Chemistry–A European Journal
Full Paper
doi.org/10.1002/chem.202005026
Chemistry Europe
European Chemical Societies Publishing

Quadruplex–Duplex Junctions | *Very Important Paper*

De Novo Design of Selective Quadruplex–Duplex Junction Ligands and Structural Characterisation of Their Binding Mode: Targeting the G4 Hot-Spot

Laura Díaz-Casado[+],[a] Israel Serrano-Chacón[+],[b] Laura Montalvillo-Jiménez,[a] Francisco Corzana,[c] Agatha Bastida,[a] Andrés G. Santana,*[a] Carlos González,*[b] and Juan Luis Asensio*[a]

In memory of Enrique Pedroso, our dear colleague and friend

Abstract: Targeting the interface between DNA quadruplex and duplex regions by small molecules holds significant promise in both therapeutics and nanotechnology. Herein, a new pharmacophore is reported, which selectively binds with high affinity to quadruplex–duplex junctions, while presenting a poorer affinity for G-quadruplex or duplex DNA alone. Ligands complying with the reported pharmacophore exhibit a significant affinity and selectivity for quadruplex–duplex junctions, including the one observed in the HIV-1 LTR-III sequence. The structure of the complex between a quadruplex–duplex junction with a ligand of this family has been determined by NMR methods. According to these data, the remarkable selectivity of this structural motif for quadruplex–duplex junctions is achieved through an unprecedented interaction mode so far unexploited in medicinal and biological chemistry: the insertion of a benzylic ammonium moiety into the centre of the partially exposed G-tetrad at the interface with the duplex. Further decoration of the described scaffolds with additional fragments opens up the road to the development of selective ligands for G-quadruplex-forming regions of the genome.

Introduction

G-quadruplexes are non-canonical DNA structures resulting from the stacking of Hoogsteen paired G-tetrads.[1] These motifs are known to be present in telomeric[2,3] and promoter regions[4] of the eukaryotic genome, controlling its stability and playing key roles in transcriptional regulation.[5,6] Indeed, during the past decades, G-quadruplexes have become attractive targets for the development of new drugs[7,8] for a number of pathologies such as cancer,[9] infective[10] and neurodegenerative diseases.[11] Unfortunately, compounds that bind G-quadruplexes with high affinity do not usually have a high selectivity for a particular topology and several approaches have been proposed to tackle this problem.[12] Among them, a promising strategy is targeting the interface formed by quadruplexes with adjacent double-stranded regions.[13] As DNA is mainly a B-form double helix in the cell, local formation of non-canonical structures entails the formation of interphases or junctions between DNA regions with different secondary structures. According to this reasoning, quadruplex–duplex junctions (QDJs) must be common in the genome and a relevant example has been identified in potential pharmacological targets, such as the viral HIV LTR-III sequence.[14] Moreover, in recent years, several structures of quadruplex–duplex junctions have been determined by NMR spectroscopy[14–17] and X-ray crystallography.[13a] Disappointingly, little progress has been made in the design of selective binders of these or other quadruplex–duplex junctions. Thus, molecular modelling studies on the potential binding modes of several G-quadruplex ligands have been described by Parkinson et al., although no experimental evidence of selective interaction was documented.[13a] While preparing the submission of this manuscript, the tight association of indoloquinoline-derived ligands to particular junction structures was reported by Weisz et al.[17] However, the structural basis for this interaction is presently unknown. This pio-

[a] L. Díaz-Casado,[+] Dr. L. Montalvillo-Jiménez, A. Bastida, Dr. A. G. Santana, Dr. J. L. Asensio
Glycochemistry and Molecular Recognition group—Dpt. Bio-Organic Chemistry
Instituto de Química Orgánica General (IQOG-CSIC)
Juan de la Cierva 3. 28006 Madrid (Spain)
E-mail: andres.g.santana@csic.es
juanluis.asensio@csic.es

[b] I. Serrano-Chacón,[+] Prof. Dr. C. González
Instituto de Química-Física Rocasolano (IQFR-CSIC)
Serrano 119. 28006 Madrid (Spain)
E-mail: cgonzalez@iqfr.csic.es

[c] F. Corzana
Department of Chemistry
Centro de Investigación en Síntesis Química
Universidad de La Rioja
Madre de Dios, 53. 26006 Logroño (Spain)

[+] These authors contributed equally to this work.

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/chem.202005026.

neering contribution constitutes the first and only experimental study on this topic reported so far.

To address this relevant issue, we have employed a multidisciplinary approach that combines NMR spectroscopy with molecular modelling and binding studies. Our design strategy was based on the observation that the centre of the G-tetrad at the interface with the duplex, a region characterised by a strongly negative electrostatic potential, is often partially accessible in junction architectures. We hypothesised that this motif, whose binding properties should be strongly modulated by the adjacent duplex extensions, could represent a hitherto unexplored hot-spot for the selective recognition of quadruplex–duplex junctions. This idea has been examined by employing all the quadruplex–duplex junction structures reported in the bibliography. Among them, some of the architectures described by Phan et al. were selected as the main model system for our study (QDJ1–QDJ5).[15a]

They comprise a two-G-tetrad antiparallel quadruplex connected with a long stem-loop hairpin, as a duplex extension, leading to a partially exposed interfacial G-tetrad suitable for ligand targeting. The G-quadruplex moiety is structurally very similar to the well-known thrombin binding aptamer (TBA)[18] in which the three-residue axial loop is replaced by a longer stem-loop hairpin.[19] We further validated our results with other junctions available in the literature, including a potential pharmacological target: the HIV-1 LTR-III sequence.[14,20–23]

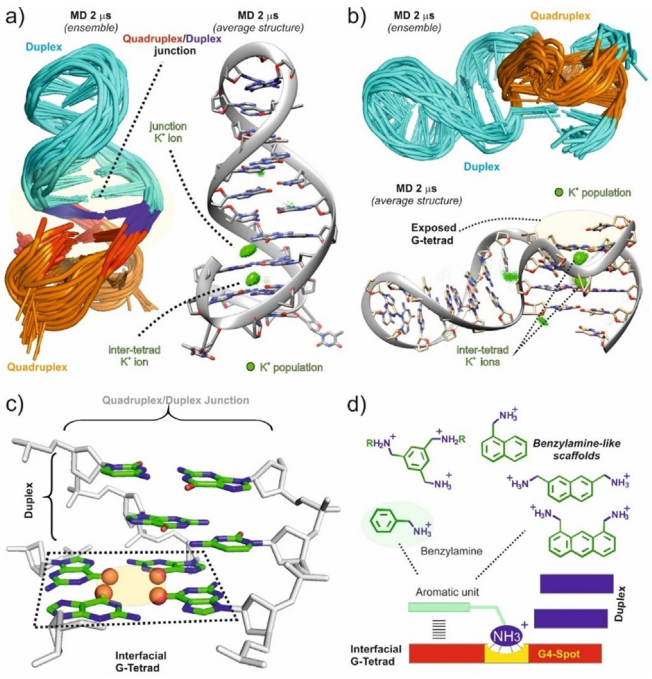

Figure 1. Representative ensembles obtained for QDJ1 (a) and QDJ5 (b) by employing MD simulations. Average structures are shown in grey. Regions characterised by a large density of potassium cations, as revealed by the analysis of the MD trajectories with Chimera, are highlighted in green. (c) Representation of the quadruplex–duplex junction structure in QDJ1. The centre of the interfacial G-tetrad is highlighted with a yellow shading, where guanine O6 atoms are represented as red spheres. (d) Schematic representation of the proposed binding mode for a benzylamine-like pharmacophore targeting the G4 spot together with examples of putative ligands.

Results and Discussion

Cation-binding properties of quadruplex–duplex interfaces and design of a selective pharmacophore

In contrast with occluded or fully exposed G-tetrads characteristic of common quadruplex regions, those located at the interface between quadruplex and duplex fragments might display distinctive electrostatic properties, and thus are potential hot-spots for the molecular recognition of properly designed cationic ligands. To test this idea, we performed molecular dynamics simulations employing two different QDJ structures reported by Phan et al. and herein referred to as QDJ1 and QDJ5 (Figure 1a and b).[15a] The relative orientation of the duplex and quadruplex moieties is co-axial in QDJ1, which determines the presence of an interfacial tetrad partially occluded by the duplex extension. On the contrary, this motif is absent in QDJ5. Long MD trajectories collected in the presence of potassium ions revealed distinct electrostatic properties in both junctions. Whereas QDJ1 displays a high density of cations at the interface between the quadruplex and the duplex stem, a low density of ions is observed in QDJ5, either at the terminal tetrads or duplex base pairs. This observation supports the notion that the electrostatic properties of terminal tetrads can be affected by co-axial duplex extensions, generating a singular recognition site. Targeting these epitopes would require the displacement of the bound potassium, which poses an energy cost on the recognition process. However, this could be compensated by additional stacking and hydrogen bonding interactions, leading to high-affinity ligands. Thus, geometrical and chemical considerations helped to conceive a general pharmacophore represented by chemically accessible highly versatile frameworks integrating an electron-rich aromatic unit attached to a methylene-amine moiety (Figure 1d). Junction recognition would benefit from two complementary modes of interaction: stacking of the aromatic platform onto the π-deficient surface of the G-tetrad at the interface with the duplex plus insertion of the ammonium moiety into the G4 centre, referred to as the G4 hot-spot throughout this manuscript (Figure 1d). To test this idea, we performed binding experiments with bare versions of the proposed pharmacophore as well as more complex synthetic derivatives thereof.

Binding experiments with model QDJs

As a proof of concept, we assayed first the association of the simplest fragment that incorporates the proposed pharmacophore, that is, benzylamine **1**, to the oligonucleotide QDJ1. We employed NMR spectroscopy, as this technique is especially suitable for the detection and characterisation of low-affinity binding processes (Figure 2a). Thus, according to titration experiments performed at 278 K, **1** induced clear chemical shift perturbations in the DNA imino region. Satisfactorily, these are consistent with the selective recognition of the quadruplex–duplex interface, with major $\Delta\delta$ values located at the junction duplex base pair G21/C7 and C20/G8. The change in $\Delta\delta$ values

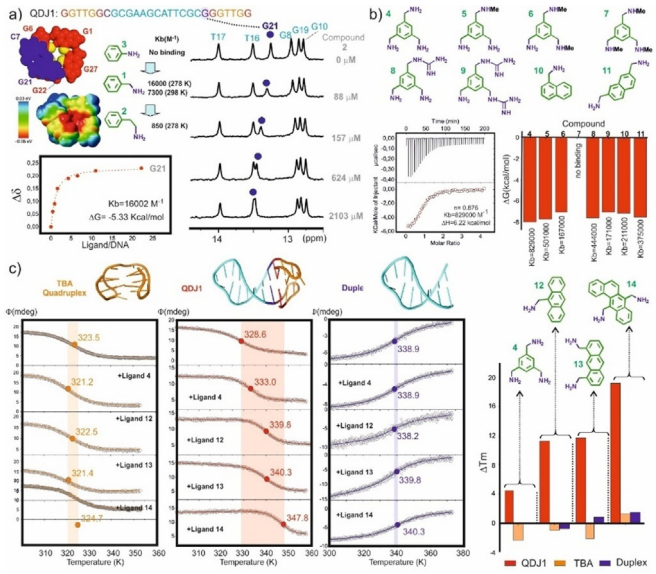

Figure 2. (a) Low-affinity scaffolds. K_b values for the association of **1**–**3** to QDJ1 at 278 K derived from NMR spectroscopy. Junction residues (CPK) and its electrostatic surface (only bases considered) are shown. Binding curve for **1** together with 1D NMR data from the corresponding titration are also represented. (b) Medium-affinity ligands. Binding data derived for the association of selected aromatic ligands incorporating benzylamine-like fragments to QDJ1. K_b (M^{-1}) values for compounds **4**–**11** were obtained from ITC (298 K) experiments (ITC data measured with **4** is shown on the left). (c) High-affinity ligands and preliminary selectivity assays. Typical CD thermal denaturation curves measured for quadruplex TBA (orange; left panels), QDJ1(red; middle panels) and the duplex fragment present in QDJ1 (violet; right panels) in the absence and presence of ligands **4** and **12**–**14** (from top to bottom). The observed shifts in melting temperatures are highlighted with a coloured shading and represented on the right.

along the titration experiments confirmed that benzylamine (**1**) binds to oligo QDJ1 with a $K_b = 16\,002\ M^{-1}$, corresponding to a ΔG value of -5.3 kcal mol^{-1} at 278 K, a significant free energy of interaction for such a limited set of contacts (values at 298 K are also shown in Figure 2a). Significantly, homologation or shortening of the aromatic/ammonium linker (as in phenethylamine **2** or aniline **3**) translates into a dramatic decrease in binding affinity, ranging from 19-fold for the case of **2** to a virtually undetectable association for **3** (Figure 2a and Figure S1 in the Supporting Information). These preliminary observations are fully consistent with the proposed binding mode and encouraged us to pursue this line of research.

Next, and as a straightforward strategy to increase the stability of the complex, we prepared a multivalent derivative equipped with three alternating methylene-amine fragments attached to a single phenyl ring (compound **4**). NMR titration assays (Figure S2 in the Supporting Information) confirmed the enhanced affinity of this scaffold, which prompted us to employ microcalorimetry to fully dissect the association process (Figure 2b). According to these data sets QDJ1 recognition by **4** is exothermic and characterised by a K_b value in the $10^6\ M^{-1}$ range. Concomitant with our rationale, *N*-methylation and *N*,*N*′-dimethylation of **4** (compounds **5** and **6**) preserves the ligand's capacity to selectively recognise the junction structure. However, the introduction of a third *N*′′-Me group (**7**) abolishes the association, thus proving the need for at least one primary NH_3^+ moiety to effectively interact with the G4 hot-spot. According to this trend, guanidinylated derivatives **8**–**9** were synthesised, showing a similar affinity pattern (see Figure 2b and Figures S3–S4 in the Supporting Information). Finally, the parallel stacking between complementary π-rich and π-deficient aromatic surfaces proved to be an important contributor to the free energy of binding. Indeed, the obtained ITC data showed that increasing the π-rich aromatic surface of the ligand enhances the complex stability. Thus, compounds **10** and **11**, with just one or two positive charges, respectively, display binding affinities comparable to that measured for **4** (Figure 2b and Figure S5 in the Supporting Information). Further extension of the aromatic platform with additional fused phenyl rings, as in **12**–**14**, renders these derivatives too hydrophobic for microcalorimetry experiments. However, CD melting experiments confirmed these derivatives bind to QDJ1 with much higher affinities than **4** (CD experiments with high-affinity ligands are shown in Figure 2c. Data sets measured for **4** are also shown for comparison purposes).

Selectivity for QDJs versus isolated quadruplex or duplex fragments

Chemical shift perturbations observed in NMR titration experiments performed with compound **4** and QDJ1 are consistent with the association of this ligand to the junction region (Figure S2 in the Supporting Information). Similar effects are observed for other compounds of the family (Figures S3–S4 in the Supporting Information). To assess the selectivity of the molecular recognition process, the interaction of the proposed pharmacophore with oligonucleotides containing the constituent duplex or quadruplex moieties present in QDJ1 was explored by different techniques. The stem-loop hairpin sequence $^{5'}$CGCGAAGCATTCGCG$^{3'}$ was used as a model for the duplex. Similarly, the TBA sequence ($^{5'}$GGTTGGTGTGGTTGG$^{3'}$) was used as a model of the quadruplex fragment considering its structural similarity with the one present in QDJ1. Most importantly, melting curves acquired with these quadruplex and duplex fragments confirmed an overwhelming preference for the junction structure. Additional evaluation of the selectivity provided by this recognition motif was derived from micro-dialysis experiments (Figures S6–S7 in the Supporting Information, and see the Experimental Section).[24] The results obtained demonstrate that **4** recognises the quadruplex–duplex interface with exquisite selectivity *with respect to the constituent duplex and quadruplex fragments alone*. Indeed, numerical simulations of this experiment indicate that K_b values for the duplex or quadruplex fragments must be $< 20\,000\ M^{-1}$, at least a 40-fold decrease in affinity with respect to QDJ1 (see Figures S6–S7 in the Supporting Information). This result is further supported by NMR titrations of the isolated quadruplex and duplex moieties, which revealed no significant association even at much higher concentrations of **4** (see Figure S8 in the Supporting Information).

The weak to null affinity for the TBA fragment exhibited by ligands **13** and **14** is intriguing as derivatives with similarly

large aromatic surfaces have been reported as reasonable quadruplex binders.[9] Taking this into account, we decided to extend our selectivity studies with these derivatives to other quadruplex topologies, including both parallel and antiparallel arrangements such as the proto-oncogene promoter KRAS,[25] the human Bcl2 promoter region,[26] a natural human telomere quadruplex[27] and the T30695 structure described by Hogan et al.[28] In Figure 3, it can be observed that ligand **13** promotes moderate stabilisations of several of the DNA fragments tested, with ΔT values up to 7.7 K in the most favourable situation. However, a clear preference for the QDJ1 structure is still apparent, providing further support for the junction-selective character of this derivative.

Regarding the more extended compound **14**, this time significant stabilisations of most of the analysed quadruplexes, in the 15–17 K range, were detected. Then again, the observed stabilisation of the QDJ1 fragment is higher (>19 K), a trend that is further accentuated with the HIV LTR-III junction (see below). Overall, the obtained results suggest that three fused aromatic units (such as those present in **13**) represent the appropriate size to provide optimal junction selectivity with respect to plain quadruplexes.

Although melting studies provide an indication of the relative complex stabilities, they do not constitute quantitative affinity measurements by themselves (which are otherwise difficult to perform with high-affinity, low-solubility compounds like **13** and **14**). Therefore, we carried out competition experiments by NMR spectroscopy (Figure 3). Accordingly, the QDJ1 (50 μM) fragment was titrated with 0.5 equivalents of **13** to generate a mixture of free and complexed junctions (50% each) with well-resolved NMR signals in the 12–14 ppm spectral region. Next, different quadruplex fragments (100 μM each) were subsequently added to the NMR sample and their influence on the QDJ1 free/bound equilibrium was determined through integration of the appropriate peaks. The obtained results demonstrate that these quadruplexes are unable to compete with QDJ1 for **13**, having a negligible impact on the fraction of complexed junction (an observation that is maintained after a 24 h equilibration). Notably, addition of the QDJ1 fragment onto a KRAS/**13** complex mixture immediately showed migration of the ligand from the quadruplex toward the junction structure (Figure S9 in the Supporting Information). Altogether the obtained results confirmed the significant preference of **13** for quadruplex–duplex junctions, regardless of the topology of the competing quadruplex fragments considered.

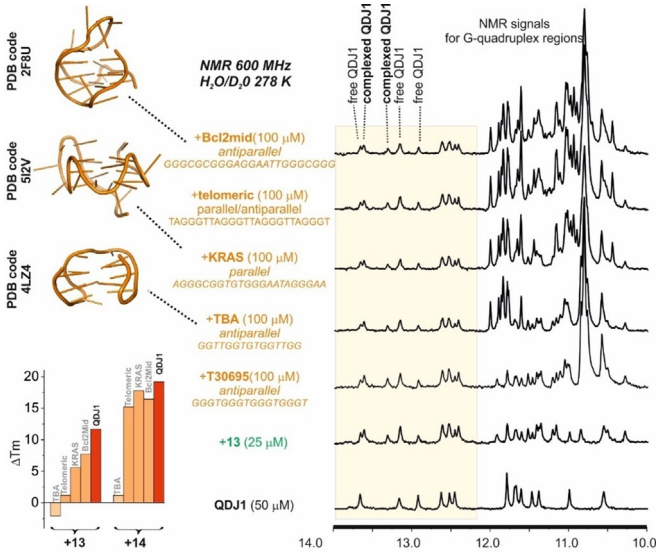

Figure 3. Selectivity assays. Left: Melting temperature perturbations measured for different quadruplex fragments (5 μM, orange) and QDJ1 (5 μM, red) in the presence of two equivalents of ligands **13** (left) or **14** (right) as determined from CD thermal denaturation measurements. Structures for selected quadruplexes are shown above. The respective PDB codes are indicated. Right: NMR competition experiments. QDJ1 (50 μM) was titrated with 0.5 equivalents of **13** to generate a mixture of free and complexed junctions (50% each) and different competing quadruplexes (100 μM each) were subsequently added. Free/bound equilibrium for the junction fragment was evaluated by integration of the appropriate NMR signals (yellow-shaded area).

Structure determination of a QDJ/ligand complex

To get further insight into the interaction mode of the analysed scaffolds with DNA QDJs, we resorted to NMR methods. Complete assignment of the NMR spectra of QDJ1/**4** and QDJ1/**13** complexes was carried out (Figures S10–S12; assignment Tables S1 and S2 are shown in Supporting Information). Chemical shift perturbations promoted by both ligands on the DNA receptor were similar (Figure S13 in the Supporting Information) and are limited to the residues in the quadruplex–duplex junction, consistently with a unique binding mode in this region. Especially illustrative is the strong change of chemical shifts of C20 base protons upon ligand binding (Figure 4a).

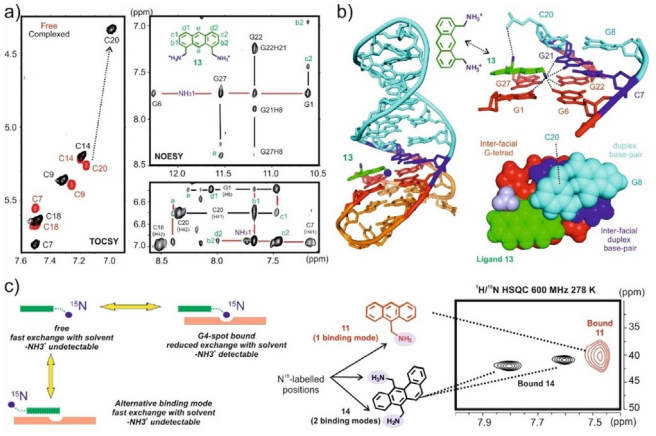

Figure 4. (a) NMR spectra acquired in potassium phosphate buffer, 20 mM KCl at pH 6.9. DNA and ligand concentrations were 500 and 600 μM, respectively. Left: Cytosine H5/H6 region in TOCSY experiments (298 K) measured with QDJ1 free (red) and complexed to **13** (black). Chemical shift perturbations detected at residue C20 have been emphasised. Right: Two different sections of a NOESY (278 K) experiment measured for the QDJ1/**13** complex are shown. NOE cross-peaks involving the ligand key ammonium group are represented. (b) NMR structure of the QDJ1/**13** complex (PDB 6FC9). Observed ligand/DNA interactions are highlighted on the top-right corner. A CPK representation of the junction region (top view) is shown on the bottom-right corner. (c) ^{1}H-^{15}N HSQC experiments for the QDJ1 complexes with ^{15}N-labelled derivatives **11** and **14** (see the main text).

Unfortunately, ligand signals in QDJ1/**4** complex could not be identified. This is probably a consequence of an intermediate exchange rate between the three equivalent poses in which a three-fold symmetric compound can be accommodated in the binding pocket. However, signals of compound **13** were clearly identified in the NMR spectra of the QDJ1/**13** complex, for which more than 20 intermolecular cross-peaks were observed (Table S3 in the Supporting Information). The solution structure of this complex was determined by restrained molecular dynamics methods on the basis of 232 experimental distance constraints. The structure is well-defined with an RMSD for the heavy atoms of 0.6 Å in less mobile regions. Complex coordinates were deposited in the PDB (code 6FC9). Statistical data are summarised in Table S4 and the conformational ensemble shown in Figure S14 (in the Supporting Information).

The obtained DNA structure is very similar to that of the free oligonucleotide, being the complex stabilised by two key interactions involving the interfacial tetrad. Although the ligand aromatic moiety stacks on the exposed guanines of the G-tetrad, one of the ligand ammonium groups points towards the tetrad central position (the proposed G4 hot-spot), interacting with the electron-rich guanine O6 groups (Figure 4b and Figure S15 in the Supporting Information). This interaction is firmly supported by a number of NOEs between the ammonium signal at 7.6 ppm with imino protons of G1, G6, G22 and G27 (Figure 4a). The fact that this ammonium signal is unusually narrow and that it is observed even at relatively high temperature (298 K) indicates that this functional group is not exposed to the solvent. Ligand contacts with the duplex region also contribute to the stability of the complex, in particular a partial stacking with C20 and a salt bridge with the phosphate of C20 (see Figure 4b). These structural data are fully consistent with the binding assays previously described with alternative ligands (a model of the QDJ1/**4** complex generated by docking plus molecular dynamics simulations, also consistent with all the experimental information herein reported as shown in Figure S16 in the Supporting Information).

To assess the validity of this general binding mode for other junction ligands, selected compounds were labelled with ^{15}N and $^{1}H/^{15}N$ HSQC experiments were recorded on the corresponding DNA complexes (Figure 4c). The rationale behind these assays is that fast exchange with solvent usually precludes detection of ammonium groups in water solution unless this process is strongly slowed down by tight binding interactions with the DNA receptor. Indeed, this behaviour would be expected for the proposed complexes, in which the ammonium moiety is inserted into the centre of the QDJ interfacial G-tetrad, establishing numerous hydrogen bonding and electrostatic interactions. On the contrary, alternative, non-specific binding modes might prove less efficient at slowing down the ammonium exchange, rendering the corresponding HSQC cross-peaks undetectable. Control assays performed with ^{15}N-labelled **11** and **14** either free or in the presence of duplex or quadruplex fragments were consistent with this view, thus failing to provide detectable cross-peaks. In contrast, clear HSQC signals were apparent for the QDJ1 complexes. Moreover, these presented downfield shifted ^{1}H and ^{15}N ammonium resonances in agreement with values reported by Plavec and coworkers for chelated intertetrad ammonium ions.[29] Interestingly, in the particular case of **14**, the obtained spectrum revealed the co-existence of a mixture of two complexes, which probably reflects the non-symmetrical nature of this ligand. Similar experiments performed with the weaker binder **4** (and derivatives) failed to detect the desired signals presumably extremely broadened by the free/bound exchange process. However, for tight binders such as **11** and **14** this simple assay provides a signature of the ammonium/G4 spot association shared for all the analysed ligands.

Binding to different QDJs

Next, we analysed the structural/topological requirements for the recognition process to happen. To this end, binding studies were performed with **4** and different junction architectures. First, we explored the alternative structures described by Phan et al. (herein referred to as QDJ2–QDJ5; see Figure 5a and Figures S17–S20 in the Supporting Information),[15a] employing both NMR and ITC. As can be observed, oligos QDJ1–QDJ4 display a similar coaxial orientation of quadruplex and duplex domains (Figure 5). However, although QDJ1 and QDJ2 exhibit a partially exposed G4 spot, this site is occluded in QDJ3 and QDJ4. Consistently, the latter provided no detectable associa-

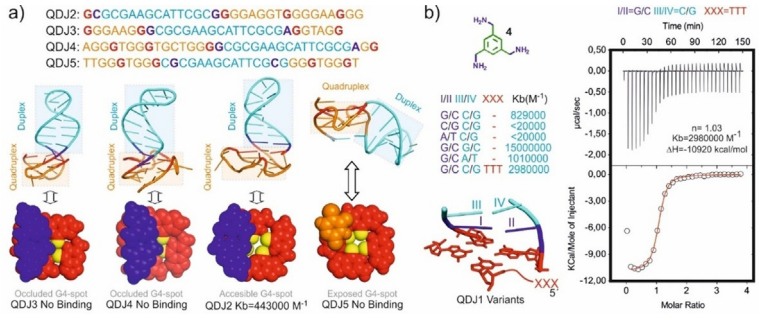

Figure 5. (a) Binding studies (298 K) performed by ITC with ligand **4** and oligos QDJ2–QDJ5 (sequences shown above). Overall topology (up, ribbon) and junction residues (down, CPK) are represented. G4 spot oxygen atoms are coloured in yellow (structures are ordered according to the increasing degree of G4 spot exposure). (b) K_b (M^{-1}, 298 K) values obtained by ITC with ligand **4** and mutated QDJ1 variants. Substituted residues at the interfacial and adjacent duplex base pairs (I/II and III/IV, respectively) are indicated. ITC data measured with a selected variant, bearing a TTT overhang at the 5′ end is represented on the left.

tion. In contrast, QDJ1 and QDJ2 displayed comparable stabilities. Regarding QDJ5, as mentioned earlier, this oligonucleotide exhibits a radically different topology, characterised by an orthogonal orientation between the duplex and the quadruplex domains. As a result, the terminal G-tetrad is exposed with its central cavity virtually unaffected by the duplex regions (see Figures 1b, 5a and S20 in the Supporting Information). Interestingly, in this particular case no association to the fully exposed G4 hot-spot was detected either.

In conclusion, neither an occluded (QDJ3-4) nor a totally exposed (QDJ5) G4 spot allows complex formation at the junction binding-site with a benzylamine-like probe. Enhancement of the G4 hot-spot binding properties by adjacent co-axial duplex regions seems to be an essential requirement for tight association, which provides significant opportunities for the design of selective binders. These results have been confirmed with other junctions available in the literature. Thus, binding of **4** to the junction studied by Karg et al.,[16] which exhibits a partially occluded G4 hot-spot, can be observed by NMR titration experiments (Figure S22 in the Supporting Information) whereas no binding is apparent with the junction studied by Russo et al.,[13a] the crystallographic structure of which shows the centre of the QDJ interfacial G-tetrad buried by a TAT triad (Figure S22 in the Supporting Information).

Following this line of thought, interfacial duplex base pairs should strongly modulate the association properties of this family of ligands to the G4 spot. Regarding the QDJ1 architecture, we have observed that swapping the G21/C7 base pair to C21/G7 or its replacement by A21/T7 led to a >40-fold decrease in binding affinity (Figure 5b and Figure S21 in the Supporting Information). Remarkably, the modulating influence of the contiguous duplex base pair (C20/G8) seems significant too, with K_b values in the $10^7\ M^{-1}$ range for the G20/C8 mutant (see Figure 5b and Figure S21 in the Supporting Information). Finally, the extension of the 5′-end (located in this case beside the QDJ interfacial G-tetrad) with a TTT overhang promotes a 3–4-fold increase in the stability of the complex ($K_b = 2980\,000\ M^{-1}$), probably by providing additional harnessing points for the ligand (Figure 5b and Figure S21 in the Supporting Information). Overall, these binding experiments point to a tighter interaction at the junction hot-spot especially when the adjacent duplex base pairs contain G-C.

Binding to biologically relevant sequences: The case of HIV-1 LTR-III

Encouraged by these results, we decided to further test the potential of the proposed pharmacophore for selective recognition of quadruplex–duplex junction architectures found in natural sequences. We focused our attention on a recently described target of great biomedical relevance: the LTR-III region of the HIV-1 virus (herein referred to as QDJ6).[14] Stabilisation of this site performed by small ligands has been proposed as a promising strategy for the inhibition of the viral transcription. Interestingly, NMR structural studies have shown that it comprises a peculiar quadruplex–duplex junction, which, in contrast to previously employed models, incorporates a highly dynamic duplex base pair at the boundary between the quadruplex and duplex fragments (Figure 6a).[14] Close inspection at the reported structure (pdb code 6H1K) revealed that this feature leads to a partially accessible interfacial G4 spot potentially targetable by the ligands, complying with the pharmacophore herein reported. Satisfactorily, NMR titration experiments clearly show that ligand **4** indeed binds to the HIV LTR-III sequence with significant affinity to form a 1:1 complex (Figure 6b). No evidence of additional binding events was observed even at high ligand/DNA ratios, which indicates that the ongoing association process is highly selective. Chemical shift changes upon complex formation were analysed on the basis of the assignments previously reported for the free DNA[14] and confirmed that protons located in the interfacial quadruplex/duplex region are the most affected by the ligand. In particular, changes in chemical shifts of base protons of C13 and T14 and imino protons of G15 and G26 are clearly observed, whereas other regions are mainly unaffected (Figure 6b,c). To demonstrate the relevance of the amino groups in the pharmacophore for QDJ6 binding, we carried out additional NMR titration experiments employing ligand **7**. As previously observed for model junction architectures QDJ1–QDJ5, *N*-methylation of the three amino moieties totally abolishes the capacity of the compound to recognise the interfacial G4 spot motif. Indeed, no association of **7** to QDJ6 was detected even at high ligand/DNA ratios. Finally, the thermodynamic

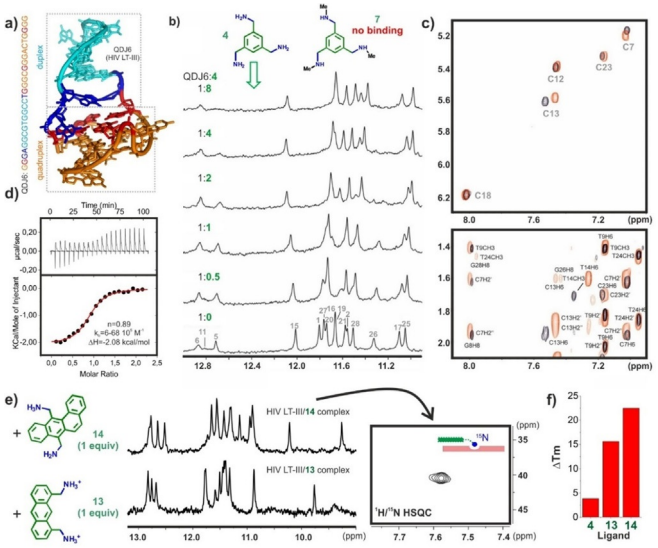

Figure 6. (a) 3D NMR structure (PDB code: 6H1K) of HIV oligonucleotide QDJ6 (ribbon representation). Non-interfacial duplex and quadruplex regions are represented in cyan and orange, respectively. Similarly, interfacial duplex base pair and G-tetrad are shown in blue and red. (b) QDJ6/**4** titration experiment. [QDJ6] = 0.1 mM, T = 298 K. (c) Top: Cytosine H5-H6 region in TOCSY experiments acquired for isolated QDJ6 (red) and 1:1 QDJ6/**4** complex (black). Bottom: NOESY spectra (150 ms mixing time) of isolated QDJ6 (red) and 1:1 QDJ6/**4** complex (black). (d) ITC profile for the titration of **4** into a solution containing QDJ6 at 298 K. (e) 1D NMR spectra for the QDJ6/**13** and QDJ6/**14** complexes (imino region) together with a $^1H/^{15}N$-HSQC spectrum for the latter acquired with a ^{15}N-labelled ligand. (f) Melting temperature perturbations deduced for QDJ6 (5 μM) in the presence of 10 μM of ligands **4**, **13** or **14** as deduced from CD thermal denaturation measurements.

features of the QDJ6/**4** complex were dissected by microcalorimetry experiments (Figure 6d). According to these data, **4** binds to QDJ6 with a significant affinity ($K_b = 6.68 \times 10^5\ M^{-1}$), corresponding to a free energy of association $\Delta G = -7.91$ kcal mol^{-1}, almost identical to that previously observed for the QDJ1/**4** interaction. However, in contrast with the latter, the binding event between **4** and QDJ6 is mainly entropy-driven, whereas it is only slightly exothermic ($\Delta H = -2.08$ kcal mol^{-1}), which might reflect the distinct dynamic behaviour exhibited by interfacial duplex base pairs in both QDJ6 and QDJ1.

Larger aromatic ligands, such as **13** and **14**, also bind to the HIV-LTIII to form 1:1 complexes. The addition of sub-stoichiometric amounts of the ligands reveal that now the exchange between free and bound DNA is slow on the chemical shift timescale. In addition, the final complexes render high quality spectra characterised by sharp and well-resolved signals in the imino region (Figure 6e). Moreover, HSQC spectra with ^{15}N-labelled ligand **14** exhibits a single cross-peak, consistent with the insertion of the ammonium group in the centre of the G-tetrad at the interface with the duplex (Figure 6e and Figure S23 in the Supporting Information). All these experimental pieces of evidence strongly suggest that both **13** and **14** present again increased affinities for the target DNA with respect to **4**. This impression was further confirmed by CD thermal denaturation experiments, which showed more significant increments in the receptor melting temperature upon formation of the complexes (Figure 6f). Indeed, the observed ΔT values are larger than those previously measured for the model junction QDJ1 (Figure 2b). In addition, they clearly surpass the thermal stabilisations previously detected with quadruplex fragments of alternative topologies (Figure 3), further supporting the junction-selective character of this pharmacophore.

Conclusion

We have described a novel binding-motif, which relies on the recognition of the strongly negative electrostatic potential of the G-tetrad centre. Although, this centre is usually occluded in most G-quadruplex structures, we have observed that it is partially exposed in some quadruplex–duplex junctions. By taking advantage of this interaction hot-spot, we have designed a family of compounds able to bind selectively to quadruplex–duplex junctions, including biomedically relevant targets as the HIV LTR-III fragment. The benzylamine-like motifs herein described are very convenient scaffolds, which can be readily conjugated with other complementary duplex or quadruplex binders, thus paving the way to the development of a whole new class of quadruplex-discriminating ligands.

Experimental Section

DNA oligonucleotides were purchased in purified desalted form from IDT (Integrated DNA technologies) and dialysed against 20 mM KCl, 20 mM K_2HPO_4 at pH 6.9 buffer before binding experiments. To ensure a proper folding of the different quadruplex, duplex or quadruplex–duplex junction structures, the corresponding oligonucleotides were subjected to an annealing protocol. For this purpose, they were placed in a water bath at 293 K, heated to 358 K for 5 min and then slowly cooled back to 293 K over a 2 h period.

Microcalorimetry (ITC) binding experiments

Binding studies were performed at 298 K, in 20 mM KCl, 20 mM K_2HPO_4 at pH 6.9, by using a VP-ITC titration calorimeter (MicroCal, LLC) with a reaction cell volume of 1.467 mL. Typically, 10 μM DNA oligonucleotide in the reaction cell was titrated with a 500–700 μM solution of the different ligands contained in a 300 μL syringe. At least 30 consecutive injections of 5 μL were applied at 5 min intervals while the DNA solution was stirred at a constant speed of 300 rpm. Dilution heats of ligand into DNA solutions (which agreed with those obtained by injections of ligands into the same volume of buffer) were subtracted from measured heats of binding. Titration curves were analysed with Origin, provided with the instrument by MicroCal LLC. For every single DNA fragment, thermodynamic parameters were derived from two independent experiments and averaged.

Nuclear magnetic resonance (NMR) binding experiments

The binding of ligands **1**–**4** to oligonucleotides QDJ1 and QDJ5 was monitored by recording 1D 600 MHz ^{1}H NMR spectra of a series of samples with variable ligand concentrations at 278 K. In all cases, the ongoing association processes were apparent from the chemical shift perturbations detected in the DNA imino region (12–15 ppm) of the receptor. The DNA samples were prepared at 100 μM concentration (calculated from the UV absorbance at 260 nm) from stock solutions in 90% H_2O:10% D_2O, 20 mM KCl, 20 mM K_2HPO_4 at pH 6.9. The ^{1}H NMR spectrum for the sample with the highest ligand/DNA ratio was recorded by dissolving the ligand (≈20 mM) in 0.5 mL of the QDJ1-QDJ5 100 μM solutions. Titration curves were obtained by adding small aliquots of this high ligand/DNA ratio sample to a ligand-free NMR sample. In all cases, the data were found to fit well assuming a 1:1 complex.

Micro-dialysis binding experiments

A solution (400 μL) containing oligonucleotide QDJ1 (50 μM) and ligand **4** (20 μM) was placed in the central compartment (herein referred to as II) of a three-chamber micro-equilibrium dialyser (Harvard apparatus) equipped with 5 kDa cut-off membranes and dialysed against the two separate DNA(50 μM)/ligand **4** (20 μM) mixtures: one containing TBA quadruplex (in chamber I) and the other with the duplex fragment present in QDJ1 (in chamber III). To achieve a complete equilibration of the three micro-dialysis compartments, this assay was left to proceed for 2 days at 298 K. Afterwards, solutions from each chamber were collected and the DNA fragments were digested with DNAase (1–2 μM) for 6 h. Next, ligand **4** present in the three samples was derivatised by treatment with ^{13}C-labelled formaldehyde (20 mM) and sodium cyanoborohydride (20 mM). After 12 h at room temperature, the solutions were loaded into NMR tubes and HSQC experiments were acquired. Integration of the single *N*,*N*-dimethyl cross-peak present in these spectra provided a simple means to determine the relative concentration of **4** in each chamber after equilibration.[24]

Micro-dialysis theoretical simulations

We performed extensive modelling studies on the micro-dialysis competition experiment described in the manuscript, employing the biochemical kinetic simulator GEPASI 21.[30] Our model com-

prised a solution, containing the quadruplex–duplex junction oligonucleotide, placed in the central chamber (herein referred to as II) of a three-compartment micro-dialysis device. Two alternative DNA solutions (one containing TBA quadruplex and the other with the duplex fragment present in QDJ1) are confined in lateral chambers I and III. Chamber volumes and DNA concentrations were fixed at 400 μL and 100 μM, respectively, in agreement with the actual conditions employed in our experimental assay. The ligand was assumed to be initially present in the three chambers at identical concentrations (20 μM). Numeric integration of the corresponding kinetic equations allowed a theoretical evaluation of the equilibrium ligand concentrations in chambers I–III by assuming a variety of binding constants for the duplex and quadruplex fragments ($K_{b\,duplex}$ and $K_{b\,quadruplex}$, respectively. For QDJ, $K_{b\,QDJ1}$ was fixed as $10^6\,M^{-1}$).

CD melting experiment

Thermal unfolding transitions were measured by employing 1 cm path length quartz cells with samples containing 5.0 μM DNA in 20 mM KCl, 20 mM K_2HPO_4 at pH 6.9. Junction-selective ligands **4** and **12**–**14** were employed at 10 μM concentration. The transitions were monitored by the decrease of the CD signal at 250 nm (duplex), 295 nm (QDJ1, HIV LTR-III quadruplex–duplex junctions and TBA, BCL2 and telomeric quadruplexes) or 265 nm (KRAS quadruplex) by using 2 nm band width. Heating rates were $20\,K\,h^{-1}$. Transitions were evaluated by using a nonlinear least squares fit assuming a two state model with pre- and post-transitional baselines sloping.

NMR spectroscopy of the complexes

Samples for NMR experiments were dissolved (in Na^+ form) in either D_2O or 9:1 H_2O/D_2O, 20 mM potassium phosphate buffer, 20 mM KCl at pH 6.9. Oligonucleotide concentration was 0.5 mM. Experiments were acquired with Bruker spectrometers operating at 600 and 800 MHz, equipped with cryoprobes and processed with the TOPSPIN software. NOESY spectra were acquired with mixing times ranging from 75 to 250 ms. TOCSY spectra were recorded with the standard MLEV-17 spin-lock sequence and a mixing time of 80 ms. In the experiments in H_2O, water suppression was achieved by including a WATERGATE[31] module in the pulse sequence prior to acquisition. The spectral analysis program SPARKY[32] was used for semiautomatic assignment of the NOESY cross-peaks and quantitative evaluation of the NOE intensities. The high similarity between the NMR spectra of free QDJ1 and its complexes with **4** and **13** allowed for a straightforward assignment of the later on the basis of the assignment published by Phan et al.[15a] Chemical shift lists are given in Tables S1–S2 (in the Supporting Information).

NMR constraints and structural calculations

Qualitative distance constraints were obtained from NOE intensities. NOEs were classified as strong, medium or weak, and distance constraints were set accordingly to 3, 4 or 5 Å. In addition to these experimentally derived constraints, hydrogen bond constrains for Watson–Crick base pairs and G-tetrads were used. Target values for distances and angles related to hydrogen bonds were set to values obtained from crystallographic data in related structures. Owing to the relatively broad line-widths of the sugar proton signals, *J*-coupling constants were not accurately measured, but only roughly estimated from DQF-COSY cross-peaks. Loose values were set for the sugar dihedral angles δ, $\nu 1$ and $\nu 2$ to constrain the deoxyribose conformation to North or South domain. A set of 232 experimental distance constraints was used to determine the three-dimensional structure by using restrained molecular dynamics methods.

Structures were calculated with the SANDER module of the molecular dynamics package AMBER 12.[33] The coordinates of the free DNA structures (PDB 2M8Z)[15a] were taken as starting points. First, a 3D model of compound **13** was built with the program Sybyl and docked manually into the duplex–quadruplex interface. Then, the docked structures were taken as starting coordinates for the AMBER refinement, consisting of an annealing protocol in vacuo, followed by trajectories of 1 ns each in which explicit solvent molecules were included. The Particle Mesh Ewald method was used to evaluate long-range electrostatic interactions.[34] The specific protocols for these calculations have been previously described.[35] The BSC1 force field was used to describe the DNA,[36] and the TIP3P model[37] was used to simulate the water molecules. Analysis of the representative structures was carried out with the program MOLMOL.[38]

Molecular dynamics simulations

The conformational and dynamical properties of selected junctions were tested through unconstrained Molecular Dynamics (MD) simulations performed with the AMBER 12 package,[33] employing the BSC1 force field.[36] MD trajectories, 1 μs long, were collected in the presence of explicit TIP3P water,[37] periodic boundary conditions and Ewald sums for the treatment of long-range electrostatic interactions.[34] DNA charges were neutralised with potassium ions including the presence of the well-known stabilising cations placed between the G-tetrads. The time-step was 1 fs in all the simulations. Finally, the distribution of potassium cations along the MD trajectories was analysed by employing the program Chimera.[39]

Acknowledgements

This investigation was supported by research grants from the Ministerio de Ciencia, Innovación y Universidades (CTQ2016-79255-P, PID2019-107476GB-I00, BFU2017-89707-P and RTI2018-099592-B-C21). A.G.S. acknowledges the Ministerio de Ciencia, Innovación *y* Universidades for a Juan de la Cierva contract. L.D.-C., I. S-R and L.M.-J. acknowledge the Ministerio de Ciencia, Innovación y Universidades for FPI contracts. NMR experiments were performed in the "Manuel Rico" NMR laboratory (LMR), a node of the Spanish Large-Scale National Facility (ICTS R-LRB).

Conflict of interest

The authors declare no conflict of interest.

Keywords: NMR structures · nucleic acids ligands · pharmacophores · quadruplex–duplex junctions · selective molecular recognition

[1] J. Y. Davis, *Angew. Chem. Int. Ed.* **2004**, *43*, 668 – 698; *Angew. Chem.* **2004**, *116*, 684 – 716.
[2] C. Schaffitzel, I. Berger, J. Postberg, J. Hanes, H. J. Lipps, A. Pluckthun, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8572 – 8577.
[3] G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, *Nat. Chem.* **2013**, *5*, 182 – 186.

[4] S. Neidle, *Nat. Rev. Chem.* **2017**, *1*, 0041.
[5] T. Tian, Y.-Q. Chen, S.-R. Wang, X. Zhou, *Chem* **2018**, *4*, 1314 – 1344.
[6] B. Onel, M. Carver, G. Wu, D. Timonina, S. Kalarn, M. Larriva, D. Yang, *J. Am. Chem. Soc.* **2016**, *138*, 2563 – 2470.
[7] G. W. Collie, G. N. Parkinson, *Chem. Soc. Rev.* **2011**, *40*, 5867 – 5892.
[8] R. Hänsel-Hertsch, M. Di Antonio, S. Balasubramanian, *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 279 – 284.
[9] It is important to note that when TBA is formed, two of the DNA bases on the top loop (G8 and T9) actually stack strongly on the G4 tetrad. This means that any G4 binder needs to displace those DNA bases to stack on the tetrad with a quite big energetic cost. See: S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat. Rev. Drug. Discovery* **2011**, *10*, 261 – 275.
[10] L. M. Harris, C. J. Merrick, *PloS Pathog.* **2015**, *11*, e1004562.
[11] A. R. Haeusler, C. J. Donnelly, G. Periz, E. A. Simko, P. G. Shaw, M. S. Kim, N. J. Maragakis, J. C. Troncoso, A. Pandey, R. Sattler, J. D. Rothstein, J. Wang, *Nature* **2014**, *507*, 195 – 200.
[12] a) A. Gupta, L.-L. Lee, S. Roy, F. A. Tanious, W. D. Wilson, D. H. Ly, B. A. Armitage, *ChemBioChem* **2013**, *14*, 1476 – 1484; b) S. Roy, K. J. Zanotti, C. T. Murphy, F. A. Tanious, W. D. Wilson, D. H. Ly, B. A. Armitage, *Chem. Commun.* **2011**, *47*, 8524 – 8526; c) S. Lusvarghi, C. T. Murphy, S. Roy, F. A. Tanious, I. Sacui, W. D. Wilson, D. H. Ly, B. A. Armitage, *J. Am. Chem. Soc.* **2009**, *131*, 18415 – 18424; d) I. G. Panyutin, M. I. Onyshchenko, E. A. Englund, D. H. Appella, R. D. Neumann, *Curr. Pharm. Des.* **2012**, *18*, 1984 – 1991; e) M. I. Onyshchenko, T. I. Gaynutdinov, E. A. Englund, D. H. Appella, R. D. Neumann, I. G. Panyutin, *Nucleic Acids Res.* **2009**, *37*, 7570 – 7580; f) T. I. Gaynutdinov, R. D. Neumann, I. G. Panyutin, *Nucleic Acids Res.* **2008**, *36*, 4079 – 4087; g) M. Tassinari, M. Zuffo, M. Nadai, V. Pirota, A. C. Sevilla Montalvo, F. Doria, M. Freccero, S. N. Richter, *Nucleic Acids Res.* **2020**, *48*, 4627 – 4642.
[13] a) I. R. Russo Krauss, S. Ramaswamy, S. Neidle, S. Haider, G. N. Parkinson, *J. Am. Chem. Soc.* **2016**, *138*, 1226 – 1233; b) T. Q. N. Nguyen, K. W. Lim, A. T. Phan, *Sci. Rep.* **2017**, *7*, 11969.
[14] E. Butovskaya, B. Heddi, B. Bakalar, S. N. Richter, A. T. Phan, *J. Am. Chem. Soc.* **2018**, *140*, 13654 – 13662.
[15] a) K. W. Lim, A. T. Phan, *Angew. Chem. Int. Ed.* **2013**, *52*, 8566 – 8569; *Angew. Chem.* **2013**, *125*, 8728 – 8731; b) D. J. Y. Tan, F. R. Winnerdy, K. W. Lim, A. T. Phan, *Nucleic Acids Res.* **2020**, *48*, 11162 – 11171; c) T. Q. N. Nguyen, K. W. Lim, A. T. Phan, *Nucleic Acids Res.* **2020**, *48*, 10567 – 10575.
[16] B. Karg, S. Mohr, K. Weisz, *Angew. Chem. Int. Ed.* **2019**, *58*, 11068 – 11071; *Angew. Chem.* **2019**, *131*, 11184 – 11188.
[17] Y. M. Vianney, P. Preckwinkel, S. Mohr, K. Weisz, *Chem. Eur. J.* **2020**, 26, 16910 – 16922.
[18] R. F. Macaya, P. Schultze, F. W. Smith, J. A. Roe, J. Feigon, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3745 – 3749.
[19] L. M. Zhu, S. H. Chou, J. D. Xu, B. R. Reid, *Nat. Struct. Biol.* **1995**, *2*, 1012 – 1017.
[20] R. Perrone, M. Nadai, I. Frasson, J. A. Poe, E. Butovskaya, T. E. Smithgall, M. Palumbo, G. Palu, S. N. Richter, *J. Med. Chem.* **2013**, *56*, 6521 – 6530.
[21] D. Piekna-Przybylska, M. A. Sullivan, G. Sharma, R. A. Bambara, *Biochemistry* **2014**, *53*, 2581 – 2593.
[22] K. W. Lim, P. Jenjaroenpun, Z. J. Low, Z. J. Khong, Y. S. Ng, V. A. Kuznetsov, A. T. Phan, *Nucleic Acids Res.* **2015**, *43*, 5630 – 5646.
[23] S. Amrane, A. Kerkour, A. Bedrat, B. Vialet, M. L. Andreola, J. L. Mergny, *J. Am. Chem. Soc.* **2014**, *136*, 5249 – 5252.
[24] E. Jiménez-Moreno, L. Montalvillo-Jiménez, A. G. Santana, A. M. Gómez, G. Jiménez-Osés, F. Corzana, A. Bastida, J. Jiménez-Barbero, F. J. Cañada, I. Gómez-Pinto, C. González, J. L. Asensio, *J. Am. Chem. Soc.* **2016**, *138*, 6463 – 6474.
[25] A. Kerkour, J. Marquevielle, S. Ivashchenko, L. A. Yatsunyk, J. L. Mergny, G. F. Salgado, *J. Biol. Chem.* **2017**, *292*, 8082 – 8091.
[26] J. Dai, D. Chen, R. A. Jones, L. H. Hurley, D. Yang, *Nucleic Acids Res.* **2006**, *34*, 5133 – 5144.
[27] A. T. Phan, V. Kuryavyi, K. N. Luu, D. J. Patel, *Nucleic Acids Res.* **2007**, *35*, 6517 – 6525.
[28] N. Jing, R. F. Rando, Y. Pommier, M. E. Hogan, *Biochemistry* **1997**, *36*, 12498 – 12505.
[29] P. Podbevsek, N. V. Hud, J. Plavec, *Nucleic Acids Res.* **2007**, *35*, 2554 – 2563.
[30] P. Mendes, *Comput. Applic. Biosci.* **1993**, *9*, 563 – 571.
[31] M. Piotto, V. Saudek, V. Sklenář, *J. Biomol. NMR* **1992**, *2*, 661 – 665.
[32] D. T. Goddard, D. G. Kneller, SPARKY, 3rd ed., University of California, San Francisco.
[33] D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Götz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M. J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P. A. Kollman, **2012**, AMBER12, University of California, San Francisco.
[34] T. Darden, D. York, L. Pedersen, *Chem. Phys.* **1993**, *98*, 10089 – 10092.
[35] R. Soliva, V. Monaco, I. Gómez-Pinto, N. J. Meeuwenoord, G. A. Marel, J. H. Boom, C. González, M. Orozco, *Nucleic Acids Res.* **2001**, *29*, 2973 – 2985.
[36] I. Ivani, P. D. Dans, A. Noy, A. Pérez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goñi, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpí, C. González, M. Vendruscolo, C. A. Laughton, S. A. Harris, D. A. Case, M. Orozco, *Nat. Methods* **2015**, *12*, 55 – 58.
[37] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926 – 936.
[38] R. Koradi, M. Billeter, K. J. Wüthrich, *Mol. Graph.* **1996**, *14*, 51 – 55.
[39] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605 – 1612.

Manuscript received: November 19, 2020
Accepted manuscript online: December 25, 2020
Version of record online: January 22, 2021