzewska, and B. Pullman. Binding of non-intercalating antibiotics to B-DNA: a theoretical study taking into account nucleic acid flexibility. *J. Biomol. Struct. Dyn.* **3**:1155–1170 (1986).
8. Caldwell, J., and P. A. Kollman. A molecular mechanical study of netropsin-DNA interactions. *Biopolymers* **25**:249–266 (1986).
9. Gago, C., C. A. Reynolds, and W. G Richards. The binding of nonintercalative drugs to alternating DNA sequences. *Mol. Pharmacol.* **35**:232–241 (1989).
10. McCammon, J. A. Computer-aided molecular design. *Science (Wash. D. C.)* **238**:486–491 (1987).
11. Wong, C. F., and J. A. McCammon. Dynamics and design of enzymes and inhibitors. *J. Am. Chem. Soc.* **108**:3830–3832 (1986).
12. Bash, P. A., U. C. Singh, R. Langridge, and P. A. Kollman. Free energy calculations by computer simulation. *Science (Wash. D. C.)* **236**:564–568 (1987).
13. Richards, W. G., P. M. King, and C. A. Reynolds. Solvation effects. *Protein. Eng.* **2**:319–327 (1989).
14. Reynolds, C. A., P. M. King, and W. G. Richards. Computed redox potentials and the design of bioreductive agents. *Nature (Lond.)* **334**:80–82 (1988).
15. Essex, J. W., C. A. Reynolds, and W. G. Richards. Relative partition coefficients from partition functions: a theoretical approach to drug transport. *J. Chem. Soc. Chem. Commun.* 1152–1154 (1989).
16. Menziani, M. C., C. A. Reynolds, and W. G. Richards. Rational drug design: binding free energy differences of carbonic anhydrase inhibitors. *J. Chem. Soc. Chem. Commun.* 853–855 (1989).
17. Singh, U. C., S. J. Weiner, J. W. Caldwell, and P. A. Kollman. *AMBER (UCSF) Version 3.1.* Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco (1988).
18. Jorgensen, W. L., J. Chandrasekhar, J. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**:926–935 (1983).
19. Singh, U. C., and P. A. Kollman. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **5**:129–145 (1984).
20. Singh, U. C., and P. A. Kollman. *QUEST (UCSF) Version 1.0.* Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco (1986).
21. Weiner, S. J., P. A. Kollman, D. T. Nguyen, and D. A. Case. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* **7**:230–252 (1986).
22. Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **23**:327–341 (1977).
23. Fehrst, A. R. The hydrogen bond in molecular recognition. *Trends Biochem. Sci.* **12**:301–304 (1987).
24. Turner, D. H., N. Sugimoto, R. Kierzek, and S. D. Dreiker. Free energy increments for hydrogen bonds in nucleic acid base pairs. *J. Am. Chem. Soc.* **109**:3783–3785 (1987).
25. van Gunsteren, W. F., H. J. C. Berendsen, R. G. Geurtsen, and H. R. J. Zwinderman. A molecular dynamics computer simulation of an eight-base-pair DNA fragment in aqueous solution: comparison with experimental two-dimensional NMR data. *Ann. N. Y. Acad. Sci.* **482**:287–303 (1986).
26. Jin, R., and K. J. Breslauer. Characterization of the minor groove environment in a drug-DNA complex: bisbenzimide bound to poly[d(AT)]·poly[d(AT)] duplex. *Proc. Natl. Acad. Sci. USA* **85**:8939–8942 (1988).
27. Bash, P. A., U. C. Singh, F. K. Brown, R. Langridge, and P. A. Kollman. Calculation of the relative change in binding free energy of a protein-inhibitor complex. *Science (Wash. D. C.)* **235**:574–576 (1987).

Send reprint requests to: W. Graham Richards, Physical Chemistry Laboratory, South Parks Road, Oxford OX1 3QZ, UK.