Docking Studies on DNA-Ligand Interactions: Building and Application of a Protocol To Identify the Binding Mode

Clarisse G. Ricci and Paulo A. Netz*

Instituto de Química, Universidade Federal do Rio Grande do Sul, av. Bento Gonçalves 9500, 91501-970, Porto Alegre, Brazil

Received April 30, 2009

Despite DNA being an important target for several drugs, most of the docking programs are validated only for proteins and their ligands. In this paper, we used AutoDock 4.0 to perform self-dockings and cross dockings between two DNA ligands (a minor groove binder and an intercalator) and four distinct receptors: 1) crystallographic DNA without intercalation gap; 2) crystallographic DNA with intercalation gap; 3) canonical B-DNA; and 4) modified B-DNA with intercalation gap. Besides being efficient in self-dockings, AutoDock is capable of correctly identifying two of the main DNA binding modes with the condition that the target possesses an artificial intercalation gap. Therefore, we suggest a default protocol to identify DNA binding modes which uses a modified canonical DNA (with gap) as receptor. This protocol was applied to dock two different Tröger bases to DNA and the predicted binding modes agree with those suggested, yet not established, by experimental data. We also applied the protocol to dock aflatoxin B_1 *exo*-8,9-epoxide, and the results are in complete agreement with experimental data from the literature. We propose that this approach can be used to investigate other ligands whose binding mode to DNA remains unknown, yielding a suitable starting point for further theoretical studies such as molecular dynamics simulations.

INTRODUCTION

As the number of biological structures in data banks rapidly increases, molecular docking is becoming an important approach to evaluate or even to elucidate the interaction between potential ligands and their macromolecular targets.[1,2] It has been shown that several docking methods described so far can correctly reproduce the binding modes of cocrystallized ligands to their protein targets (self-dockings), but none of them can be considered a universally applicable method.[2]

While there is a large number of studies reporting protein−ligand docking, much less research has been reported on docking of ligands to nucleic acids[3−13] despite DNA being an important molecular target for a wide number of antibiotics and antitumor drugs.[14] Disappointingly, most of the scoring functions have been parametrized exclusively with protein−ligands sets, and the programs have been validated only for proteins and their ligands.[2,15] As well-known, nucleic acids differ from proteins due to unique structural features such as high density charge and helix chiral geometry. Also, nucleic acids do not present a single and well-defined binding site (as occur with most of the proteins) and offer more solvent exposed binding pockets.[2] As a consequence, this leads to the question of whether docking programs validated for proteins can also produce reasonable results in ligand-DNA docking. This issue has been recently approached by Holt et al.,[13] who has shown that AutoDock and Surflex can accurately reproduce the crystal structure of several ligands (minor groove binders and intercalators) bound to DNA, within a resolution of approximately 2 Å. Certainly, this had shed new light on the potential of automated docking programs for virtual screening of DNA binding agents. However, although self-dockings (i.e., using the original crystallographic target) are considered useful as a first indication of docking accuracy, they have proved to provide little information about accuracy in real drug discovery.[2] Indeed, the employment of docking techniques to elucidate unknown DNA binding mechanisms - *without* any conclusive previous experimental data - remains a challenge.

The known fact that DNA is not rigid but rather a very flexible molecule that can assume several structurally distinct isoforms[16−18] combined with the fact that most of current automated docking programs do not take into account target flexibility leads to a so far insufficiently explored issue: when it comes to the docking of ligands whose binding mode to DNA remains unknown, which oligomer conformation should be used as target? It is likely that several structural features of the chosen oligomer conformation will strongly impact docking performance, especially when ligand and target interact through an induced-fit mechanism.

In general, interaction of small molecules with DNA occurs in two distinct ways: intercalation between base pairs or groove recognition.[19] Although the major groove offers more H-bonding donor and acceptor sites, it has been suggested that it provides a much larger and shallow binding pocket for small molecule binding than the minor groove. Indeed, most of the small ligands bind to the minor groove, while the major groove is the preferential binding site for proteins and peptides.[19,20]

Among the main DNA binding modes, intercalation is the most common way through which small and rigid aromatic molecules recognize DNA. Most of the classical and simple

* Corrresponding author fax: (++55)(51)33087304; e-mail: netz@iq.ufrgs.br.

10.1021/ci9001537 CCC: $40.75
Published on Web 08/06/2009

intercalators such as ethidium bromide and proflavine do not possess sequence selectivity since the binding of these ligands to DNA depends basically on π-stacking and stabilizing electrostatic interactions.[20,21] It is also well-known that intercalation imposes structural alterations to DNA in order to open an intercalation gap between two consecutive base pairs.[19] Therefore, intercalative binding to DNA can be considered as an *induced-fit* mechanism.

On the other hand, minor groove recognition by shape-selective agents like distamycin and netropsin is more similar to a *lock-and-key* mechanism since little or no apparent DNA distortion is observed after binding.[20] Due to their flexibility, these naturally occurring polyamides assume a curved shape that matches DNA topology[20,22] and allows a snug fit of the ligand in the minor groove. It is assumed that this interaction depends on a combination of three main factors, in which van der Waals contacts play a pivotal role assisted by hydrogen bonding and electrostatic interactions.[23,24] Also, netropsin and distamycin have attracted considerable attention due to their selective binding to AT-rich sequences in minor groove, showing that it is possible to achieve sequence-selective binding agents.[22,23,25] Netropsin, for instance, is known to interact strongly with AT-rich sequences, displaying large residence times.[26]

Considering the preference that sequence-specific ligands show for some sequences, it is important to observe that sequence selectivity can arise from i) direct readout of H-bonding pattern, ii) indirect readout of sequence flexibility (as occur with TATA-box Binding Protein[27]), or iii) a combination of both (as occur with polyamides[20]). Therefore, docking methods that do not take into account DNA flexibility are sensitive only to the first type of sequence selectivity, not to the second one. This is the main reason why we consider that, as a primary approach to docking studies of DNA-interacting agents, elucidating the binding mode may be more promising than investigating the existence of preferential binding sites.

In this context, we decided to use the software AutoDock 4.0[28] to perform not only self-dockings but also cross dockings between two ligands (ellipticine and netropsin) and four structurally distinct oligomers (two canonical and two crystallographic) in order to establish a docking protocol able to identify two of the main DNA binding modes: intercalation and minor groove recognition.

We further applied this docking protocol to predict the binding modes of two Tröger bases: a symmetric base derived from proflavine and an asymmetric base derived from proflavine and phenanthroline, each with two optical isomers. The data concerning these molecules in literature remain inconclusive. For both molecules, the (−)-isomer has proved to preferentially bind B-DNA.[29−31] However, although the binding mode is not yet elucidated, spectroscopic and biochemical experiments suggest two different binding modes for these Tröger bases: minor groove binding for the symmetric base[30] and a bimodal binding mode for the asymmetric Tröger base, with proflavine intercalated and phenanthroline residing in minor groove.[31]

Finally, we decided to test the docking protocol with a molecule that binds covalently to DNA and compare the result with that from self-docking. In this way, we chose the aflatoxin B1 *exo*-8,9-epoxide, a potent carcinogen known to intercalate to DNA forming a subsequent covalent bond with a guanine in the major groove.[32,33]

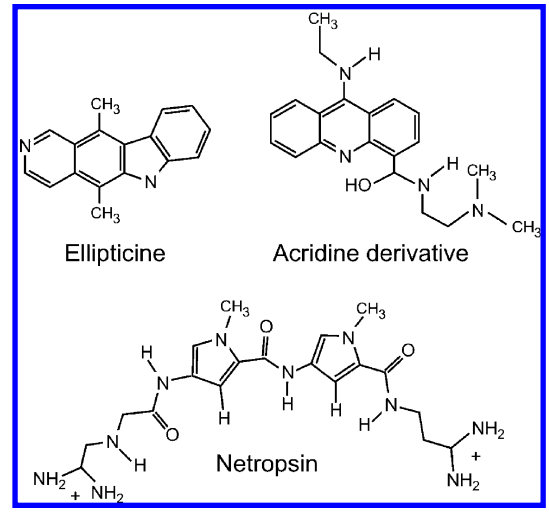

Figure 1. Examples of DNA binding agents.

If consistent with experimental data, we suggest that the tested docking protocol can be used to investigate other ligands whose binding mode to DNA remains unknown, yielding a suitable starting point for further theoretical studies as, for instance, molecular dynamics simulations in conjunction with free energy calculations.[34]

METHODOLOGY

Molecular docking experiments were performed with AutoDock 4.0, a software that uses an empirical scoring function based on the free energy of binding.[28,35] Among the stochastic search algorithms offered by the AutoDock suite, we chose the Lamarckian Genetic Algorithm (LGA) which combines global search (Genetic Algorithm alone) to local search (Solis and Wets algorithm[36]).

Genetic algorithms are based on the evolutionary concept in which the solution to an adaptative problem is spread among a genetic pool. In molecular docking, the 'solution' corresponds to the best binding position for the ligand, and it is represented by a 'chromosome' file containing translation, orientation, and torsion 'genes'. Basically, a genetic algorithm creates a randomly placed population of individuals (ligands) and then applies cycles of genetic operators (mutation and crossover) giving rise to new generations until a suitable solution is achieved.

The 'solutions' are evaluated through their free energy of binding. To achieve faster energy evaluation, AutoDock represents the macromolecule as a tridimensional grid, in which each point stores precalculated affinity potentials for all atom types of the ligand. In this way, AutoDock allows flexibility to the ligand, whereas the macromolecule is kept rigid and fixed during docking.

The Protein Data Bank[37] was searched for ligand-DNA complexes, and two structures were selected: 1Z3F (a hexamer $d(CGATCG)_2$ complexed with ellipticine) and 1DNE (a dodecamer $d(CGCGATATCGCG)_2$ complexed with netropsin). Ellipticine is a typical intercalating agent with antitumor activity,[38] whereas netropsin is a naturally occurring antibiotic that interacts with DNA through minor groove recognition.[39,40] Ellipticine and netropsin structures are shown in Figure 1.

Table 1. Summary of Tested Dockings

docking	number of runs	maximum of energy evaluations	maximum of generations
0	50	10,000,000	270,000
1	50	2,500,000	270,000
2	50	50,000,000	270,000
3	25	10,000,000	270,000
4	100	10,000,000	270,000
5	25	50,000,000	270,000
6	50	10,000,000	27,000
7	50	10,000,000	2,700
8	50	10,000,000	270
9	25	50,000,000	27,000

In a first stage, dockings with ellipticine and netropsin were performed with the oligomers from crystallographic complexes 1Z3F and 1DNE, respectively (self-dockings). After the separation of the coordinates of ligands and DNA, polar and aromatic hydrogens were added with the GROMACS package[41,42] using the GROMOS 53A6 force field,[43] and Gasteiger-Marsili charges[44] were calculated with AutoDock Tools (ADT).[45]

A grid box was created with 96 × 96 × 96 points and a resolution of 0.375 Å, in order to include the *entire* DNA fragment. After the grid box was centered in the macromolecule, grid potential maps were calculated using module AutoGrid 4.0.

Initially, each docking consisted of 50 independent runs, with an initial population of 50 individuals, a maximum number of 10^7 energy evaluations, and a maximum number of 270,000 generations. This was considered *docking zero*. Mutation and crossover were applied to the population at rates of 0.02 and 0.80, respectively. For a local search, it was set to the pseudo Solis and Wets algorithm, with a translational step size of 0.2 Å, an orientational step size of 5.0 degrees, and a torsional step size of 5.0 degrees. To the remaining parameters, default values were applied. Results differing by less than 2.0 Å in root-mean-square deviation (rmsd) were grouped in the same cluster, which is represented by the energetically most favorable conformation belonging to the cluster. Since the grid box enclosed not only the binding site but also the entire DNA fragment, we could not use rmsd as an accuracy criterion. Instead, we opted for a more subjective yet more representative criterion, which was to classify the resulting binding mode by visual inspection as intercalation, minor groove binding, or others (major groove binding, interaction with phosphate groups, etc.). We also evaluated the pattern of hydrogen bonding for the best ranked netropsin complexes using the program VMD,[46] with a donor−acceptor distance cutoff of 3.5 Å and an angle cutoff of 40°. However, as hydrogen bonding also depends strongly on the flexibility of the target, we could not see the characteristic bifurcated hydrogen bonds formed between netropsin and two consecutive DNA bases.[47]

Next, to improve docking performance while reducing computational cost, dockings were carried out with variations in the number of runs, in the maximum number of energy evaluations, or in the maximum number of generations. The remaining parameters were kept as in docking zero. The total of tested dockings is summarized in Table 1, and the results are shown in Figures 2 and 3. These dockings were carried out with netropsin and, instead of ellipticine, an intercalating acridine derivative (PDB: 2GB9[48]) whose structure is also shown in Figure 1. We chose this acridine derivative since it has more degrees of freedom than ellipticine and thus enhances the challenge of docking.

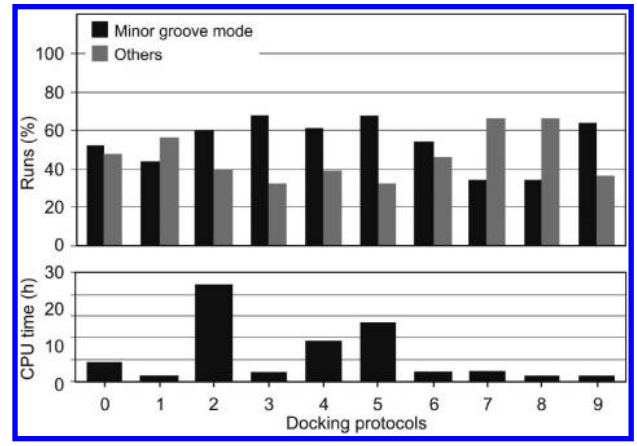

Figure 2. Results from dockings 1−9 (applied to netropsin). Above: bars indicate the percentage of runs resulting in intercalation or other binding modes. Below: CPU time associated with each protocol.

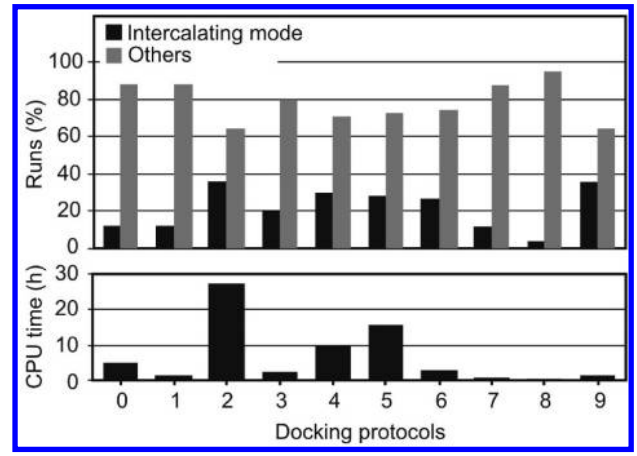

Figure 3. Results from dockings 1−9 (applied to the acridine derivative). Above: bars indicate the percentage of runs resulting in minor groove or other binding modes. Below: CPU time associated with each protocol.

Basically, it was observed that the number of runs has no clear effect upon the predictive skill of docking, although it increases the computational cost in a linear way (see dockings 3, 0, and 4 in Figures 2 and 3). In contrast, docking results are improved as the maximum number of energy evaluation increases. However, computational time increases strongly in this case (see dockings 1, 0, and 2 in Figures 2 and 3). The maximum number of generations correlates weakly with computational cost, but no prediction improvement was observed when changing from 27,000 to 270,000 generations (compare dockings 8, 7, 6, and 0 in Figures 2 and 3). Therefore, a set of standard docking parameters with the best overall performance was established, consisting of 25 runs, 50 × 10^6 energy evaluations, and 27,000 generations (docking 9). These parameters were kept fixed in further dockings with netropsin and ellipticine.

While most of the docking protocols shown in Figure 2 indicate that minor groove interaction is the most accessed binding mode for netropsin, it can also be observed that most of the docking results shown in Figure 3 point to an incorrect binding mode for the acridine derivative. However, it is

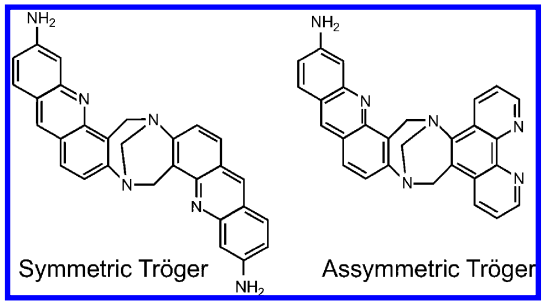

Figure 4. Structure of the Tröger bases.

worthwhile to stress that *qualitatively* the cluster profiles point to the correct binding mode, for in all docking protocols with the acridine derivative the best ranked conformations in terms of binding free energy were those resulting in intercalative binding (not shown).

In a second stage, the crystallographic receptors were replaced by canonical B-DNA with similar sequences, generated with X3DNA.[49] Since structural changes in the macromolecule are not allowed during docking, one canonical oligomer was previously modified to include an "intercalation gap". In this way, Swiss PDB Viewer[50] was used to pose an ellipticine molecule between two base pairs of canonical DNA, parallel to the base rings. After that, the complex was minimized by the steepest descent method, using the GROMACS package with the GROMOS 53A6 force field. The result was a modified B-DNA in which the base pairs flanking ellipticine were separated by 6.50 Å. Ellipticine was removed from the complex, and the modified B-DNA (see Figure 6B/F) was used as a target for docking with ellipticine. The sequences of the X3DNA generated oligomers are d(CGCAATTGCG)$_2$ (without gap) and d(CG-**GC**ATGCCG)$_2$ (with gap, indicated in bold).

We also decided to perform cross dockings: ellipticine was docked with crystallographic and canonical DNA fragments previously used as receptors for netropsin (without intercalation gap), whereas netropsin was docked with crystallographic and modified canonical DNA fragments previously used as receptors for ellipticine (containing intercalation gap).

In this way, we tested four different docking protocols, each one consisting of a docking pair, i.e., the same target oligomer docked with ellipticine and netropsin.

We also applied one of these protocols to dock two Tröger bases to DNA: a symmetric Tröger base derived from proflavine and an asymmetric Tröger base derived from proflavine and phenanthroline. Both structures are shown in Figure 4. Since these are chiral compounds, we constructed both isomers (−)-(R,R) and (+)-(S,S) using GaussView.[51] The geometries were optimized with RHF/6-31G* using Gaussian.[52] After that, polar and aromatic hydrogens were added with GROMACS using the GROMOS 53A6 force field, Gasteiger-Marsili charges were calculated using ADT, and the grid was created as described above, including the entire DNA receptor. The same default docking parameters were used for the Tröger bases, except that we opted for 100 runs instead of 25.

Finally, in order to test if this docking protocol can also be applicable to a molecule that binds covalently to DNA, we chose the carcinogen aflatoxin B1 *exo*-8,9-epoxide, an intercalator which also forms a covalent bond with the N7 position of guanine,[33] probably through a S_N2 mechanism.[32,53] The covalent adduct was obtained from the protein data bank (PDB: 1MKL[33]), and the coordinates of the ligand were separated from the coordinates of DNA of sequence d(ACAT**CG**ATCT)·(AGAT**CG**ATGT). In order to perform the docking using our generic protocol, a canonical oligomer was generated using X3DNA, with a sequence identical to the crystallographic oligomer and then modified to contain an intercalation gap as previously described. The parameters used in dockings of aflatoxin B1 *exo*-8,9-epoxide to DNA were the same as used in the dockings of the Tröger bases.

RESULTS AND DISCUSSION

Among the 25 runs performed in each docking with ellipticine and netropsin, the ten most favorable were analyzed (those presenting the most negative binding free energies). In Table 2 it is reported a summary of the results for ellipticine or netropsin dockings with the four different DNA targets. A/E are dockings with crystallographic DNA original from the ellipticine-DNA complex (with intercalation gap); B/F are dockings with modified canonical B-DNA containing an intercalation gap; C/G are dockings with crystallographic DNA from netropsin-DNA complex (without intercalation gap); and D/H are dockings with canonical B-DNA (without intercalation gap). Thus, A, B, G, and H are direct dockings, while C, D, E, and F are cross-dockings. Among direct dockings, only dockings A and G are self-dockings. Cluster profiles are shown in Figure 5, and the best ranked conformations for each docking set are illustrated in Figure 6.

Direct Dockings. As expected, the applied docking protocol proved to be very efficient to predict binding modes for ellipticine or netropsin in self-dockings (Table 2, dockings A and G). All runs with ellipticine resulted in intercalation mode, with an average binding free energy of −8.71 kcal/mol (Figure 5A), and all runs with netropsin resulted in minor groove recognition, with an average binding free energy of −9.13 kcal/mol and a very favorable best docked conformation (−9.97 kcal/mol) (Figure 5G).

However, when the original crystallographic targets are replaced by canonical DNA of similar sequence, distinct tendencies are observed for each kind of ligand (Table 2, dockings B and H). Ellipticine docking with modified B-DNA (containing an intercalation gap) shows that intercalation represents only 50% of the runs (Figure 5B), although it still corresponds to the preferential binding mode in terms of binding free energy (−8.10 kcal/mol). In contrast, there is no change in the percentage of minor groove recognition when netropsin is docked to canonical B-DNA (Figure 5H) instead of crystallographic DNA (Figure 5G). With respect to binding free energies, the cluster profile from docking H proves to be quite similar to that of docking G, with a very favorable average binding energy of −8.70 kcal/mol.

These results are in agreement with the characteristic mechanisms of each binding mode. Since ellipticine interacts with DNA through an induced-fit mechanism that is not allowed to occur during docking, it is likely that ellipticine docking to DNA will depend strongly on the target selected conformation. On the other hand, since netropsin is a very flexible ligand and its flexibility is taken into account during docking, netropsin can endure structural adaptations in order

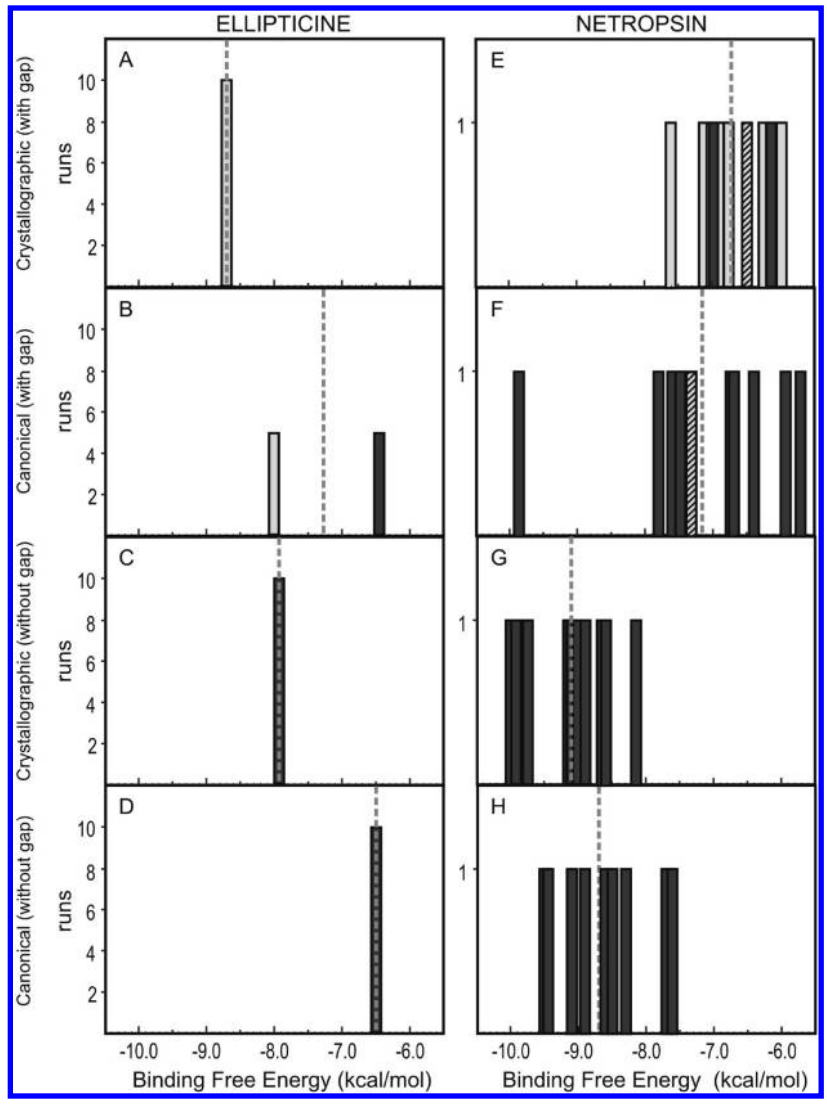

Figure 5. Cluster profiles from dockings A-H. Each cluster is represented by a bar in which the height corresponds to the number of conformations in the cluster, and the color indicates the binding mode: minor groove in black, intercalation in gray, and other in hatched pattern. The dashed line indicates the average binding free energy.

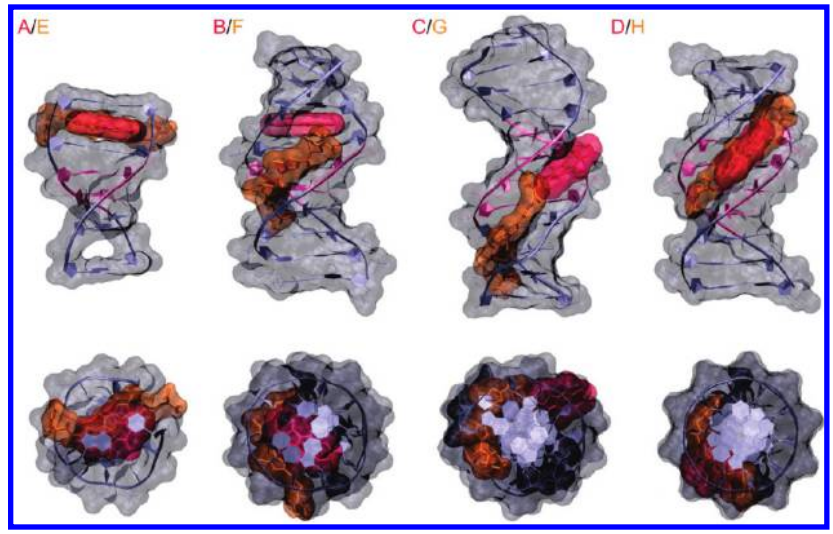

Figure 6. Best docked conformations for dockings A-H. Ellipticine is shown in pink and netropsin in orange. In DNA, CG-rich regions are shown in gray, and AT-rich regions are shown in magenta.

to successfully fit a canonical minor groove, as proved by results from docking H.

Cross Dockings. In cross dockings, the mechanistic features observed for ellipticine and netropsin binding mode

Table 2. Docking Results for Ellipticine (EL) or Netropsin (NE)[c]

docking	ligand	DNA	N_{minor} (%)	N_{int} (%)	ΔG_{best}	$\Delta G_{average}$
A	EL	crystallographic (with gap)[a]	0	100	−8.710 (int)	−8.71
B	EL	canonical (with gap)	50	50	−8.100 (int)	−7.25
C	EL	crystallographic (without gap)[b]	100	0	−7.98 (minor)	−7.98
D	EL	canonical (without gap)	100	0	−6.51 (minor)	−6.51
E	NE	crystallographic (with gap)[a]	20	70	−7.62 (int)	−6.74
F	NE	canonical (with gap)	90	0	−9.85 (minor)	−7.17
G	NE	crystallographic (without gap)[b]	100	0	−9.97 (minor)	−9.13
H	NE	canonical (without gap)	100	0	−9.47 (minor)	−8.70

[a] Original from the ellipticine-DNA complex. [b] Original from the netropsin-DNA complex. [c] N_{minor}, percentage of runs in minor groove mode; N_{int}, percentage of runs in intercalating mode; ΔG_{best}, most favorable binding free energy (kcal/mol); $\Delta G_{average}$, average binding free energy (kcal/mol).

become more evident. As expected, results from dockings C and D (Table 2) clearly show that the applied docking method cannot recognize intercalation binding mode when DNA does not possess a proper intercalating gap. This limitation of all contemporary docking methods that do not take into account receptor flexibility was already noted in the literature.[2,6] Indeed, when crystallographic (Figure 5C) or canonical DNA (Figure 5D) without intercalation gap is used as receptor, all ellipticine docking runs result in minor groove recognition, although with binding free energies not as favorable as the intercalating binding free energies from docking A.

In contrast, netropsin cross dockings E and F show that the minor groove can still be recognized as a possible binding mode for netropsin even when the DNA target presents an intercalation gap (Table 2, Figure 5E,F). However, docking E points to an intercalative binding as the preferential binding mode for netropsin both quantitatively (70% of the runs) and qualitatively (best binding free energy). Even so, it is noteworthy that the binding free energies are far less favorable than those from netropsin dockings G and H and also that the free binding energy of the best docked conformation (−7.62 kcal/mol) is not as negative as those of ellipticine intercalation in dockings A (−8.71 kcal/mol) and B (−8.10 kcal/mol).

In comparison to docking E, docking F presents a very different profile (Figure 5F) for it recognizes minor groove recognition as the preferential and most accessed binding mode for netropsin even in the presence of an intercalation gap. Of ten runs, only one did not result in minor groove recognition, and none of the runs resulted in intercalation as a possible binding mode. The average binding free energy (−7.17 kcal/mol), although not comparable to those of docking G and H, was still more favorable than the average free binding energy from docking E (−6.74 kcal/mol). Also, the minor groove binding from the best ranked run in docking F presented a very favorable binding free energy of −9.85 kcal/mol, comparable to the binding free energies from best docked conformations in dockings G and H and also far more negative than the binding free energy from the best intercalated conformation from docking E (−7.62 kcal/mol).

Conformational Analysis of Target Oligomers. In order to enlighten the docking results presented above, we used X3DNA to evaluate the double helix parameters for the four oligomers used as targets. The results are shown in Table 3.

As can be concluded by the values of twist, crystallographic DNA from ellipticine complex is more unwound (29.43°) when compared to canonical B-DNA (35.73°).

Table 3. Average Values of Structural Double Helix Parameters[d]

DNA receptor	twist (°)	P−P (Å)	rise (Å)
crystallographic (with gap)[b]	29.43	16.4	6.88[a]
canonical (with gap)	34.38	12.4	6.50[a]
crystallographic (without gap)[c]	36.22	11.1	3.42
canonical (without gap)	35.73	11.7	3.35

[a] Value for the base-pair step that contains the intercalation gap. [b] Original from ellipticine-DNA complex. [c] original from netropsin-DNA complex. [d] Twist, twist between two consecutive base pairs; P−P, minor groove distance; rise, distance between two consecutive base pairs. The analysis of these parameters was performed with X3DNA.

Indeed, it is well-known that intercalation not only requires the opening of an intercalation gap but also partially unwinds the double helix.[19,38] This occurs because base pairs are better overlapped in unwound DNA, resulting in increased π-stacking interactions between base rings and aromatic rings of the intercalator.[20] Consequently, reduced twist is probably one of the important features that favor intercalation in dockings A and E. (It is important to note that, although the classical score function of AutoDock does not explicitly take into account molecular polarizability, dockings A and B suggest that π-stacking can be sufficiently well mimetized through van der Walls interactions).

On the other hand, considering that shape selective binders possess a natural curvature compatible to B-DNA, it is likely that netropsin affinity for DNA will be strongly affected by a decrease in double helix twist. Moreover, unwinding of DNA also affects minor groove width, as denoted by P−P distances in Table 3. For shape-selective polyamines, minor groove width is an important structural feature: a narrow minor groove allows snug fit of the ligand, consequently enhancing favorable van der Waals interactions as well as electrostatic attraction between the cationic groups of the ligand and the phosphate groups of DNA.[20] Conversely, the very large minor groove of crystallographic DNA from ellipticine complex (16.4 Å) diminishes the contact surface between netropsin and minor groove walls, decreasing the netropsin-DNA interactions. In other words, these structural features - decreased twist associated with a very large minor groove - are probably the main reason behind the incorrect binding mode predicted in docking E, in which two of netropsin pirrol rings are inserted in the intercalation gap (see Figure 6).

As also revealed by Table 3, the structural parameters for crystallographic and canonical oligomers without an intercalation gap are quite similar. As a result, it is not surprising

that netropsin dockings to these oligomers show similar cluster profiles (dockings G and H). However, the average P−P distances in Table 3 must be considered carefully in order to avoid an oversimplification. It may lead to the conclusion that the crystallographic minor groove (11.1 Å) is only slightly narrower than the canonical minor groove (11.7). However, it is important to remark that minor groove dimensions in crystallographic DNA are not homogeneous as occurs in canonical DNA. As should be expected, the central region ATAT in crystallographic DNA presents a significantly narrow minor groove (10.8 Å) in comparison to the peripheral CG rich (∼13.2 Å) regions in the same oligomer. This decreased accessibility of the central region is very probably the reason why the best docked conformation in self-docking G shows netropsin docked slightly above the ATAT sequence, in a region where the minor groove is not so narrow (see Figure 6). In other words, although a narrow minor groove is considered to enhance ligand-DNA interaction in the formed complex, it may not represent the most favorable intermediate geometry for the approximation and fit of the ligand.

On the other hand, probably because canonical DNA from docking H presents uniform helix geometry with a not so narrow minor groove, docking can correctly place netropsin in the central AT-rich region (see Figure 6), which is known to be the preferential sequence for binding also because of the pattern of hydrogen bonding. In summary, it seems that in such rigid target dockings, the importance of hydrogen bonding is limited by the lack of target flexibility and plays only a secondary role in binding site prediction, which turns out to be mainly determined by helix geometry.

Moreover, this bold minor groove snuggling in the ATAT central region from crystallographic DNA is probably the reason behind the relatively favorable binding free energies in ellipticine docking C, since ellipticine is placed exactly in the region where P−P distances reaches the minimum value of 10.8 Å (see Figure 6). When ellipticine is docked to canonical DNA (dockings B and D), the binding free energies from the minor groove recognition are far less favorable than those from intercalation, clearly indicating that the former is not the preferential binding mode for ellipticine but an artifact from docking flexibility limitations.

Regarding to the modified B-DNA, Table 3 shows that it presents an almost canonical value of twist (34.38°), and P−P distances are only slightly larger (12.4 Å) than canonical minor groove distances (11.7 Å). Therefore, modified B-DNA is more similar to canonical B-DNA than to the crystallographic DNA from ellipticine complex, indicating that the artificial mechanism applied to open an intercalation gap in canonical B-DNA implies structural adaptations that are far more subtle than the changes caused by the real interaction with an intercalator. Hence, it is quite understandable that netropsin docks to modified B-DNA in a similar way as it docks to canonical and crystallographic DNA without intercalation gap, as revealed by cluster profiles from dockings F, G, and H.

Finally, as can be noted by comparing the values of rise for target oligomers, the artificially created gap (6.50 Å) is not as large as the crystallographic original gap (6.88 Å), caused by the real presence of an intercalator. This comparison contributes to explain why the results for ellipticine in docking B were not as satisfactory as in docking A and confirms the assumption that docking of intercalators to DNA is very sensitive to structural features such as the helix twist or small changes in the rise of the intercalation gap.

Comparing Docking Protocols. In order to evaluate the ability of each docking protocol in identifying binding modes, it is interesting to analyze each docking pair, i.e., the same receptor docked with two different ligands, as in Figure 6. The crystallographic oligomer with intercalation gap used in dockings A/E clearly favors intercalative binding mode, probably because of DNA strongly distorted structure (Figure 6A/E). On the other hand, oligomers without intercalation gap used in dockings C/G and D/H cannot lead to intercalation binding mode, resulting in minor groove recognition as the preferred binding mode for both ligands (Figure 6 (parts C/G and D/H)). However, it is important to remember that cluster profile for netropsin in docking E shows binding free energies that are on average less favorable when compared to minor groove binding free energies found in dockings F, G, and H or to the intercalative binding free energies of a real intercalator found in docking A. Analogous, cluster profile for ellipticine in docking D shows binding free energies that are on average less favorable when compared to intercalating binding free energies found in dockings A and B or to minor groove binding free energies of a real minor groove ligand found in docking H. Therefore, the binding free energy profiles from Figure 5 indirectly suggest that the predicted binding modes for netropsin and ellipticine in dockings E and D are not the preferential binding modes for these molecules but artifacts arising from docking methodological limitations.

Finally, the modified canonical B-DNA with intercalation gap used in dockings B/F lead, for each ligand, to the correct binding mode as the energetically most favorable (Figure 6 (parts A/E and B/F)) and thus proved to be a suitable target oligomer in DNA-ligand docking.

Predicting Tröger Bases Binding Mode. We decided to use the modified canonical B-DNA (with gap) as a receptor for docking with the Tröger bases. Therefore, four different dockings were performed: 1) (−)-(R,R) symmetric Tröger base; 2) (+)-(S,S) symmetric Tröger base; 3) (−)-(R,R) asymmetric Tröger base; 4) (+)-(S,S) asymmetric Tröger base.

Among the 100 runs performed in each docking, the 25 most favorable in terms of binding free energy were analyzed. Cluster profiles are shown in Figure 7, and the best ranked conformations for each docking are illustrated in Figure 8.

For both symmetric and asymmetric Tröger bases, cluster profiles show that binding free energies are more negative for the levorotatory form than for the dextrorotatory one, indicating an enantiospecific binding of the (−) isomer to B-DNA. This is in agreement with thermal denaturation studies[29,30] and can be explained by the intrinsic geometry of the compounds; the (−) isomer presents a right-handed helix shape which is similar to B-DNA helices, whereas the (+) isomer presents a left-handed helix shape opposite to B-DNA helices.

Indeed, shape complementarity is the reason why docking of the (−) symmetric Tröger base resulted in binding of the two proflavine moieties to the minor groove with a very favorable binding free energy (−9.90 kcal/mol), while dockings with the (+) symmetric Tröger base resulted only

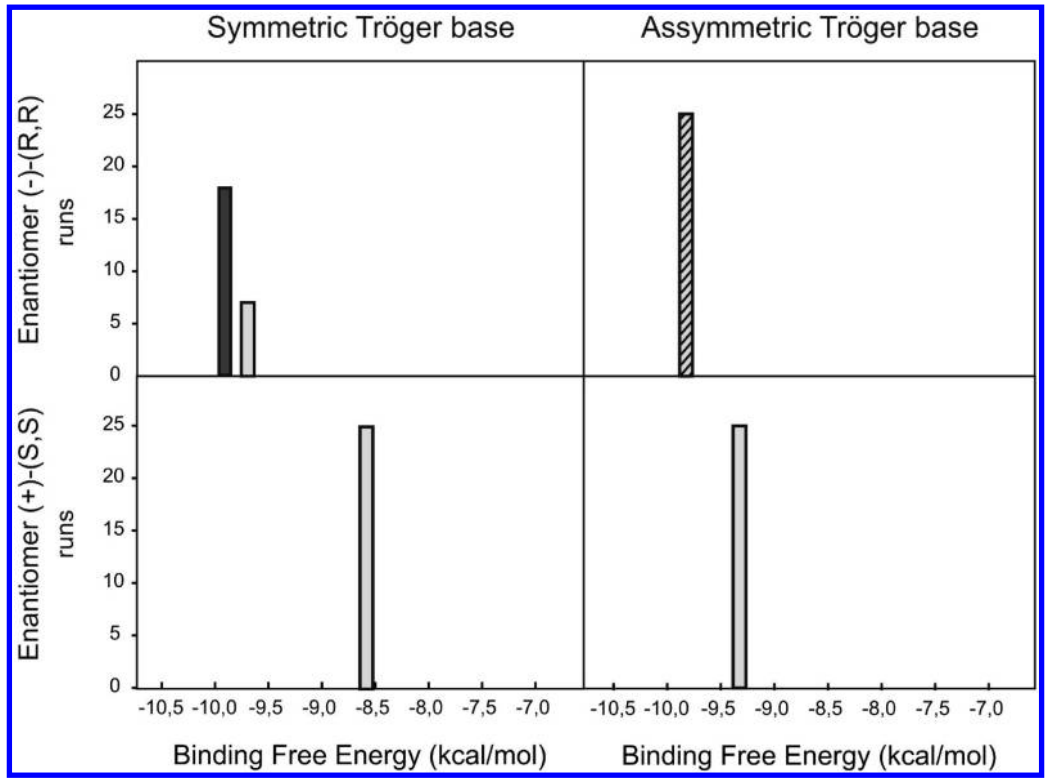

Figure 7. Cluster profiles from docking of Tröger bases to modified B-DNA. Each cluster is represented by a bar in which the height corresponds to the number of conformations in the cluster, and the color indicates the binding mode: minor groove in black, intercalation in gray, and other in hatched pattern bimodal binding mode (intercalation/minor groove recognition).

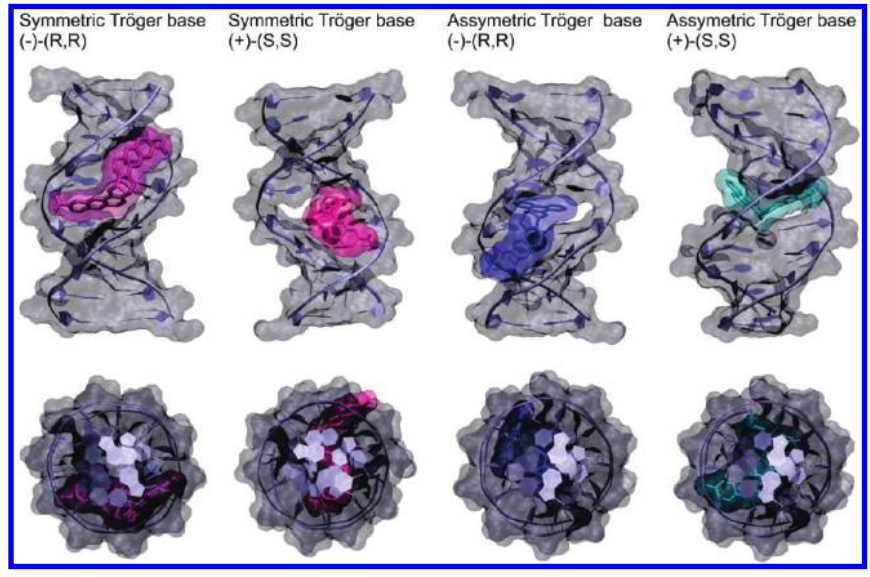

Figure 8. Best docked conformations for dockings of Tröger bases to modified B-DNA. Purple, (−)-(R,R) symmetric Tröger base; magenta, (+)-(S,S) symmetric Tröger base; dark blue, (−)-(R,R) asymmetric Tröger base; cyan, (+)-(S,S) asymmetric Tröger base.

in intercalation of one proflavine moiety while the other proflavine is projected out from the major groove (see top views in Figure 8).

Although docking with the (−) symmetric Tröger base also resulted in intercalation (−9.67 kcal/mol, 7 runs), minor groove binding seems to be the preferred binding mode qualitatively (−9.90 kcal/mol) and quantitatively (18 runs). This is in agreement with electric linear dichroism studies and with the lack of DNA unwinding activity reported by Bailly et al.,[30] who suggest that the (−) symmetric Tröger base interacts with DNA through minor groove recognition.

Moreover, DNase footprinting studies with the symmetric Tröger base proved that only the (−) enantiomer presents sequence selectivity, suggesting a different binding mode for the (+) enantiomer.[30] This assumption is also corroborated by our results which indicate that the (+) Tröger bases interact with DNA via intercalative binding (a binding mode which is usually associated with lack of sequence selectivity).

Concerning the (−) asymmetric Tröger base, all docking runs pointed to a bimodal binding mode, with a large negative binding free energy (−9.84 kcal/mol), in which the phenanthroline moiety is intercalated whereas the proflavine moiety is fitted in the minor groove (Figure 8). Again, this binding mode is only possible because ligand chirality is similar to chirality of the B-DNA double helix. On the other hand, dockings of the left-handed (+) asymmetric Tröger

base resulted only in intercalation with binding free energy (−9.35 kcal/mol) not as favorable as those from bimodal binding mode of the (−) isomer.

It has already been suggested that the (−) asymmetric Tröger base does not interact with DNA through minor groove recognition since DNase footprinting assays showed that it presents only a moderated level of sequence selectivity.[31] This moderate sequence selectivity could be explained by intercalation of one ring moiety, as was suggested by docking results and also by topoisomerase I inhibition assays reported by Baldeyrou et al.[31] However, our results only partially agree with experimental data since circular dichroism and electric linear dichroism studies suggest a bimodal binding mode in which proflavine is intercalated, while phenanthroline resides in minor groove.[31]

It is possible that docking results pointed to an alternative bimodal binding mode (with phenanthroline intercalated) due to docking flexibility limitations already mentioned. As previously discussed, the minor groove of an artificial canonical DNA containing an intercalation gap is not as large as the minor groove of a crystallographic DNA complexed with an intercalating agent. In other words, the minor groove width of an artificially modified canonical DNA may be too narrow to bear the phenanthroline moiety, which means that the target conformation may have guided the docking to this analogous binding mode in which proflavine - instead of phenanthroline - binds the minor groove. In order to investigate this hypothesis, we have also performed dockings between the (−) asymmetric Tröger base and a crystallographic DNA from the ellipticine complex (data not shown). Among the 25 most favorable runs, all pointed to the bimodal binding mode suggested by the literature (with proflavine intercalated and phenanthroline fitted in minor groove), with a binding free energy of −10.35 kcal/mol.

Testing Docking Protocol with Aflatoxin B1 *exo*-8,9-Epoxide. We decided to apply the generic docking protocol with a modified canonical B-DNA (with gap) to dock aflatoxin B1 *exo*-8,9-epoxide and compare the result with the result from the self-docking. Therefore, two different dockings were performed: A) aflatoxin and crystallographic DNA (1MKL) and B) aflatoxin and modified canonical DNA (with gap).

Among the 100 runs performed in each docking, the 25 most favorable in terms of binding free energy were analyzed. Cluster profiles are shown in Figure 9, and the best ranked conformations for each docking are illustrated in Figure 10.

The cluster profiles in Figure 9 show similar results for dockings A and B. Very promising is the fact that both dockings not only resulted in intercalation binding mode for the aflatoxin B_1 *exo*-8,9-epoxide but also placed the epoxide in close proximity and in the proper orientation to the N7 position of guanine (see Figure 10), thus consistent with the proposed S_N2 mechanism of alkylation.[32,54] In docking A, the epoxide carbon which is to be attacked is placed at 3.51 Å from the N7 position with a free binding energy of −7.61 kcal/mol, and, in docking B, the same carbon is placed at 3.09 Å from the N7 position with a binding free energy of −7.47 kcal/mol.

These very similar results are also noteworthy if we consider the structural differences between the two oligomers used as targets. As can be observed in Figure 10, the crystallographic DNA from the covalent adduct presents a significant bending toward the major groove (20° according to Giri et al.[33]), which does not occur in the modified

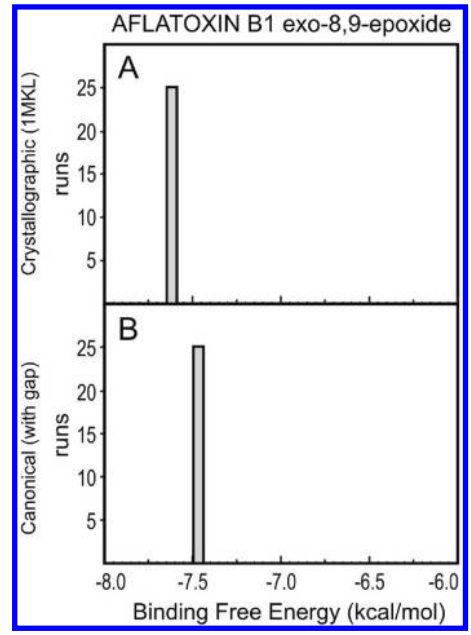

Figure 9. Cluster profiles from docking of aflatoxin B1 *exo*-8,9-epoxide with crystallographic DNA (A) and with canonical modified DNA containing intercalation gap (B). Each cluster is represented by a bar in which the height corresponds to the number of conformations in the cluster, and the color indicates the binding mode: minor groove in black, intercalation in gray, and other in hatched pattern.

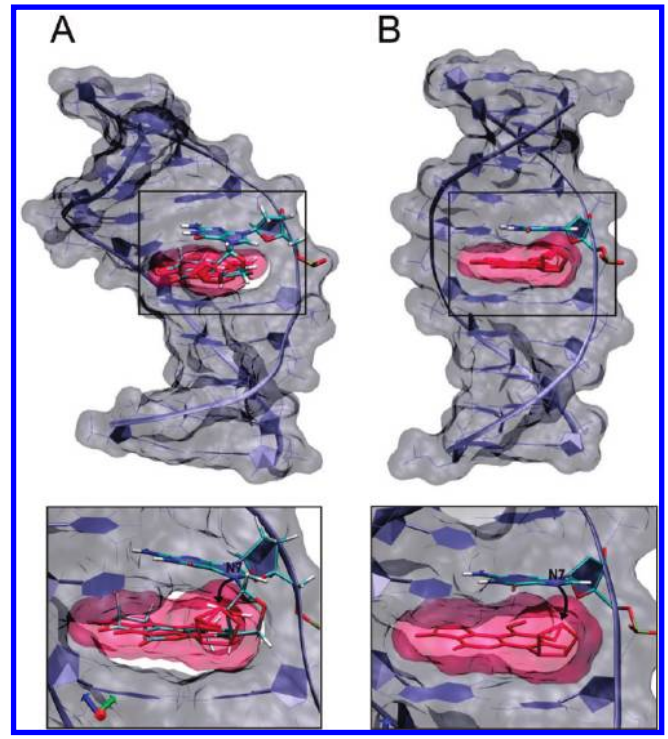

Figure 10. Best docked conformations for dockings of aflatoxin B_1 *exo*-8,9-epoxide with crystallographic DNA (A) and with modified canonical DNA with intercalation gap (B). CPK representation in DNA indicates the guanine which forms the bond with aflatoxin B_1 *exo*-8,9-epoxide. In part (A), the covalent adduct from original complex is also shown in CPK. Aflatoxin B_1 *exo*-8,9-epoxide is shown in red. Below: conformations are shown in detail, with the carbon which is to be attacked by the N7 position from guanine indicated by an arrow.

canonical B-DNA. Since such helix features already proved to be of major importance during target rigid docking, one may have expected that bending would strongly affect aflatoxin dockings and lead to different cluster profiles, which was not the case.

In this particular case, it seems that the helix bending does not result from an induced fit promoted by intercalation *per se* but arises as an effect from subsequent covalent bonding. In other words, the bending does not interfere significantly in aflatoxin dockings because it is not required for the first noncovalent interaction. This is in agreement with Giri et al., who argue that DNA helix bending arises from the change in N7 hybridization when the covalent adduct is formed.[33]

Finally, the fact that both dockings placed the epoxide in an orientation so favorable for a subsequent alkylation in the N7 position is in agreement with the hypothesis that aflatoxin B_1 was optimized during evolution to enhance its reactivity in DNA environment. This was already supported by experimental studies which showed the reaction of aflatoxin B_1 *exo*-8,9-epoxide to be more than 2000 times more efficient in DNA than in an aqueous solution with free 2′-deoxyguanosine.[54] Based also on thermodynamic analysis, Brown et al. proposed that intercalative binding plays a major role guiding the formation of the later adduct between aflatoxin B_1 and DNA,[54] which is clearly in agreement with the noncovalent complexes from dockings A and B.

CONCLUSIONS

We have performed several docking studies using ligands which interact with DNA as intercalators (ellipticine and acridine derivative) or as a minor groove binder (netropsin). Four distinct DNA structures were used as targets: crystallographic DNA from ellipticine complex, crystallographic DNA from netropsin complex, constructed B-canonical DNA, and modified B-canonical DNA containing an intercalation gap. Direct and cross dockings were carried out with an optimized set of parameters using AutoDock suite.

Docking results from validation stage are in agreement with each binding mechanism (induced-fit or lock-and-key) when the system is subjected to the flexibility limitations of the docking method: the docking with ellipticine was found to be more sensitive to the selected target conformation than the docking with netropsin. In cross dockings, this observation is more evident: netropsin can dock even to the minor groove of a structurally distorted DNA, whereas ellipticine cannot intercalate when the receptor does not possess a proper intercalation gap. Also, our results make more evident the important role that DNA features such as twist, rise, or groove width play in rigid dockings, reinforcing the necessity to treat target flexibility in dockings concerning nucleic acids.

Nevertheless, docking results also proved that the current limitations of docking methods can be overcome by a proper choice of the target conformation. We found that, provided that target DNA presents an artificial intercalation gap, dockings pointed to the correct binding mode as the energetically most favorable result, also suggesting that the AutoDock score function is efficient to evaluate ligand-DNA interactions at least in a qualitative way.

We further applied this docking protocol using a modified canonical B-DNA with an intercalation gap as receptor to predict the binding modes of two Tröger bases: a symmetric base derived from proflavine and an asymmetric base derived from proflavine and phenanthroline, each with two optical isomers.

The choice of modified canonical DNA as a receptor for docking with Tröger bases gave rise to very promising profiles. Besides resulting in very favorable binding free energies, these dockings were capable of reproducing the stronger DNA binding affinity that the Tröger levorotatory isomers possess when compared to the dextrorotatory ones. Moreover, binding modes suggested by the docking are in agreement with the binding modes suggested by experimental data found in the literature.

Finally, we tested the proposed docking protocol with aflatoxin B1 *exo*-8,9-epoxide, an intercalative binding agent that also alkylates DNA in the major groove, and compared the result with that from a self-docking. This comparison clearly shows that the choice of modified B-DNA as target not only resulted in a binding profile very similar to that from the self-docking but also correctly placed the carbon which is to be attacked in close proximity of the N7 position of guanine, in the major groove.

Therefore, we propose that a default protocol using a modified canonical DNA with an artificial intercalation gap can be successfully applied to investigate other ligands whose binding mode remains unknown, yielding a suitable starting point for further theoretical studies such as more refined dockings or molecular dynamics simulations.

ACKNOWLEDGMENT

We would like to thank Hermes L. N. de Amorim and Melina Mottin for critical review of the manuscript. We also acknowledge the financial support from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

REFERENCES AND NOTES

(1) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nat. Rev. Drug Discovery* **2004**, *3*, 935–949.
(2) Moitessier, N.; Englebienne, P.; Lee, D.,; Lawandi, J.; Corbeil, C. R. Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. *Br. J. Pharmacol.* **2008**, *153*, S7–S26.
(3) Grootenhuis, P. D. J.; Kollman, P. A.; Seibel, G. L.; DesJarlais, R. L.; Kuntz, I. D. Computerized selection of potential DNA binding compounds. *Anti-Cancer Drug Des.* **1990**, *5*, 237–242.
(4) Grootenhuis, P. D. J.; Roe, D. C.; Kollman, P. A.; Kuntz, I. D. Finding potential DNA-binding compounds by using molecular shape. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 731–750.
(5) Chen, K.; Adelstein, S. J.; Kassis, A. I. Molecular modeling of the interaction of the iodinated Hoechst analogs with DNA: implications for new radio-pharmaceutical design. *J. Mol. Struct.-THEOCHEM* **2004**, *711*, 49–56.
(6) Rohs, R.; Bloch, I.; Sklenar, H.; Shakked, Z. Molecular flexibility in *ab initio* drug docking to DNA: binding-site and binding-mode transitions in all-atom Monte Carlo simulations. *Nucleic Acids Res.* **2005**, *33*, 7048–7057.
(7) Tuttle, T.; Kraka, E.; Cremer, D. J. Docking, triggering and biological activity of dynemicin A in DNA: a computational study. *J. Am. Chem. Soc.* **2005**, *127*, 9469–9484.
(8) Evans, D. A.; Neidle, S. Virtual screening of DNA minor groove binders. *J. Med. Chem.* **2006**, *39*, 4232–4238.
(9) Sobhani, A. M.; Amini, S. R.; Tyndall, J. D. A.; Azizi, E.; Daneshtalab, M.; Khalaj, A. A theory of mode of action of azolylalkylquinolines as DNA binding agents using automated flexible ligand docking. *J. Mol. Graphics Modell.* **2006**, *25*, 459–469.
(10) Lauria, A.; Patella, C.; Ippolito, M.; Almerico, A. M. Docking and synthesis of pyrrolopyrimidodiazepinone derivatives (PPDs) and their

precursors: New scaffolds for DNA-interacting agents. *J. Mol. Struct.-THEOCHEM* **2007**, *819*, 26–31.

(11) Navarrete, J. T.; Casado, J.; Ram, F. J. Electronic spectroscopy study and molecular docking simulation of the interaction of terthiophene with DNA. *J. Mol. Struct.* **2007**, *834*, 176–181.

(12) Yan, Z.; Sikri, S.; Beveridge, D. L.; Baranger, A. M. Identification of an aminoacridine derivative that binds to RNA tetraloops. *J. Med. Chem.* **2007**, *50*, 4096–4104.

(13) Holt, P. A.; Chaires, J. B.; Trent, J. O. Molecular docking of intercalators and groove-binders to nucleic acids using AutoDock and Surflex. *J. Chem. Inf. Model.* **2008**, *48*, 1602–1615.

(14) Hurley, L. H. DNA and its associated processes as targets for cancer therapy. *Nat. Rev. Cancer* **2002**, *2*, 188–199.

(15) Sternberg, M. J. E.; Gabb, H. A.; Jackson, R. M. Predictive docking of protein-protein and protein-DNA complexes. *Curr. Opin. Struct. Biol.* **1998**, *8*, 250–256.

(16) Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M. L. The anatomy of A-, B-, and Z-DNA. *Science* **1982**, *216*, 475–485.

(17) Dickerson, R. E. DNA structure from A to Z. *Methods Enzymol.* **1992**, *211*, 67–111.

(18) Phan, A. T.; Kuryavyi, V.; Patel, D. J. DNA architecture from G to Z. *Curr. Opin. Struct. Biol.* **2006**, *16*, 288–298.

(19) Hannon, M. J. Supramolecular DNA recognition. *Chem. Soc. Rev.* **2007**, *36*, 280–295.

(20) Tse, W. C.; Boger, D. L. Sequence-Selective DNA Recognition: Natural products and Nature's lessons. *Chem. Biol.* **2004**, *11*, 1607–1617.

(21) Waring, M. J.; Bailly, C. DNA recognition by intercalators and hybrid molecules. *J. Mol. Recognit.* **1994**, *7*, 109–122.

(22) Dervan, P. B. Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.* **2001**, *9*, 2215–2235.

(23) Bailly, C.; Chaires, B. Sequence-specific DNA minor groove binders. Design and synthesis of netropsin and distamycin analogues. *Bioconjugate Chem.* **1998**, *9*, 513–538.

(24) Dolenc, J.; Borstnik, U.; Hodoscek, M.; Koller, J.; Janezic, D. Ab initio QM/MM study of the conformational stability of the complexes formed by netropsin and DNA. The importance of van der Waals interactions and hydrogen bonding. *J. Mol. Struct.-THEOCHEM.* **2005**, *718*, 77–85.

(25) Pindur, U.; Jansen, M.; Lemster, T. Advances in DNA-ligands with groove binding, intercalating and/or alkylating activity: Chemistry, DNA-binding and biology. *Curr. Med. Chem.* **2005**, *12*, 2805–2847.

(26) Bren, U.; Hodoscek, M.; Koller, J. Development and validation of empirical force field parameters for netropsin. *J. Chem. Inf. Model.* **2005**, *45*, 1546–1552.

(27) de Souza, O. N.; Ornstein, R. L. Inherent DNA curvature and flexibility correlate with TATA Box functionality. *Biopolymers* **1998**, *46*, 403–415.

(28) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using Lamarckian Genetic Algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.

(29) Tatibouët, A.; Demeunynck, M.; Andraud, C.; Collet, A.; Lhomme, J. Synthesis and study of an acridine substituted Tröger's Base: preferential binding of the (−)-isomer to B-DNA. *Chem. Commun.* **1999**, *2*, 161–162.

(30) Bailly, C.; Laine, W.; Demeuninck, M.; Lhomme, J. Enantiospecifc recognition of DNA sequences by a proflavine Tröger base. *Biochem. Biophys. Res. Commun.* **2000**, *273*, 681–685.

(31) Baldeyrou, B.; Tardy, C.; Bailly, C.; Colson, P.; Houssier, C.; Charmantray, F.; Demeuninck, M. Synthesis and DNA interaction of a mixed proflavine-phenanthroline Tröger base. *Eur. J. Med. Chem.* **2002**, *37*, 315–322.

(32) Iyer, R. S.; Coles, B. F.; Raney, K. D.; Thier, R.; Guengerich, F. P.; Harris, T. M. DNA adduct by the potent carcinogen aflatoxin B_1: mechanistic studies. *J. Am. Chem. Soc.* **1994**, *116*, 1603–1609.

(33) Giri, I.; Jenkins, M. D.; Schnetz-Boutaud, N. C.; Stone, M. P. Structural refinement of the 8,9-dihydro-8-(N7-guanyl)-9-hydroxy-aflatoxin B_1 adduct in a 5′-Cp^{AFB}G-3′ sequence. *Chem. Res. Toxicol.* **2002**, *15*, 638–647.

(34) Bren, U.; Václav, M.; Florián, J. Free energy simulations of uncatalyzed DNA replication fidelity: structure and stability of T·G and dTTP·G terminal DNA mismatches flanked by a single dangling nucleotide. *J. Phys. Chem. B* **2006**, *110*, 10557–10566.

(35) Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. A semiempirical free energy force field with charge-based desolvation. *J. Comput. Chem.* **2007**, *28*, 1145–1152.

(36) Solis, F. J.; Wets, J.-B. Minimization by random search techniques. *Math. Oper. Res.* **1981**, *6*, 19–30.

(37) http://www.rcsb.org/pdb/home/home.do (accessed July 13, 2009).

(38) Canals, A.; Purciolas, M.; Aymami, J.; Coll, M. The anticancer agent ellipticine unwinds DNA by intercalative binding in an orientation parallel to base pairs. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2005**, *61*, 1009–1012.

(39) Finlay, A. C.; Hochstein, F. A.; Sobin, B. A.; Murphy, F. X. Netropsin, a new antibiotic produced by Streptomyces. *J. Am. Chem. Soc.* **1951**, *73*, 341–343.

(40) Coll, M.; Aymami, J.; van der Marel, G. A.; van Boom, J. H.; Rich, A.; Wang, A. H. Molecular structure of the netropsin-d(CGCGATATCGCG) complex: DNA conformation in an alternating segment. *Biochemistry* **1989**, *28*, 310–320.

(41) Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* **2001**, *7*, 306–317.

(42) *GROMACS, version 3.3*; Groningen, Netherlands, 2007. Available from http://www.gromacs.org/ (accessed July 13, 2009).

(43) Oostenbrink, C.; Soares, T. A.; van der Vegt, N. F. A.; van Gunsteren, W. F. Validation of the 53A6 GROMOS force field. *Eur. Biophys. J.* **2005**, *34*, 273–284.

(44) Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity - a rapid access to atomic charges. *Tetrahedron* **1980**, *36*, 3219–3228.

(45) Sanner, M. F. Python: a programming language for software integration and development. *J. Mol. Graphics Modell.* **1999**, *17*, 57–61.

(46) Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.

(47) Coll, M.; Frederick, C. A.; Wang, A.D.-J.; Rich, A. A bifurcated hydrogen-bonded conformation in the d(A·T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8385–8389.

(48) Hopcroft, N. H.; Brogden, A. L.; Searcey, M.; Cardin, C. J. X-ray crystallographic study of DNA duplex crosslinking: simultaneous binding to two $d(CGTACG)_2$ molecules by a bis(9-aminoacridine-4-carboxamide) derivative. *Nucleic Acids Res.* **2006**, *34*, 6663–6672.

(49) Lu, X.-J.; Olson, W.-K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **2003**, *31*, 5108–5121.

(50) Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: na environment for comparative protein modeling. *Electrophoresis* **1997**, *18*, 2714–2723.

(51) Dennington, R., II; Keith, T.; Millam, J. *GaussView, Version 4.1*; Semichem, Inc.: Shawnee Mission, KS, 2007.

(52) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98 (Revision A.1x)*; Gaussian, Inc.: Pittsburgh, PA, 2001.

(53) Bren, U.; Guengerich, F. P.; Mavri, J. Guanine alkylation by the potent carcinogen aflatoxin B_1: quantum chemical calculations. *Chem. Res. Toxicol.* **2007**, *20*, 1134–1140.

(54) Brown, K. L.; Bren, U.; Stone, M. P.; Guengerich, F. P. Inherent stereospecificity in the reaction of aflatoxin B_1 8,9-epoxide with the deoxyguanosine and efficiency of DNA catalysis. *Chem. Res. Toxicol.* **2009**, *22*, 913–917.

CI9001537