

## Simulation of the Interactions between Tröger Bases and DNA

Clarisse G. Ricci and Paulo A. Netz\*

Instituto de Química, Universidade Federal do Rio Grande do Sul, 91501-970 Porto Alegre-RS, Brazil

Bases de Tröger são moléculas que, devido à sua geometria, ligam-se de modo enantiosseletivo ao DNA. Usando dinâmica molecular, foram simuladas bases de Tröger com substituintes proflavina e fenantrolina. Partindo das bases docadas no DNA, foram investigadas as distorções induzidas em dois modos de ligação: intercalação e ligação de sulco. Nos complexos de intercalação as bases apresentam tempo de residência elevado e distorcem a dupla hélice levando a um desenrolamento parcial e a valores não-canônicos de alguns ângulos torcionais. Nos complexos de ligação de sulco, elas exibem alta mobilidade e levam a uma alteração no modo de ligação, interagindo com o sulco via ponte diazocina. Os resultados sugerem a intercalação de um substituinte, com contatos adicionais no sulco, como o modo de ligação preferencial destas bases de Tröger, sendo que a ligação de sulco explica a fraca ligação dos isômeros dextrorrotatórios.

Tröger bases are a class of molecules that, due to its geometry, bind enantioselectively to DNA. Molecular dynamic simulations were performed with levorotatory isomers of proflavine and phenanthroline substituted Tröger bases. Starting with the bases docked in DNA, the distortions they promote in the double helix were investigated in two possible modes: intercalation and minor groove binding. In the intercalation complexes, they presented long residence times and distorted the double helix leading to partial unwinding and to non-canonical values of some backbone angles. In the minor groove complexes, they displayed high mobility, leading to a change in the binding mode, interacting with the minor groove mainly through the diazocin bridge. The results suggested the intercalation of one substituent (with additional contacts in the minor groove) as the preferential binding mode for these Tröger bases, while minor groove binding may explain the weaker binding observed for the dextrorotatory isomers.

**Keywords:** DNA, molecular dynamics, Tröger bases

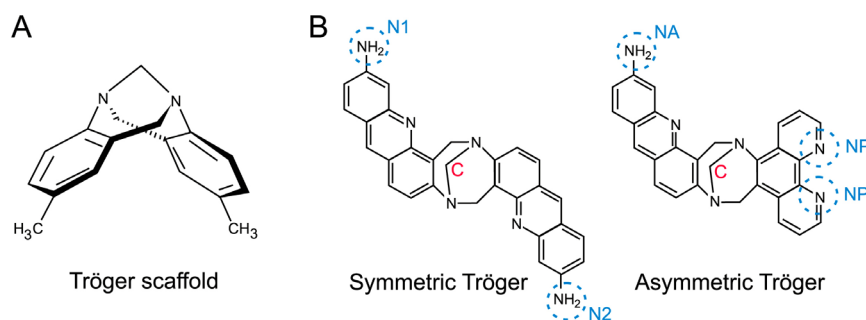
### Introduction

Molecules capable of binding nucleic acids have long been used as antibiotic and antitumoral drugs due to their cytotoxic effects upon cell growth,<sup>1-5</sup> besides being important tools in molecular biology in the form of fluorescent dyes. Since nucleic acids are directly involved in several cellular processes, great attention is focused on the discovery and design of novel DNA binding agents.<sup>5,6</sup> This is reinforced by the fact that DNA binding agents are among the most efficient drugs currently used in anticancer therapies, although they still lack selectivity enough to avoid severe and adverse side effects.<sup>5,7</sup>

In contrast with other classes of DNA binding agents, Tröger bases are a class of molecules that has received very little attention regarding its ability to bind nucleic

acids. The first Tröger base was synthesized in 1887<sup>8</sup> and further received considerable interest due to its V-shaped chiral structure (see Figure 1A).<sup>9,10</sup> Recently, Abella *et al.*<sup>11</sup> carried out some studies on Tröger bases, regarding their photophysical properties as well as regarding the detailed investigation of their mechanism of formation.<sup>12</sup> It was also only recently shown that proflavine/acridine and phenanthroline-containing Tröger bases were able to selectively bind B-DNA in a stereospecific fashion.<sup>13,14</sup> According to Valik *et al.*,<sup>10</sup> these chiral properties make these compounds of great promise as DNA probes, since they discern between right-handed and left-handed DNA isoforms. Even so, there are very few studies regarding Tröger-DNA interactions so far<sup>13-18</sup> and, despite inconclusive assumptions based on spectroscopic and biochemical data, the binding mechanism of these molecules to DNA remains unknown. It has been suggested that the levorotatory isomer of a symmetric proflavine Tröger base (Figure 1B) is more

\*e-mail: netz@iq.ufrgs.br



**Figure 1.** A) Tröger base scaffold. B) Tröger bases that interact with DNA. Some groups are highlighted (amino groups, nitrogen and carbon from diazocine bridge), since they will be used for distance analysis in the present work.

likely to interact through minor groove binding, although proflavine is a known intercalating agent.<sup>15</sup> On the other hand, similar studies concerning the levorotatory isomer of an asymmetric Tröger derivative containing both proflavine and phenanthroline (Figure 1B) have suggested a mixed mechanism consisting of intercalation of proflavine and groove binding of phenanthroline.<sup>16</sup>

Since both Tröger derivatives contain DNA binding agents as substituents, it is not clear whether the interaction follows from the intrinsic right-handed shape of the Tröger scaffold or from the intrinsic binding ability of the substituents. Thus, from the results mentioned above rises the question whether there is a common and unique mechanism of binding for Tröger bases - which would result in a new class of DNA binding agents - or if the V-shaped structure actually works as a versatile scaffold, which can combine, potentiate or even modify the substituent binding profiles. These considerations are reinforced by the fact that the levorotatory isomers (right-handed) have been shown to present higher affinity and sequence-selectivity upon binding than the dextrorotatory isomers (left-handed),<sup>14-16</sup> suggesting there may exist more than one mechanism of binding, depending on Tröger chirality. To address these issues, we recently performed docking studies with these Tröger bases, using a docking protocol which was developed for situations where the mechanism of binding (intercalation or groove binding) is unknown.<sup>19</sup> Based on docking results, we now used the docking complexes as starting structures for molecular dynamic simulations, to

access the residence time of the ligands in each binding mode (intercalation or minor groove binding) and also to investigate the effect of Tröger binding upon DNA structure and stability.

## Methods

Our previous docking studies<sup>19</sup> showed us that the interactions of the DNA with levorotatory isomers are much stronger than with dextrorotatory isomers. Among the complexes resulting from this study,<sup>19</sup> four complexes were selected for simulation, as described in Table 1. Each complex represents the energetically most favorable conformation from each docking system, according to Autodock 4.0 score function.<sup>20,21</sup> To evaluate two possible binding modes, we selected complexes in which the levorotatory isomers are docked in the minor groove of a crystallographic oligomer (PDB-ID 1DNE) or intercalated in a modified canonical B-DNA, which presents an artificial intercalation gap. The initial structures for the complexes with symmetric Tröger base are displayed in Figures 3C and 4C (0 ns), while those concerning the asymmetric Tröger base can be found in Figure 5C (0 ns) and in Supplementary Information (Figure S1, 0 ns). Also to discriminate between intrinsic DNA flexibility and ligand induced effects, each oligomer was simulated in the absence of the ligands, totalizing 6 simulated systems, as seen in Table 1.

All simulations were carried out with the GROMACS package<sup>22</sup> using AMBER 03 force field<sup>23</sup> ports, according

**Table 1.** Description of the simulated systems

System	Tröger base	DNA	Binding mode
SYM-INT	(-)-symmetric Tröger	B-canonical with gap	intercalation
SYM-GROOVE	(-)-symmetric Tröger	crystallographic without gap	minor groove binding
ASYM-INT	(-)-asymmetric Tröger	B-canonical with gap	phenanthroline intercalation
ASYM-GROOVE	(-)-asymmetric Tröger	crystallographic without gap	minor groove binding
DNA-GAP	-	B-canonical with gap	-
DNA-NOGAP	-	crystallographic without gap	-

to Sorin and Pande.<sup>24</sup> Ligand parameters and topologies were obtained from Generalized Amber Force Field,<sup>25</sup> using the AnteChamber Python Parses Interface (acpypi.py)<sup>26</sup> tool. After calculation of electrostatic potentials with the restricted Hartree Fock calculations with 6-31G\* basis set using Gaussian,<sup>27</sup> atomic charges were obtained according to RESP method,<sup>28</sup> using Gamess.<sup>29</sup>

Cubic simulation boxes were created by centering the complex or the oligomer in a box with boundaries at least 1.8 nm apart from all atoms of the solute molecules. Counterions (sodium ions) were randomly placed, followed by a short position restrained molecular dynamics in vacuum (100 ps), during which the counterions are free to migrate to the energetically most favorable positions. After this initial setup, the following procedures were applied: addition of water molecules TIP3P,<sup>30</sup> energy minimization, addition of ions leading to physiological concentration (0.154 mol L<sup>-1</sup> NaCl), energy minimization, and 100 ps of position restrained molecular dynamics with the full solvated system. After these initial steps, a heating ramp was applied, consisting of short (50 ps) consecutive simulations of 50, 100, 150, 200, 250, and 300 K. The production simulation consisted of 20 ns at 310 K. The time step was 0.002 ps, and the electrostatic interactions were calculated using PME.<sup>31,32</sup>

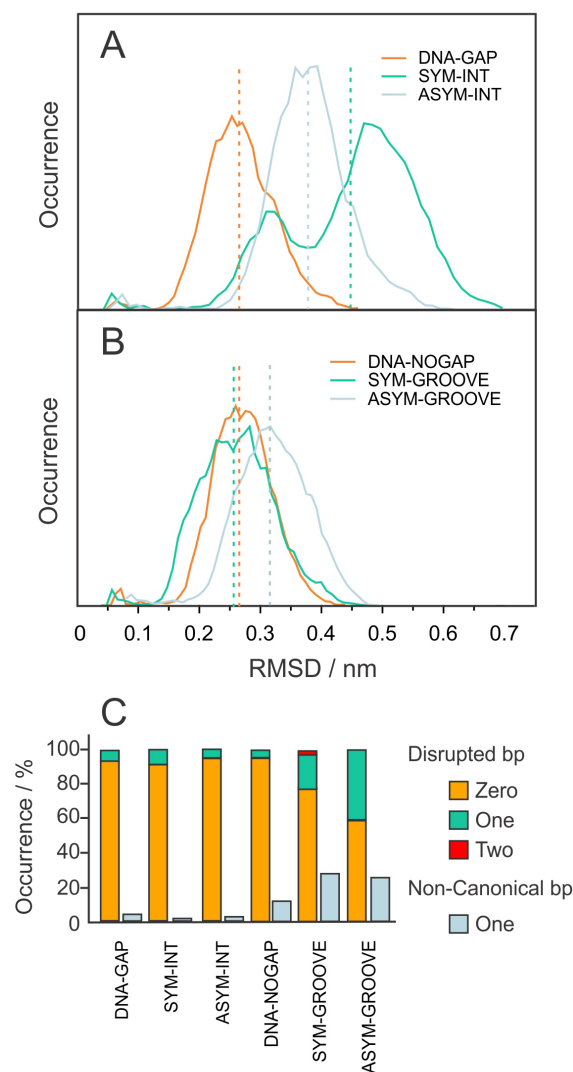
The simulated trajectories were analyzed with standard GROMACS tools in order to calculate the time evolution of the root mean square deviation (RMSD) and the distances between ligands and DNA. Structural analysis of DNA double helix (torsion angles and base-pair step parameters) was achieved with 3DNA,<sup>33</sup> which was applied to every 5 ps spaced structures extracted from the simulation trajectories. The overall double helix stability was monitored by calculating the total number of conserved base-pairs and also the number of non-canonical (non-Watson-Crick) base pairs along the simulation. Visual analysis of trajectories was achieved with the VMD package.<sup>34</sup>

## Results and Discussion

### Overall stability

In order to access the global stability of DNA, we monitored the root mean square deviation (RMSD) and also the number of base-pairs conserved through the time of the simulations (Figure 2). It is possible to observe that DNA in intercalation complexes reached higher RMSD values than DNA simulated alone, indicating that intercalation promotes significant distortion in the DNA double helix (Figure 2A). On the other hand, DNA in minor groove complexes seems to endure structural changes that are

similar in magnitude to those arising from the relaxation of DNA structure alone, since RMSD average values from SYM-GROOVE and ASYM-GROOVE systems were close to that from DNA-NOGAP (Figure 2B). Analysis of conserved base-pairs shows that all DNA oligomers are stable against denaturation (Figure 2C). Although the terminal regions eventually endured disruption of one or two base-pairs, it is interesting to note that in some cases these were replaced by the formation of non-canonical base-pairs, i.e., base-pairs that do not strictly obey the hydrogen bonding pattern proposed by Watson and Crick.<sup>35</sup> Minor groove complexes seem to display the highest degree of non-canonical pattern, but with only modest distortion of the overall structure. Therefore, one can conclude that the



**Figure 2.** A) RMSD histogram for DNA-GAP and intercalation complexes, SYM-INT and ASYM-INT. B) RMSD histogram for DNA-NOGAP and minor groove complexes, SYM-GROOVE and ASYM-GROOVE. The average RMSD values are indicated by the dashed lines. C) Pattern of canonical base-pair (bp) disruption and non-canonical base-pair formation during the simulations.

high RMSD values observed in the simulations are due to the intrinsic flexibility of DNA or due to ligand induced distortions, and do not simply follow from denaturation of DNA double helix.

### Global analysis

To access the residence time of ligands in each binding mode and also to highlight the main events that occurred during the simulations, we monitored the distances between Tröger bases and DNA base-pairs. These analysis are shown in Figures 3, 4 and 5, together with structures extracted from the trajectories. We choose three groups for each Tröger base, whose distances were measured relative to the center of mass of DNA base-pairs: the carbon from diazocine bridge (C), the proflavine nitrogens (N1 and N2, for symmetric base, and NA, for asymmetric base) and one of the phenanthroline nitrogens (NP) for the asymmetric base (see groups highlighted in Figure 1B). This global analysis is discussed in detail for each complex concerning the symmetric Tröger base, while some of the results concerning the asymmetric Tröger base are presented in the Supplementary Information.

### Intercalation complex of symmetric Tröger base (SYM-INT)

Initially, the symmetric Tröger base was centrally intercalated in DNA, with the diazocine bridge inserted in the intercalation gap (Figure 3C, 0 ns). However, it early suffered a reorientation such that one proflavine penetrated into the gap, while the other began to interact with DNA minor groove, as schematized in Figure 3A. According to the distance analysis (Figure 3B), it is possible to observe that (i) the diazocine bridge remained close to 6A-15T base-pair, (ii) the amino group N1 was approximately equidistant from 5T-16A and 6A-15T base-pairs, thus inserted in the intercalation gap, and (iii) the other amino group (N2) kept closer interactions with 5T-16A and 4T-17A base-pairs, although the higher distances indicate a loose attachment.

For all three groups monitored, the greatest fluctuations occurred regarding the 6A-15T base-pair and were due to the movement of the base-pair itself, since the distances concerning other base-pairs remained stable. Moreover, the fluctuation pattern of 6A-15T base-pair is the same regarding N2 and C groups, but inverse when it concerns the N1 group (see red lines in Figure 3B). Therefore, while the center of mass of 6A-15T base-pair approaches N2 and the diazocine bridge, it departs from N1 group and vice-versa, indicating that the 6A-15T base-pair moves in the base-pair plane. It is also likely that this movement

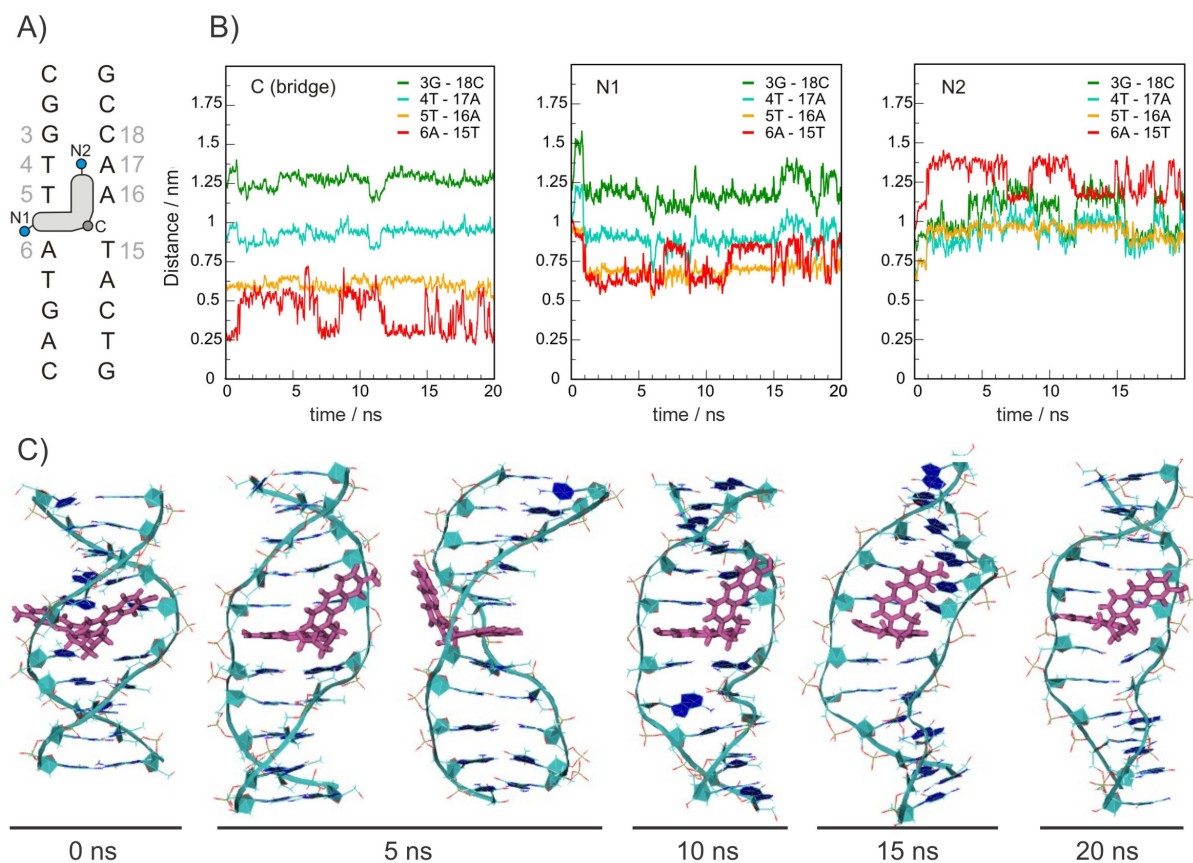
arises from a decrease in DNA twist angle, which can lead to an enhancement of  $\pi$ -stacking interactions between the aromatic rings from Tröger base and the nitrogen bases of DNA.<sup>36</sup> Indeed, unwinding of DNA can be observed in the structures extracted from simulation (Figure 3C) and was further quantified by base-pair step analysis (see Figure 9).

Also noteworthy is the fact that very early in the simulation (ca. 1 ns) the Tröger base endured a reorientation inside the gap leading to an abrupt decrease in the distance between N1 group and the center of mass of all monitored base-pairs (Figure 3B). This indicates that N1 group penetrated inside the intercalation gap, approaching all base-pairs simultaneously, and is confirmed by structure extracted from 5 ns of simulation (Figure 3C). In this stage, the region flanking the intercalation gap already presents local unwinding, which seems to propagate to other regions of the oligomer after 15 ns. Figure 3C also shows that DNA backbone suffers severe distortions near the intercalation gap, that may elapse from the reorientation of the 6A-15T base-pair. Finally, structures extracted from the last 5 ns of simulations show that the bases belonging to the terminal base-pair are piled as if they were consecutive base-pairs, an artifact arising from the finite and short extent of the DNA segment used in the simulation. The final structure, at 20 ns, shows an unwound double helix presenting severe zigzag distortions in the backbone.

A similar profile was observed for the intercalation complex of asymmetric Tröger base, which also remained intercalated in the gap during the simulation time. Further details upon this complex are described in Supplementary Information (Figure S1). In both intercalation complexes, it were observed long residence times (at least 20 ns), together with a decrease in the twist angle, which is a well established signature for the intercalation binding mode.<sup>6,36-38</sup>

### Minor groove complex of symmetric Tröger base (SYM-GROOVE)

As schematically shown in Figure 4A, the Tröger base was initially docked in the minor groove, with the diazocine bridge projected outside. Although the Tröger base interacted with DNA during the entire simulation, distance analysis show that it endured significant reorientation in the minor groove (Figure 4B). According to the distances regarding the diazocine bridge, it is possible to identify four well defined transitions, indicated by the dashed lines in Figure 4B. In this way, the graphic can be divided in five regions (I, II, III, IV, and V), each one consisting of stable distances between the diazocine bridge and DNA base-pairs. As expected, some transitions regarding the

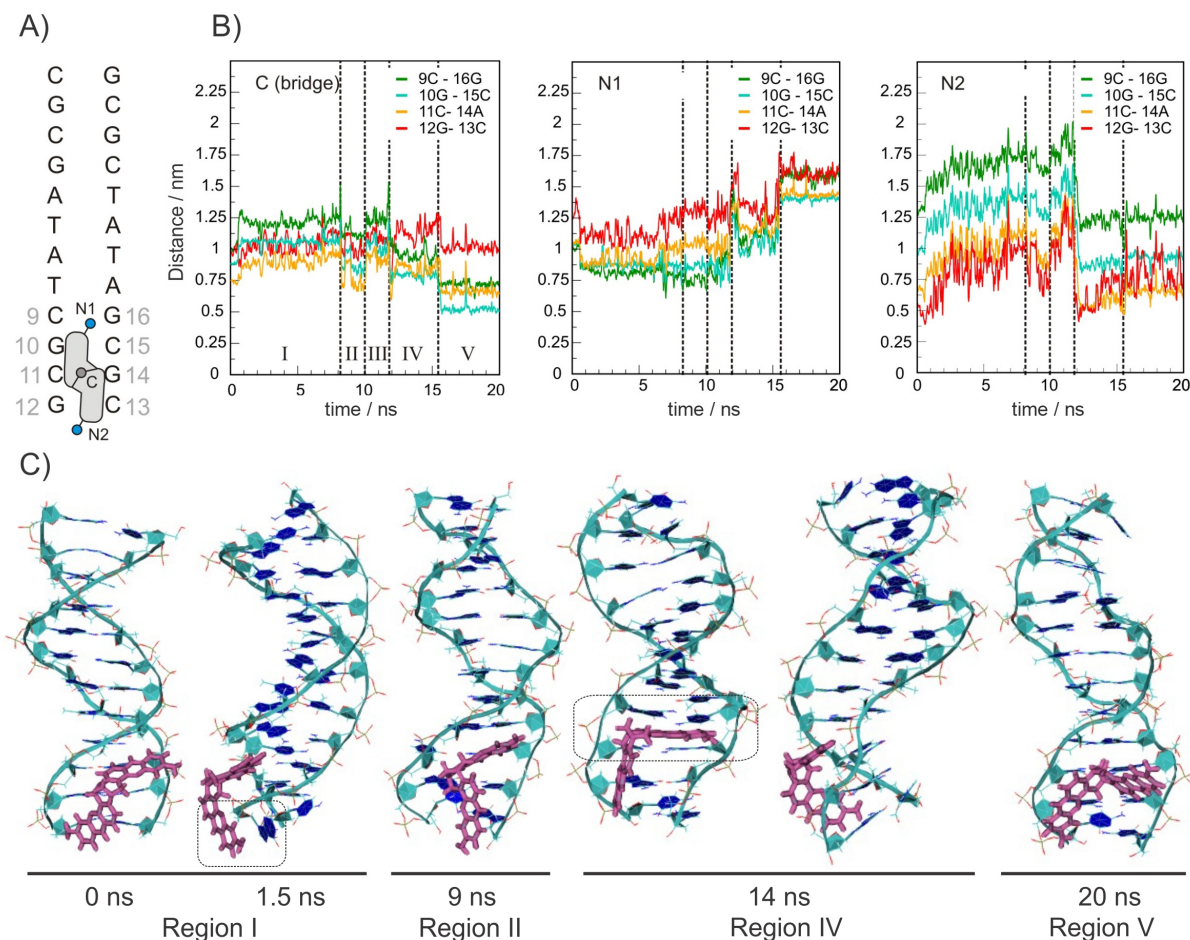


**Figure 3.** Time evolution of SYM-INT complex. A) Schematic representation of binding site and binding mode. B) Distance of diazocin bridge (C) or amino groups (N1 and N2) from the center of mass of the 4 base pairs belonging to the binding site. C) Structures extracted from simulated trajectories.

diazocin bridge are correlated with transitions observed for the amino groups, reflecting the structural rigidity of the Tröger scaffold. Considering that these transitions may reflect a change in the binding site or in the binding mode, it is interesting to perform a detailed analysis of each region, as described hereafter.

In the three first regions, the diazocin bridge remained closer to the 11C-14G base pair (orange line in Figure 4B) and the transitions in the diazocin bridge distances were not correlated with transitions in the distances concerning the amino groups. Therefore, one can conclude that both the binding site and the binding mode remained essentially the same during these stages of the simulation. Also noteworthy in region I is the fact that eventually one of the terminal base-pairs disrupted and started to interact with the aromatic rings of the Tröger base, as expected to occur in intercalative binding (see Figure 4C, 1.5 ns, dashed area). Obviously, this event is an artifact arising from the short extent of the oligomer and from the close proximity of Tröger base to the oligomer terminal region. Despite being artificial, however, this effect may be closely related to the intrinsic intercalative ability of proflavine moiety. In region

IV, distance analysis highlights a significant reorientation of Tröger base, which can be derived from the following observations: (i) the diazocin bridge becomes closer of 10G-15C instead of 11C-14G base-pair; (ii) the N1 amino group departed from 9C-16G and 10G-15C base-pairs, and (iii) the N2 amino group simultaneously approached all four base-pairs, indicating a movement of this group towards the central region of the oligomer. According to the structures extracted from region IV, this movement provided a better fit of the N2 proflavine into the minor groove, while it also positioned the N1 proflavine parallel to the 9C-16G and 10G-15C base-pairs (see Figure 4C, 14 ns, dashed area). In a side view of this same structure, it is possible to observe that this new orientation is very favorable to intercalative binding. Although it last for about 2 ns, intercalation did not occur. This observed behavior is consistent with a mechanism of intercalation similar to the proposed by Mukherjee *et al.*<sup>39</sup> for the case of daunomycin. According to this mechanism, there is a previous minor groove-bound state which is separated from the intercalated state by a free energy barrier of about 12 kcal mol<sup>-1</sup>. This could also explain why the event of spontaneous gap opening would



**Figure 4.** Time evolution of SYM-GROOVE complex. A) Schematic representation of binding site and binding mode. B) Distances of diazocine bridge (C) or amino groups (N1 and N2) from the center of mass of the 4 base-pairs belonging to the binding site. C) Structures extracted from simulated trajectories.

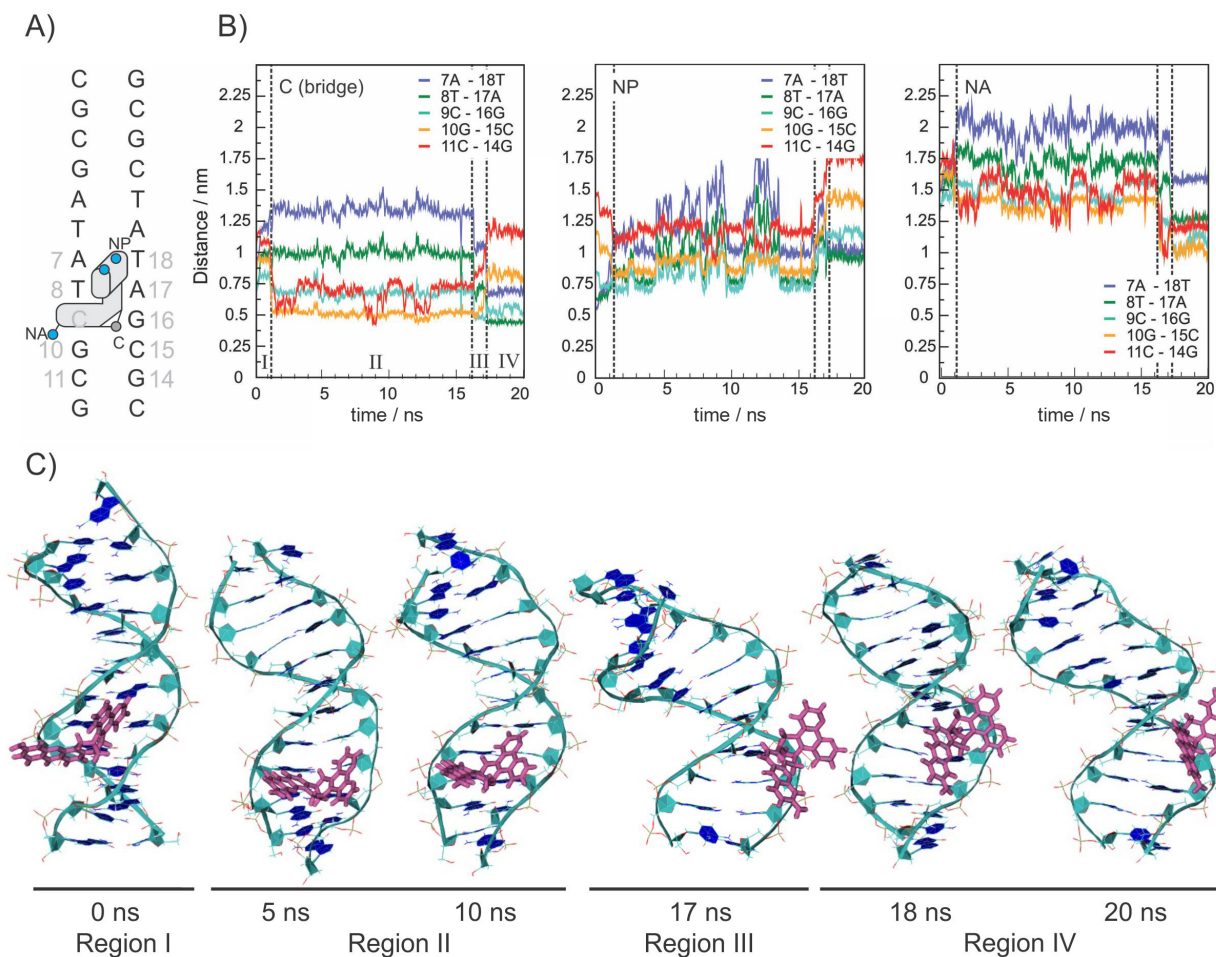
be expected to be extremely rare in the time scale of our simulations.

At 15.4 ns, it occurred an abrupt approach of diazocine bridge towards DNA - as can be seen from the simultaneous decrease in distances regarding the four monitored base-pairs - and an abrupt departure of N1 amino group (region V in Figure 4B). The resulting binding mode is illustrated in the structure at 20 ns (Figure 4C), in which the Tröger base interacts with DNA mainly through the diazocine bridge, while the amino groups from proflavine are projected outside. It is noteworthy that the diazocine bridge is now turned inside the minor groove, and not projected outside as it was in the beginning of the simulation.

Concerning the DNA structure, it was observed some local distortions, suggesting the groove binding may also perturbs DNA conformation (Figure 4C). In 1.5 ns, for instance, DNA endured a bending towards the major groove, while in 14 ns occurred an apparent enlargement of minor groove, which refer to A-DNA isoform. These distortions seem to be reversible, since they do not appear in the final structure (20 ns).

#### Minor groove complex of asymmetric Tröger base (ASYM-GROOVE)

In the beginning of the simulation, the Tröger base was partially docked in the minor groove, as shown in Figure 5A. As with the SYM-GROOVE complex, there was a significant reorientation of the asymmetric Tröger base inside the minor groove, which is better described by the distances regarding the diazocine bridge (Figure 5B). Initially (region I), the diazocine bridge was closer to the 9C-16G base-pair (cyan line in Figure 5B), at approximately 0.8 nm. After 1 ns (region II), it occurred a reorientation in which: (i) the diazocine bridge approached 9C-16G, 10G-15C and 11C-14G base-pairs, and (ii) the diazocine bridge turned to interact preferentially with 10G-15C base-pair, at approximately 0.5 nm. These observations clearly indicate changes not only in the binding site but also in the binding mode. Indeed, structures extracted from 5 and 10 ns show that, although the phenanthroline moiety remains close to the minor groove, the diazocine bridge is now projected inside the



**Figure 5.** Time evolution of ASYM-GROOVE complex. A) Schematic representation of binding site and binding mode. B) Distances of diazocine bridge (C), amino group (NA) and phenanthroline nitrogen (NP) from the center of mass of the 5 base-pairs belonging to the binding site. C) Structures extracted from simulated trajectories.

minor groove (Figure 5C). This binding mode remained stable for about 15 ns, with some mobility regarding the amino group (NA) and the phenanthroline nitrogen (NP), as can be observed by the distance fluctuations in Figure 7B.

After 15 ns (region III), again the diazocine bridge interacted preferentially with the 9C-16G base-pair, but in closer proximity as compared to region I. This transition is correlated with simultaneous transitions in other groups (NA and NP), indicating an approach of proflavine and a departure of phenanthroline moiety relatively to DNA. This new binding mode is illustrated in structure from 17 ns, where it can be clearly observed that the phenanthroline moiety is now projected outside the double helix, while the Tröger base interact with DNA mainly through the diazocine bridge (Figure 5C).

Soon after 17 ns (region IV in Figure 5B), there was a last reorientation, in which the diazocine bridge turned to interact closely with the 8T-17A base-pair, in closer proximity as compared with previous stages of the

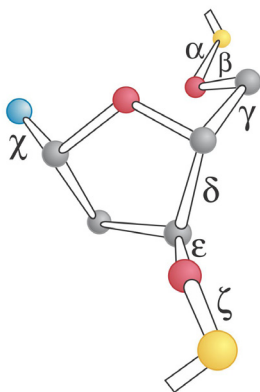
simulation. Simultaneously, there was a further departure of phenanthroline from DNA, so that phenanthroline finished the simulation significantly farther from the minor groove than it was in the initial structure. Summarizing, the distance analysis showed a trend of the Tröger base to interact with DNA mainly through the diazocine bridge (at 0.5 nm), with little contribution from the nitrogen groups to the interaction (at 1 nm), as can be observed in the structures from 18 and 20 ns (Figure 5C). It is noteworthy that this binding mode is very similar to the resulting binding mode from SYM-GROOVE complex.

Regarding the DNA double helix, the asymmetric Tröger base seems to induce an enlargement of minor groove and a shortening of the double helix. These structural changes could arise from local bending or from conformational changes towards A-DNA (Figure 5C), since structure from 17 ns shows plain resemblance to the A-DNA isoform, with a shallow minor groove and deep major groove. The torsion angles, together with base-pair step analysis, however, could not confirm neither a conformational change towards

A-DNA nor a consistent bending towards any groove (see Figure S7 in Supplementary Information). It is more likely, therefore, that the observed local distortions are due to the intrinsic flexibility of this oligomer.

### Backbone conformational analysis

To better evaluate the distortions previously observed in the simulation structures, we also monitored some backbone torsional angles. DNA backbone encompasses six torsional angles (Figure 6), which makes the torsional analysis for nucleic acids much more complex than for proteins.<sup>40,41</sup> Some of these angles, however, are of more prominent role than others in the conformational description of DNA. The combined analysis of  $\alpha$  and  $\gamma$  angles, for instance, is specially relevant for the description of low-twist conformations, which are important during DNA recognition by proteins and small molecules.<sup>42</sup> Moreover, the accumulation of non-canonical  $\alpha/\gamma$  conformations may lead to strong distortions in the double helix and thus monitoring these angles becomes an established stability criterion in molecular dynamic simulations.<sup>43,44</sup>



**Figure 6.** Backbone ( $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$  and  $\zeta$ ) and glycosidic ( $\chi$ ) torsion angles, adapted from reference 40. Phosphorus atoms are depicted in yellow, oxygen atoms in red, carbon (united) atoms in gray and nitrogen atoms in blue.

Figure 7 shows as scattered points the values of the  $\alpha$  and  $\gamma$  angles sampled in 4000 snapshots along simulations, for the base-pairs belonging to the symmetric Tröger binding site. The same results concerning the asymmetric Tröger base are very similar and thus are shown in Figure S2 (in Supplementary Information). The global minimum is defined in the area enclosed by the solid line and local minima are defined in the areas enclosed by dashed lines.<sup>42</sup> According to Várnai *et al.*,<sup>42</sup> these non-canonical  $\alpha/\gamma$  regions are essentially associated to altered values of roll and twist parameters, which describe curvature and winding of DNA, respectively.

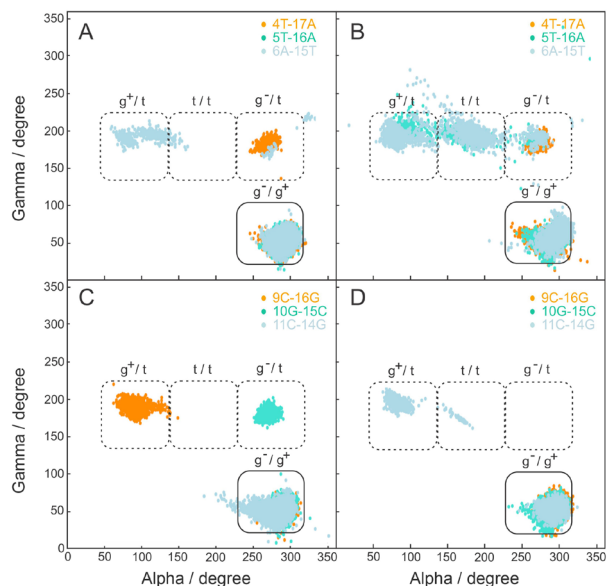
In all simulated systems (Figure 7A-D), the most densely populated region corresponds to the  $g-/g+$  ground state, although DNA also samples other local minima ( $g+/t$  and  $g-/t$ ). Indeed, the parm99 AMBER force field - upon which AMBER 03 force field is based - is known for overly favor these local minima, but the accumulation of such transitions only lead to severe and irreversible distortions in very long simulations ( $> 20$  ns).<sup>43</sup> Moreover, DNA did not get trapped in these conformations, since most of the transitions were reversible (Figure S3 - Supplementary Information).

In the intercalation complex, however, it is very clear that intercalation also favors  $\alpha/\gamma$  transitions to  $t/t$  region (Figure 7B), which is not populated when DNA is simulated alone (Figure 7A). Although  $t/t$  region is considered a local minimum, this conformation is usually found in A-DNA and is associated with a low-twist profile.<sup>45,46</sup> In agreement with the distortions previously observed in 3C. Since parm99 force field is known for presenting a low energy barrier between regions  $g+/t$  and  $t/t$ ,<sup>42</sup> this suggests that the chosen force field (AMBER 03) may favor  $\alpha/\gamma$  transitions into  $t/t$  region, thus facilitating the distortions required for intercalation binding mode. However, in a previous work we evaluated AMBER 03 torsional profiles and showed that, although  $\gamma:t$  is indeed a local minimum,  $\alpha:t$  conformation corresponds to a high energy barrier (30 kJ mol<sup>-1</sup>).<sup>44</sup> Although this estimate is only rough - the actual value of torsion angles also depend on the non-bonded interactions - it makes unlikely that  $t/t$  region is a local minimum in AMBER 03 force field, at least for B-DNA isoform. Therefore, it seems plausible to discard the hypothesis that intercalation is being overly stabilized by AMBER 03 torsional parameters.

On the other hand, minor groove binding does not significantly induce further  $\alpha/\gamma$  transitions in DNA, which remains mainly in  $g-/g+$  canonical region and in  $g+/t$  local minimum (Figure 7D). Apparently, transitions to  $g-/t$  local minimum are inhibited by minor groove binding, which may indicate a sort of stabilizing effect of the Tröger base upon DNA backbone.

Besides  $\alpha$  and  $\gamma$ , it were also monitored the  $\epsilon$  and  $\zeta$  angles, whose combination defines the two isoforms of B-DNA known as BI (more frequent) and BII (less frequent).<sup>45</sup> The  $\epsilon$  and  $\zeta$  angles are those who present the greater variability among the backbone torsional angles - thus being the most flexible.<sup>47</sup> In this way, monitoring the  $\epsilon/\zeta$  angles is not exactly a stability criterion, but rather a way to access DNA flexibility. Considering that flexibility is a very important feature not only in ligand discrimination but also to allow the double helix to adapt upon binding, it is crucial that the simulations reflect a balanced equilibrium between BI and BII isoforms.

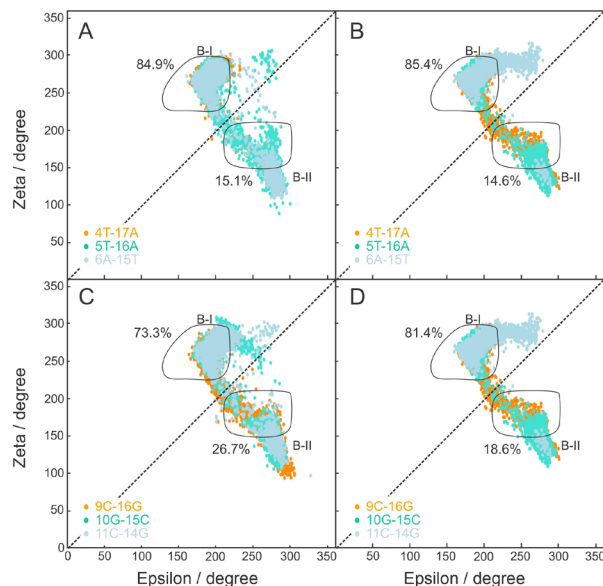




**Figure 7.**  $\alpha/\gamma$  sampling for nucleotides belonging to the symmetric Tröger binding site. A) DNA-GAP. B) SYM-INT. C) DNA-NOGAP. D) SYM-GROOVE. The global (solid line) and local (dashed line) minima are indicated, according to literature.<sup>42</sup>

Figure 8 shows as scattered points the values of the  $\epsilon$  and  $\zeta$  angles sampled in 4000 snapshots along simulations, for the base-pairs belonging to the symmetric base binding site. The results concerning the asymmetric Tröger base are very similar and thus are shown in Figure S4 (Supplementary Information). Regions corresponding to BI ( $\epsilon:t/\zeta:g-$ ) and BII ( $\epsilon:g-/\zeta:t$ ) conformations according to Schneider *et al.*<sup>45</sup> are shown. Since these angles display large fluctuations, however, the diagonal line can also be used as a coarse separation between BI ( $\epsilon - \zeta < 0$ ) and BII ( $\epsilon - \zeta > 0$ ) regions. Although it depends on the sequence, it is usually established that equilibrium between these two isoforms occurs in a proportion of 85:15 (BI:BII).<sup>47,48</sup>

As can be clearly seen in Figure 8, all systems display a balanced equilibrium between the two canonical isoforms, indicating that DNA intrinsic flexibility is being correctly sampled during simulations with AMBER 03 force field. Moreover, there is a recognizable “path” of points connecting BI and BII regions, which suggests the existence of a preferential energy path involving concerted changes in  $\epsilon$  and  $\zeta$  angles. Apparently, minor groove binding slightly restricts BII sampling, as can be observed by comparing Figures 8C and 8D. This difference, however, is of the same magnitude as that observed between the two oligomers simulated alone (Figures 8A and C and also in the Supplementary Information, Figures S4C and S4D), which is probably related to sequence properties alone. Therefore, it is reasonable to conclude that neither binding mode (intercalation nor minor groove binding) significantly enhances or restricts DNA backbone flexibility.



**Figure 8.**  $\epsilon/\zeta$  sampling for the nucleotides belonging to the symmetric Tröger binding site. A) DNA-GAP. B) SYM-INT. C) DNA-NOGAP. D) SYM-GROOVE. The BI and BII regions are indicated, as well as the percentual occurrence of each B- canonical conformation during the simulation time.

#### Base-pair step parameters

Among the several base-step parameters that describe local distortions in DNA double helix (for a review, see Dickerson and co-workers<sup>40,41</sup>), we chose to monitor rise and twist, which describe the distance and the torsion angle between two consecutive base-pairs, respectively. Both are good indicators of intercalative binding signature, since intercalation involves the opening of an intercalation gap and consequent unwinding of the double helix.<sup>6,37,38</sup>

Figure 9 shows as scattered points the distribution of the base-pair steps belonging to the symmetric Tröger base binding site in a rise/twist conformational space. Again, as the results concerning the asymmetric Tröger base showed very similar profiles, these can be found in Figure S5 (Supplementary Information).

As can be seen in Figure 9, all nucleotides remain close to the B-DNA canonical conformation, except for the base-pair step containing the intercalation gap. As already expected, this base-pair step occupies a different region in the rise/twist conformational space, presenting a high rise (7Å) and a decreased twist ( $-10^\circ$  to  $25^\circ$ ) - characteristics of the presence of an intercalation gap (Figure 9B).

In the DNA-GAP system (Figure 9A), it becomes evident that, although DNA presents an artificial intercalation gap in the beginning of the simulation, it quickly returns to the B-canonical conformation in the absence of an intercalative agent. Therefore, although there is a small sampling of a high-rise region in the rise/twist conformational space

- corresponding to snapshots from the beginning of the simulation - the most densely populated region corresponds to the B-canonical conformation (see orange points in Figure 9A).

It is also worth noting that, even for systems in which the torsion angles remained mainly in the B-canonical ground state (DNA-GAP, DNA-NOGAP and SYM-GROOVE), there was significant variations in the values of rise and twist (Figures 9A, 9C and 9D). This means that DNA intrinsic flexibility is not completely dependent on changes in the torsion angles. Moreover, the variations observed in these parameters during the simulations are greater than the differences between A and B-DNA, reinforcing the difficulties usually found in discerning between different DNA isoforms. In all systems, the average twist is underestimated ( $30^\circ$ ), remaining closer to A-DNA ( $32^\circ$ ) than to B-DNA ( $36^\circ$ ). Alone, however, this effect does not indicate a B to A transition, but is rather due to the force field itself, which is known to underestimate twist in  $3\text{--}4^\circ$  when compared to crystallographic values.<sup>49</sup>

We also performed a similar analysis regarding roll and slide parameters (Supplementary Information). Since these parameters influence the width and the depth of DNA grooves,<sup>50</sup> respectively, they can also be used to discern between A and B-DNA. Moreover, positive values of roll indicate a local bending towards major groove while negative values indicate a local bending towards minor groove.<sup>51</sup> Our results show that, in average, both roll and slide remain around the B-canonical minimum in the roll/slide conformational space (Supplementary Information, Figures S6 and S7). Although minor groove binding

promotes some perturbations in roll and slide (Figures S6D and S7D), these perturbations do not seem coherent enough to promote a consistent bending towards any groove.

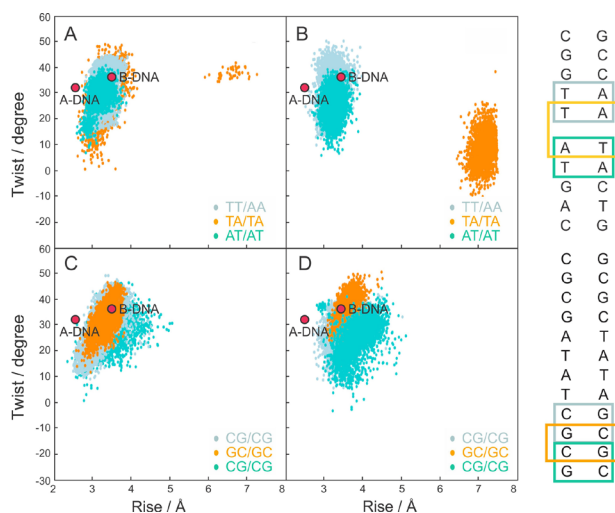
Starting from docked conformations, we performed molecular dynamics simulations with four different complexes involving the levorotatory isomers of two Tröger bases - two intercalation complexes and two minor groove complexes - in order to investigate the residence times of these two possible binding modes and also the structural changes each binding mode promotes on DNA double helix.

While intercalation complexes presented long residence times (at least 20 ns), minor groove complexes resulted in much shorter residence times, with changes both in the binding site and in the binding mode. From visual inspection, it was already noticeable that intercalative binding induces severe structural changes in DNA double helix, while the distortions observed in minor groove complexes are much more subtle and possibly due to the intrinsic DNA flexibility.

The distortions observed in intercalative binding were also confirmed by the analysis of torsion angles and base-pair step parameters. When intercalated, Tröger bases significantly affect the backbone conformation, leading DNA to non-canonical low-twist conformations. This occurs mainly through enhanced  $t/t$  sampling in the  $\alpha/\gamma$  conformational space for nucleotides containing the intercalation gap. Probably, these torsional changes reflect an adaptation of DNA geometry to accommodate the intercalation gap. Since the  $\epsilon/\zeta$  profile is not significantly affected by the presence of the Tröger bases, it is reasonable to assume that the distortions observed in the DNA backbone in intercalation complexes are correlated to  $\alpha$  and  $\gamma$  angles, not to  $\epsilon$  and  $\zeta$ . As expected, significant conformational changes were also observed in the rise/twist conformational space, showing that intercalation does affect the rise of the intercalation gap and the double helix torsion angle. On the other hand, the same analysis performed for minor groove complexes could not prove that this binding mode can induce any structural effect overcoming the intrinsic DNA flexibility. These observations, together with the longer residence times observed for intercalation, may be an evidence that minor groove is not as tight as intercalation binding mode.

Very interestingly, both intercalation complexes lead to similar binding modes in the end of the simulations, as also occurred with the minor groove complexes.

The intercalation complex of the symmetric Tröger base (SYM-INT), for instance, rapidly converged to a mixed binding mode very similar to the binding mode observed for the intercalation complex of asymmetric Tröger base (ASYM-INT) - with one substituent intercalated in the gap while the other kept further contacts in the minor



**Figure 9.** Rise/twist distribution for the nucleotides belonging to the symmetric Tröger binding site. A) DNA-GAP. B) SYM-INT. C) DNA-NOGAP. D) SYM-GROOVE. The regions corresponding to canonical A- and B-DNA are indicated. The base-pair steps analyzed in each oligomer are also depicted (right).

groove. Although the residence times of Tröger bases in the intercalation gap were long (at least 20 ns) and the distortions observed in the DNA were precisely those expected for an intercalating agent, these results alone do not prove that intercalation is the preferential binding mode of the Tröger levorotatory isomers. To prove this, it would be necessary not only to show that these molecules remain in the gap for long residence times, but also to prove that they can induce the opening of an intercalation gap in canonical DNA.

It is worth noting that, during the simulation of symmetric Tröger base with the oligomer that did not present intercalation gap (SYM-GROOVE), the aromatic rings of proflavine remained parallel to one base-pair step for about 2 ns, in a orientation very favorable to intercalate (see Figure 4C, 14 ns). Considering the main structural changes required for the opening of a intercalation gap ( $\alpha/\gamma$  transitions accompanied by significant changes in rise and twist), it is likely that this event involves an overly high energy barrier, which makes very small the probability of such an event being sampled in a nanosecond time scale. In addition to this, proflavine alone is a known intercalating agent and thus capable of induce an intercalation gap. If, even inserted in the Tröger scaffold, proflavine could assume an orientation favorable to intercalate, it is indeed likely that intercalation did not occur only due to the limited phase space sampling.

Another way to address this issue is to reason that, if proflavine alone is capable of intercalate, one can think of two main reasons for which intercalation would not be the preferential binding mode: (i) somehow, the Tröger scaffold sterically hinders proflavine intercalation or (ii) there exists an alternative binding mode which is energetically more stable than intercalation.

The first hypothesis seems not so plausible, since the simulation of SYM-GROOVE complex sampled a conformation very favorable to intercalation, for about 2 ns. Moreover, both intercalation complexes presented high residence times in the intercalating mode, which would not be expected if the Tröger scaffold represented an steric hindrance to intercalative binding.

The remaining hypothesis, thus, is that minor groove binding could be energetically more favorable than intercalation, being the preferential binding mode. Geometrically, however, the V-shaped Tröger scaffold provides a sharp bend which is, a priori, incompatible with the more mild bend of DNA minor groove. Even so, considering DNA intrinsic flexibility, one might expect that minor groove binding could induce a bending of DNA towards major groove, adapting minor groove to the almost 90° bend of Tröger scaffold. Clearly, this was not

observed in the simulations, since all structural analysis showed that DNA remained, in average, little altered during the simulations of minor groove complexes. Moreover, both minor groove complexes converged to a new and unexpected binding mode, in which the interaction occurs almost exclusively through the diazocin bridge, with both substituents projected outside the double-helix.

Counter-intuitively, this new binding mode does not take into account the potential complementarity between the B-DNA minor groove and the levorotatory Tröger base - both right-handed - and hence does not depend neither on the substituents nor on the Tröger chirality. It can not explain, therefore, the stereoselectivity observed for the levorotatory isomers, but may consists in a binding mode intrinsic of the Tröger scaffold and responsible for the weak and nonspecific binding observed for the dextrorotatory isomers.<sup>13-16</sup>

## Conclusions

According to our data, it is very unlikely that minor groove binding is the preferential binding mode for the two Tröger bases simulated in this study, which decreases the chances that another binding mode could overcome intercalation as the preferred binding mode. In this way, results from simulations suggest, although far from conclude, that the preferential binding mode of these Tröger base consists in the intercalation of one substituent, in accordance with the biochemical evidences found by Baldeyrou *et al.*<sup>16</sup> To confirm this hypothesis, meanwhile a crystallographic structure of a Tröger base complexed with DNA is still missing, it would be very interesting to develop a method for mapping the energy barrier associated with the opening of an intercalation gap.

## Acknowledgments

We acknowledge the financial support from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) - 474810/2008-9, Edital Universal and Bolsa de Produtividade.

## Supplementary Information

Supplementary Information (Figures S1-S7) is available free of charge at <http://jbcs.sbq.org.br> as a PDF file.

## References

1. Waring, M. J.; Wakelin, L. P. G.; *Nature* **1974**, 252, 653.

2. Stubbe, J.; *Chem. Rev.* **1987**, *87*, 1107.
3. Arcamone, F. M.; Animati, F.; Barbieri, B.; Configliacchi, E.; D'Alessio, R.; Geroni, C.; Giuliane, F. C.; Lazzari, E.; Menozzi, M.; Nicola, M.; Penco, S.; Verini, M. A.; *J. Med. Chem.* **1989**, *32*, 774.
4. Minotti, G.; Menna, P.; Salvatorelli, E.; Cairo, G.; Gianni, L.; *Pharmacol. Rev.* **2002**, *56*, 185.
5. Gurova, K.; *Future Oncol.* **2009**, *5*, 1685.
6. Hannon, M. J.; *Chem. Soc. Rev.* **2007**, *36*, 280.
7. Hurley, L. H.; *Nat. Rev. Cancer* **2002**, *2*, 188.
8. Tröger, J.; *J. Prakt. Chem.* **1887**, *36*, 225.
9. Spielman, M. A.; *J. Am. Chem. Soc.* **1935**, *57*, 583.
10. Valik, M.; Strongin, R. M.; Král, V.; *Supramol. Chem.* **2005**, *17*, 347.
11. Abella, C. A. M.; Rodembusch, F. S.; Stefani, V.; *Tetrahedron Lett.* **2004**, *45*, 5601.
12. Abella, C. A. M.; Benassi, M.; Santos, L. S.; Eberlin, M. N.; Coelho, F.; *J. Org. Chem.* **2007**, *72*, 4048.
13. Yashima, E.; Akashi, M.; Miyauchi, N.; *Chem. Lett.* **1991**, *20*, 1017.
14. Tatibouët, A.; Demeunynck, M.; Andraud, C.; Collet, A.; Lhomme, J.; *Chem. Commun.* **1999**, 161.
15. Bailly, C.; Laine, W.; Demeunynck, M.; Lhomme, J.; *Biochem. Biophys. Res. Commun.* **2000**, *273*, 681.
16. Baldeyrou, B.; Tardy, C.; Bailly, C.; Colson, P.; Houssier, C.; Charmantray, F.; Demeunynck, M.; *Eur. J. Med. Chem.* **2002**, *37*, 315.
17. Valik, M.; Malina, J.; Palivec, L.; Foltýnová, J.; Tkadlecová, M.; Urbanová, M.; Brabec, V.; Král, V.; *Tetrahedron* **2006**, *62*, 8591.
18. Claessens, N.; Pierard, F.; Bresson, C.; Moucheron, C.; Mesmaeker, A. K. D.; *J. Inorg. Biochem.* **2007**, *101*, 987.
19. Ricci, C. G.; Netz, P. A.; *J. Chem. Inf. Model.* **2009**, *49*, 1925.
20. Morris, G. M.; Googsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J.; *J. Comput. Chem.* **1998**, *19*, 1639.
21. Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S.; *J. Comput. Chem.* **2007**, *28*, 1145.
22. Lindahl, E.; Hess, B.; van der Spoel, D.; *J. Mol. Model.* **2001**, *7*, 306.
23. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P.; *J. Comput. Chem.* **2003**, *24*, 1999.
24. Sorin, E. J.; Pande, V. S.; *Biophys. J.* **2005**, *88*, 2472.
25. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A.; *J. Comput. Chem.* **2004**, *25*, 1157.
26. Batista, P. R.; Wilter, A.; Durham, E. H. A. B.; Pascutti, P. G.; *Cell Biochem. Biophys.* **2006**, *44*, 395.
27. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; 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.; Gonzalez, C.; Pople, J. A.; *Gaussian 03, Revision C.02*, Gaussian, Inc.: Wallingford, CT, 2004.
28. Baili, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A.; *J. Phys. Chem.* **1993**, *97*, 10269.
29. Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. J.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A.; *J. Comput. Chem.* **1993**, *14*, 1347.
30. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L.; *J. Chem. Phys.* **1983**, *79*, 926.
31. Darden, T.; York, D.; Pedersen, L. G.; *J. Chem. Phys.* **1993**, *98*, 10089.
32. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G.; *J. Chem. Phys.* **1995**, *103*, 8577.
33. Lu, X. J.; Olson, W. K.; *Nucleic Acids Res.* **2003**, *31*, 5108.
34. Humphrey, W.; Dalke, A.; Schulten, K.; *J. Mol. Graph.* **1996**, *14*, 33.
35. Watson, J. D.; Crick, F. H. C.; *Nature* **1953**, *171*, 964.
36. Tse, W. C.; Boger, D. L.; *Chem. Biol.* **2004**, *11*, 1607.
37. Canals, A.; Purciolas, M.; Aymami, J.; Coll, M.; *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2005**, *61*, 1009.
38. Waring, M. J.; Bailly, C.; *J. Mol. Recognit.* **1994**, *7*, 109.
39. Mukherjee, A.; Lavery, R.; Bagchi, B.; Hynes, J. T.; *J. Am. Chem. Soc.* **2008**, *130*, 9747.
40. Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M. L.; *Science* **1982**, *216*, 475.
41. Dickerson, R. E.; *Methods Enzymol.* **1992**, *211*, 67.
42. Várnai, P.; Djuranovic, D.; Lavery, R.; Hartman, B.; *Nucleic Acids Res.* **2002**, *30*, 5398.
43. Pérez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham III, T. E.; Laughton, C. A.; Orozco, M.; *Biophys. J.* **2007**, *92*, 3817.
44. Ricci, C. G.; de Andrade, A. S. C.; Mottin, M.; Netz, P. A.; *J. Phys. Chem. B* **2010**, *114*, 9882.
45. Schneider, B.; Neidle, S.; Berman, H. M.; *Biopolymers* **1997**, *42*, 113.
46. Gao, Y.-G.; Robinson, H.; Wang, A. H.-J.; *Eur. J. Biochem.* **1999**, *261*, 413.

47. Bertrand, H.-O.; Ha-Duong, T.; Femandjian, S.; Hartmann, B.; *Nucleic Acids Res.* **1998**, *26*, 1261.
48. Hartmann, B.; Piazzola, D.; Lavery, R.; *Nucleic Acids Res.* **1993**, *21*, 561.
49. Olson, W. K.; Zhurkin, V. B.; *Curr. Opin. Struct. Biol.* **2000**, *10*, 286.
50. Calladine, C.; Drew, H. R.; *J. Mol. Biol.* **1984**, *178*, 773.
51. Dickerson, R. E.; Chiu, T. K.; *Biopolymers* **1997**, *44*, 361.

*Submitted: November 25, 2011*  
*Published online: June 26, 2012*

# Supplementary Information

## Simulation of the Interactions between Tröger Bases and DNA

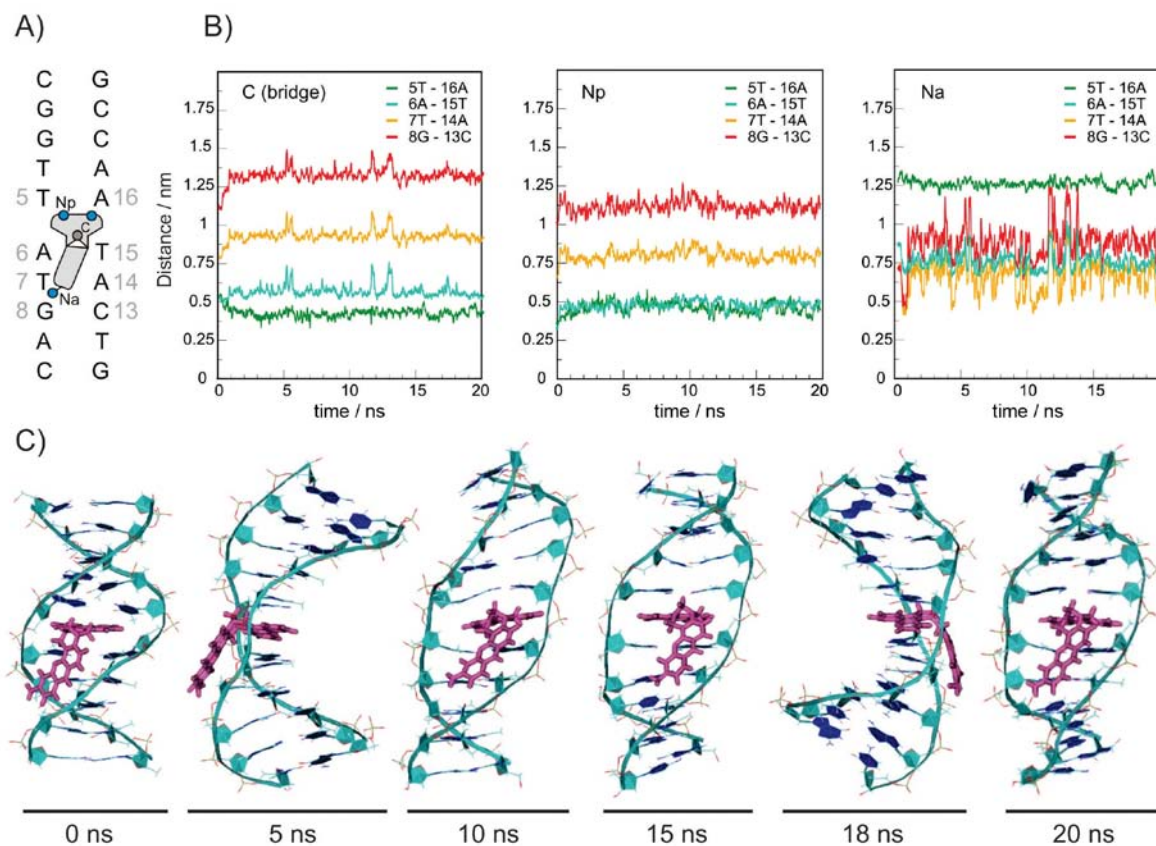
Clarisse G. Ricci and Paulo A. Netz\*

Instituto de Química, Universidade Federal do Rio Grande do Sul, 91501-970 Porto Alegre-RS, Brazil

### Intercalation Complex of Asymmetric Troger Base (ASYM-INT)

During the simulation of this complex, the phenanthroline moiety remained intercalated in the gap while the proflavine kept interactions with DNA minor groove, as schematized in Figure S1A. According to the distances (Figure S1B), both the diazocin bridge (C) and the Np group were stably equidistant from the base-pairs containing the gap, which confirms the intercalation (see cyan and green lines in Figure S1B).

Structures extracted from the simulation show that DNA double helix was relatively well conserved, with less zigzag distortions in backbone as compared to those observed in the intercalation complex of symmetric Troger base (see Figure 5C). Eventually, a slight bending towards the major groove can be observed (see structures at 5 and 18 ns in Figure S1C). As with SYM-INT complex, it was observed a decrease in the twist angle, which can be clearly observed at 10, 15, and 20 ns. In this case, however, unwinding was restricted to the base-pairs flanking the intercalation gap, without propagation to the terminal regions of the oligomer.



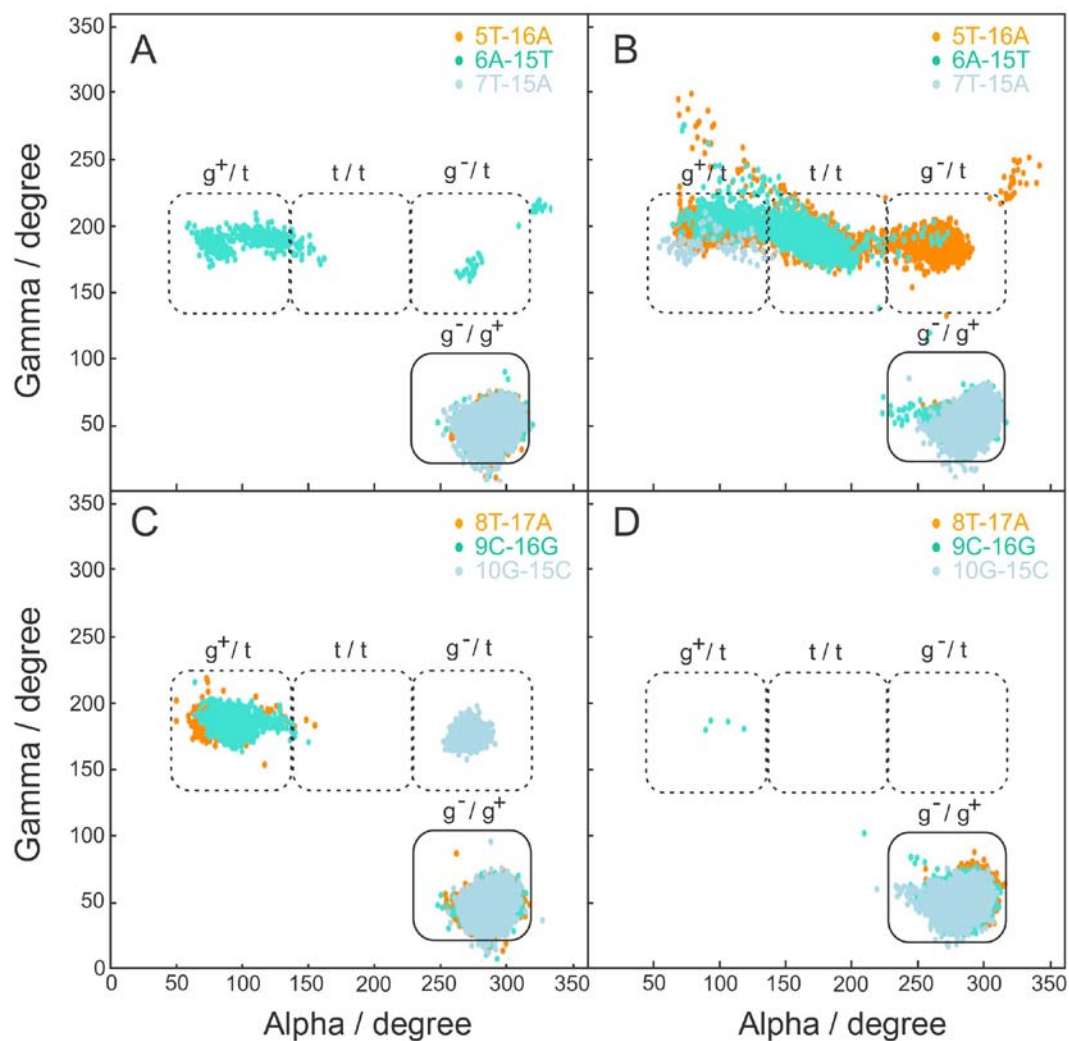
**Figure S1.** Time evolution of ASYM-INT complex. A) Schematic representation of binding site and binding mode. B) Distances of diazocine bridge (C), amino group (NA) and phenanthroline nitrogen (NP) from the center of mass of 4 base-pairs belonging to the binding site. C) Structures extracted from simulated trajectories.

\*e-mail: netz@iq.ufrgs.br

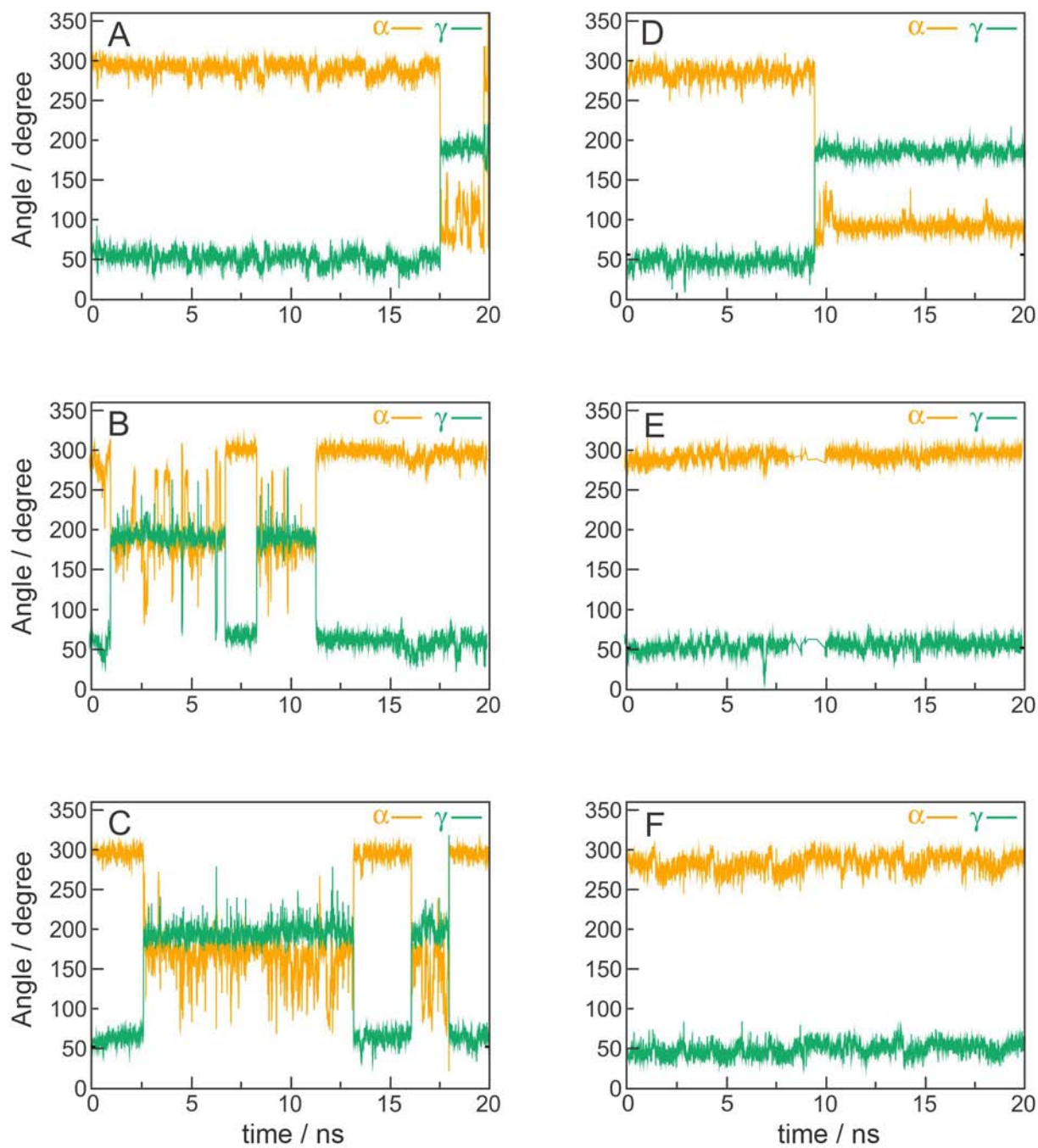
$\alpha/\gamma$  Analysis of ASYM-INT and ASYM-GROOVE

As occurred with SYM-INT complex, intercalation of asymmetric Tröger base also induces several  $\alpha/\gamma$  transitions to the t/t region, confirming this low-twist conformation is

important during small-molecule intercalative binding. In the minor groove binding, a contrary effect was observed, since the Tröger base seems to inhibit  $\alpha/\gamma$  transitions into non-canonical conformations and DNA remains mainly in the g<sup>-</sup>/g<sup>+</sup> ground state during the entire simulation.

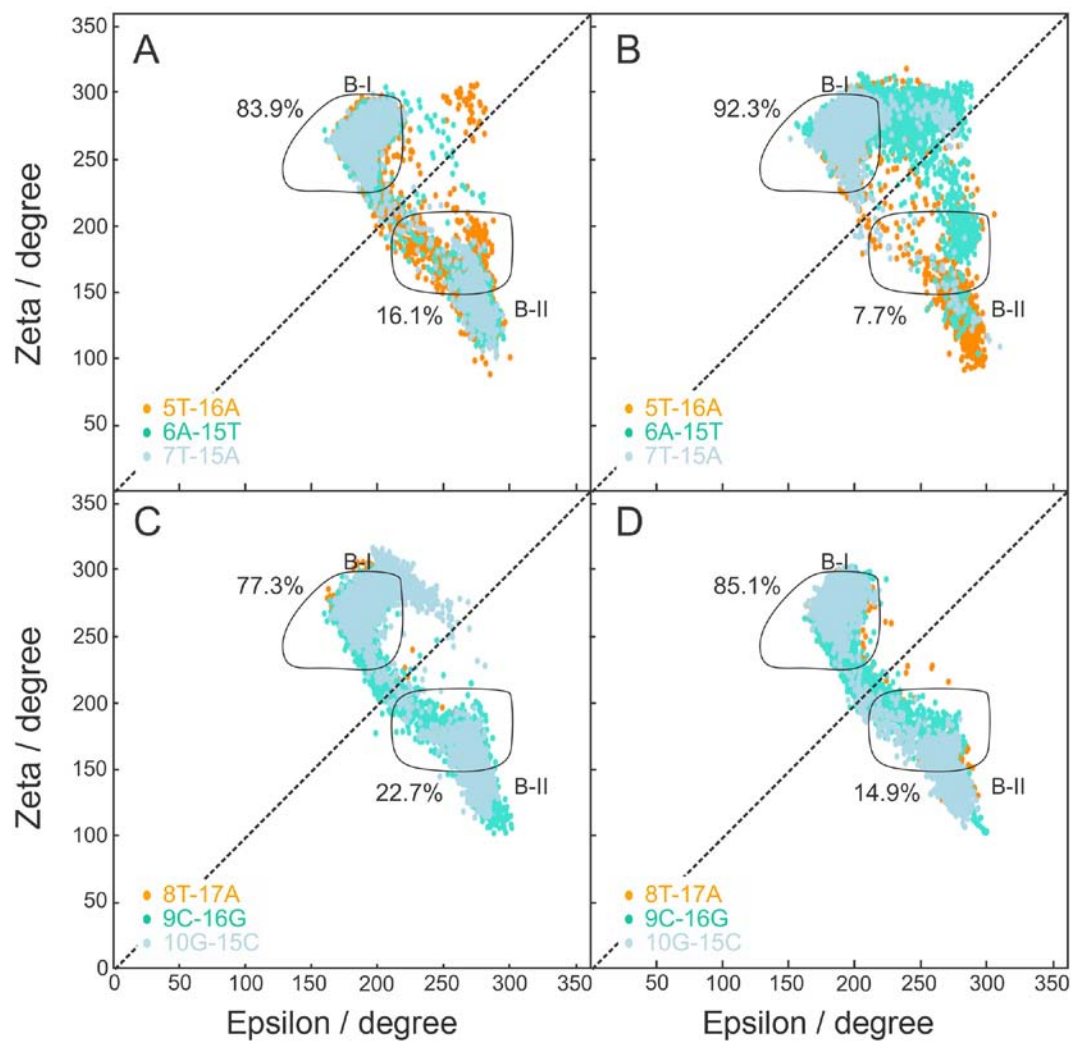


**Figure S2.**  $\alpha/\gamma$  sampling for the nucleotides belonging to the asymmetric Tröger binding site. A) DNA-GAP; B) ASYM-INT; C) DNA-NOGAP; D) ASYM-GROOVE. The global (solid line) and local (dashed line) minima are indicated.



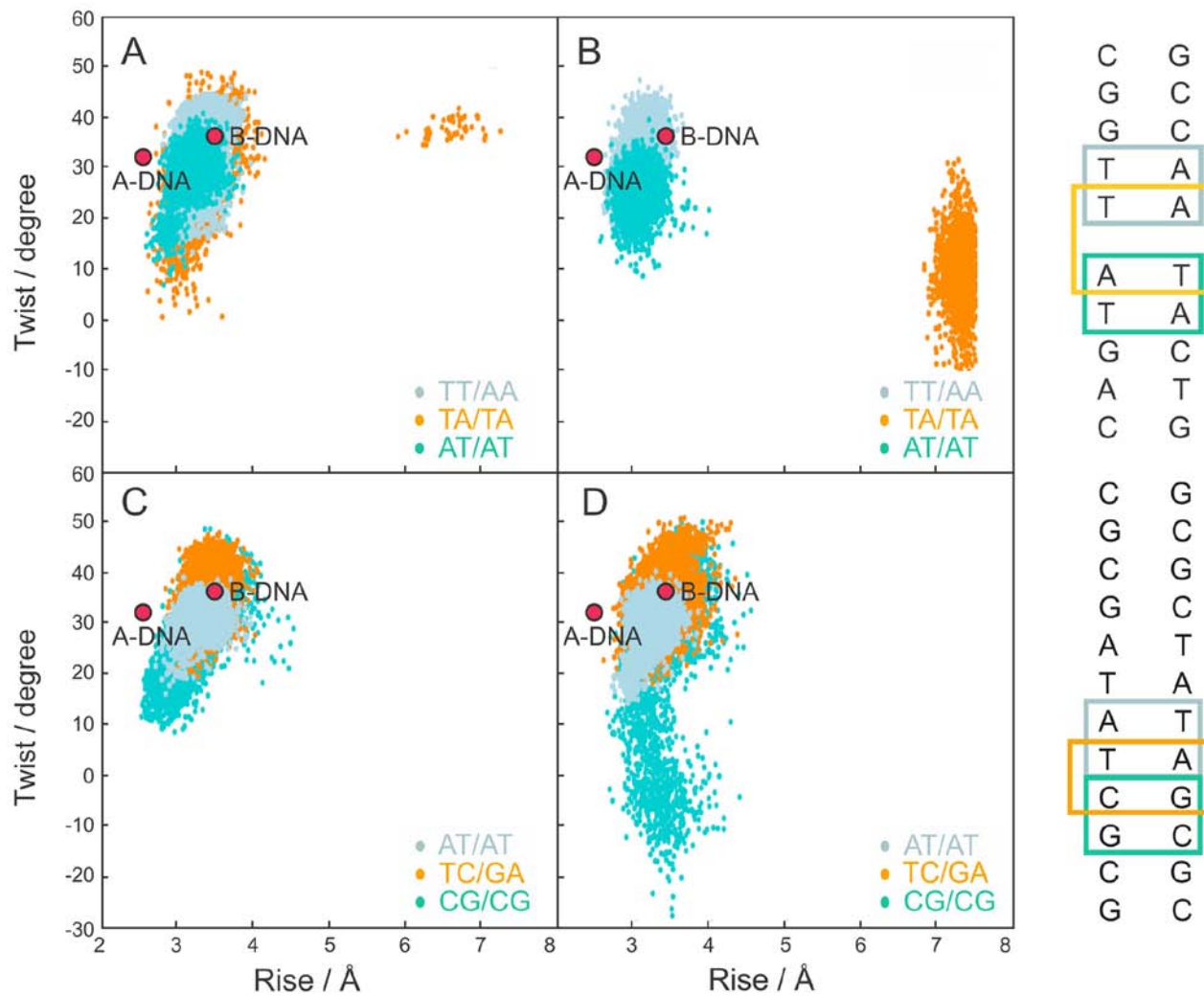
**Figure S3.** Temporal analysis of  $\alpha$  (orange) and  $\gamma$  (green) transitions. The left column shows transitions for 6A nucleotide in DNA-GAP (A); SYM-INT (B); and ASYM-INT (C). The right column shows the transitions for 16G nucleotide in DNA-NOGAP (D); SYM-GROOVE (E); and ASYM-GROOVE (E).



$\epsilon/\zeta$  Analysis of ASYM-INT and ASYM-GROOVE

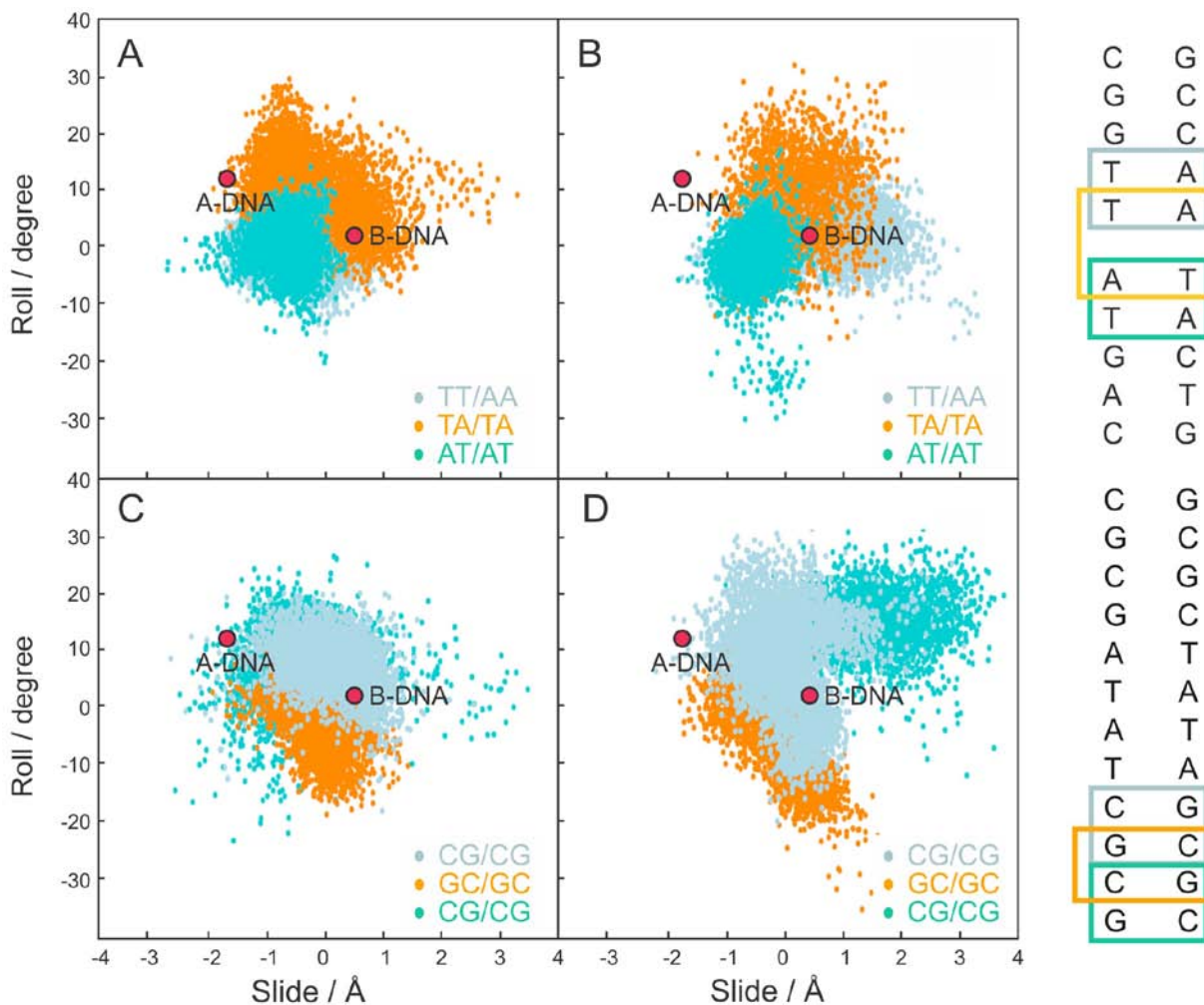
**Figure S4.**  $\epsilon/\zeta$  sampling for the nucleotides belonging to the asymmetric Tröger binding site. A) DNA-GAP; B) ASYM-INT; C) DNA-NOGAP; D) ASYM-GROOVE. The BI and BII regions are indicated, as well as the percentual occurrence of each canonical conformation during the simulation time.

## Rise/twist analysis of ASYM-INT and ASYM-GROOVE

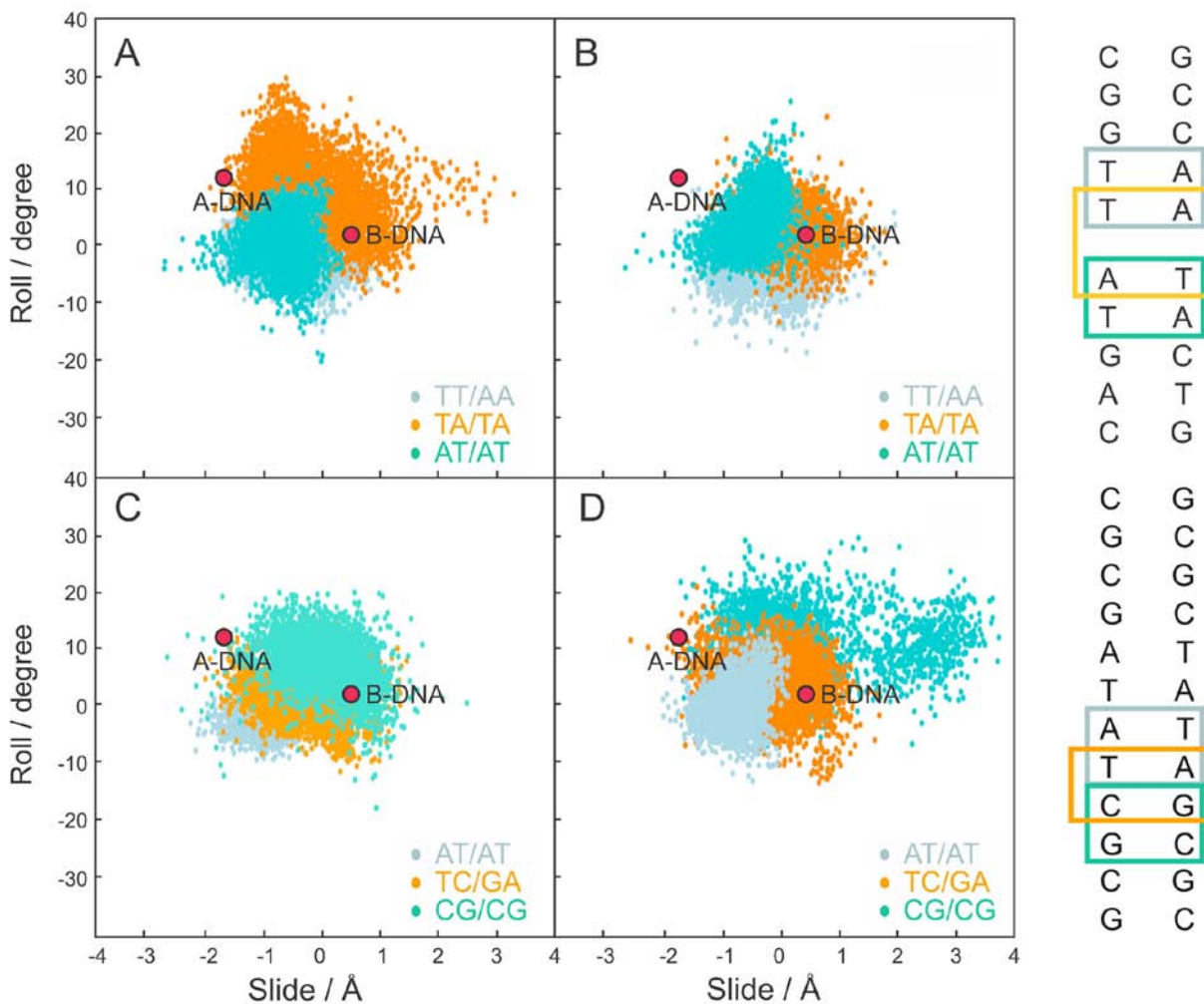


**Figure S5.** Rise/twist distribution for the base-pair steps belonging to the asymmetric Troger binding site. A) DNA-GAP; B) ASYM-INT; C) DNA-NOGAP; D) ASYM-GROOVE. The regions corresponding to canonical A- and B-DNA are indicated.

## Roll/Slide Analysis



**Figure S6.** Slide/roll distribution for the base-pair steps belonging to the symmetric Tröger binding sites. A) DNA-GAP; B) SYM-INT; C) DNA-NOGAP; D) SYM-GROOVE. The regions corresponding to canonical A- and B-DNA are indicated.



**Figure S7.** Slide/roll distribution for the base-pair steps belonging to the asymmetric Troger binding sites. A) DNA-GAP; B) ASYM-INT; C) DNA-NOGAP; D) ASYM-GROOVE. The regions corresponding to canonical A- and B-DNA are indicated.