Published online 22 February 2012 *Nucleic Acids Research, 2012, Vol. 40, No. 11* ***5115–5128***
doi:10.1093/nar/gks067

Double threading through DNA: NMR structural study of a bis-naphthalene macrocycle bound to a thymine–thymine mismatch

Muriel Jourdan[1,*], Anton Granzhan[2], Regis Guillot[3], Pascal Dumy[1] and Marie-Paule Teulade-Fichou[2,*]

[1]CNRS UMR5250, ICMG FR2607, Département de Chimie Moléculaire, Université Joseph Fourier, 570 rue de la Chimie, 38041 Grenoble Cedex 9, [2]CNRS UMR176, Institut Curie, Centre de Recherche and [3]CNRS UMR8182, Institut de Chimie Moléculaire et des Matériaux d'Orsay, Université Paris-Sud XI, 91405 Orsay, France

Received November 29, 2011; Revised January 17, 2012; Accepted January 18, 2012

ABSTRACT

The macrocyclic bis-naphthalene macrocycle (2,7-BisNP), belonging to the cyclobisintercalator family of DNA ligands, recognizes T–T mismatch sites in duplex DNA with high affinity and selectivity, as evidenced by thermal denaturation experiments and NMR titrations. The binding of this macrocycle to an 11-mer DNA oligonucleotide containing a T–T mismatch was studied using NMR spectroscopy and NMR-restrained molecular modeling. The ligand forms a single type of complex with the DNA, in which one of the naphthalene rings of the ligand occupies the place of one of the mismatched thymines, which is flipped out of the duplex. The second naphthalene unit of the ligand intercalates at the A-T base pair flanking the mismatch site, leading to encapsulation of its thymine residue via double stacking. The polyammonium linking chains of the macrocycle are located in the minor and the major grooves of the oligonucleotide and participate in the stabilization of the complex by formation of hydrogen bonds with the encapsulated thymine base and the mismatched thymine remaining inside the helix. The study highlights the uniqueness of this cyclobisintercalation binding mode and its importance for recognition of DNA lesion sites by small molecules.

INTRODUCTION

Mismatched, or non-complementary, base pairs in DNA are permanently formed in living cells due to misincorporation of nucleotides during DNA synthesis, inclusion of chemically damaged nucleotides, or inclusion of normal nucleotides opposite damaged bases in the DNA template strand (1). The rate of generation of DNA mismatches is greatly increased by exogenous factors, such as genotoxic chemicals or UV radiation. DNA mismatches disrupt the genetic information and are therefore highly toxic for the organism. As a consequence, all organisms possess enzyme systems responsible for the detection and repair of damaged DNA (DNA repair systems) (2,3). The primary task of the DNA repair system, i.e. detection and excision of damaged and mismatched bases from the DNA, is met by the DNA glycosylases that locate the damaged or 'wrong' nucleobases embedded in the massive excess of undamaged genome much like a needle in a haystack. The mechanism, by which DNA glycosylases accomplish this formidable task, extremely efficiently and rapidly, is still a matter of debate (1,4–9), but ultimately its deeper understanding will allow us to control the DNA repair processes in normal and tumor tissues.

Notably, modern cancer therapy is largely based on the agents (drugs or radiation) killing cancer cells more efficiently than normal cells, and DNA-damaging drugs represent the primary therapy in most cases. However, due to the activity of the DNA repair system, a large part of the drug-induced lesions is repaired before they cause cell death and stop proliferation. Therefore, the efficacy of the DNA damage-based cancer therapy can be increased by inhibition of such repair pathways, and DNA repair enzymes represent a topic of an immense interest in this context (10–15). Apart from a few known inhibitors of DNA repair enzymes such as APE1 (16–20), PARP (21–23) or UDG (24,25), it may be assumed that small molecules, interfering with repair enzymes by binding to

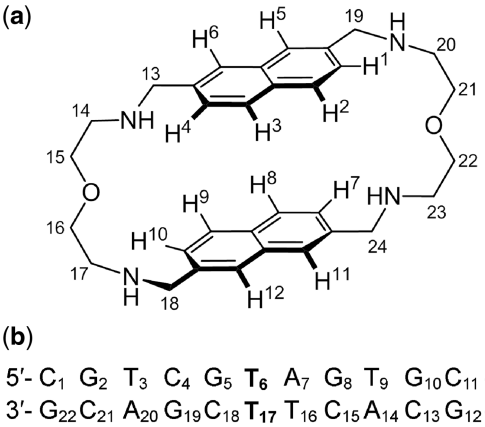
*To whom correspondence should be addressed. Tel: +33 4 56 52 08 39; Fax: +33 4 56 52 08 05; Email: muriel.jourdan@ujf-grenoble.fr
Correspondence may also be addressed to Marie-Paule Teulade-Fichou Tel: +33 1 69 86 30 86; Fax: +33 1 69 86 53 81;
Email: mp.teulade-fichou@curie.fr

© The Author(s) 2012. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

their substrate, i.e. DNA lesion sites, could lead to modulation of DNA repair pathways and enhance of the anti-cancer activity of DNA-damaging agents.

Along these lines, a few families of small molecules that selectively bind to DNA lesions, e.g. to mismatched base pairs in the DNA, have been described during the last decade. Among them, derivatives of naphthyridine have been shown to bind to G–G, G–A, A–A or C–C mismatched base pairs by insertion into the duplex and formation of hydrogen bonds with the mismatched and neighboring bases (26–30). The solution structure of a complex of naphthyridine–azaquinolone dimer with an oligonucleotide having A–A mismatch in the CAG/CAG context was characterized by NMR; it features the binding of two molecules of ligand in the mismatch site, accompanied by ejection of two flanking cytosine residues from the duplex (31). The second important class of mismatch binders is represented by bulky metal complexes, such as $[Rh(bpy)_2(chrysi)]^{3+}$ and $[Co(phen)_2(HPIP)]^{3+}$, which do not form hydrogen bonds with the nitrogenous bases, but insert into mismatched and abasic sites taking advantage of their higher accessibility, as compared to the Watson–Crick base pairs (32–37). Structures of $[Rh(bpy)_2(chrysi)]^{3+}$ bound to A–C and A–A mismatches were determined in the solid state using X-ray diffraction analysis (38,39), whereas the structure of a complex with a C–C mismatch was recently characterized in solution using NMR spectroscopy (40). Notably, in all three cases, the binding of the metal complex to the mismatch site leads to ejection of both mispaired bases out of the stack. However, intercalation of $[Rh(bpy)_2(chrysi)]^{3+}$ between two well-matched base pairs was also observed in the solid state (38). Most importantly from the therapeutic point of view, mismatch-selective metal complexes inhibit the growth of mismatch-repair-deficient cancer cell lines, and not that of mismatch repair-proficient ones, which supports the hypothesis that recognition of base mismatches by small molecules and interference with the DNA repair enzymes is possible in cells as well (41,42).

In another approach, we have shown that certain macrocyclic ligands containing two aromatic planes connected by two flexible polyamine chains (also termed cyclobisintercalators, or CBI) (43) can selectively bind to DNA lesions such as abasic sites (44,45) and homopyrimidine mismatches (46,47). The initial studies were performed with the 2,7–substituted bis-acridine (BisA) and 2,6-substitued bis-naphthalene (2,6-BisNP) macrocycles. Later, however, screening of a panel of 20 homo- and heterodimeric macrocyclic ligands led to the identification of the 2,7–substituted bis-naphthalene isomer (2,7-BisNP, Scheme 1) as the ligand with the highest binding selectivity for mismatched versus fully matched DNA duplexes (48). The selective binding of this family of ligands to DNA lesion sites is, presumably, due to their peculiar constrained conformation and the short distance between the two aromatic planes, which does not allow their simultaneous intercalation into regular DNA, as it would violate the neighbor-exclusion principle (49). However, it appears that at sites of weaker thermodynamic stability, such as mismatches and abasic



Scheme 1. (a) Chemical structure and proton numbering scheme of 2,7-BisNP. (b) Sequence and numbering of the TT-DNA duplex. The T-T mismatch is highlighted in boldface.

sites, this principle can be violated, allowing the highly efficient binding of cyclobisintercalators. In fact, the binding of the cyclobisacridine BisA to a DNA duplex with an abasic site has been characterized by a high-resolution NMR and molecular modeling study, which showed that one of the acridine units of BisA occupies the abasic pocket, while the second one intercalates at the neighboring base pair step, in a threading bisintercalation fashion (45). Thus, we speculated that binding of 2,7-BisNP to T–T mismatch may proceed in a similar fashion. In addition, it was postulated that one of the mismatched thymines may be extruded (flipped out) from the helix, as this would be in agreement with their higher susceptibility to permanganate oxidation observed upon binding of CBI macrocycles (46). However, the structural details of the recognition event have not been available, so far.

In this study, we report the NMR characterization and molecular modeling of 2,7-BisNP binding to an 11-mer oligonucleotide duplex containing a thymine–thymine mismatch target (TT-DNA, Scheme 1). The central part of TT-DNA (5′-TCG**T**AGT-3′/5′-ACT**T**CGA-3′) corresponds to that of 17-mer duplex, used in our previous studies (46,46); the terminal GC base pairs were added to increase the thermal stability of the short duplex. Moreover, the asymmetry of the binding site with respect to the neighboring base pairs (G–C versus A–T) allowed us to hope that one type of ligand–DNA complex would be present exclusively, unlike in the case of BisA–abasic site recognition, where the symmetry of the binding site (two flanking G–C pairs) gave rise to two complexes (45). In addition, the solid-state conformation of 2,7-BisNP alone was obtained from X-ray diffraction analysis and served for comparison with the one observed upon binding to the DNA mismatch in solution.

MATERIALS AND METHODS

Solid-state structure of 2,7-BisNP

The synthesis and characterization data of 2,7-BisNP × 4 HCl were reported previously (48). Crystals suitable

for X-ray diffraction analysis were obtained by vapor diffusion of Et_2O or *t*-BuOMe into solutions of 2,7-BisNP × 4 HCl in MeOH. A satisfactory dataset was obtained with a crystal grown form a MeOH–Et_2O solvent pair. Diffraction data were collected by using a Bruker Kappa X8 APEX II diffractometer with graphite-monochromated Mo $K\alpha$ radiation ($\lambda = 0.71073$ Å) at 100 ± 1 K. The structure was solved by direct methods using SHELXS-97 and refined by full-matrix least-squares techniques using SHELXL–97 with anisotropic displacement parameters for all non-hydrogen atoms. X-ray data for 2,7-BisNP × 4 HCl × 4 MeOH (CCDC-847671): $C_{36}H_{60}Cl_4N_4O_6$, $M = 786.68\ g\,mol^{-1}$, space group $P2_1/c$ (no. 14), $a = 32.786\,6(9)$, $b = 5.524\,80\,(10)$, $c = 23.056\,6(6)$ Å, $\alpha = \gamma = 90°$, $\beta = 105.943\,0(10)°$, $V = 4015.81(17)\ Å^3$, $Z = 4$, $\mu = 0.342\ mm^{-1}$, $F(000) = 1680$, crystal size $0.20 \times 0.11 \times 0.06\ mm^3$. A total of 98 941 reflections ($1.29° \leq \theta \leq 33.30°$) were collected, of which 15 469 were unique ($R_{int} = 0.0470$) and 10 758 observed [$I > 2\sigma(I)$] were used in calculations. The structure was refined on F^2 to $R = 0.0820$ [$I > 2\sigma(I)$] for 483 refined parameters and 6 restraints. Largest difference peak and hole 1.741 and −1.264 e $Å^{-3}$.

NMR studies

Sample preparation. The oligonucleotides (purified by HPLC and ion exchange chromatography) were purchased from Eurogentec (Belgium). The two strands d(CGTCGTAGTGC) and d(GCACTTCGACG) were dissolved and mixed in 500 µl of 10 mM sodium phosphate buffer (pH 6) containing 50 mM NaCl, heated to 80°C and slowly cooled to room temperature to form the duplex. The sample was lyophilized once from water, twice from 99.9% D_2O and then redissolved in 500 µl of 99.99% D_2O. For NMR experiments involving exchangeable protons, the oligomers were dissolved in 90% H_2O/10% D_2O (v/v). The final duplex concentration was about 1.5 mM. For the 2,7-BisNP–TT-DNA complex, the ligand was dissolved in the NMR buffer at 19 mM and added progressively to the free DNA in the NMR tube until a 1:1 complex was obtained, as confirmed by 1D NMR spectroscopy.

NMR experiments. ^{1}H NMR experiments were carried out with a Varian Unity Plus 500 MHz spectrometer or a Varian Unity 800 MHz spectrometer equipped with a cryo-probe. Chemical shifts were calibrated relative to sodium 3–(trimethylsilyl)propionate-2,2,3,3-d_4 (TSP-d_4). Data were processed with VNMR software (Varian) and displayed and analyzed with the CCPNMR program (50).

The NOESY spectra of the free DNA and of the ligand–TT-DNA complex were recorded in D_2O buffer at 10°C using mixing times of 75, 100, 150 and 250 ms. The TOCSY spectra were recorded with a spin-lock time of 40 and 60 ms using a MLEV17 pulse train of 8 kHz for the free DNA and with a spin-lock time of 60 or 110 ms for the complex. The experiments were recorded with 512 (t_1) × 2048 (t_2) complex data points. Ninety-six scans were acquired over a spectral width of 5000 Hz (for experiments at 500 MHz) or 64 scans were acquired over a spectral width of 8000 Hz (for experiments at 800 MHz) in both dimensions. Water resonance was suppressed by a standard presaturation during the relaxation delay and during the mixing time for NOESY. The DQFCOSY spectrum was run with 2048 points in t_2, 256 points in t_1, and 96 scans for each t_1 value. For a sample in 90% H_2O, NOESY spectra were acquired at 10 or 5°C using the sculpting gradients scheme with selective pulses of 3–4 ms for water suppression. The mixing time was set to 350 ms. The experiments were recorded with 340 (t_1) × 4092 (t_2) complex data points and 160 scans for each FID. 1D proton-decoupled ^{31}P spectra and 2D ^{1}H-^{31}P HSQC were recorded at 10°C at 202 MHz; chemical shifts were referenced to trimethyl phosphate.

For the free 2,7-BisNP, 1D NMR spectra, 2D NOESY with mixing time of 0.8 s and DQFCOSY were performed in a D_2O buffer containing 10 mM sodium phosphate (pH 6) and 50 mM NaCl at 10°C.

Distance restraints and dihedral angles. The NOE intensities measured from the NOESY spectra in D_2O (mixing time of 150 ms) were converted into distance restraints in the CCPNMR program and calibrated to the fixed reference cytosine H5–H6 = 2.45 Å. The lower and upper boundaries were set to 0.8 and 1.2 multiples of calculated distances, respectively. For dihedral angles restraints, a qualitative analysis of the DQFCOSY cross-peaks intensities was used to evaluate the approximate values of the $J_{H1'-H2'}$, $J_{H2''-H3'}$ and $J_{H3'-H4'}$ coupling constants and deduce the pseudorotation angles *P*. The *P*-values were then used in the Antechamber module (AMBER 10.1 software, University of California, San Francisco) to generate the sugar pucker restraints. For the ligand–DNA complex, intermolecular NOEs between the ligand and DNA were measured and classified semi-quantitatively as distances into three categories: strong (2.0–3.0 Å), medium (3.0–4.0 Å) and weak (4.0–5.0 Å).

Molecular modeling

The B-DNA duplex was built using the Biopolymer module of Biosym Insight II software, saved as a pdb file and exported into xLEaP module of the AMBER program. Sodium counterions were added. Structure calculations of the free DNA and of the ligand–TT-DNA complex were performed with the SANDER module of AMBER10.1 using the ff99SB force field. A 2 fs time step was used for all the calculations, and the SHAKE algorithm was applied to remove bond stretching. The tetraprotonated 2,7-BisNP ligand was constructed in Insight II and minimized with the AMBER force field, which lead to a U-shaped conformation with a near-parallel orientation of naphthalene units. The partial charges were calculated using HF/6-31G* (Gaussian 03 Rev. A.1, Gaussian, Inc., Pittsburgh, PA, USA) and the RESP module in AMBER. A new AMBER library containing the ligand parameters was then created.

Structure of the free TT-DNA. Five random starting structures were prepared using 3–5 ps unrestrained

molecular dynamics at elevated temperature ranging from 600 to 1000 K. Each of these structures was subjected to 100 ps restrained simulated annealing protocol with implicit solvent (Born solvation model) using 232 experimental interproton distances, 100 sugar dihedral angles, 91 backbone torsion angles which maintained the right-handed character of the helix and 27 angles which preserved base pairing as restraints. The temperature was rapidly increased to 400 K and maintained over the first 20 ps. The system was then gradually cooled from 400 to 100 K for 70 ps and finally to 0 K for the last 10 ps. The TAUPT parameter was set to 0.4 ps for the first 20 ps, and then during cooling it was increased to 4.0 ps. From 90 to 95 ps, TAUPT was decreased to 1 ps and to 0.1 ps for the last 5 ps. The weight of NMR restraints (REST parameter) increases from 0.1 to 1 during the first 10 ps (3.2 to 32 kcal mol^{-1} Å^{-2} for distance restraints and 50 kcal mol^{-1} rad^{-2} for torsion angle restraints) and kept at this value until the end of the simulation. The five final structures were averaged and solvated in a periodic TIP3 water box that extended to a distance of 5 Å from any solute atom (approximate dimensions $35 \times 35 \times 50$ Å, 1766 water molecules). One thousand steps of minimization of the solvent and counterions only were applied; the whole system was then subjected to 5000 steps of minimization. A 20 ps NVT dynamics were then performed at 300 K on the water only (with the DNA and sodium ions weakly constrained) before 100 ps of NPT dynamics at 300 K were performed on the whole system without NMR restraints, and then 100 ps of NPT dynamics at 300 K with NMR restraints. The structures over the last 5 ps were analyzed. The structures have been deposited to the Protein Data Bank (code 2LL9).

Structure of the 2,7-BisNP–TT-DNA complex. The structure of the free TT-DNA obtained from the above protocol was modified in the Biopolymer module of Biosym Insight II software to create an intercalation site at the A_7pG_8 and $C_{15}pT_{16}$ steps. The backbone angles of the $T_6pA_7pG_8$ and $C_{15}pT_{16}pT_{17}$ steps were manually changed, and T_6 and T_{17} were pushed toward the major groove. 2,7-BisNP was manually docked into the DNA. The complex was solvated with a new periodic TIP3 water box extended to a distance of 4.5 Å from any solute atom (approximate dimensions $40 \times 40 \times 55$ Å, 2400 water molecules). One thousand steps of minimization of the solvent and counterions only were applied, and then the whole system was subjected to 5000 steps of minimization using NMR restraints with a force constant of 32 kcal mol^{-1} Å^{-2}. Then 20 ps of equilibration followed by 40 ps of simulated annealing molecular dynamics were performed. No force constants were introduced to ensure the A_7–T_{16} base pairing. The final structure has been deposited to the Protein data bank (code 2LLJ).

DNA melting experiments

Thermal denaturation studies were performed in triplicate according to the previously publishing protocol in 10 mM sodium cacodylate buffer containing 50 mM NaCl, pH 6 (48). The concentration of duplex TT-DNA or TA-DNA (5′-CGT CGT AGT GC-3′/5-GCA CTA CGA CG-3′) was 2 μM, and the concentration of ligand was varied from 0 to 8 μM.

RESULTS

Structure of 2,7-BisNP in the solid state

The preliminary X-ray diffraction data obtained with several crystals grown form MeOH–Et_2O or MeOH–*t*-BuOMe solvent pairs indicated the same space group and unit cell volume, which provided evidence of the identical structure. A detailed analysis was performed with a crystal of 2,7-BisNP × 4 HCl × 4 MeOH grown from MeOH–Et_2O. Two crystallographically non-equivalent cations of the macrocycle with similar geometric parameters were observed in the asymmetric unit. The structure (Figure 1a and b) reveals the non-overlapping, anti-parallel orientation of the naphthalene units within each cation. The planes of the aromatic units are separated by a distance of 6.6–6.7 Å. Two molecules of MeOH are inscribed in the macrocycle, hold by the hydrogen bonds between the OH groups, chloride ions, and the protonated amino groups of the side chains of 2,7-BisNP, as well as by van der Waals contacts between the methyl groups and naphthalene units. Two other solvent molecules and chloride ions reside in the intermolecular cavities (Figure 1c). Importantly, the naphthalene units do not participate in intra- or intermolecular π-stacking interactions. The structure is hold together by a network of hydrogen bonds between chloride anions, methanol molecules and protonated amino groups of the side chains.

Although the anti-parallel conformation of 2,7-BisNP in the solid state does not feature a confined space between the naphthalene units, it may be expected that in solution the macrocycle is more dynamic giving rise to numerous possible conformers. Indeed, a molecular dynamics study of 2,7-BisNP in a water box identified several conformers differing mainly by a more or less open orientation of their naphthalene units (48). In addition, in a previous X-ray structural study, the *amphi* isomer (2,6-BisNP) was found to adopt a semi-closed conformation with a parallel orientation of the two aromatic units upon binding of an aromatic dicarboxylate guest (51). It is therefore clear that the conformation of the macrocycle, free or in a complex, is influenced both by the attachment position of the two linkers on the aromatic moieties (48) and by the nature of the guest trapped inside the macrocyclic scaffold, which points out to the ability of this class of macrocycles to adapt to a large variety of guests.

DNA melting experiments

To confirm the selective binding of 2,7-BisNP to the 11-mer TT-DNA and get a rough idea of the binding stoichiometry and affinity, thermal denaturation experiments were performed at variable ligand-to-duplex ratios. In conditions comparable to those of the NMR experiments ([Na^+] = 60 mM, pH 6), the duplex TT-DNA alone denatured at T_m = 33.0°C. The addition of 2,7-BisNP

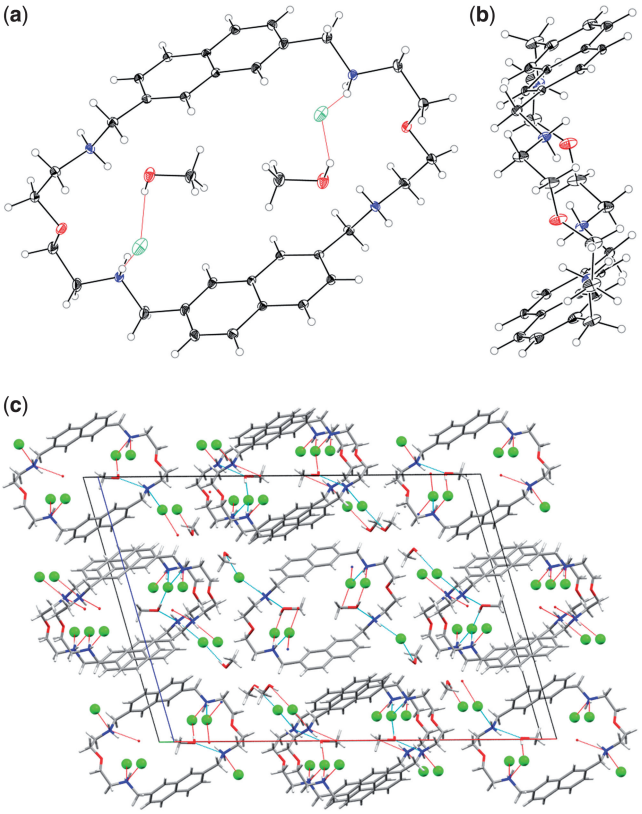


Figure 1. Structure of 2,7-BisNP × 4 HCl in the solid state. (**a**) ORTEP view along the crystallographic *b* axis of 2,7-BisNP with the enclosed solvent molecules and anions. The thermal ellipsoids of the non-hydrogen atoms are shown with 30% probability. The oxygen atoms of the side chains occupy two alternate positions; the ones with the higher occupancy factor (0.83) are shown. (**b**) ORTEP view along the crystallographic *a* axis without solvent molecules and anions. (**c**) Crystal packing (CPK coloring scheme).

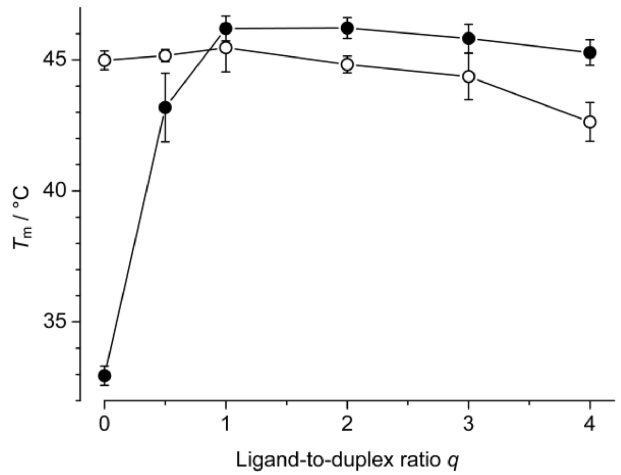


Figure 2. Melting temperatures of TT-DNA (filled circles) and TA-DNA (empty cirlces) in the absence and in the presence of 2,7-BisNP (0–4 equivalents). Conditions: c_{DNA} = 2 μM in cacodylate buffer, pH 6.0, [Na^+] = 60 mM.

led to significant stabilization of TT-DNA, as demonstrated by large increase of the melting temperature (Figure 2). The maximal effect was observed at a 1:1 ligand-to-DNA ratio (ΔT_m = 13.3°C), whereas at higher ligand loadings, the stabilization effect leveled off and slightly decreased to 12.3°C. This minor decrease is attributable to weak binding of the excess ligand to the single-stranded form, which may slightly counterbalance the strong stabilizing effect induced by duplex binding. At a ligand-to-DNA of 0.5:1, a broad melting curve was observed, indicative of incomplete saturation of the binding site (52).

To confirm the selective binding of 2,7-BisNP to mismatched DNA, the same experiment was performed with the fully paired duplex TA-DNA. The denaturation profile of TA-DNA was hardly influenced by the presence of 2,7-BisNP (Figure 2), and only at high ligand loadings, a decrease in melting temperature was observed (ΔT_m = −2.4°C at a 4:1 ligand-to-DNA ratio), indicative of weak binding to the single DNA strands as described for analogous macrocycles (53,54). Interestingly, melting temperature of the free duplex TA-DNA (T_m = 45.0°C) is very close, within the experimental error, to the melting temperature of TT-DNA in the presence of one molar equivalent of ligand (46.2°C), indicating that binding of 2,7-BisNP leads to stabilization of the T–T-mismatched duplex to the level of a regular T–A duplex. Altogether, thermal denaturation experiments speak for the preferential binding of the macrocycle at the T–T base pair site.

Structure of TT-DNA in the absence of ligand

Proton and ^{31}P assignments. The base and sugar protons of TT-DNA were identified on the basis of analysis of 2D NOESY, TOCSY and DQFCOSY spectra using the sequential assignment strategy, well described for right-handed DNA duplexes (55). All cross-peak intensities were consistent with a B-like DNA conformation and the bases adopting an *anti*-conformation. Qualitative analysis of the proton–proton coupling constants in the DQFCOSY spectra (no detectable H3′–H4′ and H2″–H3′ cross-peaks corresponding to very small $J_{H3'-H4'}$ and $J_{H2''-H3'}$ coupling constants, but characteristic H1′–H2′ pattern) along with intensities in the NOESY spectra indicate that all the sugars, except C_1, C_{15} and G_{22}, are in a C2′-*endo* conformation. For these sugars, a *P*-value in the range 162–180° was used. C_1, C_{15} and G_{22} showed large $J_{H3'-H4'}$ coupling constant and small but detectable $J_{H2''-H3'}$, corresponding to a pseudorotation phase angle *P* ranging from 90 to 120°. For the others sugars, a *P*-value in the range 162–180° was used.

Around the mismatched base pair T_6–T_{17}, the bases are well-stacked and both T_6 and T_{17} remain inside the helix, as confirmed by the sequential NOE connectivities observed between T_6H6 and G_5H1′, G_5H2′ and G_5H2″, A_7H8 and T_6H1′, T_6H2′ and T_6H2″, T_{17}H6 and T_{16}H1′, T_{16}H2′ and T_{16}H2″, C_{18}H6 and T_{17}H1′, T_{17}H2′ and T_{17}H2″, as well as characteristic H8/H6 cross-peaks (G_5H8/T_6H6, T_6H6/A_7H8, T_{16}H6/T_{17}H6, T_{17}H6/C_{18}H6).

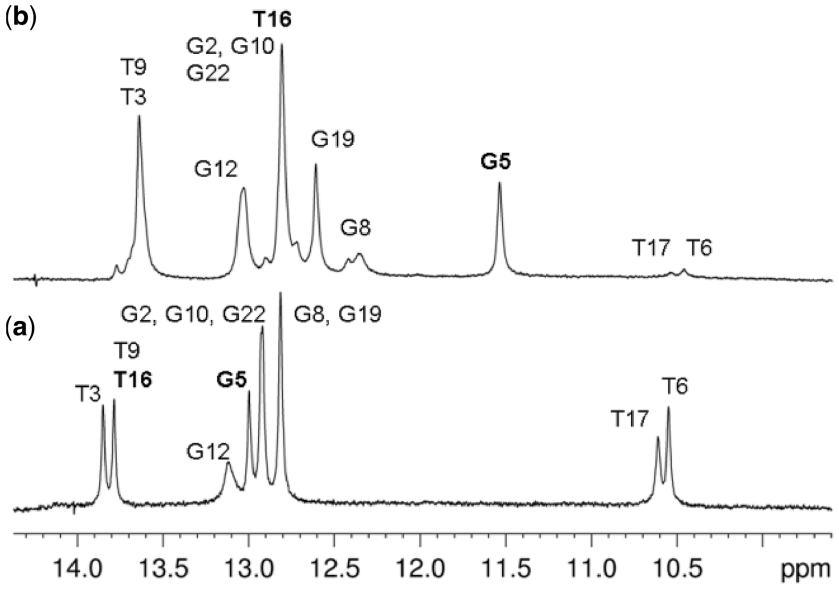


Figure 3. 1D imino spectra of (**a**) free TT-DNA at 10°C and (**b**) 2,7-BisNP–TT-DNA complex at 5°C. T_{16} and G_5, which undergo most significant shifts, are highlighted in bold.

The exchangeable imino protons were assigned with the 1D and 2D NOESY spectrum recorded in 90% H_2O at 10 or 5°C. The 1D spectrum shows three main sets of imino proton: thymine imino protons in the range 13.5–14.0 ppm, guanine imino proton in the range 12.8–13.1 ppm and two resonances at 10.56 ppm and 10.61 ppm, corresponding to the mismatched thymines (Figure 3a). It is known that non-hydrogen-bound imino protons, such as those of bulged bases (56), bases facing abasic sites (57), or non-Watson–Crick base pairs like base mismatches (58–62) or wobble base pairs (63,64) resonate in this range. These resonances were assigned to the imino protons of T_6 and T_{17}, respectively, based on the cross-peaks observed in the 2D spectra. However, it was not possible to deduce the hydrogen bonding pattern between T_6 and T_{17}, as the imino protons of these bases were almost coincident. Nevertheless, we can conclude that the T–T mismatch does not introduce dramatic distortions to the duplex structure, as chemical shifts of the DNA protons are in the common range observed for regular DNA.

^{31}P chemical shifts were used to obtain information about the conformation of the phosphodiester backbone, as they are sensitive to the local distortions. The 1D ^{31}P spectrum was recorded at 10°C (Supplementary Figure S1), and the signals were assigned from a ^{1}H-^{31}P HSQC experiment (Supplementary Table S1). The ^{31}P chemical shifts were observed in the range from −3.72 to −5.14 ppm, clearly indicating that all torsion angles are in the BI conformation (65).

Structural modeling of the free TT-DNA. The structural analysis of the free TT-DNA was carried out using a simulated annealing protocol within the AMBER 10.1 program. A set of five random conformations was generated to ensure that each of the starting structures was in different regions of the conformational space and that the final conformation is independent of the initial structure. A 100 ps restrained dynamics in implicit water were then performed followed by 100 ps of restrained dynamics in explicit solvent (see Materials and Methods). An ensemble of the structures obtained over the last 5 ps of the dynamics is shown in Figure 4. The DNA adopts a regular B-DNA-like structure with well-formed Watson–Crick alignments of all canonical base pairs. The bases T_6 and T_{17} remain inside the helix, and the major and minor groove size are normal around the mismatch. Thus, the T–T mismatch does not introduce dramatic distortion to the structure, in accordance with what has been observed in earlier NMR studies (58–62). A good agreement with the NMR experimental restraints was obtained: 13 restraints were violated by 0.2–0.3Å, 14 by 0.3–0.5 Å and four by more than 0.5 Å. The mutual root mean square deviation (rmsd on all atoms except the terminal base pairs) between the last five structures from the dynamic are in the range 0.32–0.39 Å, evidencing no major fluctuations between the structures and a good convergence. This successful convergence to similar conformation is indicative of the performance of the restrained molecular dynamics. Most significant differences are observed at the terminal base pairs and can probably be attributed to end fraying.

Structure of the 2,7-BisNP–TT-DNA complex

The titration of 2,7-BisNP to TT-DNA was monitored by 1D ^{1}H NMR spectroscopy (Supplementary Figure S2). Upon the addition of the ligand new resonances appeared and increased while the resonances of the free oligonucleotide decreased, disappearing completely when a 1:1 ratio was reached. This indicates a slow exchange regime on NMR timescale. Line broadening was also observed, and some of the aromatic protons of the ligand were significantly shifted upfield. These

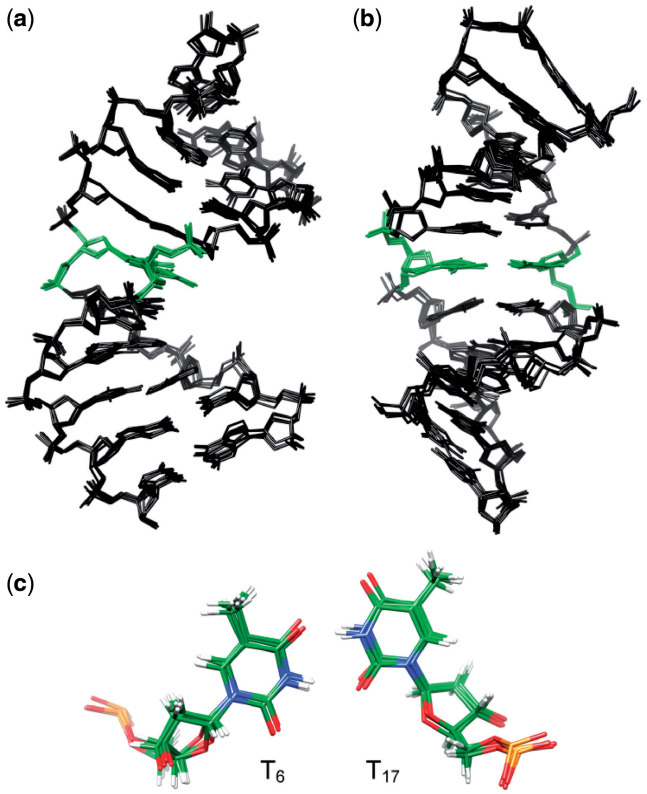


Figure 4. Structure of TT-DNA: superposition of the five last structures emergent from the restraint molecular dynamics viewed (**a**) from the side and (**b**) into the major groove. The mismatched thymine residues T_6 and T_{17} are depicted in green, hydrogen atoms are not shown. (**c**) Close-up down view of the T_6–T_{17} mismatch (O: red, N: blue, H: white, P: orange).

observations are consistent with a specific binding and insertion of the ligand inside the DNA.

Proton and ^{31}P assignments of the 2,7-BisNP–TT-DNA complex. Most of the non-exchangeable protons of the oligonucleotide were assigned using a combination of 2D NOESY, TOCSY and DQFCOSY. In the NOESY, the observed sequential internucleotide connectivities H6/8 and H1′ as well as H6/8 and H2′/H2″ allowed the assignment of most of the protons. However, this sequential walk was interrupted: no NOEs were observed between G_8H8 and A_7-sugar protons and between T_{16}H6 and C_{15}-sugar protons, which gives a clear evidence of intercalation of an aromatic unit between the base pairs A_7–T_{16} and G_8-C_{15}, as these NOEs were clearly present in the free oligonucleotide. Adenine A_7 was difficult to identify, as its signal was not observed in the typical range of chemical shifts normally observed for adenine (downfield), but was significantly shifted upfield (δ = 7.88 ppm) and overlapped with five H8 or H6 base protons. This also undoubtedly points out to binding of the ligand in the neighborhood of A_7.

Weak NOE signals were observed between T_6H6 and G_5-sugar protons, but not between T_{17}H6 and T_{16}-sugar protons. Similarly, no NOE were present between C_{18}H8 and T_{17}-sugar protons. Unfortunately, it was not possible to unambiguously determine if A_7H8 gave NOE with T_6-sugar proton, as the peak of A_7H8 was strongly overlapped with other protons. These observations suggest ejection of the mismatched thymine T_{17} from the double helix, with the opposite T_6 remaining closer to its flanking bases G_5 and A_7 than T_{17} does to C_{18} and T_{16}. The intranucleotide NOEs indicate that even after binding of 2,7-BisNP to DNA, the bases maintain their *anti*-conformation. The DQFCOSY cross peak pattern suggests that the sugars mainly adopt the C2′-*endo* puckering, even the thymines T_6 and T_{17}, as it was observed for the oligonucleotide in the absence of the ligand.

The imino and amino protons were assigned on the basis of the 2D-NOESY spectrum of the complex recorded in H_2O at 10 or 5°C. Usual correlations for Watson–Crick base pair were observed between G-NH imino protons and C-NH_2 amino protons and between T-NH and A-H2 protons. Imino–imino NOEs between flanking base pairs were also detected throughout the duplex. G_5NH and T_{16}NH were markedly shifted upfield (1.5 ppm and 1.0 ppm, respectively), which gave evidence for intercalation of 2,7-BisNP at the G_5pT_6 and $C_{15}pT_{16}$ steps. The T_6NH and T_{17}NH were easily detected in the 1D spectrum as broad signals at 10.49 and 10.56 ppm (Figure 3b), but could not be unambiguously assigned even in the 2D spectra due to the absence of correlation peaks. This is consistent with the fact that they are extruded from the duplex and more exposed to the solvent.

The protons of 2,7-BisNP were assigned following the analysis of the NOESY, TOCSY and DQFCOSY datasets in D_2O solution. The aromatic region of the DQFCOSY spectrum displayed four sets of vicinal coupling connectivities corresponding to H1/H2, H3/H4 and H7/H8, H9/H10. The protons H1, H4, H7 and H10 were identified on the basis of their long-range coupling constant with H5, H6, H11 and H12, respectively, and their NOE with the methylene proton of side chains (H19, H13, H24 and H18). The other methylene protons of the linker chains were identified by combined use of TOCSY and NOESY cross-peak. Unfortunately, unambiguous assignment of all the protons of the side chains was not possible due to broad signals and severe overlaps. Upon binding to the DNA, the methylene protons become non-equivalent and have different, but very similar, chemicals shifts leading to numerous overlapping signals which were difficult to assign.

Finally, the ^{31}P spectra of the 2,7-BisNP–TT-DNA complex (Supplementary Figure S1) show similar chemical shift dispersion compared to the free TT-DNA, indicating no major backbone deformation upon binding of the ligand.

Chemical shift mapping upon binding of 2,7-BisNP to TT-DNA. Changes in the chemical shift upon binding represent a valuable tool for the localization of the binding site of the ligand. The interaction of 2,7-BisNP with the TT-DNA duplex resulted in line broadening and significant variation of chemical shifts relative to the free DNA (Figure 5). As can be seen, the central part of the sequence $T_6pA_7pG_8$ previously identified as the possible binding site undergoes the most pronounced changes. Indeed, most of

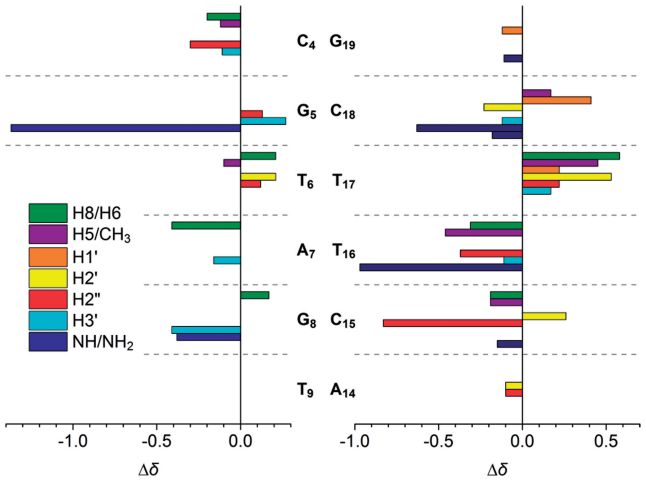


Figure 5. Changes in ^{1}H chemical shifts of TT-DNA ($\Delta\delta = \delta_{bound} - \delta_{free}$) upon binding of 2,7-BisNP. Only significant changes ($|\Delta\delta| \geq 0.10$ ppm) are shown.

the protons of the base pairs, which flank the T_6–T_{17} mismatch are shifted upfield due to the intercalation of the ligand and the associated π-stacking effect. For example, A_7H8 is shifted upfield by 0.41 ppm, T_{16}H6 is shifted by 0.31 ppm, T_6CH3 is shifted by 0.46 ppm and almost all the protons of C_{15} are shifted upfield. The most striking effect is observed for the exchangeable protons: G_5NH is shifted by −1.37 ppm, T_{16}NH by −0.97 ppm, G_8NH by −0.38 ppm and $C_{18}NH_2$ protons are shifted by −0.63 and −0.18 ppm. This provides a strong evidence of the insertion of the ligand into DNA. The large shift experienced by G_5NH and, to a lesser extent, by $C_{18}NH_2$ together with the shift of the non-exchangeable protons of A_7–T_{16} pair clearly reveals the position of one of the aromatic ring of 2,7-BisNP inside the helix in the plane of T_6–T_{17}. Indeed, such shift would be inconsistent with an intercalation occurring between the base pairs T_6–T_{17} and A_7–T_{16}.

The signals of some protons were, instead, shifted downfield. Notably, most protons of the mismatched thymines T_6 and especially T_{17} exhibited higher chemical shifts compared with the free TT-DNA, providing evidence of their ejection from the base stack. No significant changes ($|\Delta\delta| \geq 0.10$ ppm) were detected for the base pairs other than C_4–G_{19}, G_5–C_{18}, T_6–T_{17}, A_7–T_{16}, G_8–C_{15} and T_9–A_{14}, which speaks for the localization of the binding site in the central part of the duplex. The signals of the protons of 2,7-BisNP also undergo significant upfield shifts ($\Delta\delta$ up to −1.36 ppm; Table 1) which are characteristic of its intercalation into the DNA.

(3) Intermolecular NOEs between 2,7-BisNP and TT-DNA. The assignment of the ligand and DNA protons provided the basis for identification of the ligand–DNA contacts, which gave insight into the mode of binding of 2,7-BisNP. Twenty-three specific intermolecular ligand–TT-DNA contacts (Table 1) were detected in the 2D NOESY spectrum (Figure 6), which allowed, in combination with the observations described previously, to identify more precisely the location of the binding site and the binding mode. Numerous NOE contacts were observed between the two aromatics cores of 2,7-BisNP and the T_{16} base, consistent with the 'sandwiching' of T_{16} between the aromatic planes of the ligand (Figure 7). Some NOEs were also detected between one naphthalene ring (protons H1 and H2) and the sugar protons of T_{17}. NOEs between A_7H1′and the side-chain protons CH_2-24 indicate that this chain lies in the minor groove. On the contrary, the NOE observed with the other side chain, namely between CH_2-18 and $T_{16}CH_3$, requires the second chain to be placed in the major groove. This is only possible if the binding of 2,7-BisNP occurs via a threading intercalation mode, with the linking chains occupying both the minor and the major grooves. The NOE cross-peaks between H5 and H6 of the ligand and the imino proton of the guanine (G_5NH1) as well as between H5 and the cytosine (C_{18}H5) flanking the mismatch site give another evidence of insertion of the ligand at the level of T_6–T_{17} base pair.

Structure of the 2,7-BisNP–TT-DNA complex from restrained molecular dynamics calculations. The 23 intermolecular distance restraints determined from the NOESY spectra were used for molecular dynamics calculation in order to obtain a model of 2,7-BisNP binding to TT-DNA that would agree with NMR data. The structure obtained after 40 ps of dynamics in explicit water is displayed in Figures 7 and 8. It shows the threading of the macrocyclic ligand through TT-DNA and the specific binding around the central thymine–thymine mismatch site. One of the naphthalene rings of 2,7-BisNP faces T_6 and occupies the place of T_{17}, which is extruded toward the major groove. This naphthalene is almost perpendicular to the main axis of A_7–T_{16} base pair, which allows a good stacking with T_{16}, but not with A_7 (Figure 8c). The T_6 remains stacked with the flanking G_5 residue. The sandwiched base pair A_7–T_{16} forms a buckle of around 38° to allow the ligand to snugly fit inside the helix. The second naphthalene ring of 2,7-BisNP also stacks with T_{16}, leading to the 'sandwiching' of the latter in the cavity delineated by the two naphthalene units. The side chains of 2,7-BisNP are located in the middle of each groove. Stabilization of the complex occurs, in addition to the aforementioned π–stacking interactions, due to formation of the hydrogen bonds (Figure 8d) between the protonated amino groups of the ligand and carbonyl groups of thymines T_6 and T_{16} (C13-NH_2^+ and T_6O4, C19-NH_2^+ and T_{16}O2, C24-NH_2^+ and T_{16}O2). In particular, the hydrogen bond with T_6 may explain the specificity of the ligand for T-X mismatch, while the two hydrogen bonds with the 'encapsulated' thymine T_{16} largely contribute to the stabilization of the complex. In addition, four hydrogen bonds between the positively charged amino groups and water molecules were detected (66). One of the water molecules mediates the bonding between C18–NH_2^+ and T_{16}O4, while the other one mediates the interaction between C19–NH2 and two phosphate groups ($pT_{16}p$).

Table 1. [1]H chemical shifts (ppm) of 2,7-BisNP in the free form and in the 2,7-BisNP–TT-DNA complex and observed intermolecular NOEs[a]

Proton[b]	δ_{free}	δ_{bound}	$\Delta\delta$[c]	2,7-BisNP–TT-DNA NOEs
H1	7.43	6.86	−0.57	$T_{16}H1'$ (m), $T_{16}CH_3$ (w), $T_{17}H1'$ (m), $T_{17}H2'$ (m), $T_{17}H2''$ (m)
H2	7.81	6.72	−1.09	$T_{16}H1'$ (w), $T_{16}CH_3$ (m), $T_{17}H1'$ (s), $T_{17}H2'$ (s), $T_{17}H2''$ (m)
H3	7.81	7.15	−0.66	
H4	7.43	7.25	−0.18	
H5	7.78	6.65	−1.19	G_5NH1 (m), $C_{18}H5$ (w)
H6	7.78	6.59	−1.13	G_5NH1 (m)
H7	7.43	6.59	−0.84	$T_{16}H1'$ (s), $T_{16}H2'$ (s), $T_{16}H2''$ (s), $T_{16}CH_3$ (m)
H8	7.81	6.67	−1.14	$T_{16}H1'$ (m), $T_{16}H2'$ (m), $T_{16}H2''$ (m), $T_{16}CH_3$ (m)
H9	7.81	7.15	−0.66	
H10	7.43	7.12	−0.31	
H11	7.78	6.86	−0.92	
H12	7.78	6.42	−1.36	
CH_2-13	4.22	4.07/4.16	−0.15/−0.06	
CH_2-14	3.40	NA	NA	
CH_2-15	3.88	3.44/3.42	−0.44/−0.46	
CH_2-16	3.88	NA	NA	
CH_2-17	3.40	NA	NA	
CH_2-18	4.22	4.07	−0.15	$T_{16}CH_3$ (w)
CH_2-19	4.22	3.93	−0.29	
CH_2-20	3.40	NA	NA	
CH_2-21	3.88	NA	NA	
CH_2-22	3.88	3.22	−0.66	
CH_2-23	3.40	3.43	0.03	
CH_2-24	4.22	3.89/4.07	−0.33/−0.15	A_7H1' (m)

[a]Chemical shifts at 10°C.
[b]For the (arbitrary) proton numbering of 2,7-BisNP, see Scheme 1.
[c]$\Delta\delta = \delta_{bound} - \delta_{free}$.
s, strong; m, medium; w, weak; NA, Not assigned.

DISCUSSION

Thermal denaturation data strongly suggest the formation of a complex with a 1:1 stoichiometry upon binding of the bis-naphthalene macrocycle to the 11-mer duplex containing a TT mismatch. Accordingly in the NMR-monitored titration, clean formation of the ligand–DNA complex was observed at a 1:1 ratio and no excess of the ligand was necessary. Finally, 1D and 2D NMR data give evidence that only one type of complex between 2,7-BisNP and TT-DNA is formed.

The binding model, deduced from the numerous ligand–DNA NOE interactions, indicates that the macrocycle is located inside the double helix and sandwiches the thymine T_{16} of the A–T base pair in the vicinity of the T_6–T_{17} mismatched base pair. Presumably, the weaker stability of the A_7–T_{16} base pair compared with the G_5–C_{18} pair which flanks the T–T mismatch on the other side, leads to formation of a single complex. It is likely that the hydrogen bonding between A_7 and T_{16} is weakened due to the strong kink (38°) observed between the planes of these bases. The two polyammonium linkers of the macrocyclic scaffold lie both in the minor and the major grooves in a threading mode. In addition, the binding features two hydrogen bonds with the encapsulated thymine ($T_{16}O2$) and one hydrogen bond with one of the mismatched thymines (T_6O4). Both mismatched thymines are strongly displaced as compared to their position in the free duplex, especially T_{17} which is flipped out of the duplex toward the major groove. Thus, binding of the macrocycle induces a local distortion of the mismatch site, and interactions are observed over the four base pairs from C_5–G_{18} to G_8–C_{15}. Altogether, the ligand–DNA complex results from a complex array of multiple interactions involving the mismatched base pair and the three base pairs in the vicinity, which is based on hydrogen bonding, electrostatic interactions and π–stacking. It is clear that the macrocycle takes benefit of the local thermodynamic instability of the mismatch site and of the neighboring base pairs to enter inside the DNA and accommodate a thymine ring in its cavity. The binding of the ligand necessitates firstly the disruption of the Watson–Crick base pairing between A_7 and T_{16} and of the wobble pairing between T_6 and T_{17}. These events should be energetically compensated by the final stability of the complex, as evidenced by the large increase in melting temperature. In this regard, the three hydrogen bonds established between the ammonium groups of the ligand linkers and the carbonyl groups of the two thymine bases (T_6 and T_{16}) should provide significant contribution. Additionally, partial restoration of hydrogen bonding between A_7 and T_{16}, which is likely (although not evidenced by NMR), may contribute to stabilization of the complex. Another important contribution is due to strong electrostatic interaction of the highly charged (+4) ligand and DNA phosphate backbone, which is highly favorable since both ligand linkers occupy the grooves of the DNA. We can also anticipate that the entropic factors also largely participate to the stabilization effect, since, first, binding of the ligand leads to release of water molecules from the DNA grooves and from the solvation sphere of the ligand and,

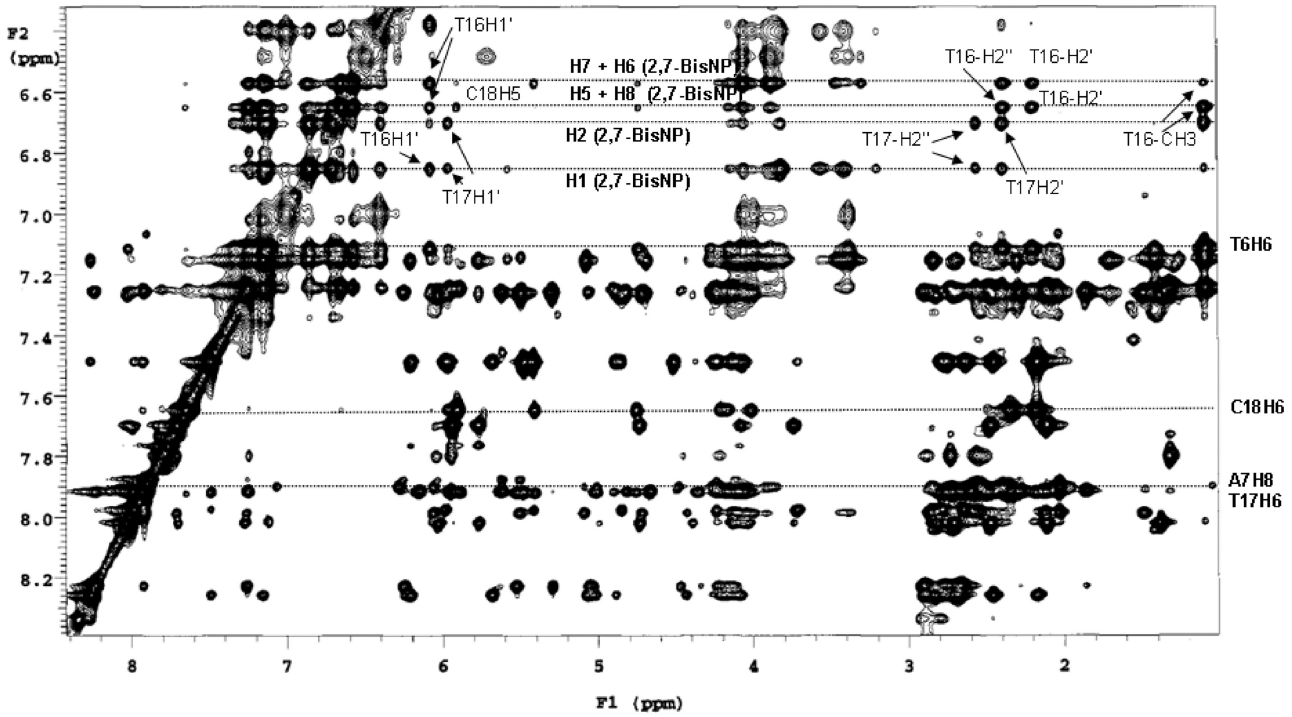


Figure 6. Expansion of the 250 ms NOESY spectrum of the 2,7-BisNP–TT-DNA complex in D_2O at 10°C. Correlations are between the aromatic protons and the sugar protons. Selected intermolecular 2,7-BisNP–TT-DNA cross-peaks are indicated.

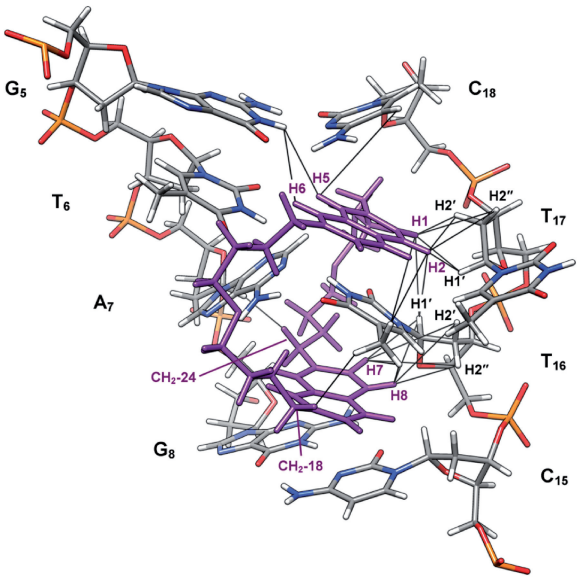


Figure 7. Structure of the complex of 2,7-BisNP (purple) with TT-DNA (CPK colors) from NMR-restrained molecular dynamics calculations showing the binding site and the NOEs observed in NMR spectra (black lines).

second, the ligand is preorganized due to its macrocyclic nature, which means low entropy cost for adaptation to the binding site.

The cyclobisintercalation or pseudocatenation binding mode observed in the present study is remarkable in many aspects.

(1) This very particular and unusual DNA binding behavior is evidenced here for the second time with a CBI macrocycle and a DNA duplex containing a single-base pairing defect (abasic site or a mismatch). Indeed, in a previous study, we have shown that a macrocyclic bis-acridine BisA binds via a similar threading double intercalation mode to an apurinic site with encapsulation of a pyrimidine base (cytosine) of the G–C base pair flanking the lesion (45). However, in that case, two complexes in an 85:15 ratio were observed: in the major complex, both acridine units were intercalated, sandwiching the G–C base pair flanking the lesion, whereas in the minor complex only one acridine was inserted in the abasic pocket, the location of the second one being not defined. This is different in the present study since the asymmetry of the base pairs flanking the mismatch site (A–T versus G–C) led to formation of only one complex, through the encapsulation of the more labile A_7–T_{16} base pair. Thus, this very particular binding mode involving double threading of the double helix at a base pairing defect with encapsulation of the neighboring base pair seems to be a unique feature of the CBI macrocycles. Nonetheless, it should be noted that threading bisintercalation has also been characterized in the case of a macrocyclic bis-naphthalenediimide (67) or proposed in the case of a macrocyclic bis-acridine (68,69). However, both derivatives were built with two long linker chains (> 10 Å) and were found to bind to regular double-stranded DNA with spanning over 2–4 bp. This binding behavior is in

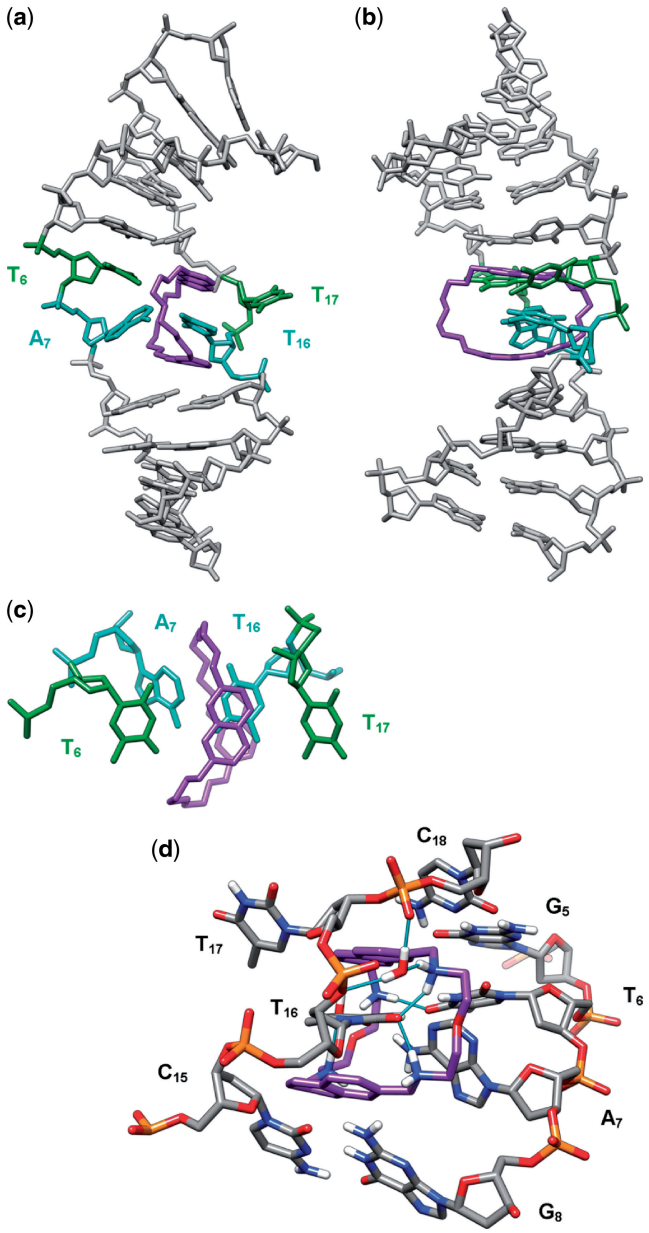


Figure 8. Threading intercalation binding mode of 2,7-BisNP (purple) to TT-DNA from NMR-restrained molecular dynamics calculations, viewed (**a**) from the major groove and (**b**) from the side of the flipped-out T_{17}. (**c**) Top view of the naphthalene rings of 2,7-BisNP stacking with the encapsulated thymine T_{16}. (**d**) Close-up view of the binding site (view from the minor groove, only heteroatom-bound hydrogen atoms are shown) showing direct and water-mediated hydrogen bonds (cyan sticks) between 2,7-BisNP and TT-DNA.

stark contrast with that of our macrocyclic bis-naphthalene that does not bind fully paired duplex DNA.

(2) As we anticipated (46–48), the characteristic DNA binding behavior of cyclobisintercalators studied by our group is obviously determined by the short length of the polyammonium linkers that imposes a constrained and short distance between the two aromatic moieties. Indeed, an interchromophoric distance of ~6.7–6.8 Å can be measured for 2,7-BisNP bound to the T-T mismatched duplex. On the one hand, this short distance prevents intercalation at contiguous Watson–Crick base pairs for steric reasons (neighbor-exclusion principle) (49); on the other hand, it is ideally suited for the encapsulation of aromatic guests, particularly nucleosides and nucleotides (70). This structural particularity governs both the binding preference for the T–T mismatch over the regular duplex and the sandwiching of the neighboring pyrimidine base (T_{16}) by π-stacking interactions.
Interestingly, the short distance between the two naphthalene planes is roughly the same in the solid state (6.6–6.7 Å). Conversely, the X-ray analysis indicates a Z-shaped conformation with antiparallel orientation of the two aromatic units, which differs from the almost parallel orientation of these units (U-shaped conformation) in the DNA-bound macrocycle as shown by the NMR analysis. This indicates a certain degree of conformational flexibility of the macrocyclic scaffold, resting mainly on the free rotation about the C_{ar}–C and C–N bonds between the aromatic units and the spacers, and is consistent with previous data obtained from the molecular dynamics study of the free 2,7-BisNP in a water box (48). This is an evidence of the conformational adaptability of the bis-naphthalene macrocycle that enables fitting with its DNA target and is likely a key factor of the binding interaction.

(3) Finally, the observed binding mode sheds light on the high structural deformability (plasticity) of duplex DNA which, albeit well-known, is often underestimated when designing DNA-interactive agents. In this regard, a comparison with the two other classes of mismatch-binding agents, i.e. naphthyridine derivatives (31) and bulky metal complexes (38–40), shows that in all cases binding proceeds via ejection of one of the mismatched bases, as in the present study. Thus, all three families of ligands take advantage of the local thermodynamic instability of the mismatch sites, compared to Watson–Crick base pairs. However, the uniqueness of the CBI series is represented by their two-point binding mode (mismatched base pair and intercalation pocket at the neighboring base pair step) and their conformational restrictions, which do not allow binding to fully paired DNA regions. In summary, the observed threading interaction is thus the result of a subtle interplay between the conformational flexibility of the macrocyclic ligand and the local dynamics of the DNA mismatched site.

CONCLUSIONS

In the present study, we characterized the structural details of a bis-naphthalene macrocycle bound to a T–T DNA mismatch via a rare cyclobisintercalation binding mode, which involves encapsulation of the neighboring thymine base between the aromatic planes of the macrocycle and ejection of one of the mispaired thymines. This binding mode is unique since it provides high specificity for DNA base pairing defects with respect to

regularly paired DNA, and allows recognition of unpaired thymine due to the hydrogen bonding pattern of the 2,7-BisNP ligand. Interestingly, the structure of the ligand–DNA complex obtained in the present study shows that the macrocyclic scaffold of the ligand allows the introduction of additional chemical functionalities, either at the linking chains or in the positions of the naphthalene units which are not in close contact with the DNA, without perturbing the binding specificity. This might enable us to endow the macrocycle with reporter groups (fluorescent, redox units) (71) for detection of the pairing defect. It may be expected that this pseudocatenated interaction can result in a slow dissociation of the ligand–DNA complex, which may dramatically increase the residence time of the macrocycle inside the DNA, as compared with classical intercalators or minor groove binders and interfere with the processing of mismatch repair by enzymes. Altogether, the present work highlights that bisnaphthalene 2,7-BisNP and, more largely, bisintercalators of the CBI type hold a great potential for therapeutic and diagnostic applications.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–4, Supplementary Figures 1–3.

ACKNOWLEDGMENTS

We thank Marie-Louise Dheu-Andries and Sebastien Morin for their help with the AMBER 10 program, Dr. Adrien Favier for his help with NMR experiments, and Dr. Hélène Jamet (UJF Grenoble) for Gaussian calculations. Prof. Elmar Weinhold (RWTH Aachen) is greatly acknowledged for helpful discussions. The figures in this manuscript were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco.

FUNDING

The Centre national de la recherche scientifique [TGE RMN THC Fr3050 (NMR studies at 800 MHz) and the interdisciplinary action 'Interface physique, biologie et chimie: soutien à la prise de risque' (PIR051 'DNA Lesion Binders')]. Funding for open access charge: CNRS.

Conflict of interest statement. None declared.

REFERENCES

1. Iyer,R.R., Pluciennik,A., Burdett,V. and Modrich,P.L. (2006) DNA mismatch repair: functions and mechanisms. *Chem. Rev.*, **106**, 302–323.
2. Jackson,S.P. and Bartek,J. (2009) The DNA-damage response in human biology and disease. *Nature*, **461**, 1071–1078.
3. Fukui,K. (2010) DNA mismatch repair in eukaryotes and bacteria. *J. Nucleic Acids*, Article ID 260512.
4. Verdine,G.L. and Bruner,S.D. (1997) How do DNA repair proteins locate damaged bases in the genome? *Chem. Biol.*, **4**, 329–334.
5. Rajski,S.R., Jackson,B.A. and Barton,J.K. (2000) DNA repair: models for damage and mismatch recognition. *Mutat. Res.*, **447**, 49–72.
6. Stivers,J.T. and Jiang,Y.L. (2003) A mechanistic perspective on the chemistry of DNA repair glycosylases. *Chem. Rev.*, **103**, 2729–2759.
7. Zharkov,D.O. and Grollman,A.P. (2005) The DNA trackwalkers: principles of lesion search and recognition by DNA glycosylases. *Mutat. Res.*, **577**, 24–54.
8. Yang,W. (2008) Structure and mechanism for DNA lesion recognition. *Cell Res.*, **18**, 184–197.
9. Friedman,J.I. and Stivers,J.T. (2010) Detection of damaged DNA bases by DNA glycosylase enzymes. *Biochemistry*, **49**, 4957–4967.
10. Helleday,T., Petermann,E., Lundin,C., Hodgson,B. and Sharma,R.A. (2008) DNA repair pathways as targets for cancer therapy. *Nat. Rev. Cancer*, **8**, 193–204.
11. Kelley,M.R. and Fishel,M.L. (2008) DNA repair proteins as molecular targets for cancer therapeutics. *Anticancer Agents Med. Chem.*, **8**, 417–425.
12. Zhu,Y., Hu,J., Hu,Y. and Liu,W. (2009) Targeting DNA repair pathways: a novel approach to reduce cancer therapeutic resistance. *Cancer Treat. Rev.*, **35**, 590–596.
13. Jorgensen,T.J. (2009) Enhancing radiosensitivity: targeting the DNA repair pathways. *Cancer Biol. Ther.*, **8**, 665–670.
14. Tell,G. and Wilson,D.M. III (2010) Targeting DNA repair proteins for cancer treatment. *Cell. Mol. Life Sci.*, **67**, 3569–3572.
15. Abbotts,R. and Madhusudan,S. (2010) Human AP endonuclease 1 (APE1): from mechanistic insights to druggable target in cancer. *Cancer Treat. Rev.*, **36**, 425–435.
16. Bapat,A., Glass,L.S., Luo,M., Fishel,M.L., Long,E.C., Georgiadis,M.M. and Kelley,M.R. (2010) Novel small-molecule inhibitor of apurinic/apyrimidinic endonuclease 1 blocks proliferation and reduces viability of glioblastoma cells. *J. Pharmacol. Exp. Ther.*, **334**, 988–998.
17. Madhusudan,S., Smart,F., Shrimpton,P., Parsons,J.L., Gardiner,L., Houlbrook,S., Talbot,D.C., Hammonds,T., Freemont,P.A., Sternberg,M.J. *et al.* (2005) Isolation of a small molecule inhibitor of DNA base excision repair. *Nucleic Acids Res.*, **33**, 4711–4724.
18. Seiple,L.A., Cardellina,J.H. II, Akee,R. and Stivers,J.T. (2008) Potent inhibition of human apurinic/apyrimidinic endonuclease 1 by arylstibonic acids. *Mol. Pharmacol.*, **73**, 669–677.
19. Simeonov,A., Kulkarni,A., Dorjsuren,D., Jadhav,A., Shen,M., McNeill,D.R., Austin,C.P. and Wilson,D.M. III (2009) Identification and characterization of inhibitors of human apurinic/apyrimidinic endonuclease APE1. *PLoS One*, **4**, e5740.
20. Wilson,D.M. III and Simeonov,A. (2010) Small molecule inhibitors of DNA repair nuclease activities of APE1. *Cell. Mol. Life Sci.*, **67**, 3621–3631.
21. Haince,J.F., Rouleau,M., Hendzel,M.J., Masson,J.Y. and Poirier,G.G. (2005) Targeting poly(ADP-ribosyl)ation: a promising approach in cancer therapy. *Trends Mol. Med.*, **11**, 456–463.
22. Ferraris,D.V. (2010) Evolution of poly(ADP-ribose) polymerase-1 (PARP-1) inhibitors. From concept to clinic. *J. Med. Chem.*, **53**, 4561–4584.
23. Underhill,C., Toulmonde,M. and Bonnefoi,H. (2011) A review of PARP inhibitors: from bench to bedside. *Ann. Oncol.*, **22**, 268–279.
24. Jiang,Y.L., Ichikawa,Y. and Stivers,J.T. (2002) Inhibition of uracil DNA glycosylase by an oxacarbenium ion mimic. *Biochemistry*, **41**, 7116–7124.
25. Huang,H., Stivers,J.T. and Greenberg,M.M. (2009) Competitive inhibition of uracil DNA glycosylase by a modified nucleotide whose triphosphate is a substrate for DNA polymerase. *J. Am. Chem. Soc.*, **131**, 1344–1345.
26. 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.
27. 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.

28. Kobori,A., Horie,S., Suda,H., Saito,I. and Nakatani,K. (2004) The SPR sensor detecting cytosine[bond]cytosine mismatches. *J. Am. Chem. Soc.*, **126**, 557–562.
29. 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.
30. Nakatani,K. (2009) Recognition of mismatched base pairs in DNA. *Bull. Chem. Soc. Jpn.*, **82**, 1055–1069.
31. Nakatani,K., Hagihara,S., Goto,Y., Kobori,A., Hagihara,M., Hayashi,G., Kyo,M., Nomura,M., Mishima,M. and Kojima,C. (2005) Small-molecule ligand induces nucleotide flipping in $(CAG)_n$ trinucleotide repeats. *Nat. Chem. Biol.*, **1**, 39–43.
32. Jackson,B.A. and Barton,J.K. (1997) Recognition of DNA base mismatches by a rhodium intercalator. *J. Am. Chem. Soc.*, **119**, 12986–12987.
33. Jackson,B.A. and Barton,J.K. (2000) Recognition of base mismatches in DNA by 5,6-chrysenequinone diimine complexes of rhodium(III): a proposed mechanism for preferential binding in destabilized regions of the double helix. *Biochemistry*, **39**, 6176–6182.
34. Junicke,H., Hart,J.R., Kisko,J., Glebov,O., Kirsch,I.R. and Barton,J.K. (2003) A rhodium(III) complex for high-affinity DNA base-pair mismatch recognition. *Proc. Natl Acad. Sci. USA*, **100**, 3737–3742.
35. Zeglis,B.M., Pierre,V.C. and Barton,J.K. (2007) Metallo-intercalators and metallo-insertors. *Chem. Commun.*, 4565–4579.
36. Chen,H., Yang,P., Yuan,C. and Pu,X. (2005) Study on the binding of base-mismatched oligonucleotide $d(GCGAGC)_2$ by cobalt(III) Complexes. *Eur. J. Inorg. Chem.*, 3141–3148.
37. Chen,H., Dou,C., Wu,Y., Li,H., Xi,X. and Yang,P. (2009) Structure-specific binding of $[Co(phen)_2(HPIP)]^{3+}$ to a DNA duplex containing sheared G:A mismatch base pairs. *J. Inorg. Biochem.*, **103**, 827–832.
38. 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.
39. Zeglis,B.M., Pierre,V.C., Kaiser,J.T. and Barton,J.K. (2009) A bulky rhodium complex bound to an adenosine-adenosine DNA mismatch: general architecture of the metalloinsertion binding mode. *Biochemistry*, **48**, 4247–4253.
40. 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.
41. 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.
42. Ernst,R.J., Song,H. and Barton,J.K. (2009) DNA mismatch binding and antiproliferative activity of rhodium metalloinsertors. *J. Am. Chem. Soc.*, **131**, 2359–2366.
43. Teulade-Fichou,M.-P. and Vigneron,J.-P. (2003) Interactions of macrocyclic compounds with nucleic acids. In: Demeunynck,M., Bailly,C. and Wilson,W.D. (eds), *Small Molecule DNA and RNA Binders: From Synthesis to Nucleic Acid Complexes*. Wiley-VCH, Weinheim, pp. 278–314.
44. 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.
45. 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.
46. 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.
47. Bahr,M., Gabelica,V., Granzhan,A., Teulade-Fichou,M.P. and Weinhold,E. (2008) Selective recognition of pyrimidine-pyrimidine DNA mismatches by distance-constrained macrocyclic bis-intercalators. *Nucleic Acids Res.*, **36**, 5000–5012.
48. Granzhan,A., Largy,E., Saettel,N. and Teulade-Fichou,M.-P. (2010) Macrocyclic DNA-mismatch-binding ligands: structural determinants of selectivity. *Chemistry*, **16**, 878–889.
49. Crothers,D.M. (1968) Calculation of binding isotherms for heterogenous polymers. *Biopolymers*, **6**, 575–584.
50. Vranken,W.F., Boucher,W., Stevens,T.J., Fogh,R.H., Pajon,A., Llinas,M., Ulrich,E.L., Markley,J.L., Ionides,J. and Laue,E.D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*, **59**, 687–696.
51. Paris,T., Vigneron,J.-P., Lehn,J.-M., Cesario,M., Guilheim,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.
52. McGhee,J.D. (1976) Theoretical calculations of the helix-coil transition of DNA in the presence of large, cooperatively binding ligands. *Biopolymers*, **15**, 1345–1375.
53. Slama-Schwok,A., Peronnet,F., Hantz-Brachet,E., Taillandier,E., Teulade-Fichou,M.P., Vigneron,J.P., Best-Belpomme,M. and Lehn,J.-M. (1997) A macrocyclic bis-acridine shifts the equilibrium from duplexes towards DNA hairpins. *Nucleic Acids Res.*, **25**, 2574–2581.
54. Teulade-Fichou,M.P., Fauquet,M., Baudoin,O., Vigneron,J.P. and Lehn,J.-M. (2000) DNA double helix destabilizing properties of cyclobisintercaland compounds and competition with a single strand binding protein. *Bioorg. Med. Chem.*, **8**, 215–222.
55. Wüthrich,K. (1986) *NMR of Proteins and Nucleic Acids*. John Wiley & Sons, New York.
56. Kalnik,M.W., Norman,D.G., Li,B.F., Swann,P.F. and Patel,D.J. (1990) Conformational transitions in thymidine bulge-containing deoxytridecanucleotide duplexes. Role of flanking sequence and temperature in modulating the equilibrium between looped out and stacked thymidine bulge states. *J. Biol. Chem.*, **265**, 636–647.
57. Coppel,Y., Berthet,N., Coulombeau,C., Coulombeau,C., Garcia,J. and Lhomme,J. (1997) Solution conformation of an abasic DNA undecamer duplex d(CGCACXCACGC) × d(GCGTGTGTGCG): the unpaired thymine stacks inside the helix. *Biochemistry*, **36**, 4817–4830.
58. Arnold,F.H., Wolk,S., Cruz,P. and Tinoco,I. Jr (1987) Structure, dynamics, and thermodynamics of mismatched DNA oligonucleotide duplexes $d(CCCAGGG)_2$ and $d(CCCTGGG)_2$. *Biochemistry*, **26**, 4068–4075.
59. Kouchakdjian,M., Li,B.F., Swann,P.F. and Patel,D.J. (1988) Pyrimidine • pyrimidine base-pair mismatches in DNA. A nuclear magnetic resonance study of T•T pairing at neutral pH and C•C pairing at acidic pH in dodecanucleotide duplexes. *J. Mol. Biol.*, **202**, 139–155.
60. Gervais,V., Cognet,J.A.H., Le Bret,M., Sowers,L.C. and Fazakerley,G.V. (1995) Solution structure of two mismatches A•A and T•T in the *K-ras* gene context by nuclear magnetic resonance and molecular dynamics. *Eur. J. Biochem.*, **228**, 279–290.
61. Trotta,E. and Paci,M. (1998) Solution structure of DAPI selectively bound in the minor groove of a DNA T.T mismatch-containing site: NMR and molecular dynamics studies. *Nucleic Acids Res.*, **26**, 4706–4713.
62. Bhattacharya,P.K., Cha,J. and Barton,J.K. (2002) ^{1}H NMR determination of base-pair lifetimes in oligonucleotides containing single base mismatches. *Nucleic Acids Res.*, **30**, 4740–4750.
63. Patel,D.J., Kozlowski,S.A., Marky,L.A., Rice,J.A., Broka,C., Dallas,J., Itakura,K. and Breslauer,K.J. (1982) Structure, dynamics, and energetics of deoxyguanosine. thymidine wobble base pair formation in the self-complementary d(CGTGAATTC GCG) duplex in solution. *Biochemistry*, **21**, 437–444.
64. Allawi,H.T. and SantaLucia,J. Jr (1997) Thermodynamics and NMR of internal G.T mismatches in DNA. *Biochemistry*, **36**, 10581–10594.
65. Gorenstein,D.G. (1994) Conformation and dynamics of DNA and protein–DNA complexes by ^{31}P NMR. *Chem. Rev.*, **94**, 1315–1338.

66. Mills,J.E. and Dean,P.M. (1996) Three-dimensional hydrogen-bond geometry and probability information from a crystal survey. *J. Comput.-Aided Mol. Des.*, **10**, 607–622.
67. Chu,Y., Hoffman,D.W. and Iverson,B.L. (2009) A pseudocatenane structure formed between DNA and a cyclic bisintercalator. *J. Am. Chem. Soc.*, **131**, 3499–3508.
68. Zimmerman,S.C., Lamberson,C.R., Cory,M. and Fairley,T.A. (1989) Topologically constrained bifunctional intercalators: DNA intercalation by a macrocyclic bisacridine. *J. Am. Chem. Soc.*, **111**, 6805–6809.
69. Veal,J.M., Li,Y., Zimmerman,S.C., Lamberson,C.R., Cory,M., Zon,G. and Wilson,D.M. (1990) Interaction of a macrocyclic bisacridine with DNA. *Biochemistry*, **29**, 10918–10927.
70. 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.
71. Granzhan,A. and Teulade-Fichou,M.-P. (2009) Synthesis of mono- and bibrachial naphthalene-based macrocycles with pyrene or ferrocene units for anion detection. *Tetrahedron*, **65**, 1349–1360.