(c) H. M. Berman, S. S. C. Chu, and G. A. Jeffrey, *Science*, **157**, 1576 (1967).
(9) C. Romers, C. Altona, N. R. Buys, and E. Havinga, *Top. Stereochem.*, **4**, 39 (1969).
(10) D. Horton and W. N. Turner, *J. Org. Chem.*, **30**, 3387 (1965).
(11) C. V. Holland, D. Horton, and J. S. Jewell, *J. Org. Chem.*, **32**, 1818 (1967).
(12) S. Wolfe, M.-H. Whangbo, and D. J. Mitchell, *Carbohydr. Res.*, in press; private communication.
(13) G. A. Jeffrey, J. A. Pople, and L. Radom, *Carbohydr. Res.*, **25**, 117 (1972).
(14) G. A. Jeffrey, J. A. Pople, and L. Radom, *Carbohydr. Res.*, **38**, 81 (1974).
(15) G. A. Jeffrey, J. A. Pople, J. S. Binkley, and S. Vishveshwara, *J. Am. Chem. Soc.*, **100**, 373 (1978).
(16) R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.*, **54**, 724 (1971).
(17) C. C. Costain, *J. Chem. Phys.*, **29**, 864 (1958).
(18) R. L. Hilderbrandt and J. D. Wieser, *J. Chem. Phys.*, **55**, 4648 (1971).
(19) J. S. Rollett, "Computing Methods in Crystallography", Pergamon Press, Elmsford, N.Y., 1965, p 55.
(20) Convergence for the β-Cl model conformer was obtained by constructing, at each cycle, not an extrapolated density matrix, but one representing 0.7 of the previous density matrix and 0.3 of the current density matrix. This is equivalent to applying a damping factor to the change in each array element of 0.3.
(21) This anomeric energy may be an overestimate for the glycosyl chlorides, since it assumes that the CH_3–O–CH_2–Cl torsion angle is unrestricted and can have the optimum value of 74°. In a pyranose ring, this angle is constrained by the ring closure to be closer to 60°; in the two α glycosyl chlorides for which experimental data are available (see Table III), the C(5)–O(5)–C(1)–Cl torsion angles are 68 and 72°.
(22) These values represent energy differences between the actual minima and are a refinement of those reported previously (ref 14 and 15), where the standard torsion angles of 60 and 180° gave energy differences of 3.0 and 2.4 kcal/mol for CH_3–O–CH_2–OH and CH_3–O–CH_2–OCH_3, respectively.
(23) S. J. Angyal, *Aust. J. Chem.*, **21**, 2737 (1968).
(24) R. U. Lemieux, *Pure Appl. Chem.*, **25**, 527 (1971).
(25) This is presently not practical for us, owing to program and computer limitations.
(26) E. L. Sutton, Ed., "Tables of Interatomic Distances", *Chem. Soc., Spec. Publ.*, **No. 18** (1965).
(27) Very different potential energy curves for dimethoxymethane and methoxymethyl chloride were calculated using a lower level semiempirical molecular orbital approximation (CNDO/2) by I. Ivaroska and T. Bleha, *Tetrahedron Lett.*, 249 (1975); *J. Mol. Struct.*, **24**, 249 (1975).
(28) J. T. Edward, *Chem. Ind. (London)*, 1102 (1955).
(29) Since the minimal STO-3G basis set is well-known to underestimate the steric repulsions, these results are consistent with the reversal in anomeric energies for the methanol fluoride and chloride on going to the higher level approximation, shown in Table II. An underestimate of the V_1 potential for methoxymethyl chloride could easily reverse our observed F > Cl anomeric energies.
(30) L. Radom, W. A. Lathan, W. J. Hehre, and J. A. Pople, *J. Am. Chem. Soc.*, **95**, 693 (1973).
(31) H. Sacchse, *Ber.*, **23**, 1363 (1890).

Theoretical Studies of Drug–Dinucleotide Interactions. Empirical Energy Function Calculations on the Interaction of Ethidium, 9-Aminoacridine, and Proflavin Cations with the Base-Paired Dinucleotides GpC and CpG

Merrill E. Nuss, Frederick J. Marsh, and Peter A. Kollman*

Contribution from the Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, California 94143. *Received April 12, 1978*

Abstract: We present empirical potential function calculations on the interaction of the cationic intercalators ethidium, proflavin, and 9-aminoacridine with the base-paired dinucleotide phosphates GpC and CpG. The calculations find reasonable base-paired structures for the GpC- and CpG-intercalator complexes and relative binding energies in the same order as the experimental free energies of association to DNA for ethidium and proflavin. The calculations are consistent with a low activation energy (<20 kcal/mol) for "opening up" the dinucleotide phosphate base pair from the DNA structure to one into which the intercalator can fit. The calculations are able to reproduce and rationalize the pyrimidine-3′,5′-purine specificity of these intercalators.

Introduction

One of the best characterized prototypal "drug-receptor" interactions is that between planar drugs and nucleotides. There is a wealth of physical chemical information on the interaction of planar aromatic dyes with DNA, RNA, and other nucleotides supporting an intercalated structure[1] for the dye–nucleotide complex.

Compared to the relatively large number of empirical potential energy calculations on peptides and nucleotides, there have been relatively few such studies of nucleotide–intercalator interactions. Sobell has used his X-ray structures of drug–nucleotide complexes as a basis for building molecular models of the drug–dinucleotide complex.[2] Alden and Arnott have used constrained model building to construct a stereochemically reasonable drug–decanucleotide complex.[3] Pack and Loew have carried out quantum mechanical calculations on fragments of the dinucleotide intercalator complex.[4] A very important study by Gilbert and Claverie,[5] using empirical potential functions, compared the interaction energy of proflavin, 9-aminoacridine, and acridine using a fixed geometrical model and focusing on base–intercalator interactions.

To our knowledge, no one has carried out an empirical potential function calculation on a *base-paired* dinucleotide phosphate and its complex with an intercalator varying all the torsional (14) and intermolecular (12) degrees of freedom. We describe such a study below.

The first goal of this study was to see whether these empirical energy calculations could find qualitatively reasonable structures for both base-paired dinucleotides and their intercalated complexes. The results of our calculations indicate that such methods can be successfully used in structure prediction of intercalated complexes.

A second goal was to calculate the energetics of drug intercalation, and, by examining the components of this process, gain insight into the important interactions which determine the kinetics and thermodynamics of drug intercalation. We discuss these insights below.

We carried out our calculations on the interaction of the

0002-7863/79/1501-0825$01.00/0

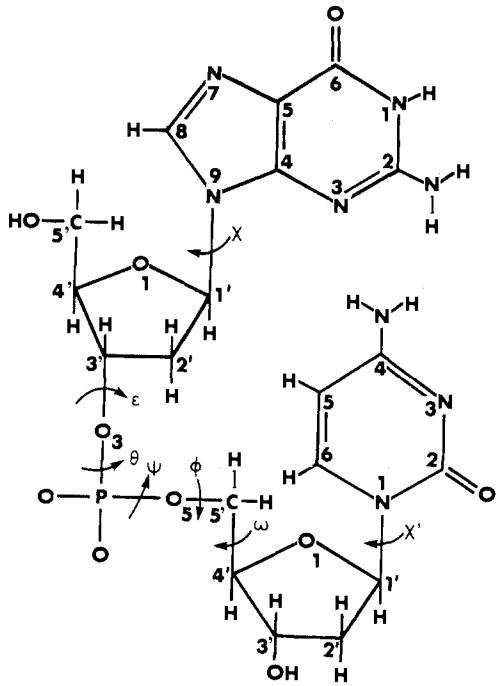
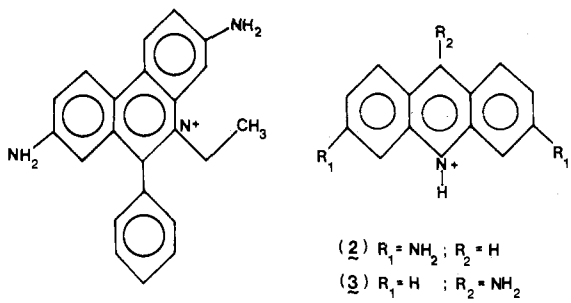


Figure 1. Dihedral angles in dinucleoside phosphate illustrated for GpC: χ = C6N1C1′C2′ (CpG) and C8N9C1′C2′ (GpC); ϵ = C4′C3′O3P; θ = C3′O3′PO5; ψ = O3′PO5′C5′; ϕ = PO5′C5′C4; ω = O5′C5′C4′C3′; χ' = C2′C1′N9C4 (CpG) and C2′C1′N1C2 (GpC).

base-paired dinucleoside phosphates

$$\begin{matrix} G\cdots C \\ {}^{3'}_{5'}\langle \quad \rangle^{5'}_{3'} \\ C\cdots G \end{matrix} \quad \text{and} \quad \begin{matrix} C\cdots G \\ {}^{3'}_{5'}\langle \quad \rangle^{5'}_{3'} \\ G\cdots C \end{matrix}$$

with ethidium (**1**), proflavin (**2**), and 9-aminoacridine (**3**). Our



(1)

(2) $R_1 = NH_2$; $R_2 = H$

(3) $R_1 = H$; $R_2 = NH_2$

reason for this choice is the well-established preference of ethidium and proflavin for pyrimidine-3′,5′-purine sequences in solution[6,7] and in the solid state.[2] Thus, the third goal of this work was to see whether the calculations could reproduce and rationalize this preference.

Computational Details

We used a potential function of the following type:

$$V = \sum_{i\neq j}^{n}\left(\frac{q_i q_j}{\epsilon R_{ij}} + B_{ij}e^{-C_{ij}R_{ij}} - \frac{A_{ij}}{R_{ij}^6}\right) + \sum_{k=1}^{\text{ntor}}\sum_{l=1}^{3} V_l^k \cos l\phi \quad (1)$$

The sum over i and j extended over all atom pairs separated by more than three bonds or on different molecules and contains contributions from electrostatic, exchange repulsion, and dispersion attraction. We use the final term in the potential to describe the torsional energy for rotation around the single bonds in the molecule, of which there are seven for each dinucleoside phosphate (Figure 1). The dispersion, exchange repulsion, and torsional terms were taken from literature values[8] (Table I). Electrostatic charges on the nucleotide atoms were taken from our ab initio calculations;[9] those for the drug atoms came from CNDO/2 calculations[10] carried out by us. The dielectric constant ϵ was taken as 1.

Table I. Potential Function Parameters Employed

atom pair	A_{ij}[a]	B_{ij}[b]	C_{ij}[c]
	Nonbonded Terms		
HH	19.05	1643.0	3.76
HC	120.45	11 297.0	3.67
CC	746.8	92 430.0	3.6
CN	746.8	92 430.0	3.6
NN	746.8	92 430.0	3.6
CO	746.8	49 029.0	3.6
NO	746.8	49 029.0	3.6
OH[d]	120.45	11 297.0	3.67
NH[d]	120.45	11 297.0	3.67
OO	746.8	42 156.0	3.6
PO	2400.0	256 774.0	3.6
PN	2400.0	305 369.0	3.6
PC	2400.0	305 369.0	3.6
PH	383.4	3466.0	3.6
	Torsional Barriers		
bond	V_1	V_2	V_3
C-N	0.0	0.9	0.0
C-P	0.0	0.0	1.21
C-C	0.0	0.0	1.16
O-P	0.0	1.5	1.5

[a] In kcal/mol·Å.[6] [b] In kcal/mol. [c] In Å^{-1}. [d] Note that A_{ij}, B_{ij}, and C_{ij} for the H-bonded hydrogen · · · N or O interactions are set to zero.

Two features of our potential function deserve further emphasis. First, we used a combination V_2 and V_3 Fourier potential function for the P-O bonds (ψ and θ), which we and others[11] have shown enables one to reproduce the tendency for R-O-P-O groups to have a gauche dihedral angle. Secondly, we treat hydrogen-bonded H's as Hagler, Huler, and Lifson did in their empirical potential function studies on amide crystals,[12] assigning them no dispersion or exchange repulsion interaction with the electron-donor atom. Hagler, Huler, and Lifson found that this approach allowed a successful analysis of amide crystal energies and structures, with no special H-bonded potential required.

The internal angles and bond angles for the BDNA structure were taken from Arnott[13] and those for drug-nucleotide complex from Sobell.[2] Hydrogen atoms were added to these structures in standard orientations. The Sobell structure was built for CpG; we constructed the GpC isomer by switching the purine and pyrimidine while retaining the geometry of the sugar-base connection from Sobell's structure.

The initial dihedral angles for the BDNA dinucleoside phosphate were taken from Arnott[13] and the initial dihedral angles for drug-dinucleoside phosphates from Sobell.[2] Ethidium (**1**) was placed initially as in the Sobell structure and the initial orientations for 9-aminoacridine (**3**) and proflavin (**2**) on top of the ethidium chromophore.

The energy was minimized using the Fletcher-Powell-Davidon method on numerically calculated derivatives.[14] The calculations continued until energy changes were less than 0.1 kcal/mol in an iteration and the gradients were all less than 1 kcal/Å or 0.1 kcal/deg. Some runs were minimized more accurately than this and suggested that the above criteria led to reasonable stopping points for the calculation.

Results and Discussion

A. Structures. 1. Torsional Angles. Table II summarizes the values of the torsional angles that define the sugar-phosphate backbone and sugar-base orientation. It should be noted that in the drug-nucleic acid complex the normal C3′-exo (C2′-

Table II. Calculated Torsional Angles for GpC and CpG in DNA and DNA-Intercalator Conformations[a]

			GpC						
angle[a]	BDNA[b]	EB[b]	BDNA	EB	EB[c]	PF[d] (major)	PF[d] (minor)	9AA[e] (minor)	9AA[e] (major)
χ',g	215.0	214.4	205.8	231.6	236.1	235.3	231.7	233.0	232.4
ω	329.1	318.4	317.8	277.8	269.0	269.7	269.4	273.5	274.0
ϕ	151.4	167.3	154.8	165.3	169.7	165.0	164.8	162.7	161.6
ψ	39.2	33.1	62.0	70.3	69.7	73.6	73.4	71.3	72.3
θ	98.9	80.8	100.5	68.1	64.2	61.3	64.7	62.2	62.7
ϵ	200.9	136.8	182.5	148.3	156.3	167.9	157.7	153.9	153.2
χ'	35.0	90.5	56.9	87.6	99.1	94.2	99.4	104.8	105.5
χ	215.0	214.4	205.6	235.4	236.8	243.3	231.7	234.3	234.0
ω	329.1	318.4	317.9	276.9	268.1	275.6	269.3	272.9	273.2
ϕ	151.4	167.3	154.7	166.0	169.9	173.6	165.3	164.2	162.6
ψ	39.2	33.1	61.7	71.0	70.3	70.4	73.3	70.9	72.5
θ	98.9	80.8	100.2	66.2	62.8	65.2	65.0	61.4	62.1
ϵ	200.9	136.8	182.3	145.6	159.6	161.2	157.4	154.5	154.3
χ'	35.0	90.5	57.6	91.2	97.9	83.1	98.7	104.1	104.8
			CpG						
angle	BDNA[b]	EB[b]	BDNA	EB	EB[c]	PF (minor)	PF (major)	9AA (minor)	9AA (major)
χ	215.0	214.4	205.6	217.0	232.9	223.5	223.5	218.4	220.7
ω	329.1	318.4	316.5	288.9	270.9	275.7	277.3	280.6	290.2
ϕ	151.4	167.3	154.7	149.8	166.2	152.3	150.4	89.7	138.1
ψ	39.2	33.1	67.1	73.0	70.9	73.4	70.6	110.0	77.1
θ	98.9	80.8	101.5	65.3	62.3	65.6	64.3	53.9	59.6
ϵ	200.9	136.8	178.9	152.7	159.7	166.3	164.3	174.4	156.2
χ'	35.0	90.5	52.4	93.4	97.5	99.5	98.6	101.5	104.5
χ	215.0	214.4	205.8	219.0	233.7	224.8	227.0	234.6	230.4
ω	329.1	318.4	316.8	291.1	271.5	274.8	276.0	285.7	283.3
ϕ	151.4	167.3	154.1	149.5	166.9	152.8	152.4	133.7	153.0
ψ	39.2	33.1	66.2	73.6	70.1	73.5	68.8	80.8	73.1
θ	98.9	80.8	100.4	65.5	61.5	65.9	64.4	58.0	61.6
ϵ	200.9	136.8	180.8	147.2	160.0	166.7	168.7	153.4	160.9
χ'	35.0	90.5	51.2	96.0	96.1	99.7	100.0	102.3	90.9

[a] The torsional angle is defined in terms of four consecutive atoms, ABCD; the positive sense of rotation is counterclockwise from A to D while looking down the BC bond. See Figure 1 and caption for definition of angles. [b] Starting values; BDNA from ref 11; EB from ref 2. BNA ≡ dinucleoside phosphate complex; 9AA ≡ 9-aminoacridine-dinucleoside phosphate complex; PF ≡ proflavin-dinucleoside phosphate complex; EB ≡ ethidium-dinucleoside phosphate complex. [c] Ethidium bromide without the ethyl and phenyl side chains. [d] Proflavin major has the N^+ group of proflavin pointing in toward the nucleoside; proflavin minor (see Figure 4) has the N^+ group pointing away. [e] 9-Aminoacridine (major) has the N^+ group pointing in toward the nucleoside; 9-aminoacridine (minor) has the N^+ group pointing away from the nucleoside. [f] The first seven angles refer to the first dinucleoside phosphate, the second to its H-bonded partner. [g] Our torsional angles are related to those of Sundaralingham[32] in the following way: our χ = Sundaralingham's $\chi + 120$; $\omega = 360 - \psi$; $\phi = 360 - \psi$; $\psi = 360 - \omega'$; $\theta = 360 - \omega$; $\epsilon = 360 - \phi'$; $\chi' = \chi' - 180$.

endo) deoxyribose sugar ring puckering of BDNA is altered to a mixed sugar puckering of the type C3′-endo(3′-5′)C2′-endo.[2] The conformations of the sugar rings, however, were not varied during the minimization calculations. Figure 2 compares the BDNA and energy optimized structures for the base-paired GpC and CpG structures and Figure 3 compares the Sobell model and energy optimized structures for ethidium GpC and ethidium CpG complexes.

As one might expect, the corresponding torsional angles of the two dinucleotides are almost identical for each complex since the drug-nucleic acid or BDNA complexes have a dyad or "pseudo"-dyad axis. In addition, the dihedral angles for most of the drug-nucleic acid complexes are similar indicating that the different drug-dinucleotide structures are closely related. The one exception is proflavin, with the C10 atom of the chromophore pointing toward the major groove. The starting conformation for this structure has the chromophore stacked between the base pairs of the dinucleotides with the amine groups pointing directly at the phosphate backbone. In this orientation the repulsive contribution for the interaction of the amine groups of the drug with the sugar-phosphate backbone of the nucleic acid dominates the total energy of the complex. To relieve the steric repulsion the dinucleotides are forced to move away from each other. This complex is the least stable of all the drug-nucleic acid structures because of this steric repulsion of the amine groups with the backbone.

Most of the conformational changes that occur when a drug intercalates between the base pairs of the nucleic acids result from (1) repuckering of the sugar rings and (2) a change in the dihedral angles of the sugar-phosphate backbone. The distance between the bases must increase from 3.4 to 6.8 Å when a drug intercalates between the base pairs of DNA. Our results indicate that by changing the pucker of the sugar rings of DNA, the backbone of DNA can be extended approximately 3.4 Å to accommodate the intercalator without a "large" change in the torsional angles. For example, a comparison of the differences in the torsional angles (excluding χ and χ') of the optimized BDNA structure with the optimized drug-nucleic acid structures, excluding proflavin (major), shows that the largest change in the dihedral angle is approximately 40° for ω and ψ. For the other torsional angles the changes are generally less than 20°. Previous work by Alden and Arnott[3] suggests that the angle ψ should increase by 120° if both the sugar puckers are C3′endo when a drug molecule intercalates. The same angle, in our calculations, only increases by approximately 40°. In addition Alden and Arnott[3] constrained their model building of hexa- and decanucleotides to those structures whose backbone and base structure closely resembled normal DNA for

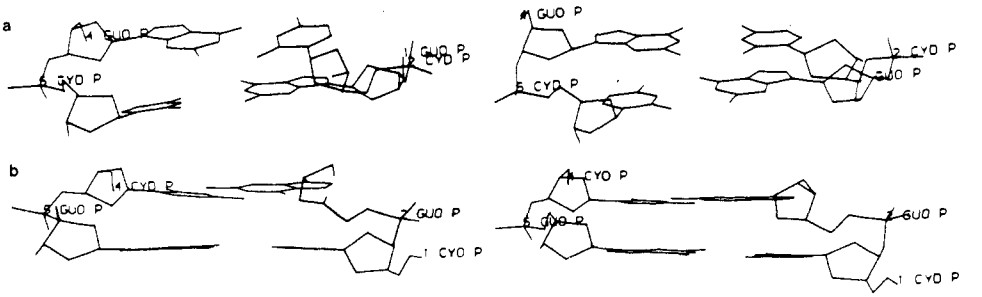


Figure 2. (a) Calculated and observed (ref 11) BDNA structures for GpC. (b) Calculated and observed (ref 11) BDNA structures for CpG.

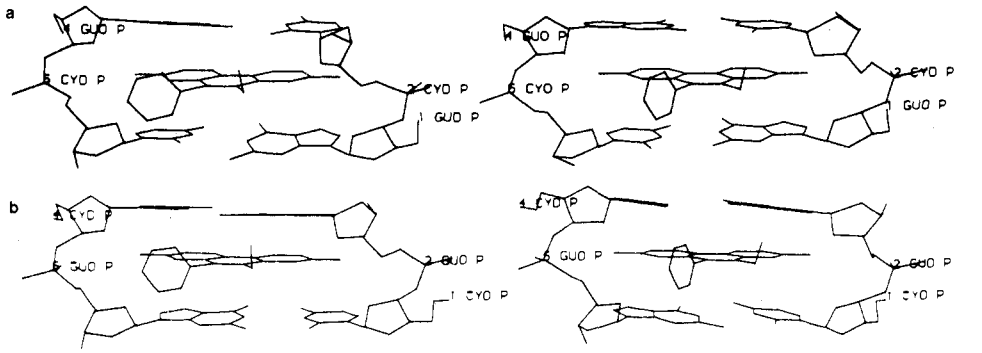


Figure 3. (a) Calculated and observed (ref 2) EtBr–GpC complex. (b) Calculated and observed (ref 2) EtBr–CpG complex.

those nucleotides adjacent to the intercalation site. That is why we must do our energy calculations on at least hexanucleotides for a complete comparison to be made.

The values of χ and χ' do not vary by much with the different drug–nucleic acid complexes. Their angle is essentially determined by the base stacking interactions which are very similar for all of the intercalators. There is a change, however, in the value of χ' between the optimized BDNA structure and the BDNA angles taken from Arnott,[13] the value changing from 35.0 to 71.3°.

2. Hydrogen Bond Distances. One important test of the validity of our potential energy function is the length of the hydrogen bond between the bases of the two dinucleoside monophosphates that are base paired in the BDNA and drug–nucleic acid structures. Table III gives the length of the six hydrogen bonds that are formed between the two self-complementary dinucleotides. All of the hydrogen bonds, with the exception of the proflavin (major) structure, have "acceptable" distances.

In each complex there is a small variation in the length of the different hydrogen bonds. This effect is accentuated if ethidium bromide is the intercalator. It is likely that the side chains of ethidium, especially the phenyl ring, interact with the base in such a way as to shorten one hydrogen bond and lengthen the other. When the side chains are removed from ethidium, the hydrogen bond lengths for the drug–nucleic acid complex are very similar to the other intercalator structures which also do not have side chains. In addition, the difference in the bond lengths is more pronounced for the first set "top" base pairs of three hydrogen bonds than it is for the bottom base pairs (GpC). It is possible that the positive side chain (ethyl group) of ethidium bromide, which has the methyl group pointing up, interacts with the bases in such a way as to distort the first set of hydrogen bonds more than the second set.

The calculated hydrogen bond energy we find for G–C is roughly half that found by Pullman[15] using similar empirical functions. Thus, it is interesting that our optimized structure has such reasonable H-bond lengths and structure, since the full negative charge of the two phosphates (18 Å apart) might be expected to drive the dinucleoside phosphates apart. We also carried out calculations on the GpC base pairs with Na^+ ions bifurcating the PO_2^- groups and these had structures very little altered from those in which Na^+ was not included.[16]

3. Comparison of Experimental Structures with Energy-Optimized Complexes. a. Ethidium (1). A comparison between the crystal and calculated structures of ethidium bromide with dinucleosides is important to test the reliability of the potential energy functions for these types of calculations. Table IV compares the values of the 14 torsional angles that are variables and the six hydrogen bond lengths for a crystalline complex of ethidium bromide with 5-iodocytidyl(3′–5′)guanosine (CpG) and our optimized structure of the drug with cytidyl(3′–5′)guanosine monophosphate (CpG). The deviations of the angles and bond lengths are relatively small indicating that our potential energy functions can (1) reproduce the characteristic changes (torsional angles) necessary for the sugar–phosphate backbone to extend from 3.4 to 6.8 Å and (2) maintain the hydrogen bonds that help stabilize the DNA and drug–nucleic acid structures.

It is important to note that our calculations give intercalation structures in closer agreement with the Sobell crystal structure (Table IV) than with the Sobell model (Table II). The latter was built to "fit" into double-stranded DNA; it is thus not surprising that our calculations, which do not have this constraint, come closer to the crystal structure coordinates.

b. Proflavin (2). Sobell has recently suggested[2] that intercalators may be divided into three classes depending on the direction of entrance into DNA. Actinomycin D and ethidium bromide enter from the minor groove, while drugs such as daunomycin and proflavin enter from the wide groove. Some drugs (9-aminoacridine) can enter from either side. Our work has provided a convenient check on some of Sobell's assumptions. Proflavin, for steric conditions, has to bind in the major groove. However, there are two orientations of the drug when it approaches from this direction: (1) that illustrated in Figure 4 (proflavin minor) or (2) that with N5, N1, and N7 pointing

Table III. Hydrogen Bond Lengths

				GpC					
H bonds	BDNA[a]	EB[a]	BDNA	EB	EB[b]	PF (minor)	PF (major)	9AA (minor)	9AA (major)
N3-N1[c]	2.85	2.98	2.77	2.75	2.77	2.80	2.76	2.80	2.81
N4-O6[c]	2.91	2.91	2.86	3.00	2.78	2.81	2.86	2.77	2.81
O2-N2[c]	2.84	2.72	2.82	3.25	2.93	2.87	3.01	2.88	2.87
N1-N3[d]	2.85	2.98	2.76	2.62	2.81	2.81	2.74	2.79	2.80
O6-N4[d]	2.91	2.91	2.86	3.01	2.77	2.82	2.82	2.78	2.81
N2-O2[d]	2.84	2.72	2.83	3.36	2.91	2.87	2.79	2.91	2.89
				CpG					
H bonds	BDNA	EB[a]	BDNA	EB	EB[b]	PF (minor)	PF (major)	9AA (minor)	9AA (major)
N3-N1[c]	2.85	2.89	2.85	2.83	2.83	2.84	3.63	2.90	2.83
N4-O6[c]	2.78	2.65	2.81	2.82	2.77	2.74	4.61	2.92	2.80
O2-N2[c]	2.79	3.01	2.79	2.73	2.90	2.83	3.21	2.76	2.77
N1-N3[d]	2.86	2.89	2.84	2.83	2.80	2.84	3.46	3.11	2.84
O6-N4[d]	2.79	2.65	2.83	2.84	2.78	2.73	4.40	3.19	2.86
N2-O2[d]	2.80	3.01	2.76	2.69	2.93	2.84	3.18	2.90	2.71

[a] Starting values; see Table II. [b] Ethidium bromide without the ethyl and phenyl side chains. [c] Top base pairs. [d] Bottom base pairs.

Table IV. Comparison of Model Calculations with Crystal Structure

	crystal structure[a]	calcd model[b]	deviation
	Torsional Angles[c] (deg)		
	A. First Dinucleoside		
χ	199	217	18
ω	288	289	1
ϕ	150	150	0
ψ	74	73	1
θ	79	65	14
ϵ	134	153	19
χ'	91	93	2
	B. Second Dinucleoside		
χ	191	219	28
ω	305	291	14
ϕ	136	150	14
ψ	69	74	5
θ	69	66	3
ϵ	135	147	12
χ'	96	96	0
	Hydrogen Bond Lengths (Å)		
N3-N1	2.88, 2.91	2.83, 2.83	0.06[d]
O4-N6	2.89, 3.01	2.82, 2.84	0.12[d]
O2-N2	2.81, 3.02	2.73, 2.70	0.20[d]

[a] H. M. Sobell, private communication, ref 2. [b] These results, CpG. [c] The torsional angles is defined in terms of four consecutive atoms, ABCD; the positive sense of rotation is counterclockwise from A to D while looking down BC bond. [d] Average deviation between calculated and observed H-bond distance.

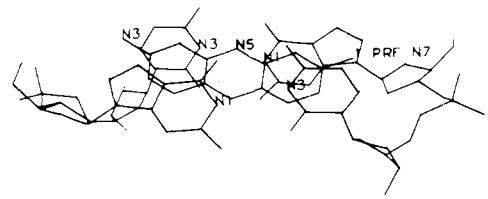


Figure 4. Calculated structure for proflavin (minor)-GpC complex looking down from the top of the nucleoside.

down rather than "up" as they are in Figure 4 (proflavin major). Our calculated energies (Table V, comparison of proflavin (minor) and proflavin (major)) suggest very strongly that the amines are pointing away from the helix. This allows the chromophore to slide between the base pairs of DNA to maximize the base stacking interaction. If the amines are pointing toward the helix the bases will not be able to stack effectively with the chromophore, which will result in a higher energy for the complex. Our results show that the drug-base and drug-backbone energy is much higher for the amines pointing toward the helix than for the opposite orientation. In this context, it is of interest that Patel and Canuel[7] have noted that the proflavin intercalation geometry which places the intercalator directly beneath the bases (with the relative orientations of the sugar-base and C-NH_2 bonds similar to our Figure 4) is more consistent with his NMR results (approximately equal upfield chemical shifts for all the proflavin protons) than the model proposed in ref 3, in which the proflavin is moved to an intercalation geometry in which the the hydrogens on the side of the molecule with the NH_2 and N groups would be expected to have rather different chemical shifts than the H's on the other side. The X-ray structure of a ribonucleotide CpG proflavin complex recently reported is also qualitatively consistent with our structure,[18] although they found, interestingly, a C3'endo pucker for both sugars and our calculations assume the mixed sugar pucker proposed by Sobell.[2]

c. 9-Aminoacridine (3). Our calculated results suggest a lack of preferential orientation of 9-aminoacridine, consistent with Sobell's assumption that the chromophore can enter and bind from either groove. There is no steric consideration with the drug entering from either side of the helix because it does not have amine groups on the side of the chromophore. Also, there is apparently no preference in terms of the interaction energy (Table V) for the chromophore to point toward the major or minor groove for GpC, although for CpG there is a definite preference for the minor groove structure. Note also (Table II) that the torsional angles for the structure where 9-NH_2(minor) interacts with base-paired CpG are rather different that those found for 9-NH_2(major) interacting with base-paired CpG or with 9-NH_2(major) or (minor) interacting with base-paired GpC. This is fascinating especially in view of the fact that Sakore et al.[19] found *two intercalated structures* for 9-aminoacridine interacting with iodo-CpG. One is *symmetric* with the 9-aminoacridine symmetrically disposed between the base pairs, and the other is asymmetric, with the intercalator closer to one nucleotide. This asymmetry appears

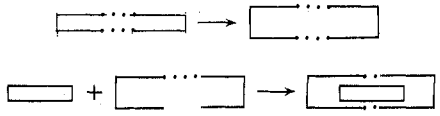
Table V. Energy Component Analysis (kcal/mol)

	GpC								
	BDNA[a]	EB[a]	BDNA	EB	EB[b]	9AA (minor)	9AA (major)	PF (minor)	PF (major)
				A. Intramolecular Components[c]					
base-base	83.45	108.21	85.66	108.53	107.94	107.88	107.79	107.99	107.40
base-bk	54.55	67.48	58.07	59.02	58.23	59.47	59.72	59.89	57.08
bk-bk	−47.37	−59.17	−64.06	−62.78	−62.10	−62.83	−63.00	−62.49	−62.04
torsional	−7.25	−8.77	−13.12	−15.05	−15.34	−14.80	−14.64	−15.21	−16.22
				B. Intermolecular Components					
base-base	−25.04	−18.26	−27.16	−20.43	−21.65	−21.66	−22.30	−22.55	−22.97
base-bk	8.28	6.02	8.06	7.59	8.54	8.73	8.56	8.20	7.65
bk-bk	10.83	10.06	11.04	10.68	10.79	10.92	10.93	10.90	10.46
drug-base		−76.39		−79.01	−79.95	−79.07	−77.54	−76.48	−75.11
drug-bk		−60.23		−83.51	−79.34	−73.04	−72.12	−81.09	−70.43
				C. Total Components					
E	77.45	−31.05	58.49	−74.96	−72.88	−64.40	−62.60	−70.84	−64.18
ΔE (without drug)		28.12		29.07	27.92	29.22	28.57	28.24	22.87
ΔE		108.50		−133.45	−131.37	−122.89	−121.09	−129.33	−122.67
	CpG								
	BDNA[a]	EB[a]	BDNA	EB	EB[b]	9AA (minor)	9AA (major)	PF (minor)	PF (major)
				A. Intramolecular Components[c]					
base-base	97.29	109.02	95.83	109.07	107.99	109.66	109.43	109.77	109.84
base-bk	59.52	65.05	69.84	66.24	59.03	67.05	66.68	67.54	67.32
bk-bk	−47.12	−59.17	−64.92	−64.18	−62.20	−70.18	−64.62	−63.00	−62.69
torsional	−7.23	−8.78	−13.57	−15.09	−15.71	−11.04	−14.72	−15.74	−15.75
				B. Intermolecular Components[c]					
base-base	−34.51	−23.28	−33.30	−23.19	−21.81	−21.12	−24.00	−24.50	−15.42
base-bk	0.68	7.17	1.83	5.17	8.47	3.87	4.22	4.81	3.12
bk-bk	10.91	10.07	11.17	10.76	10.79	13.47	11.09	11.14	11.54
drug-base		−77.84		−79.09	−80.39	−73.89	−76.82	−75.54	−63.74
drug-bk		−57.83		−81.81	−78.93	−86.99	−71.36	−81.55	−81.59
				C. Total Components					
E	89.54	−35.59	66.88	−72.12	−72.76	−69.17	−60.10	−67.07	−53.61
ΔE (without drug)		10.54		21.90	19.68	24.83	21.20	23.14	24.84
ΔE		−125.13		−139.00	−139.64	−136.05	−126.98	−133.95	−120.49

[a] Starting values; see Table II. [b] Ethidium bromide without the phenyl or ethyl side chains. [c] Base refers to atoms of guanine and cytosine of dinucleoside; bk refers to all other atoms of dinucleoside monophosphate (sugar and phosphate group); drug refers to atoms of the intercalator.

in our calculated torsional angles for 9-aminoacridine (minor)-CpG; the other three 9-aminoacridine structures we find are symmetric. It is also encouraging that our 9-NH_2 minor and 9-NH_2 major structures for CpG have the 9-amino groups as found by Sakore et al.[19] for the asymmetric structure (9-NH_2 group pointing toward the base-sugar bond) and symmetric structure (9-NH_2 group pointing away from the sugar-base bond).

B. Interaction Energies. The interaction of a drug with two base-paired dinucleoside monophosphates (GpC) can be envisaged as a two-step process in the gas phase. A schematic diagram of the sequence of events for the formation of a drug-nucleic acid complex is shown below. The first step is the



(a)

(b)

unstacking of the base-paired nucleic acid followed by the insertion of a drug chromophore between the unstacked base pairs. The first step is an endothermic process, i.e., energy is required to separate the bases. The second step, the insertion of the drug between the base pairs, must be an exothermic process that is larger in magnitude than the initial process if intercalation is to occur.

Table VI. Drug Specificities for CpG Base Pair ΔE (CpG + drug → complex)

energy component[a]	ethidium[b]	proflavin[c]	9-AA[d]
	A. Intramolecular[e]		
base-base	13.24	13.94 (14.01)	13.83
base-backbone	−3.60	−2.30 (−2.52)	−2.79
backbone-backbone	0.74	1.92 (2.23)	−5.26
torsional	−1.52	−2.17 (−2.18)	2.53
	B. Intermolecular[f]		
base-base	10.11	8.80 (17.88)	12.18
base-backbone	3.34	2.98 (1.29)	2.04
backbone-backbone	−0.41	−0.03 (0.37)	2.30
drug-base	−79.09	−75.54 (−63.74)	−73.89
drug-backbone	−81.81	−81.55 (−81.59)	−86.99
	C. Total		
	−139.00	−133.95 (−120.49)	−136.05

[a] In kcal/mol. [b] **1.** [c] **2.** In parentheses is proflavin with its amines pointing toward the major groove. [d] **3.** [e] Intramolecular energies involve those interactions within the dinucleoside phosphates. [f] Intermolecular energies include both dinucleoside phosphate-dinucleoside phosphate intermolecular interactions and those between the drug and the dinucleoside phosphate.

Table V gives a breakdown (component analysis) of the different contributions to the interaction energy of a drug in-

Table VII. Base Sequence Specificity of Intercalators. Differences in Interaction Energies of Drugs with CpG and GpC

energy component[a]	ethidium[b]	ethidium[c]	proflavin[d]	9-AA[e]
		A. Intramolecular[f]		
base-base	−9.65	−13.03	−8.39	−8.53
base-backbone	−4.55	−16.54	−4.12	−4.81
backbone-backbone	−0.52	−0.25	0.35	−0.76
torsional	0.41	−0.03	−0.08	0.37
		B. Intermolecular[g]		
base-base	3.38	3.45	4.19	4.44
base-backbone	3.81	8.75	2.84	1.89
backbone-backbone	−0.05	0.07	0.11	0.03
drug-base	−0.08	−1.45	0.94	0.72
drug-backbone	1.70	2.40	−0.46	0.76
		C. Total		
	−5.55	−16.63	−4.62	−5.89

[a] Energies in kcal/mol; a negative value implies a stronger interaction with base-paired CpG than GpC. [b] Ethidium cation with side chain (**1**). [c] Ethidium cation in initial Sobell geometry. [d] Proflavin (**2**); used minor groove structure for comparison. [e] 9-Aminoacridine (**3**); used major groove structure for comparison. [f] Intramolecular energies involve those interactions within the dinucleoside phosphate. [g] Include both nucleoside-nucleoside and nucleoside-drug intermolecular interactions.

teracting with two base-paired self-complementary dinucleoside monophosphates (GpC and CpG). The components are divided into two sections: (1) intramolecular effects, base-base, base-backbone (bk), and backbone-backbone, and torsional terms, and (2) intermolecular contributions resulting from base-base, base-backbone, backbone-backbone, drug-base, and drug-backbone interactions.

The designation for base-base contributions refers to individual atom-atom interactions where both atoms are on one of the four bases of the two base-paired dinucleosides. An intramolecular base-base interaction refers to interactions between base atoms in the same base or between bases on the same dinucleoside phosphate. An atom-atom interaction such as drug-bk (backbone) gives the contribution to the energy based on the drug-sugar phosphate backbone interaction. The atoms of the drug refer to the different intercalators. The absolute values for the energies are not important because they include many atom-atom interactions involving atoms whose relative orientations do not change during the minimization (e.g., N_1 and C_8 on guanine). Table VI contains the *difference* in energy components during the intercalation process base-paired CpG + intercalator → complex. Table VII compares the differences in energy components for the intercalation process with base-paired CpG with those components for base-paired GpC. In each case, we have separately calculated the contribution to complex formation of electrostatic, dispersion, and exchange repulsion atom-atom terms and these are tabulated in detail in ref 33.

1. Step (a). Unstacking of Nucleotides. In evaluating the component analysis data two points become apparent regarding the interaction energy of the drug-nucleic acid complexes we have studied. First, all of the drugs that form "reasonable" drug-dinucleoside complexes have essentially the same magnitude for the first step (the unstacking of the bases of the dinucleoside) for the intercalation process. To form an intercalation complex the distances between the bases of the dinucleoside must be increased from 3.4 to 6.8 Å. This requires a conformational change in the sugar-phosphate backbone (it must be extended), and results in an increase in energy as the basis of the dinucleoside monophosphates becomes "unstacked". This increase is ~15 kcal/mol loss of dispersion for both GpC and CpG base-paired dinucleoside phosphates. In addition, there is a change in electrostatic energies for the unstacked dinucleotide. For GpC, the unstacked dinucleotide is 4 kcal/mol higher in electrostatic energy; for CpG, the unstacked form is 3 kcal/mol lower. These changes are all *intramolecular* and the difference in electrostatic components is an important reason why it is easier to unstack CpG than GpC dinucleoside phosphate. The intermolecular energy contribution to unstacking GpC comes mainly from the less optimum interbase H bonds in the unstacked structure. In CpG, there is a substantial dispersion attraction (between the two guanines in the stacked structure), which is lost on unstacking. Thus, the calculated net destabilization for step (a) is 15-30 kcal/mol for the dinucleotide drug interactions in Table V.

It appears from experimental kinetic measurements that the unstacking free energy of activation is about 15 kcal/mol for proflavin-DNA intercalation.[20] Our calculated activation energies cannot be compared rigorously to these, but the fact that the numbers are in the same range is encouraging.

2. Step (b). Drug Intercalation. The results of Table VII show that the intercalation complex is formed because of the strong interaction between the drug and the two dinucleosides. For examples, there is a gain of ~150-160 kcal when ethidium bromide interacts with dinucleoside base pairs in their "unstacked" conformation (step (b)). The total difference, therefore, is a net stabilization in the gas phase of ~130-140 kcal/mol ethidium bromide-nucleic acid structure. There is a large variation in the interaction energy of step (b) for the different intercalation complexes.

In comparing the interaction energy of different intercalators with GpC and CpG it is clear that the different binding energies result mainly from the interaction of the drug with the sugar-phosphate backbone and the bases of the dinucleosides. With this type of analysis one would also not expect there to be a great difference in the binding energy between proflavin and ethidium without the side chains. The difference in the interaction energy for these two molecules is indeed very small. The difference in binding energy for the two ethidium molecules, with and without side chains, is also small. The relatively high energy of the proflavin (major) orientation is a result of the amines actually being too close to the phosphate backbone and distorting the nucleic acid structure.

One would also like to understand the nature of the effects which change the gas-phase ΔE of −130 to −140 kcal/mol to the observed $\Delta G \sim -10$ kcal/mol for ethidium-DNA interactions.[1] A qualitative estimate of the energetics of solvation of nucleotide, ethidium, and complex comes from the reaction field stabilization for ions:[21]

$$E = \frac{\epsilon - 1}{2\epsilon} \frac{q^2}{a} \qquad (2)$$

where q is the charge of the molecule, $\epsilon = 80$ (H_2O), and a is

the cavity radius. If we use the dimensions of the dinucleotide perpendicular to the helix axis (roughly 18 × 10 Å ($r \sim 7$ Å)) for the dinucleotide and its complex and an average "radius" of 6 Å for ethidium one calculates solvation estimates for dinucleoside phosphate ($q = -2$), ethidium ($q = 1$), and complex ($q = -1$) of −95, −27, and −24 kcal/mol and a contribution of ΔG (solvation) to the intercalation process of ∼100 kcal/mol. Estimating the rotation and translational entropy change on association as Page and Jencks[22] (ΔG trans-rot ∼20 kcal/mol) have leads to a net ΔG of ~ -10 to −20 kcal/mol for the interaction of cationic intercalations into DNA. Although the above estimates are very crude, they do suggest that our gas-phase ΔE's are not unreasonable.[23]

3. Drug Specificities. Table VI compares the interaction energies of the three drugs with base-paired CpG. As one can see, the order of binding affinities is in the order ethidium ≳9-aminoacridine > proflavin. There is some direct experimental evidence which suggests that ethidium bromide binds more tightly to DNA than does proflavin and 9AA is also more tightly bound than proflavin.[26] It is premature, however, to compare directly our results with the experimental work. DNA has ten different base pair combinations with which the two chromophores can bind. It is expected that each of the ten drug–base pair combinations will have a different equilibrium constant because of the different drug–base interactions.

4. Base Specificity. In Table VII we compare the interaction energies for drug GpC interactions with drug CpG interactions. Krugh was the first to note a preference of ethidium for the pyrimidine-3′,5′-purine isomer.[6] Patel has noted such a preference for proflavin as well.[7] Our calculated interaction energies for ethidium and proflavin are consistent with these observations. Since we used the Sobell CpG structure[2] as the basis for our sugar pucker and initial dihedral angles, a more negative calculated ΔE for the experimentally favored CpG sequence might be suspect because this geometry choice could have biased the results. Clearly, calculations allowing more geometrical flexibility in the sugar are required to sort out this point.[27]

However, it is encouraging that the calculations are able to reproduce this energy difference. Whether one uses the initial Sobell or energy optimized geometries, the main factors favoring pyrimidine-3′,5′-purine specificity are (1) the difference (mainly electrostatic) in intramolecular base–base interaction, which favors the unstacked structure by 3 kcal/mol in CpG and the stacked structure by 4 kcal/mol in GpC and (2) the intramolecular 3′-base–backbone interaction (mainly electrostatic), which destabilizes the unstacked structure by ∼1 kcal/mol in GpC and stabilizes it by ∼4 kcal/mol in CpG.

Pack and Loew[4] have previously focused attention on the latter interaction as playing the most important role in causing pyrimidine-3′,5′-purine sequence specificity, and in our calculations using the Sobell and initial BDNA geometry, effect (2) is slightly larger than (1). On the other hand, base stacking calculations by Pullman[15] find that the GpC base-paired structure is 2.8 kcal/mol harder to unstack (from the BDNA structure to ∞ separation) than the CpG base pair and this supports our finding of factor (1) (base–base interaction) as a reason that intercalation is easier between base-paired CpG. We conclude that both factors (1) and (2) are important; the fact that the electrostatic component is the predominant one in both base–base and base–backbone interactions suggests experiments of base specificity as a function of ionic strength or with esterified phosphates. The relative sequence specificity due to base–backbone interactions should be more sensitive to the extent of anionic character of the phosphate than the backbone–backbone interactions, which are more shielded from solvent, so one might be able to experimentally assess the relative roles of factors (1) and (2) in determining base sequence specificities.

It is clear from studies that the main reason for the pyrimidine-3′,5′-purine specificity is *intramolecular* and not dependent on the drug–nucleic acid interaction. One could envisage purine-3′,5′-pyrimidine drug preferences (e.g., actinomycin)[2] but only if the drug–nucleotide interactions overcome the inherently smaller energy to open up pyrimidine-3′,5′-purine sequences.

Conclusion and Further Considerations

Our calculations are the first attempt to use a empirical potential energy function to study the interaction of drugs with small nucleic acid components to determine the origin, nature, and magnitude of the forces that dictate the conformation and intermolecular structure of such a complex. We have minimized the complex energy *varying* all torsional (14) and intermolecular (12) degrees of freedom. Comparing hydrogen bond lengths and torsional angles of the sugar–phosphate backbone in the calculated and crystal structure of ethidium bromide–dinucleoside monophosphate gives us confidence that our program is capable of producing reasonable geometries for this type of complex. The proflavin intercalated structure we find to be of lower energy is closer to that observed by X-ray crystallography than the alternative structure we considered. Thus, we feel that our first goal has been realized. Secondly, we have found reasonable energies for base unstacking and drug intercalation. An extremely crude solvation estimate has indicated that our calculated interaction energies are qualitatively reasonable and points out the importance of solvation–desolvation effects in determining intercalation energies for cationic drugs. Thirdly, our calculations explain the observed pyrimidine-3′,5′-purine sequence specificity of ethidium and proflavin in terms of *intramolecular* base–base and base–backbone interactions and suggest similar specificity for 9-aminoacridine, for which, to our knowledge, no solution-phase specificities have been determined. Thus, we feel that we have reached the goals we set for ourselves in the Introduction. However, it is necessary to make some cautionary remarks about the results we have obtained. First, we started our minimization calculations out at or *near* the Sobell or BDNA structures; in a number of cases, we checked to see that different starting guesses converged to the same structure. However, it is clear that if we started the two dinucleoside phosphates far enough apart the energy minimizer would never bring them together because of the electrostatic repulsion which leads to an energy barrier required to bring two anions in contact. One might change this by using distance dependence dielectric functions,[23] but these have not yet been adequately calibrated for our use here. In any case, we have no guarantee that we have found the global minima in our calculations, only that we have found physically reasonable local minima near observed or model-built geometries.

Secondly, our potential functions need to be further validated.[28] Our calculations include only electrostatic, exchange repulsion, and dispersion effects. Based on the results of ref 12 and 29, we feel that the neglect of polarization and charge transfer energies is a reasonable one for this system.[30] These energies may be of significant magnitude, but are unlikely[12] to affect the structure of the complexes in a major way. A more serious approximation lies in the crude way in which solvation is handled here, and we feel that it is worth expending considerable effort on improved solvation models such as those of ref 25 and 31.

Thirdly, our use of a fixed sugar pucker needs to be examined and the recent work of Levitt and Warshel[27] puts us in a position to do so. However, the essential point of the Levitt–Warshel paper[27] is the extreme flexibility of the furanose ring. This suggests that the 15–30 kcal/mol barrier we calculate for "opening up" the dinucleotide (step (a) in the intercalation process) is likely to be an upper bound for this process.

Fourthly, the extension of these studies to tetra- and hexanucleotides is necessary to be sure of the "relevance" of these model calculations to intercalation into polynucleotides. The fact that our dinucleotide calculations find a minimum energy near BDNA and our dinucleotide–intercalator calculations find a minimum near the Sobell model built intercalation geometry is encouraging in this regard. Of particular future interest is the neighbor exclusion rule, in which intercalation occurs at every other site, which has been explained[2] on the basis of sugar pucker changes.

Finally, there are a number of other questions to which we plan to devote future attention: (1) What is the role of sugar repuckering in the intercalation process? (2) What is the kinetic pathway and the energetics along this pathway for the intercalation process? (3) We have only considered two of the ten dinucleotide intercalation sites in this study. What does one find for the remaining eight sites? (4) How do the results reported here change when one considers intercalation into larger nucleotide fragments, e.g., tetra- and hexanucleotide base-paired sequences? (5) What differences does one expect between ribonucleotides and the deoxyribonucleotides studied here?

Acknowledgments. Peter A. Kollman would like to thank the Academic Senate, the NIH (GM-20564), and the NSF (CHE-81718) for research support and the NIH (GM-70718) for a Career Development Award. We would also like to thank H. M. Sobell, G. H. Loew, G. L. Pack, D. Patel, A. Warshel, and M. Levitt for communicating their results to us prior to publication. We are especially grateful to Professor R. Langridge for useful discussions and the use of the UCSF Computer Graphics Laboratory, supported by NIH Grant RR-1081, and to to Professor W. Olson for her helpful comments on this manuscript.

References and Notes

(1) V. A. Bloomfield, D. Crothers, and I. Tinoco, "Physical Chemistry of Nucleic Acids", Harper and Row, New York, N.Y., 1974, p 72.
(2) H. Sobell and S. C. Jain, *J. Mol. Biol.*, **114**, 301, 317, 333 (1977); **68**, 21 (1972).
(3) C. J. Alden and S. Arnott, *Nucleic Acids Res.*, **2**, 1701 (1975); **4**, 3855 (1977).
(4) G. R. Pack and G. H. Loew, *Int. J. Quantum Chem., Quantum Biol. Symp.* **4**, 87 (1976); "Origins of the Specificity in the Intercalation of Ethidium into Nucleic Acids: A Theoretical Analysis", submitted for publication.
(5) M. Gilbert and P. Claverie in "Molecular Associations in Biology", B. Pullman, Ed., Academic Press, New York, 1968, N.Y., p 350.
(6) T. R. Krugh and C. G. Reinhardt, *J. Mol. Biol.*, **97**, 133 (1975).
(7) D. J. Patel and L. Canuel, *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 2624 (1977).
(8) We used the CC and HH nonbonded values from the Warshel–Levitt molecular mechanics program (Quantum Chemistry Program Exchange No. 247) and scaled the repulsive terms for O and N taking into account the van der Waals radii (e.g., OO should have a shorter van der Waals minimum than NN by 0.2 Å). For phosphorus, we also scaled the dispersion attraction (A_{ij}) to take into account the greater polarizability of phosphorus.
(9) M. E. Nuss and P. A. Kollman, *J. Med. Chem.*, submitted.
(10) J. A. Pople and D. L. Beveridge, "Approximate Molecular Orbital Theory", McGraw-Hill, New York, N.Y., 1970.
(11) D. M. Hayes, S. Rothenberg, and P. A. Kollman, *J. Am. Chem. Soc.*, **99**, 2150 (1977); G. Govil, *Biopolymers*, **15**, 2303 (1976).
(12) A. T. Hagler, E. Huler, and S. Lifson, *J. Am. Chem. Soc.*, **96**, 5319 (1974); see also J. Caillet and P. Claverie, *Acta Crystallogr., Sect. A*, **31**, 448 (1975).
(13) S. Arnott, P. J. Cambell-Smith, and P. Chandrasekaran in "CRC Handbook of Biochemistry", Vol. II, "Nucleic Acids", 3rd ed., G. D. Fasman, Ed., CRC Press, Cleveland, Ohio, 1976, p 411.
(14) D. J. Wilde and C. S. Beightler, "Foundations of Optimization", Prentice-Hall, Englewood Cliffs, N.J., 1967, p 331.
(15) B. Pullman in "Molecular Associations in Biology", B. Pullman, Ed., Academic Press, New York, N.Y., 1968.
(16) Using STO-3G partial charges, one might expect that one would underestimate the polarity of the key H-bonded groups (see ref 17), so we tried adding 0.4 e^- to O_2 and N_4 of cytosine and to O_2 and N_2 of guanine and subtracted 0.4 e^- from C_2, the two N_4 hydrogens (0.2 e^- each) of cytosine and C_6, and the two N_2 hydrogens (0.2 e^- each) of guanine. A calculation employing these parameters effectively doubled the H-bond strength, but did not have a major change on the H-bond structure, decreasing the H-bond distances by an average of 0.1 Å.
(17) D. M. Hayes and P. A. Kollman, *J. Am. Chem. Soc.*, **98**, 3335 (1976).
(18) S. Neidle, A. Achari, G. L. Taylor, H. Berman, H. Carrell, J. Glucker, and W. Stallings, *Nature (London)*, **269**, 304 (1977).
(19) T. D. Sakore, S. C. Jain, C. C. Tsai, and H. M. Sobell, *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 188 (1977).
(20) H. J. Li and D. Crothers, *J. Mol. Biol.*, **39**, 461 (1969).
(21) D. L. Beveridge and G. Schnuelle, *J. Phys. Chem.*, **79**, 2562 (1975).
(22) M. I. Page and W. P. Jencks, *Proc. Natl. Acad. Sci. U.S.A.*, **68**, 1678 (1971).
(23) The calculated ΔE's are still too large, probably because of our use of $\epsilon = 1$ for all atom pairs. The use of a larger ϵ cuts down the electrostatic interaction energy and one could scale ϵ to bring the calculated ΔE's into agreement with the experimental ΔG's. However, when $\epsilon = 2$ is used in the energy minimization one finds much too large H-bond distances for the base pairs. One could imagine using a distance-dependent ϵ, as indicated by the results of Hopfinger[24] and Warshel and Levitt.[25] Such an approach might lead to a good H-bond structure, but reduced electrostatic backbone–intercalator and backbone–backbone interactions. Further studies on this approach are in progress.
(24) A. Hopfinger, "Conformational Properties of Macromolecules", Academic Press, New York, N.Y., 1973.
(25) A. Warshel and M. Levitt, *J. Mol. Biol.*, **103**, 227 (1976).
(26) F. Quadrifoglio, V. Crescenzi, and V. Giancotti, *Biophys. Chem.*, **1**, 319 (1974).
(27) The interesting recent study by Levitt and Warshel (*J. Am. Chem. Soc.*, **100**, 2607 (1978)) has derived an analytical potential for ribose puckering which we are currently incorporating into our intercalator calculations.
(28) For example, we know from the crystal structure study by Sakore et al.[19] that the intercalated structures of 9-aminoacridine major and minor into CpG should be close to equoenergetic. Our calculations (Table V) find that they differ in energy by ~9 kcal/mol. By scaling the H-bond and dispersions attractions, we hope to find a set of parameters that will find the two structures to be of approximately equal energy.
(29) J. Caillet and B. Pullman, *Acta Crystallogr., Sect. A*, **31**, 448 (1975).
(30) It should be noted that the potential we employed is largely parametrized by fitting a large number of crystal structures and energies and thus may indirectly include some polarization and charge transfer effects in the dispersion term. Compared to other force fields, its repulsive term is comparatively soft.
(31) M. J. Huron and P. Claverie, *J. Phys. Chem.*, **78**, 1853, 1862 (1974).
(32) M. Sundaralingham, *Biopolymers*, **7**, 821 (1969).
(33) M. E. Nuss, Ph.D. Thesis, University of California, San Francisco, 1977.