Journal of Computer-Aided Molecular Design, **13:** 123–137, 1999.
KLUWER/ESCOM

Major versus minor groove DNA binding of a bisarginylporphyrin hybrid molecule: A molecular mechanics investigation

Nohad Gresh[a,*] & Martine Perrée-Fauvet[b]
[a]*Laboratoire de Pharmacochimie Moléculaire et Structurale, CNRS-URA 1500, INSERM U266, Université Paris 5, 4 Avenue de l'Observatoire, F-75270 Paris Cedex 06, France;* [b]*Laboratoire de Chimie Bioorganique et Bioinorganique, CNRS-URA 1384, Université Paris 11, F-91405 Orsay Cedex, France*

Received 24 May 1998; Accepted 18 August 1998

Key words: DNA intercalating ligand, energy minimisation, nucleic acid–oligopeptide interactions

Summary

On the basis of theoretical computations, we have recently synthesised [Perrée-Fauvet, M. and Gresh, N., Tetrahedron Lett., 36 (1995) 4227] a bisarginyl conjugate of a tricationic porphyrin (BAP), designed to target, in the major groove of DNA, the d(GGC GCC)$_2$ sequence which is part of the primary binding site of the HIV-1 retrovirus site [Wain-Hobson, S. et al., Cell, 40 (1985) 9]. In the theoretical model, the chromophore intercalates at the central d(CpG)$_2$ step and each of the arginyl arms targets O_6/N_7belonging to guanine bases flanking the intercalation site. Recent IR and UV-visible spectroscopic studies have confirmed the essential features of these theoretical predictions [Mohammadi, S. et al., Biochemistry, 37 (1998) 6165]. In the present study, we compare the energies of competing intercalation modes of BAP to several double-stranded oligonucleotides, according to whether one, two or three *N*-methylpyridinium rings project into the major groove. Correspondingly, three minor groove binding modes were considered, the arginyl arms now targeting N_3, O_2 sites belonging to the purine or pyrimidine bases flanking the intercalation site. This investigation has shown that: (i) in both the major and minor grooves, the best-bound complexes have the three *N*-methylpyridinium rings in the groove opposite to that of the phenyl group bearing the arginyl arms; (ii) major groove binding is preferred over minor groove binding by a significant energy (29 kcal/mol); and (iii) the best-bound sequence in the major groove is d(GGC GCC)$_2$ with two successive guanines upstream from the intercalation. On the other hand, due to the flexibility of the arginyl arms, other GC-rich sequences have close binding energies, two of them being less stable than it by less than 8 kcal/mol. These results serve as the basis for the design of derivatives of BAP with enhanced sequence selectivities in the major groove.

Introduction

The majority of protein–DNA interactions occur in the major groove and several of these involve hydrogen bonds between the amino acid side chains and well-defined sites of the bases, namely O_6/N_7 for guanine, N_6/N_7 for adenine, N_4 for cytosine, and O_4 and the 5-methyl for thymine [1–3]. The onset of such interactions, earlier put forth by Seeman et al. [4] and Hélène [5], was highlighted in numerous recent high-resolution X-ray [6–9] and NMR [10] studies of protein–DNA interactions. The knowledge gained by these structural studies recently allowed the design and selection of mutant Zn-fingers having modified DNA-binding properties and, for this class of proteins, the derivation of a 'code' for their interactions with DNA [11–18]. The de novo design of high-affinity, sequence-selective DNA-binding ligands could have important pharmacological consequences. This was the incentive for several endeavours in our laboratories in which we used intercalators as anchoring units at a d(CpG)$_2$ step, linked to oligopeptides of different se-

*To whom correspondence should be addressed.
Supplementary material available from the authors: three tables listing the intermolecular ligand–DNA H-bonds of the complexes of BAP in the major groove in modes B and C, and in the minor groove.

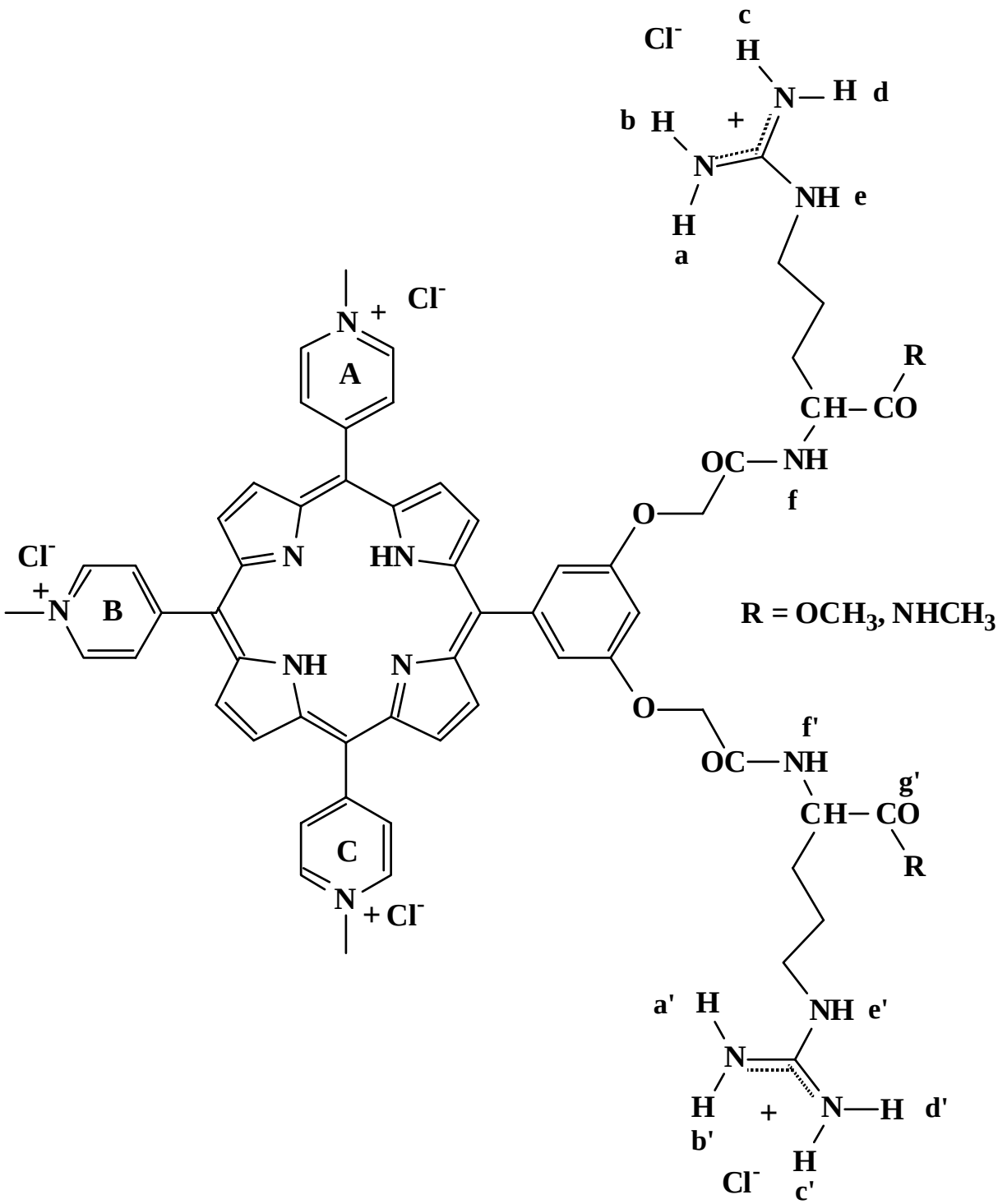

Figure 1. Molecular structure of BAP.

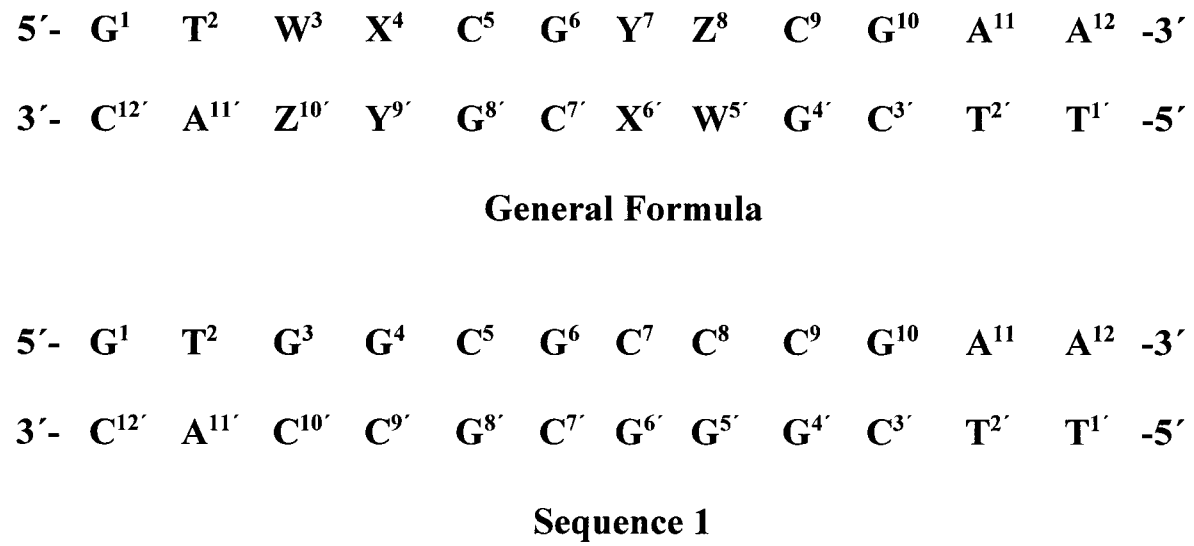

Figure 2. Base numbering of the double-stranded DNA dodecamers investigated for the binding of BAP. Upper line: the 'Watson' strand along the 5′→3′ direction; lower line: the 'Crick' strand along the 3′→5′ direction. General formula, sequence **1**: sequence d(GTGGCGCCCGAA)·d(TTCGGGCGCCAC), part of the PBS of HIV-1.

quences, lengths and chiralities, targeting the purines flanking the intercalation site [19–23]. Among these, a bisarginyl conjugate of a tricationic porphyrin (BAP) (Figure 1, R = OCH_3), designed to target the palindromic d(GGC GCC)$_2$ sequence of oncogenes [24–28] and of the primary binding site (PBS) of the HIV-1 retrovirus [29,30] was recently synthesised on the basis of our theoretical computations and its binding mode to DNA investigated by means of FTIR and UV-visible spectroscopy [31]. These experimental studies confirmed the principal predictions of our computations. We discussed in Reference 31 the rationale for the selection of the tetracationic *meso*-tetrakis(*N*-methyl-4-pyridiniumyl)porphyrin (H_2TMPyP-4) ring as an intercalator [32–34] leading to the design of BAP. We investigate here in detail the different competing modes of BAP binding to several double-stranded DNA dodecamers in the major groove as well as in the minor groove. The BAP structure investigated in this theoretical study is represented in Figure 1 (R = $NHCH_3$). The structural and energetical results will be used as a basis for the design of BAP derivatives spanning a greater (n > 6) number of DNA bases.

To our knowledge, apart from the experimental studies carried out jointly with molecular modelling [21,31], recognition of extended DNA sequences in the major groove by alternative approaches remains virtually limited to the category of synthetic antisense oligonucleotides [35–39], peptide nucleic acids [40], covalently bound oligopeptides [41] and peptide complexes of iron terpyridyl [42] and rhodium phenanthroline [43]. The sequences of these oligopeptides are generally closely related to those of the active site of the transcriptional activator proteins that they simulate. In contrast, most antitumour drugs, whether intercalating [44,45] or non-intercalating [46,47], bind DNA in the *minor* groove. This is also the case for *virtually all* synthetic sequence-specific oligopeptides, whether non-intercalating [48–51] or covalently bound to an intercalator [52–57].

Computational procedure

As in our earlier studies on the interactions of intercalator-oligopeptide conjugates with DNA [19–23,58,59], the computations were performed with the JUMNA molecular mechanics procedure [60,61]. The standard calibration of the method was used. As discussed in the original papers, solvation effects were implicitly accounted for by the use of a sigmoidal dielectric function, and, to account for screening effects, the phosphodiester group has a net charge of −0.5. We followed a similar strategy as previously described [23,59]; namely: (i) using JUMNA, creation of an intercalation site at the d(CpG)$_2$ site of the oligonucleotide with an interplanar separation of 6.8 Å; (ii) manual docking of the ligand using the BioSym graphics software (BioSym Technologies,

Table 1. Binding energies (kcal/mol) of bisarginylporphyrin BAP in mode A with sequences d(GTWXC GYZCGAA)·d(TTCGWXC GYZAC)

Binding of arginyl arms	In the major groove										In the minor groove	
Sequence WXC GYZ	GGC GCC	GTC GAC	GAC GTC	GCC GGC	TGC GCA	AGC GCT	CGC GCG	CCC GGG	TAC GTA		TAC GTA	
E_{int}	−327.0	−317.4	−306.6	−300.0	−299.2	−320.1	−297.3	−301.4	−293.0	−257.0		
ΔE_{DNA}	104.4	98.0	103.6	104.0	97.3	110.2	87.1	79.6	83.8	61.0		
ΔE_{ligand}	14.6	14.6	12.6	16.0	20.7	13.3	14.8	21.1	16.0	17.3		
$\Delta E = E_{int} + \Delta E_{DNA} + \Delta E_{ligand}$	−208.0	−204.8	−190.4	−180.0	−181.2	−196.6	−195.4	−200.7	−193.2	−178.7		
$\delta\Delta E$	0	3.2	17.6	28.0	26.8	11.4	12.6	7.3	14.8	29.3		

San Diego, CA, U.S.A.) (this was done by inserting the porphyrin ring in the intercalation site, while each of the *N*-methylpyridinium rings was located in the appropriate groove according to the binding mode considered, i.e. A, B or C described below); (iii) one round of restrained energy minimisation, in which hydrogen-bonding distances are imposed between the guanidinium group of the arginine side chain and the appropriate sites (O_6 and N_7 of guanines for the major groove recognition of GC-rich sequences, O_2 of thymines and N_3 of adenines for the minor groove of the TACGTA sequence, O_4 of thymines and N_7 of adenines for its major groove) belonging to the two successive base pairs flanking the intercalation site on both DNA strands; and (iv) unconstrained energy minimisation using the restrained minimum-energy position as a starting point. Further details are reported in Reference 23. A 100 ps molecular dynamics simulation was performed on the complex of BAP with the best-bound sequence using the BioSym software and the AMBER force field.

Results and discussion

For both major and minor groove binding, three distinct modes of binding of BAP were compared:

(1) mode A – with the phenyl group (bearing the arginyl arms) in one groove and the three *N*-methylpyridinium rings in the opposite groove;

(2) mode B – with the phenyl group and the two *N*-methylpyridinium rings flanking it in the same groove, while the third *N*-methylpyridinium ring is in the opposite groove;

(3) mode C – with the phenyl group and one of the *N*-methylpyridinium rings flanking it in the same groove, while the other two *N*-methylpyridinium rings are in the opposite groove.

Several DNA double-stranded dodecamers were investigated, all having in common a $d(C^5pG^6)\cdot d(C^{7'}pG^{8'})$ site, where the chromophore is intercalated (Figure 2). Owing to the symmetrical disposition of the two arms with respect to the chromophore plane, our comparison of the binding energies of BAP is limited for modes A and B to six-base-pair palindromic sequences. These sequences all have in common d(GT)·d(AC) at the 5′ end and d(CGAA)·d(TTCG) at the 3′ end. They thus encompass the hexanucleotide $d(WXCGYZ)_2$ palindrome, in which W-Z and X-Y denote complementary bases. W=X=G and Z=Y=C hence correspond to the site of the HIV-1 PBS which

BAP was designed to target preferentially (Figure 2, sequence **1**). In order to evaluate the sensitivity of the binding energy of BAP to the location of the targeted guanines, each of the two guanines of sequence **1**, $G^4/G^{6'}$ (one nucleotide upstream from the intercalation site) and $G^3/G^{5'}$ (two nucleotides upstream from the intercalation site), will be replaced in succession by a thymine, an adenine or a cytosine. This will be followed by asimultaneous replacement of G^3-$G^4/G^{5'}$-$G^{6'}$ either by two cytosines or by a TA dinucleotide sequence.

Major groove binding

Mode A

The results of our computations are reported in Table 1 which lists, for the investigated sequences: the intermolecular DNA–BAP interaction energy E_{int}; the values of the DNA (ΔE_{DNA}) and ligand (ΔE_{ligand}) conformational energy changes with respect to their preferred conformational energies in the absence of interaction; the resulting energy balance ΔE; and the difference (δ) of energy balances with respect to the best ΔE value taken as energy zero.

Table 1 indicates that:

(1) A significant energy preference occurs in favour of major groove binding, the best minor groove complex of BAP being 14.5 kcal/mol less stable than the corresponding major groove complex, and 29.3 kcal/mol less stable than the best major groove complex.

(2) The most stable complex is the one with sequence **1**, the targeted HIV site, having two guanines upstream from the intercalation site.

(3) Because of the flexibility of the arginyl arms, out of the nine sequences considered, there are two sequences whose complexes with BAP come to within 7 kcal/mol of it and two additional sequences, d(AGC GCT)$_2$ and d(CGC GCG)$_2$, whose complexes are less stable than it by $<$13 kcal/mol.

The preferential binding of the arginine side chains to guanine is summarised by the following two rankings of affinities:

– for the sequences having in common a guanine two steps upstream from the intercalation site:

GGC G > GTC G > GAC G > GCC G

$\delta\Delta E$ 0 > 3.2 > 17.6 > 28.0

– for the sequences having in common a guanine one step upstream from the intercalation site:

GGC G > AGC G > CGC G > TGC G

$\delta\Delta E$ 0 > 11.4 > 12.6 > 26.8

Replacement of both guanines by adenine/thymine, as is the case with TAC G, results in a 15 kcal/mol loss of the binding affinities of BAP, major groove binding still remaining favoured over minor groove binding for this sequence.

Mode **A** complexes are stabilised by close (3 Å) electrostatic interactions between the quaternary nitrogen of the two *N*-methylpyridinium rings flanking the phenyl group and the oxygen O_2 of two phosphate groups belonging to both DNA strands (Tables 2a–2b). Thus, in the complex of BAP with sequence **1**, four such interactions can be characterised. The first *N*-methylpyridinium ring interacts with the phosphate groups between C^5 and G^6 and between $C^{10'}$ and $A^{11'}$. The second ring interacts with the phosphate groups between C^8 and G^9 and between $C^{7'}$ and $G^{8'}$.

Consideration of the structures of representative complexes sheds light on the factors governing these energetical orderings. The intermolecular ligand–DNA H-bonds of the complexes are reported in Tables 2a–2b. Additional data and the Cartesian coordinates of the investigated complexes are available upon request.

The complex of BAP with sequence **1**, represented in Figure 3, is stabilised by a much more localised array of symmetrical H-bonds. On each arginine side chain, the amino protons of the guanidinium group, H_c and $H_{c'}$, bind respectively to N_7 of guanines G^3 and $G^{5'}$ which are located two steps upstream from the intercalation site; H_d and $H_{d'}$ are respectively H-bonded to O_6 of guanines G^4 and $G^{6'}$, one step upstream from the intercalation site. Interactions of the arginine guanidinium group with two successive guanines were recently observed in the crystal structure of the complex of the DNA-binding domain of yeast RAP1 with a telomeric DNA [9]. The amino protons H_e/ $H_{e'}$, cis to $H_d/H_{d'}$, and the amide protons of the arginine $H_f/H_{f'}$ are simultaneously H-bonded to O_6 of bases $G^{8'}/G^6$ of the intercalation site on both strands of DNA. In addition, the amino protons H_d and H_e on the first arginyl arm bind respectively to O_6 and N_7 of $G^{8'}$, whereas on the second arginyl arm the amino proton $H_{e'}$ binds to N_7 of $G^{6'}$. An intermolecular interaction is also seen between the amino proton $H_{b'}$ and O_1P belonging to $G^{4'}$, three steps upstream from the intercalation site (see Table 2a).

Replacement of $G^4/G^{6'}$, one nucleotide upstream from the intercalation site. The intermolecular ligand-DNA H-bonds of the complexes are listed in Table 2a.

Table 2a. Binding of bisarginylporphyrin BAP in the major groove in mode A. Values of the ligand-oligonucleotide distances (Å) in the optimised complexes of BAP with sequences:
1: d(GT**GGC GCC**CGAA)·d(TTCG**GGC GCC**AC) (HIV PBS Sequence)
2: d(GT**GTC GAC**CGAA)·d(TTCG**GTC GAC**AC
3: d(GT**GAC GTC**CGAA)·d(TTCG**GAC GTC**AC)
4: d(GT**GCC GGC**CGAA)·d(TTCG**GCC GGC**AC)

Sequence	**1**		**2**		**3**		**4**	
First	H_b-$N_7(G^3)$	2.8	H_b-$N_7(G^3)$	2.8				
arginyl arm	H_c-$N_7(G^3)$	2.0	H_c-$N_7(G^3)$	2.3	H_c-$N_7(G^3)$	2.1	H_c-$N_7(G^3)$	2.2
H-bonds	H_d-$O_6(G^4)$	2.2	H_d-$O_4(T^4)$	2.3				
	H_d-$O_6(G^{8'})$	2.0	H_d-$O_6(G^{8'})$	2.1	H_d-$O_6(G^{8'})$	2.0	H_d-$O_6(G^{9'})$	2.0
	H_e-$N_7(G^{8'})$	2.2	H_e-$N_7(G^{8'})$	2.3	H_e-$N_7(G^{8'})$	2.0	H_e-$N_7(G^{8'})$	2.1
	H_e-$O_6(G^{8'})$	2.3	H_e-$O_6(G^{8'})$	2.2			H_e-$O_6(G^{8'})$	2.6
	H_f-$O_6(G^{8'})$	2.6	H_f-$O_6(G^{8'})$	2.6	H_f-$N_7(G^{8'})$	2.3	H_f-$N_7(G^{8'})$	2.4
Second	$H_{b'}$-O_1P ($G^{4'}$)	1.9			$H_{b'}$-O_1P ($G^{4'}$)	1.9		
arginyl arm	$H_{c'}$-$N_7(G^{5'})$	2.3	$H_{c'}$-$N_7(G^{5'})$	2.0	$H_{c'}$-$N_7(G^{5'})$	2.6	$H_{c'}$-$O_6(G^7)$	2.6
H-bonds	$H_{d'}$-$O_6(G^{6'})$	1.9	$H_{d'}$-$O_4(T^{6'})$	2.0			$H_{d'}$-$N_7(G^6)$	2.2
	$H_{e'}$-$N_7(G^{6'})$	2.6	$H_{d'}$-$O_6(G^6)$	2.5				
	$H_{e'}$-$O_6(G^6)$	2.3	$H_{e'}$-$O_6(G^6)$	2.0	$H_{e'}$-$O_6(G^6)$	2.0	$H_{e'}$-$O_6(G^6)$	2.0
	$H_{f'}$-$O_6(G^6)$	2.1	$H_{f'}$-$O_6(G^6)$	2.2	$H_{f'}$-$O_6(G^6)$	2.0	$H_{f'}$-$O_6(G^6)$	2.1
	$O_{g'}$-$H_2(N_4)$ ($C^{7'}$)	2.0			$O_{g'}$-$H_2(N_4)$ ($C^{7'}$)	2.0		
Ionic	N_A-O_2P (C^5)	3.3	N_A-O_2P (C^5)	3.2	N_A-O_2P (C^5)	3.3	N_A-O_2P (C^5)	3.3
interactions	N_A-O_2P ($C^{10'}$)	3.6	N_A-O_2P ($C^{10'}$)	3.0	N_A-O_2P ($C^{10'}$)	3.6	N_A-O_2P ($C^{10'}$)	3.0
	N_C-O_2P (C^8)	2.9	N_C-O_2P (C^8)	2.9	N_C-O_2P (C^8)	2.9	N_C-O_2P (C^8)	2.8
	N_C-O_2P ($C^{7'}$)	3.0	N_C-O_2P ($C^{7'}$)	3.1	N_C-O_2P ($C^{7'}$)	3.0	N_C-O_2P ($C^{7'}$)	3.0

(1) Replacement by thymine bases incurs a loss of only 3.2 kcal/mol of the total binding energy. Such a small energy separation is due to compensation between a 9.6 kcal/mol weaker ligand–DNA binding energy on the one hand and a 6.4 kcal/mol smaller DNA conformational energy change on the other hand. The same pattern of ligand–DNA H-bonds is found as with the PBS sequence, O_4 of $T^4/T^{6'}$ replacing O_6 of $G^4/G^{6'}$. A slightly more reduced pattern of H-bonds occurs on the second arginyl arm. The limited loss of binding energies incurred by the $G^4 \rightarrow T^4$ and $G^{6'} \rightarrow T^{6'}$ replacements can be related to the results of a survey of the occurrences of arginine–DNA base interactions, indicating O_4 of thymine as being the third most frequently contacted site by arginine after O_6 and N_7 of guanine [17].
(2) Replacement by adenines results in a loss of 17.6 kcal/mol which is dominated by E_{int}. Whereas the H-bond interactions of the arginine arms with G^6 and $G^{8'}$ of the intercalation site, as well as with G^3 and $G^{5'}$, are preserved, A^4 and $A^{6'}$ are not implied in the binding. Such a loss will be seen to be 6.2 kcal/mol more severe than the one resulting from the corresponding replacement of G^3 and $G^{5'}$, one more step upstream (see below). This possibly translates the greater sensitivity of the guanidinium groups to the electrostatic potential of N_7 of G rather than N_7 of A, when these sites are closer to the intercalation site.
(3) Replacement by cytosines results in a larger (28 kcal/mol) loss again dominated by E_{int}. The arginine side chains can now H-bond not only to the guanine bases G^6 and $G^{8'}$ of the intercalation site, but also to G^7 and $G^{9'}$ one step downstream from the intercalation site. However, these interactions do not appear to compensate for the repulsive guanidinium–cytosine interactions.

Replacement of $G^3/G^{5'}$, two nucleotides upstream from the intercalation site. The intermolecular ligand–DNA H-bonds of the complexes are reported in Table 2b.

Table 2b. Binding of bisarginylporphyrin BAP in the major groove in mode A. Values of the ligand-oligonucleotide distances (Å) in the optimised complexes of BAP with sequences:
5: d(GT**TGC GCA**CGAA).d(TTCG**TGC GCA**AC)
6: d(GT**AGC GCT**CGAA).d(TTCG**AGC GCT**AC)
7: d(GT**CGC GCG**CGAA).d(TTCG**CGC GCG**AC)
8: d(GT**CCC GGG**CGAA).d(TTCG**CCC GGG**AC)
9: d(GT**TAC GTA**CGAA).d(TTCG**TAC GTA**AC)

Sequence	**5**		**6**		**7**		**8**		**9**	
First	H_b-$O_4(T^3)$	2.1					H_b-$O_6(G^{9'})$	2.1		
arginyl arm	H_c-$O_4(T^3)$	2.7	H_c-$N_7(A^3)$	2.2	H_c-$N_7(G^4)$	2.1	H_c-$N_7(G^{9'})$	2.0	H_c-$O_4(T^3)$	2.1
H-bonds	H_c-$O_6(G^4)$	2.4	H_d-$O_6(G^4)$	2.2	H_c-$O_6(G^4)$	2.7			H_d-$O_4(T^{9'})$	2.2
	H_d-$O_6(G^{8'})$	2.0	H_d-$O_6(G^{8'})$	2.0	H_d-$O_6(G^{8'})$	2.0	H_d-$N_7(G^{8'})$	2.5	H_d-$O_6(G^{8'})$	2.0
	H_e-$N_7(G^{8'})$	2.0	H_e-$N_7(G^{8'})$	2.3	H_e-$N_7(G^{8'})$	2.4	H_e-$N_7(G^{8'})$	2.2	H_e-$N_7(G^{8'})$	2.1
	H_f-$O_6(G^{8'})$	2.7	H_e-$O_6(G^{8'})$	2.2	H_e-$O_6(G^{8'})$	2.3			H_e-$O_6(G^{8'})$	2.6
	H_f-$N_7(G^{8'})$	2.5	H_f-$O_6(G^{8'})$	2.7	H_f-$N_7(G^{8'})$	2.5	H_f-$N_7(G^{8'})$	2.8		
Second	$H_{b'}$-O_1P ($G^{4'}$)	2.1			$H_{b'}$-$O_6(G^7)$	2.1	$H_{b'}$-$O_4(T^{5'})$	2.1		
arginyl arm	$H_{d'}$-N_7 (G^6)	2.6	$H_{c'}$-$N_7(A^{5'})$	2.1			$H_{c'}$-$O_6(G^7)$	2.1	$H_{c'}$-$O_4(T^7)$	2.3
H-bonds	$H_{d'}$-$O_6(G^6)$	2.0	$H_{d'}$-$O_6(G^{6'})$	1.9	$H_{d'}$-$O_6(G^{6'})$	1.9	$H_{d'}$-$N_7(G^6)$	2.8	$H_{d'}$-$N_7(G^6)$	2.2
	$H_{e'}$-$N_7(G^{6'})$	2.0	$H_{e'}$-$O_6(G^6)$	2.3	$H_{e'}$-$O_6(G^6)$	2.1	$H_{e'}$-$N_7(G^6)$	2.1	$H_{e'}$-$O_6(G^6)$	2.0
	$H_{f'}$-$O_6(G^6)$	2.2	$H_{f'}$-$O_6(G^6)$	2.0	$H_{f'}$-$O_6(G^6)$	2.1	$H_{e'}$-$O_6(G^6)$	2.4	$H_{f'}$-$O_6(G^6)$	2.1
Ionic	$O_{g'}$-$H_2(N_4)$ ($C^{7'}$)	2.0	$O_{g'}$-$H_2(N_4)$ ($C^{7'}$)	2.0			$H_{f'}$-$O_6(G^6)$	2.0	$O_{g'}$-$H_2(N_4)$ ($C^{7'}$)	2.0
interactions	N_A-O_2P (C^5)	3.3	N_A-O_2P (C^5)	3.2	N_A-O_2P (C^5)	3.2	N_A-O_2P (C^5)	3.5	N_A-O_2P (C^5)	3.3
	N_A-O_2P ($A^{10'}$)	3.5	N_A-O_2P ($T^{10'}$)	3.1	N_A-O_2P ($G^{10'}$)	2.9	N_A-O_2P ($G^{10'}$)	3.0	N_A-O_2P ($A^{10'}$)	3.8
	N_C-O_2P (A^8)	2.9	N_C-O_2P (T^8)	3.1	N_C-O_2P (G^8)	3.1	N_C-O_2P (G^8)	3.0	N_C-O_2P (A^8)	2.8
	N_C-O_2P ($C^{7'}$)	3.0	N_C-O_2P ($C^{7'}$)	3.0	N_C-O_2P ($C^{7'}$)	3.0	N_C-O_2P ($C^{7'}$)	3.3	N_C-O_2P ($C^{7'}$)	3.1

(1) Replacement by thymines results in a larger (26.8 kcal/mol) loss of the binding energy due to a less favourable ligand-DNA E_{int}. This is in marked contrast to the situation occurring in the $G^4/G^{6'} \rightarrow T$ replacement. It suggests that, due to the steric hindrance caused by the 5-methyl group, arginine–thymine H-bond interactions may be critically dependent upon T location and O_4 accessibility along the DNA sequence. In the present case, such H-bonds do occur on one arginyl arm but, with respect to the preceding sequence, at the cost of higher ligand conformational energy.
(2) Replacement by adenines results in a 11.4 kcal/mol loss of the total binding energy due to both a less favourable ligand–DNA interaction and a higher DNA conformational energy term. The same pattern of ligand–DNA H-bonds is found as with the PBS sequence, N_7 of $A^3/A^{5'}$ replacing N_7 of $G^3/G^{5'}$, but providing a weaker interaction energy, consistent with the weaker attraction exerted on a positively charged ligand by this site in adenine than in guanine [62].
(3) Replacement by cytosines results in a similar (12.6 kcal/mol) loss of the binding energy. This is due now to a much weaker ligand–DNA interaction energy term (29.7 kcal/mol as compared to the PBS sequence) compensated by a less severe DNA conformational energy term. A much more reduced number of H-bonds can be seen, limited to G^6 and $G^{8'}$ of the intercalation site and to the two immediately upstream guanines G^4 and $G^{6'}$.

Replacement of G^3-$G^4/G^{5'}$-$G^{6'}$, one and two nucleotides upstream from the intercalation site. The intermolecular ligand–DNA H-bonds of the complexes are listed in Table 2b.
(1) Replacement by cytosines to yield sequence **8** with the hexameric d(CCC GGG)$_2$ palindrome results in a 7.3 kcal/mol loss in the total binding energy. For this sequence again, the ligand-DNA intermolecular interaction energy is much less favourable (25.6 kcal/mol) than with the PBS sequence, but this is compensated by a much smaller (24.8 kcal/mol) DNA conforma-

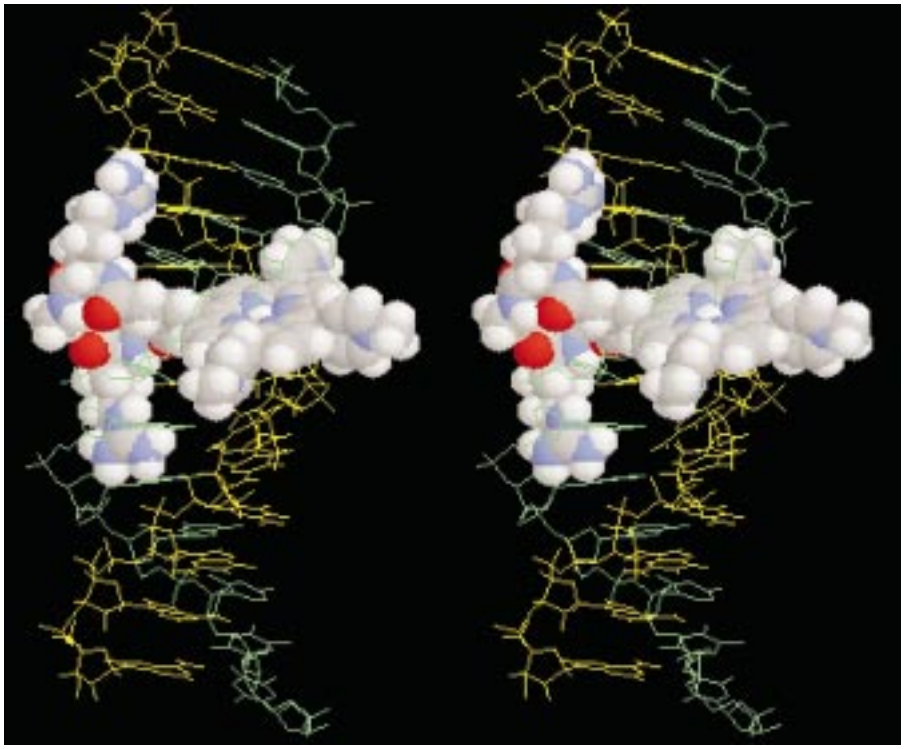

Figure 3. Stereoview of the major groove complex of BAP with sequence d(GTGGCGCCCGAA)·d(TTCGGGCGCCAC) in mode A (hydrogens are omitted for clarity).

tional energy loss, counteracted, however, by a more costly (6.5 kcal/mol) ligand conformational energy. The complex is stabilised by a much more localised array of H-bonds limited to N_7/O_6 of guanines G^6 and $G^{8'}$ of the intercalation site and of guanines G^7 and $G^{9'}$ one step downstream from the intercalation site.
(2) Replacement by T^3-$A^4/T^{5'}$-$A^{6'}$ to yield sequence **9** with the hexameric d(TAC GTA)$_2$ palindrome results in a greater (14.8 kcal/mol) loss of the binding energy. Although the E_{int}value is the weakest one among mode A major groove complexes, it is compensated to some extent by a smaller DNA conformational energy loss. In addition to the H-bonds of arginine with G^6 and $G^{8'}$ of the intercalation site common to the other sequences, the complex is stabilised by H-bonds occurring across the major groove between two amino protons of each arginine side chain respectively bound to O_4 of $T^3/T^{9'}$ and O_4 of $T^{5'}/T^7$.

In order to evaluate the extent to which alternative modes of binding of BAP to sequence **1** could be populated, we performed a sampling of the accessible energy surface by a standard 100 ps molecular dynamics protocol at a temperature of 300 K. Inclusion of an explicit water solvation box on oligonucleotides of this size is beyond our present computational capabilities. A value of $\epsilon = 4r$ was selected to implicitly account for solvation effects. For compatibility purposes with JUMNA, the phosphates were allotted a net charge of -0.5 to account for counterion effects. The AMBER force field was used, since it was recently shown to provide interaction energies closely similar to those computed with the FLEX force field of JUMNA [63]. In order to prevent unrealistically large deformations of DNA due to the absence of explicit counterions and solvent molecules, the DNA conformation was kept rigid but BAP was fully relaxed. Analysis of the 100 structures clearly showed the persistence of BAP–DNA binding involving hydrogen bonds between the arginine side chains and both bases G^3/G^4 and $G^{5'}/G^{6'}$ upstream from the intercalation site and the proximity of the side chains to $G^6/G^{8'}$ of the intercalation site. All the structures conformed to this pattern, and had total energies differing by less than 5 kcal/mol. This is illustrated in Figure 4, in which nine representative structures from the run are superimposed. We have also performed a molecular dynamics run in which the DNA was relaxed, but the distance constraints maintaining the pattern of inter-base hydrogen

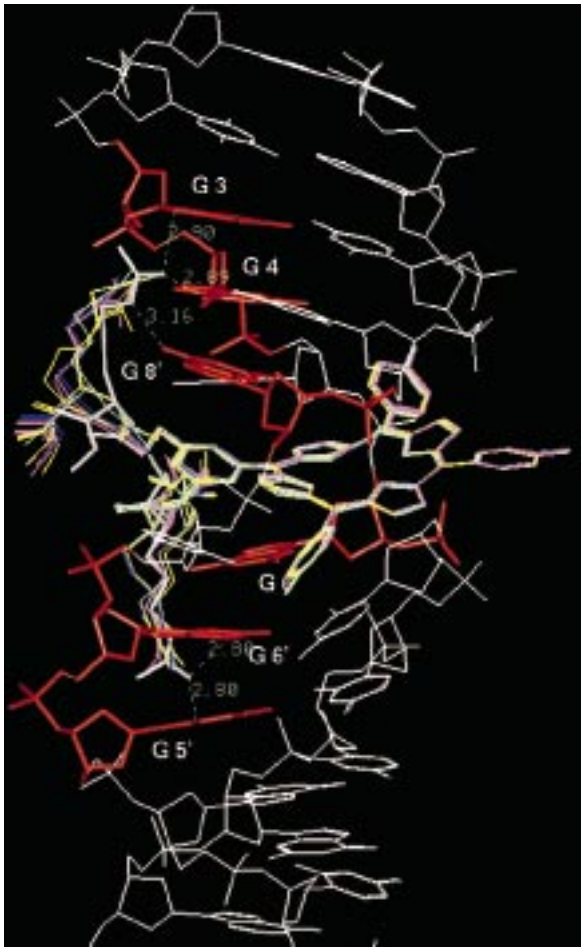

Figure 4. Superimposition of nine complexes of BAP with sequence d(GTGGCGCCCGAA)·d(TTCGGGCGCCAC) in mode A, at the outcome of a 100 ps molecular dynamics run. Bound guanines are represented in red (hydrogens are omitted for clarity).

bonds were imposed. It was observed that even though DNA underwent severe deformations at both its 5′ and 3′ ends, the interactions of the arginine side chains persisted with all but one of the six above-mentioned guanine bases (unpublished). This illustrates the stability of guanine–arginine interactions, but also shows that more exhaustive treatments of environmental effects on DNA conformation are warranted for large-scale molecular dynamics simulations of ligand–DNA complexes.

Mode B

The results of the energy computations are reported in Table 3 and the intermolecular ligand–DNA H-bonds of the complexes are available as supplementary material. Mode B complexes are stabilised by only three close (3.2 Å) electrostatic interactions between the two *N*-methylpyridinium rings flanking the phenyl group and the phosphate groups. The best-bound sequence is again the PBS one, with a stabilisation energy that is 8.8 kcal/mol less favourable than in mode A. One arginine side chain is bound to O_6 and N_7 of G^4 upstream from the intercalation site. The other is bound to O_6 and N_7 of $G^{5'}$ and to O_6 and N_7 of $G^{6'}$ upstream. Additional interactions involve the anionic oxygens O_1P of G^1 on one strand and O_1P of $C^{3'}$ on the other. This complex is represented in Figure 5.

Replacement of $G^4/G^{6'}$, one nucleotide upstream from the intercalation site. Whatever the replacing bases, closely similar structures were obtained after energy minimisation, except that no H-bonds occurred with bases 4 and 6′ nor with their complementary bases 9′ and 7. However, an additional H-bond can be seen which involves G^3. Thus, in contrast with the situation found in mode A, O_4 of $T^4/T^{6'}$ is not implied in the binding when the replacing bases are thymines, and the corresponding sequence is now the least favourably bound among the three, whereas it was the best replacing sequence in mode A. With respect to the latter mode, the ranking of the four sequences with one G two steps upstream is as follows:

	GGC G	>	GAC G	>	GCC G	>	GTC G
δΔE	8.8	>	38.9	>	43.9	>	51.4

Replacement of $G^3/G^{5'}$, two nucleotides upstream from the intercalation site. Replacement by adenine bases gives rise to the smallest loss of binding energies, and the ranking of the four sequences having one G one step upstream is, with respect to mode A:

	GGC G	>	AGC G	>	CGC G	>	TGC G
δΔE	8.8	>	28.7	>	56.2	>	75.1

As in the PBS sequence, H-bonds occur with the guanines $G^4/G^{6'}$, one step upstream. Despite the H-bonds involving O_4 of its $T^3/T^{5'}$ bases and the symmetrical features of its complex with BAP, sequence TGC G ranks last in this scale. Consistent with the results in mode A, this indicates that the limited accessibility of O_4 can adversely affect its affinity for the arginines, unless it is appropriately positioned as was the case with sequence GTC G in such a mode.

Replacement of G^3-$G^4/G^{5'}$-$G^{6'}$, one and two nucleotides upstream from the intercalation site. The sequence with a CCCGGG hexamer has a 54.4 kcal/mol less favourable binding energy than the PBS one

Table 3. Binding energies (kcal/mol) of bisarginylporphyrin BAP in mode B with sequences d(GTWXC GYZCGAA)·d(TTCGWXC GYZAC)

Binding of arginyl arms	In the major groove										In the minor groove	
Sequence WXC GYZ	GGC GCC	GTC GAC	GAC GTC	GCC GGC	TGC GCA	AGC GCT	CGC GCG	CCC GGG	TAC GTA	TAC GTA		
E_{int}	−296.6	−271.3	−283.1	−279.0	−239.9	−278.9	−255.2	−285.5	−247.8	−246.7		
ΔE_{DNA}	82.1	88.4	83.5	87.5	87.7	77.4	79.8	117.7	87.6	67.1		
ΔE_{ligand}	15.3	26.3	30.5	27.4	19.3	22.2	23.6	23.0	23.8	11.6		
$\Delta E = E_{int} + \Delta E_{DNA} + \Delta E_{ligand}$	−199.2	−156.6	−169.1	−164.1	−132.9	−179.3	−151.8	−144.8	−136.4	−168.0		
$\delta\Delta E$	8.8	51.4	38.9	43.9	75.1	28.7	56.2	63.2	71.6	40.0		

in mode B. This is due principally to a higher DNA conformational energy loss. Such a less favourable binding energy occurs in spite of the interactions of the arginine side chains with the guanines which are downstream from the intercalation site. The sequence with a TACGTA hexamer is less favourably bound than the CCCGGG one by 8.4 kcal/mol, and is less favourably bound in mode B than in the minor groove (see below).

The major groove complexes in modes A and B are characterised by large DNA deformation energies. These ΔE_{DNA} values are in the range of those found in a theoretical study of the major groove binding of a proline-rich derivative of 9-aminoacridine [23]. Thus, the BAP complex with the PBS sequence in mode A has the greatest number of hydrogen-bond interactions and is also the one with the highest ΔE_{DNA} value. This could reflect the onset of more extensive DNA conformational rearrangements induced by the arginine arms. We have compared the values of the DNA torsional angles in the representative complexes of BAP with the representative sequences encompassing the GGC GCC, CCC GGG, and TAC GTA hexamers (available upon request). Comparable values were found in the BAP complexes with the latter two sequences. The BAP complex with the PBS sequence had a similar pattern of torsional angles as with its CCC GGG isomer, except for base $G^{8'}$ with ϵ and ζ torsional angles that are g^- and *t*, respectively, instead of the reverse situation with this isomeric sequence. We also note that the sugar pucker pattern is $C_{2'}$ endo for the four bases of the intercalation site, except for the BAP complex with the TAC GTA sequence, in which the sugars of bases G^6 and $G^{8'}$ are $C_{1'}$ exo.

Mode C

An unsymmetrical location of the arginine side chains in the major groove can be anticipated in mode C, due to the location in this groove of one *N*-methylpyridinium ring cis to the phenoxy group, displacing the peptide closer to one strand than to the other one. For that purpose, in addition to the sequence with the hexameric $d(GGCGCC)_2$ palindrome, we will consider a sequence with a d(CCCGCC)·d(GGCGGG) hexamer instead. It has all its four guanines flanking the intercalation site on the same strand. We note that it also corresponds to the *Sp1* oncogenic site [64]. The results, reported in Table 4, show the *Sp1* sequence to be 13 kcal/mol more favoured than the PBS one. It is endowed, however, with a 26.2 kcal/mol less favourable binding energy

Table 4. Binding energies (kcal/mol) of bisarginylporphyrin BAP in the major groove in mode C with sequences:
1: d(GT**GGC GCC**CGAA)·d(TTCG**GGC GCC** AC), (HIV PBS sequence);
10: d(GT**CCC GCC**CGAA)·d(TTCG**GGC GGG** AC)

Sequence	**1**	**10**
E_{int}	−263.5	−252.3
ΔE_{DNA}	70.3	59.8
ΔE_{ligand}	24.4	10.7
$\Delta E = E_{int} + \Delta E_{DNA} + \Delta E_{ligand}$	−168.8	−181.8
$\delta\Delta E$	39.2	26.2

than the best binding mode A with the PBS sequence. It is observed that whereas in mode C, E_{int} favours the PBS site over the *Sp1* one, this is counteracted by both DNA and ligand conformational energies, leading to an overall preference for the latter.

The intermolecular ligand–DNA H-bonds of the complexes are given as supplementary material. In its complex with the PBS site, BAP has one arginine side chain simultaneously bound on one strand to both O_6 and N_7 of G^4 and to the anionic oxygen O_1 belonging to thymine T^2. On the other strand, the second arginine side chain is bound simultaneously to O_6 of $G^{6'}$ and to N_7 of $G^{5'}$. In its complex with the *Sp1* site, BAP has one arginine side chain bound simultaneously to $G^{9'}$ and $G^{10'}$ downstream from the intercalation site and the other one to $G^{6'}$ and $G^{5'}$ upstream. All four guanines flanking the intercalation site are thus H-bonded to BAP. In addition, one amino proton is also H-bonded to N_3 of C^3 on the opposite strand, an interaction resulting in a deformation of the C^3-$G^{10'}$ base pair without, however, disrupting the H-bond between N_3 of C^3 and N_1 of $G^{10'}$. This complex is represented in Figure 6.

As a conclusion for this section on major groove binding, the ranking of the three binding modes is A (one ring in the major groove) > B (three rings) > C (two rings). The PBS-encompassing sequence has the best binding energies for BAP in modes A and B. Replacing G^3-G^4/$G^{5'}$-$G^{6'}$, upstream from the intercalation site, by another base is more costly for G^4 and $G^{6'}$ which are the closest to the intercalation site, except in the case of the $G^4/G^{6'} \rightarrow T$ replacement.

Minor groove binding

Investigation of minor groove binding was restricted to the sequence encompassing the $d(TAC\ GTA)_2$ hexamer, on account of the more attractive electrostatic potential exerted on positive charge(s) in the minor groove by AT-rich sequences than by GC-rich ones [62]. The results of the computations are reported in Tables 1 and 3 for modes A and B along with the major groove ones. Again, mode A is seen to be the most favourable one. The minor groove complex of BAP in mode A is represented in Figure 7. Table 1 shows the binding energy of BAP to be 16.6 kcal/mol less favourable than major groove binding for the sequence considered and 29.3 kcal/mol less favourable than for the best sequence encompassing the PBS site. Comparison with the latter's energy of complexation shows a very great energy difference in terms of E_{int} (70 kcal/mol), compensated by a much smaller (43.4 kcal/mol) DNA conformational energy term. In fact, minor groove binding in mode A gives rise to the smallest DNA conformational energy rearrangement, namely 61 kcal/mol. In mode B, minor groove binding appears to be more favourable than the corresponding major groove binding mode of this $d(TAC\ GTA)_2$ sequence.

The intermolecular ligand–DNA H-bonds of the complexes are given as supplementary material. The minor groove complex of BAP in mode A is stabilised by H-bonds between each arginine side chain and N_3, O_2, and $O_{1'}$ sites with the guanidinium group cross-linking both DNA strands. In mode B, each arginine side chain is H-bonded on one strand to the successive TA bases, one and two steps downstream from the intercalation site. A related pattern was also observed in theoretical studies of the DNA complexes of other intercalator-linked oligopeptides [23]. In addition and similar to the situation found with BAP bound in modes A and B, the minor groove binding of these derivatives led to smaller DNA deformation energies than their major groove binding. The ΔE_{DNA} value is smaller in the case of the minor groove BAP complex, but it is in the range of values found in a theoretical study of the binding of a porphyrin–netropsin derivative to double-stranded B-DNA dodecamers [59]. The intercalation site has the standard mixed sugar pucker pattern. The conformational changes are more localised in the vicinity of the intercalation site than in the BAP major groove complexes. This could provide an explanation for the fact that the ΔE_{DNA} value is smaller than in the latter complexes.

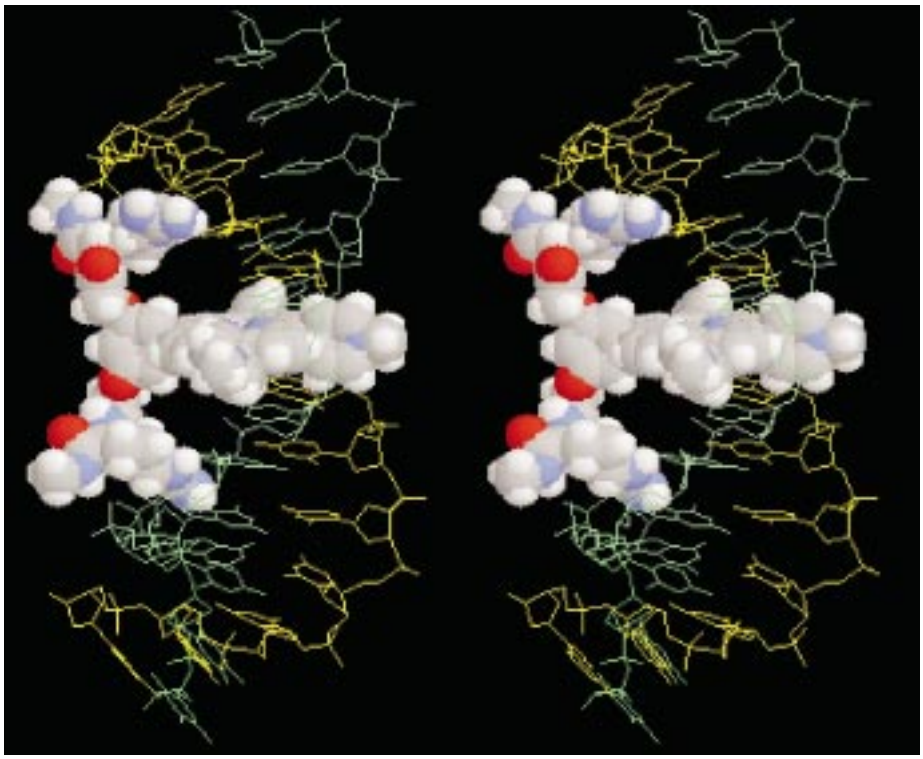

Figure 5. Stereoview of the major groove complex of BAP with sequence d(GTGGCGCCCGAA).d(TTCGGGCGCCAC) in mode B (hydrogens are omitted for clarity).

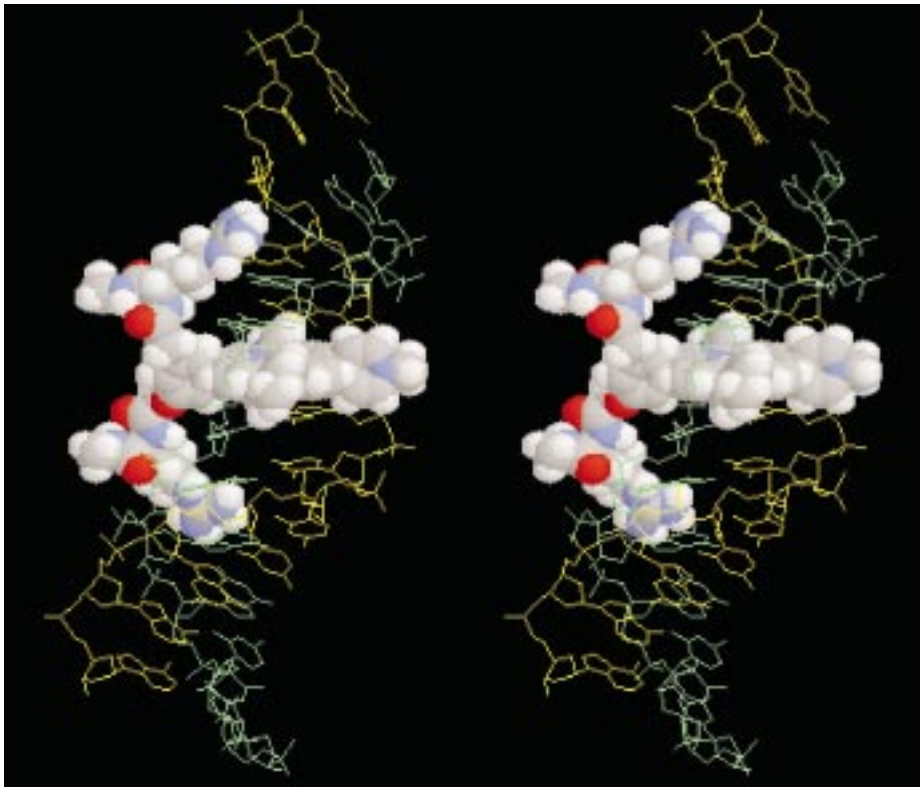

Figure 6. Stereoview of the major groove complex of BAP with sequence d(GTCCCGCCCGAA)·d(TTCGGGCGGGAC) in mode C (hydrogens are omitted for clarity).

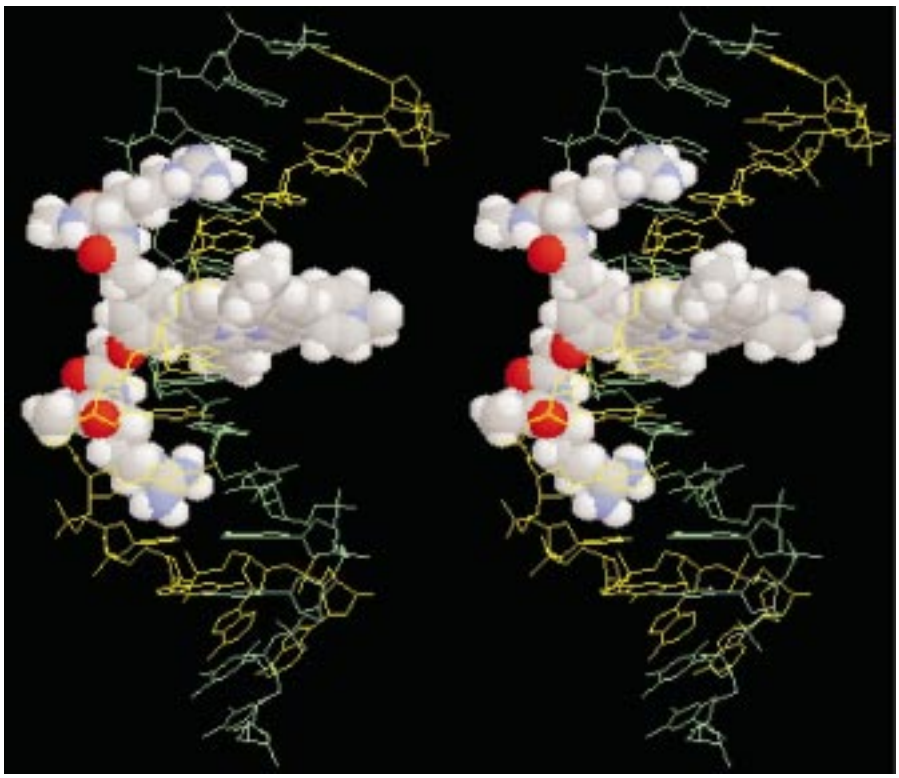

Figure 7. Stereoview of the minor groove complex of BAP with sequence d(GTTACGTACGAA)·d(TTCTAGCGTAAC) in mode A (hydrogens are omitted for clarity).

Conclusions

In this study, we have investigated the modes of binding to representative double-stranded DNA dodecamers of a bisarginyl derivative of a tricationic porphyrin (BAP) which was designed to target in the major groove the d(GGC GCC)$_2$ sequence of the PBS of the HIV-1 retrovirus. The theoretical computations showed the major groove to be preferred over the minor groove by a very large energy difference and the PBS sequence to be the most favourably bound sequence. The best binding mode is mode A, with the three *N*-methylpyridinium rings in one groove and the phenyl ring bearing the arginyl arms in the opposite groove. Mode A of major groove binding was more tolerant to replacements of the guanine bases flanking the intercalation site than alternative mode B, which locates the two *N*-methylpyridinium rings adjacent to the phenyl ring in the major groove. Thus, in mode A, two such replacements entailed losses of only 3 and 7 kcal/mol with respect to the binding energy with the PBS sequence, and three additional replacements entailed losses smaller than 15 kcal/mol. In mode B, all the replacements resulted in energy losses greater than 20 kcal/mol, and five complexes are less stable than the corresponding PBS one by more than 40 kcal/mol. This could be due to the fact that in mode A, the arginine side chains being closer to the intercalation site, invariably interact with the guanines of this site, rendering the complexes less sensitive to replacements of the upstream guanines by the other bases. Binding to the PBS sequence in mode C results in a 39.2 kcal/mol loss of total energy compared to mode A. In fact, because of the unsymmetrical disposition of the *N*-methylpyridinium rings in mode C, the nonpalindromic sequence d(CCCGCC)·d(GGCGGG), which is part of the *Sp1* oncogene [64], was found to be more favourable for BAP binding than the PBS sequence. In its complex with the *Sp1* sequence, the arginine side chains were found to interact with all five guanines on the same strand. The present results also suggest that limiting the flexibility of the arginine side chains should increase the preference of BAP for the PBS sequence, by limiting their accessibility to the competing sequences.

We have confirmed these theoretical investigations by carrying out spectroscopic studies on some of the most representative complexes of BAP [31]. IR spectroscopy gave evidence of the binding of the arginine side chains in the major groove of sequences encompassing the GGC GCC, CCC GGG and TAC GTA hexamers, whereas UV-visible thermal denaturation measurements proved that the binding of BAP occurs preferentially with the double-stranded oligonucleotides encompassing the PBS sequence GGC GCC.

Thus, the de novo designed compound BAP constitutes one of the very rare intercalators which, as for the antitumour drugs mitoxantrone [65,66] and ditercalinium [67,68] and their derivatives, binds DNA in the major groove rather than in the minor groove. Furthermore, preliminary investigations indicated its ability to inhibit the early phases of HIV-1 metabolism at concentrations as low as 5×10^{-8} M (Subra et al., private communication).

As a continuation of the present study, we aim to target a greater number of base pairs in the major groove, up to 12. For that purpose, the present structural results are used as starting points to design higher order homologues. In these, the C-terminal ends of each arginine peptide are extended by oligopeptides having an alternation of hydrophilic and hydrophobic residues and their ends are linked to an intercalator (A. Kossanyi et al., work in progress). Such oligopeptide-linked trisintercalators could be endowed with both increased selectivities and affinities.

Refinements to the present theoretical models will be undertaken in parallel, in which the effects of solvation are explicitly taken into account by means of a Continuum reaction field representation [69,70] using an approach recently followed by us in a study of the binding of metalloprotease inhibitors to the active site of native and mutated thermolysin [71]. It will be instructive to evaluate, in a coherent framework, the contribution of solvation in the energy balances governing the major versus minor groove preferences. This has never been attempted, to our knowledge, in the domain of DNA–ligand interactions.

Acknowledgements

This work was supported by the Agence Nationale de Recherche sur le Sida (ANRS).

References

1. Steitz, T.A., Q. Rev. Biophys., 23 (1990) 205.
2. Brennan, R.G., Curr. Opin. Struct. Biol., 2 (1992) 100.
3. Harrison, S., Nature, 353 (1991) 715.
4. Seeman, N., Rosenberg, J. and Rich, A., Proc. Natl. Acad. Sci. USA, 73 (1976) 804.
5. Hélène, C., FEBS Lett., 74 (1977) 10.
6. Pavletich, N.P. and Pabo, C.O., Science, 252 (1991) 809.
7. Pavletich, N.P. and Pabo, C.O., Science, 261 (1993) 1701.
8. Marmorstein, R., Carey, M., Ptashne, M. and Harrison, S.C., Nature, 356 (1992) 408.
9. Konig, P., Giraldo, R., Chapman, L. and Rhodes, D., Cell, 85 (1996) 125.
10. Baleja, J.D., Marmorstein, R., Harrison, S.C. and Wagner, G., Nature, 356 (1992) 450.
11. Desjarlais, J.R. and Berg, J.M., Proteins Struct. Funct. Genet., 12 (1992) 101.
12. Desjarlais, J.R. and Berg, J.M., Proc. Natl. Acad. Sci. USA, 89 (1992) 7345.
13. Choo, Y. and Klug, A., Proc. Natl. Acad. Sci. USA, 91 (1994) 11163 and 11168.
14. Jamieson, A.C., Kim, S.-H. and Wells, J.A., Biochemistry, 33 (1994) 5689.
15. Lustig, M. and Jernigan, R.L., Nucleic Acids Res., 23 (1995) 4707.
16. Rebar, E.J. and Pabo, C.O., Science, 263 (1994) 671.
17. Suzuki, M., Gerstein, M. and Yagi, N., Nucleic Acids Res., 22 (1994) 3397.
18. Wu, H., Yang, W.R. and Barbas III, C.F., Proc. Natl. Acad. Sci. USA, 92 (1995) 344.
19. Gresh, N. and Kahn, P.H., J. Biomol. Struct. Dyn., 7 (1990) 1141.
20. Gresh, N. and Kahn, P.H., J. Biomol. Struct. Dyn., 8 (1991) 827.
21. Gresh, N., René, B., Hui, X., Barsi, M.-C., Roques, B.P. and Garbay, C., J. Biomol. Struct. Dyn., 12 (1994) 91.
22. Perrée-Fauvet, M. and Gresh, N., Tetrahedron Lett., 36 (1995) 4227.
23. Gresh, N., J. Biomol. Struct. Dyn., 14 (1996) 255.
24. Tsimanis, A., Bichko, V., Dreilina, D., Meldrais, J., Lozha, V., Kukaine, R. and Gren, E., Nucleic Acids Res., 11 (1983) 6079.
25. Hong, F.D., Huang, H.-J.S., To, H., Young, L.-J.S., Oro, A., Bookstein, R., Lee, E.Y.-H.P. and Lee, W.-H., Proc. Natl. Acad. Sci. USA, 86 (1989) 5502.
26. Dvorak, M., Urbanek, P., Bartunek, P., Paces, V., Vlach, J., Pecenka, V., Arnold, L., Travnicek, M. and Riman, J., Nucleic Acids Res., 17 (1989) 5651.
27. Smith, S., Baker, D. and Jardines, L., Biochem. Biophys. Res. Commun., 160 (1989) 1397.
28. Timsit, Y. and Moras, D., J. Mol. Biol., 251 (1995) 629.
29. Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S. and Alizon, M., Cell, 40 (1985) 9.
30. Ratner, L., Haseltine, W., Patarca, R., Livak, K.J., Starcich, B., Josephs, S.F., Doran, E.R., Rafalski, J.A., Whitehorn, E.A., Baumeister, K., Ivanoff, L., Petteway Jr., S.R., Peaerson, M.L., Lautenberger, J.A., Papas, T.S., Ghrayeb, J., Chang, N.T., Gallo, R.C. and Wong-Staal, F., Nature, 313 (1985) 277.
31. Mohammadi, S., Perrée-Fauvet, M., Gresh, N., Hillairet, K. and Taillandier, E., Biochemistry, 37 (1998) 6165.
32. Fiel, R.J., Howard, J.C., Mark, E.H. and Datta Gupta, N., Nucleic Acids Res., 6 (1979) 3093.

33. Pasternack, R.F. and Gibbs, E.J., In Tullius, T. (Ed.) Metal DNA Chemistry, American Chemical Society, Washington, DC, 1989, pp. 59–73.
34. Pasternack, R.F. and Gibbs, E.J., In Sigel, A. and Sigel, H. (Eds.) Metal Ions in Biological Systems, Vol. 33, Marcel Dekker, New York, NY, 1996, pp. 367–397.
35. Moser, H.E. and Dervan, P.B., Science, 238 (1987) 645.
36. Le Doan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J.-L., Thuong, N.T., Lhomme, J. and Hélène, C., Nucleic Acids Res., 15 (1987) 7749.
37. Hélène, C. and Toulmé, J.J., Biochim. Biophys. Acta, 1049 (1990) 99.
38. De Mesmaeker, A., Häner, R., Martin, P. and Moser, H.E., Acc. Chem. Res., 28 (1995) 366.
39. Escudé, C., Nguyen, C.H., Mergny, J.L., Sun, J.S., Bisagni, E., Garestier, T. and Hélène, C., J. Am. Chem. Soc., 117 (1995) 10212.
40. Hyrup, B. and Nielsen, P.E., Bioorg. Med. Chem., 4 (1996) 5.
41. Park, C., Campbell, J.L. and Goddard III, W.A., J. Am. Chem. Soc., 117 (1995) 6287.
42. Cuenoud, B. and Schepartz, A., Science, 259 (1993) 510.
43. Terbrueggen, R.H. and Barton, J.K., Biochemistry, 34 (1995) 8227.
44. Arcamone, F., Doxorubicin, Anticancer Antibiotics, Academic Press, New York, NY, 1981.
45. Lee, S.H. and Goldberg, I.H., Biochemistry, 28 (1989) 1019.
46. Zein, N., Poncin, M., Nilakantan, R. and Ellestad, G.A., Science, 244 (1989) 697.
47. Ho, S.N., Boyer, S.H., Schreiber, S.L., Danishefshky, S.J. and Crabtree, G.R., Proc. Natl. Acad. Sci. USA, 91 (1994) 9203.
48. Zimmer, C. and Wahnert, U., Prog. Biophys. Mol. Biol., 41 (1986) 31.
49. Mrksich, M., Wade, W.S., Dwyer, T.J., Geierstanger, B.H., Wemmer, D.E. and Dervan, P.B., Proc. Natl. Acad. Sci. USA, 89 (1992) 7586.
50. Dwyer, T.J., Geierstanger, B.H., Bathini, Y., Lown, J.W. and Wemmer, D.E., J. Am. Chem. Soc., 114 (1992) 5911.
51. Nikolaev, V.A., Grokhovsky, S.L., Surovaya, A.N. and Gursky, G.V., J. Biomol. Struct. Dyn., 14 (1996) 31.
52. Bailly, C., Helbecque, N., Hénichart, J.-P., Colson, P., Houssier, C., Rao, K.E., Shea, R.G. and Lown, J.W., J. Mol. Recog., 3 (1990) 26.
53. Bailly, C. and Hénichart, J.-P., Bioconj. Chem., 2 (1991) 379.
54. Anneheim-Herbelin, G., Perrée-Fauvet, M., Gaudemer, A., Hélissey, P., Giorgi-Renault, S. and Gresh, N., Tetrahedron Lett., 34 (1993) 7263.
55. Goulaouic, H., Carteau, S., Subra, F., Mouscadet, J.-F., Auclair, C. and Sun, J.-S., Biochemistry, 33 (1994) 1412.
56. Bourdouxhe-Housiaux, C., Colson, P., Houssier, C., Waring, M.J. and Bailly, C., Biochemistry, 35 (1996) 4251.
57. Hélissey, P., Bailly, C., Vishwakarma, J.N., Auclair, C., Waring, M.J. and Giorgi-Renault, S., Anti-Cancer Drug Des., 11 (1996) 527.
58. Hui, X. and Gresh, N., J. Biomol. Struct. Dyn., 11 (1993) 333.
59. Perrée-Fauvet, M. and Gresh, N., J. Biomol. Struct. Dyn., 11 (1994) 1203.
60. Lavery, R., In Wells, R.D. and Harvey, S.C. (Eds.), Unusual DNA Structures, Springer, New York, NY, 1988, pp. 189–206.
61. Lavery, R., Adv. Comput. Biol., 1 (1994) 69.
62. Pullman, B. and Pullman, A., Q. Rev. Biophys., 14 (1981) 289.
63. Flatters, D., Zakrzewska, K. and Lavery, R., J. Comput. Chem., 18 (1997) 1043.
64. Jones, K.A., Kadonaga, J.T., Luciw, P.A. and Tjian, R., Science, 232 (1986) 755.
65. Murdock, K.C., Child, R.C., Fabio, P.F., Angier, R.B., Wallace, R.E., Durr, F.E. and Citarella, R.V., J. Med. Chem., 22 (1979) 1024.
66. Wallace, R.E., Murdock, K.C., Angier, R.B. and Durr, F.E., Cancer Res., 39 (1979) 1570.
67. Garbay-Jaureguiberry, C., Esnault, C., Delepierre, M., Laugaa, P., Laalami, S., Le Pecq, J.-B. and Roques, B.P., Drugs Exp. Clin. Res., XIII (1987) 353.
68. Garbay-Jaureguiberry, C., Barsi, M.-C., Jacquemin-Sablon, A., Le Pecq, J.-B. and Roques, B. P., J. Med. Chem., 35 (1992) 72.
69. Langlet, J., Claverie, P., Caillet, J. and Pullman, A., J. Phys. Chem., 92 (1988) 1631.
70. Langlet, J., Gresh, N. and Giessner-Prettre, C., Biopolymers, 36 (1995) 765.
71. Gresh, N. and Roques, B.P., Biopolymers, 47 (1997) 145.