ChemPubSoc Europe

DOI: 10.1002/chem.201904923

CHEMISTRY
A European Journal
Communication

DNA Structures | *Very Important Paper*

VIP Intercalation of a Heterocyclic Ligand between Quartets in a G-Rich Tetrahelical Structure

Anita Kotar[+],[a] Vojč Kocman[+],[a] and Janez Plavec*[a, b, c]

Abstract: YES G-rich oligonucleotide VK2 folds into an AGCGA-quadruplex tetrahelical structure distinct and significantly different from G-quadruplexes, even though it contains four G_3 tracts. Herein, a bis-quinolinium ligand 360A with high affinity for G-quadruplex structures and selective telomerase inhibition is shown to strongly bind to VK2. Upon binding, 360A does not induce a conformational switch from VK2 to an expected G-quadruplex. In contrast, NMR structural study revealed formation of a well-defined VK2–360A complex with a 1:1 binding stoichiometry, in which 360A intercalates between GAGA- and GCGC-quartets in the central cavity of VK2. This is the first high-resolution structure of a G-quadruplex ligand intercalating into a G-rich tetrahelical fold. This unique mode of ligand binding into tetrahelical DNA architecture offers insights into the stabilization of an AGCGA-quadruplex by a heterocyclic ligand and provides guidelines for rational design of novel VK2 binding molecules with selectivity for different DNA secondary structures.

Guanine-rich DNA oligonucleotides can fold into tetrahelical structures called G-quadruplexes characterized by stacked G-quartets, cyclic arrangements of four guanine residues held together by Hoogsteen hydrogen bonds and stabilized by a central (Na^+/K^+) cations.[1] In the human genome, G-rich segments are localized in regions implicated in processes such as DNA replication, telomere maintenance, gene expression and genetic instability.[2]

Probing the human genome with polymerase stop assay experiments, designed to detect higher-order structures in DNA, demonstrated that many stable structures are formed by G-rich segments that do not comply with the basic G-quadruplex folding motif [that is, 5′-d(GGG(N_{1-7})GGG(N_{1-7})GGG(N_{1-7})GGG)-3′] and cannot be simply assumed to form G-quadruplexes containing long loops or single-nucleotide bulges.[3] In this respect, it was recently reported that sequences that adhere to a 5′-d(AGCGA(N_{1-20})AGCGA(N_{1-20})AGCGA(N_{1-20})AGCGA)-3′ folding motif and are also G-rich can adopt structures, significantly different from G-quadruplexes, named AGCGA-quadruplexes (Figure 1 a,b).[4] They are characterized by G–G base pairs in N1-carbonyl symmetric geometry, G–A base pairs in N1–N7, carbonyl-amino and G–C base pairs in Watson–Crick geometries (Figure 1 c). In some of these structures the G–A and G–C base pairs were shown to form GAGA- and GCGC-quartets, respectively.[4b]

Oligonucleotide VK2, 5′-d(G_3AGCGAG_3AGCGAG_3AGCGA G_3AGCG)-3′, originating from regulatory region of the PLEKHG3 human gene, related to autism, is a representative member of the AGCGA-quadruplex family.[4a,5] VK2 adopts a monomeric

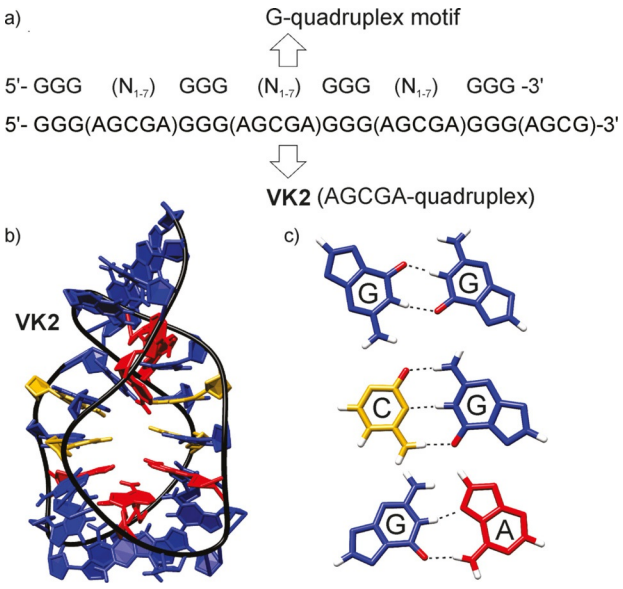

Figure 1. a) Comparison of G-quadruplex motif with VK2 oligonucleotide sequence, which adopts an AGCGA-quadruplex. b) Cartoon presentation of the VK2 structure. c) G–G base pair in N1-carbonyl symmetric geometry, G–C base pair in *Watson–Crick* and G–A base pair in N1–N7, carbonyl-amino geometries. Guanines are colored in blue, adenines in red and cytosines in yellow. Dashed lines represent hydrogen-bonds between residues.

[a] *Dr. A. Kotar,[+] Dr. V. Kocman,[+] Prof. Dr. J. Plavec*
National Institute of Chemistry
Hajdrihova 19 1000 Ljubljana (Slovenia)
E-mail: janez.plavec@ki.si

[b] *Prof. Dr. J. Plavec*
Faculty of Chemistry and Chemical Technology, University of Ljubljana
Večna pot 113, Ljubljana (Slovenia)

[c] *Prof. Dr. J. Plavec*
EN-FIST Center of Excellence
Trg Osvobodilne fronte 13, 1000 Ljubljana (Slovenia)

[+] *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.201904923.

structure with a C_2 axis of symmetry (Figures 1; Figure S1, Supporting Information).[4a,6] Intriguingly, its sequence adheres to folding motifs required for both G- and AGCGA-quadruplexes.

Considering that G-rich regions have been implicated in many vital cellular processes, numerous ligands have been developed that bind to, and stabilize, G-quadruplexes as potential drugs, or probes for in vivo visualization.[7] Critically, G-quadruplex ligands have to be highly selective over other types of DNA structures, most crucially double helices. The majority of known ligands bind through π–π stacking on outer G-quartets, which are unique structural features of G-quadruplexes.[8] In comparison, molecules that interact with grooves and loops of G-quadruplexes are much less common, whereas intercalation between (inner) G-quartets has not been observed so far.[9]

Determining how well-known G-quadruplex ligands interact with alternative G-rich tetrahelical structures, such as the VK2 AGCGA-quadruplex, is important considering that a given ligand could act as a conformational switcher by unfolding an AGCGA- quadruplex and forming a G-quadruplex. Alternatively, a ligand could utilize different binding modes to selectively bind to VK2 AGCGA-quadruplex. Furthermore, ligands that bind to VK2 could prove a valuable starting point in design of AGCGA-quadruplex binders important for targeting regulatory regions of genes connected to neurological disorders, cancer and abnormalities in bone and cartilage development.[4b]

Compound 360A, 3,3′-[(pyridine-2,6-dicarbonyl)bis(azanediyl)]bis(1-methylquinolin-1-ium), is a bis-quinolinium ligand displaying strong affinity and selectivity for G-quadruplex structures and selective telomerase inhibition.[10] Our initial experiments showed that 360A binds to, and thermally stabilizes, the fold adopted by VK2 as proven by the approximately 18 °C increase of T_m for the VK2–360A complex compared to VK2 alone (Figure 2a). Potential structural transformation of VK2 from AGCGA- to a putative G-quadruplex in the presence of 360A was not observed even after annealing and/or in the presence of 100 mM K^+ ions (Figure S2, Supporting Information).

NMR titration experiments revealed that already the addition of 0.25 equivalents of 360A to a VK2 solution resulted in the appearance of a new set of signals in ^{1}H NMR spectra, which indicated the formation of VK2–360A complex (Figures 2b, Figure S3, Supporting Information). Separate and resolved signals for VK2 alone and in the complex demonstrate slow exchange between the two species on the ^{1}H NMR chemical shift timescale. During further titration to up to 1:2 VK2/360A ratio, signals of the complex became more intense, whereas signals of VK2 alone were diminished to baseline. No new signals were detected at 1:3 and 1:4 ratios.

Well-resolved signals of A8H2″ of VK2 alone and when bound with 360A allowed us to monitor the formation of the complex (Figure 2b). The intensity of A8H2″ signal of VK2–360A complex reached a plateau at 1.5 molar equivalents of ligand and suggested the existence of one binding site for 360A on VK2. The plot of the A8H2″ signal intensity in the complex as a function of the total ligand concentration fitted to 1:1 binding mode yielded a K_D value of 7 ± 1 μM (Figure S4, Supporting Information).

Proton signal line shapes of 360A were analyzed at VK2:360A ratios from 1:2 to 1:20 and revealed variations in chemical exchange rates from fast to slow regime of ligand binding (Figures 2c, Figure S5, Supporting Information). Broadening and decrease in intensity of ligand signals was observed at 1:20 and 1:10 VK2/360A ratios. Line broadening is severe at 1:7 VK2/360A ratio, but at the same time it is possible to detect a new set of distinct signals of bound ligand. All proton signals of 360A in 1:1 VK2–360A complex are uniformly shifted upfield by $\Delta\delta$ of 0.7–1.0 ppm compared to the chemical shifts of 360A alone, which suggests that the ligand intercalates into the structure of VK2.

The imino proton resonances of VK2 alone and in the VK2–360A complex were unambiguously assigned by using 1D ^{15}N-edited HSQC spectra acquired on 10% residue-specific ^{13}C,^{15}N-labeled oligonucleotides (Figure S6, Supporting Information). The two subunits of VK2 numbered as G1 to G15, which are connected formally with A16 (Figure 3), exhibit isochronous ^{1}H NMR signals. A single set of imino signals is observed for residues G1–G15 in VK2 quadruplex alone, except for G15 with two resolved imino signals. Further assignments of aromatic and sugar protons were based on in-depth analysis of correlations in NOESY spectra (Figures 4, Figure S7, Supporting Information). However, similar fingerprints of cross-peaks in sugar-

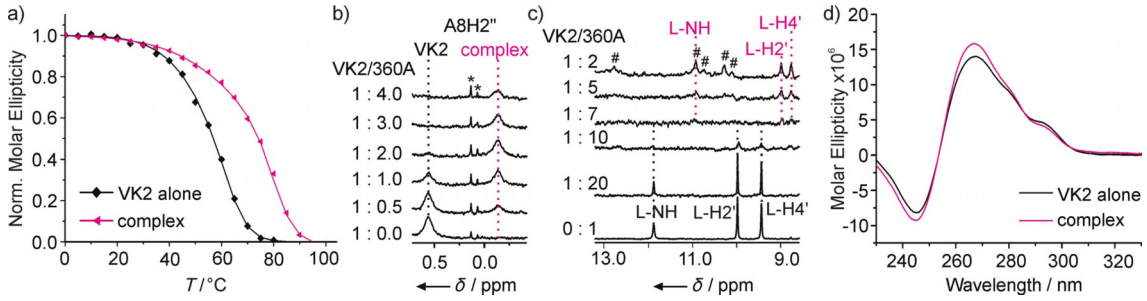

Figure 2. Binding of 360A to VK2. a) Normalized CD melting profiles (0–95 °C) of VK2 alone (black) and VK2–360A complex (magenta) monitored at 266 nm. b) A8H2″ signals in ^{1}H NMR spectra of VK2 as 360A is titrated into solution. Signals of unknown impurities are labeled with asterisks. c) Selected region of ^{1}H NMR spectra of 360A alone and in the presence of VK2. Signals of VK2 are labeled with the symbol "#". Assignments of b) VK2 and c) 360A signals are shown above the corresponding spectra and are marked in black and magenta for VK2 or 360A alone and in the complex, respectively. The molar ratio between VK2 and 360A is indicated on the left-hand side of NMR spectra. The spectra were recorded at b) 0.2 and c) 0.01 mM VK2, 100 mM LiCl, pH 6.0, 0 °C at 600 MHz. d) CD spectra of VK2 alone (black) and VK2–360A complex (magenta) at 0.025 mM VK2, 0.1 mM 360A, 100 mM LiCl, pH 6.0 at 0 °C.

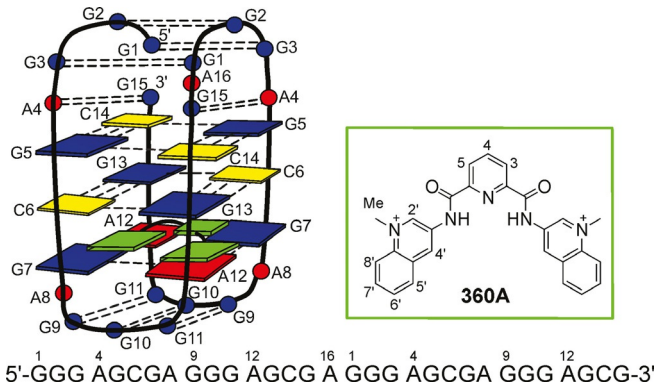

Figure 3. Topology of complex VK2–360A, oligonucleotide sequence of VK2 and chemical structure of 360A with the residue and atom numbering used in this study. Guanine residues are colored in blue, adenines in red, cytosines in yellow and 360A in green. Dashed lines represent hydrogen bonds between residues.

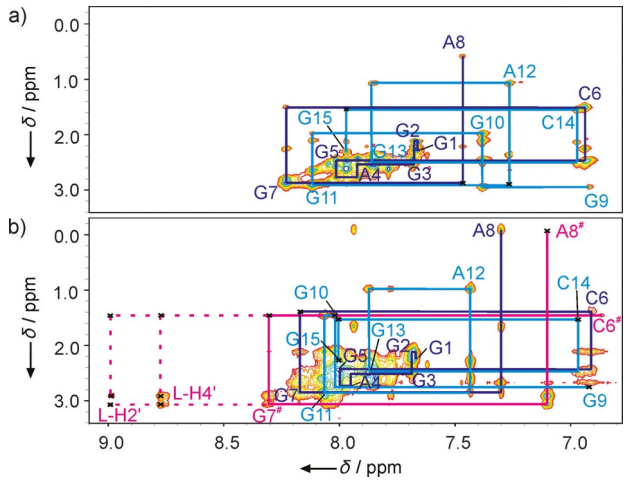

Figure 4. Aromatic-H2′/2″ region of ^{1}H-^{1}H NOESY spectra (τ_m 200 ms) of a) VK2 alone and b) 1:1 VK2–360A complex. The G1–A8 and G9–G15 sequential walks are depicted in blue and cyan, respectively. Sequential assignment of new VK2 cross-peaks are marked with "#" and connections between VK2 and 360A are shown in magenta. NMR spectra were recorded at 0.35 mM VK2 per strand, 100 mM LiCl, pH 6.0, 0 °C, 1.4 mM 360A (b), at 800 MHz.

NOE contacts between 360A and aromatic and sugars protons of both C6 and G7 clearly show that the ligand is located between planes defined by two C6-G13 and two G7-A12 base pairs inside the central cavity of VK2 (Figures 3, Figures S8 and S9, Supporting Information). Specifically, quinoline rings of 360A are positioned near sugar moieties of C6 and G7. NOE contacts between A12 and G13 residues and 360A are impossible to assign due to severe overlap of respective aromatic protons (Figure S10, Supporting Information). Observation of multiple weak NOE interactions between 360A and A8 indicate that two A8 residues are tightly and asymmetrically stacked on the two G7–A12 base pairs.

The high-resolution structure of 1:1 VK2–360A complex was calculated with a simulated-annealing method based on NOE-derived distance, H-bond, and torsion-angle restraints (Figure 5 a, Table S3, Supporting Information). Calculations resulted

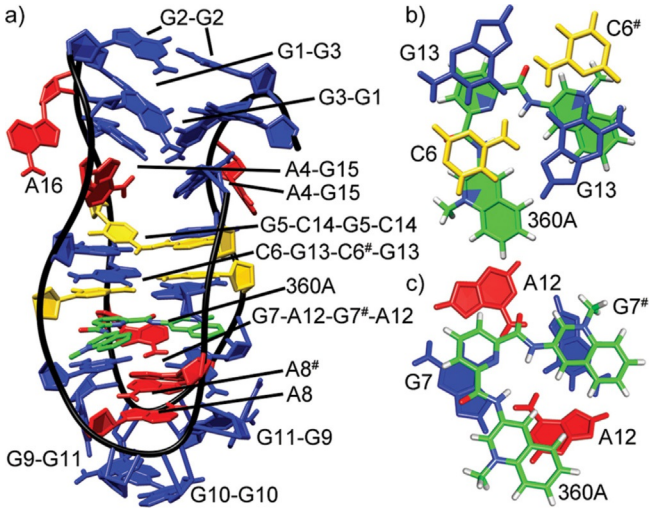

Figure 5. Cartoon representation of the structure of the 1:1 VK2–360A complex (PDB ID: 6SX3). a) A wide groove view of the complex. Stacking of 360A over b) C6G13C6#G13-quartet and c) G7A12G7#A12-quartet. Residues marked with "#" indicate parts of the structure where the symmetry of VK2 is lost due to binding of 360A. Guanines are colored in blue, adenines in red, cytosines in yellow, and 360A in green.

aromatic regions of NOESY spectra of VK2 alone and VK2–360A complex clearly suggest that the VK2 topology remained unchanged upon ligand binding. In full agreement, CD spectra showed that AGCGA-quadruplex structure is retained in the VK2–360A complex with its characteristic minimum at 245 nm, maximum at 266 nm, and two shoulders at 280 and 290 nm (Figure 2 d). However, local structural rearrangements occur in the VK2–360A complex compared to VK2 alone as is revealed by upfield shifts of imino protons, especially of G9 and G11 (Table S1, Supporting Information). Interestingly, two resolved imino signals in VK2–360A complex are observed for G7, which points to loss of local structural symmetry. Similar symmetry braking is reflected in two separate NOESY sequential walks for the C6-G7-A8 residues in the VK2–360A complex. Importantly, the G5-A8 segments form the characteristic central cavity in VK2 alone (Figure 1 c).

in a well-converged family of ten structures of VK2–360A complex (overall r.m.s.d.: 2.6 Å, Figure S11, Supporting Information). The position of 360A in VK2 is well-defined and supported by 37 intermolecular NOE-based distance restraints. In the calculated structure of VK2–360A complex, two C6–G13 base pairs form a C6G13C6#G13-quartet through their association through major grooves (Figure 5 b). Similarly, the two G7–A12 base pairs are closer in the complex than in VK2 alone and form a G7A12G7#A12-quartet (Figure 5 c). Ligand 360A is intercalated between G7A12G7#A12- and C6G13C6#G13-quartets (Figure 5). This serves as a platform for the formation of a G5C14G5C14-quartet. Therefore, in the complex compared to VK2 alone the hydrophobic core of G9–G11, G10–G10 and G11–G9 base pairs around two A8 residues is extended with a GAGA- and two GCGC-quartets. Although the conformational freedom in the center of VK2 alone is a major source of the en-

tropic contribution to the stability of the structure,[6] we postulate that the VK2–360 complex is stabilized in the center by a hydrophobic core and entropically by conformational freedom of G1–G3, G2–G2, G3–G1 and two A4–G15 base pairs on top of the complex (Figures S11 and S12, Supporting Information).

Our results show how 360A utilizes G–A and G–C base pairs to form GAGA- and GCGC-quartets to intercalate into VK2 AGCGA-quadruplex and, by doing so, it stabilizes the complex. Although 360A is a well-established G-quadruplex stabilizer, which is presumed to stack on outer G-quartet(s),[10,11] no structural transformation of VK2 into a G-quadruplex was observed upon binding of 360A. The aromatic scaffold of 360A is small and flexible enough to enter into the central cavity of VK2 and thus intercalate between G7A12G7#A12- and C6G13C6#G13-quartets. Two positive charged groups of 360A additionally contribute to its binding to VK2 by favorable electrostatic interactions. Despite reducing the conformational flexibility of the central cavity, folding topology of VK2 remains unchanged after binding of 360A. In contrast, upon interacting with a telomeric G-quadruplex, 360A induces a change in topology.[12] Our results suggest that the binding into the central cavity of VK2 is crucially regulated by the size of the ligand. Preliminary results showed significantly lower affinity and much less effective thermal stabilization of VK2 for other bis-quinolinium ligands, such as Piridostatin and Phen-DC3 (Figures S13–S17, Supporting Information). Piridostatin containing short side-chains on quinolinium rings and Phen-DC3 comprising phenanthroline instead of pyridine central ring seem to be too bulky for efficient intercalation into VK2.

In summary, the high-resolution structure of VK2–360A complex represents the first report of intercalation of a heterocyclic ligand into a quadruplex structure. The detailed structural insights into interactions of 360A into AGCGA-quadruplexes uncover alternative structural elements that could bind 360A in addition to G-quartets in the context of G-rich genomic regions.

Acknowledgements

This project was financial supported by the Slovenian Research Agency (P1-0242, J1-1704). The authors acknowledge Prof. Marie-Paule Teulade-Fichou for initial batches of 360A and Phen-DC3, which enabled the preliminary tests, and Dr. Štefan Možina for help with preliminary experiments.

Conflict of interest

The authors declare no conflict of interest.

Keywords: 360A complex · AGCGA-quadruplex · G-quadruplex · G-rich regions · NMR

[1] a) R. Hänsel-Hertsch, M. Di Antonio, S. Balasubramanian, *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 279–284; b) M. Marušič, P. Šket, L. Bauer, V. Viglasky, J. Plavec, *Nucleic Acids Res.* **2012**, *40*, 6946–6956; c) T. Tian, Y.-Q. Chen, S.-R. Wang, X. Zhou, *Chem.* **2018**, *4*, 1314–1344; d) N. V. Hud, J. Plavec in *Quadruplex Nucleic Acids* (Eds.: S. Neidle, S. Balasubramanian), Royal Society of Chemistry, Cambridge, **2006**.

[2] a) J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res.* **2005**, *33*, 2908–2916; b) D. Rhodes, H. J. Lipps, *Nucleic Acids Res.* **2015**, *43*, 8627–8637.

[3] a) V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith, S. Balasubramanian, *Nat. Biotechnol.* **2015**, *33*, 877–881; b) N. Maizels, L. T. Gray, *PLoS Genet.* **2013**, *9*, e1003468; c) G. Marsico, V. S. Chambers, A. B. Sahakyan, P. McCauley, J. M. Boutell, M. Di Antonio, S. Balasubramanian, *Nucleic Acids Res.* **2019**, *47*, 3862–3387.

[4] a) V. Kocman, J. Plavec, *Nat. Commun.* **2014**, *5*, 5831; b) V. Kocman, J. Plavec, *Nat. Commun.* **2017**, *8*, 15355.

[5] a) A. J. Griswold, D. Ma, S. J. Sacharow, J. L. Robinson, J. M. Jaworski, H. H. Wright, R. K. Abramson, H. Lybæk, N. Øyen, M. Cuccaro, J. Gilbert, M. A. Pericak-Vance, *Autism Res.* **2011**, *4*, 221–227; b) T. T. Nguyen, W. S. Park, B. O. Park, C. Y. Kim, Y. Oh, J. M. Kim, H. Choi, T. Kyung, C.-H. Kim, G. Lee, K. M. Hahn, T. Meyer, W. D. Heo, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 10091–10096.

[6] S. Hadži, V. Kocman, D. Oblak, J. Plavec, J. Lah, *Angew. Chem. Int. Ed.* **2019**, *58*, 2387–2391; *Angew. Chem.* **2019**, *131*, 2409–2413.

[7] a) S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat. Rev. Drug Discovery* **2011**, *10*, 261–275; b) E. Largy, A. Granzhan, F. Hamon, D. Verga, M. P. Teulade-Fichou, *Top. Curr. Chem.* **2012**, *330*, 111–177; c) A. Kotar, B. Wang, A. Shivalingam, J. Gonzalez-Garcia, R. Vilar, J. Plavec, *Angew. Chem. Int. Ed.* **2016**, *55*, 12508–12511; *Angew. Chem.* **2016**, *128*, 12696–12699; d) D. Monchaud, M.-P. Teulade-Fichou, *Org. Biomol. Chem.* **2008**, *6*, 627–636; e) R. Vilar, *Met. Ions Life Sci.* **2018**, *18*, 325–349, https://doi.org/10.1515/9783110470734-012; f) S. Neidle, *Nat. Chem. Rev.* **2017**, *1*, 41.

[8] a) S. M. Haider, S. Neidle, G. N. Parkinson, *Biochimie* **2011**, *93*, 1239–1251; b) G. N. Parkinson, F. Cuenca, S. Neidle, *J. Mol. Biol.* **2008**, *381*, 1145–1156; c) M. Bončina, C. Podlipnik, I. Piantanida, J. Eilmes, M. P. Teulade-Fichou, G. Vesnaver, J. Lah, *Nucleic Acids Res.* **2015**, *43*, 10376–10386.

[9] a) B. Pagano, I. Fotticchia, S. De Tito, C. A. Mattia, L. Mayol, E. Novellino, A. Randazzo, C. Giancola, *J. Nucleic Acids* **2010**, 247137; b) M. Cavallari, A. Garbesi, R. Di Felice, *J. Phys. Chem. B* **2009**, *113*, 13152–13160; c) E. Gavathiotis, R. A. Heald, M. F. G. Stevens, M. S. Searle, *J. Mol. Biol.* **2003**, *334*, 25–36.

[10] a) G. Pennarun, C. Granotier, L. R. Gauthier, D. Gomez, F. Hoffschir, E. Mandine, J. F. Riou, J. L. Mergny, P. Mailliet, F. D. Boussin, *Oncogene* **2005**, *24*, 2917–2928; b) C. Granotier, G. Pennarun, L. Riou, F. Hoffschir, L. R. Gauthier, A. De Cian, D. Gomez, E. Mandine, J.-F. Riou, J.-L. Mergny, P. Mailliet, B. Dutrillaux, F. D. Boussin, *Nucleic Acids Res.* **2005**, *33*, 4182–4190.

[11] C. Saintomé, P. Alberti, N. Guinot, P. Lejault, J. Chatain, P. Mailliet, J. F. Riou, A. Bugaut, *Chem. Commun.* **2018**, *54*, 1897–1900.

[12] a) A. Marchand, A. Granzhan, K. Iida, Y. Tsushima, Y. Ma, K. Nagasawa, M. P. Teulade-Fichou, V. Gabelica, *J. Am. Chem. Soc.* **2015**, *137*, 750–756; b) M. Bončina, F. Hamon, B. Islam, M.-P. Teulade-Fichou, G. Vesnaver, S. M. Haider, J. Lah, *Biophys. J.* **2015**, *108*, 2903–2911.

Manuscript received: October 29, 2019
Revised manuscript received: November 20, 2019
Accepted manuscript online: November 21, 2019
Version of record online: January 9, 2020

