Journal Name

COMMUNICATION

HMGB1 binds to *KRAS* promoter G-quadruplex: a new player in oncogene transcriptional regulation?†

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Jussara Amato,[a] Thushara W. Madanayake,[b] Nunzia Iaccarino,[a] Ettore Novellino,[a] Antonio Randazzo,[a] Laurence H. Hurley[b] and Bruno Pagano*[a]

This communication reports on a possible distinct role of HMGB1 protein. Biophysical studies revealed that HMGB1 binds and stabilizes the G-quadruplex-forming sequence of *KRAS* promoter element responsible for most of the transcriptional activity. Biological data showed that inhibition of HMGB1 increases *KRAS* expression. These results suggest that HMGB1 could play a role in the gene transcriptional regulation via the functional recognition of G-quadruplex.

The interaction between DNA and chromosomal proteins regulates the structure, dynamics, and function of chromosomes.[1] The high mobility group (HMG) proteins are a class of chromosomal proteins that are being recognized as essential actors in the execution and regulation of nuclear functions.[2] Among the HMGs, the high mobility group box (HMGB) protein family is the most abundant one, and comprises 4 members: HMGB1, HMGB2, HMGB3 and HMGB4.[3] HMGB1 is the most highly expressed of all the HMG family members (there are about 10^6 HMGB1 molecules per cell), and it has been extremely conserved during evolution.[3] In all cells, normal HMGB1 accumulates in nuclei to bind chromatin; however, it can shuttle between the nucleus and cytoplasm. Given its mobility, HMGB1 has been found in the cytosol, cellular membrane, and extracellular space when its nuclear localization signal is modified. The subcellular location of HMGB1 changes depending on cell type, tissue, and stress signals.[3] Nuclear HMGB1 acts as a DNA chaperone with DNA binding and bending activities and it is involved in a number of DNA activity-associated events.[3,4] Noteworthy, it regulates chromatin structure, DNA repair, recombination, and transcription. In addition to its nuclear functions, cytoplasmic HMGB1 binds many proteins involved in autophagy and cancer progression, while extracellular HMGB1 plays significant roles in inflammation, immunity, and cell growth, proliferation, and death.[3] Interestingly, HMGB1 binds rather weakly to the common B-form DNA and has practically no sequence specificity. On the other hand, it binds with high affinity to bent or distort DNAs and to non-B DNA structures like 4-way junctions and DNA bulges.[3,5-9] Moreover, HMGB1 was recently identified as telomeric G-quadruplex (G4) DNA binding partner.[10] This suggests that HMGB1 might recognize specific unusual DNA conformations to exert some of its biological functions.

Vast literature over the last years has demonstrated that certain non-B DNA conformations are transiently formed in chromosomes at specific loci, substantiating the intriguing idea that a shift in DNA structural conformation could be another layer of non-genetic or epigenetic regulator of gene expression and thereby an important determinant of cell fate.[11] Undoubtedly, most of the attention in non-B DNA-forming sequences is currently paid to G-rich DNA with the potential to form G4s.[12] Recent studies have provided evidence for the occurrence of G4s in cells and their key role in the regulation of many cellular processes and in a number of human diseases.[13] In particular, the presence of putative G4s in the promoter regions of several genes and their hypothesized functional role in transcriptional regulation may have significant biomedical implications.[14] Indeed, G4-forming sequences are found in the promoters of oncogenes implicated in cancer development and progression, and recognized as the hallmarks of cancer.[15]

One of these genes is the human *KRAS* proto-oncogene, which encodes a small GTPase transductor protein - called KRAS - involved in the regulation of cellular processes such as proliferation, differentiation, and survival.[16] *KRAS* is upregulated or mutationally activated in a multitude of cancers, including pancreatic and colon cancers. It plays a central role in controlling tumor metabolism by orchestrating multiple metabolic changes like stimulation of glucose uptake, differential channeling of glucose intermediates, reprogrammed glutamine metabolism, increased autophagy,

[a] *Department of Pharmacy, University of Naples "Federico II", Via D. Montesano 49, 80131 Naples, Italy. E-mail: bruno.pagano@unina.it*
[b] *College of Pharmacy, University of Arizona, 1703 East Mabel Street, Tucson, Arizona 85721, United States.*

† Electronic Supplementary Information (ESI) available: Detailed experimental methods and Figures S1-S4. See DOI: 10.1039/x0xx00000x

and macropinocytosis. Understanding the mechanisms underlying regulation of expression of this oncogene is of primary importance, also for possible therapeutic interventions. The *KRAS* gene promoter region, located upstream of the transcription start site, includes a nuclease hypersensitive element (NHE) that is essential for transcription regulation and that contains a G-rich strand with propensity to form G4 structures.[17] These structures are in equilibrium with the conventional duplex DNA and are facilitated by the negative supercoiling induced by transcriptional activity and other factors.[18] The biological consequence of G4 formation in the promoter element is the gene silencing.[17,18] Therefore, these noncanonical structures play an essential role in the complex mechanism of transcriptional regulation. The G4-forming sequence of *KRAS* promoter is the target of nuclear proteins,[19] which are thought to modulate the formation of G4 to control gene expression. To date, little is known about these proteins; indeed, most of G4-interacting proteins, as well as their functional relevance, remain elusive.[18] Previous studies have shown that heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) is involved in the *KRAS* transcription regulation.[20] Interestingly, this protein was found to recognize the folded G4 DNA conformation and to unfold the structure, thus facilitating the assembly of the NHE strands into the duplex. These findings support a model for transcription mechanism in which hnRNP A1 should function as a G4 destabilizing protein, thereby favoring transcription activation. In this context, it is plausible that other G4-sensing proteins may also be involved in the *KRAS* transcriptional regulation. These proteins might work to inhibit gene expression via G4 stabilization or, conversely, to recruit helicases that unfold the noncanonical structure.

In this communication, we propose that one of these *KRAS* G4 binding proteins is HMGB1. Numerous studies have reported the involvement of HMGB proteins in the transcriptional regulation of genes via their interaction with transcription regulators and chromatin.[21] Herein, we discovered that the human nuclear HMGB1 binds specifically to *KRAS* promoter G4 *in vitro* and does not unfold the DNA structure upon interaction. Preliminary biological analysis also suggested that this protein could really play a role in the oncogene transcriptional regulation. The 32-nucleotides long G-rich sequence of the NHE contained in the human *KRAS* promoter (32G, Fig. 1), and responsible for most of the transcriptional activity,[22] was selected to study in detail the hypothesized interaction. Our first goal was to evaluate the binding of HMGB1 to the G4-forming sequence 32G, in comparison with the complementary C-rich sequence (32C, Fig. 1) and the duplex formed by the two complementary strands (32ds). The structure adopted in solution by each DNA sample was first verified by circular dichroism (CD) measurements (Fig. 1). CD spectroscopy is extremely sensitive to conformational changes of nucleic acids and provides characteristic spectra depending on the structure adopted. In agreement with the presence of a parallel G4 topology, the CD spectrum of 32G at 20 °C showed a positive band at 263 nm and a negative band at around 242 nm, which are distinctive of this structure, while 32ds provides a CD spectrum with a positive band around 270 nm and a negative

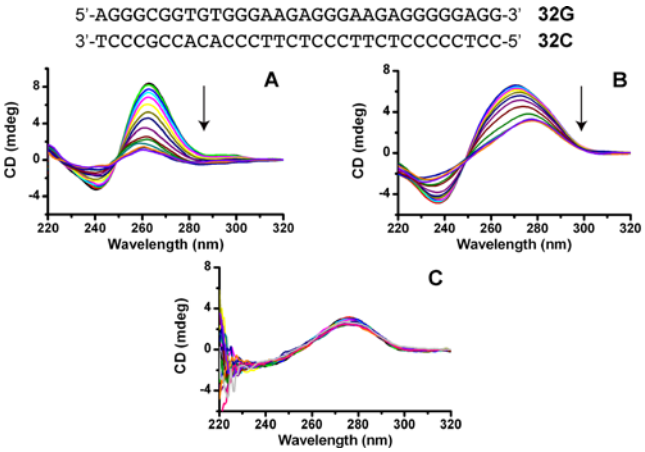

Fig. 1. Sequences of the human *KRAS* promoter investigated herein and temperature-dependent CD spectra (from 20 to 95 °C with temperature increase of 5 °C) of (A) 32G (11 μM), (B) 32ds (10 μM), and (C) 32C (10 μM) in 50 mM Tris-HCl buffer (pH 7.4) containing 100 mM K^+ (32G and 32C) or 100 mM Li^+ (32ds).

one around 238 nm, characteristic of G-rich/C-rich duplexes.[23] On the other hand, the CD spectrum of 32C was similar to that of other single-stranded oligonucleotides, e.g. $d(AC)_{12}$, with a weak positive band at 275-280 nm and a negative one of comparable intensity at about 240 nm,[24] thus showing that, as expected, this oligonucleotide is unfolded under the experimental condition used. CD melting experiments for 32G and 32ds showed sigmoidal transition curves still indicative of the formation of a G4 and duplex structure, respectively (Figs. S1A and B, ESI†). Conversely, no distinct transition was observed for 32C (Fig. S1C, ESI†), thus confirming that it remains unfolded at pH 7.4 and requires more acidic conditions to fold into an i-motif-like structure (see Figs. S1D-F, ESI†).

Surface plasmon resonance (SPR) was first used to simultaneously evaluate the binding of HMGB1 to *KRAS* promoter G4, complementary single-stranded, and duplex DNAs (Fig. 2). No detectable protein binding to both 32C single-

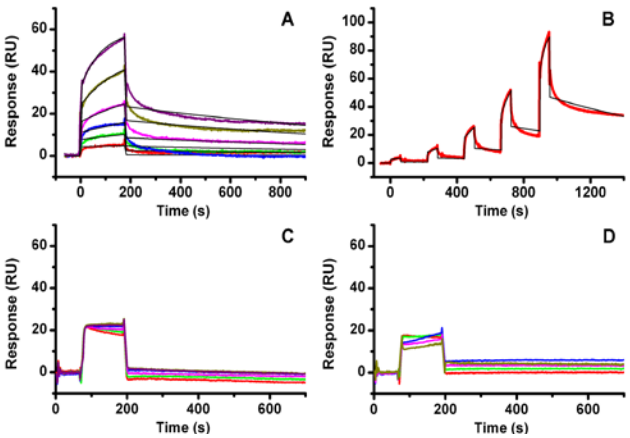

Fig. 2. Binding affinity of human HMGB1 for (A, B) 32G G-quadruplex, (C) 32C single strand, and (D) 32ds duplex measured by SPR analysis at 25 °C. Both multi-cycle (A) and single-cycle (B) kinetics experiments were performed for the analysis of 32G interaction. HMGB1 was immobilized on the SPR chip by amine coupling chemistry. Oligonucleotides were injected at a flow rate of 30 μl/min, with contact time of 120 and 60 s for multi-cycle and single-cycle kinetics, respectively. Sensograms were obtained in the oligonucleotide concentration range of 0.1–10 μM, using 50 mM Tris-HCl buffer (pH 7.4) containing 100 mM KCl as running solution for 32G and 32C, and 50 mM Tris-HCl (pH 7.4) for 32ds. Binding parameters for 32C and 32ds (C, D) were not measurable. Sensograms are shown as colored lines and their respective fits as black lines. Data are representative of at least two independent experiments.

strand and 32ds duplex was observed. Indeed, SPR sensorgrams clearly show that, in both cases, the SPR response does not change by increasing the concentration of DNA molecules (the responses are superimposable to the bulk effect). On the other hand, SPR experiments on 32G, performed employing both multi-cycle and single-cycle kinetics methods (Figs. 2A and B, respectively), showed a response proportional to DNA concentration, indicative of a specific interaction with the G4. To evaluate the affinity, the obtained binding curves were fitted to a simple 1:1 binding model, giving a calculated dissociation constant (K_d) value of 4.0 (±1.0) x 10^{-7} M. Interestingly, although the affinity for 32G was not as high as the affinity displayed for other types of non-B DNA structures,[5,8] it was found to be similar to that previously observed for the specific binding of HMGB1 to cisplatin-modified DNA (K_d of 3.7 x 10^{-7} M),[25] and two orders of magnitude higher than for the nonspecific binding to B-form DNA (K_d of about 5 x 10^{-5} M).[25]

To verify the interaction observed on a surface, spectroscopic methods that could allow us to further characterize the binding in solution were subsequently employed. In particular, fluorescence titration experiments were performed by monitoring the changes in the intrinsic fluorescence of protein upon binding to DNA.[26] Fluorescence emission spectra of HMGB1 in the absence and presence of increasing amounts of 32G were recorded (Fig. 3A). As a result of complex formation, a significant quenching of fluorescence intensity on increasing G4 concentration was observed until saturation was reached. The fraction of bound protein (α) at each point of the titration was calculated and plotted against the DNA concentration to obtain an isotherm binding curve. The curve was fitted by non-linear regression to an independent and equivalent binding sites model, giving an equilibrium dissociation constant (K_d) of 6.0 (±1.0) x 10^{-7} M and a stoichiometry of 1:1. Moreover, microscale thermophoresis (MST) measurements were performed by titrating FAM-labeled 32G (F32G), 32ds (F32ds) and 32C (F32C) with increasing concentrations of HMGB1 (Figs. 3B and S2). The results of these experiments indicated that HMGB1 binds to F32G and F32ds with a K_d of 4.7 (±1.4) x 10^{-7} M and 3.2 (±1.7) x 10^{-5} M, respectively. Whereas no significant

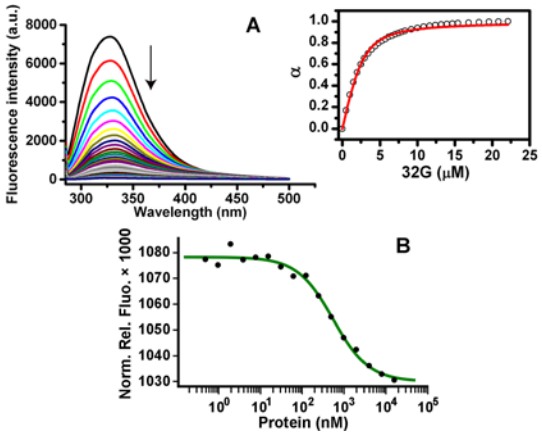

Fig. 3. (A) Fluorescence emission spectra of HMGB1 (3 μM) in the absence and presence of stepwise additions (5 μL) of 32G (100 μM) at 25 °C. The inset shows the titration curve obtained by plotting the fraction of bound protein (α) versus DNA concentration. (B) MST measurement of 5'FAM-labeled 32G (23 nM) recorded by adding increasing concentrations of HMGB1 (0.48 nM-15.7 μM).

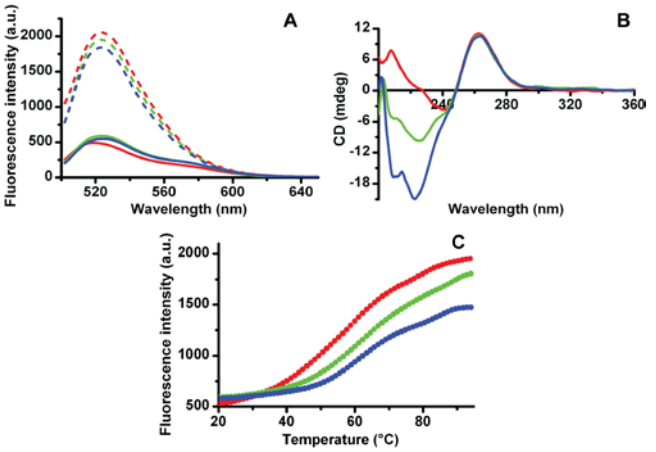

Fig. 4. (A) Fluorescence emission spectra of labeled F32GT (0.1 μM) obtained exciting FAM at 490 nm in the absence (red) and presence of 1 (green) or 2 (blue) molar equivalents of HMGB1, recorded at 20 °C (solid lines) and 100 °C (dashed lines). (B) CD spectra of 32G (1 μM) alone (red) and of 1:1 (green) and 1:2 (blue) DNA/HMGB1 mixtures. (C) FRET-melting curves of F32GT (0.1 μM) in the absence (red) and presence of 1 (green) or 2 (blue) equivalents of HMGB1.

change of the thermophoretic signal was observed for 32C (Fig. S2B, ESI†), thus indicating no binding.

Nuclear proteins have been identified that bind to *KRAS* promoter G4 and exhibit a destabilizing effect on the DNA structure.[20] To gain insight into the effect of HMGB1 on the G4 stability and conformation, Förster resonance energy transfer (FRET) experiments were performed by using the G4-forming sequence 32G dually labeled with the donor FAM and the acceptor TAMRA at the 5' and 3' ends, respectively (F32GT).[26] The emission intensity of the fluorescent probe depends on its distance from the acceptor, and this distance significantly changes in the G4 folding/unfolding process. When F32GT is in the G4 conformation, the two dyes are close to each other and the fluorescence emission of FAM is efficiently quenched by TAMRA. When the G4 unfolds, the distance between the two dyes increases and the intensity of FAM emission increases. FRET experiments were carried out by adding increasing amounts of HMGB1 to a F32GT solution at 20 °C and monitoring the fluorescence emission spectrum of FAM (Fig. 4A). No significant change in the fluorescence spectrum was observed even in the presence of two equivalents of protein, suggesting that HMGB1 does not unfold the G4 upon binding. On the other hand, considerable changes in fluorescence intensity are observed, as expected, recording the same spectra at 100 °C, temperature at which the G4 is unfolded.

CD experiments were also performed to further monitor the unfolding or any possible change in the secondary structure of 32G that might be induced by the interaction with the protein (Fig. 4B). In agreement with FRET experiments, CD analysis showed that no unfolding or structural alterations of DNA G4 occur upon binding to HMGB1. Indeed, the presence of the protein (up to 2 molar equivalents) does not affect the CD spectrum of 32G within the wavelength range 240–360 nm. In this range, the CD spectrum of the complex retains the same shape and signal intensity observed for 32G alone.

Furthermore, to evaluate if the protein has a stabilizing effect on the G4 structure, FRET and CD melting experiments were

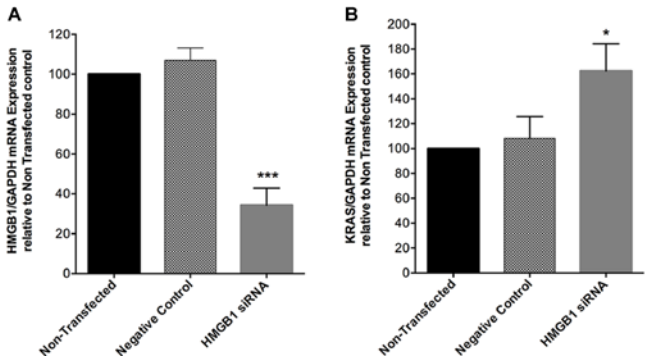

Fig. 5. The effect of HMGB1 siRNA on *KRAS* mRNA expression. (A) HMGB1 depletion with HMGB1 siRNA transfection. (B) *KRAS* mRNA expression in response to HMGB1 knockdown. mRNA expressions are relative to GAPDH and normalized to non-transfected control. Data are from mean values of four experiments. Statistical significance relative to non-transfected control is indicated by *P<0.05, ***P<0.001.

carried out. When HMGB1 was added to the solution of the labeled G4, FRET (Fig. 4C) and CD (Fig. S3, ESI†) melting curves showed an increase in the thermal stability of the DNA structure, thus confirming the nature of the interaction. FRET melting experiments were also performed on FAM-TAMRA dual-labeled 32ds (F32dsT) for comparison (Fig. S4, ESI†). No change in duplex thermal stability was observed, proving once again the weak/nonspecific interaction with HMGB1.

Finally, to determine the effect of HMGB1 reduction on *KRAS* mRNA expression, HMGB1 siRNA was transiently transfected into Panc1 cells. The transfection was resulted 50-80% reduction of HMGB1 relative to GAPDH (Fig. 5A) as shown by qRT-PCR. The knockdown of HMGB1 was significant relative to non-transfected cell control. KRAS levels were significantly upregulated (~60%) in the HMGB1 mRNA reduced samples compared to the non-transfected control (Fig. 5B). These results suggest that the inhibition of HMGB1 gene increases the *KRAS* gene expression at mRNA level.

It has been shown that the *KRAS* promoter element has 3 sequential G4-forming regions,[27] which on the basis of their proximity to the transcriptional start site have been termed the Near- (-129 to -159), Mid- (-174 to -228) and Far-regions (-238 to -275). Some of us recently proposed these could serve as a buffer (Near-region), active switch (Mid-region) and brake (Far-region).[18] Although it cannot be ruled out that HMGB1 may act as a mediator that detects the G4 and recruits helicases or other proteins, the results of this study suggest that HMGB1 might play a complementary role to hnRNP A1, which unfolds the here investigated Near-region G4. This supports the idea that this G4 might actually act as a buffer element controlled by functional proteins.

In summary, our data provide the first evidence that HMGB1 is able to bind to and stabilize the Near-region *KRAS* promoter G4, thus suggesting an active role for this protein in the mechanism underlying the regulation of oncogene transcription. These findings lay the basis for further studies aiming to elucidate the role of protein/G4 functional interactions in the control of oncogene expression in cancer cells.

This work was supported by the Italian Association for Cancer Research (IG-16730 to BP and IG-18695 to AR) and the National Institutes of Health (5RO1CA153821 to LHH). We are grateful to Prof. Michael T. Lotze for inspiring this work.

Conflicts of interest

LHH has a financial interest in Reglagene.

Notes and references

1 G. S. Stein, T. C. Spelsberg and L. J. Kleinsmith, *Science*, 1974, **183**, 817; S. C. Elgin and H. Weintraub, *Annu. Rev. Biochem.*, 1975, **44**, 725.
2 M. E. Bianchi and A. Agresti, *Curr. Opin. Genet. Dev.*, 2005, **15**, 496.
3 R. Kang, R. Chen, Q. Zhang, W. Hou, S. Wu, L. Cao, J. Huang, Y. Yu, X. G. Fan, Z. Yan, X. Sun, H. Wang, Q. Wang, A. Tsung, T. R. Billiar, H. J. 3rd Zeh, M. T. Lotze and D. Tang, *Mol. Aspects Med.*, 2014, **40**, 1.
4 A. Agresti and M. E. Bianchi, *Curr. Opin. Genet. Dev.*, 2003, **13**, 170.
5 M. E. Bianchi, M. Beltrame and G. Paonessa, *Science*, 1989, **243**, 1056.
6 D. M. J. Lilley, *Nature*, 1992, **357**, 282.
7 C. Gaillard and F. Strauss, *Science*, 1994, **264**, 433.
8 S. Jaouen, L. de Koning, C. Gaillard, E. Muselíková-Polanská, M. Stros and F. Strauss, *J. Mol. Biol.*, 2005, **353**, 822.
9 K. Furuita, S. Murata, J. G. Jee, S. Ichikawa, A. Matsuda and C. Kojima, *J. Am. Chem. Soc.*, 2011, **133**, 5788.
10 B. Pagano, L. Margarucci, P. Zizza, J. Amato, N. Iaccarino, C. Cassiano, E. Salvati, E. Novellino, A. Biroccio, A. Casapullo and A. Randazzo, *Chem. Commun.*, 2015, **51**, 2964.
11 R. D. Wells, R. W. Blakesley, S. C. Hardies, G. T. Horn, J. E. Larson, E. Selsing, J. F. Burd, H. W. Chan, J. B. Dodgson, K. F. Jensen, I. F. Nes and R. M. Wartell, *CRC Crit. Rev. Biochem.*, 1977, **4**, 305; J. Zhao, A. Bacolla, G. Wang and K. M. Vasquez, *Cell. Mol. Life Sci.*, 2010, **67**, 43; M. Bansal, A. Kumar and V. R. Yella, *Curr. Opin. Struct. Biol.*, 2014, **25**, 77.
12 M. L. Bochman, K. Paeschke and V. A. Zakian, *Nat. Rev. Genet.*, 2012, **13**, 770.
13 D. Rhodes and H. J. Lipps, *Nucleic Acids Res.* 2015, **43**, 8627; N. Maizels, *EMBO Rep.*, 2015, **16**, 910.
14 T. Tian, Y.-Q. Chen, S.-R.Wang and X. Zhou, *Chem.*, 2018, **4**, 1314.
15 T. A. Brooks, S. Kendrick and L. Hurley, *FEBS J.*, 2010, **277**, 3459.
16 Y. Pylayeva-Gupta, E. Grabocka and D. Bar-Sagi, *Nat. Rev. Cancer*, 2011, **11**, 761.
17 S. Cogoi and L. E. Xodo, *Nucleic Acids Res.*, 2006, **34**, 2536.
18 C. E. Kaiser, N. A. Van Ert, P. Agrawal, R. Chawla, D. Yang and L. H. Hurley, *J. Am. Chem. Soc.*, 2017, **139**, 8522.
19 S. Cogoi, M. Paramasivam, B. Spolaore and L. E. Xodo, *Nucleic Acids Res.*, 2008, **36**, 3765.
20 M. Paramasivam, A. Membrino, S. Cogoi, H. Fukuda, H. Nakagama and L. E. Xodo, *Nucleic Acids Res.*, 2009, **37**, 2841.
21 T. Ueda and M. Yoshida, *Biochim. Biophys. Acta*, 2010, **1799**, 114.
22 F. Yamamoto and M. Perucho, *Oncogene Res.*, 1988, **3**, 125.
23 M. Vorlíčková, I. Kejnovská, K. Bednářová, D. Renčiuk and J. Kypr, *Chirality*, 2012, **24**, 691.
24 C. L. Clark, P. K. Cecil, D. Singh and D. M. Gray, *Nucleic Acids Res.*, 1997, **25**, 4098.
25 P. M. Pil and S. J. Lippard, *Science*, 1992, **256**, 234; S. S. Lange and K. M. Vasquez, *Mol. Carcinog.*, 2009, **48**, 571.
26 C. Giancola and B. Pagano, *Top. Curr. Chem.*, 2013, **330**, 211.
27 R. K. Morgan, H. Batra, V. C. Gaerig, J. Hockings and T. A. Brooks, *Biochim. Biophys. Acta - Gene Regul. Mech.*, 2016, **1859**, 235.

