Ligand binding mode to duplex and triplex DNA assessed by combining electrospray tandem mass spectrometry and molecular modeling.

Frédéric Rosu[1], Nguyen Chi-Hung[2], Edwin De Pauw [1]

& Valérie Gabelica[1*]

1: Laboratoire de Spectrométrie de Masse, Université de Liège, 3 allée de la chimie, 4000 Liège (Belgique).

2: UMR 176 CNRS-Institut Curie, Laboratoire de Pharmacochimie, Section de Recherche, Batiment 110, Centre universitaire, 91405 Orsay, France.

* Correspondance should be addressed to: v.gabelica@ulg.ac.be

Tel: +32-4 3663432 Fax: +32-4 3663413

ABSTRACT

In this paper we report the analysis of seven benzopyridoindole and benzopyridoquinoxaline drugs binding to different duplex DNA and triple helical DNA, using an approach combining electrospray ionization mass spectrometry (ESI-MS), tandem mass spectrometry (MS/MS), and molecular modeling. The ligands were ranked according to the collision energy (CE_{50}) necessary to dissociate 50% of the complex with the duplex or the triplex in tandem MS. To determine the probable ligand binding site and binding mode, molecular modeling was used to calculate relative ligand binding energies in different binding sites and binding modes. For duplex DNA binding, the ligand-DNA interaction energies are roughly correlated with the experimental CE_{50}, with the two benzopyridoindole ligands more tightly bound than the benzopyridoquinoxaline ligands. There is however no marked AT vs. GC base preference in binding, as supported both by the ESI-MS and the calculated ligand binding energies. Product ion spectra of the complexes with triplex DNA show only loss of neutral ligand for the benzopyridoquinoxalines, and loss of the third strand for the benzopyridoindoles, the ligand remaining on the duplex part. This indicates a higher binding energy of the benzopyridoindoles, and also shows that the ligands interact with the triplex via the duplex. The ranking of the ligand interaction energies compared with the CE_{50} values obtained by MS/MS on the complexes with the triplex clearly indicates that the ligands intercalate via the minor groove of the Watson-Crick duplex. Regarding triplex vs. duplex selectivity, our experiments have demonstrated that the most selective drugs for triplex share the same heteroaromatic core.

INTRODUCTION

DNA-binding reagents exhibit a high potential as chemotherapeutic drugs, and many approved anticancer therapies include molecules that are DNA binders. These agents interfere with gene replication or transcription in proliferating cells like cancer cells [1]. However, classical chemotherapies are still very toxic for healthy cells as well, and cause many unwanted effects. One of the most challenging goals is therefore the design of molecules which bind to nucleic acids with high structural and sequence selectivity in order to target specific disease-related genes. Triplex DNA formation is one such strategy to target specific DNA sequences [2-6]. Triple helices of polyribonucleotides were first observed in 1957 [7]. DNA triplexes are formed when a DNA strand called the antigene binds to a duplex (the gene) in its major groove. Combining triplex-specific ligands with antigenes is therefore a means of targeting genes more specifically than with duplex-binding agents [8,9] .

The interaction between the drugs and the nucleic acids can be studied when the physical properties of the ligand molecule change upon binding and are easily monitored. Spectroscopic techniques (UV-Visible absorption spectroscopy, fluorescence, circular dichroism) [10,11], equilibrium dialysis [12,13], surface plasmon resonance, [14,15] or calorimetric techniques [16,17] are commonly used to study ligand-DNA interactions. Electrospray ionization mass spectrometry has also shown its potential in the evaluation of DNA-ligand binding for more than a decade [18-26]. The study of the interaction of families of ligands with particular DNA structure like duplex and quadruplex structures can be made quickly. In these ESI-MS assays, the relative intensities of the free DNA and of the complexes is taken as a picture of the relative abundances of these species in solution [27], and this approach has proven valid when the ESI-MS data are compared with classical spectroscopic techniques [28-32].

The present paper reports for the first time the use of ESI-MS and ESI-MS/MS to study triplex DNA with noncovalently bound ligands. The ligand structures are shown in Scheme 1. The benzopyridoindole (BPI) ligands **1** and **2** are representatives of the benzo[*e*]pyridoindole (BePI, **1**) and of the benzo[*g*]pyridoindole (BgPI, **2**) families [33,34]. These molecules interact with duplex DNA [35], triplex DNA [34,36,37], and inhibit topoisomerases I and II [38]. Benzo[*f*]pyridoquinoxaline derivatives (BPQs, **3**-**5)** were developed in order to target triplex DNA more specifically [39,40], while pyridoquinoxaline derivatives (PQ, **7**) served to assess the role of the extra benzo ring. All these ligands are DNA intercalators [34,41,42]. Encouraged by the promising results obtained with BPQs, a new benzoquinoquinoxaline (BQQ, **6**), was designed to optimize triplex interaction, synthesized, and successfully tested [43].

Our goal here was to evaluate binding affinity and selectivity of these ligands for DNA duplexes of varying GC content, and for a triplex DNA, and to evaluate how tandem mass spectrometry (MS/MS) combined with molecular modeling can be used in order to obtain information on the ligand binding site. MS/MS data are interpreted using the transition state theory concepts [44-46]. The dissociation rate of the complex depends (1) on an enthalpy term which is the energy difference between the reactant (i.e. intact complex) and the transition state (the state at which dissociation becomes irreversible) (2) an entropy term which reflects the probability of the dissociation pathway and which depends on the mechanism and (3) on the internal energy. Consequently, when internal energy and dissociation pathways are the same, the dissociation rate should be related to the interaction energy in the complex. For small complexes, these interactions energies can be calculated by quantum chemical methods, but for large complexes like those DNA complexes studied here, molecular mechanics approaches are more practical. We will show here that, when the experimental MS/MS relative dissociation rates are compared to relative interaction energies calculated for different structural models, some structural interpretation can be made on the ligand binding mode.

MATERIALS AND METHODS

Materials

The self-complementary duplexes Dk33 $(dCGTAAATTTACG)_2$, Dk66 $(dCGCGAATTCGCG)_2$, and Dk100 $(dCGCGGGCCCGCG)_2$ were prepared in 100 mM aqueous NH_4OAc. All stock solutions were diluted to 50 μM at neutral pH. The triple helical DNA was prepared in 150 mM NH_4OAc acidified with acetic acid (pH = 5.5), as previously described [47]. The triplex was formed from single strands dCCTTTTCTCTTTCC (T1), dGGAAAGAGAAAAGG (T2), which constitute a Watson-Crick duplex (T1•T2), and the strand dCCTTTCTCTTTTCC (T3) which is the antigene strand. Sequence T3 is the reverse of sequence T1. The oligonucleotides were purchased from Eurogentec (Seraing, Belgium) and used without further purification.

The synthesis of the benzo[*e*]pyridoindole **1** (CA: 125974-68-7) and the benzo[*f*]pyridoindole **2** (CA: 170890-29-6) was described in 1990 [33], the synthesis of the benzo[*f*]pyridoquinoxalines **3** (CA: 165548-08-3), benzo[*h*]pyridoquinoxalines **4** (CA: 165548-10-7) and **5** (CA: 165548-11-8) and of the pyridoquinoxaline **7** (CA: 165548-02-7) was described in 1995 [39], and the synthesis of the benzo[*f*]quinoquinoxaline **6** (CA: 373595-26-7) was reported in 1998 [43]. The drug stock solutions were 100 or 200 μM in bi-distillated water.

For ESI-MS, drug–duplex mixtures of 10 μM DNA and 15 μM drug were prepared in 150 mM NH_4OAc and 15% methanol. The small proportion of methanol used (15%) does not disturb the particular DNA structures used in our assay as verified by CD (data not shown) and the high ionic strength used prevents the AT-rich duplex and the triplex from unfolding. The concentrations of the DNA stock solutions were checked no more than 3 days before the ESI-MS experiments by UV absorbance measurements.

Mass Spectrometry

ESI-MS experiments were performed on a LCQ instrument (Finnigan LCQ ion trap instrument (ThermoFinnigan, SanJose, CA) equipped with its standard heated capillary electrospray source or a Q-TOF Ultima Global (Micromass, now Waters, Manchester, UK) with its standard ESI source. On the LCQ, the needle voltage was set to -3.9 kV. The capillary was heated to 180 °C and the applied potential was -10 V. The skimmer was at ground potential and the tube lens offset was maintained at 40V. On the Q-TOF, in negative ion mode the capillary voltage was set to −2.2 kV and the cone voltage to -35 V and the RF Lens1 to -70 V. The hexapole collision voltage of 10 V was used for full scan MS. The affinity of the drug for a given structure is deduced from the concentration of bound ligand per DNA molecule [27,30]. The concentration of bound ligand per DNA molecule is calculated from the relative intensities of the free DNA and of the complexes using the following equation (1):

$$[\text{Bound Ligand}] = C_0*(I_{(1:1)}+2I_{(2:1)}+3I_{(3:1)})/(I_{(DNA)}+I_{(1:1)}+I_{(2:1)}+I_{(3:1)}) \quad (1)$$

where C_0 is the starting DNA concentration (expressed in duplex or triplex concentration), $I_{(DNA)}$ is the relative intensity of the free DNA, and $I_{(n:1)}$ are relative intensities of the complexes (*n* drug molecules bound to one DNA structure). The concentrations of bound ligand were found independent on the instrument used.

MS/MS experiments were all performed on the Q-TOF Ultima Global. This choice was done for two reasons. First, the mass range of the LCQ is limited to 2000 m/z and some triple helix fragments are observed m/z > 2000. Second, thanks to the higher-energy collision regime of the Q-TOF, direct noncovalent bond breaking is favored compared to rearrangement reactions giving neutral base losses [48]. The parent ion of interest was selected in the first quadrupole, and the hexapole collision voltage was varied. The argon pressure in the collision hexapole (3.0×10^{-5} mbar ± 5%) and the source

pressure (2.70 mbar) were carefully kept constant. Source block and desolvation temperatures were set to 70 °C and 100 °C, respectively.

Molecular modeling

Hyperchem 7.5 software (Hypercube, Inc.) has been used with AMBER99 force field. The starting duplex d(CGCGAATTCGCG)$_2$ was the solution structure deposited in the Protein Data Bank with code 1GIP [49]. As these ligands are known to bind DNA by intercalation [34,41,42], three different intercalation sites on the duplex d(CGCGAATTCGCG)$_2$ have been independently generated as follows. First, an increase of the space between two base pair by 3.4 Å was performed. Second, the twist angle of the helix (including one of the two base pair involved in the intercalation site) was decrease by 10 degree to compensate the change in the distance between the sugar-phosphate backbones. These parameters were chosen based on the structures of complexes between duplex DNA and intercalators found in PDB entries 2DES [50] and 1D10 [51]. Finally, local geometry optimization including the two base pairs and the backbone involved the intercalation site and the sugar-phosphate surrounding the site was performed to relax the system. Local geometry optimization was performed to relax the backbone surrounding the intercalation site. The intercalation sites were CG*CG, CG*AT and AT*AT. The global effect of the intercalation site on the duplex was the increase of the dimension of the grooves due to the decrease of the twisting angle of the two base pairs of the intercalating site. Each drug was manually docked in the intercalation site in different orientations as described in the text below, and energy minimized in the force field generated by the duplex until an energy gradient of 0.05 kcal/(mol.Å) was reached (Polak-Riebiere conjugate gradient algorithm). The maximum gradient of the generated ligand-DNA was never higher than 0.3 kcal/(mol.Å) which is satisfactory (the DNA without drug has at a gradient of 0.2 kcal/(mol.Å) and indicated that no steric clash/problem is encountered. The same methodology as for the duplex was used to generate the intercalation sites and calculate the

interaction energies between the drugs and the triplex. The starting triplex model was generated using a smaller triple helical DNA (7 triplets) based on the NMR solution structure of a pyrimidine-purine-pyrimidine DNA triplex (PDB entry 149D) [52]. Two intercalation sites between the base-triplet were generated: CGC-TAT and TAT-TAT.

The interaction energy was calculated using the following equation (2):

$$E_{int} = E_{(complex)} - E_{(Ligand)} - E_{(DNA)} \quad (2)$$

Where $E_{(complex)}$ is the complex energy, $E_{(ligand)}$ is the energy of the ligand with its geometry in the complex, and $E_{(duplex)}$ is the energy of the duplex with its geometry in the complex.

RESULTS AND DISCUSSIONS

ESI-MS determination of ligand selectivity.

The sequence and structural selectivities of the seven compounds were first probed using ESI-MS (Figure 1). Figure 1(a) shows the ESI-MS spectra obtained with the duplex d(CGCGGGCCCGCG)$_2$ (DK100) or the triplex and ligand **6**. <The 1:1 and 2:1 complex are detected. The amount of ligand **6** bound to the triplex under the experimental conditions is higher than to the duplex. We have previously shown a methodology to quantify the amount of bound ligand in DNA-ligand mixtures using ESI-MS [27]. The methodology is based on the assumption that the relative intensities of the free DNA and of the complexes are proportional to the relative abundances in solution. This assumption has been shown valid in the case of double-stranded DNA complexed with minor groove binders [27,53]. The equilibrium binding constants for the 1:1 and 2:1 complexes with **6** and the triplex measured by ESI-MS were $K_1 = 1.2 \times 10^5$ M^{-1}, $K_2 = 3.9 \times 10^4$ M^{-1}, respectively. In

comparison the values of the binding constants for the 1:1 and 2:1 complexes with **6** and the duplex were $K_1 = 3.2 \times 10^4$ M^{-1} and $K_2 = 3.7 \times 10^4$ M^{-1}. This example shows the potentiality of ESI-MS to detect and quantify multiple drug-DNA stoichiometries present at equilibrium in solution.

The calculated amounts of ligand bound to the different DNA structures are summarized in Figure 1(b). All drugs interact with the three DNA duplexes. Although intercalators usually show a preference for GC base pairs, this is not the case here, where even a small preference for AT-rich sequences is observed for some ligands, in agreement with footprinting studies [36]. The two benzopyridoindoles **1** and **2** have the highest duplex and triplex affinities compared to the other drugs. However, the BPI ligands (**1** and **2**) show no marked selectivity for the triple helix compared to duplex DNAs. Ligands **4**, and **6** show substantial selectivity for the triplex while ligand **5** shows moderate selectivity. The aromatic rings of these three ligands share a common tetracyclic core. Ligands **4** and **5** belong to the 8-amino-benzo[*f*]pyrido[3,4-b]quinoxaline family and are much more selective than **3** (11-amino-benzo[*f*]pyrido[4,3-b]quinoxaline) where the benzene ring is linked to the other side on the quinoxaline. Ligand **7**, which has the smallest heteroaromatic ring, has the lowest affinity for both DNA structures. Taken together, these observations suggest that the shape of the heteroaromatic ring system of these drugs therefore plays an important role in stacking interactions with the base triplets.

ESI-MS/MS and molecular modeling on the (1:1) duplex-ligand complexes

In the case of duplex-drug complexes, we have previously classified drugs in three categories, based on their dissociation pattern in tandem mass spectrometry experiments [54]. The first group includes drugs for which the complex dissociates mainly via the loss of a neutral drug. For the second group, the complex dissociates mainly via the loss of a negatively charged drug and for the third group via the separation of the strands (noted ss) which share the available charges and some drug molecule could stick on both strands. Here we have performed MS/MS experiments on the 1:1 (Drug-Dk66)

complex (Figure 2). All the complexes with each of the seven drugs dissociate mainly via the loss of a neutral drug.

$$[DK66+Drug]^{5-} \rightarrow [Dk66]^{5-} + Drug \quad (3)$$

In the case where the dissociation mechanism is the same for a series of ligands, the CE_{50} values are directly related to the activation energy of dissociation. This condition is fulfilled here as (1) the dissociation pathways are identical and (2) the parallel dissociation curves indicate similar activation entropies. Moreover, if the dissociation involves the loss of neutral ligand from the charged duplex, which is the case here, the activation energy is more likely to be proportional to the ligand binding energy [54].

The dissociation pattern observed for all ligands is the same as for the drugs of the first group like cryptolepine, proflavine, daunomycin and doxorubicin [54]. Figure 2(a) clearly shows the larger collision energy necessary to dissociate the complexes with the BgPI **2** compared to the BPQ **3**. At 14 eV, the complex with **3** is nearly completely dissociated while only 15% of the complex with **2** is dissociated. Figure 2(b) represents the MS/MS dissociation curves obtained for all ligands.

The percentage of intact complex is calculated using the following equation (4):

$$\%Complex = \frac{I_{([DK66+Drug]^{5-})}}{I_{([DK66+Drug]^{5-})} + I_{(DK66^{5-})}} \quad (4)$$

The competitive dissociation of the duplex into the single strands (ss^{3-} and ss^{2-}) is only observed at collision energies > 18 eV, so this fragmentation channel does not perturb the ordering of CE_{50}. The two BPI ligands **1** and **2** leave the duplex at higher energy than the quinoxaline derivatives (**3**-**7**).

Among the quinoxaline derivatives, **6** has the largest interaction energy with its binding site on the duplex (since higher collision energy is needed to dissociate the complex as compared to the other BPQ drugs).

We performed molecular modeling to compare the interaction energies of a drug with different binding site sequences on duplex d(CGCGAATTCGCG)$_2$ and to correlate the calculated interaction energies to the relative CE_{50} values obtained by MS/MS. Several groups have examined the molecular recognition and the dynamics of drug-DNA interactions using molecular modeling [55-57]. As the energetic parameters calculated by molecular modeling are related to the gas phase, the calculation using molecular modeling of the energetic of drug-DNA complexes allows a direct comparison with MS/MS experiments.

The intercalation binding mode of these ligands was demonstrated previously [34,41,42]. Here two configurations were studied for each intercalation site: the alkyl chain of the drug can be oriented in two directions (Figure 3(a)). The central intercalation site is symmetric and only one orientation of the alkyl chain was studied for the CG*CG intercalation site (to prevent the aminoalkyl chain from going outside the duplex). The planar aromatic rings of the drugs were inserted into the duplex, with the alkyl chain laying in the minor groove allowing interaction between the side chain and the groove of the helix [34,41,42]. Small molecules generally interact via the DNA minor groove because it is narrower and more electronegative than the major groove, providing more favorable interactions with the ligand [58]. Only a small number of drugs like the bis-intercalator ditercalinium [59], or drugs that have large bulky groups [60], bind to duplex DNA through the major groove. Figure 3(b) shows the geometries of the generated complexes obtained between **2** and the duplex Dk66 after energy minimization. The values of the interaction energies obtained using Equation 2 are summarized in Table 1, together with the CE_{50} values obtained by MS/MS.

It is noteworthy that the interaction energies calculated for the intercalation site CG-CG, AT-AT and CG-AT follow roughly the same ranking order as in the MS/MS experiments. Only the E_{int} obtained for the intercalation site CG-AT with the aminoalkyl tail oriented toward the GC-side ("up") does not follow the experimental trend. The absence of a significant dependence of interaction energies on the base sequence is in agreement with the absence of sequence selectivity observed for the duplex in full scan ESI-MS: figure 1(b) shows that the amount of bound ligand does not change according with the GC content of the duplex. In the population of complexes selected for MS/MS, the probability of the ligand being in GC-rich sites is therefore supposed to be proportional to the fraction of these sites.

ESI-MS/MS and molecular modeling on the (1:1) triplex-ligand complexes

Figure 4(a) shows the ESI-MS/MS spectra obtained at different collision energies for the BPQ **3** (left) and the BgPI **2** (right). These two families of compounds show two different fragmentation patterns, summarized in Scheme 2. DNA complexes with quinoxaline derivatives dissociate principally via the loss of a neutral drug, while DNA-BPI complexes dissociate mainly via the loss of the antigene strand T3. Some ligand remains bound to the remaining duplex, but free duplex is detected as well. For ligands **3**-**7**, loss of neutral ligand is almost complete before antigene loss starts, but for ligands **1**-**2** antigene loss occurs at lower energies than loss of neutral ligand. There is therefore a competition between loss of neutral ligand and loss of antigene strand $[T3]^{3-}$. Loss of neutral ligand occurs at lower collision voltage for ligands **3**-**7**, while loss of antigene strand occurs at lower collision voltages for ligands **1**-**2**.

It must be pointed out that no drug is observed on the antigene strand, suggesting a stronger interaction with the Watson-Crick duplex part of the triplex. Figure 4(b) shows the dissociation curves

obtained for each drug. The percentage of intact complex is calculated using Equation (5), which is valid for both cases (intensities are replaced by zero when appropriate).

$$\%\text{Complex} = \frac{I_{(\text{Triplex}+\text{Drug}^{7-})}}{I_{(\text{Triplex}+\text{Drug}^{7-})} + I_{(\text{Triplex}^{7-})} + \frac{I_{(\text{T1}\bullet\text{T2}^{4-})} + I_{(\text{T1}\bullet\text{T2}+\text{Drug}^{4-})} + I_{(\text{T3}^{3-})}}{2}} \quad (5)$$

The two different fragmentation pathways do not necessarily imply that the binding modes are different. The breakdown curves of the duplex-ligand complexes also showed a large difference in dissociation threshold between ligands **1**-**2** and **3**-**7**. The different dissociation pathways with the triplex can simply be due to the fact that the ligand loss threshold is well above the antigene loss threshold for ligands **1**-**2**, and well below it for ligands **3**-**7**. However, the fact that the breakdown curves are not strictly parallel is a stronger experimental argument indicating that different ligands could have different binding modes. Non-parallel curves indicate that the activation entropy of dissociation is different, suggesting that the dissociation mechanisms are different. This could in turn be due to a different binding mode. Note for example the crossing of the breakdown curves of ligands **1** and **2** in Figure 4(b), left. Among the pyridoquinoxaline drugs **3**-**7**, ligand **6** needs the highest collision energy to be expelled, suggesting it is the best triplex binder of the series.

For the molecular modeling, two intercalation sites were generated using the same procedure as for duplex: CGC-TAT and TAT-TAT. Four starting intercalation geometries were generated for each site: two with the drug docked with the alkyl chain in the major groove with the aminoalkyl chain going “up” (ie: amino group of the alkyl chain toward the 3’ end of the antigen strand) and “down” (aminoalkyl chain toward the 5’ end of the antigen strand) and two by intercalating the drug with the alkyl chain in the minor groove, and testing the two orientations of the chain (Figure 5(a)). A total of eight binding modes were therefore tested for each drug. We checked that the optimized conformations of the triplex with ligand **4** and **6** were stable in vacuo by running lengthy dynamics simulation (total of

5 ns) at the temperature of 300 K (data not shown). Although the helical conformation is slightly distorted (especially the 3' end of the antigene strand), hydrogen bonding in the base triplets involved in the ligand binding site is well conserved, and ligand binding mode remains the same as shown in Figure 5(c), thereby validating our modelling approach.

The calculated interaction energies for the different binding sites are summarized in Table 2 with the corresponding experimental voltage to observe 50 % dissociation (CE_{50}) obtained by tandem mass spectrometry. The calculated interaction energies are the most favorable for the TAT-TAT binding site with the aromatic rings of the drugs inserted through the minor groove of the Watson-Crick duplex and the aminoalkyl chain in the minor groove. For that binding site, the interaction energies follow the same ranking as the experimentally determined CE_{50} (Table 2). Figure 5(b) shows the structure of the complex with ligand **2** in that binding mode, and Figure 5(c) shows the superimposition of the ligand and the base triplet for ligands **4**-**7**. The only exception is ligand **1**, for which a significantly lowest energy site corresponds to the aminoalkyl chain in the major groove. In previous structure-affinity relationships and molecular modeling for BePI and BgPI derivatives, it was already suggested that the aminoalkyl chain was located in the minor groove of the duplex for BgPIs like **2** and in the major groove together with the antigene strand for BePIs like **1** [34]. Our calculations are also in line with this model and, coming back the MS/MS results, the fact that the breakdown curves of ligands **1** and **2** are not parallel can now be interpreted as an influence of the ligand binding mode on the activation entropy of dissociation. This illustrates that (1) a careful interpretation of the breakdown curves allows guessing that two of the ligands might have different binding modes, and (2) molecular modeling calculations, even at the modest level used here, could allow identifying the correct binding mode.

CONCLUSION

The major conclusions of the present work are as follows. The benzopyridoindoles (**1** and **2**) have larger interaction energies than the pyridoquinoxalines (**3**-**7**). The nature of the heteroaromatic rings therefore plays a major role in the affinity of the ligand for the target. However, the benzopyridoindoles do not show any selectivity for the triplex compared to the duplex (see Figure 1). Molecular modeling indeed shows that BPI drugs interact preferentially with the Watson-Crick duplex and have small interaction with the antigen strand. Among the benzopyridoquinoxalines, the three most selective ligands (**4**, **5** and **6**) all share the same heteroaromatic core, which best stacks on a base triplet (see Figure 5). Ligand **4** has stronger interaction with the base triplet than isomers **7** and **5**. Our results consistently suggest that ligand **6** is the most promising lead for triplex-selective ligands. It has (i) the best triplex specificity, as shown by the ESI-MS screening (Figure 1), (ii) the highest interaction energy of all BPQs with the triplex, as shown by ESI-MS/MS (Figure 4 and Table 2), and (iii) the most favorable stacking configuration on a base triplet, as shown by molecular modeling (Figure 5(c)). This study demonstrates for the first time the use of ESI-MS and ESI-MS/MS to study drug-triplex interactions, and underlines the utility of molecular modeling for the interpretation of tandem mass spectrometry data in structural terms.

ACKNOWLEDGMENTS

This work was supported by the FRS-FNRS (Fonds de la Recherche Scientifique-FNRS) via a Research Associate position to VG and a Post-doctoral fellowship to FR. The authors would like to thank Dr. Christian Bailly, Dr. Jean-Louis Mergny, and Dr. Patrizia Alberti for helpful discussion.

TABLE 1: Interaction energies calculated for each drug and intercalation site in the duplex $(CGCGAATTCGCG)_2$.

Drug	CE_{50} (V)	E_{int} (kcal/mol)			
		CG-CG site	**AT-AT site**	**CG-AT site Tail down**	**CG-AT site Tail up**
1	16.6	**-55.4**	**-57.2**	-50.3	-42.7
2	16.2	**-55.5**	-51.2	**-55.9**	-43.0
6	11.1	**-51.1**	**-51.2**	-49.3	-43.2
3	10.4	**-53.5**	**-49.3**	-42.7	-46.8
4	9.6	**-47.1**	**-46.3**	-45.9	-41.0
5	9.3	**-47.5**	**-45.7**	-45.6	-43.4
7	9.2	**-42.2**	**-41.9**	-37.6	-40.8

TABLE 2: Interaction energies calculated for each drug and intercalation site in the Triplex (CTTTTCC*GAAAAGG-CTTTTCC).

Drug	CE_{50} (V)	E_{int} (kcal/mol)							
		CGC-TAT Major groove Tail down	CGC-TAT Major groove Tail up	CGC-TAT Minor groove Tail down	CGC-TAT Minor groove Tail up	TAT-TAT Major groove Tail down	TAT-TAT Major groove Tail up	TAT-TAT Minor groove Tail down	TAT-TAT Minor groove Tail up
1	16.3	-33.5	**-68.0**	-53.4	-47.3	-46.1	-41.0	-66.4	-55.5
2	17.0	-39.2	-44.0	-49.3	-44.7	-39.2	-48.2	**-64.8**	-62.9
6	12.0	-45.6	-55.0	-59.5	-49.6	-46.5	-38.0	**-64.4**	-54.4
3	10.5	-43.2	-52.1	-51.7	-48.2	-51.3	-38.8	**-58.1**	-47.8
4	10.0	-43.5	-50.6	-50.2	**-56.0**	-47.5	-46.3	**-55.5**	-45.8
5	9.3	-52.2	-46.8	**-50.5**	-47.2	-41.9	-46.5	**-49.4**	-48.7
7	9.2	-33.7	-43.3	**-53.8**	-47.9	-38.6	-51.5	**-51.9**	-43.8

Energy minimization to 0.05 kcal/(mol.A) using Polak-Ribiere gradient conjugate optimization algorithm.

REFERENCES

1. Hurley, L. H. DNA and its associated processes as targets for cancer therapy. *Nature Rev. Cancer* **2002**, *2*, 188-200.

2. Hélène, C. The anti-gene strategy: control of gene expression by triplex-forming-oligonucleotides. *Anticancer Drug Des* **1991**, *6*, 569-584.

3. Plum, G. E.; Pilch D. S.; Singleton S. F.; Breslauer K. J. Nucleic acid hybridization: triplex stability and energetics. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 319-350.

4. Jenkins, T. C. Targeting multi-stranded DNA structures. *Curr. Med. Chem.* **2000**, *7*, 99-115.

5. Opalinska, J. B.; Gewirtz A. M. Nucleic-acid therapeutics: basic principles and recent applications. *Nature Rev. Drug Discov.* **2002**, *1*, 503-514.

6. Besch, R.; Giovannangeli C.; Degitz K. Triplex-forming oligonucleotides - Sequence-specific DNA ligands as tools for gene inhibition and for modulation of DNA-associated functions. *Current Drug Targets* **2004**, *5*, 691-703.

7. Felsenfeld, G.; Davis D. R.; Rich A. Formation of a three-stranded polynucleotide molecule. *J. Am. Chem. Soc.* **1957**, *79*, 2023-2024.

8. Marchand, C.; Bailly C.; Nguyen C. H.; Bisagni E.; Garestier T.; Helene C.; Waring M. J. Stabilization of triple helical DNA by a benzopynidoquinoxaline intercalator. *Biochemistry* **1996**, *35*, 5022-5032.

9. Zain, R.; Marchand C.; Sun J. S.; Nguyen C. H.; Bisagni E.; Garestier T.; Helene C. Design of a triple-helix-specific cleaving reagent. *Chem. Biol.* **1999**, *6*, 771-777.

10. Eftink, M. R. Fluorescence methods for studying equilibrium macromolecule-ligand interactions. *Methods Enzymol.* **1997**, *278*, 221-257.

11. Chaires, J. B. Analysis and interpretation of ligand-DNA binding isotherms. *Methods Enzymol.* **2001**, *340*, 3-23.

12. Chaires, J. B. Structural selectivity of drug-nucleic acid interactions probed by competition dialysis. *Dna Binders and Related Subjects* **2005**, *253*, 33-53.

13. Chaires, J. B. Competition dialysis: an assay to measure the structural selectvity of drug-nucleic acid interactions. *Curr. Pharm. Des.* **2005**, *11*, 0-00.

14. Piehler, J.; Brecht A.; Gauglitz G.; Zerlin M.; Maul C.; Thiericke R.; Grabley S. Label-free monitoring of DNA-ligand interactions. *Anal. Biochem.* **1997**, *249*, 94-102.

15. Myszka D. G.; Jonsen M. D.; Graves B. J. Equilibrium analysis of high affinity interactions using BIACORE. 1998 In , 326-30.

16. Freire, E.; Mayorge O. L.; Straume M. Isothermal titration calorimetry. *Anal. Chem.* **1990**, 950 A-959 A.

17. Breslauer, K. J.; Freire E.; Straume M. Calorimetry: a tool for DNA and ligand-DNA studies. *Methods Enzymol.* **1992**, *211*, 533-567.

18. Gale, D. C.; Goodlett D. R.; Light-Wahl K. J.; Smith R. D. Observation of duplex DNA-drug non-covalent complexes by electrospray ionization mass spectrometry. *J. Am. Chem. Soc.* **1994**, *116*, 6027-6028.

19. Gale, D. C.; Smith R. D. Characterization of non-covalent complexes formed between minor groove binding molecules and duplex DNA by electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1154-1164.

20. Pocsfalvi, G.; Di Landa G.; Ferranti P.; Ritieni A.; Randazzo G.; Marorni A. Observation of non-covalent interactions between bauvericin and oligonucleotides usin electrospray ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 265-272.

21. Triolo, A.; Arcamone F. M.; Raffaelli A.; Salvadori P. Non-covalent complexes between DNA-binding drugs and double-stranded deoxyoligonucleotides : a study by ionspray mass spectrometry. *J. Mass Spectrom.* **1997**, *32*, 1186-1194.

22. Gabelica, V.; De Pauw E.; Rosu F. Interaction between Antitumor Drugs and Double-stranded DNA Studied by Electrospray Ionization Mass Spectrometry. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.

23. Kapur, A.; Beck J. L.; Sheil M. M. Observation of daunomycin and nogalamycin complexes with duplex DNA using electrospray ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2489-2497.

24. Wan, K. X.; Shibue T.; Gross M. L. Non-covalent complexes between DNA-binding drugs and double-stranded oligodeoxynucleotides: a study by electrospray ionization mass spectrometry. *J. Am. Chem. Soc.* **2000**, *122*, 300-307.

25. Reyzer, M.; Brodbelt J. S.; Kerwin S. M.; Kumar D. Evaluation of complexation of metal-mediated DNA-binding drugs to oligonucleotides via electrospray ionization mass spectrometry. *Nucleic Acids Res.* **2001**, *29*, e103.

26. Beck, J.; Colgrave M. L.; Ralph S. F.; Sheil M. M. Electrospray ionization mass spectrometry of oligonucleotide complexes with drugs, metals, and proteins. *Mass Spectrom. Rev.* **2001**, *20*, 61-87.

27. Rosu, F.; Gabelica V.; Houssier C.; De Pauw E. Determination of affinity, stoichiometry and sequence selectivity of minor groove binder complexes with double-stranded oligodeoxynucleotides by electrospray ionization mass spectrometry. *Nucleic Acids Res.* **2002**, *30*, e82.

28. Guittat, L.; De Cian A.; Rosu F.; Gabelica V.; De Pauw E.; Delfourne E.; Mergny J. L. Ascididemin and meridine stabilise G-quadruplexes and inhibit telomerase in vitro. *Biochim. Biophys. Acta* **2005**, *1724*, 375-384.

29. Guittat, L.; Alberti P.; Rosu F.; Van Miert S.; Thetiot E.; Pieters L.; Gabelica V.; De Pauw E.; Ottaviani A.; Riou J.-F.; Mergny J.-L. Interactions of cryptolepine and neocryptolepine with unusual DNA structures. *Biochimie* **2003**, *85*, 535-547.

30. Rosu, F.; De Pauw E.; Guittat L.; Alberti P.; Lacroix L.; Mailliet P.; Riou J.-F.; Mergny J.-L. Selective interaction of ethidium derivatives with quadruplexes. An equilibrium dialysis and electrospray ionization mass spectrometry analysis. *Biochemistry* **2003**, *42*, 10361-10371.

31. Carrasco, C.; Rosu F.; Gabelica V.; Houssier C.; De Pauw E.; Garbay-Jaureguiberry C.; Roques B.; Wilson W. D.; Chaires J. B.; Waring M. J.; Bailly C. Tight binding of the antitumor drug ditercalinium to quadruplex dna. *Chem. Bio. Chem.* **2002**, *3*, 100-106.

32. Mazzitelli, C. L.; Brodbelt J. S.; Kern J. T.; Rodriguez M.; Kerwin S. M. Evaluation of binding of perylene diimide and benzannulated perylene diimide ligands to DNA by electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 593-604.

33. Nguyen, C. H.; Lhoste J. M.; Lavelle F.; Bissery M. C.; Bisagni E. Synthesis and Antitumor-Activity of 1-[[(Dialkylamino)Alkyl]Amino]-4-Methyl-5H-Pyrido[4,3-B]Benzo[E])Indole and 1-[[Dialkylamino)Alkyl]Amino]-4-Methyl-5H-Pyrido[4,3-B]-Benzo[G])Indole - A New Class of Antineoplastic Agents. *J. Med. Chem.* **1990**, *33*, 1519-1528.

34. Escudé, C.; Nguyen C. H.; Mergny J.-L.; Sun J.-S.; Bisagni E.; Garestier T.; Hélène C. Selective stabilization of DNA triple helices by benzopyridoindole derivatives. *J. Am. Chem. Soc.* **1995**, *117*, 10212-10219.

35. Pilch, D. S.; Waring M. J.; Sun J.-S.; Rougée M.; Nguyen C. H.; Bisagni E.; Garestier T.; Hélène C. Characterization of a triple helix-specific ligand BePI [3-metoxy-7h-8-methyl-11-[(3'-amino)propylamino]-benzo[e]pyrido[4,3-b]indole] intercalates into both double-helical and triple-helical DNA. *J. Mol. Biol.* **1993**, *232*, 926-946.

36. Bailly, C.; Marchand C.; Nguyen C. H.; Bisagni E.; Garestier T.; Helene C.; Waring M. J. Localized Chemical-Reactivity in Double-Stranded Dna Associated with the Intercalative Binding of Benzo[E]Pyridoindole and Benzo[G]Pyridoindole Triple-Helix-Stabilizing Ligands. *European Journal of Biochemistry* **1995**, *232*, 66-76.

37. Mergny, J.-L.; Duval-Valentin G.; Nguyen C. H.; Perrouault L.; Faucon B.; Rougée M.; Montenay-Garestier T.; Bisagni E.; Hélène C. Triple helix-specific ligands. *Science* **1992**, *256*, 1681-1684.

38. Riou, J. F.; Fosse P.; Nguyen C. H.; Larsen A. K.; Bissery M. C.; Grondard L.; Saucier J. M.; Bisagni E.; Lavelle F. Intoplicine (Rp-60475) and Its Derivatives, A New Class of Antitumor Agents Inhibiting Both Topoisomerase-I and Topoisomerase-Ii Activities. *Cancer Research* **1993**, *53*, 5987-5993.

39. Nguyen, C. H.; Fan E.; Riou J.-F.; Bissery M.-C.; Vrignaud P.; Lavelle F.; Bisagni E. Synthesis and biological evaluation of amino-substituted benzo[f]pyrido[4,3-b] and pyrido[3,4-b]quinoxalines: a new class of antineoplastic agents. *Anticancer Drug Des* **1995**, *10*, 277-297.

40. Marchand, C.; Bailly C.; Nguyen C. H.; Bisagni E.; Garestier T.; Hélène C.; Waring M. J. Stabilization of triple helical DNA by a benzopyridoquinoxaline intercalator. *Biochemistry* **1996**, *35*, 5022-5032.

41. Pilch, D. S.; Martin M.-T.; Nguyen C. H.; Sun J.-S.; Bisagni E.; Garestier T.; Hélène C. Self-association and DNA-binding properties of two triple helix-specific ligands: comparison of a benzo[e]pyridoindole ans a benzo[g]pyridoindole. *J. Am. Chem. Soc.* **1993**, *115*, 9942-9951.

42. Nabiev, I.; Chourpa I.; Riou J. F.; Nguyen C. H.; Lavelle F.; Manfait M. Molecular-Interactions of Dna Topoisomerase-I and Topoisomerase-Ii Inhibitor with Dna and Topoisomerases and in Ternary Complexes - Binding Modes and Biological Effects for Intoplicine Derivatives. *Biochemistry* **1994**, *33*, 9013-9023.

43. Escudé, C.; Nguyen C. H.; Kukreti S.; Janin Y.; Sun J.-S.; Bisagni E.; Garestier T.; Hélène C. Rational design of a triple helix-specific intercalating ligand. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 3591-3596.

44. Vékey, K. Internal energy effects in mass spectrometry. *J. Mass Spectrom.* **1996**, *31*, 445-463.

45. Hase, W. L. Some recent advances and remaining questions regarding unimolecular rate theory. *Acc. Chem. Res.* **1998**, *31*, 659-665.

46. Lifshitz, C. Some recent aspects of unimolecular gas phase ion chemistry. *Chem. Soc. Rev.* **2001**, *30*, 186-192.

47. Rosu, F.; Gabelica V.; Houssier C.; Colson P.; De Pauw E. Triplex and quadruplex DNA structures studied by electrospray mass spectrometry. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1729-1736.

48. Gabelica, V.; De Pauw E. Comparison of the collision-induced dissociation of duplex DNA at different collision regimes: evidence for a multistep dissociation mechanism. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 91-98.

49. Kuszewski, J.; Schwieters C.; Clore G. M. Improving the Accuracy of NMR Structures of DNA by Means of a Database Potential of Mean Force Describing Base-Base Positional Interactions. *J. Am. Chem. Soc.* **2001**, *123*, 3903-3918.

50. Cirilli, M.; Bachechi F.; Ughetto G.; Colonna F. P.; Capobianco M. L. Interactions Between Morpholinyl Anthracyclines and Dna - the Crystal-Structure of A Morpholino Doxorubicin Bound to D(Cgtacg). *Journal of Molecular Biology* **1993**, *230*, 878-889.

51. Frederick, C. A.; Williams L. D.; Ughetto G.; Van der Marel G. A.; Van Boom J. H.; Rich A.; Wang A. H. Structural comparison of anticancer drug-DNA complexes: adriamycin and daunomycin. *Biochemistry* **1990**, *29*, 2538-2549.

52. Radhakrishnan, I.; Patel D. J. Solution Structure of A Pyrimidine-Center-Dot-Purine-Center-Dot-Pyrimidine Dna Triplex Containing T-Center-Dot-At, C+Center-Dot-Gc and G-Center-Dot-Ta Triples. *Structure* **1994**, *2*, 17-32.

53. Gabelica, V.; Galic N.; Rosu F.; Houssier C.; De Pauw E. Influence of response factors on determining equilibrium association constants of non-covalent complexes by electrospray ionization mass spectrometry. *J. Mass Spectrom.* **2003**, *38*, 491-501.

54. Rosu, F.; Pirotte S.; De Pauw E.; Gabelica V. Positive and negative ion mode ESI-MS and MS/MS for studying drug–DNA complexes. *Int. J. Mass Spectrom.* **2006**, *253*, 156-171.

55. Bostock-Smith, C. E.; Harris S. A.; Laughton C. A.; Searle M. S. Induced fit DNA recognition by a minor groove binding analogue of Hoechst 33258:: fluctuations in DNA A tract structure investigated by NMR and molecular dynamics simulations. *Nucleic Acids Res.* **2001**, *29*, 693-702.

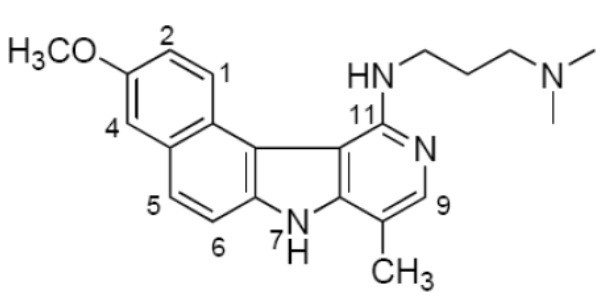
56. Chaires, J. B.; Satyanaraana S.; Suh D.; Fokt I.; Przewloka T.; Priebe W. Parsing the free energy of antracycline antibiotic binding to DNA. *Biochemistry* **1996**, *35*, 2047-2053.

57. Harris, S. A.; Gavathiotis E.; Searle M. S.; Orozco M.; Laughton C. A. Cooperativity in Drug-DNA recognition: a molecular dynamics study. *J. Am. Chem. Soc.* **2001**, *123*, 12658-12663.

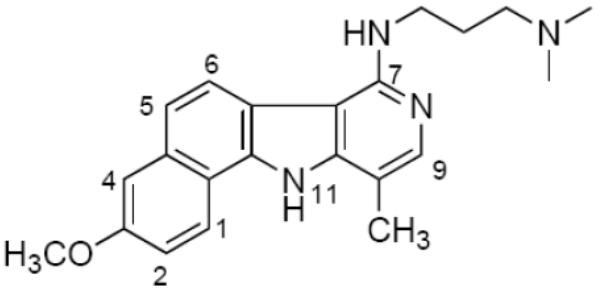
58. Hannon, M. J. Supramolecular DNA recognition. *Chem. Soc. Rev.* **2007**, *36*, 280-295.

59. Gao, Q.; Williams L. D.; Egli M.; Rabinovitch D.; Chen S.-L.; Quigley G. J.; Rich A. Drug-induced DNA repair: X-ray structure of a DNA-ditercalinium complex. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2422-2426.

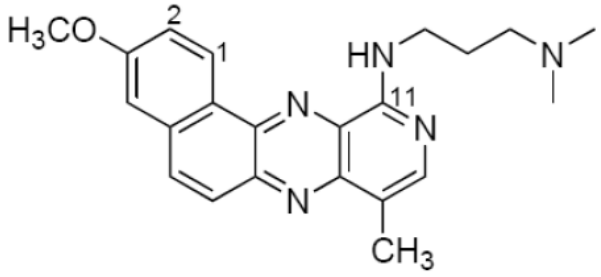
60. Jin, E.; Katritch V.; Olson W. K.; Kharatisvili M.; Abagyan R.; Plich D. S. Aminoglycoside binding in the major groove of duplex RNA: the thermodynamic and electrostatic forces that govern recognition. *J. Mol. Biol.* **2000**, 95-110.



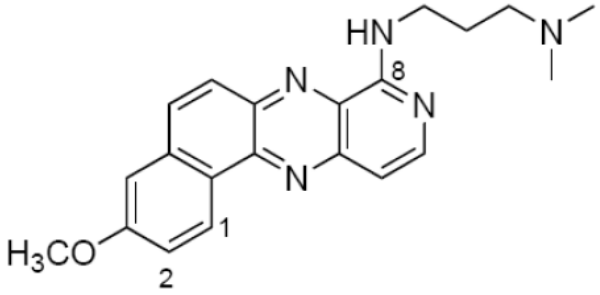
1



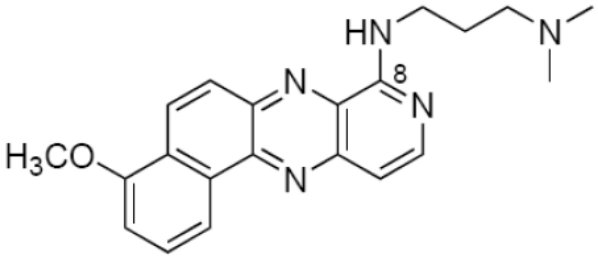
2



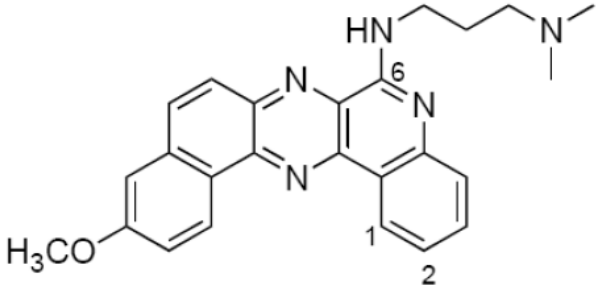
3



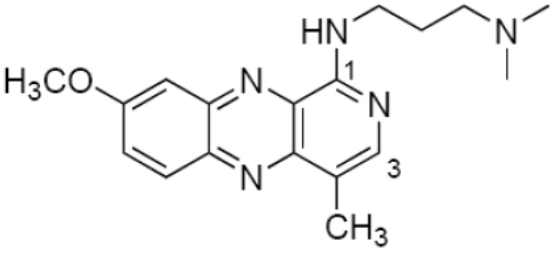
4



5

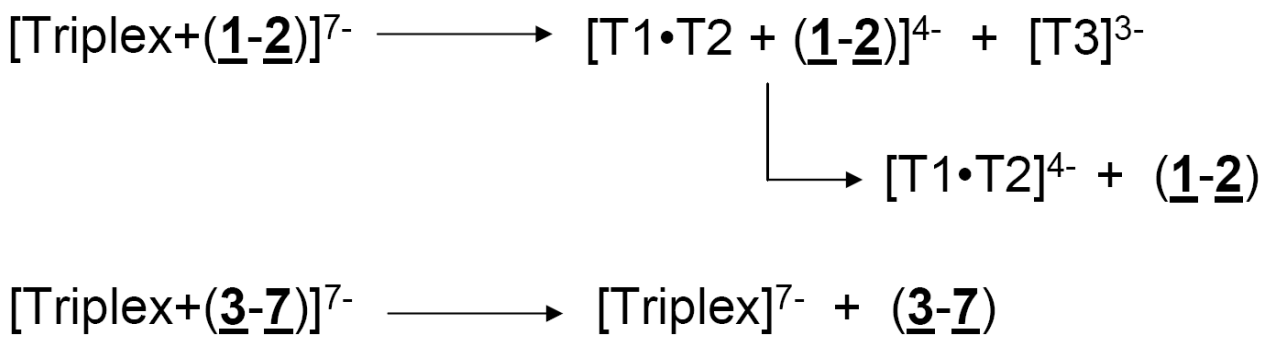


6



7

Scheme 1.



Scheme 2.

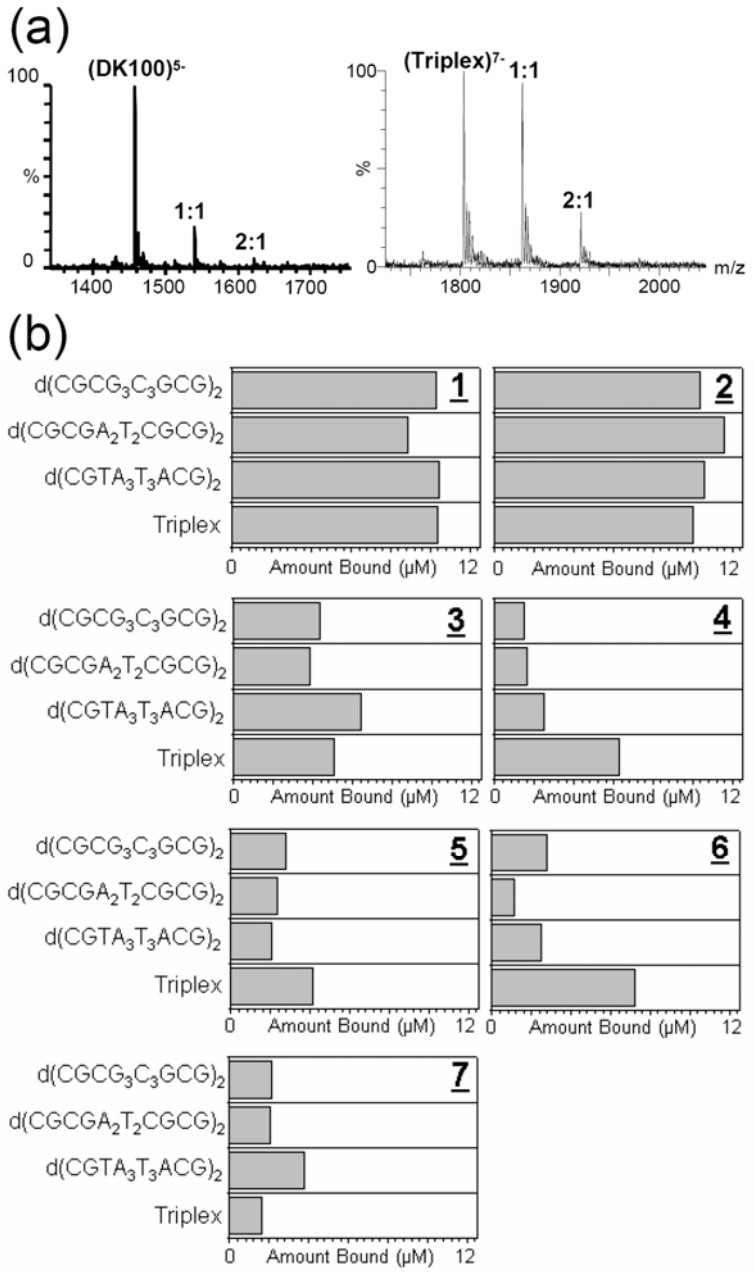


Figure 1. ESI-MS screening of ligand-duplex and ligand-triplex binding. (a) Representative ESI-MS spectra with a duplex and triplex DNA: ESI-MS spectra of solution containing 10 μM DNA (duplex DK100 on the left and triplex on the right, see materials and methods for the sequences) and 15 μM of ligand **6**. (b) Relative ligand affinities for the different DNA structures: concentration of ligand bound (in μM bound out of 15 μM total ligand added to 10 μM DNA) to three duplexes and the triplex.

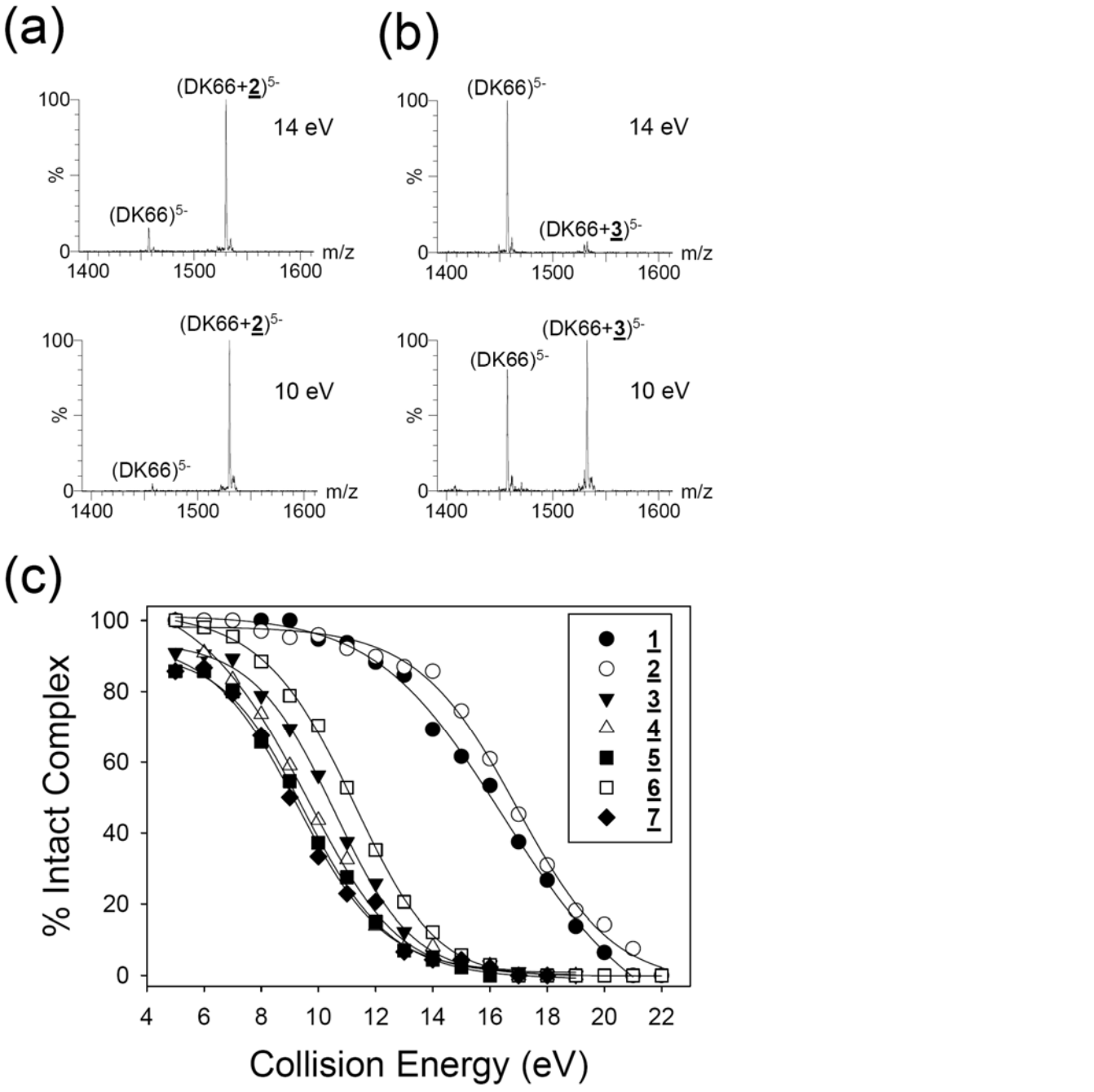


Figure 2: ESI-MS/MS experiments on $[1:1]^{5-}$ complexes with the duplex DK66. (a) Representative MS/MS spectra obtained with ligands **2** (left) and **3** (right), using collision energies of 10 and 14 eV. The argon pressure inside the hexapole collision cell was 3.0 10^{-5} mbar. (b) MS/MS breakdown curves of different intercalators drugs with the $d(CGCGAATTCGCG)_2$ (Dk66) duplex. The % of intact complex was calculated using equation (4) for each collision energy.

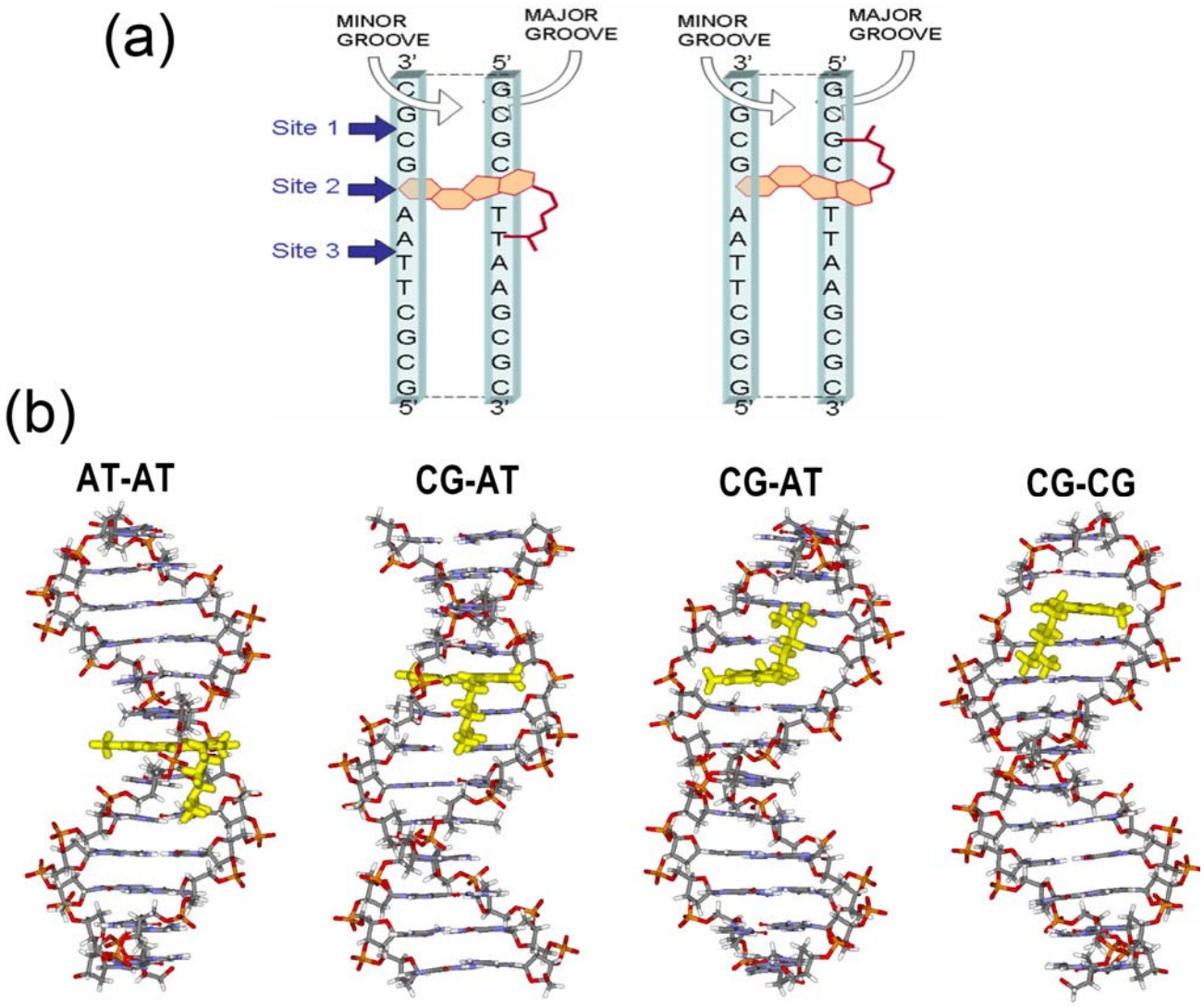


Figure 3: Modeling of duplex-ligand binding sites. (a) Localization of the three intercalation sites used in the molecular modeling. A drug is represented in site 2, with two possible orientations (alkyl chain toward 5' (down) or 3' end (up) in respect to the first stand of the duplex). (b) Modeling of the complex between **2** and the duplex Dk66 in the intercalation site AT-AT, CG-AT with the aminoalkyl chain "down", CG-AT with the chain "up" and CG-CG. The drug is shown in yellow.

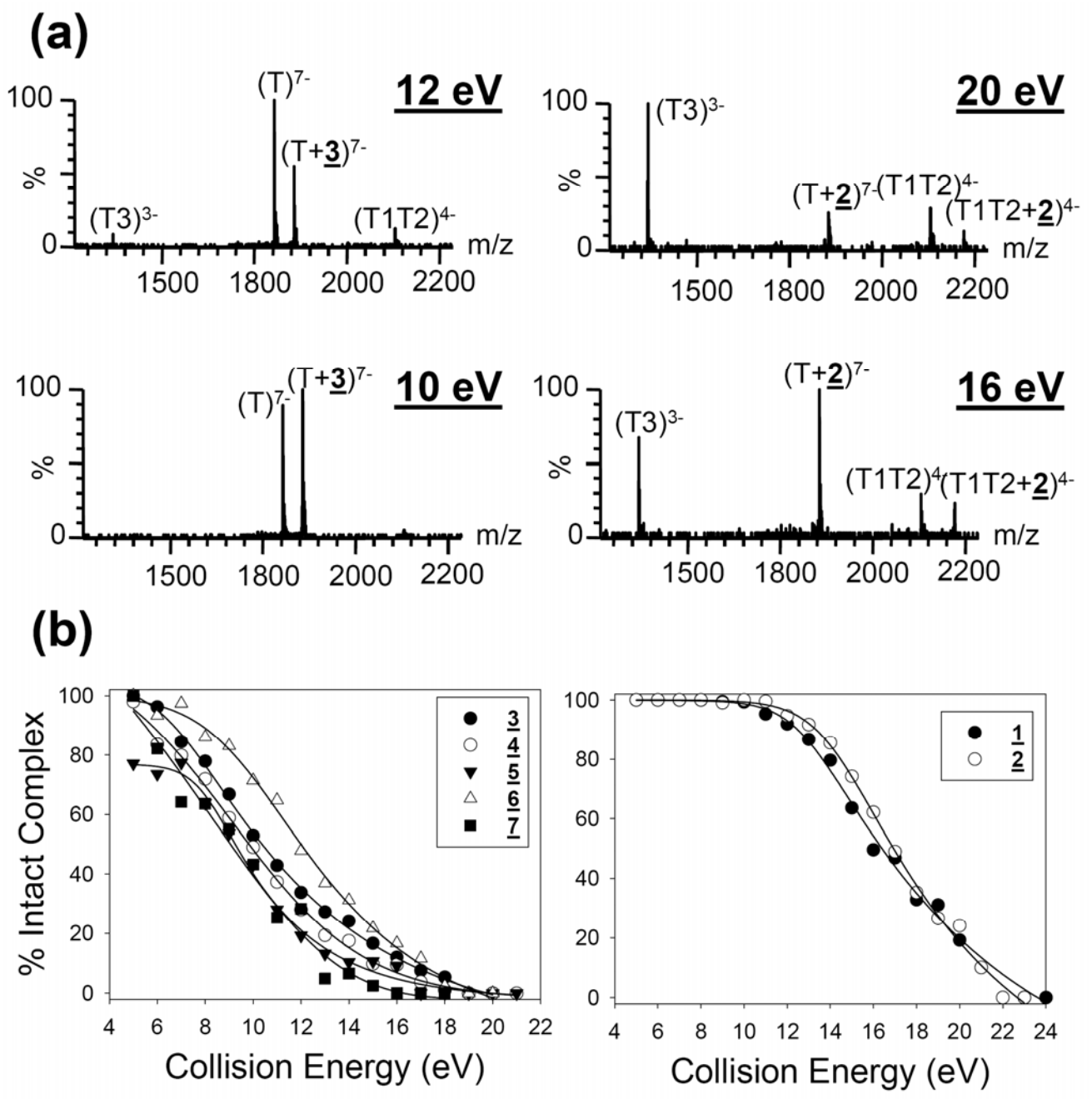


Figure 4: ESI-MS/MS experiments on $[1:1]^{5-}$ complexes with the triplex. (a) Representative spectra obtained for quinoxaline drivatives (left, shown for ligand **3** at 10 and 12 eV collision energy) and BPIs (right, shown for ligand **2** at 16 and 20 eV collision energy). The argon pressure inside the hexapole collision cell was 3.0×10^{-5} mbar. (b) MS/MS breakdown curves of the different drugs with the triplex. The % of intact complex is calculated using Equation (5) for each collision energy.

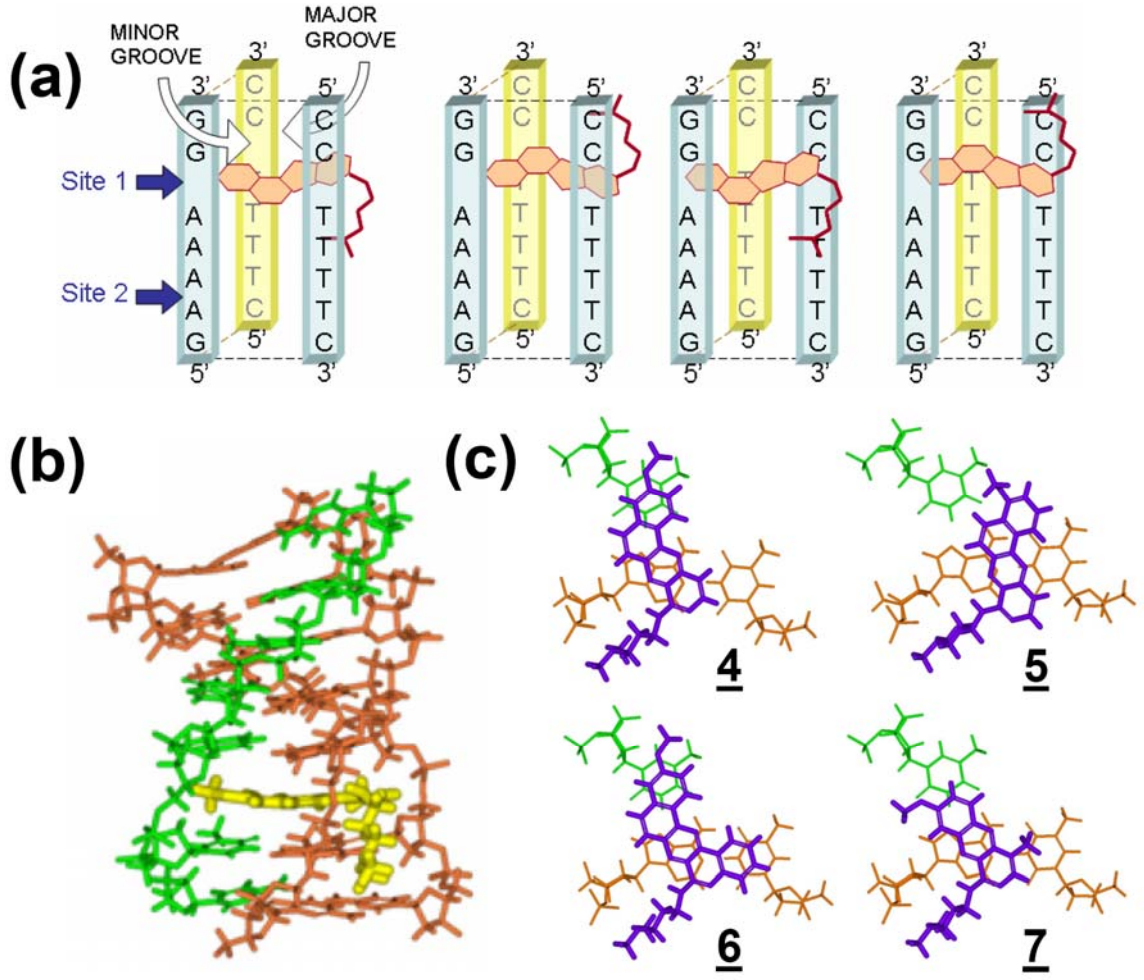


Figure 5: Modeling of duplex-ligand binding sites. (a) Localization of the two intercalation sites used in the molecular modeling. A drug is represented for site 1. Two orientations of the drug are possible (alkyl chain toward 5' or 3' end in respect to the G-rich strand). The drug was docked either by the minor and the major groove of the helix. (b) Energy minimized structure of ligand **2** bound to the triplex (CTTTTCC*GAAAAGG-CTTTTCC), in site 2 (TAT-TAT) with the aminoalkyl chain "down". The drug is colored in yellow, the Watson-Crick duplex in brown, and the antigene strand in green. (c) Zoom of the Energy-minimized models of **4**, **5**, **6**, **7** intercalated in the TAT-TAT base triplets with the aminoalkyl chain "down". The planar ring system of the ligands is shown stacked with the base triplet . Only the bottom triplet is shown for clarity. The drug is colored in mauve, the Watson-Crick duplex in brown and the antigen strand in green.