A theoretical study of anthracene and phenanthrene derivatives acting as A-T specific intercalators

Kai-Xian Chen*, Nohad Gresh and Bernard Pullman

Institut de Biologie Physico-Chimique, Laboratoire de Biochimie Théorique associé au CNRS, 13 rue Pierre et Marie Curie, 75005 Paris, France

Received 10 September 1986; Accepted 27 October 1986

ABSTRACT

Theoretical computations are performed on the comparative A-T versus G-C binding selectivities of two DNA intercalating molecules recently synthesized by Wilson et al. [1]. These are derivatives of phenanthrene and anthracene with side chains containing an hydroxy group bound to its C α carbon and a cationic amino group bound to its C β carbon. We have optimized the binding energies of these phenanthrene and anthracene derivatives (**1** and **2**, respectively) to the double-stranded tetramers d(ATAT)₂ and d(GCGC)₂, the intercalation occurring in the central pyrimidine (3'-5') purine sequence. The sum of the intercalator-oligonucleotide intermolecular interaction energy plus the conformational energy variation of the intercalator upon binding were computed by the SIBFA procedures, which use empirical formulas based on ab initio SCF computations. Both compounds are found to bind more favourably to the AT sequence than to the GC one. Moreover, the affinity of **1** for the AT oligomer is computed to be larger than that of **2**, whereas conversely that of **2** is larger than that of **1** for the GC oligomer. The AT versus GC binding selectivity of **1** is significantly larger than that of **2**. These results are in excellent agreement with the experimental findings of Wilson et al. [1]. However, contrary to the suggestion of these authors the α -hydroxy group of the side chain of the intercalators does not seem to play a decisive role in determining the A-T specificity.

INTRODUCTION.

Whereas a number of intercalator molecules are endowed with high binding affinities to DNA, often correlated with distinctive antitumor or antibiotic potencies, they may differ widely in terms of their preferential base sequence selectivities. Hence, to name a few, whereas actinomycin [2] and mitoxantrone [3] are G-C specific, tilorone [4, 5] and the neocarzinostatin chromophore [6] are A-T specific. A more involved situation can occur as in the case of daunomycin and adriamycin, for which we have shown [7, 8] that the actual sequence selectivity must be expressed at the level of a triplet of mixed base pairs.

Altogether, it is nevertheless generally considered that intercalators prefer G-C sites. Wilson et al. [1] have recently undertaken a search for intercalators which would definitely exhibit a strong preference for A-T sites. They succeeded in finding them by synthesizing three compounds among which deriva-

Figure 1 : Compounds 1 and 2. Molecular structure and atom numbering.

tives of phenanthrene and anthracene both with side chains containing an hydroxyl group in the α position and an amino group connected to the C_β carbon. The molecular structures of these phenanthrene and anthracene derivatives (denoted 1 and 2, respectively, hereafter) are given in figure 1, together with the atom numbering. A combination of NMR measurements, spectrophotometric titrations, and thermal denaturation studies have shown [1] that both compounds display a definite A-T versus G-C selectivity. Furthermore, whereas the affinity of 1 for poly (dA-dT) . poly (dA-dT) was found to be distinctly larger than that of 2, conversely, that of 2 for poly (dG-dC) . poly (dG-dC) was found to be larger than that of 1. Compound 1 was, in fact, reported to be the most A-T specific intercalator known today.

These results have stimulated the present investigation, the purpose of which is to account by theoretical computations for the overall A-T selectivity of 1 and 2 and the greater preferential A-T selectivity of 1 than of 2 and to determine the origin of this phenomenon. We will also attempt to account for the reversal of relative binding affinities between 1 and 2 upon passing from the A-T polymer, favouring 1, to the G-C polymer, favouring 2.

This research will be done by performing theoretical computations of the compared binding affinities of 1 and 2 with the model double-stranded tetramers d(ATAT)₂ and d(GCGC)₂. In which intercalation is considered to take place in the central pyrimidine (3'-5') purine sequence (see figure 2 for base numbering). The choice of this sequence rather than of the isomeric purine (3'-5') pyrimidine sequence is based on the results of previous computations [9-12] of comparative

Figure 2 : The investigated tetramers $d(ATAT)_2$ and $d(GCGC)_2$. Base numbering.

energy balances for the binding of representative intercalators, which significantly favoured the former type of isomer.

For minor groove binding, both $d(ATAT)_2$ and $d(GCGC)_2$ sequences will be investigated. For major groove binding, only the $d(GCGC)_2$ sequence will be investigated. This is justified by the results of our former computations devoted to DNA binding of cationic ligands [13], showing that the major groove of the A-T sequences is significantly disfavoured with respect to the minor groove.

PROCEDURE

Similar to our previous studies of intercalative binding to oligonucleotides [7, 8, 14, 15], the interactions are computed with rigid models of the oligonucleotides fixed in the appropriate conformation (see below), but allow a large flexibility to the drug, considered protonated on its nitrogen atom.

The variations of the conformational energy of the intercalator upon complex formation are computed with the SIBFA (Sum of Interactions Between Fragments Ab initio computed) procedure [16]. Within this methodology, the intercalator is built of elementary constitutive fragments separated by single bonds, and the variation of the intramolecular energy upon a conformational change is obtained as the variable part of the sum of interactions between the fragments expressed as :

$$(1) \delta E_{\text{conf}} = E_{\text{MTP}} + E_{\text{pol}} + E_{\text{rep}} + E_{\text{disp}} + E_{\text{tor}}$$

The intermolecular oligonucleotide-intercalator interaction energies are

computed by the SIBFA 2 procedure [17] as a sum of five terms :

$$(2) \Delta E = E_{\text{MTP}} + E_{\text{pol}} + E_{\text{rep}} + E_{\text{disp}} + E_{\text{CT}}$$

In expressions (1) and (2), E_{MTP} and E_{pol} denote respectively the electrostatic and polarization contributions, computed using a multipolar expansion of the *ab initio* SCF molecular wave functions of the fragments, and E_{rep} and E_{disp} are the repulsion and dispersion contributions, respectively. E_{tor} is a torsional energy contribution, calibrated in [16] for elementary rotations along C-C and C-O bonds, and E_{CT} is a charge-transfer contribution (see [16, 17] for details). The fragments utilized for the intercalators are a C₁₀ methylated phenanthrene or anthracene ring, a protonated methyldiethylammonium and HOH molecules, using standard geometries [18]. To allow rotations along the C₁₀-C_α, C_β-N, N-C_γ and C_γ-C_δ bonds the aromatic and ammonium groups were then subdivided into sub-fragments : aromatic ring plus methane and ammonium group plus methane (resp.) in such a way as to preserve the same multipolar expansion as in the original fragments. The conformational energies of the two intercalators either free or in the complex were minimized as a function of all their dihedral angles.

The DNA tetramer is built out of elementary fragments, following a procedure previously utilized to compute the molecular electrostatic potential of large biomolecules [19]. The fragments are the nucleic acid bases, the phosphodiester linkage and deoxyribose, with the same internal geometry as in standard B-DNA [20].

To construct the geometry of the tetramer susceptible to accept an intercalating drug, we have resorted to the results of conformational analysis of intercalated DNA due to Miller et al [21, 22], using their AGNAS procedure [23, the label standing for Algorithm to Generate Nucleic Acids Structure). Consistent with the experimental observation of an unwinding angle of 17° for DNA upon binding this class of compounds [1, 24], we have adopted, in analogy to our study devoted to mitoxantrone [15], the conformation denoted (IIB, ΔX = 0.0) in ref. (21, 22), characterized by this value of unwinding angle. Similar to our preceding studies, we will assume also that the same tetramer conformation is retained for all sequences. The *ab initio* SCF computations on the constitutive fragments necessary for the application of the SIBFA procedure [16, 17] are performed using our usual basis set with a *dzeta* exponent of 1.5 on the ammonium hydrogen [25].

The search for the optimal configurations for the oligonucleotide-intercalator complexes are performed by energy-minimization [26] of the sum of ΔE

plus $\delta E_{\text{conf.}}$. The variables involved in the minimization are the conformational angles of the intercalator together with the intermolecular variables defining the relative orientation of the intercalator with respect to the tetramer.

RESULTS AND DISCUSSION.

The results of the computations are reported in tables I and II. Table I lists the values of the intermolecular oligonucleotide - intercalator interaction energies, ΔE , their different components, the variation of the conformational energy of the intercalator in the complex with respect to the energy of its most stable conformation in the absence of interaction $\delta E_{\text{conf.}}$, the values of the unstacking energies of the CpG and TpA sequences $\delta E_{\text{unstack}}$ computed as in [7], the resulting energy balance $\delta E = \Delta E - \delta E_{\text{conf.}} - \delta E_{\text{unstack}}$ and the value, δ , of the energy differences with respect to the most favourable value of δE taken as energy zero.

Table II lists the most significant interatomic distances between the oligonucleotide and the intercalator and the optimized values of the torsion angles of the intercalator, prior to interaction, as well as in the respective complexes.

TABLE I
Values of the binding energies of phenanthrene and anthracene derivatives 1 and 2 to d(ATAT)₂ and d(GCGC)₂ (see text for definitions) Energies in kcal/mole.

	MINOR GROOVE		BINDING		MAJOR GROOVE BINDING	
	d (A T A T) ₂		d (G C G C) ₂		d (G C G C) ₂	
	Phenanthrene derivative 1	Anthracene derivative 2	Phenanthrene derivative 1	Anthracene derivative 2	Phenanthrene derivative 1	Anthracene derivative 2
ΔE	- 263.2	- 258.0	- 242.7	- 242.2	- 250.5	- 252.0
E_{rep}	- 227.6	- 223.8	- 205.9	- 207.5	- 223.4	- 229.4
E_{pol}	- 14.4	- 13.2	- 12.5	- 12.6	- 13.1	- 13.0
E_{CT}	- 4.2	- 3.8	- 2.9	- 3.6	- 4.2	- 3.1
E_{disp}	- 70.9	- 63.8	- 67.4	- 61.8	- 65.8	- 51.5
E_{rep}	53.9	46.6	45.9	43.3	56.0	44.9
$\delta E_{\text{conf.}}$	20.8	17.9	21.4	15.1	19.5	16.6
$\delta E_{\text{unstack}}$	9.5	9.5	10.5	10.5	10.5	10.5
δE	- 232.9	- 230.6	- 210.8	- 216.6	- 220.5	- 224.9
δ	0.0	2.3	22.1	16.3	12.4	8.0

TABLE II.

Values of the most significant interatomic distances in the complexes of 1 and 2 with $d(ATAT)_2$ and $d(GCGC)_2$ and values of their optimized torsional angles prior to interaction, and in the complexes. Distances in Angstroms, torsional angles in degrees.

		MINOR GROOVE BINDING				MAJOR GROOVE BINDING			
		d (ATAT) ₂		d (GGGC) ₂			d (GGGC) ₂		
		1	2	1	2		1	2	
O ₂ (π ₂ or C ₂) - HC _β		1.92	1.94	2.18	2.16	O ₁ (PO ₄) - H C _{γ1} - H(O ₂ H)	1.94	1.96 1.92	
O ₂ (π ₂ or C ₂) - H ₁ C _{γ1} - H C _{γ2}		2.64	2.81	3.96	3.97	H ₄ (C ₂) - H (O ₁ H)	2.32	2.55	
		2.62	3.04	4.02	4.09				
H ₃ (A ₃ or C ₃) - H (O ₁ H)		2.49	2.17	2.25	2.03				
O _{1,3} (A ₃ or C ₃) - H(O ₂ H) - H ₂ C _{γ1}		2.33	2.62	4.00	4.19				
		2.85	2.77	2.80	2.88				
	1	2	1	2	1	2		1	2
C ₁₂ -C ₁₀ -C _α -C _β	141.4	143.8	153.0	159.7	162.0	167.1		132.9	161.1
C ₁₀ -C _α -C _β -H	50.6	300.2	183.5	180.8	172.4	170.4		183.7	177.7
C _α -C _β -H-H-C _{γ1}	116.7	122.8	67.8	67.5	81.9	83.5		92.7	90.0
C _β -H-H-C _{γ1} -C ₈₁	275.3	250.3	186.7	188.5	197.7	193.8		196.0	195.5
H-C _{γ1} -C ₈₁ -O ₈₁	302.3	309.1	197.2	196.7	198.9	196.1		205.5	199.5
C _β -H-H-C _{γ2} -C ₈₂	150.6	146.1	52.7	48.7	62.1	64.1		66.3	64.4
H-C _{γ2} -C ₈₂ -O ₈₂	56.4	304.1	43.1	40.5	272.9	277.2		272.6	274.0
C ₁₀ -C _α -O ₁ -H	61.8	44.9	66.7	66.8	64.1	55.3		57.2	46.2
C _{γ1} -C ₈₁ -O ₈₁ -H	179.9	175.4	310.7	309.3	309.8	307.0		307.5	306.4
C _{γ2} -C ₈₂ -O ₈₂ -H	179.6	177.0	180.2	171.6	189.6	183.4		178.0	180.0

1) MINOR GROOVE BINDING.

Table I shows that the affinities of both 1 and 2 are significantly larger for the $d(ATAT)_2$ oligomer than for the $d(GCGC)_2$ oligomer. The energy differences, favouring the former, amount to 22.1 and 14 kcal/mole for 1 and 2 respectively. The corresponding differences in term of ΔE are 20.5 and 15.8 kcal/mole, respectively. Such a strong destabilization of the minor groove complexes of $d(GCGC)_2$ with respect to those of $d(ATAT)_2$ is due to the repulsive electrostatic effect exerted on the cationic side chain of 1 and 2 by the 2-amino group of the guanines, a situation also formerly encountered with the corresponding intercalation complexes of daunomycin [7] and adriamycin [8]. This is translated principally at the level of the E_{MTP} contribution and it is significant to observe that the corresponding differences of the E_{MTP} values are of 21.7 and 16.3 kcal/mole for 1 and 2 respectively, close to the corresponding differences in ΔE 's. It is interesting to note also that, owing to the increased oligonucleotide - intercalator interatomic distances entailed by this effect, the short-range repulsion contribution, E_{rep} , is actually smaller in the $d(GCGC)_2$ complex than in the $d(ATAT)_2$ complex.

Figure 3 : Representation of the minor groove complexes of $d(ATAT)_2$ with :
(a) 1 ; (b) 2.

- Binding to $d(ATAT)_2$

The affinity of 1 for $d(ATAT)_2$ is larger than that of 2 by an amount of 2.3 kcal/mole. This preference is imposed by the intermolecular term ΔE , favouring 1 over 2 by an amount of 5.2 kcal/mole, but opposed by a less favourable intercalator conformational energy $8E_{conf}$ for 1 than for 2, of 2.9 kcal/mole. The preference for 1 in terms of ΔE stems predominantly from the E_{MTP} term, which is 3.8 kcal/mole more favourable for 1 than for 2.

The complexes of $d(ATAT)_2$ with 1 and 2 are represented in figures 3a and 3b respectively.

The most significant interatomic interactions and distances are as follows :

- A hydrogen bond between O_2 of thymine T_2 of the intercalation site and one H atom linked to the C_β carbon, immediately preceding the ammonium nitrogen, on which the positive charge is partially delocalized. This H-bond is only slightly shorter (by 0.03 Å) with 1 than with 2.
- Two elongated H-bonds between O_2 of the same T_2 and hydrogens linked to the two C atoms, $C_{\gamma 1}$ and $C_{\gamma 2}$, immediately following the ammonium nitrogen. The corresponding distances are 0.2-0.4 Å shorter with 1 than with 2 implying that the side chain of 1 is located deeper in the core of the minor groove, where the electrostatic potential is the most attractive [27] (see also [28]).
- An H-bond between N_3 of adenine A_3 of the intercalation site and the hydroxyl hydrogen linked to C_α . This H-bond is shorter in the complex of 2 than in the complex of 1. It was proposed in ref. [1] that an H-bond involving this hydroxyl

TABLE III
Values of the binding energies of the phenanthrene derivative 1 and of its des-hydroxy analog (Energies in kcal/mole).

	MINOR GROOVE BINDING				MAJOR GROOVE BINDING	
	$d(ATAT)_2$		$d(GCGC)_2$		$d(GCGC)_2$	
	Phenanthrene derivative 1	Des-OH Phenanthrene derivative	Phenanthrene derivative 1	Des-OH Phenanthrene derivative	Phenanthrene derivative 1	Des-OH Phenanthrene derivative
ΔE	- 263.2	- 261.3	- 242.7	- 234.0	- 250.5	- 246.9
E_{HBP}	- 227.6	- 225.6	- 205.9	- 198.3	- 223.4	- 219.5
E_{pol}	- 14.4	- 13.7	- 12.5	- 11.7	- 13.1	- 12.4
E_{CT}	- 4.2	- 4.2	- 2.9	- 2.4	- 4.2	- 3.9
E_{disp}	- 70.9	- 68.3	- 67.4	- 64.8	- 65.8	- 63.9
E_{rep}	53.9	50.5	45.9	43.3	56.0	52.7
$\delta E_{\text{conf.}}$	20.8	19.4	21.4	24.4	19.5	20.5
$\delta E_{\text{unstack}}$	9.5	9.5	10.5	10.5	10.5	10.5
δE	- 232.9	- 232.4	- 210.8	- 199.1	- 220.5	- 215.9
δ	0.0	0.5	22.1	33.8	12.4	17.0

hydrogen and O_2 of a thymine was an important determinant in conferring AT versus GC specificity. We find instead that this hydroxyl group interacts with N_3 of adenine, the binding to O_2 of thymine being effected as indicated above essentially by the hydrogens of the cationic trimethylammonium moiety of the side chain. In fact, we find that the H-bonds between the C_α hydroxyl hydrogen and the purine N_3 are actually shorter in the $d(GCGC)_2$ complex than in the $d(ATAT)_2$ complex (see below). In fact, independent computations carried out on the example of the appropriate phenanthrene derivatives show furthermore that the removal of the α -hydroxyl group (and its replacement by an H atom) does not affect and actually enhances the A-T versus G-C preference of the compounds (Table III). We must therefore conclude that the presence of this proton donating group at C_α has no bearing for this preference.

- d) An elongated H-bond involving O_1 of the sugar bound to A_3 and one of the terminal hydroxyl hydrogens. This H-bond is shorter in the complex of 1 than in the complex of 2.
- e) Lastly, a somewhat elongated H-bond between O_1 of the same sugar and the other H atom, H_2 , linked to the $C_{\gamma 1}$ carbon.

Figure 4 : Minor groove complexes of $d(GCGC)_2$ with: (a) 1 ; (b) 2.

The conformations of the two intercalator molecules in their $d(ATAT)_2$ complexes are similar (see table II). They are characterized by trans conformations along bonds $C_{12}C_{10}-C_{\alpha}C_{\beta}$, $C_{10}C_{\alpha}-C_{\beta}N$, $C_{\beta}N-C_{\gamma1}C_{\delta1}$, $N-C_{\gamma1}-C_{\delta1}O_{\delta1}$ and $C_{\gamma2}C_{\delta2}-O_{\delta2}H$, g^+ conformations along bonds $C_{\alpha}C_{\beta}-NC_{\gamma1}$, $C_{\beta}N-C_{\gamma2}C_{\delta2}$, $NC_{\gamma2}-C_{\delta2}O_{\delta2}$ and $C_{10}C_{\alpha}-O_1H$ and g^- along bond $C_{\gamma1}C_{\delta1}-O_{\delta1}H$.

Prior to oligonucleotide binding, the intrinsically preferred conformations of 1 and 2 correspond to folded forms, in which the cationic side chain is folded towards the electron-rich phenanthrene and anthracene rings. In the anthracene derivative, the vicinity of the C_4H and C_5H aromatic bonds to the side chain interferes with an optimal folding of the latter towards the aromatic ring. The less favourable δE_{conf} found for 1 as compared to 2 upon elongating the structure to optimize the binding translates thus the inherently greater stabilization of the folded conformation of 1 as compared to that of 2.

- Binding to $d(GCGC)_2$.

Contrary to the situation with $d(ATAT)_2$, it is now the binding of 2 that is favoured over that of 1. Whereas the intermolecular energy term, ΔE , favours 1 only by a small amount (0.5 kcal/mole), the δE_{conf} term is much more unfavourable for 1 than for 2, resulting in an overall preference for the latter of 5.8 kcal/mole.

The complexes of $d(GCGC)_2$ with 1 and 2 are represented in figures 4a and 4b respectively.

They are stabilized by :

- a) A hydrogen bond between O_2 of the cytosine C_2 of the intercalation site and one H atom linked to C_{β} . The repulsive effect exerted by the neighboring 2-amino

group of base G_3 , results in an increased equilibrium distance of $2.16 \text{ \AA}$ as compared to the corresponding $H \dots O_2(T)$ distance of $1.92 \text{ \AA}$ with $d(ATAT)_2$.

- b) An H-bonding interaction between N_3 of base G_3 and the hydroxyl hydrogen linked to $C_{\alpha'}$.
- c) An elongated H-bond between O_1 , of the sugar linked to G_3 and one H linked to $C_{\gamma 1}$.

The elongated H-bonds involving the H-atoms linked to $C_{\gamma 1}$ and $C_{\gamma 2}$ and O_2 of the pyrimidine in the 2 position (T_2), which were present in the $d(ATAT)_2$ complex, have now disappeared, as well as the elongated H-bond involving one terminal hydroxyl hydrogen and O_1 , of the sugar bound to the purine in the 3-position, A_3 or G_3 .

The torsional angles of the two intercalators in their $d(GCGC)_2$ complexes are similar to each other, and also similar to those in the corresponding $d(ATAT)_2$ complexes, with the exception of $NC_{\gamma 2} - C_{82} - O_{82}$ which is now g^- instead of g^+ (see table II).

2) MAJOR GROOVE BINDING TO $d(GCGC)_2$

For both 1 and 2, the interaction energies are significantly larger for major groove than for minor groove binding to $d(GCGC)_2$, although they remain less favourable than for minor groove binding to $d(ATAT)_2$. The global energy differences favouring the latter binding amount to 12.4 kcal/mole for 1 and 5.7 kcal/mole for 2. On the other hand 2 is preferred over 1 for $d(GCGC)_2$ major groove

Figure 5 : Major groove complexes of $d(GCGC)_2$ with: (a) 1 ; (b) 2.

binding by 4.4 kcal/mole. This last more favourable δE value for 2 stems from both the intermolecular and the conformational energy terms.

The major groove complexes of $d(GCGC)_2$ with 1 and 2 are represented in figures 5a et 5b respectively. In the complex of 2, two hydrogen bonds are found involving as proton acceptor one anionic oxygen of a central phosphate group and as proton donors one $C_{\gamma}H$ and one terminal hydroxyl group. An additional, elongated H-bonding interaction involves the C_{α} hydroxyl hydrogen, located above N_4 atom of C_2 . In the complex of 1, only the terminal hydroxyl hydrogen is H-bonded to the central anionic oxygen, in addition to the interaction involving the C_{α} hydroxyl hydrogen and N_4 . The conformational angles of the intercalator molecules with $d(GCGC)_2$ in the major groove remain closely similar to those in the minor groove (see table II).

CONCLUSIONS.

The present study, carried out on the model double-stranded tetranucleotides $d(ATAT)_2$ and $d(GCGC)_2$, accounts satisfactorily for the AT versus GC sequence selectivity of 1 and 2, by means of intermolecular and conformational energy computations. For minor groove binding, the energy difference, δ , favours the AT tetramer over the GC tetramer by an amount to 22 kcal/mole for 1 and 14 kcal/mole for 2. The preference for minor groove binding to $d(ATAT)_2$ versus major groove binding to $d(GCGC)_2$ amounts to 12.4 kcal/mole for 1 and 5.7 kcal/mole for 2. For either minor or major groove binding to $d(GCGC)_2$, a significantly greater AT versus GC sequence preference is found for 1 over 2, in agreement with the experimental results of Wilson et al. [1].

The present computations also account for the larger affinity of 1 than of 2 for AT binding (by 2.3 kcal/mole) and conversely, for the larger affinity of 2 than of 1 for GC binding (of 5.8 and 4.4 kcal/mole for minor and major groove binding, respectively).

In the minor groove of the AT sequence, the complexes are stabilized essentially by H-bonding interactions between the trimethylammonium moiety of the side chain and O_2 of one thymine, T_2 , of the intercalation site and between the C_{α} hydroxyl hydrogen and N_3 of the adenine A_3 of the intercalation site, on the same strand. Interactions involving O_1 , of the sugar linked to A_3 also contribute to the binding, but to a somewhat lesser extent. The preference of 1 over 2 for $d(ATAT)_2$ in terms of ΔE can be ascribed to a more adequate matching of the shape of the phenanthrene derivative in the minor groove, enabling its hydrogen atoms linked to $C_{\gamma 1}$ and $C_{\gamma 2}$, on which the positive charge is partly delocalized, to be closer to O_2

than is the case with 2. This preference is counteracted to some extent by the conformational energy term δE_{conf} , more unfavourable for 1 than for 2. This latter situation is due to an inherently more favourable interaction between the cationic side chain and the aromatic ring, in the folded conformation prior to interaction, in 1 than in 2, since the folding is less hindered by the proximate aromatic C_4H and C_5H bonds in the former.

Upon binding to $\text{d}(\text{GCGC})_2$ in the minor groove, the electrostatic repulsions exerted on the side chain by the 2-amino group of the guanines, result in an increased separation of the side chain from O_2 of the central cytosine C_2 . This levels off the values of ΔE for 1 and 2, so that the outcome of the energy balance is now imposed by δE_{conf} , leading to a distinct preference for 2. In the major groove, the preference for 2 is ascribed to the additional H-bond between one anionic phosphate oxygen and a hydrogen atom bound to the C_γ carbon.

It is gratifying to observe that despite the approximations involved in the present study (neglect of solvation effects, assumption of the same conformation for the two tetranucleotides in their different complexes), these computations enable us to account for the essential features of the experimental results. It must nevertheless be noticed that they disprove the suggestion of the experimentalists as to the origin of the A-T preference supposed by them to be due to an H-bond between the OH at C_α and O_2 of thymine. Following our findings, the presence of this OH is irrelevant to the observed preference.

ACKNOWLEDGEMENTS.

Dr. Kai-Xian Chen expresses his deep gratitude to the Association for International Cancer Research (Great Britain) for a scholarship permitting him to work at the Institut de Biologie Physico-Chimique, Paris. All the authors express their deep thanks to the National Foundation for Cancer Research (Bethesda, U.S.A.) for supporting this research and the CIRCE - CNRS, Orsay (France) for computational facilities.

*Permanent address: Shanghai Institute of Materia Medica, Academia Sinica, Shanghai, China

REFERENCES.

- (1) Wilson, W.D., Wang, Y., Kusuma, S., Chandrasekaran, S., Yang, N. and Boykin, D. (1985) J. Am. Chem. Soc. 107, 4989-4995.
- (2) Muller, W., and Crothers, D. (1968) J. Mol. Biol. 35, 251-290.
- (3) Lown, J., Morgan, A., Yen, S., Wang, Y. and Wilson, W.D. (1985) Biochemistry 24, 4028-4035.
- (4) Strum, J. Schreiber, L. and Daune, M. (1981) Biopolymers 20, 765-785.
- (5) Strum, J. (1982) Biopolymers 21, 1189-1206.

-
- (6) Dasgupta, D. and Goldberg, I. (1986) *Nucleic Acids Res.* 14, 1089-1105.
 - (7) Chen, K.-X., Gresh, N. and Pullman, B. (1985) *J. Biomol. Struct. Dynam.* 3, 445-466.
 - (8) Chen, K.-X., Gresh, N. and Pullman, B. (1986) *Nucleic Acids Res.* 14, 2251-2267.
 - (9) Pack, G. and Loew, G. (1978) *Biochim. Biophys. Acta* 519, 163-172.
 - (10) Ornstein, R. and Rein, R. (1979) *Biopolymers* 18, 1277-1291.
 - (11) Ornstein, R. and Rein, R. (1979) *Biopolymers* 18, 2821-2847.
 - (12) Nuss, M., Marsh, F. and Kollman, P. (1979) *J. Am. Chem. Soc.* 101, 825-833.
 - (13) Gresh, N. and Pullman, B. (1983) *Int. J. Quantum Chem.* 24, 491-507.
 - (14) Chen, K.-X., Gresh, N. and Pullman, B. *Mol. Pharmacol.* (in press).
 - (15) Chen, K.-X., Gresh, N. and Pullman, B. (1986) *Nucleic Acids Res.* 14, 3799-3812.
 - (16) Gresh, N., Claverie, P. and Pullman, A. (1984) *Theoret. Chim. Acta* 66, 1-20.
 - (17) Gresh, N., Claverie, P. and Pullman, A. (1986) *Int. J. Quantum Chem.* 29, 101-118.
 - (18) *Molecular Structures and Dimensions*, Vol. A1, Crystallographic Data Center, University Chemical Laboratory, Cambridge and International Union of Crystallography (1972).
 - (19) Pullman, A., Zakrzewska, K. and Perahia, D. (1979) *Int. J. Quantum Chem.* 16, 395-403.
 - (20) Arnott, S. and Hukins, D. (1973) *J. Mol. Biol.* 81, 93-105.
 - (21) Miller, K. and Pycior, J. (1979) *Biopolymers* 18, 2683-2719.
 - (22) Miller, K. (1981) *Proceedings of the second SUNYA conversation in the discipline biomolecular stereodynamics* Vol. 2, Sarma, R., Ed. Adenine Press, New York 469-486.
 - (23) Miller, K. (1979) *Biopolymers* 18, 959-980.
 - (24) Jones, R., Lanier, A., Keel, R. and Wilson, W. (1980) *Nucleic Acids Res.* 8, 1613-1624.
 - (25) Pullman, B., Gresh, N., Berthod, H. and Pullman, A. (1977) *Theoret. Chim. Acta* 44, 151-163.
 - (26) Fletcher, R. (1972) *FORTRAN Subroutines for Minimization by Quasi-Newton Method* AERE Report R7125.
 - (27) Pullman, A. and Pullman, B. (1981) *Quart. Rev. Biophysics* 14, 289-380.
 - (28) Gresh, N. and Pullman, B. (1984) *Mol. Pharmacol.* 25, 452-458.