5000–5012 *Nucleic Acids Research, 2008, Vol. 36, No. 15* *Published online 25 July 2008*
doi:10.1093/nar/gkn392

Selective recognition of pyrimidine–pyrimidine DNA mismatches by distance-constrained macrocyclic bis-intercalators

Matthias Bahr[1], Valérie Gabelica[2], Anton Granzhan[3], Marie-Paule Teulade-Fichou[3,*] and Elmar Weinhold[1]

[1]Institute of Organic Chemistry, RWTH Aachen University, Landoltweg 1, D-52056 Aachen, Germany, [2]Department of Chemistry, Mass Spectrometry Laboratory, University of Liège, B-4000 Liège, Belgium and [3]Institut Curie, UMR176 CNRS, Centre Universitaire, F-91405 Orsay, France

Received April 28, 2008; Revised May 30, 2008; Accepted June 4, 2008

ABSTRACT

Binding of three macrocyclic bis-intercalators, derivatives of acridine and naphthalene, and two acyclic model compounds to mismatch-containing and matched duplex oligodeoxynucleotides was analyzed by thermal denaturation experiments, electrospray ionization mass spectrometry studies (ESI-MS) and fluorescent intercalator displacement (FID) titrations. The macrocyclic bis-intercalators bind to duplexes containing mismatched thymine bases with high selectivity over the fully matched ones, whereas the acyclic model compounds are much less selective and strongly bind to the matched DNA. Moreover, the results from thermal denaturation experiments are in very good agreement with the binding affinities obtained by ESI-MS and FID measurements. The FID results also demonstrate that the macrocyclic naphthalene derivative BisNP preferentially binds to pyrimidine–pyrimidine mismatches compared to all other possible base mismatches. This ligand also efficiently competes with a DNA enzyme (M.TaqI) for binding to a duplex with a TT-mismatch, as shown by competitive fluorescence titrations. Altogether, our results demonstrate that macrocyclic distance-constrained bis-intercalators are efficient and selective mismatch-binding ligands that can interfere with mismatch-binding enzymes.

INTRODUCTION

Mismatched base pairs in DNA can arise by several processes. One of their most important sources are the replication errors, i.e. direct misincorporation of bases due to statistical errors during DNA replication, or due to damages, or lesions, in the parental strand, which lead to incorrect recognition and thus incorporation of wrong bases in the newly replicated strand (1–3). Another mechanism is based on the formation of a heteroduplex between two homologous DNA molecules during the recombination process: if the two DNA strands differ slightly in their sequence, mismatched base pairs may be formed (4). Mismatches may also be generated in hairpins that are formed between imperfect palindromes existing in repeat-containing sequences, such as microsatellites and trinucleotide repeats (5,6). An additional special but important path for the formation of mismatched base pairs is deamination of 5-methylcytosine, a modified base that is present in the DNA of many organisms. In this case, the deamination product is thymine and a GT mismatch is generated (7).

Base mispairs can be hazardous to the cell in altering its ability to transfer the information content of DNA. In particular, mismatches may result in point mutations that are potentially harmful, depending on where they occur in the genome. Consequently, every organism has evolved a variety of control and repair strategies based on complex enzymatic machineries responsible for the maintenance of DNA integrity. In particular, mismatches are recognized and repaired by specific enzymes, which constitute the mismatch repair (MMR) system. The MMR system is based on a complex network of interactions between various enzymes and is closely related to the other systems recognizing and signaling lesions (base-excision repair, nucleotide-excision repair, methyl-directed repair, etc.) (8).

In addition, it is well established that most human tumors develop through a succession of genetic and epigenetic changes and that high frequency of mutations is

*To whom correspondence should be addressed. Tel: +33 169 86 30 86; Fax: +33 169 07 53 81; Email: marie-paule.teulade-fichou@curie.fr
Correspondence may also be addressed to Elmar Weinhold. Email: elmar.weinhold@oc.rwth-aachen.de

clearly correlated to cancerogenesis. For instance, mutations of oncogenes or tumor suppressor genes, such as p53 and BRCA (9), and sequence instability of microsatellites (10,11) are directly related to the occurrence of inherited cancers. Finally, the importance of DNA repair in the prevention of carcinogenesis has been recently highlighted by finding a direct correlation between the defective DNA MMR and hereditary colon cancer (12).

In terms of molecular recognition, repair of DNA mismatches requires that the correct base in the mispair is distinguished from the incorrect one. Since both bases in a mispair are regular constituents of DNA, this recognition is much more challenging compared with the recognition of damaged bases, which are structurally altered compared with the common DNA bases. For instance, the specific recognition of GT mispairs by the MutS enzyme is fascinating, because they are thermodynamically only slightly less stable than the Watson–Crick base pairs (13) and induce only a slight structural modification of duplex DNA (14–18). Therefore, studies aimed at a deeper understanding of the recognition of mismatches by repair enzymes have raised continued attention for more than a decade. Several models have been proposed to rationalize the mechanisms of mismatch recognition, but these are still poorly understood (14). Given the complexity of these processes the task is highly challenging and requires several approaches, such as genetic, biochemical and chemical ones. In particular, a chemical tool for studying mismatch recognition is represented by small molecules that, similar to the mismatch-recognizing enzymes, can bind base mispairs with a high selectivity over fully paired DNA. Such mismatch-binding ligands (*mismatch binders*) may eventually interfere with the repair systems with negative or positive consequences, leading to inhibition (19) or promotion of repair, and thus display high therapeutic potential.

DNA mismatches constitute sites of reduced thermodynamic stability compared with Watson–Crick base pairs (20–22), which leads to enhanced conformational dynamics (‘*breathing*’) and may facilitate the insertion of classical intercalators and various drugs into these sites (23–25). The design of mismatch binders usually takes advantage of this local conformational suppleness, although it is only modest, restricted to two to three adjacent base pairs and depends on the type of mismatch. Importantly, beyond the specific mismatch recognition motifs efficient mismatch binders should feature unfavorable interactions with Watson–Crick base pairs as well as with grooves, to ensure low binding to the fully paired DNA and high selectivity for the mismatches. Thus, in the past decade several series of mismatch-recognizing agents have emerged. Among these are molecular systems that operate via intercalation, such as rhodium-based metalloinsertors, which preferentially bind to CC and CA mismatches (19,26,27), or via bis-intercalation, such as bis-naphthyridine derivatives that selectively bind to GG and GA mismatches (28–31). Minor groove binders such as imidazole-rich polyamides have also been shown to selectively bind to the GT mismatched sites, recognizing the geometric modifications of the grooves induced by this mispairing (32–34). In another approach, we have shown that a macrocyclic bis-acridine compound (**BisA**, Chart 1) recognizes base-pairing defects, like abasic sites (35,36) and thymine-containing mismatches, such as TT, TC and, to a lesser extent, TG-mismatched base pairs (37), via a putative threading bis-intercalation mode. Moreover, insertion of **BisA** at TX-mismatch sites induces a displacement of the thymine into an extrahelical position, as shown by the susceptibility of the flipped-out thymine to oxidation by $KMnO_4$. This established for the first time that base flipping, which is characteristic of DNA methyltransferases and DNA glycosylases (38,39), can also be performed by a small molecule (37). Subsequently, several studies have shown that other mismatch binders, such as naphthyridine dimers (40,41) and bulky rhodium-containing metalloinsertors

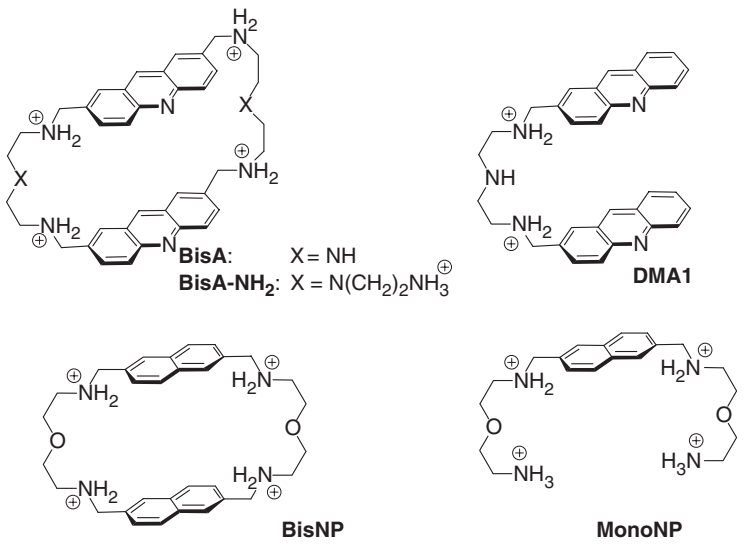

Chart 1. Structures of macrocyclic ligands and acyclic control compounds used in this study. The charges correspond to the presumable protonation sites at pH 6.0. All ligands were handled as hydrochloride salts.

(27,42) are also able to displace bases into extrahelical positions.

Beyond their potential usefulness for interfering with the MMR system, the mismatch binders may also serve in genetic diagnostics for the detection of point mutations and single-nucleotide polymorphism typing (43,44). Towards this objective, several mismatch-binding ligands have been endowed with cross-linking (45), photoactive (46,47) or fluorescent groups (48). Such molecular probes could complement the number of chemical, physical and enzymatic methods (49–51) developed for mismatch detection since most exhibit drawbacks and/or questionable reliability (52).

In order to get deeper insight into the interaction of macrocyclic compounds, such as **BisA**, with mismatches, we carried out a systematic study aimed at the determination of structural factors that determine the binding, as well as the stoichiometric and thermodynamic parameters of the binding event. To achieve this goal, we extended the macrocyclic series by several analogues of **BisA** and studied their mismatch-binding properties by a number of biochemical and spectroscopic methods. First, we prepared a derivative with amine-terminated side chains in the linkers between the two acridine units (**BisA-NH$_2$**, Chart 1), which provides a higher cationic charge as well as possibility for further functionalization. Second, to evaluate the importance of stacking interactions the acridine units were replaced by smaller naphthalene moieties (**BisNP**). Finally, we prepared the control compounds **DMA1**, containing two acridine units linked in an acyclic framework, and **MonoNP**, containing one naphthalene unit endowed with two side chains identical to the linkers used in **BisNP**. In the current work, we report the binding properties of this novel series of ligands towards thymine-containing mismatched DNA duplexes and their fully matched counterparts, studied with a set of biophysical methods, i.e. thermal denaturation, mass spectrometry and fluorescent intercalator displacement (FID) experiments. We show that the most efficient ligand of the series, namely **BisNP**, has strong preference for pyrimidine–pyrimidine mismatches, binds them with nanomolar affinity and is able to significantly interfere with binding of the DNA methyltransferase M.TaqI, which binds to a TT mismatch in its recognition sequence.

MATERIALS AND METHODS

Ligands and chemicals

The details of synthesis and characterization of compounds **BisA**, **BisNP** and **DMA1** have been published elsewhere (53–55). The preparation and characterization of compounds **BisA-NH$_2$** and **MonoNP** are given in the Supplementary Data. The identity and purity of all samples have been confirmed by ^{1}H and ^{13}C NMR, mass-spectrometric analysis, microanalysis and/or HPLC data. Stock solutions (c = 2 mM) of the ligands (hydrochloride salts) were prepared in purified water and stored at 4°C. All buffers were prepared from biochemistry-grade chemicals (Fluka AG, Buchs, Switzerland).

DNA

For thermal denaturation studies, the 12-mer oligonucleotides **I** (5′-GTTCGTAGTAAC-3′) and **II$_X$** (5′-GTT ACT**X**CGAAC-3′; **X** = T, C, G or A) (synthesis scale 200 nmol, purified by RP-HPLC) were purchased from Eurogentec (Seraign, Belgium). The quality of oligonucleotides was confirmed by mass-spectrometric analysis data provided by the manufacturer. The 14-mer oligonucleotides **III** (5′-GGGGTCGTAGTGGC-3′) and **IV$_X$** (5′-GCCACT**X**CGACCCC-3′, **X** = T, C, G or A), used in electrospray ionization mass spectrometry (ESI-MS) experiments, were obtained from Eurogentec in OliGold quality. The lyophilized oligonucleotides were dissolved in the appropriate buffer to a strand concentration of 500 μM.

For FID and DNA enzyme competition experiments, the 17-mer oligonucleotides **V$_Y$** (5′-CCAGTTCG**Y**AGT AACCC-3′; **Y** = T, C, G or A) and **VI$_X$** (5′-GGGTT ACT**X**CGAACTGG-3′; **X** = T, C, G or A) were purchased from MWG-Biotech (Ebersberg, Germany) and purified by RP-HPLC. The 16-mer oligonucleotides **VII** (5′-GCTCTGCTCG**2**TGCCG-3′; **2** = 2-aminopurine) and **VIII** (5′-CGGCATCG**M**GCAGAGC-3′; **M** = N^6-methyladenine) were synthesized using a 392 DNA/RNA synthesizer (Applied Biosystems, Forster City, CA, USA) and purified by RP-HPLC. Duplexes were formed by heating fully complementary or single-mismatched oligonucleotides in the same buffer as used for M.TaqI binding experiments [20 mM Tris–acetate, 10 mM Mg(OAc)$_2$, 50 mM KOAc, 1 mM DTT, 0.01% Triton X-100, pH 7.9], but without Triton X-100, to 95°C and slow cooling to room temperature within 2–3 h. Annealing of ODN **V$_Y$** and **VI$_X$** yielded the 17-bp duplex oligonucleotides **17-YX** and annealing of ODN **VII** and **VIII** yielded the 16-bp duplex oligonucleotide **16-2T**.

DNA thermal denaturation studies

The thermal denaturation profiles were recorded on a UVIKON XL double-beam spectrophotometer equipped with a thermoelectric temperature controller and an inert-gas port for measurements below the dew-point temperature. The oligonucleotides **I** and **II$_X$** (3 μM each) and indicated amount of ligands in sodium cacodylate buffer (10 mM $NaAsMe_2O_2$, 50 mM NaCl, pH 6.0) were heated from ambient temperature to 80°C at a rate of 2.5°/min, kept at 80°C for 5 min, cooled to 5°C at a rate of 1.0°/min, kept at 5°C for 10 min to anneal the duplexes (**12-TX**) and finally heated to 80°C at a rate of 0.5°/min. During the latter ramp, the absorbance was monitored at 260 and 270 nm. The temperatures of DNA-melting transitions, T_m, were determined from the first-derivative plots of absorbance versus temperature. In the case of 17-bp duplexes **17-TX**, the experiments were performed in a buffer solution of lower ionic strength (10 mM $NaAsMe_2O_2$, 10 mM NaCl, pH 6.0) and the concentration of duplexes in samples was 6 μM.

Mass spectrometry

ESI-MS experiments were performed with a Q-TOF Ultima Global mass spectrometer (Micromass/Waters).

The ESI source was operated in negative-ion mode with a voltage on the spray capillary of −2.2 kV. The negative ions produced by the source were transferred to the mass analyzer using the following experimental settings: cone and RF lens 1 voltages of 100 and 35 V, respectively; source and desolvation temperatures of 70 and 100°C, respectively; source backing pressure of 2.77 mbar; hexapole collision voltage of 10 V. These settings ensured that the 14-mer mismatch-containing duplexes and their complexes remained intact, and that the relative intensities of the complexes were proportional to the relative concentrations in solution. Unless stated otherwise, the injected solutions contained 5 μM of **14-TX** duplexes. The 14-mer duplexes were hybridized by mixing both strands at 50 μM concentration and heating to 80°C for 5 min followed by overnight cooling before storage at 4°C. The duplex and ligand were mixed and diluted to a concentration of 5 μM each, in a mixture of aqueous 150 mM ammonium acetate buffer, pH 6.0 (85%) and methanol (15%), which was added immediately prior to sample injection to increase the ion signals. Knowing the total DNA concentration, the concentrations of free duplex, 1 : 1 and 2 : 1 complexes (two ligand molecules bound to one duplex) were determined from the relative peak areas, A (Equations 1–3), and the amount of free ligand was calculated from the mass-balance equation (Equation 4). The equilibrium-binding constants for the (1 : 1)-type complex, K_a, were then calculated using Equation 5.

$$[\text{duplex}]_{\text{free}} = [\text{duplex}]_{\text{total}} \times \frac{A(\text{duplex})}{A(\text{duplex}) + A(1:1) + A(2:1)} \quad \mathbf{1}$$

$$[1:1] = [\text{duplex}]_{\text{total}} \times \frac{A(1:1)}{A(\text{duplex}) + A(1:1) + A(2:1)} \quad \mathbf{2}$$

$$[2:1] = [\text{duplex}]_{\text{total}} \times \frac{A(2:1)}{A(\text{duplex}) + A(1:1) + A(2:1)} \quad \mathbf{3}$$

$$[\text{ligand}]_{\text{free}} = [\text{ligand}]_{\text{total}} - [1:1] - 2 \times [2:1] \quad \mathbf{4}$$

$$K_a = \frac{[1:1]}{[\text{duplex}]_{\text{free}} \times [\text{ligand}]_{\text{free}}} \quad \mathbf{5}$$

Fluorescence titrations

Fluorescence titrations were performed with a Varian Cary Eclipse fluorescence spectrophotometer equipped with a thermo-controlled cell holder. For the FID titrations, the excitation and emission wavelengths were set to 520 nm and 615 nm, whereas for titrations with the 2-aminopurine-labeled DNA they were set to 320 and 384 nm, respectively. The spectral bandwidth was 5 nm for the excitation and 10 nm for the emission. The titrations were performed at 25°C in a Suprasil QS cell (Hellma) with an optical path of 1 cm. The DNA and ethidium bromide or **BisNP** concentrations were kept constant throughout the whole titration by adding solutions with the same DNA and ethidium bromide or **BisNP** concentrations in addition to macrocyclic ligands or M.TaqI. The FID titrations were carried out with **17-YX** duplexes (100 nM) and ethidium bromide (333 nM or 1 μM) in M.TaqI-binding buffer. The titrations with M.TaqI were performed with either (I) **16-2T** (100 nM), (II) **16-2T** (100 nM) and **17-TT** (100 nM) or (III) **16-2T** (100 nM), **17-TT** (100 nM) and **BisNP** (2 μM) in M.TaqI-binding buffer. All titration curves were fitted analytically with Excel Solver, using the appropriate equilibrium equations and a numerical simulation of titration (III) was accomplished with Maple (Supplementary Data). The fluorescence decrease, shown in Figures 5 and 8, was calculated from the fluorescence intensity $F_{1.2}$ after addition of 1.2 eq. of the ligand with regard to the background fluorescence $F_{\text{background}}$ (i.e. the fluorescence of the solution without any DNA) and the fluorescence in the beginning of the titration F_0 (i.e. the fluorescence of the DNA-containing solution before addition of the ligand) (Equation 6).

$$\text{Fluorescence decrease (\%)} = \left(1 - \frac{F_{1.2} - F_{\text{background}}}{F_0 - F_{\text{background}}}\right) \times 100\% \quad \mathbf{6}$$

The relative fluorescence intensity change, shown in Figure 6, was calculated from the fluorescence intensity F with regard to the fluorescence at the beginning of the titration, F_0, and the fluorescence at the end of the titration, F_{end} (Equation 7).

$$\text{Relative fluorescence intensity change} = \left(\frac{F - F_0}{F_{\text{end}} - F_0}\right) \quad \mathbf{7}$$

DNA enzyme M.TaqI

The DNA adenine-N^6 methyltransferase from *Thermus aquaticus* (M.TaqI) was overexpressed and purified as reported previously (56, 57). The enzyme was stored in buffer (10 mM Tris–HCl, 300 mM KCl, 0.1 mM EDTA, 1 mM DTT, 50% glycerol, pH 7.4) at −20°C. The M.TaqI concentration was estimated by UV absorption at 280 nm using an extinction coefficient of 86 600 L/(cm mol).

RESULTS

Thermal denaturation studies

Thermal denaturation of DNA is a rapid and straightforward method for the determination of the stabilization effect of ligands on duplex DNA. The extent of this stabilization provides a semi-quantitative evaluation of ligand affinity towards duplex DNA. Base mismatches significantly reduce the thermodynamic stability of a DNA duplex, and the extent of this destabilization is dependent on the duplex length. As a consequence, the stabilization effect of mismatch-binding ligands becomes more pronounced under conditions of low duplex stability (low

12-TX
5′-GT TCG**T**AGT AAC-3′
3′-CA AGC**X**TCA TTG-5′

14-TX
5′-GGGG TCG**T**AGT GGC-3′
3′-CCCC AGC**X**TCA CCG-5′

17-YX
5′-CCAGT TCG**Y**AGT AACCC-3′
3′-GGTCA AGC**X**TCA TTGGG-5′

16-2T
5′-GCTCTGCTCG**2**TGCCG-3′
3′-CGAGACG**M**GC**T**ACGGC-5′

Chart 2. Sequences of duplex oligodeoxynucleotides used in this study. **X**, **Y** = A, C, G or T; **2** = 2-aminopurine; **M** = N^6-methyladenine.

ionic strength, short duplexes). Hence, to improve comparison between the various compounds, we chose to use the dodecamer duplexes **12-TX**, significantly shorter than the 17-mer duplexes **17-TX** used in our previous study (37) and in the FID assay of the present study (Chart 2). However, in both cases the central 7-bp region of the sequence, containing the 4-bp recognition sequence of M.TaqI (5′-TCGY-3′, where Y = A for the natural substrate), was kept constant in order to avoid any misinterpretations that might arise from the sequence selectivity of the ligands and influence of the neighboring base pairs on the stability of mismatched base pairs.

At the employed conditions (pH 6, $[Na^+] = 60\,mM$), the fully matched duplex **12-TA** had a melting temperature of 36.0°C, whereas the mismatch-containing duplexes **12-TG, 12-TT** and **12-TC** denaturated at significantly lower temperatures (27.3, 20.8 and 17.9°C, respectively). The ligand-induced changes of the melting temperatures (ΔT_m) at several ligand-to-duplex ratios (q) are listed in Table S1 (Supplementary Data) and represented as a bar graph in Figure 1, to facilitate direct comparison between the effect of the various ligands.

All tested macrocyclic ligands strongly stabilize the mismatch-containing duplexes, whereas their effect on the fully matched duplex **12-TA** is much less pronounced. Both bis-acridine derivatives **BisA** and **BisA-NH$_2$** stabilize all three mismatch-containing duplexes, with **BisA-NH$_2$** showing larger increases of melting temperature, which can be attributed in part to its higher cationic charge (6+ versus 4+ for **BisA** at pH 6.0). Notably, the TC-mismatched duplex **12-TC** was preferentially stabilized with these bis-acridine ligands and in most cases a significant increase in ΔT_m was observed when the ligand concentration was increased from $q = 1$ to $q = 2$, which indicates that binding was not saturated with one equivalent of ligand. Remarkably, both compounds induce no stabilization of the fully matched duplex **12-TA**, even at the highest concentration ($q = 3$). Moreover, a small decrease of T_m of **12-TA** was observed with **BisA-NH$_2$** ($\Delta T_m = -0.6$°C at $q = 2$), which may be attributed to a certain affinity for the single strands.

Likewise, the bis-naphthalene macrocycle **BisNP** induced a pronounced stabilization of the three mismatch-containing duplexes but with a clear preference for the TT-mismatched duplex **12-TT** which was stabilized

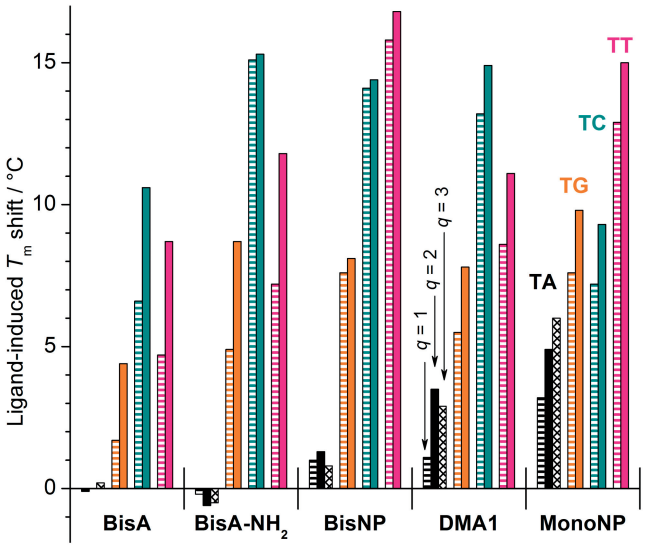

Figure 1. Ligand-induced changes of melting temperature (ΔT_m) of the fully matched (**12-TA**: black) and mismatch-containing duplexes (**12-TG**: orange, **12-TC**: cyan, **12-TT**: magenta bars) at ligand-to-duplex ratios of $q = 1$ (horizontally hatched bars), $q = 2$ (filled bars) and $q = 3$ (cross-hatched bars, only for **12-TA** duplex); [**12-TX**] = 3 μM; estimated error in T_m determination is ±1.0°C.

better than the TC analogue ($\Delta T_m = +15.8$°C and +14.1°C at $q = 1$ for **12-TT** and **12-TC**, respectively). In this case, the increase of ligand-to-duplex ratio ($q = 1$ to $q = 2$) resulted in almost no significant increase of ΔT_m, which indicates that binding saturation was almost achieved with one equivalent of **BisNP**, revealing a high affinity for the mismatched duplexes. Additionally, only small stabilization of the fully matched duplex **12-TA** was observed even at the highest ligand concentration ($\Delta T_m = +1.0$°C at $q = 3$) indicating the poor ability of **BisNP** to associate with the matched duplex.

Finally, both acyclic control compounds, **DMA1** and **MonoNP** stabilize the mismatch-containing duplexes to an extent comparable to that of the macrocycles and with a similar preference: the acridine derivative **DMA1** preferentially stabilizes **12-TC**, whereas the naphthalene derivative **MonoNP** has the most pronounced effect on **12-TT**. Most importantly, both control compounds, especially **MonoNP**, show strong stabilization of the fully matched duplex **12-TA** ($\Delta T_m = +1.1$ and +3.2°C at $q = 1$ for **DMA1** and **MonoNP**, respectively) which was further increased as the ligand concentration was raised (for **MonoNP** $\Delta T_m = +6.0$°C at $q = 3$). Altogether, thermal denaturation data indicate that both control compounds have much lower selectivity for the mismatched over the matched duplexes as compared with the macrocyclic ligands.

In order to confirm the non-selective binding of the control compounds, thermal denaturation experiments were repeated for **BisNP**, **DMA1** and **MonoNP** with the longer duplexes **17-TX** (Chart 2) that offer additional non-specific matched binding sites. The results (Figure 2; Supplementary Table S2) showed that the macrocycle **BisNP** retained pronounced selectivity for the mismatched

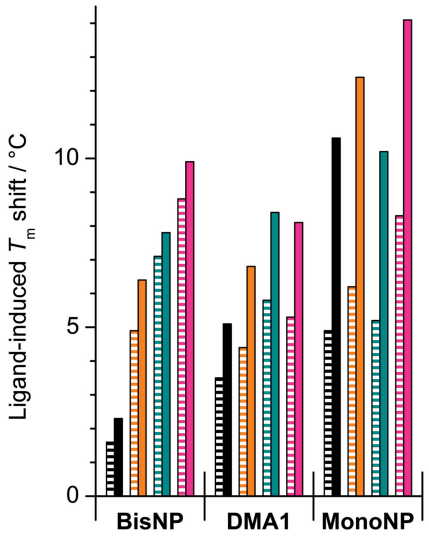

Figure 2. Ligand-induced changes of melting temperature (ΔT_m) of fully matched and mismatch-containing duplexes **17-TX** (for sequence see Chart 2); [**17-TX**] = 6 μM; estimated error in T_m determination is ±0.5°C. For assignment of data sets see Figure 1 caption.

duplexes, and showed the same behaviour observed with the short duplexes (preference for the TT mismatch and binding saturation almost complete at $q = 1$). In contrast, **DMA1** and especially **MonoNP** displayed significant stabilization of the fully matched duplex **17-TA** ($\Delta T_m = +3.5$ and $+4.9$°C at $q = 1$, respectively). Especially in the case of **MonoNP** ΔT_m values were strongly increased in the presence of two equivalents of ligand, presumably due to pronounced non-selective binding that is consistent with the higher cationic charge of **MonoNP** (4+) compared to **DMA1** (2+). Thus, **MonoNP** was no longer able to discriminate between mismatched and fully matched 17-bp duplexes. Since the charge state and aromatic units of **MonoNP** and **BisNP** are identical, these data highlight that the macrocyclic scaffold is required for suppression of non-selective binding to the fully matched DNA duplexes. Finally, it is worth noticing that the same trends in terms of ligand binding can be inferred from measurements with the two sets of duplexes **12-TX** and **17-TX** (Supplementary Figure S1), which is important because the **17-TX** set was used for further FID experiments (see subsequently).

ESI-MS studies

ESI-MS is a powerful method for monitoring the binding of small molecules to short DNA duplexes (58–60). Mass spectrometry analysis of complex mixtures requires only minute amounts of sample and, compared with spectrophotometric techniques, provides direct information on the stoichiometry of the interaction for non-covalent complexes. Furthermore, the affinity and selectivity of DNA and RNA ligands may be reliably determined by ESI-MS, since the relative peak intensities are proportional to the relative abundance of the corresponding species in solution (61–63).

The ESI-MS experiments with mismatch-containing and fully matched duplexes were performed at conditions as close as possible to those of the thermal denaturation experiments, i.e. at a duplex concentration of 5 μM and a ligand concentration of 5 μM ($q = 1$). However, the 14-bp duplexes **14-TX** (Chart 2), containing additional GC base pairs at both termini and having higher thermodynamic stability, were used instead of **12-TX** to minimize dissociation of the duplex into single strands. Duplexes and duplex–ligand complexes were detected at charge states 6– and 5–. It should be noted that variation of the charge of ligand species (e.g. 6+ for **BisA-NH$_2$** versus 4+ for **BisA**) hardly influences the charge of the duplex–ligand complex, since the charge state or the distribution of charge states observed in ESI-MS depend more on the total size of the complex than on the spatial distribution of charges within a complex (60). Representative ESI-MS spectra of **14-TX** in the presence of **BisA**, **BisA-NH$_2$** and **BisNP** are shown for the charge state 6– in Figure 3. In the case of all mismatch-containing duplexes, the major peak corresponds to the ligand–oligonucleotide complex with a 1 : 1 stoichiometry. The relative abundances of the 1 : 1 complex and of the free duplex vary significantly with the type of mismatch and the nature of the ligand, which reflects differences in the stability of the ligand–DNA complexes. However, in the case of the fully matched duplex (**14-TA**), the 1 : 1 species are much less abundant and barely detectable with **BisA-NH$_2$** even at ligand concentration of 12 μM, i.e. at $q = 2.4$ (Figure 3B, lower spectrum). This observation indicates that the macrocyclic ligands have much lower affinity to the fully matched duplex than to the duplexes with a mismatch.

The mass spectra also reveal low–abundant 2 : 1 complexes (peaks marked with arrows in Figure 3) in the case of **BisNP** with all the duplexes and for **BisA** with **14-TA** and **14-TG** duplexes. This indicates the existence of a secondary binding site. No 2 : 1 complex was detected with **BisA-NH$_2$** (Figure 3B). Finally, both **BisA** and **BisNP** also bind to the remaining single strands (peaks marked by circles in Figure 3). This was confirmed by ESI-MS spectra of single strands in the presence of ligands (data not shown). In the case of control compounds **DMA1** and **MonoNP** a general lower prevalence of the 1 : 1 complexes relative to the free duplexes was observed for the three mismatched duplexes (Supplementary Figure S2).

The peak areas in the mass spectra were used for the calculation of the equilibrium affinity constants, K_a, of the duplex–ligand complexes (Equations 1–5, Materials and methods section), which are summarized in Table 1. The K_a values for the macrocyclic ligands, i.e. **BisA**, **BisA-NH$_2$** and **BisNP**, and the duplex with a TC or TT mismatch, i.e. **14-TC** or **14-TT**, are in the range 10^6–10^7 M^{-1}, whereas the affinity to the fully matched duplex **14-TA** is significantly lower, especially in the case of **BisA-NH$_2$** (K_a around 10^4 M^{-1}). Moreover, the difference in binding affinity to various mismatched duplexes agrees quite well with the selectivity observed in the thermal denaturation experiments. Thus, the less thermodynamically stable TT and TC mismatches are preferentially bound by the

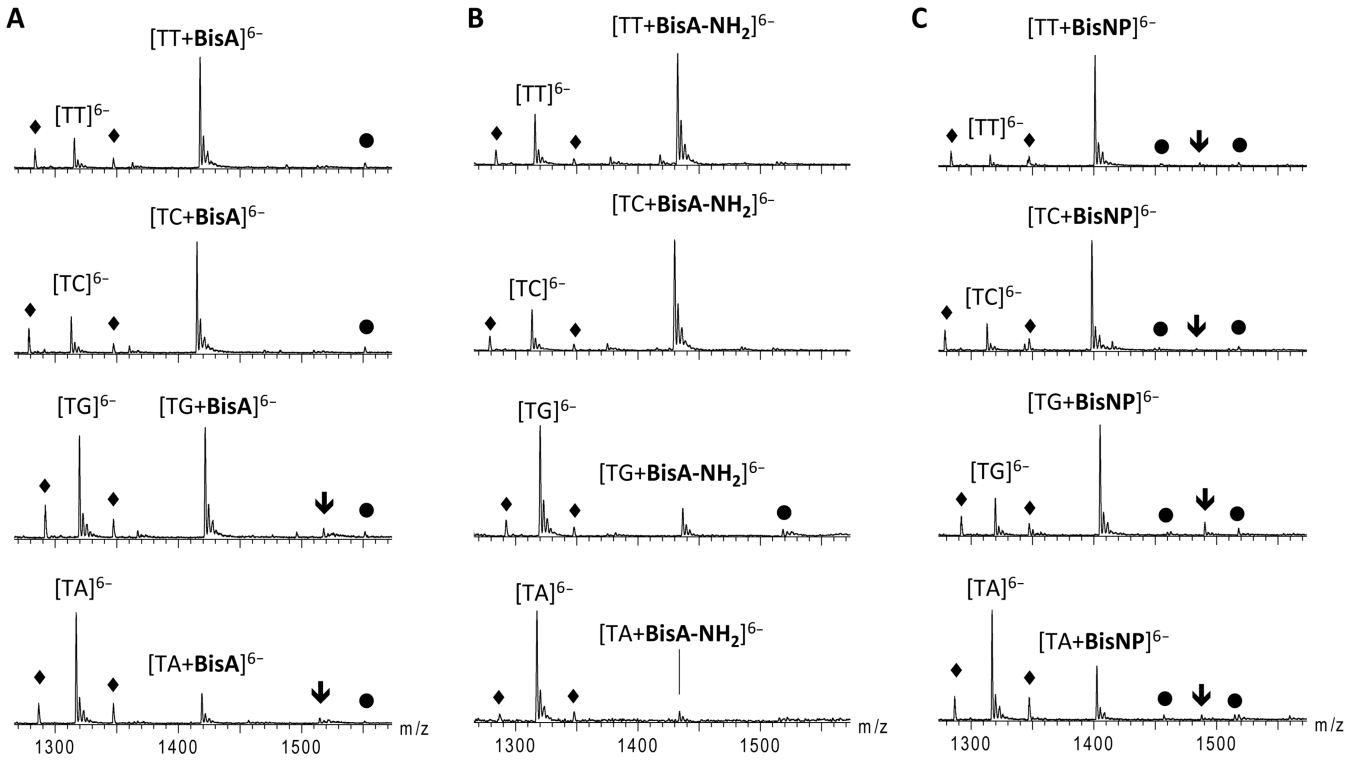

Figure 3. ESI-MS spectra of duplexes **14-TX** in the presence of macrocycles **BisA** (**A**), **BisA-NH$_2$** (**B**) and **BisNP** (**C**). [**14-TX**] = [ligand] = 5 μM, except for **14-TA** and **BisA-NH$_2$** (B, lower spectrum), where [ligand] = 12 μM. The free duplexes are labeled as [**TX**]$^{6-}$ and the 1 : 1 complexes as [**TX** + ligand]$^{6-}$. The diamonds indicate peaks corresponding to the triply charged single strands, the circles indicate triply charged 1 : 1 complexes with single strands when detected, and the arrows indicate 2 : 1 complexes [**TX** + 2 L]$^{6-}$ when detected.

Table 1. Affinity constants for binding of ligands to duplexes **14-TX** determined by ESI-MS, and selectivities for the mismatched duplexes compared to the fully matched duplex[a]

	$K_a/10^6$ M^{-1} (Selectivity[b])			
	X = A	**X** = G	**X** = C	**X** = T
BisA	0.10 (1.0)	0.32 (3.2)	1.8 (18)	2.0 (20)
BisA-NH$_2$	0.0080[c] (1.0)	0.063 (7.9)	2.0 (250)	1.6 (200)
BisNP	0.20 (1.0)	2.0 (10)	3.2 (16)	14 (70)
DMA1	0.063 (1.0)	0.063 (1.0)	0.63 (10)	0.32 (5.0)
MonoNP	0.083 (1.0)	0.25 (3.0)	0.25 (3.0)	0.56 (6.8)

[a]Experimental conditions: [**14-TX**] = [ligand] = 5 μM.
[b]defined as K_a(TX)/K_a(TA).
[c][ligand] = 12 μM.

ligands, whereas the TG mismatch is recognized with variable efficiency, depending on the ligand. The highest affinity for TG, TT and TC mismatches was observed with **BisNP**, which is consistent with the T_m data.

To estimate the selectivity of the ligands for the mismatched duplexes over the fully matched one, the K_a(TX)/K_a(TA) ratios were calculated (Table 1). The acyclic control compounds **DMA1** and **MonoNP** display both low affinity and low specificity for the mismatched duplexes. However, a small binding preference of **DMA1** for **14-TC** and **MonoNP** for **14-TT** may be noticed, in agreement with the T_m data (Figure 1). In contrast, the macrocyclic ligands **BisA** and **BisNP** showed high selectivity for the TT and TC mismatched duplexes **14-TT** and **14-TC** (16- to 70-fold increased binding affinity compared to **14-TA**). This selectivity was even more pronounced in the case of **BisA-NH$_2$** (200- to 250-fold increased binding affinity). Compared to the other ligands, this exceptionally high selectivity is rather due to the much lower affinity for the fully matched duplex than to a higher affinity for the mismatched duplexes. This low affinity to the fully matched DNA was rather unexpected given the high positive charge of **BisA-NH$_2$**.

FID studies

The FID assay is used to investigate binding of ligands to various DNA structures and can provide apparent binding constants (64–66). In this assay, the DNA is fluorescently stained with a dye, typically ethidium bromide or thiazole orange, which has a low fluorescence or is almost non-fluorescent in the unbound state but becomes strongly fluorescent upon intercalation into DNA. Binding of ligands to the DNA–intercalator complex leads to a displacement of the bound fluorophore that is accompanied by a decrease in fluorescence intensity. The fluorescence decrease as a function of added ligand enables monitoring the binding event and determination of ligand affinity towards DNA. In the present study, we used the FID assay to determine the binding affinities of macrocyclic ligands to the TT mismatch-containing duplex **17-TT**.

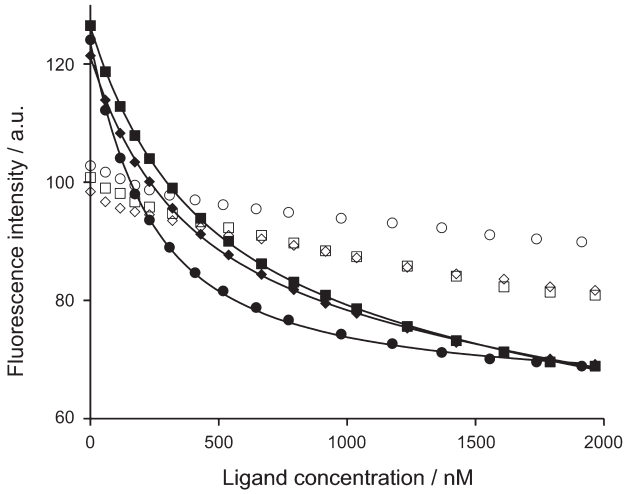

Figure 4. FID titrations (excitation at 520 nm and emission at 615 nm) of **17-TT** (filled symbols) and **17-TA** (empty symbols) ($c = 100$ nM each) in the presence of ethidium bromide ($c = 333$ nM) with **BisA** (squares), **BisA-NH$_2$** (diamonds) or **BisNP** (circles). The solid lines represent the analytical fits for a 1:1 binding model.

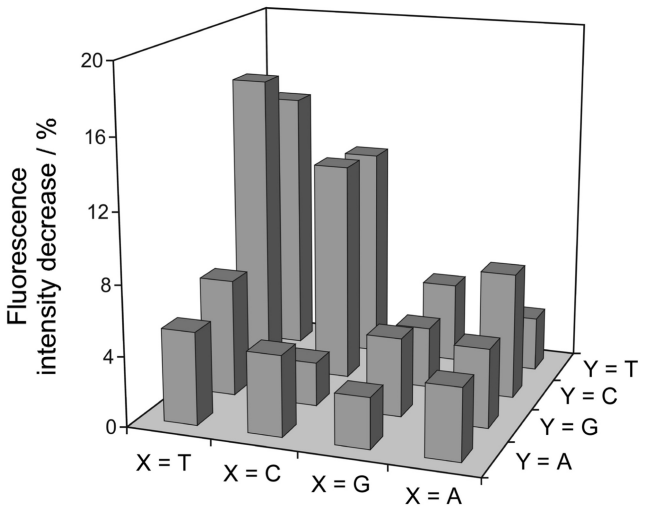

Figure 5. Relative fluorescence intensity decrease (excitation at 520 nm and emission at 615 nm) upon addition of **BisNP** ($c = 120$ nM) to all sixteen **17-YX** duplexes (100 nM each) in the presence of ethidium bromide (1 μM).

We chose ethidium bromide as fluorescent intercalator, because its low binding affinity towards double-stranded DNA (about $10^5\,M^{-1}$) allows quantitative analysis of titration data without the risk of underestimating binding affinities in the nanomolar range (67,68). The FID assay was performed in a buffer (pH 7.9, 10 mM Mg^{2+} and 50 mM K^+) which is also suitable for the DNA enzyme competition experiments described subsequently.

Binding of the macrocyclic ligands to the TT mismatch-containing duplex **17-TT**, as monitored by the decrease of the fluorescence signal from ethidium bromide, is shown in Figure 4. Addition of all ligands to the **17-TT**–ethidium bromide complex caused a distinct decrease of fluorescence intensity and yielded apparent affinity constants K_a of 2.9, 3.5 and $6.7 \times 10^6\,M^{-1}$ for **BisA**, **BisA-NH$_2$** and **BisNP**, respectively, which are similar to the values obtained from the ESI-MS experiments (Table 1). In contrast, only a slight decrease in fluorescence intensity was observed when the macrocyclic compounds were added to the fully matched **17-TA** duplex. No reliable binding affinities could be determined from these titrations because of the low fluorescence decrease. It should be noted that the initial fluorescence intensity of the **17-TA**–ethidium bromide complex is about 20% lower compared with the **17-TT**–ethidium bromide complex which may reflect a lower quantum yield and/or a slightly lower affinity of ethidium bromide for the fully matched duplex.

All three assays (i.e. thermal denaturation, ESI-MS and FID) demonstrate that the naphthalene derivative **BisNP** has the highest mismatch affinity among all investigated macrocycles, and it thus was chosen for further FID screening experiments with all 16 possible base combinations at the variable position. Figure 5 shows the relative fluorescence decrease after addition of 1.2 eq. of **BisNP** to all possible **17-YX**–ethidium bromide complexes (the corresponding numeric values are given in Supplementary Table S3). The largest decrease was observed with the DNA duplexes **17-TT**, **17-TC**, **17-CT** and **17-CC**, indicating a strong binding preference for pyrimidine–pyrimidine mismatches.

Binding competition with a DNA enzyme

A competition assay with the DNA adenine-N^6 methyltransferase M.TaqI (57) was performed to test whether the small ligand **BisNP** is able to interfere with DNA binding of an enzyme. M.TaqI, like other DNA methyltransferases and DNA glycosylases, flips its target base out of the DNA helix for catalysis and DNA binding can be conveniently monitored by replacing the target adenine with the fluorescent base analogue 2-aminopurine (2Ap) within the 5′-TCGA-3′ recognition sequence. 2Ap forms a Watson–Crick-like base pair with thymine and thus does not alter the structure of DNA significantly. The fluorescence of 2Ap is efficiently quenched when it is stacked inside the DNA double helix, whereas binding of M.TaqI leads to a strong fluorescence increase caused by destacking (base flipping) (56,69,70). This is illustrated in Figure 6A (I) and a titration curve of the DNA duplex **16-2T** containing 2Ap at the target position with M.TaqI is shown in Figure 6B (curve I). In the presence of the TT mismatch-containing DNA duplex **17-TT**, however, M.TaqI preferentially binds to the mismatch-containing DNA (Figure 6A, II), which leads to a retarded 2Ap fluorescence increase and the observed sigmoid-shaped titration curve reflects competition between the labeled and unlabeled duplex (Figure 6B, curve II) (71). High-affinity binding of M.TaqI to DNA with a mismatch at the target position is not surprising, since such a behavior has been observed with other DNA methyltransferases before and was attributed to the energetic ease of flipping nucleobases from these thermodynamically weakened sites (72,73). Regression analysis of this competitive titration curve (Figure 6B, curve II) gave an affinity constant K_a for M.TaqI and **17-TT** of $280 \times 10^6\,M^{-1}$, which is

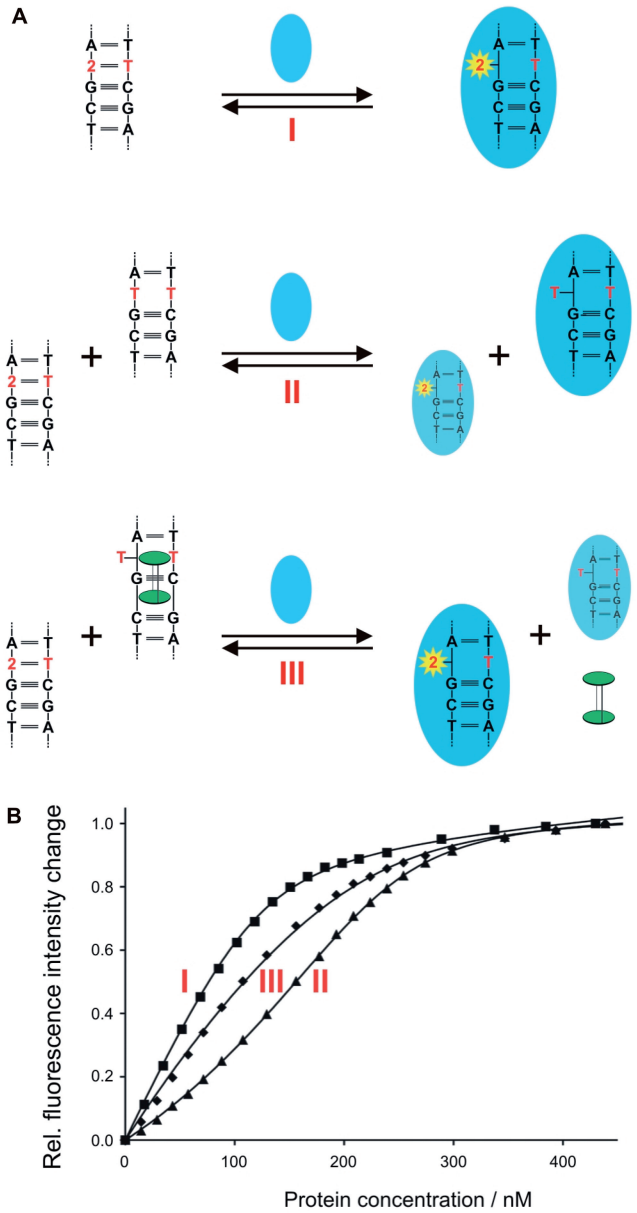

Figure 6. Competition of **BisNP** (green) and M.TaqI (blue) for a TT-mismatch in DNA. (**A**) Principle of the competitive binding assay: (I) binding of M.TaqI to a DNA duplex with 2-aminopurine (**2**) at the target position leads to base flipping and a fluorescence increase (yellow) of 2-aminopurine; (II) in the presence of a DNA duplex with a TT-mismatch at the target position, the 2-aminopurine fluorescence increase is retarded; (III) adding **BisNP** to the two DNA duplexes leads to a partial reversal of the fluorescence increase retardation. (**B**) Corresponding fluorescence titration curves (excitation at 320 nm and emission at 384 nm), with I: **16-2T** (100 nM), II: **16-2T** and **17-TT** (100 nM each), III: **16-2T** and **17-TT** (100 nM each) in the presence of **BisNP** (2 μM). Solid lines: curve fittings according to a one equilibrium (I) or a two equilibria (II, III) binding model for M.TaqI.

higher than the K_a value for M.TaqI and **16-2T** of $100 \times 10^6\,M^{-1}$ obtained by direct titration (Figure 6B, curve I). When this competition experiment was repeated in the presence of **BisNP** at a concentration of 2 μM (Figure 6A, III), the increase of 2Ap fluorescence intensity was much less retarded and the shape of the curve was no longer sigmoidal (Figure 6B, curve III). This demonstrates that the binding affinity of M.TaqI to the mismatched duplex had apparently decreased. Regression analysis of this titration curve in the presence of **BisNP** gave an *apparent* affinity constant K_a^{app} for M.TaqI and **17-TT** duplex of $55 \times 10^6\,M^{-1}$, which is about five times lower than in the absence of **BisNP**. Since the presence of **BisNP** does not significantly influence the binding of M.TaqI to **16-2T** (Supplementary Figure S3), this effect can be directly attributed to **BisNP** successfully competing with M.TaqI for the mismatched TT-binding site. In fact, the shape of curve III in Figure 6B can be well predicted from the affinity constants of M.TaqI to the DNA duplexes and the measured binding affinity of **BisNP** to the TT-mismatch ($K_a = 6.7 \times 10^6\,M^{-1}$, from FID) by numerical fitting (Supplementary Figure S4). When the competitive titration was performed with the fully matched duplex **17-TA** instead of **17-TT**, no apparent reduction in binding affinity was observed (Supplementary Figure S5).

DISCUSSION

We previously reported that the macrocyclic bis-acridine derivative **BisA** selectively binds the TX mismatches with high affinity (37). Using this ligand as a prototype, we investigated the interaction of two macrocyclic analogues and two acyclic control compounds, having one or two aromatic units, with various mismatch-containing duplexes using three independent biophysical methods. In addition, screening of ligand binding to all possible base combinations was carried out. Finally, we examined the potential of the most promising mismatch binder **BisNP** for interference with a DNA-binding enzyme (M.TaqI).

Ligand structural prerequisites for selective mismatch-binding

Thermal denaturation and ESI-MS studies both demonstrate the crucial role of the cyclic topology of the ligand to obtain high mismatch selectivity. The macrocyclic bis-intercalators are known to adopt a semi-closed conformation with a constrained distance of ~7 Å between the two aromatic moieties (36,54). This provides a structural explanation for the observed discrimination between fully matched and mismatched DNA: insertion into fully matched DNA is disfavored because of steric hindrance due to the macrocyclic scaffold. In addition, the short distance between the two intercalator units does not allow bis-intercalation between regular base pairs in line with the nearest-neighbor exclusion principle (74), as shown for a structurally related phenanthridine dimer (75). However, it has been previously shown that **BisA** exhibits strong affinity for nucleobases and phosphate residues (53). When offered a site of low thermodynamic stability (i.e. a base mispair), these macrocyclic compounds obviously take advantage and thread into the helix, as has been observed in the NMR structure of **BisA** bound to an abasic-site-containing duplex (36). This threading bis-intercalation mode can also be expected for the binding

of the studied macrocycles to DNA with a base mismatch, as schematized in Figure 6A (III).

Interestingly, the macrocycle **BisNP** shows a considerable enhancement in mismatch affinity compared to the parent bis-acridine **BisA**, which bears the same tetracationic charge. Also the replacement of the central NH by an oxygen atom in the linker was shown to have no influence on mismatch stabilization (data not shown). The main structural difference between the two compounds resides in the size of the cavity delineated by the two intercalators which is smaller in the case of **BisNP** and might be more favorable to a tight trapping of nucleic bases and in particular pyrimidines (53).

Finally, it is clear that the two control compounds do not feature the same structural advantages as those of the macrocycles because **DMA1** is an acylic dimer of high flexibility and **MonoNP** has a simple flat monomeric structure. As a result, both compounds bind efficiently to the fully matched and mismatched DNA, although retaining a low selectivity for TT and TC mismatches.

Comparison of data obtained by the different methods

Binding of the macrocyclic ligands to TT-mismatched duplexes was analyzed by thermal denaturation, ESI-MS and FID titrations and the results are compared in Figure 7. All methods show that among the three macrocycles, **BisNP** is the best mismatch binder. Taking into account that ESI-MS and FID are two entirely different methods to obtain affinity constants and that relative affinities are evaluated by the T_m method, the good qualitative agreement between all three methods and the quantitative agreement between the ESI-MS and FID data for the different compounds are remarkable. A possible source for minor deviations are the conditions of the FID experiments, which were performed at higher pH (7.9) than the ESI-MS and T_m measurements (pH 6.0). This choice of conditions was justified from the viewpoint of compatibility with the enzyme-binding experiments. Since the pK_a values of the benzylic nitrogens of analogous macrocycles fall in the range 7–9 (76), it may be assumed that the macrocyclic ligands exist predominantly as tetracationic species at pH 6.0 and as tricationic species at pH 7.9. Thus, the conditions of the FID assay are *a priori* less favorable for ligand–DNA interaction due to reduced cationic charge of the ligands, which minimizes electrostatic contributions. Nonetheless, it is remarkable that even in such conditions strong binding of all three macrocycles was achieved. The absence of pH-dependency of the interaction indicates that mismatch recognition by macrocycles is governed by nonionic forces, likely π–π stacking between the two aromatic units of the ligand and the DNA bases.

Selectivity for pyrimidine–pyrimidine mismatches

Binding of **BisNP** to the four possible TX base combinations was studied by the three methods and data are compared in Figure 8. As the FID assay showed only little response with duplexes **17-TA** and **17-TG** (Figure 4), the binding constants could not be determined and the decrease of fluorescence intensity of ethidium bromide upon addition of 1.2 equivalents of **BisNP** was used to allow comparison of ligand affinity. As a general trend, a preference for TT and TC mismatches was observed though the affinity for the TC-mismatched duplex is less obvious from the ESI-MS data. Finally, binding to the TG-mismatched DNA was clearly weaker albeit a significant stabilization was obtained in thermal denaturation experiments.

The interaction of **BisNP** with other mismatches was further analyzed by FID screening of all the 16 possible base combinations (Figure 5). This experiment confirmed the high selectivity for homopyrimidine mismatches (TT, CC and TC), which is rather irrespective of the strand orientation (TC versus CT). These data should be interpreted taking into account the relative stabilities of the different mismatches. Indeed, the stability of internal mismatches is strongly dependent on the sequence context and in particular on the two flanking base pairs. In our case, the studied XY mismatches are in the G**X**A/C**Y**T

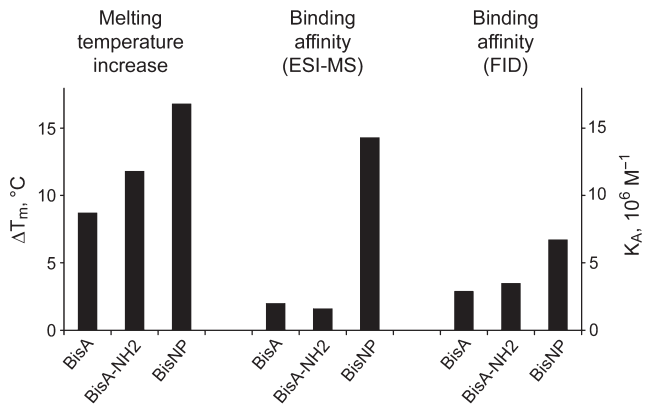

Figure 7. Comparison between ligand-induced melting temperature increase of duplex DNA (left, $q = 2$) and binding affinity constants determined by ESI-MS (center) and by FID experiments (right) for binding of **BisA**, **BisA-NH$_2$** and **BisNP** to TT mismatch-containing DNA duplexes **12-TT** (T_m), **14-TT** (ESI-MS) or **17-TT** (FID).

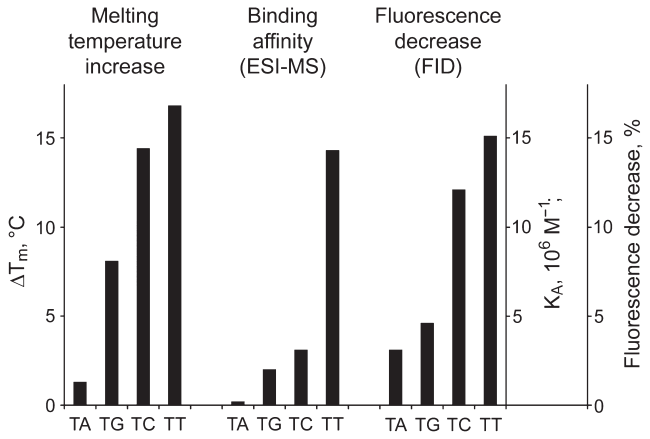

Figure 8. Comparison between **BisNP** induced shift of melting temperature of **12-TX** (left, $q = 2$), binding constants of **BisNP** to **14-TX** determined by ESI-MS (center) and background-corrected decrease of ethidium bromide fluorescence upon addition of 1.2 eq. **BisNP** to **17-TX** (right).

triplet context and, according to the nearest-neighbor model, the following stability order can be assumed: GG > GA > TG > GT ≈ AG > AA > CT ≈ CA ≈ TT ≈ CC > AC > TC (77).

Plotting of the FID data as a function of the calculated thermodynamic stabilities (T_m) of the various mismatches (Supplementary Figure S6) allows to draw several conclusions. Firstly, as previously observed (37), the recognition of the three TX mispairs is inversely correlated with their respective stability. This holds also for the CC and CT mispairs (Figures 5 and S6). Similarly, the poor recognition of the homopurine mismatches (GG, GA, AG and AA) which is comparable to that of TG and GT, is consistent with their higher stability compared with the homopyrimidine mispairs. Finally, the correlation between binding affinity and stability is not verified with the purine-pyrimidine mispairs CA and AC whose stability is quite low. This observation indicates that mismatch stability alone is not sufficient to rationalize all binding selectivities observed for the macrocyclic ligand **BisNP**. Apparently, steric factors like H-bonding patterns and the geometry specific to each mismatch as well as local dynamics also contribute to the mismatch recognition process (78).

BisNP is able to compete with a DNA-binding enzyme

To investigate whether small macrocyclic mismatch binders are able to interfere with DNA-binding of much larger endogenous ligands, such as enzymes, we designed a competition experiment using the DNA adenine-N^6 methyltransferase M.TaqI. M.TaqI was chosen based on our precedent work in which M.TaqI, a paradigm for a base flipping enzyme, was used to establish base flipping from mismatches for **BisA** (37). Although mismatches are not natural substrates of DNA methyltransferases, it was shown that introduction of a mismatch at the target site within the recognition sequence can even increase the protein-binding affinity (72,73). This is explained by the fact that the mismatched site is thermodynamically weakened compared to the canonical base pair and can facilitate the base-flipping process (73). A double-competition experiment was devised using a fluorescently 2Ap-labeled duplex and a TT mismatch duplex. Our results demonstrate that **BisNP** is able to weaken the binding of M.TaqI to the TT-mismatched duplex, whereas the binding of the enzyme to the fully matched duplex was not affected. This indicates that **BisNP** interferes with the formation of the M.TaqI–DNA complex by binding to the mismatched DNA. Given the affinity of M.TaqI to the TT-mismatched duplex ($K_a = 280 \times 10^6\,M^{-1}$), these results are very encouraging for further studies on the interference of macrocyclic bis-intercalator-type ligands with the DNA binding of mismatch-recognizing enzymes such as MutS, whose affinity for the mismatched sites is significantly lower ($K_a = 0.2 \times 10^6 - 5 \times 10^6\,M^{-1}$) (15). Finally, this concept of inhibiting DNA repair enzymes by blocking their DNA substrates with small molecules is of pharmacological interest because these enzymes often counteract the action of clinically used anti-tumor agents (79).

CONCLUSIONS

Taken together, our results suggest that the challenging discrimination of TX mispairs versus fully matched TA base pairs may be addressed by distance-constrained macrocyclic bis-intercalators such as **BisA**, **BisA-NH$_2$** and **BisNP**. This is shown by thermal denaturation, ESI-MS and FID studies and the very good agreement of these three methods, which allows us to select **BisNP** as the best mismatch binder from the series. Screening experiments with all possible base combinations confirm the strong preference of **BisNP** for homopyrimidine mismatches. In addition, it is demonstrated that the **BisNP** ligand is able to efficiently compete with a DNA-binding enzyme for a single TT-mismatch site.

Recognition of DNA mismatches is under a focus of interest as this may give clues about the initial DNA recognition event(s) triggering the complex repair process. In addition, binding of mismatches by small molecules may provide novel therapeutic alternatives to anticancer therapies. The macrocyclic family studied in the present work thus represents a new class of very efficient and selective DNA mismatch binders. Therefore, these remarkable properties make our compounds valuable candidates to further investigate potential interference with repair enzymes that directly bind mismatched DNA.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The *Centre National de la Recherche Scientifique* (CNRS) and the European FP6 program 'CIDNA' (NMP-CT-2003-505669) are greatly acknowledged for the financial support to A.G. M.-P.T.-F. and V.G. acknowledge the support of the CGRI (Belgium) and EGIDE (France) for a Tournesol (now Hubert Curien) exchange program. Funding to pay the Open Access publication charges for this article was provided by RWTH Aachen University and the CNRS.

Conflict of interest statement. None declared.

REFERENCES

1. Echols,H. and Goodman,M.F. (1991) Fidelity mechanisms in DNA replication. *Annu. Rev. Biochem.*, **60**, 477–511.
2. Goodman,M.F. (1997) Hydrogen bonding revisited: geometric selection as a principal determinant of DNA replication fidelity. *Proc. Natl Acad. Sci. USA*, **94**, 10493–10495.
3. Mendelman,L.V., Boosalis,M.S., Petruska,J. and Goodman,M.F. (1989) Nearest neighbor influences on DNA polymerase insertion fidelity. *J. Biol. Chem.*, **264**, 14415–14423.
4. Lee,A.M., Xiao,J. and Singleton,S.F. (2006) Origins of sequence selectivity in homologous genetic recombination: insights from rapid kinetic probing of RecA-mediated DNA strand exchange. *J. Mol. Biol.*, **360**, 343–359.
5. Gacy,A.M., Goellner,G., Juranic,N., Macura,S. and McMurray,C.T. (1995) Trinucleotide repeats that expand in human disease form hairpin structures in vitro. *Cell*, **81**, 533–540.
6. Nakayabu,M., Miwa,S., Suzuki,M., Izuta,S., Sobue,G. and Yoshida,S. (1998) Mismatched nucleotides may facilitate expansion

of trinucleotide repeats in genetic diseases. *Nucleic Acids Res.*, **26**, 1980–1984.
7. Frederico,L.A., Kunkel,T.A. and Shaw,B.R. (1993) Cytosine deamination in mismatched base pairs. *Biochemistry*, **32**, 6523–6530.
8. Modrich,P. and Lahue,R. (1996) Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.*, **65**, 101–133.
9. Eshleman,J.R. and Markowitz,S.D. (1995) Microsatellite instability in inherited and sporadic neoplasms. *Curr. Opin. Oncol.*, **7**, 83–89.
10. Buermeyer,A.B., Deschênes,S.M., Baker,S.M. and Liskay,R.M. (1999) Mammalian DNA mismatch repair. *Annu. Rev. Genet.*, **33**, 533–564.
11. Ninomiya,H., Nomura,K., Satoh,Y., Okumura,S., Nakagawa,K., Fujiwara,M., Tsuchiya,E. and Ishikawa,Y. (2006) Genetic instability in lung cancer: concurrent analysis of chromosomal, mini- and microsatellite instability and loss of heterozygosity. *Br. J. Cancer*, **94**, 1485–1491.
12. Kolodner,R.D. and Alani,E. (1994) Mismatch repair and cancer susceptibility. *Curr. Opin. Biotechnol.*, **5**, 585–594.
13. Werntges,H., Steger,G., Riesner,D. and Fritz,H.-J. (1986) Mismatches in DNA double strands: thermodynamic parameters and their correlation to repair efficiencies. *Nucleic Acids Res.*, **14**, 3773–3790.
14. Rajski,S.R., Jackson,B.A. and Barton,J.K. (2000) DNA repair : models for damage and mismatch recognition. *Mutation Res.*, **447**, 49–72.
15. Brown,J., Brown,T. and Fox,K.R. (2001) Affinity of mismatch-binding protein MutS for heteroduplexes containing different mismatches. *Biochem. J.*, **354**, 627–633.
16. Burkovics,P., Szukacsov,V., Unk,I. and Haracska,L. (2006) Human Ape2 protein has a 3'-5' exonuclease activity that acts preferentially on mismatched base pairs. *Nucleic Acids Res.*, **34**, 2508–2515.
17. Nag,N., Rao,B.J. and Krishnamoorthy,G. (2007) Altered dynamics of DNA bases adjacent to a mismatch: a cue for mismatch recognition by MutS. *J. Mol. Biol.*, **374**, 39–53.
18. Porello,S.L., Williams,S.D., Kuhn,H., Michaels,M.L. and David,S.S. (1996) Specific recognition of substrate analogs by the DNA mismatch repair enzyme MutY. *J. Am. Chem. Soc.*, **118**, 10684–10692.
19. Hart,J.R., Glebov,O., Ernst,R.J., Kirsch,I.R. and Barton,J.K. (2006) DNA mismatch-specific targeting and hypersensitivity of mismatch-repair-deficient cells to bulky rhodium(III) intercalators. *Proc. Natl Acad. Sci. USA*, **103**, 15359–15363.
20. Li,Y., Zon,G. and Wilson,W.D. (1991) Thermodynamics of DNA duplexes with adjacent G·A mismatches. *Biochemistry*, **30**, 7566–7572.
21. Peyret,N., Seneviratne,P.A., Allawi,H.T. and SantaLucia,J. Jr. (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A·A, C·C, G·G, and T·T mismatches. *Biochemistry*, **38**, 3468–3477.
22. Allawi,H.T. and SantaLucia,J. Jr. (1997) Thermodynamics and NMR of Internal G·T Mismatches in DNA. *Biochemistry*, **36**, 10581–10594.
23. Yamashita,K., Sata,S., Takamiya,H., Takagi,M. and Takenaka,S. (2001) Analysis of the complex of oligonucleotide duplexes with ligands by MALDI-TOF mass spectroscopy. *Chem. Lett.*, **30**, 680–681.
24. Zhong,M., Rashes,M.S., Marky,L.A. and Kallenbach,N.R. (1992) T-T Base mismatches enhance drug-Binding at the branch site in a 4-arm DNA junction. *Biochemistry*, **31**, 8064–8071.
25. Liu,C. and Chen,F.M. (1996) Actinomycin D binds strongly and dissociates slowly at the dGpdC site with flanking T/T mismatches. *Biochemistry*, **35**, 16346–16353.
26. Jackson,B.A. and Barton,J.K. (1997) Recognition of DNA base mismatches by a rhodium intercalator. *J. Am. Chem. Soc.*, **119**, 12986–12987.
27. Pierre,V.C., Kaiser,J.T. and Barton,J.K. (2007) Insights into finding a mismatch through the structure of a mispaired DNA bound by a rhodium intercalator. *Proc. Natl Acad. Sci. USA*, **104**, 429–434.
28. Nakatani,K., Sando,S., Kumasawa,H., Kikuchi,J. and Saito,I. (2001) Recognition of guanine-guanine mismatches by the dimeric form of 2-amino-1,8-naphthyridine. *J. Am. Chem. Soc.*, **123**, 12650–12657.
29. Nakatani,K., Sando,S. and Saito,I. (2001) Scanning of guanine-guanine mismatches in DNA by synthetic ligands using surface plasmon resonance. *Nat. Biotechnol.*, **19**, 51–55.
30. Hagihara,S., Kumasawa,H., Goto,Y., Hayashi,G., Kobori,A., Saito,I. and Nakatani,K. (2004) Detection of guanine-adenine mismatches by surface plasmon resonance sensor carrying naphthyridine-azaquinolone hybrid on the surface. *Nucleic Acids Res.*, **32**, 278–286.
31. Kobori,A., Horie,S., Suda,H., Saito,I. and Nakatani,K. (2004) The SPR sensor detecting cytosine-cytosine mismatches. *J. Am. Chem. Soc.*, **126**, 557–562.
32. Yang,X.L., Hubbard,R.B., Lee,M., Tao,Z.F., Sugiyama,H. and Wang,A.H.J. (1999) Imidazole-imidazole pair as a minor groove recognition motif for T : G mismatched base pairs. *Nucleic Acids Res.*, **27**, 4183–4190.
33. Lacy,E.R., Cox,K.K., Wilson,W.D. and Lee,M. (2002) Recognition of T·G mismatched base pairs in DNA by stacked imidazole-containing polyamides: surface plasmon resonance and circular dichroism studies. *Nucleic Acids Res.*, **30**, 1834–1841.
34. Lacy,E.R., Nguyen,B., Le,M., Cox,K.K., O'Hare,C., Hartley,J.A., Lee,M. and Wilson,W.D. (2004) Energetic basis for selective recognition of T·G mismatched base pairs in DNA by imidazole-rich polyamides. *Nucleic Acids Res.*, **32**, 2000–2007.
35. Berthet,N., Michon,J., Lhomme,J., Teulade-Fichou,M.-P., Vigneron,J.P. and Lehn,J.M. (1999) Recognition of abasic sites in DNA by a cyclobisacridine molecule. *Chem. Eur. J.*, **5**, 3625–3630.
36. Jourdan,M., Garcia,J., Lhomme,J., Teulade-Fichou,M.-P., Vigneron,J.P. and Lehn,J.M. (1999) Threading bis-intercalation of a macrocyclic bisacridine at abasic sites in DNA: nuclear magnetic resonance and molecular modeling study. *Biochemistry*, **38**, 14205–14213.
37. David,A., Bleimling,N., Beuck,C., Lehn,J.M., Weinhold,E. and Teulade-Fichou,M.P. (2003) DNA mismatch-specific base flipping by a bisacridine macrocycle. *ChemBioChem*, **4**, 1326–1331.
38. Roberts,R.J. and Cheng,X.D. (1998) Base flipping. *Annu. Rev. Biochem.*, **67**, 181–198.
39. Klimasauskas,S., Kumar,S., Roberts,R.J. and Cheng,X.D. (1994) Hhal methyltransferase flips its target base out of the DNA helix. *Cell*, **76**, 357–369.
40. Peng,T. and Nakatani,K. (2005) Binding of naphthyridine carbamate dimer to the $(CGG)_n$ repeat results in the Disruption of the G-C Base Pairing. *Angew. Chem., Int. Ed. Engl.*, **44**, 7280–7283.
41. Nakatani,K., Hagihara,S., Goto,Y., Kobori,A., Hagihara,M., Hayashi,G., Kyo,M., Nomura,M., Mishima,M. and Kojima, (2005) Small-molecule ligand induces nucleotide flipping in $(CAG)_n$ trinucleotide repeats. *Nat. Chem. Biol.*, **1**, 39–43.
42. Cordier,C., Pierre,V.C. and Barton,J.K. (2007) Insertion of a bulky rhodium complex into a DNA cytosine-cytosine mismatch: an NMR solution study. *J. Am. Chem. Soc.*, **129**, 12287–12295.
43. Kwok,P.Y. (2001) Methods for genotyping single nucleotide polymorphisms. *Annu. Rev. Genomics Hum. Genet.*, **2**, 235–258.
44. Nakatani,K. (2004) Chemistry challenges in SNP typing. *ChemBioChem*, **5**, 1623–1633.
45. Petitjean,A. and Barton,J.K. (2004) Tuning the DNA reactivity of cis-platinum: conjugation to a mismatch-specific metallointercalator. *J. Am. Chem. Soc.*, **126**, 14728–14729.
46. Hart,J.R., Johnson,M.D. and Barton,J.K. (2004) Single-nucleotide polymorphism discovery by targeted DNA photocleavage. *Proc. Natl Acad. Sci. USA*, **101**, 14040–14044.
47. Zeglis,B.M. and Barton,J.K. (2007) DNA base mismatch detection with bulky rhodium intercalators: synthesis and applications. *Nat. Protoc.*, **2**, 357–371.
48. Zeglis,B.M. and Barton,J.K. (2006) A mismatch-selective bifunctional rhodium–Oregon Green conjugate: a fluorescent probe for mismatched DNA. *J. Am. Chem. Soc.*, **128**, 5654–5655.
49. Cotton,R.G. (1989) Detection of single base changes in nucleic acids. *Biochem. J.*, **263**, 1–10.
50. Ferrari,M., Carrera,P. and Cremonesi,L. (1996) Different approaches of molecular scanning of point mutations in genetic diseases. *Pure Appl. Chem.*, **68**, 1913–1918.
51. Brown,J., Brown,T. and Fox,K.R. (2003) Cleavage of fragments containing DNA mismatches by enzymic and chemical probes. *Biochem. J.*, **371**, 697–708.
52. Bui,C.T., Rees,K., Lambrinakos,A., Bedir,A. and Cotton,R.G.H. (2002) Site-selective reactions of imperfectly matched DNA with small chemical molecules: applications in mutation detection. *Bioorg. Chem.*, **30**, 216–232.

53. Teulade-Fichou,M.-P., Vigneron,J.P. and Lehn,J.M. (1995) Molecular recognition of nucleosides and nucleotides by a water-soluble cyclo-bis-intercaland receptor based on acridine subunits. *Supramol. Chem.*, **5**, 139–147.
54. Paris,T., Vigneron,J.-P., Lehn,J.-M., Cesario,M., Guilhem,J. and Pascard,C. (1999) Molecular recognition of anionic substrates. Crystal structures of the supramolecular inclusion complexes of terephthalate and isophthalate dianions with a bis-intercaland receptor molecule. *J. Inclusion Phenom. Macrocyclic Chem.*, **33**, 191–202.
55. Amrane,S., De Cian,A., Rosu,, F. Kaiser,M., De Pauwn,E., Teulade-Fichou,M.-P. and Mergny,J.-L. (2008) Identification of trinucleotide repeat ligands with a FRET melting assay. *ChemBioChem*, **9**, 1229–1234.
56. Holz,B., Klimasauskas,S., Serva,S. and Weinhold,E. (1998) 2-Aminopurine as a fluorescent probe for DNA base flipping by methyltransferases. *Nucleic Acids Res.*, **26**, 1076–1083.
57. Goedecke,K., Pignot,M., Goody,R.S., Scheidig,A.J. and Weinhold,E. (2001) Structure of the N6-adenine DNA methyltransferase M.TaqI in complex with DNA and a cofactor analog. *Nat. Struct. Biol.*, **8**, 121–125.
58. Banoub,J.H., Newton,R.P., Esmans,E., Ewing,D.F. and Mackenzie,G. (2005) Recent developments in mass spectrometry for the characterization of nucleosides, nucleotides, oligonucleotides, and nucleic acids. *Chem. Rev.*, **105**, 1869–1915.
59. Hofstadler,S.A. and Griffey,R.H. (2001) Analysis of noncovalent complexes of DNA and RNA by mass spectrometry. *Chem. Rev.*, **101**, 377–390.
60. Rosu,F., De Pauw,E. and Gabelica,V. (2008) Electrospray mass spectrometry to study drug-nucleic acids interactions. *Biochimie.*, **90**, 1074–1097.
61. Rosu,F., Gabelica,V., Houssier,C. and De Pauw,E. (2002) Determination of affinity, stoichiometry and sequence selectivity of minor groove binder complexes with double-stranded oligodeoxynucleotides by electrospray ionization mass spectrometry. *Nucleic Acids Res.*, **30**, e82.
62. Gabelica,V., Galic,N., Rosu,F., Houssier,C. and De Pauw,E. (2003) Influence of response factors on determining equilibrium association constants of non-covalent complexes by electrospray ionization mass spectrometry. *J. Mass Spectrom.*, **38**, 491–501.
63. Sannes-Lowery,K.A., Griffey,R.H. and Hofstadler,S.A. (2000) Measuring dissociation constants of RNA and aminoglycoside antibiotics by electrospray ionization mass spectrometry. *Anal. Biochem.*, **280**, 264–271.
64. Boger,D.L., Fink,B.E. and Hedrick,M.P. (2000) Total synthesis of distamycin A and 2640 analogues: a solution-phase combinatorial approach to the discovery of new, bioactive DNA binding agents and development of a rapid, high-throughput screen for determining relative DNA binding affinity or DNA binding sequence selectivity. *J. Am. Chem. Soc.*, **122**, 6382–6394.
65. Boger,D.L., Fink,B.E., Brunette,S.R., Tse,W.C. and Hedrick,M.P. (2001) A simple, high-resolution method for establishing DNA binding affinity and sequence selectivity. *J. Am. Chem. Soc.*, **123**, 5878–5891.
66. Tse,W.C. and Boger,D.L. (2004) A fluorescent intercalator displacement assay for establishing DNA binding selectivity and affinity. *Acc. Chem. Res.*, **37**, 61–69.
67. Hernandez,L.I., Zhong,M., Courtney,S.H., Marky,L.A. and Kallenbach,N.R. (1994) Equilibrium analysis of ethidium binding to DNA containing base mismatches and branches. *Biochemistry*, **33**, 13140–13146.
68. Boger,D.L. and Tse,W.C. (2001) Thiazole orange as the fluorescent intercalator in a high resolution fid assay for determining DNA binding affinity and sequence selectivity of small molecules. *Bioorg. Med. Chem.*, **9**, 2511–2518.
69. Ward,D.C., Reich,E. and Stryer,L. (1969) Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J. Biol. Chem.*, **244**, 1228–1237.
70. Allan,B.W. and Reich,N.O. (1996) Targeted base stacking disruption by the EcoRI DNA methyltransferase. *Biochemistry*, **35**, 14757–14762.
71. Beuck,C., Singh,I., Bhattacharya,A., Hecker,W., Parmar,V.S., Seitz,O. and Weinhold,E. (2003) Polycyclic aromatic DNA-base surrogates: High-affinity binding to an adenine-specific base-flipping DNA methyltransferase. *Angew. Chem. Int. Ed.*, **42**, 3958–3960.
72. Yang,A.S., Shen,J.-C., Zingg,J.-M., Mi,S. and Jones,P.A. (1995) *Hha*I and *Hpa*II DNA methyltransferases bind DNA mismatches, methylate uracil and block DNA repair. *Nucleic Acids Res.*, **23**, 1380–1387.
73. Klimasauskas,S. and Roberts,R.J. (1995) MHhaI binds tightly to substrates containing mismatches at the target base. *Nucleic Acids Res.*, **23**, 1388–1395.
74. McGhee,J.D. and von Hippel,P.H. (1974) Theoretical aspects of DNA-protein interactions: Co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. *J. Mol. Biol.*, **86**, 469–489.
75. Malojcic,G., Piantanida,I., Marinic,M., Zinic,M., Marjanovic,M., Kralj,M., Pavelic,K. and Schneider,H.J. (2005) A novel bis-phenanthridine triamine with pH controlled binding to nucleotides and nucleic acids. *Org. Biomol. Chem.*, **3**, 4373–4381.
76. Dhaenens,M., Lehn,J.-M. and Vigneron,J.-P. (1993) Molecular recognition of nucleosides, nucleotides and anionic planar substrates by a water-soluble bis-intercaland-type receptor molecule. *J. Chem. Soc. Perkin Trans. 2*, 1379–1381.
77. Estimated with HyTher software (http://ozone2.chem.wayne.edu/).
78. Tikhomirova,A., Beletskaya,I.V. and Chalikian,T.V. (2006) Stability of DNA duplexes containing GG, CC, AA, and TT Mismatches. *Biochemistry*, **45**, 10563–10571.
79. Schärer,O.D. (2003) Chemistry and biology of DNA repair. *Angew. Chem.*, **115**, 3052–3082; *Angew. Chem. Int. Ed.*, **42**, 2946–2974.