G-Quadruplexes

International Edition: DOI: 10.1002/anie.201606877
German Edition: DOI: 10.1002/ange.201606877

NMR Structure of a Triangulenium-Based Long-Lived Fluorescence Probe Bound to a G-Quadruplex

*Anita Kotar, Baifan Wang, Arun Shivalingam, Jorge Gonzalez-Garcia, Ramon Vilar, and Janez Plavec**

Abstract: *An NMR structural study of the interaction between a small-molecule optical probe (DAOTA-M2) and a G-quadruplex from the promoter region of the c-myc oncogene revealed that they interact at 1:2 binding stoichiometry. NMR-restrained structural calculations show that binding of DAOTA-M2 occurs mainly through π–π stacking between the polyaromatic core of the ligand and guanine residues of the outer G-quartets. Interestingly, the binding affinities of DAOTA-M2 differ by a factor of two for the outer G-quartets of the unimolecular parallel G-quadruplex under study. Unrestrained MD calculations indicate that DAOTA-M2 displays significant dynamic behavior when stacked on a G-quartet plane. These studies provide molecular guidelines for the design of triangulenium derivatives that can be used as optical probes for G-quadruplexes.*

Guanine-rich oligonucleotides can readily fold into four-stranded helical structures known as G-quadruplexes, which have received significant attention over the past few years since their formation has been associated with a number of essential biological processes such as gene expression, replication, and telomere biology.[1] While the formation of G-quadruplexes in vitro under physiological conditions has been unambiguously demonstrated, their formation and functions in vivo have proven more difficult to confirm. Landmark studies using labeled antibodies in fixed cells in ciliate macronuclei and more recently in human cells have provided the most direct evidence for the existence of G-quadruplexes in vivo.[2] However, these imaging studies have some limitations, such as a lack of dynamic information and the fact that fixation artifacts could impair visualization.[3] An alternative approach to G-quadruplex imaging is the use of small-molecule optical probes.[4] The majority of G-quadruplex optical probes reported to date rely on changes in their emission intensity upon interaction with DNA.[5] Recently, we reported a small-molecule optical probe based on triangulenium (DAOTA-M2, Figure 1a), and its interaction with G-quadruplexes can be imaged in live cells using fluorescence-lifetime imaging microscopy (FLIM).[6] The key feature of this probe is that it shows significantly longer fluorescence lifetimes when bound to G-quadruplex DNA than when it binds to other topologies such as duplex or single-stranded DNA.[7]

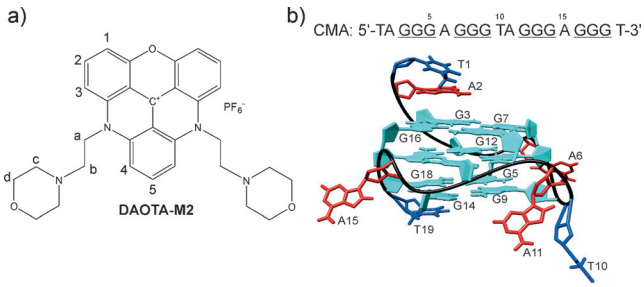

Figure 1. Triangulenium based probe DAOTA-M2 and the oligonucleotide CMA. a) The chemical structure of DAOTA-M2 with the atom numbering used in this study. b) The sequence and the corresponding CMA G-quadruplex structure in KCl solution (PDB ID: 2LBY[9a]). Guanines (for clarity, those involved in the central G-quartet are not marked) are shown in cyan, adenines in red, and thymines in blue.

With the aim of further rationalizing the behavior of DAOTA-M2 and to aid us in the development of even better and more selective triangulenium-based probes, we herein present detailed NMR studies to characterize the interaction between DAOTA-M2 and an intramolecular G-quadruplex DNA structure. We focused on the modified 19-nucleotide sequence d[TAGGGAGGGTAGGGAGGGT] originating from the NHE III_1 region upstream of the P1 promoter of the oncogene *c-myc*, the overexpression of which has been implicated in a wide variety of human cancers.[8] This sequence contains four guanine tracts and has been shown to form a well-defined parallel-stranded G-quadruplex (named CMA) under physiologically relevant K^+ ion concentrations (Figure 1b).[9] Initial titration experiments suggested that DAOTA-M2 binds to CMA with a 1:2 stoichiometry.

Twelve well-resolved imino proton signals in the ^{1}H NMR spectrum of CMA between δ 11.07 and 12.09 ppm are consistent with the formation of a single G-quadruplex. As DAOTA-M2 was titrated into a solution of CMA, signals in the ^{1}H NMR spectrum became broader and a new set of distinct signals appeared in the imino, aromatic, and methyl

[*] A. Kotar, Dr. B. Wang, Prof. Dr. J. Plavec
Slovenian NMR Center, National Institute of Chemistry
Hajdrihova 19, 1000 Ljubljana (Slovenia)
E-mail: Janez.plavec@ki.si
Prof. Dr. J. Plavec
EN-FIST Center of Excellence, Trg OF 13, 1000 Ljubljana (Slovenia)
and
Faculty of Chemistry and Chemical Technology
University of Ljubljana, Večna pot 113, Ljubljana (Slovenia)
Dr. A. Shivalingam, Dr. J. Gonzalez-Garcia, Prof. Dr. R. Vilar
Department of Chemistry, Imperial College London
South Kensington, London SW7 2AZ (UK)

Supporting information for this article can be found under: http://dx.doi.org/10.1002/anie.201606877.

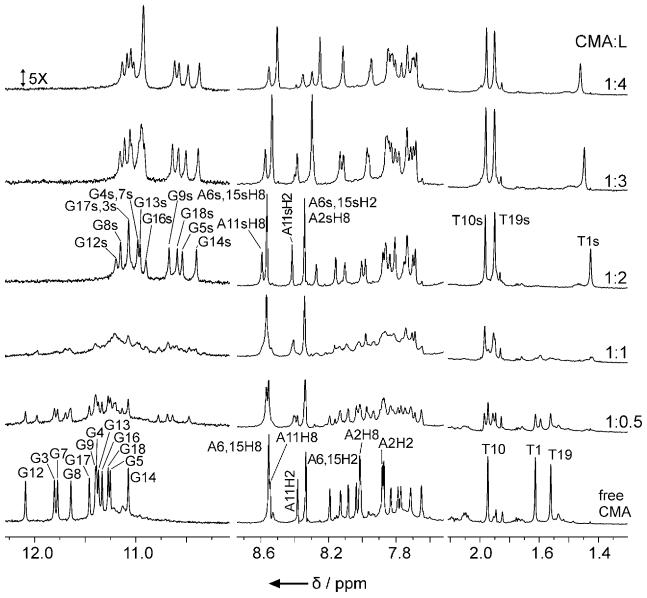

Figure 2. The imino, aromatic, and methyl regions of the 1D ^{1}H NMR spectra of CMA as DAOTA-M2 (L) is titrated into the solution. The molar ratio between CMA and the ligand is indicated on the right-hand side of spectra. Assignments of the H1, H8/H2, and Me protons of free CMA and the 1:2 CMA–ligand complex (denoted by the final letter "s") are shown above the corresponding spectra. The vertical scale of the imino region of all spectra was increased 5-fold. The NMR spectra were recorded at 0.2 mM of CMA, 100 mM KCl, pH 7.0, 25 °C on a 600 MHz NMR spectrometer.

At a 1:2 CMA/ligand ratio, the ^{1}H NMR signals of DAOTA-M2 were broader in comparison to those of the G-quadruplex and could not be unambiguously detected. Reverse titration by adding CMA to a solution of DAOTA-M2 at identical salt and pH conditions caused the signals for the ligand to become broader, less intense, and slightly shifted upfield (Figures 3, and Figures S10, S11).

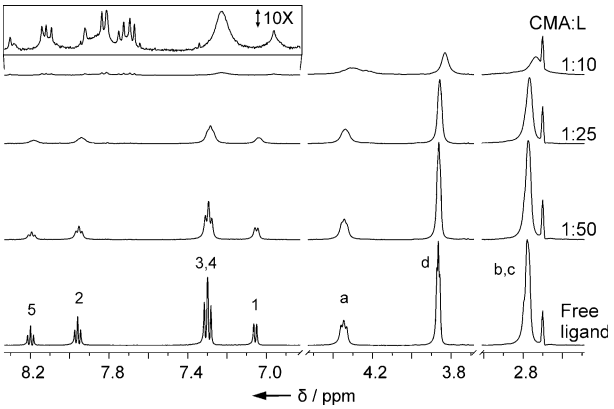

Figure 3. 1D ^{1}H NMR spectra of DAOTA-M2 alone and in the presence of CMA at CMA/ligand ratios of 1:50 to 1:10. The NMR spectra were recorded at 1.0 mM of DAOTA-M2, 100 mM KCl, pH 7.0, 25 °C on a 600 MHz NMR spectrometer.

regions upon the addition of 0.5 molar equivalents of DAOTA-M2 (Figure 2).

The observation of new signals upfield of the CMA signals indicates the formation of a CMA–ligand complex. Exchange between the two species is slow on the ^{1}H NMR chemical-shift timescale. Upon reaching a CMA/ligand ratio of 1:1, the line broadening became severe and the intensity of the signals for free CMA decreased. At a 1:2 CMA/ligand ratio, the signals became sharp and well-resolved, which suggests the formation of a well-defined complex. No new signals were detected in the imino and methyl regions at CMA/ligand ratios of 1:3 and 1:4, thus suggesting that there are only two binding sites for DAOTA-M2 on CMA.

Unambiguous spectral assignment of protons in the 1:2 CMA–ligand complex was achieved with the use of ^{13}C- and ^{15}N-edited experiments on residue-specific (8%) ^{13}C,^{15}N-labelled G8, G12, G14, and G18 of CMA (Figure S1 in the Supporting Information). Further assignments were made on the basis of exchange cross-peaks in NOESY spectra (Figures S2–S6, Table S1). Similar fingerprints in NOESY and CD spectra (Figure S7) of the complex and free CMA suggest that its parallel topology remained unchanged upon ligand binding. The largest $\Delta\delta$ values were observed for the H1 protons of guanines at the 3′ and 5′ G-quartets (Table S2). Furthermore, the protons of T1, A2, and T19 showed much larger $\Delta\delta$ values than the protons of A6, T10, A11, and A15, thus suggesting localization of DAOTA-M2 close to the outer G-quartets, while residues in loops are not significantly involved in interactions (Figures S8, S9).

Off-diagonal cross-peaks in the imino-imino region of the NOESY spectrum at a 1:0.5 CMA/ligand ratio demonstrate that CMA is involved in exchange between free and three different ligand-bound states (CMA3, CMA5, and CMAB; Figure 4). The exchange cross-peaks between CMA and CMAB as well as between CMA3 and CMA5 were relayed and were observed at only mixing times longer than 80 ms (Figure S12). Integration of resolved cross-peaks of the G12 residue showed that the population of CMA3 was nearly twice as high as CMA5 at a given concentration of the ligand. Dissociation constants (K_D) were in the micromolar range for formation of CMA3 and CMA5. Once the first ligand molecule is bound, the affinity of CMA for the other molecule increases and the formation of CMAB from either CMA3 or CMA5 occurs with a K_D value of 0.1 ± 0.4 μM ($R^2 = 0.9928$). Different binding kinetics at the 3′ and 5′ G-quartets was also indicated by variable-temperature NMR experiments. A decrease in temperature shifted the exchange regime of DAOTA-M2 at the 5′ G-quartet to an intermediate timescale, while it remained slow on the ^{1}H NMR chemical-shift timescale for the 3′ G-quartet (Figure S13).

High-resolution structures of the 1:2 CMA–ligand complex were calculated with a simulated-annealing method based on NOE-derived distance, H-bond, and torsion-angle restraints (Table S3, Figures S14, S15). The position of the ligand molecules at both outer G-quartets is well-defined and supported by 38 intermolecular NOE restraints (Figure 5a). The binding of DAOTA-M2 occurs mainly through π–π stacking between the triangulenium core and the guanine residues of the 5′ and 3′ G-quartets and possibly through cation–π and electrostatic interactions resulting from the

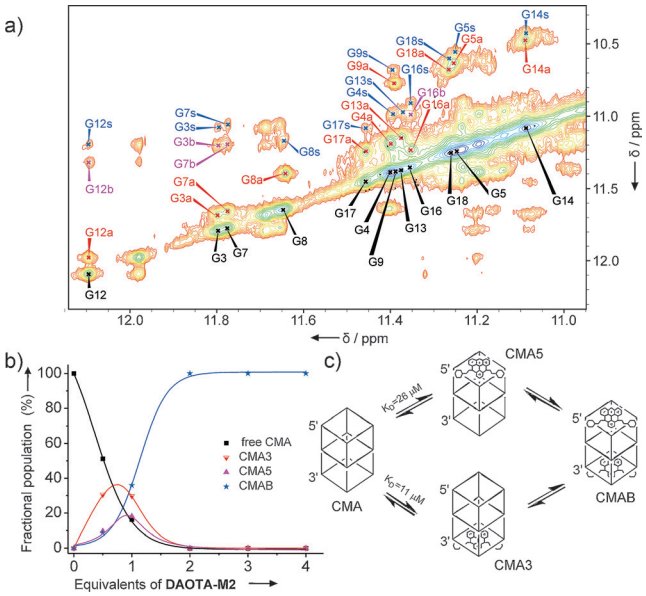

Figure 4. Binding of DAOTA-M2 to CMA. a) The imino-imino region of the NOESY spectrum of CMA at a 1:0.5 CMA/ligand ratio. Diagonal cross-peaks corresponding to free CMA are labeled in black. Cross-peaks arising from chemical exchange between free CMA and CMA3, CMA5, CMAB are labeled in red, magenta, and blue and are denoted by the final letters "a", "b" and "s", respectively. The spectrum was recorded with mixing time of 80 ms at 25 °C. b) The fractional population (±3 unit%) of free CMA and its complexes in the presence of DAOTA-M2, obtained by integration of the NOESY cross-peaks. The experimental data were fitted assuming equilibria in panel (c), with $R^2 > 0.95$ and $\chi^2_{red} < 5$. c) A schematic presentation of the formation of the 1:2 CMA–ligand complex.

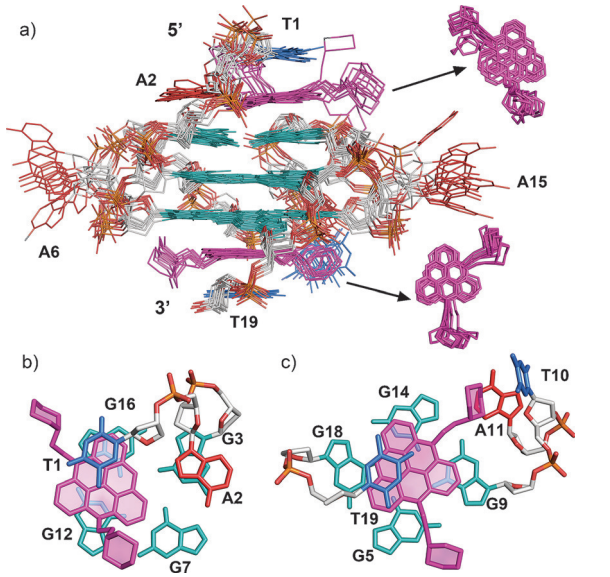

Figure 5. The structure of the 1:2 CMA–ligand complex (PDB ID: 5LIG). a) A structure ensemble of the 10 lowest-energy structures with the smallest restraints violations. The position of the lowest-energy structure of the ligand bound at the 5′ (b) and 3′ (c) ends. DAOTA-M2 is shown in magenta, guanines in cyan, adenines in red, and thymines in blue.

positively charged triangulenium core. At the 5′ end, the ligand stacks over the G-quartet with its triangulenium core coplanar with A2 and further flanked by T1 (Figure 5b). The orientation of the morpholino moieties is defined by six NOEs to the T1, G7, and G12 residues. Upon ligand binding, a considerable rearrangement is observed for the T1 and A2 residues. A2 sterically hinders central positioning of the ligand at the 5′ G-quartet and it seems that coplanar orientation of A2 with respect to the triangulenium core of the ligand is favorable. In full agreement, replacement of A2 with the propyl spacer C3 to link T1 and G3 also reveals that T1 becomes more flexible, thereby leading to unfavorable binding of the ligand to CMA. However, in both CMA and C3-modified G-quadruplexes, the terminal T1 residue is rearranged to stack over the triangulenium core (Figure S16). At the 3′ end of the CMA G-quadruplex, the triangulenium core of DAOTA-M2 is oriented centrally and covers all four guanines (Figure 5c). Upon ligand binding, the flanking T19 residue rearranges its conformation to stack over the triangulenium core and forms a sandwich-like structure. The morpholino moieties interact with G5, T10, A11, and T19 and may contribute to the stronger binding affinity at the 3′ end.

Molecular dynamics (MD) simulation of the 1:2 CMA–ligand complex showed that the ligands display significant dynamic behavior when bound to CMA. In spite of this, both ligands remained bound to individual G-quartets during three independent 100 ns MD simulations. Structures of DOATA-M2 bound to the 5′ end could be grouped into four major clusters, while two main clusters were established at the 3′ end (Figure 6, Tables S4, S5).

The ligand at the 5′ end exhibits much higher flexibility in comparison to the 3′ end, which is in good agreement with the NOE data (Figure S17). The two ethylene linkers of DAOTA-M2 can adopt *syn* and *anti* orientations, with the latter showing a higher population in MD simulations of the 1:2 CMA–ligand complex (Figure S18). The relative conformation of the morpholino groups with respect to the triangulenium core of the free and CMA-bound ligand could reduce the intramolecular quenching to differed extents and may explain the enhancement in fluorescence intensity when DAOTA-M2 is bound to G-quadruplexes.[6]

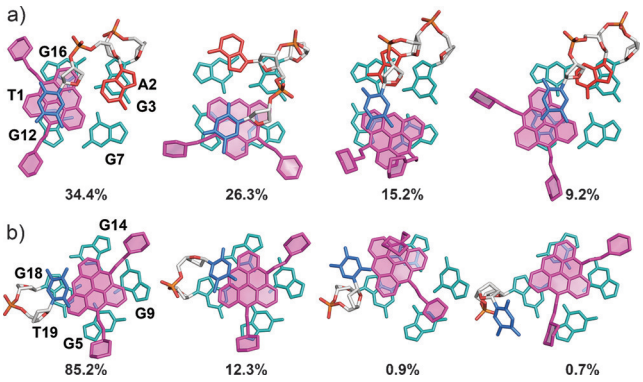

Figure 6. Representative structures of the binding modes of DAOTA-M2 with respective populations at the 5′ (a) and 3′ (b) ends of CMA in the four most populated clusters from MD simulations. DAOTA-M2 is shown in magenta, guanines in cyan, adenines in red, and thymines in blue.

Our current study provides a structural rationale for the hypothesis[6,7] that DAOTA-M2 has two different excited-state lifetimes. The favorable geometric features of DAOTA-M2 offer extensive stacking in comparison to most reported G-quadruplex ligands that cover only one or two guanines in the G-quartet.[10] The apparent differences between the binding affinities and dynamics of interaction of DAOTA-M2 could be attributed to factors such as rearrangement of extra-quartet residues that influence the position of the ligand on outer G-quartets. We observe similar rearrangement of terminal residues to those reported for complexes of G-quadruplexes with acridine BSU6039 and quindoline.[10f,k] On the other hand, we did not observe the stacking of another molecule of DAOTA-M2 and subsequent dimerization, as was observed in the daunomycin complex,[11] most probably due to the positive charge of the triangulenium core. Our studies show that the local base composition affects affinity and could therefore be used to tailor structure specificity. This could improve the fluorescence selectivity of new triangulenium derivatives by eliminating weaker binding sites and/or binding modes that are not G-quadruplex specific. Thus, modifications to DAOTA-M2 can be rationally suggested based on these structural details. Extension of the ethylene linker by 1 to 3 atoms could reduce the conformational flexibility of the morpholino groups through specific interactions with the grooves.[12] Substitution of the morpholino moiety with a more basic group such as a piperazine ring could also improve interactions with electronegative areas on the surface of the G-quadruplex (Figure S19), although caution should be exercised since the morpholino groups play an important role in defining the unique optical properties of DAOTA-M2. In summary, these results provide important insight into the binding mode of DAOTA-M2 and therefore should aid in the further development of probes with improved selectivity, affinity, and fluorescence response for the recognition and imagining of G-quadruplexes in vitro and in vivo.

Acknowledgements

This work was supported by the Slovenian Research Agency (ARRS, grants P1-242, J1-6733), the UK's Engineering and Physical Sciences Research Council (EP/H005285/1) and a Newton Fellowship for J.G.-G.

Keywords: G-quadruplexes · molecular dynamics · NMR spectroscopy · optical probes · structure elucidation

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 12508–12511
Angew. Chem. **2016**, *128*, 12696–12699

[1] D. Rhodes, H. J. Lipps, *Nucleic Acids Res.* **2015**, *43*, 8627.
[2] a) G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, *Nat. Chem.* **2013**, *5*, 182; b) C. Schaffitzel, I. Berger, J. Postberg, J. Hanes, H. J. Lipps, A. Pluckthun, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8572; c) A. Henderson, Y. Wu, Y. C. Huang, E. A. Chavez, J. Platt, F. B. Johnson, R. M. Brosh, Jr., D. Sen, P. M. Lansdorp, *Nucleic Acids Res.* **2014**, *42*, 860.
[3] a) L. Schmiedeberg, P. Skene, A. Deaton, A. Bird, *PLoS ONE* **2009**, *4*, e4636; b) U. Schnell, F. Dijk, K. A. Sjollema, B. N. G. Giepmans, *Nat. Methods* **2012**, *9*, 152.
[4] a) E. Largy, A. Granzhan, F. Hamon, D. Verga, M. P. Teulade-Fichou, *Top. Curr. Chem.* **2013**, *330*, 111; b) T. Y. Tseng, C. H. Chien, J. F. Chu, W. C. Huang, M. Y. Lin, C. C. Chang, T. C. Chang, *J. Biomed. Opt.* **2013**, *18*, 101309; c) B. R. Vummidi, J. Alzeer, N. W. Luedtke, *ChemBioChem* **2013**, *14*, 540; d) A. J. Stevens, H. L. Kennedy, M. A. Kennedy, *Biochemistry* **2016**, *55*, 3714.
[5] a) S. N. Georgiades, N. H. Abd Karim, K. Suntharalingam, R. Vilar, *Angew. Chem. Int. Ed.* **2010**, *49*, 4020; *Angew. Chem.* **2010**, *122*, 4114; b) D. Monchaud, M.-P. Teulade-Fichou, *Org. Biomol. Chem.* **2008**, *6*, 627; c) S. A. Ohnmacht, S. Neidle, *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2602; d) G. F. Salgado, C. Cazenave, A. Kerkour, J.-L. Mergny, *Chem. Sci.* **2015**, *6*, 3314; e) A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11593; f) A. Spinello, G. Barone, J. Grunenberg, *Phys. Chem. Chem. Phys.* **2016**, *18*, 2871.
[6] A. Shivalingam, M. A. Izquierdo, A. L. Marois, A. Vysniauskas, K. Suhling, M. K. Kuimova, R. Vilar, *Nat. Commun.* **2015**, *6*, 8178.
[7] A. Shivalingam, A. Vysniauskas, T. Albrecht, A. J. White, M. K. Kuimova, R. Vilar, *Chem. Eur. J.* **2016**, *22*, 4129.
[8] C. V. Dang, *Cell* **2012**, *149*, 22.
[9] a) R. I. Mathad, E. Hatzakis, J. Dai, D. Yang, *Nucleic Acids Res.* **2011**, *39*, 9023; b) M. Trajkovski, E. Morel, F. Hamon, S. Bombard, M.-P. Teulade-Fichou, J. Plavec, *Chem. Eur. J.* **2015**, *21*, 7798.
[10] a) N. H. Campbell, G. N. Parkinson, A. P. Reszka, S. Neidle, *J. Am. Chem. Soc.* **2008**, *130*, 6722; b) A. T. Phan, V. Kuryavyi, H. Y. Gaw, D. J. Patel, *Nat. Chem. Biol.* **2005**, *1*, 167; c) N. H. Campbell, M. Patel, A. B. Tofa, R. Ghosh, G. N. Parkinson, S. Neidle, *Biochemistry* **2009**, *48*, 1675; d) C. Bazzicalupi, M. Ferraroni, F. Papi, L. Massai, B. Bertrand, L. Messori, P. Gratteri, A. Casini, *Angew. Chem. Int. Ed.* **2016**, *55*, 4256; *Angew. Chem.* **2016**, *128*, 4328; e) C. Hounsou, L. Guittat, D. Monchaud, M. Jourdan, N. S. Saettel, J.-L. Mergny, M.-P. Teulade-Fichou, *ChemMedChem* **2007**, *2*, 655; f) E. Gavathiotis, R. A. Heald, M. F. G. Stevens, M. S. Searle, *J. Mol. Biol.* **2003**, *334*, 25; g) J. M. Nicoludis, S. T. Miller, P. D. Jeffrey, S. P. Barrett, P. R. Rablen, T. J. Lawton, L. A. Yatsunyk, *J. Am. Chem. Soc.* **2012**, *134*, 20446; h) G. N. Parkinson, F. Cuenca, S. Neidle, *J. Mol. Biol.* **2008**, *381*, 1145; i) T. Wilson, P. J. Costa, V. Félix, M. P. Williamson, J. A. Thomas, *J. Med. Chem.* **2013**, *56*, 8674; j) S. Sparapani, S. M. Haider, F. Doria, M. Gunaratnam, S. Neidle, *J. Am. Chem. Soc.* **2010**, *132*, 12263; k) J. Dai, M. Carver, L. H. Hurley, D. Yang, *J. Am. Chem. Soc.* **2011**, *133*, 17673.
[11] G. R. Clark, P. D. Pytel, C. J. Squire, S. Neidle, *J. Am. Chem. Soc.* **2003**, *125*, 4066.
[12] M. Micco, G. W. Collie, A. G. Dale, S. A. Ohnmacht, I. Pazitna, M. Gunaratnam, A. P. Reszka, S. Neidle, *J. Med. Chem.* **2013**, *56*, 2959.

Received: July 15, 2016
Published online: August 31, 2016

