Structural Optimization of Azacryptands for Targeting Three-Way DNA Junctions

Angélique Pipier, Titouan Chetot, Apollonia Kalamatianou, Nicolas Martin, Maëlle Caroff, Sébastien Britton, Nicolas Chéron, Lukáš Trantírek, Anton Granzhan, David Monchaud

GDCh Research Article Angewandte International Edition Chemie www.angewandte.org

How to cite: *Angew. Chem. Int. Ed.* **2024**, *63*, e202409780
doi.org/10.1002/anie.202409780

DNA Junctions

Structural Optimization of Azacryptands for Targeting Three-Way DNA Junctions

Angélique Pipier, Titouan Chetot, Apollonia Kalamatianou, Nicolas Martin, Maëlle Caroff, Sébastien Britton, Nicolas Chéron, Lukáš Trantírek, Anton Granzhan, and David Monchaud**

Abstract: Transient melting of the duplex-DNA (B-DNA) during DNA transactions allows repeated sequences to fold into non-B-DNA structures, including DNA junctions and G-quadruplexes. These noncanonical structures can act as impediments to DNA polymerase progression along the duplex, thereby triggering DNA damage and ultimately jeopardizing genomic stability. Their stabilization by *ad hoc* ligands is currently being explored as a putative anticancer strategy since it might represent an efficient way to inflict toxic DNA damage specifically to rapidly dividing cancer cells. The relevance of this strategy is only emerging for three-way DNA junctions (TWJs) and, to date, no molecule has been recognized as a reference TWJ ligand, featuring both high affinity and selectivity. Herein, we characterize such reference ligands through a combination of *in vitro* techniques comprising affinity and selectivity assays (competitive FRET-melting and TWJ Screen assays), functional tests (qPCR and *Taq* stop assays) and structural analyses (molecular dynamics and NMR investigations). We identify novel azacryptands **TrisNP-*amphi*** and **TrisNP-*ana*** as the most promising ligands, interacting with TWJs with high affinity and selectivity. These ligands represent new molecular tools to investigate the cellular roles of TWJs and explore how they can be exploited in innovative anticancer therapies.

Introduction

In 2022, the telomere-to-telomere (T2T) consortium reported on the sequence of a truly complete genome,[1] gathering 3.05 Gbp of nuclear DNA that cover 22 chromosomes plus the X-chromosome. The *tour-de-force* was to include satellite DNA, consisting of highly repeated DNA sequences known to be reluctant to sequencing. This unique whole-genome coverage indicated that *ca.* 54 % of the genome is composed of repeated elements, comprising both tandem repeats (simple repeats and satellites) and interspersed repeats (short or long interspersed elements, SINEs or LINEs, respectively).[2]

Tandem repeats, or satellites *sensu lato*, are the repeats for which the genomic prevalence and functional relevance are the most studied. The term *satellite* was coined in 1961 due to the distribution of these sequences above and below the band of the bulk DNA during equilibrium sedimentation experiments.[3] Satellites are classified by their size: *i*) microsatellites, or short tandem repeats (STRs), are both short (two to six bp-long sequence per pattern) and abundant (they cover *ca.* 3 % of our genome), a representative example being the telomeric microsatellite d[TTAGGG]$_n$, with repeats > 10 kb; *ii*) minisatellites are *ca.* 15 bp-long sequence/pattern with arrays of highly variable length (from 0.5 to 30 kb); *iii*) satellites (*ca.* 200 bp-long sequence/pattern) constitute the bulk of centromeres and both pericentromeric and subtelomeric regions, among which α-satellites are the most abundant (representing *ca.* 50 % of satellite DNA, 10 % of all DNA repeats); and *iv*) macrosatellites (> 1 kb-long sequence/pattern) represent large chromosomal regions.[4]

The distinguishing feature of DNA repeats stands in their ability to fold into DNA structures that deviate from

[*] Dr. A. Pipier, Dr. D. Monchaud
Institut de Chimie Moléculaire, ICMUB CNRS UMR6302, 9, Avenue Alain Savary, 21078 Dijon, France
E-mail: david.monchaud@cnrs.fr

T. Chetot, A. Kalamatianou, N. Martin, Dr. A. Granzhan
Chemistry and Modelling for the Biology of Cancer (CMBC), CNRS UMR9187, INSERM U1196, Institut Curie, Université Paris Saclay, 91405 Orsay, France
E-mail: anton.granzhan@curie.fr

M. Caroff, Dr. S. Britton
Institut de Pharmacologie et de Biologie Structurale (IPBS), Université de Toulouse, CNRS, Université Toulouse III – Paul Sabatier (UT3), Toulouse, France

Dr. N. Chéron
PASTEUR, Département de chimie, École Normale Supérieure (ENS), PSL University, Sorbonne Université, CNRS UMR8640, 75005 Paris, France

Dr. L. Trantírek
Central European Institute of Technology, Masaryk University, Kamenice 753/5, 625 00 Brno, Czech Republic

the classical Watson–Crick duplex (or B-DNA), consequently termed non-B-DNA structures.[5] A single-molecule real-time (SMRT) sequencing polymerization kinetics study demonstrated that *ca.* 13 % of the human genome could fold into non-B-DNA structures, with important consequences for genetic stability since non-B-DNA structures can alter polymerization kinetics and increase error rates.[6] When translated into cells, the derailment of polymerases in charge of DNA transactions (replication, transcription and repair) results in DNA damage, which is coped with by activation of the DNA damage response (DDR) machinery.[7] The formation and persistence of non-B-DNA structures thus threaten genomic stability.[8]

The nature of non-B-DNA structures is encoded in their sequences. To date, the alphabet of DNA structures is ripe with *ca.* 20 letters, from A- to Z-DNA (Figure 1), but the cellular existence and prevalence have been shown only for a handful of them.[9] The low-complexity direct repeats (DRs) are involved in the formation of G-quadruplex-DNA (G-DNA or G4, *e.g.*, the guanine (G)-rich telomeric d-$[TTAGGG]_n$ repeats),[10] i-motifs (iMs, or i-DNA, *e.g.*, the cytosine (C)-rich d$[TCCCCC]_n$ repeats),[11] or hairpins (also termed three-way DNA junctions, or TWJs, slipped- or S-DNA, *e.g.*, in d$[CAG]_n$ or d$[CGG]_n$ trinucleotide repeats). More complex repeats include the mirror repeats (MRs), where the second half of the sequence is the mirror image of the first half: they are involved in the formation of triplex-DNA (also referred to as hinged DNA or H-DNA). Inverted repeats (IRs), where the second half of the sequence is reverse complementary to the first half, are involved in the formation of four-way DNA junctions (or FWJs, also termed cruciform DNA, or C-DNA).

The genomic distribution of satellite DNA is uneven: IRs are by far the most abundant repeats, with a median occurrence of 206 motifs/100 kb *vs.* 91 and 35 motifs/100 kb for DRs and MRs, respectively (IR size > 6 nt, with intervening sequence up to 100 nt; DR size from 10 to 300 nt; and MR size > 10 nt, with intervening sequence < 8 nt).[12] Given that non-B-DNA structures jeopardize genetic stability, their high genomic prevalence is of critical strategic relevance in the global context of cancers. This was exploited through the use of small molecules (ligands) designed to specifically target non-B-DNA structures (Figure 1), with the aim of blocking DNA polymerase progression and, in doing so, inflicting severe DNA damage to targeted cells.[8a] This approach is well documented for ligands targeting G4s[13] and emerging for those targeting TWJs.[9,14]

Over the past years, significant efforts have been invested to identify and characterize promising TWJ ligands. After the initial discovery of TWJ-binding properties of azacryptands,[15] we *i*) screened > 1,200 compounds in the aim of characterizing the most promising scaffolds,[16] *ii*) assessed their antiproliferative activity, and *iii*) demonstrated their ability to trigger DNA damage and their synergistic relationship with DDR inhibitors, a strategy referred to as chemically induced synthetic lethality. These studies provided proof of the use of TWJ ligands as DNA damage-inducing agents.[17] In parallel with our efforts, Brabec and colleagues demonstrated that TWJ-targeting iron metallohelicates efficiently induce DNA damage in cells,[18] while Vasquez-Lopez and colleagues showed that copper(II) metallohelicates trigger oxidative DNA damage at TWJ sites,[19] thus providing additional support to this concept. We also chemically modified one of the TWJ ligands in order to perform bioorthogonal labelling and gain insights into the cellular uptake and distribution of TWJ ligands.[20]

To go a step further, we must now address the most critical issue regarding TWJ ligands: their specificity for TWJs over other DNA structures, including B-DNA and G4. This specificity is required for investigating TWJ biology in depth *via* classical chemical biology means. Indeed, while our prototype ligands **TrisNP** and **TrisPOB** (see below) showed preferential affinity for TWJs in competitive FRET-melting, mass spectrometry (ESI-MS) and competition

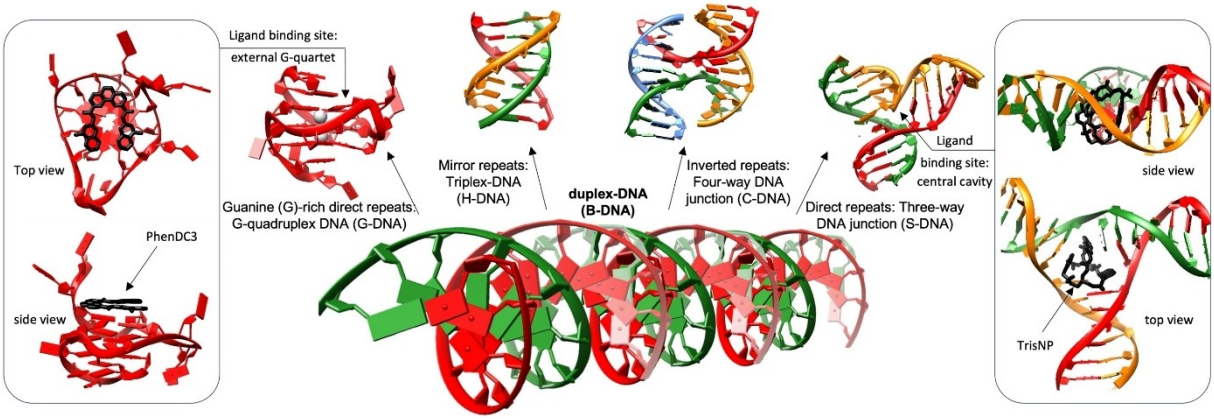


Figure 1. Schematic representation of the structure of the canonical duplex-DNA (B-DNA) and of the non-canonical structures originating in the folding of repeated sequences, including G-quadruplex-DNA (G-DNA), triplex-DNA (H-DNA), four-way DNA junction (or cruciform DNA, C-DNA) and three-way DNA junction (or slipped DNA, S-DNA). Left and right panels: structure of G-DNA (left) or S-DNA (right) in interaction with one of their respective ligands: **PhenDC3** (left, PDB ID: 2MGN) and **TrisNP** (right, structural model from molecular dynamics). DNA structures were rendered with the UCSF Chimera package.

equilibrium dialysis assays, their interaction with G4s *in vitro* was non-negligible.[15,20–21] Even if their cellular effects were clearly distinct from that of G4 ligands,[20] the interaction with G4s questions how they actually mediate their cellular outcomes and highlights the need for truly specific TWJ ligands.

To tackle this issue, we report herein on the design, synthesis, *in vitro* and *in silico* evaluations of a series of 17 azacryptands including **TrisNP**, **TrisPOB** and 15 novel derivatives, aiming at optimizing their TWJ affinity and selectivity. These investigations include FRET-melting, TWJ-screen and the newly developed quantitative PCR (qPCR) stop and *Taq* polymerase stop assays to investigate the ability of the candidates to impede polymerase processivity, along with molecular dynamics (MD) to gain insights into their binding mode, NMR to demonstrate target engagement in the presence of competing cellular components, and cytotoxicity screenings to investigate their cellular activity.

Results and Discussion

Molecular design and synthesis. Our leading TWJ ligands, **TrisNP** and **TrisPOB**, belong to the family of azacryptands, *i.e.*, macrobicyclic cage-like compounds containing three (hetero)aromatic units connected with polyamine linkers.[22] Herein, we extended the exploration of this scaffold in a search of better TWJ ligands. Specifically, 15 novel azacryptands (Figure 2) were designed by a systematic variation of *i*) the nature of the aromatic units, *i.e.*, naphthalene (**TrisNP** and derivatives **1**–**5**) *vs.* bis-benzene (**TrisPOB**, **6** and **7**), anthracene (**8**) and benzene (**9**–**15**) derivatives; *ii*) the substitution pattern of the aromatic units in the naphthalene (**TrisNP** and **1**–**5**) and benzene series (**5** and **9**), which has a strong impact on the overall geometry of the azacryptand; and *iii*) the nature of the C_3-symmetric polyamine linkers connecting the aromatic units in the naphthalene series (**TrisNP** *vs.* compounds **1** and **2**).

Most azacryptands were obtained through a well-established [3+2]-condensation route from the corresponding aromatic dialdehydes and C_3-symmetric triamines, with or without isolation of the corresponding hexaimine intermediates followed by their reduction with $NaBH_4$ (Schemes S1 and S2). Azacryptands containing bulky aromatic units (**5**, **8**, **11**) could not be obtained through this method since the initial [3+2]-condensation step was inefficient.[23] In this case, the first step employed a [2+2]-condensation, followed by *in situ* reduction to give macrocyclic polyamine intermediates. These intermediates were subsequently made to react with the third equivalent of the dialdehyde before *in situ* reduction and purification (Schemes S3–S5). Hybrid azacryptand **10** was also obtained by this method, using tetrafluoroterephthalaldehyde in the first and terephthalal-

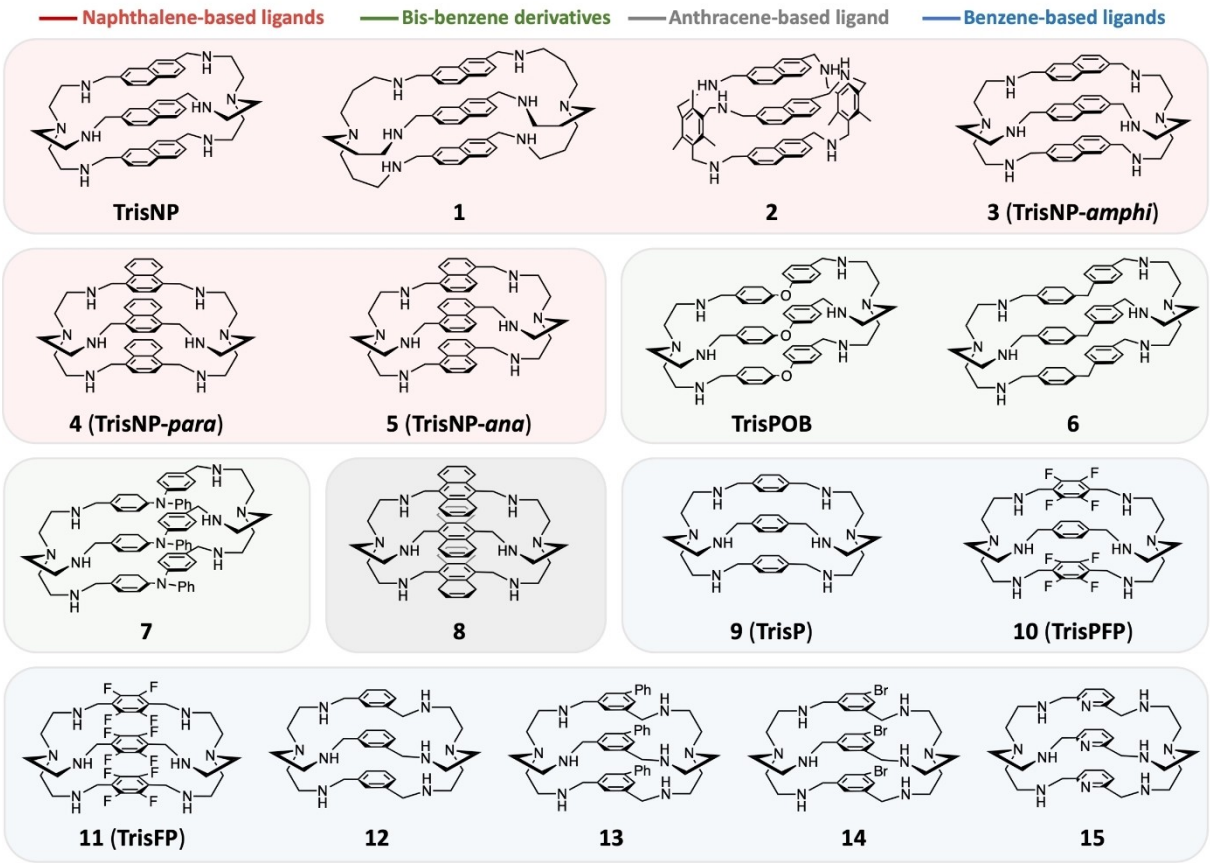


Figure 2. Chemical structures of established TWJ ligands (**TrisNP** and **TrisPOB**) and novel TWJ ligand candidates (**1**–**15**) used in this study.

dehyde in the second condensation steps (Scheme S5). All azacryptands, except for **7**, were converted into water-soluble hydrochloride or hydrobromide (**3**, **12**) salts. The detailed synthetic procedures and characterization data are provided in the Supporting Information.

***In vitro* TWJ stabilization and selectivity**. The TWJ affinity and selectivity of the novel azacryptands (along with **TrisNP** and **TrisPOB** as references) were first evaluated by FRET-melting experiments,[16] performed with TWJ-, G4- and double-stranded (DS)-DNA (also referred to as duplex) labelled with a FAM on their 5′-end and a TAMRA on their 3′-end (FAM-TWJ-TAMRA, FAM-G4-TAMRA and FAM-DS-TAMRA, respectively, Table S1). In the absence of ligand, the thermal denaturation of these structures occurred at $T_{1/2}=49.6$, 53.6 and 55.5 °C, respectively, indicating a comparable thermodynamic stability. This enables the use of ligand-induced stabilization ($\triangle T_{1/2}$) as a proxy of ligand's affinity to one or another structure and, this way, to identify the most TWJ-selective candidates.

The results shown in Figure 3A and Table S2 indicate that the TWJ stabilization imparted by the azacryptands is fair to good for all compounds except for two candidates (**7**

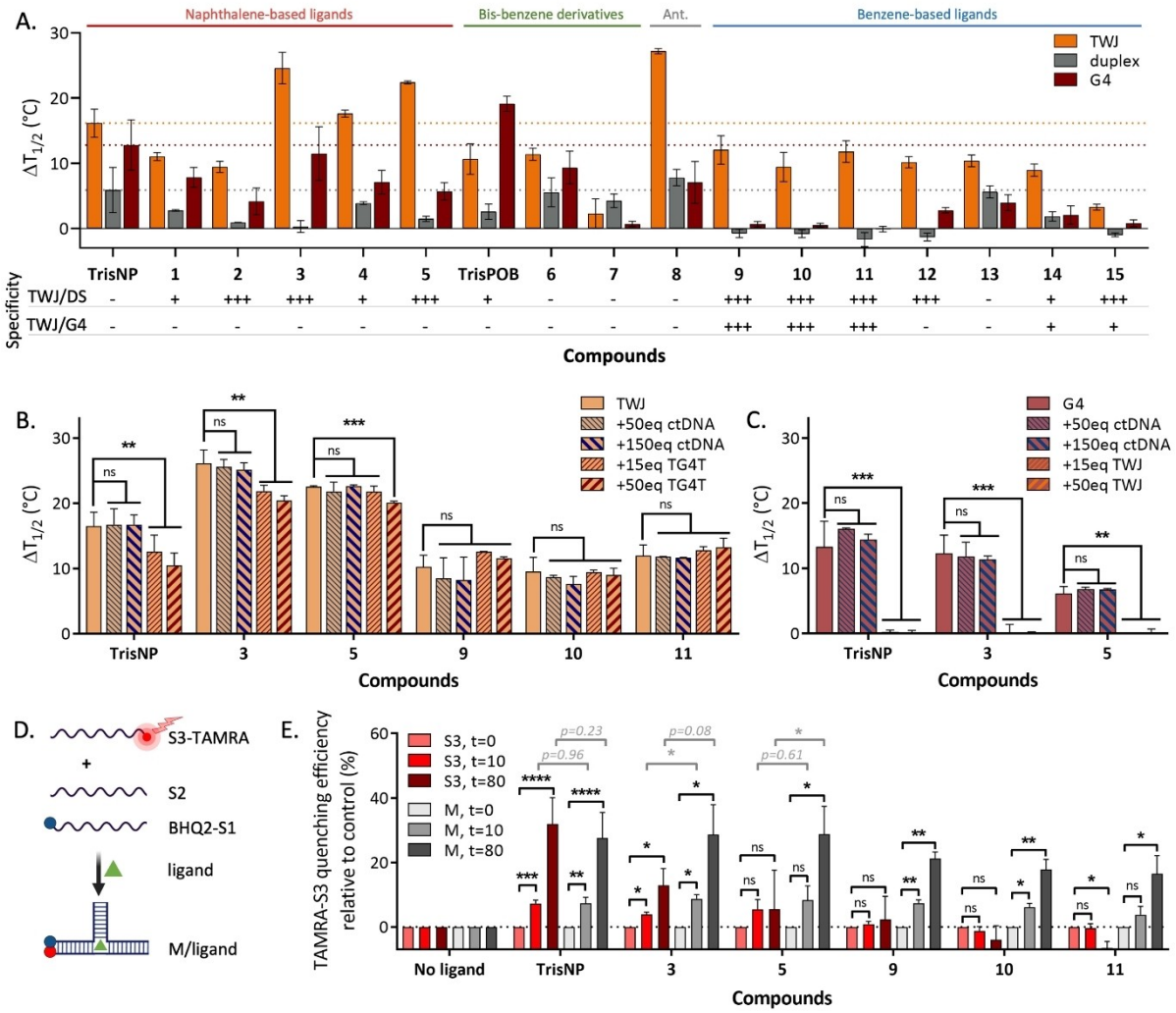


Figure 3. A) Evaluation of the apparent affinity of 17 azacryptands by FRET-melting assays performed with TWJ, DS and G4. Assays were performed with 0.2 μM of each double-labelled DNA structure and 1.0 μM of each ligand in appropriate buffer. The specificity score is estimated on the basis of the $\Delta T_{1/2}{}^{TWJ}/\Delta T_{1/2}{}^{G4}$ or $\Delta T_{1/2}{}^{TWJ}/\Delta T_{1/2}{}^{DS}$ ratio (+ + +: ≥ 10; +: ≥ 4) or the lack of DS/G4 stabilization ($\Delta T_{1/2}<4$ °C). Dotted lines correspond to **TrisNP** results. B) Evaluation of the TWJ-selectivity of the five most promising candidates plus **TrisNP** by competitive FRET-melting experiments performed with FAM-TWJ-TAMRA in the absence or presence of 10 and 30 μM ctDNA, or 3 and 10 μM TG_4T as competitors. The increase in $\Delta T_{1/2}$ values in the presence of an excess of competitors (*e.g.*, **9** and **11**) is often observed for highly selective ligands and likely originates in an increase of ionic strength in the experimental medium. C) Evaluation of G4-DNA binding of **3**, **5** and **TrisNP** by competitive FRET-melting experiments performed with FAM-G4-TAMRA in the absence or presence of 10 and 30 μM ctDNA, or 3 and 10 μM TWJ as competitors. D) Principle of the modified TWJ-Screen assay (blue sphere = BHQ2; red sphere = TAMRA; green triangle = ligand). E) Evaluation of the ability of the five most promising candidates plus **TrisNP** to promote TWJ formation by the TWJ-Screen assay, quantified by the relative quenching of S3-TAMRA fluorescence when mixed with BHQ2-S1 and S2 (M, 0.2 μM) and the ligand (1.0 μM, the mixture of S3-TAMRA + ligand being used as a control) as a function of the time. Statistical analyses were performed with multiple unpaired *t*-tests: ns: $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

and **15**; $\triangle T_{1/2}^{TWJ} = 2.3–3.3\,°C$) and reaches very high levels of stabilization for five of them (**TrisNP**, **3–5** and **8**; $\triangle T_{1/2}^{TWJ} > 15\,°C$). They also indicate that nine candidates (**2**, **3**, **5**, **9–12**, **14** and **15**) have little effect on DS-DNA ($\triangle T_{1/2}^{DS} \leq 2\,°C$), while only five candidates (**7**, **9–11** and **15**) have little impact on G4 ($\triangle T_{1/2}^{G4} \leq 1\,°C$). Altogether, this first screen led to the identification of five azacryptands (**3**, **5** and **9–11**) having significant apparent affinity for TWJ ($\triangle T_{1/2}^{TWJ} \geq 10\,°C$) and good selectivity over DS- (**3**, **5**) or both DS- and G4-DNA (**9–11**).

To assess the selectivity of these ligands, we performed competitive FRET-melting experiments with **3**, **5** and **9–11** along with **TrisNP** as a reference and both calf thymus DNA (ctDNA) and TG_4T as competitors representing the diversity of cellular DNA structures (double-stranded DNA and G4s); their thermal stability was largely superior to that of TWJ, enabling their use as competitors (TG_4T, $T_{1/2} = 76\,°C$ *vs.* 49 °C for FAM-TWJ-TAMRA, Figure S1). As seen in Figure 3B and Tables S3 and S4, the experiments performed in the presence of a large excess of either ctDNA (50 and 150 mol. equiv., expressed in nucleotide content) or G4-DNA (15 or 50 mol. equiv. TG_4T, expressed in G4 units) show that TWJ stabilization imparted by these ligands is largely maintained in a competitive context. The selectivity factor ($^{FRET}S = (\triangle T_{1/2}^{TWJ}{}_{(+competitor)} / \triangle T_{1/2}^{TWJ}{}_{(no\ competitor)}) \times 100$, in %) over DS-DNA is excellent for **TrisNP**, **3**, **5** and **11** ($^{FRET}S \geq 90\,\%$), and good for **9** and **10** ($\geq 75\,\%$); the selectivity over G4-DNA is excellent for **9–11** ($\geq 90\,\%$), good for **3** and **5** ($\geq 75\,\%$), but moderate for **TrisNP** ($\geq 60\,\%$). Given the non-negligible interaction of **TrisNP**, **3** and **5** with G4s ($\triangle T_{1/2}^{G4} \geq 5\,°C$, further confirmed with both Myc and Kit G4s, Figure S1 and Table S5), we performed a reverse competition assay, using FAM-G4-TAMRA in the presence of a large excess of either ctDNA (50 and 150 mol. equiv.) or TWJ (15 or 50 mol. equiv., expressed in TWJ motif units). The results seen in Figure 3C and Table S3 confirm the preferential TWJ affinity of all three ligands, as the G4 stabilization effect is lost ($^{FRET}S < 2\%$) in the presence of TWJ, while being strongly maintained ($^{FRET}S \geq 90\,\%$) in the presence of ctDNA, confirming their specificity for non-B-DNA structures.

We next evaluated the ability of these candidates to promote TWJ folding *via* a modified version of the TWJ-Screen assay.[24] In this assay, two of the three strands constituting an intermolecular TWJ were labelled with TAMRA on the 3′-end of the S3 strand and its quencher BHQ2 on the 5′-end of the S1 strand, while the S2 strand was left unmodified (Figure 3D). The three strands used herein are short enough (14 nt-long, Table S1) not to fold spontaneously into a stable TWJ in the condition of this assay. The presence of a ligand promotes TWJ folding, which can be monitored in real time by observing the FRET between BHQ2 and TAMRA, brought together upon TWJ formation. Of note, the interaction of the ligands with S3-TAMRA alone was also monitored to exclude the possible interference of the ligands with the label, unrelated to TWJ folding.

The results seen in Figure 3E and Table S6 indicate that both **TrisNP** and **3** interact with TAMRA-S3 alone (as TAMRA fluorescence intensity (△FI) decreased by *ca.* 32 and 13 %, respectively, after 80 min) while **5** barely interacted with it (ΔFI *ca.* 6 %) and **9–11** did not interact with this control ($\triangle FI \leq 2\,\%$). When all strands were mixed (M, mixture), all ligands triggered a significant decrease in TAMRA fluorescence intensity (△FI observed after 80 min between 17 to 29 %), evidencing the TWJ assembly. The comparison of the results obtained with S3-TAMRA alone and M highlights the excellent TWJ-folding ability of **9–11**, good for **5** (as it had a marginal interaction with the control strand), while the effects of **3** and **TrisNP** cannot be reliably interpreted due to their strong interaction with S3-TAMRA.

In vitro* stalling of DNA synthesis at non-B-DNA sites**. To get closer to biological applications, we assessed whether ligand-stabilized TWJs impede DNA transactions *in vitro.* To this end, we adapted a qPCR stop assay (previously exploited to study the effects of G4 stabilization)[25] and applied it to a selection of the most promising candidates, *i.e.*, the naphthalene-based **3** (or **TrisNP-*amphi) and **5** (or **TrisNP-*ana***), the benzene-based **10**, along with **TrisNP** and **TrisPOB** as references (the results obtained with **9**, **11** and the G4 ligand **PhenDC3** as control can be seen in Figure S2). Briefly, a TWJ-forming sequence (along with a G4-forming sequence and a scrambled sequence as controls) was inserted in the middle of a 64-nt strand (27 nt on the 5′ side, 37 nt on the 3′ side) used by Sabouri *et al.*[25a] and its amplification by *Taq* DNA polymerase was followed through a typical qPCR procedure. We investigated the extent to which the different concentrations of ligands inhibited DNA polymerization, which translates into a decrease in DNA amplification (expressed as fold change relative to the amplification obtained without ligand).

As seen in Figures 4, S2 and S3 and Table S7, the observed effects vary from ligand to ligand: both **TrisNP** and **TrisPOB** inhibited the amplification of the G4-containing matrix ($IC_{50} = 0.51$ and 1.22 μM, respectively) more efficiently than the TWJ-containing matrix ($IC_{50} = 4.27$ and 9.14 μM, respectively; *ca.* a 8-fold difference in IC_{50}) and were also quite active against the scrambled matrix ($IC_{50} =$ 4.70 and 19.0 μM, respectively), in line with FRET-melting results (Figure 3A). In contrast, two other naphthalene derivatives, **3** and **5**, displayed a higher activity against TWJ *vs.* G4 ($IC_{50} = 0.58$ *vs.* 2.32 μM for **3**; 0.41 *vs.* >2 μM for **5**). Even if they also inhibited amplification of the scrambled matrix ($IC_{50} = 1.99$ and 1.50 μM, respectively), this inhibition occurred at high concentrations and an interesting specificity for TWJ could be observed at concentrations close to the IC_{50} values (~0.5 μM, with a percentage of inhibition of 21 and 0 % for **3** against the G4-containing and scrambled sequences, respectively, and 19 and 4 % for **5**). Finally, the benzene derivative **10** was poorly efficient in this assay, being non-significantly more active with the G4 matrix (which was also observed with **9** and **11**, Figure S2). It is important to note that the interaction of the ligands with the scrambled sequence might originate in the long-known ability of polyamines containing aromatic fragments to interact with single-stranded DNA;[26] this admittedly tough control was however included to be representative of the diversity of the nucleic acids that can be found in cells. We

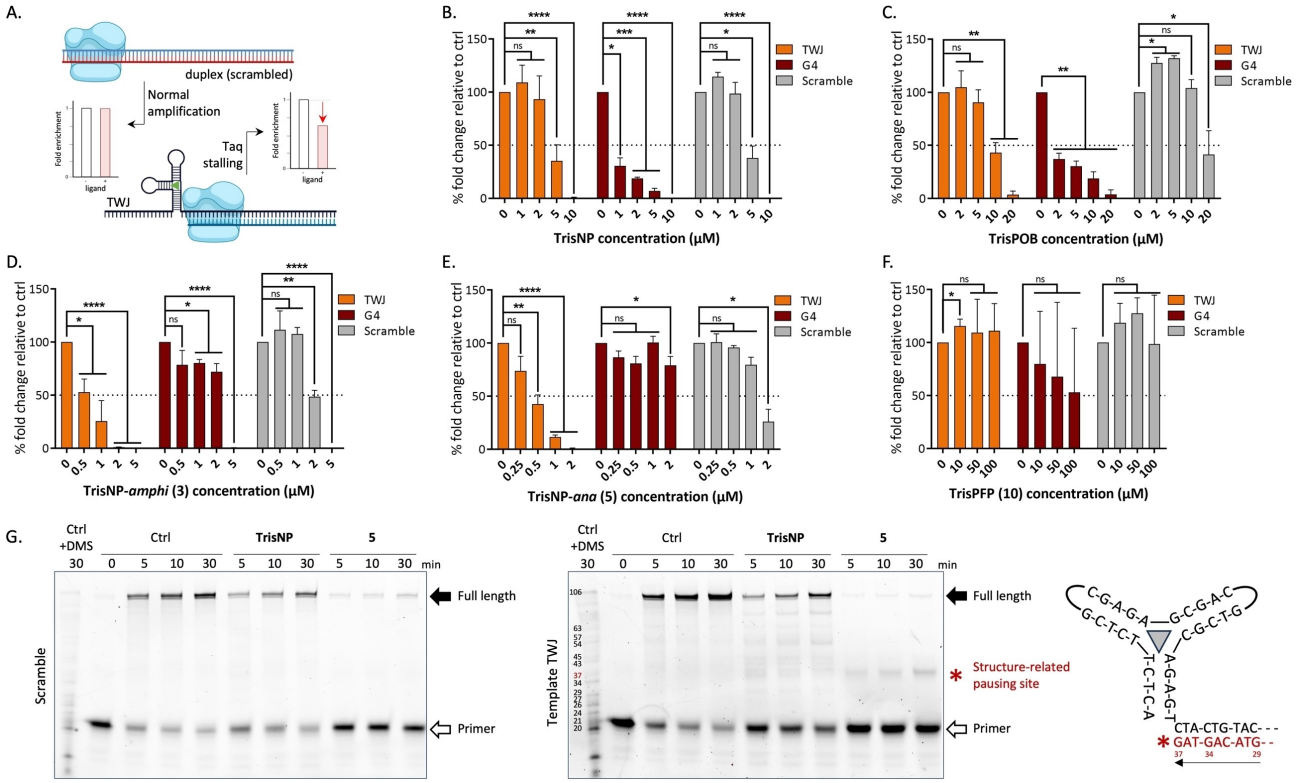
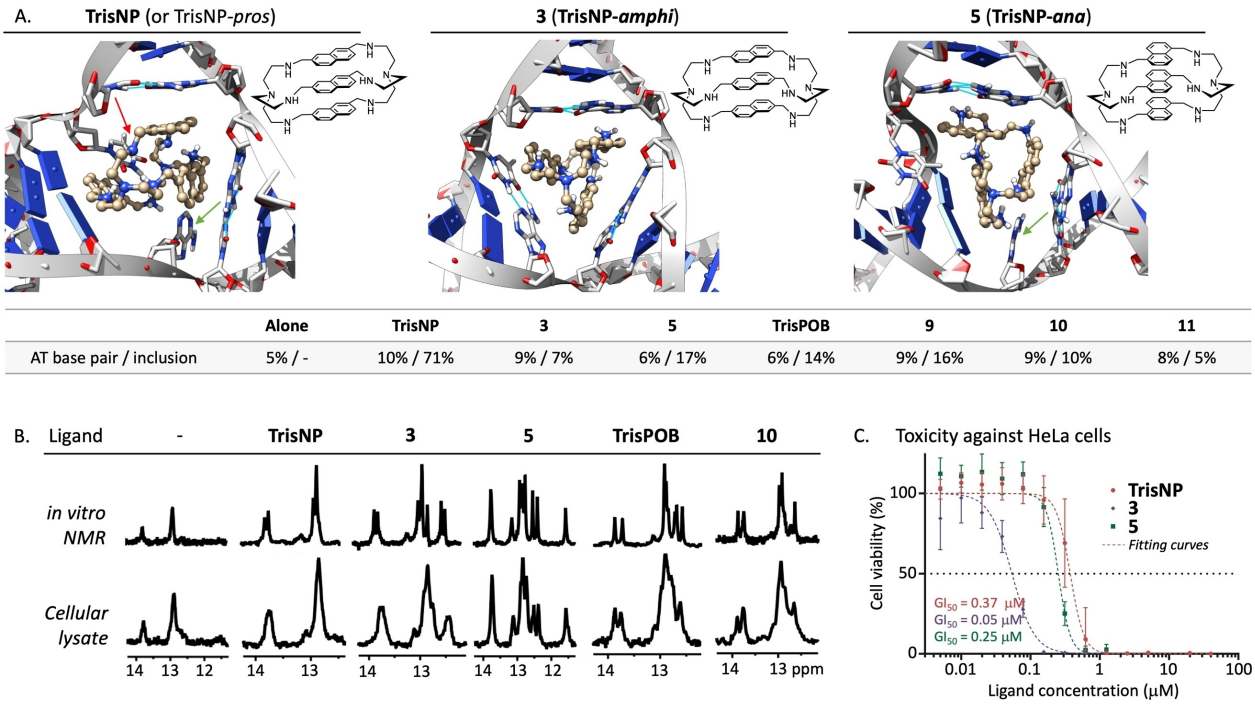


Figure 4. A) Principle of the qPCR stop assay (created with BioRender). B–F) Evaluation of the ability of **TrisNP** (B), **TrisPOB** (C) (as references) and three out of five most promising candidates (D: compound **3**; E: compound **5**; F: compound **10**; results for compounds **9** and **11** in Figure S1) to inhibit *in vitro* polymerization by the qPCR Stop assay, quantified as relative amplification (% fold change) of structured (TWJ or G4) or control (Scramble) templates (2 pM). Data are means ± s.d. from at least 3 independent experiments performed in duplicates; statistical analyses were performed with multiple unpaired *t*-tests: ns: $p > 0.05$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$. G) *Taq* polymerase stop assays performed with the scramble and TWJ-containing templates (the G4 template in Figure S4) without or with ligands (**TrisNP** and **5**); reactions were performed at 50 °C, stopped after 5, 10 or 30 min and then loaded on a denaturing 12 % (v/v) polyacrylamide gel (quantification in Figure S4). Right panel: schematic representation of a ligand-stabilized TWJ that acts as a structural roadblock able to pause polymerization (the pausing site is labelled with a *, loops are T_6).

also investigated the exact pausing site by *Taq* polymerase stop assay:[25a] as seen in Figures 4G, S4 and S5, no pausing sites were detected with the scrambled sequence in the presence of both **TrisNP** and **5**, while a stop site was clearly identified one nucleotide before the TWJ structure (indicated with a * on both gel and the schematic representation on the right, Figure 4G), particularly with **5**. This assay thus confirmed that a ligand-stabilized TWJ does indeed form an efficient roadblock to the polymerase progression.

Structural insights into the TWJ/ligand interaction. To rationalize these results, we aimed at gaining accurate structural insights into how these ligands, *i.e.*, **3** (**TrisNP-*amphi***), **5** (**TrisNP-*ana***) and **10** (**TrisPFP**), along with **TrisNP** and **TrisPOB** as controls, interact with TWJs. To this end, molecular dynamics (MD) simulations were undertaken, after docking the structurally optimized ligand in a TWJ structure (built from PDB ID: 2ET0).[27] Ten independent classical-MD simulations of 2 μs were performed for each ligand. One of them was clustered; the representative structures are shown in Figure 5A (and Figure S6). Several metrics (averaged over 20 μs) were used to characterize the TWJ/ligand interactions (Table S8). The TWJ structure alone was found to be highly dynamic, as illustrated by the constant breathing of the base pairs at the junction point: the two G≡C pairs persisted during 25–26 % of the simulation time, while the A=T pair during 5 % only. The presence of a ligand within the cavity affects these dynamics: all ligands stabilize the cavity but not to the same extent, as demonstrated by the higher persistence rate of the A=T pair, which increased to 6 % only with **TrisPOB** and **5** and to 8–10 % for the other ligands. The two G≡C pairs persisted during 30–50 % of the simulations with the ligands. Importantly, some of the ligands induced a base pair disruption (indicated by green arrows in Figure 5A) with the inclusion of one of the nucleobases inside the ligand (red arrows, quantified by the inclusion rate, %), which originates in the ability of some azacryptands to interact with unpaired nucleotides. The inclusion rate is very high for **TrisNP** (71 %), significant for **5**, **9** and **TrisPOB** (14–17 %) and weak for **3**, **10** and **11** (5–10 %). The comparison between **TrisNP** and two of its isomers **3** and **5** is striking (Figure 5A), demonstrating how the modulation of the naphthalene connection (*pros-* vs. *amphi-* vs. *ana-*)[28] impacts the TWJ/ligand interactions (cavity-preserving *vs.* -disrupting binding mode). Interestingly, a good negative correlation ($R = -0.8$, Figure S7) was found between the root mean square



	Alone	TrisNP	3	5	TrisPOB	9	10	11
AT base pair / inclusion	5% / -	10% / 71%	9% / 7%	6% / 17%	6% / 14%	9% / 16%	9% / 10%	8% / 5%

Figure 5. A) Representative conformations of **TrisNP**, **3** and **5** bound in a TWJ obtained *via* molecular dynamics (MD, representative structure after clustering one 2-µs-long simulation), along with the quantification ($n=10$) of AT base pair and of the inclusion complex persistence as a percentage of the total MD simulation time. The transient inclusion complex in which nucleobases are sandwiched inside **TrisNP** is indicated with a red arrow, the unpaired nucleobases with green arrows (the structures of the complexes with **TrisPOB**, **9**–**11** can be seen in Figure S5). B) Imino regions of ^{1}H NMR spectra of 100 µM TWJ before and after the addition of one molar equivalent of ligands (**TrisNP**, **TrisPOB**, **3**, **5**, **10**). The spectra were acquired at 37 °C *in vitro* (130 mM KCl, 20 mM KPOi, pH 7.2) or in crude homogenate from HeLa cells. C) Antiproliferative activity of **TrisNP**, **3** and **5** against cervical cancer cells (HeLa) evaluated by the SRB assay ($n \geq 3$).

deviation (RMSD) of ligand-stabilized TWJs (which measures the spatial difference with respect to a reference structure along the MD trajectories) and $\triangle T_{1/2}$ values (*i.e.*, thermal stability imparted by the ligand). This indicates that the best ligands (*i.e.*, highest $\triangle T_{1/2}$ values) are those whose binding minimally distort the overall TWJ structure (*i.e.*, lowest RMSD values), even if they minimally or partly disrupt base pairs at the cavity.

Target engagement in the presence of cellular components. We then assessed the capacity of selected ligands to bind to TWJs in the presence of competing cellular metabolites by NMR spectroscopy. To this end, ^{1}H NMR spectra of 1:1 ligand:TWJ mixtures were first recorded in a buffer solution (130 mM KCl, 20 mM KPOi, pH 7.2) and then in crude cellular homogenates from HeLa cells.[29] The experiments were performed with **3** (**TrisNP-*amphi***), **5** (**TrisNP-*ana***), **9**, **10**, **11**, along with **TrisNP** and **TrisPOB** as controls. In the absence of ligands, the TWJ spectrum exhibits two degenerate signals at ~12.9 (the characteristic position of imino protons involved in G≡C base pairs) and 13.8 ppm (the imino protons engaged in A=T base pairs) (Figure 5C). The broad, unresolved nature of these signals aligns with the results from MD simulations, indicating the highly dynamic nature of TWJ stems, likely originating from base-pair breathing associated with imino proton chemical exchange. Upon interaction with ligands, a higher number of signals appears, indicating that the molecules stabilize the structure (or, better put, decrease its dynamics), confirming strong interaction in these conditions.

Interestingly, the results obtained in crude cell lysates fully parallel those collected in the buffer: the only difference is the resolution of the signals, found to be lower in cell extracts, which is due to the increase in both viscosity and crowding (proteins, genomic DNA, metabolites, etc.) in cell extracts. More importantly, the presence of cellular competitors does not trigger a loss of NMR signals, thereby advocating for the excellent specificity of these azacryptands for TWJ. These results were completed with a series of NMR experiments performed with TWJs in which the three base pairs that form the walls of the central cavity were systematically modified (Table S9). The results seen in Figure S8 indicate that the ligands bind efficiently to all TWJs, irrespective of the nature of the base pairs that form their cavity. However, a more accurate analysis revealed some changes (number, size and/or shape of some peaks) indicative of subtle differences in TWJ-binding. This was further investigated by FRET-melting experiments performed with a set of four different TWJs comprising a different GC/AT base pair ratio (from 3/0 to 0/3) at the cavity site (Table S1). The results seen in Figure S9 (Tables S10 and S11) confirm these differences, with a globally lower affinity for homo- (3/0 and 0/3) *vs.* hetero-cavities (2/1 and 1/2). This trend was weak for good binders (**3** and **5**, average $\Delta T_{1/2}=9.1$ *vs.* 11.6 °C for homo- and

hetero-cavities, respectively; 22 % difference) to notable for weak binders (**9**–**11**, average $\Delta T_{1/2} = 1.6$ *vs.* 4.3 °C for homo- and hetero-cavities, respectively; 63 % difference). A possible explanation could be that the cavity formed only by AT base pairs is too flexible and the one formed only by GC base pairs is too rigid to well accommodate a ligand, whereas a fine balance between flexibility (AT) and stability (GC) results in cavities well-suited to host a ligand.

First insights into the cellular properties. Identified as the most promising ligands, **TrisNP-*amphi*** (**3**) and **TrisNP-*ana*** (**5**) were selected for a first assessment of their cellular activity. To this end, their cytotoxicity, along with **TrisNP** as control, was tested in HeLa cancer cells (Figure 5C) as well as non-cancer cells (immortalized human fibroblasts BJ-hTERT and WI38-hTERT, Figure S10 and Table S12). The collected results indicate that both **3** and **5** are more active against HeLa cells than **TrisNP** ($GI_{50} = 0.05$ and 0.25 *vs.* 0.37 μM, respectively) and that the cancer/non-cancer activity ratios are in favor of **3** (up to 18-fold difference) and **5** (30-fold) *vs.* **TrisNP** (9-fold), which emphasizes the interest of these new derivatives.

Conclusions

Research in the biology of repetitive DNA sequences, abundant in our genome, is hampered by their propensity to fold into non-B-DNA structures known to be reluctant to DNA sequencing. This property makes the very nature of these DNA sequences the biggest challenge to their own genomic analysis. Chemical biology uniquely provides an opportunity for probing them in a cellular context, *e.g.*, to assess their involvement in cellular circuitries and, consequently, their relevance as putative targets for a therapeutic intervention. For this approach to be successful, efficient molecular tools must be used, eliciting both high affinity and high selectivity for their DNA targets. While sequence-specific targeting with small molecules is challenging because of their limited contacts with DNA, structure-specific targeting is strategically wiser as non-B-DNA structures offer a wider range of structurally well-defined binding sites for small molecules.

DNA junctions, more precisely TWJs, are ideal candidates for such a chemical biology quest: they fold from sequences that are not only abundant (*ca.* 90 motifs per 100 kb for a size ranging from 10 to 300 nt) but also located in genomic regions of interest (being enriched in heterochromatin domain comprising transcriptionally silent arrays of DNA satellites from centromeric and pericentromeric regions); they display a well-defined 3D shape readily targetable by small molecules (notably within the prismatic-shaped cavity formed at the junction of the three duplex arms) and ligand-stabilized TWJs were shown to trigger extensive DNA damage that can be potentiated by DDR inhibitors (chemically induced synthetic lethality). TWJs stabilization by small molecules is thus a promising strategy to block DNA transactions and induce toxic DNA damage and genetic instability in cancer cells.

For this strategy to become a reality, well-designed and fully characterized TWJ ligands must be used to exploit the cellular outcomes and make them amenable to mechanistic interpretations in a reliable manner. We report here on our efforts to identify such ideal TWJ ligands, evaluated through a series of orthogonal and complementary techniques comprising affinity (competitive FRET-melting and TWJ Screen assays) and functional tests (qPCR and *Taq* polymerase stop assays), along with structural analyses (MD and NMR investigations). From these diverse assays, novel naphthalene-based ligands **TrisNP-*amphi*** (**3**) and **TrisNP-*ana*** (**5**) turn out to be the most promising candidates, as they positively scored to all *in vitro* investigations implemented here and interact with TWJs according to a well-defined and cavity-preserving binding mode.

Beyond identifying these two prototypes, the results reported herein also disclose interesting structure–activity trends. Thus, large aromatic units such as anthracene (**8**) and naphthalene (**TrisNP**, **1–5**) lead to the highest TWJ affinity (Figure 3), presumably due to their ability to stack efficiently with the base pairs at the central cavity, but also, in some cases, to undesirable interactions with G4 (making them "dual-targeting" agents)[20] and even with DS-DNA (**8**, making it a non-specific DNA-binding agent).[30] In contrast, benzene-based ligands (**9**–**14**), in particular para-substituted derivatives (**9**–**11**), seem to offer the best trade-off between TWJ affinity and selectivity over both DS and G4 DNA; however, the imparted stabilization of the TWJ is insufficient to hamper *Taq* processivity (Figure 4). Also, and quite surprisingly, the small size of both benzene and tetrafluorobenzene units (F being an isostere of H) makes them fit on one side of the cavity, where they can disrupt one base pair (up to 16 % of inclusion for **9**, Figure 5). Finally, the introduction of bulky phenyl moieties (**13**) or bromine atoms (**14**) into the benzene rings, or their replacement with pyridine analogues (**15**), reduces TWJ affinity and/or leads to interaction with both DS and G4, making these derivatives poor TWJ ligands. In addition to this enhanced selectivity, our first assessment of **TrisNP-*amphi*** (**3**) and **TrisNP-*ana*** (**5**) cytotoxic activity in cancer *vs.* non-cancer cell lines revealed an increased selectivity towards transformed cell lines which could be related to genomic features, such as repeated sequence amplification and aneuploidy. We have now to investigate in more detail the basis of this cellular activity to understand how they suppress cell proliferation or lead to cell death. Further studies will assess the reliability of the transition from *in vitro* to cell-based investigations and validate the predictive potential of the combined experimental and theoretical workflow described herein (as demonstrated for instance by the interesting–yet to be confirmed–correlation between qPCR stop and cell viability assays). This is mandatory to uncover ever more efficient TWJ ligands, in the aim of making them new and reliable entities within the therapeutic armamentarium to fight against cancers.

Supporting Information

The authors have cited additional references within the Supporting Information (Ref. [31–61]) and provided detailed methods and protocols plus additional information and results (Tables S1–S12), supporting Figures (Figures S1–S10) and synthesis (Schemes S1–S5) and characterization data of azacryptands (pages S26–S54).

Acknowledgements

This work was supported by the CNRS (for S.B., N.C., A.G. and D.M.), the *Agence Nationale de la Recherche* (ANR-22-CE44-0039-01) and iSITE BFC (PIA2, grant n° UB21018.-MUB.IS). This work was granted access to the HPC resources of CINES and IDRIS for N.C. under the allocation 2023–077156 made by GENCI. L.T. was supported by the project Fragment-Screen (grant ID: 101094131) and the National Institute for Cancer Research project (Program EXCELES, ID Project No. LX22NPO5102)—Funded by the European Union. We thank Dr. O. Sordet (CRCT, Toulouse, France) and Pr. R.A. Weinberg (Whitehead Institute, Cambridge, USA) for sharing cell lines and Lenka Smilkova for assistance with NMR measurements.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: DNA junctions · ligands · *in vitro* assays · molecular dynamics · NMR

[1] S. Nurk, S. Koren, A. Rhie, M. Rautiainen, A. V. Bzikadze, A. Mikheenko et al., *Science* **2022**, *376*, 44–53.
[2] S. J. Hoyt, J. M. Storer, et al., *Science* **2022**, *376*, eabk3112.
[3] S. Kit, *J. Mol. Biol.* **1961**, *3*, IN711–IN712.
[4] J. Thakur, J. Packiaraj, S. Henikoff, *Int. J. Mol. Sci.* **2021**, *22*, 4309.
[5] G. Wang, K. M. Vasquez, *Nat. Rev. Genet.* **2023**, *24*, 211–234.
[6] W. M. Guiblet, M. A. Cremona, M. Cechova, R. S. Harris, I. Kejnovska, E. Kejnovsky, K. Eckert, F. Chiaromonte, K. D. Makova, *Genome Res.* **2018**, *28*, 1767–1778.
[7] A. N. Blackford, S. P. Jackson, *Mol. Cell* **2017**, *66*, 801–817.
[8] a) J. Zell, F. Rota Sperti, S. Britton, D. Monchaud, *RSC Chem. Biol.* **2021**, *2*, 47–76; b) G. Wang, K. M. Vasquez, *DNA Repair* **2014**, *19*, 143–151; c) C. Mellor, C. Perez, J. E. Sale, *Crit. Rev. Biochem. Mol. Biol.* **2022**, *57*, 412–442; d) G. Matos-Rodrigues, J. A. Hisey, A. Nussenzweig, S. M. Mirkin, *Mol. Cell* **2023**, *83*, 3622–3641.
[9] K. T. McQuaid, A. Pipier, C. J. Cardin, D. Monchaud, *Nucleic Acids Res.* **2022**, *50*, 12636–12656.
[10] Y. Wang, D. J. Patel, *Biochem.* **1992**, *31*, 8112–8119.
[11] a) K. Gehring, J.-L. Leroy, M. Guéron, *Nature* **1993**, *363*, 561–565; b) S. Tao, Y. Run, D. Monchaud, W. Zhang, *Trends Genet.* **2024**, https://doi.org/10.1016/j.tig.2024.05.011
[12] I. Georgakopoulos-Soares, S. Morganella, N. Jain, M. Hemberg, S. Nik-Zainal, *Genome Res.* **2018**, *28*, 1264–1271.
[13] a) S. Neidle, *J. Med. Chem.* **2016**, *59*, 5987–6011; b) H. Xu, L. H. Hurley, *Bioorg. Med. Chem. Lett.* **2022**, *77*, 129016.
[14] E. Ivens, M. M. D. Cominetti, M. Searcey, *Bioorg. Med. Chem.* **2022**, *69*, 116897.
[15] J. Novotna, A. Laguerre, A. Granzhan, M. Pirrotta, M.-P. Teulade-Fichou, D. Monchaud, *Org. Biomol. Chem.* **2015**, *13*, 215–222.
[16] K. Duskova, J. Lamarche, S. Amor, C. Caron, N. Queyriaux, M. Gaschard, M.-J. Penouilh, G. de Robillard, D. Delmas, C. H. Devillers, A. Granzhan, M.-P. Teulade-Fichou, M. Chavarot-Kerlidou, B. Therrien, S. Britton, D. Monchaud, *J. Med. Chem.* **2019**, *62*, 4456–4466.
[17] K. Duskova, P. Lejault, É. Benchimol, R. Guillot, S. Britton, A. Granzhan, D. Monchaud, *J. Am. Chem. Soc.* **2020**, *142*, 424–435.
[18] J. Malina, H. Kostrhunova, P. Scott, V. Brabec, *Nucleic Acids Res.* **2023**, *51*, 7174–7183.
[19] A. Alcalde-Ordóñez, N. Barreiro-Piñeiro, B. McGorman, J. Gómez-González, D. Bouzada, F. Rivadulla, M. E. Vázquez, A. Kellett, J. Martínez-Costas, M. V. López, *Chem. Sci.* **2023**, *14*, 14082–14091.
[20] J. Zell, K. Duskova, L. Chouh, M. Bossaert, N. Chéron, A. Granzhan, S. Britton, D. Monchaud, *Nucleic Acids Res.* **2021**, *49*, 10275–10288.
[21] a) A. Granzhan, D. Monchaud, N. Saettel, A. Guédin, J.-L. Mergny, M.-P. Teulade-Fichou, *J. Nucleic Acids* **2010**, *2010*, 460561; b) A. Pruška, J. A. Harrison, A. Granzhan, A. Marchand, R. Zenobi, *Anal. Chem.* **2023**, *95*, 14384–14391.
[22] V. Amendola, G. Bergamaschi, A. Miljkovic, *Supramol. Chem.* **2018**, *30*, 236–242.
[23] S. S. R. Namashivaya, A. S. Oshchepkov, H. Ding, S. Förster, V. N. Khrustalev, E. A. Kataev, *Org. Lett.* **2019**, *21*, 8746–8750.
[24] L. Guyon, M. Pirrotta, K. Duskova, A. Granzhan, M.-P. Teulade-Fichou, D. Monchaud, *Nucleic Acids Res.* **2018**, *46*, e16.
[25] a) J. Jamroskovic, I. Obi, A. Movahedi, K. Chand, E. Chorell, N. Sabouri, *DNA Repair* **2019**, *82*, 102678; b) J. Mitteaux, P. Lejault, F. Wojciechowski, A. Joubert, J. Boudon, N. Desbois, C. P. Gros, R. H. E. Hudson, J.-B. Boulé, A. Granzhan, D. Monchaud, *J. Am. Chem. Soc.* **2021**, *143*, 12567–12577.
[26] a) M. Gasiorek, H.-J. Schneider, *Chem. Eur. J.* **2015**, *21*, 18328–18332; b) M.-P. Teulade-Fichou, M. Fauquet, O. Baudoin, J.-P. Vigneron, J.-M. Lehn, *Bioorg. Med. Chem.* **2000**, *8*, 215–222; c) A. Pipier, A. De Rache, C. Modeste, S. Amrane, E. Mothes-Martin, J.-L. Stigliani, P. Calsou, J.-L. Mergny, G. Pratviel, D. Gomez, *Dalton Trans.* **2019**, *48*, 6091–6099.
[27] A. Oleksi, A. G. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M. J. Hannon, M. Coll, *Angew. Chem. Int. Ed.* **2006**, *45*, 1227–1231.
[28] H. Wales, *Ind. Eng. Chem.* **1922**, *14*, 317–318.
[29] P. Viskova, D. Krafcik, L. Trantirek, S. Foldynova-Trantirkova, *Curr. Protoc. Nucleic Acid Chem.* **2019**, *76*, e71.
[30] A. Granzhan, E. Largy, N. Saettel, M. P. Teulade-Fichou, *Chem. Eur. J.* **2010**, *16*, 878–889.

Manuscript received: May 23, 2024
Accepted manuscript online: June 14, 2024
Version of record online: August 1, 2024