Volume 12 Number 16 1984 Nucleic Acids Research

The solvation contribution to the binding energy of DNA with non-intercalating antibiotics

Krystyna Zakrzewska, Richard Lavery 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 29 May 1984; Accepted 17 July 1984

ABSTRACT

The influence of the solvent on the binding energies to DNA of six non-intercalating antibiotics - netropsin, distamycin-3, distamycin-2, SN 18071, berenil and stilbamidine - is evaluated by combining the effect of the first hydration shell with that of bulk water. The first effect is computed by a methodology based on a spherical/point dipole model of water and limited to electrostatic interaction energies. Hydration shells are obtained which are energy optimized with respect to both water-solute and water-water interactions for the complexes and for the isolated DNA oligomers and ligands. The method allows even very large complexes to be studied in reasonable computation times. The second effect is introduced via a cavity treatment. It is shown that if the vacuum interaction energies already predict correctly the preference of the ligands for the minor groove of AT squences of B-DNA, the introduction of the solvation effect is indispensable for reproducing the order of affinity of the ligands and for bringing the values of the complexation energies into close agreement with experimental data.

INTRODUCTION

The interaction of DNA with non-intercalating ligands has been widely studied experimentally for many years (1, 2, 3). Such studies present a multiple interest, since many of this class of ligands possess antitumoral or antiviral activity and, in addition, in the case of peptidic ligands, can provide information upon the nature of general nucleic acid-protein recognition (4).

Several attemps have also been carried out theoretically (5-7) in view of determining the factors responsible for the binding and, especially, for the marked specificity of the great majority of these ligands for the minor groove of the AT sequences of B-DNA. We have explored extensively from this point of view netropsin and SN 18071 (7) (Fig. 1) and have been able to show that this specificity is governed by the steric fit of the ligands to this groove and their strong electrostatic interactions with DNA within it, independently of whether hydrogen bonds may or may not be formed with the purine or pyrimidine bases.

Our previous computations have been carried out for interactions in vacuum and the results obtained indicate that the specificity for the minor groove of A-T sequences is thus an intrinsic characterstic of these interactions. Two

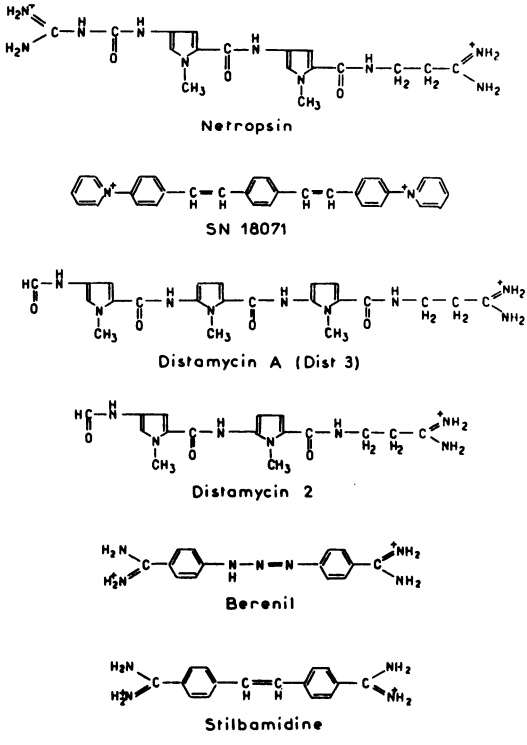

FIGURE 1. Chemical formulas of the ligands studied.

problems remained, however, to be solved. In the first place the interaction energies obtained in vacuum do not necessarily reproduce correctly the order of the ligand affinities observed experimentally. This aspect of the problem will become particularly evident later in this paper when the results of computations for four more compounds will be presented. In the second place, the theoretical interaction energies seem much higher than indicated by the experimental value available for one of the compounds investigated (vide infra).

The most obvious reason for these discrepancies appears to be the omission of water in the theoretical evaluations and this paper is intended to remedy this situation by including explicitly the solvent effect in the computations.

METHODOLOGY

That water must play an important role in complex formation becomes clear when it is recalled that sites on the nucleic acid and on the ligand which can strongly interact with one another are often also sites that can be readily hydrated.

Since the formation of the complex involves at least a partial dehydration of both the nucleic acid and the ligand, one may foresee that the final complexation energy will be a delicate balance of opposing terms. The inclusion of solvent effects therefore seems an inevitable step towards a full understanding of the mechanisms governing DNA-ligand interactions.

The methodology presently proposed for this sake belongs to the category of the mixed "discrete-continuum" procedures (8-10). It involves first a discrete representation of water molecules which are directly bound to the entities in interaction, completed by the introduction of the remaining water as a bulk effect. For the discrete representation, since we are dealing with very large systems, containing of the order of 1000 atoms (a turn of double stranded DNA and a relatively large ligand), we have adopted a "supermolecule" procedure, based on two assumptions : firstly, that the interactions between water molecules and between water and the solute may be approximated as purely electrostatic and, secondly, that the water molecule may be represented as a sphere having at its center a point dipole. This basis enables the construction of a methodology which can be used for macromolecules with several thousand atoms and can provide an energy-optimized first hydration shell with respect to both water-solute and water-water interactions. The effect of the bulk solvent was introduced by a cavity treatment following the method of Halicioglu and Sinanoglu (11).

The procedure involves two principal steps :

(1) In the first one a "hydration envelope" is constructed around the solute molecule by the superposition of spheres centered on each of its non-hydrogen atoms. The radii of these spheres for each type of atom (given in table 1) are chosen as follows : for atoms capable of forming a hydrogen bond with water the radius is the usual hydrogen bond distance ; for other atoms it is the distance for Van der Waals contacts. A single water sphere is then rolled over the entire solute so that its center describes the hydration envelope of the solute and at each position of the center its energy of interaction with the solute is calculated after aligning the water dipole with the local field vector generated by the solute molecule. The interaction energy is then simply given as a product of the water dipole moment by the field intensity. The water sphere is, in reality, not moved continuously over the surface, but rather over a uniformly spaced grid of points on the surface of each atom. All points on the envelope giving rise to interaction energies more negative than a threshold value, taken to be -5 kcal/mole, are considered as possible sites for direct water binding. For smaller absolute values it is assumed that a water molecule would prefer to return to its bulk water environment. Note that although the water sphere is placed in contact with each accessible atom of the solute in turn, its

TABLE 1. Hydration radii for different classes of atoms and their energy correction (see text).

ATOM	HYDRATION RADIUS (Å)	ENERGY CORRECTION (KCAL/MOLE)
C_{aliph}	3.5	0.
$C_{arom.}$	3.2	0.
$-N\begin{matrix}H\\H\end{matrix}$	3.	0.
$-N(H)-$	3.	0.
$-N=$	3.0	-3.
$O=$	2.8	-2.
$-O-$	2.8	-3.
O^-	2.8	0.
P	3.5	0.
Na^+	2.4	0.

position may often correspond to contacts with several atoms and thus "bridging" hydration positions between adjacent interaction sites are by no means excluded.

The field created by the solute molecule is obtained from an expansion of atomic monopoles calculated by a modified Hückel-Del Re procedure, described previously (12), which has been shown to be capable of reproducing satisfactorily the electrostatic properties of molecules obtained by much more detailed multipole expansions.

The resulting energies are now sorted into ascending order and a first guess at the hydration shell is made by placing waters on the envelope following this order, but without occupying sites which would lead to the intersection of water spheres. When as many waters as possible have been added by this procedure, the second step begins.

(2) In order to optimize the total energy of the water-solute system an iterative cycling is now carried out. Considering each water molecule in turn, we search for the position on the hydration envelope which leads to an optimum energy for the system, the water dipole being aligned at each such point with the net field generated by the solute and by the other water molecules. This search is limited, for each water, to points on the hydration envelope belonging to that solute atom on which the water molecule was originally placed. This simplification was introduced

after preliminary tests had shown that free displacement of water molecules onto neighbouring atoms during cycling did not lead to significant changes in the results obtained, but, on the other hand, increased considerably the computation time. After all water molecules have been considered in this way, which we call a "macro-cycle", the procedure is repeated until convergence of the system energy to better than 1 kcal/mole. However, at the end of each macro-cycle, water molecules which become bound to the solute by less than the threshold value (-5 kcal/mole) are eliminated from the system. Care must be taken, nevertheless, not to eliminate at one time any two water molecules which are close together on the hydration envelope. The final hydration energy of the solute is then simply calculated as the sum of the individual water-solute interaction energies.

The methodology is completed by a consideration of the bulk water effects and this was done using the classical Halicioglu-Sinanoglu cavity treatment (11). The surface area of the solute necessary for this step is obtained in a natural way as the total area of the hydration envelope described above. The formula used is :

$$C = K_1 (\bar{\phi}^{1/3}) \; S \gamma \left[1 - \frac{\partial \ln \gamma}{\partial \ln T} - \frac{2}{3} A T \right]$$

where S is the total area of the hydration envelope, γ is the water surface tension and $K_1 (\phi^{-1/3})$ is a constant depending on the volume fraction, T is the absolute temperature, and A denotes a coefficient of thermal expansion appropriate to the cavity, approximated as in (11) by the expansion coefficient of the solvent. The numerical values of the constants A and γ were taken from (11).

In order to introduce with this method the solution correction to the macromolecule-ligand binding energies we first obtain an optimum ligand complex geometry, using for this sake the semi-empirical interaction methodology described previously (7). This procedure calculates the macromolecule-ligand interaction energy E_V as a sum of pairwise additive Lennard-Jones (6-12) terms and of electrostatic interactions between the monopole expansions of the charge density of the macromolecule and of the ligand. Moreover, it allows an easy manipulation of the ligand upon approaching the macromolecule with the aid of force calculations. Ligand flexibility is also permitted.

Once the most favorable resulting interaction energy E_V is obtained, subsequent hydration of the complex by the method described above leads to a complex hydration energy H_{comp} and a cavity energy C_{comp}. We then separate the ligand from the DNA oligomer and hydrate each species separately ; let us term the resulting hydration and cavity energies H_{DNA} and C_{DNA} and H_{lig} and C_{lig}, respectively, the isolated ligand being held in the geometry it adopts within the DNA complex.

The terms described above allow the binding energy in water E_W to be obtained in the following way :

$$E_W = E_V + \Delta H + \Delta C$$

with

$$\Delta H = H_{comp} - (H_{DNA} + H_{lig})$$

$$\Delta C = C_{comp} - (C_{DNA} + C_{lig})$$

where ΔH is the hydration energy change on forming the complex, generally positive since waters are displaced from DNA and from the ligand when they are brought together, and ΔC is the cavity energy change on forming the complex, generally negative since the complex occupies less space than its separated components.

This procedure, intended for application to nucleic acids, was first tested on the hydration of nucleic acid components, bases, phosphates and sugars and for isolated sodium ions. Good correlation was obtained with precedent polyhydration models based on quantum mechanical calculations or electrostatic calculations more precise than those presently employed (9, 10, 13, 14). The practical realisation of the procedure for the bases (15) entails, however, the introduction of two correction terms. The first is a damping of the base fields out of the plane of the base, which arises from the neglect of higher multipoles in the representation of the base electrostatic properties. A simple repulsive energy addition with a dependance of $|2z|$ was found to be appropriate, where z is the displacement (in angstroms) of the point studied perpendicular to the base plane. The second concerns the interaction energy of water dipoles with base atoms which bind water through one of its protons. In these cases a point dipole placed at roughly the position of the water oxygen atom underestimates the binding energy. Table 1 gives the values of the corrections for the various types of atoms. These constant corrections were applied to all water molecules situated on the hydration spheres of the atoms concerned.

The water model adopted is a sphere of radius 1.6 Å containing a point dipole of 2.1 Debye obtained from an SCF calculation in the same basis set as that employed in determining our monopole expansions. This value is a little greater than the experimental dipole moment of water but it does not seem that, in the case of a method employing a point dipole model, there is a necessity or an advantage in using the experimentally measured value.

The polynucleotide models used in our calculations consist of 22 5'-nucleotides forming double stranded oligomers of 11 base pairs. Two homopolymeric base sequences are considered, $(dA)_{11}.(dT)_{11}$ and $(dG)_{11}.(dC)_{11}$

representing one helical turn of poly(dA).poly(dT) and poly(dG).poly(dC), respectively. The geometry of these helices is that of Arnott et al. (16). A simple model screening by sodium cations is employed in which the ion is placed at the bridge position of every phosphate at a distance of 2.15 Å from each anionic oxygen, in the OPO^- plane (17).

The geometries of the ligands where constructed with standard bond lengths and angles (18). In the case of netropsin the complexes with the different grooves and base sequences of DNA employed were those of our previous study (7). For SN 18071, however, the three central rings of the ligand were not maintained co-planar, as in our previous work, but were allowed to turn. Torsions of roughly 30° for the C-C bonds adjacent to the double bonds were found to optimize the ligand binding energy, in accord with neutron diffraction and ab initio theoretical studies of the conformation of the related molecule stilbene (19, 20).

RESULTS AND DISCUSSION

We will present, firstly, a description of the hydration shells found for the DNA oligomers used subsequently as substrates for ligand binding. We will then give a description of the hydration changes occuring after ligand binding, for the example of netropsin, before passing to a comparison of the overall results for the other ligands investigated : SN 18071, distamycin-2, distamycin-3, berenil and stilbamidine (see figure 1).

(1) The hydration of B-DNA helical turns of poly(dA).poly(dT) and poly(dG).-poly(dC) (table 2).

The number of water molecules positioned in the first hydration sphere of a helical turn of poly(dA).poly(dT) is 222. This corresponds to an average of 10 molecules per nucleotide. The experimental results (21 and reference therein) estimate this number as 11-12 waters per nucleotide increasing to 20 waters with an increase in the relative humidity. These 20 waters, however, are not all in direct

TABLE 2. Hydration of model oligonucleotides (H_{DNA} and C_{DNA} in kcal/mole).

SOLUTE		NUMBER OF WATER MOLECULES	H_{DNA}	C_{DNA}
DNA CONFORMATION	DNA SEQUENCE			
B	AT	222	-2589.4	745.1
	GC	217	-2659.1	702.1
A	AT	177	-2513.8	726 2
	GC	174	-2534 1	728 3

contact with DNA, but 8 or 9 of them are bound through other water molecules. Our results therefore seem to be satisfactory on this point. From our 222 water molecules, 169 are positioned on the phosphates and cations, which gives an average of 7-8 water molecules per screened-phosphate. The experimental results obtained in very low humidity give this number as 5-6 per phosphate (22). The water molecules positioned on the phosphates have the average interaction energy of about -13 kcal/mole. There are 37 water molecules positioned on the bases with the total binding energy of -316.1 kcal/mole. From them 14 are situated in the minor groove (on Thy(02) and Adn(N3)) with the average energy of -9.8 kcal/mole and 23 in the major groove (Thy(04) and Adn(N7)) with an average energy of -7.8 kcal/mole. Thus one can see that the minor groove has fewer water molecules than the major groove, but with higher interaction energy per molecule. Finally there are 16 water molecules located by our model on the sugar moieties with an average binding energy of -6.1 kcal/mole. Most of them are situated on the sugar esteric oxygen, in the DNA minor groove.

These results can be qualitatively compared with the results obtained by Dickerson and collaborators (23, 24) by a crystallographic analysis of water positions in a monocrystal of a B-DNA dodecamer (CGCGAATTCGCG). In this study an average of 2 water molecules per base pair were located in the major groove, which correlates well with our findings. In the minor groove a spine of water was discovered in the AT part of the dodecamer formed by water molecules bridging the adenine N3 and thymine 02 atoms. These waters are in turn bridged by a second layer of molecules that confers an approximately tetrahedral bond coordination on the oxygen of the first water shell. In the ideal spine there are two water molecules per base pair but half of these do not interact directly with the nucleic acid atoms. Therefore our result of 14 water molecules for 11 base pairs is in quite good agreement with these bindings. In contrast to our results, Kopka et al. (24) locate only 3 water molecules per phosphate group, which is probably connected with the high lability of phosphate oxygens rendering their localization in the crystal difficult. Another interesting point to which these authors draw attention is the role of sugar 0'1 atoms, which according to them play an important role in the hydration of the minor groove. It is interesting to note that in our model the sugar 01' atoms are also rather strongly hydrated.

In the case of a helical turn of poly(dG).poly(dC) 217 water molecules are located in our first hydration shell (table 2). Out of these 157 are hydrating the phosphates and cations with an average energy of -13.4 kcal/mole. 44 water molecules interact with the bases, from these, 30 are located in the major groove (Gua(O6), Gua(N7) and Cyt(N4)) with an average energy varying between -13.7

kcal/mole for Gua(O6), -9.0 kcal/mole for Gua(N7) and much weaker -5.3 kcal/mole for Cyt(N4). The GC minor groove is much less hydrated with 14 water molecules located on Gua(N3) and Cyt(O2). They have, however, relatively strong average interaction energies, -10.3 kcal/mole for Gua(N3) and -8.8 kcal/mole for Cyt(O2). Finally, there are 16 water molecules located on the sugars, mainly in the minor groove in the vicinity of esteric oxygens, with an average energy of -6.8 kcal/mole. The comparison of our results with those of Kopka et al. (24) is difficult because the GC base pairs are positioned at the ends of the dodecamer studied, where intermolecular packing occludes the minor groove for the first two or three base pairs from each end of the helix. The major groove of these base pairs is also perturbed by spermine. In consequence it is not surprising that the strong hydration of Gua(N7) which we find was not seen in the dodecamer studied.

The following conclusions may be drawn from a comparison of the hydration schemes of the two model oligomers studied. Firstly, the hydration of the phosphate-sugar backbone is very similar in both cases as concerns the number of water molecules interacting and their average energy. Secondly, and in contrast, there are big differences in the hydration energetics of the minor and major grooves of the two sequences. In the case of poly(dA).poly(dT) model both grooves are rather similar, with total hydration energies of -136.7 kcal/mole for the minor groove and -179.4 kcal/mole for the major groove. In the case of poly(dG).-poly(dC) model, however, the major groove is much more hydrated than the minor groove, the total hydration energies being -308.8 kcal/mole and -139.9 kcal/-mole, respectively. This result can be correlated with our earlier calculations concerning the electrostatic fields on the surface envelope of B-DNA, which also showed that the strongest groove fields were associated with the GC major groove and, also, that the difference in maximum fields between the major and minor grooves was much more pronounced for the GC sequence than for the AT sequence (25).

(2) An example of the effect of complex formation on the hydration scheme : the case of netropsin.

When netropsin is bound to the minor groove of the $(dA)_{11}.(dT)_{11}$ oligomer, a total of 209 water molecules are located in the first hydration shell of the complex. The screened-phosphates of DNA are found to have lost 18 waters compared to the free oligomer, but the average interaction energy per water molecules is the same (-13.4 kcal/mole). On the contrary, the sugars now bind 5 more waters although there is a slight decrease in their average binding energy (-7.0 kcal/mole). This shift of few water molecules from the phosphates to the sugars can be interpreted as due to a change in the electrostatic field distribution upon

introducing the charged ligand (+2) into the minor groove. The number of water molecules located in the major groove remains unchanged after complexation, but their average interaction energy is reduced by 1.0 kcal/mole, being now -6.9 kcal/mole. The reduction in electrostatic field in the major groove, which may be deduced from this binding is in line with our previous observation (26) of a reduction of the electrostatic potential in the major groove upon the binding of positively charged ligands in the minor groove.

In the minor groove itself the number of water molecules bound to the base atoms is reduced by 4, with the slight increase in the average binding energy to -10.2 kcal/mole. In addition to those located on the atoms of DNA, there are 4 water molecules bound to netropsin carbonyl oxygens, which point outwards from the groove. In contrast, the charged ends of netropsin are completely dehydrated. Netropsin, which when isolated binds 23 water molecules, can thus retain only very few of them once it enters the groove of the nucleic acid.

Detailed description of the situation found with the other ligands will be given separately. A general remark which can be made, however, from this example, is that the changes in hydration occuring upon complexation can be quite subtle. The introduction of a ligand into a groove of DNA does not simply displace water molecules due to the resulting steric hindrances, it may also displace or, at

TABLE 3. Hydration of DNA-ligand complexes and of the separated ligand (H_{comp}, C_{comp}, H_{lig} and C_{lig} are given in kcal/mole).

LIGAND	DNA BINDING			HYDRATION OF COMPLEX			HYDRATION OF LIGAND		
	Conformation	Sequence	Groove	Number of water molecule	H_{comp}	C_{comp}	Number of water molecule	H_{lig}	C_{lig}
Netropsin	B	AT	Min.	209	-2456.8	708.3	23	-248.1	122.3
			Maj.	208	-2535.9	769.2	25	-268.7	122.6
		GC	Min.	202	-2498.5	677.7	23	-248.1	122.3
			Maj.	216	-2600.4	745.0	27	-277.5	122.6
	A	AT	Min.	182	-2523.1	776.5	24	-258.2	121.6
			Maj.	166	-2447.8	711.4	25	-261.7	121.6
		GC	Min.	186	-2642.7	785.4	24	-260.7	121.7
			Maj.	168	-2488.0	725.5	25	-261.7	140.8
SN 18071	B	AT	Min.	213	-2524.1	742.9	20	-149.3	124.8
		GC	Min.	213	-2620.0	731.2	20	-149.3	124.8
Distamycin-3	B	AT	Min.	215	-2531.7	718.7	22	-207.8	137.3
		GC	Min.	209	-2609.4	700.9	19	-183.6	137.3
Distamycin-2	B	AT	Min.	213	-2536.1	726.7	15	-159.1	107.3
		GC	Min.	210	-2619.0	712.0	18	-178.6	106.9
Berenil	B	AT	Min.	210	-2492.9	732.5	16	-174.0	82.4
Stilbamidine	B	AT	Min.	213	-2524.1	742.9	20	-149.3	124.8

least, weaken the binding of rather distant water molecules due to its electrostatic influence and it may similarly enhance binding at other sites.

(3) The effect of solvation on the energetics of binding.

Table 3 contains the details of the hydration and cavity energies of the various complexes and of the isolated ligands. Table 4 gives the overall hydration and cavity energy changes upon complexation and also the vacuum ligand binding energies, which together allow the final energy of ligand binding in aqueous solution to be calculated.

Let us first consider this last component to the overall energy. To the previously established values of the in vacuum interaction energy for netropsin and SN 18071 we have now added results for four more compounds. In all cases in which the computation were carried out for both AT and GC sequences, the results confirm the preference for binding to the minor groove of the AT sequence of B-DNA which appears thus to be an intrinsic property of these ligands. We have considered this to be also the case for berenil and stilbamidine, for which we have therefore only evaluated the binding energies corresponding to this interaction.

We may next consider the comparison of the strength of the complexes formed with DNA, for the preferred binding sites in the poly(dA).poly(dT) minor groove of the six ligands. From table 4 it may be seen that the order of binding in the absence of water is :

Netropsin > Berenil > Stilbamidine > Distamycin-3 > SN 18071 > Distamycin-2.

Now, although exact experimental values for the complexation energies of all these ligands are unknown, an indirect indication of their relative affinities with respect to AT sequences may be obtained from the so-called C_{50} values determined by Baguley and his collaborators (1). These values are defined as the micromolar ligand concentrations required to halve the observed fluorescence of ethidium molecules bound to DNA. Among the compounds investigated here their values range from 0.07 for distamycin A (the greatest affinity) to 1.8 for stilbamidine (the smallest affinity), passing through 0.20 for SN 18071, 0.55 for netropsin and 0.90 for berenil. A comparison with the E_V energies of table 4 shows immediately that no correlation exists between the order of these energies and the order of affinities deducible from the C_{50} values.

We may now consider the effect upon the energetics of binding of the inclusion of the solvent effect. In the first place we may remark a certain number of properties common to all the ligands. Thus, the hydration energy change ΔH upon complexation is always positive, corresponding to the considerable loss of hydration sites when DNA and the ligand are brought together. Similarly, the cavity energy

TABLE 4. Effect of solvation on the energetics of DNA-ligand binding (all values in kcal/mole).

Ligand	DNA			E_V				ΔH	ΔC	E_W
	Conformation	Sequence	Groove	Electrostatic	Lennard Jones	Ligand conform. energy	Total			
Netropsin	B	AT	Min.	-203.5	-42.3	7.6	-238.2	380.7	-159.2	- 16.7
			Maj.	-104.4	-19.5	15.3	-108.7	322.2	- 98.5	115.0
		GC	Min.	-155.3	-34.1	8.2	-181.2	408.7	-147.5	80.0
			Maj.	-138.0	-15.6	17.2	-136.4	336.2	- 80.5	119.3
	A	AT	Min.	- 74.4	-15.6	15.4	- 74.6	248.9	- 71.3	103.0
			Maj.	- 40.2	-22.3	13.9	- 48.7	327.7	-136.4	142.6
		GC	Min.	- 60.5	-14.0	16.0	- 58.6	152.1	- 64.6	28.9
			Maj.	- 53.1	-25.4	13.9	- 64.6	307.8	-143.6	99.6
SN 18071	B	AT	Min.	- 86.4	-31.1	9.8	-107.7	214.6	-127.0	- 20.1
			Maj.	- 49.8	-21.3	9.8	- 61.3	193.0	- 89.0	42.7
		GC	Min.	- 72.2	-25.8	9.8	- 88.2	188.4	- 96.4	3.8
			Maj.	- 65.6	-21.3	9.8	- 77.1	147.8	- 65.1	5.6
Distamycin-3	B	AT	Min.	- 95.9	-48.8	15.3	-129.4	265.5	-163.7	-27.6
		GC	Min.	- 77.8	-42.4	18.0	-102.2	233.3	-139.3	- 8.2
Distamycin-2	B	AT	Min.	- 94.1	-36.7	16.6	-114.2	212.4	-125.7	-27.5
		GC	Min.	- 75.6	-30.9	18.5	- 88.0	218.7	- 97.8	32.9
Berenil	B	AT	Min.	-174.3	-25.2	15.2	-184.3	270.5	- 95.0	- 8.8
Stilbamidine	B	AT	Min.	-143.5	-26.5	5.2	-164.8	249.0	- 89.2	- 5.0

change upon complexation ΔC is always negative, since the complex has a smaller surface area than the sum of the isolated species. However, the essential result for the present study appears when the hydration and cavity variations, ΔH and ΔC, are added to the result for E_V. The order of the binding energies, in water, E_W now becomes :

Distamycin-3 ≈ Distamycin-2 > SN 18071 > Netropsin > Berenil > Stilbamidine.

It is immediately obvious that this time the order correlates completely with the affinity order indicated by the C_{50} values (the C_{50} value for distamycin-2 is not known). Moreover, it may also be noted that the values of the binding energies are reduced from the order of 1-2 hundred (s) kcal/mole to only tens of kcal/mole in water, values much more in line with what may be expected experimentally. Indeed, in the case of netropsin an experimental binding enthalpy has been obtained (-9.2 kcal/mole, reference 27) which is at least of the order of magnitude of our calculated value of -16.7 kcal/mole. Thus, the introduction of the solvent effect, while not affecting the specificity features of the interaction as found in vacuum, seems indispensable and sucessfull for reproducing the quantitative aspects of the order of affinities and the values of the binding energies.

It may be now interesting to reconsider, in the light of our results, the explanations for the binding strengths of different non-intercalating ligands which are generally put forward. Preeminent among these is the role attributed to

TABLE 5. A qualitative comparison of the ligands studied.

LIGAND	TERMINAL AMIDINIUM GROUPS	POTENTIAL HYDROGEN DONATING GROUPS ALONG THE CENTRAL PART OF THE LIGAND	NET CHARGE
Netropsin	2	4	+ 2
SN 18071	0	0	+ 2
Distamycin-3	1	4	+ 1
Distamycin-2	1	3	+ 1
Berenil	2	1(*)	+ 2
Stilbamidine	2	0	+ 2

(*) This is the NH group of the central diazoamino linkage. The H of this group, situated on the convexe side of the molecule points, however, in the complex, towards the exterior and is thus not engaged in binding to the substrate (5).

hydrogen bond formation between the ligand and sites of DNA. We present in table 5 a qualitative comparison of the ligands studied. If we consider their hydrogen bond forming capabilities and compare them with the vacuum binding energies that we have obtained, a certain correlation appears : for example, netropsin would in vacuum bind considerably stronger than SN 18071, the latter ligand being incapable of forming hydrogen bonds. In fact, a second factor modulates this order, namely the charge on the ligand. The fact that berenil and stilbamidine are doubly charged leads to their stronger binding in vacuum than that of distamycins (singly charged), even though the former can form fewer hydrogen bonds. Similarly, SN 18071 binds more strongly than distamycin-2, despite the considerable hydrogen forming possibilities of the latter ligand. On the other hand, if we pass to the final binding results, in water, there seems to be almost no correlation with the ability of hydrogen bond formation. The present study shows rather that no single property of the ligands allows an easy prediction of their binding order. Hydrogen bond formation certainly plays a role, but so do the charges on the ligand, its size and its flexibility. It is a subtle balance of all these features that leads to the final binding energy.

We may make a final remark concerning the possibility of the binding of the ligands investigated to the minor groove of the GC sequences. The values of E_W in table 4 suggest that this should be feasible for distamycin-3. The corresponding high C_{50} value does not support this conclusion, but more recent experimental studies by Zimmer et al. (28) have shown that distamycin-3 can indeed bind to poly(dG).poly(dC). It is also interesting to note that the latter study concluded that distamycin-2 could not bind to this sequence, which is again in correspondance with our theoretical predictions.

CONCLUSIONS

We have proposed a method for studying the hydration of DNA oligomer-ligand complexes which requires only reasonable computation expense. Its application to a range of non-intercalating ligands has led to a very encouraging improvement of the correlation with experimental results for binding in solution. Comparisons of both the specificity of given ligands for different grooves and different base sequences of DNA and of the relative binding energies of different ligands are well reproduced.

Several points of general interest concerning complexation in solution have come to light as a result of this study. Firstly, the change in the hydration of DNA as a result of ligand binding does not correspond to a simple steric displacement of water. The electrostatic effects of the ligand must be considered, especially when it is charged, and these effects can be quite long range. Secondly,

It seems that no single property of the ligand, such as its hydrogen bond forming capabilities, can be used to predict its binding strength. Lastly, the final complexation energy in water appears to be a very subtle balance of many effects. Strong ligand binding in "vacuum" need not imply strong binding in solution since the dehydration of DNA and of the ligand are of considerable importance as is the reduction of the cavity consequent to complexation.

ACKNOWLEDGMENT

This work was supported by the National Foundation for Cancer Research (Bethesda, M.D., USA) to which the authors wish to express their thanks.

REFERENCES

(1) Baguley, B.C. (1982) Molecular and Cellular Biochemistry 43, 167-181
(2) Zimmer, Ch. (1975) Progress Nucl. Acid. Res. and Mol. Biol. 15, 285-318
(3) Zimmer, Ch. (1984) in Specificity in Biological Interactions (C. Chagas and B. Pullman Eds), The Vatican Press and Reidel Publishing Company, in press.
(4) Gursky, G.V., Tumanyan, V.G., Zasedatelev, A.S., Zhuze, A.L., Grokhovsky, S.L. and Gottikh, B.P. (1977) Nucleic Acid-Protein Recognition, Ed. H.J. Vogel, Academic Press, 189-217.
(5) Gresh, N. and Pullman, B., Molecular Pharmacology, in press.
(6) Gresh, N. and Pullman, B. (1984) Theoret. Chim. Acta 64, 383-395
(7) Zakrzewska, K., Lavery, R. and Pullman, B. (1983) Nucl. Acid Res. 11, 8825-8839
(8) Pullman, A. (1976) in The New World of Quantum Chemistry, (B. Pullman and R. Parr Eds.), Reidel Publishing Company, 149-187.
(9) Pullman, A. (1980) in Quantum Theory of Chemical Reactions, (R. Daudel, A. Pullman, L. Salem and A. Veillard Eds.), Reidel Publishing Company, 1-24.
(10) Claverie, P., Daudey, J.P., Langlet, J., Pullman, B. and Piazzola, D. (1978) J. Phys. Chem. 82, 405-418.
(11) Halicioglu, T. and Sinanoglu, O. (1969) Ann. N.Y. Acad. Sci. 158, 308-317.
(12) Lavery, R. Zakrzewska, K. and Pullman, A., J. Comput. Chem., in press.
(13) Pullman, A., Pullman, B. and Berthod, H. (1978) Theoret. Chim. Acta 47, 175-192
(14) Pullman, B., Pullman, A. and Berthod, H. (1978) Int. J. Quant. Chem., Quantum Biol. Symp. 5, 79-90
(15) Pullman, B. Miertus, S. and Perahia, D. (1979) Theoret. Chim. Acta 50, 317-325
(16) Arnott, S., Chandrasekaran, R., Birdsall, D.L., Leslie, A.G.W. and Ratliff, R.L. (1980) Nature (London) 283, 743-745.
(17) Berthod, H. and Pullman, A. (1965) J. Chim. Phys. 62, 942-946.
(18) Pople, J.A., Beveridge, D.L. (1970) Approximate Molecular Orbital Theory, Mc Graw-Hill.
(19) Traettberg, M., Frantsen, E.B., Mijlhoff, R.C. and Hoekstra, A. (1975) J. Mol. Struct. 26, 57-68.
(20) Wolf, A., Schmidtke, H.H. and Knop, J.V. (1978) Theoret. Chim. Acta 48, 37-45.
(21) Saenger, W. (1983) Principles of Nucleic Acid Structure, C.G. Cantor Eds.
(22) Tunis, M.J.B. and Hearst, J.E. (1965) Biopolymers 3, 57-68

(23) Drew, H.R. and Dickerson, R.E. (1981) J. Mol. Biol. 151, 535-556
(24) Kopka, M.L., Fratini, A.V., Drew, H.R. and Dickerson, R.E. (1983) J. Mol. Biol. 163, 129-146
(25) Lavery, R., Pullman, A. and Pullman, B. (1982) Theoret. Chim. Acta 62, 93-106
(26) Zakrzewska, K. and Pullman, B. (1983) Nucl. Acids Res. 11, 8841-8860
(27) Marky, L.A., Blumenfeld, L.S. and Breslauer, K.J. (1983) Nucl. Acids Res. 11, 2857-2870
(28) Zimmer, Ch., Luck, G., Birch-Hirschfeld, E., Weiss, R., Arcamone, F. and Guschlbauer, W. (1983) Biochem. Biophys. Acta 741, 15-22