Small-molecule affinity capture of DNA/RNA quadruplexes and their identification *in vitro* and *in vivo* through G4RP protocol

Isaline Renard,[1] Michael Grandmougin,[1] Apolline Roux,[1] Sunny Y. Yang,[2] Pauline Lejault,[1] Marc Pirrotta,[1] Judy M. Y. Wong[2] and David Monchaud[1,*]

[1] Institut de Chimie Moléculaire, ICMUB CNRS UMR6302, UBFC Dijon, France.
[2]Faculty of Pharmaceutical Sciences, The University of British Columbia, Vancouver, Canada.

I. Material and methods.

All chemicals were purchased from Sigma-Aldrich, except for Boc-^{PNA}G(Z)-OH purchased from ASM Research Chemicals and Biotin-PEG4-CO2H purchased from Iris Bioetch, and used without further purification. NMR spectra were recorded with a Bruker 300 Avance III NanoBay spectrometer (300 MHz for ^{1}H and 75 MHz for ^{13}C). Chemical shifts are reported in δ ppm, using [D] solvents as standards. The following abbreviations are used: s: singlet, d: doublet, t: triplet, q: quartet, m: multiplet, br: broad. MALDI-TOF (Matrix-Assisted Laser Desorption/Ionization - Time Of Flight) mass spectrometry was carried out using a Bruker Ultraflex II LRF 2000 spectrometer and ESI (ElectroSpray Ionization) mass spectrometry was carried out using a Bruker Daltonics AmaZon SL spectrometer. High Resolution Mass Spectrometry (HRMS) experiments were performed on a LTQ Orbitrap XL (Thermo Scientific) spectrometer equipped with Ion Max source and HESI-II probe. Purification: Flash chromatography was performed using the Automatic Reveleris® Flash Chromatography System (GRACE) equipped with multiple channel detection (UV (201 and/or 214 nm) and ELSD (Evaporative Light Scattering Detection)) with Reveleris® C18 RP 80 g cartridges. Linear gradient was used CH_3CN/0.1% TFA and H_2O/0.1% TFA as eluents (at 0% CH_3CN/0.1% TFA (12 min) from 0 to 10% CH_3CN/0.1% TFA (10 min), from 10 to 40% CH_3CN/0.1% TFA (5 min) then from 40 to 55% CH_3CN/0.1% TFA (5 min), at a flow rate of 60 mL/min). Semi-preparative RP-HPLC was performed using a Thermo Beta Basic C18 column, 5 µm, 30 $\times$ 150 mm with a UV-Vis detection thanks to an Ultimate 3000 diode array detector at 201, 214, 222, 260 nm. Linear gradient was used CH_3CN/0.1 % HCO_2H and H_2O/0.1 % HCO_2H as eluents (from 5

to 50 % CH_3CN/0.1 % HCO_2H (45 min), at a flow rate of 15 mL/min). HPLC-analytic was performed on a Thermo-Dionex ultimate 3000 instrument using a Phenomenex Kinetex C18 column, 2.6 µm, 2.1 × 50 mm equipped with a diode array detector (Thermo-Dionex DAD 3000-RS). Linear gradient was used CH_3CN/0.1 % HCO_2H and H_2O/0.1 % HCO_2H as eluents (from 0 to 20 % CH_3CN/0.1 % HCO_2H (5 min) then 100% CH_3CN (1.5min), at a flow rate of 0.5 mL/min).

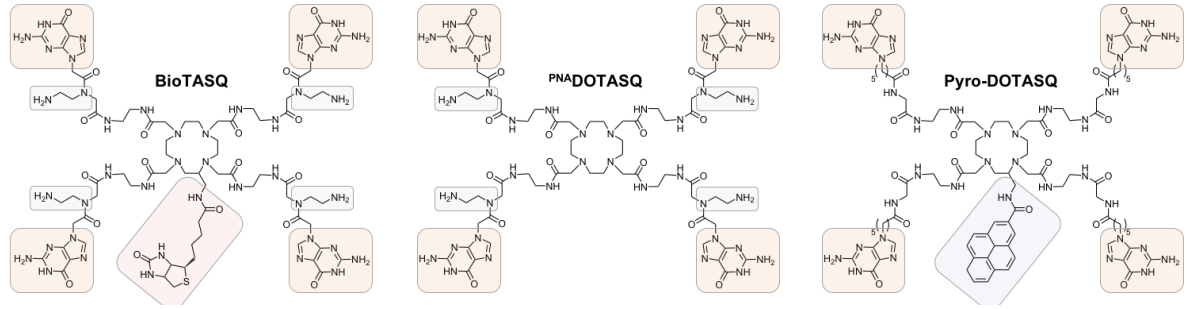

Figure S1. Chemical structures of BioTASQ v.1, ^{PNA}DOTASQ and Pyro-DOTASQ.

II. Design of BioTASQ v.2.

The design of **BioTASQ v.1** (Figure S1, left) was directly inspired by two previously reported TASQs, *i.e.*, **^{PNA}DOTASQ** (Figure S1, center, cf. Haudecoeur, R. et al. *J. Am. Chem. Soc.* **2013**, *135*, 550-553), considered here as the parent compound, and **Pyro-DOTASQ** (Figure S1, right, cf. Laguerre, A. et al. *Chimia* **2015**, *69*, 530-536), in which a pyrene moiety was grafted directly on the DOTA template. The synthesis of **BioTASQ v.1** was fully described in Yang, S.Y. et al. *Nat. Commun*. **2018**, *9*, 4730. The **BioTASQ v.2** differs from **BioTASQ v.1** in the length of the biotin appendage: while the biotin is directly linked to the aminomethylcyclen (or **AMC**) core in the **BioTASQ v.1** structure (Figures S1 and S2), we selected a $(PEG)_4$-type linker in between the biotin and the **AMC** core for the design of **BioTASQ v.2** (Figure S2). To this end, we used the commercially available **Biotin-PEG4-CO2H** (Iris Biotech), resulting in a 23-membered flexible linker between the biotin and DOTA moieties. This approach was designed to discard the possible steric clash that might occur upon MagneSphere® addition (Figure S2).

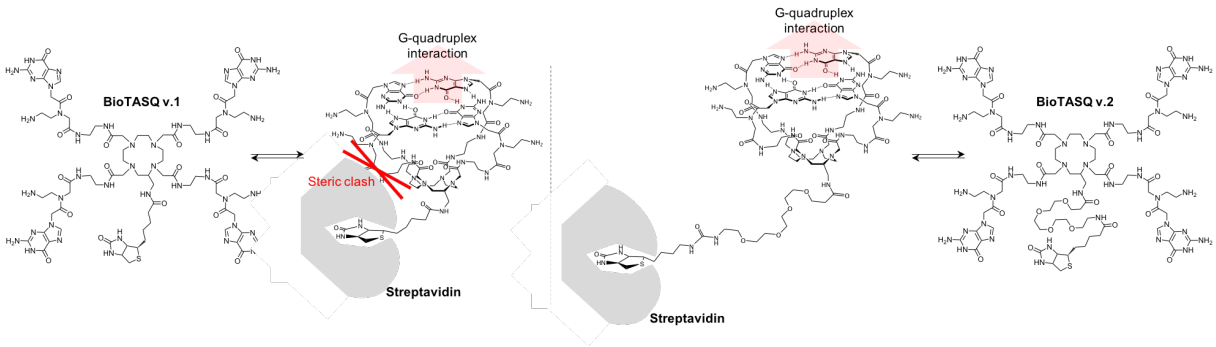

Figure S2. Left panel: the possible steric clash between BioTASQ v.1 and the streptavidin upon biotin/streptavidin binding. Right panel: the structure of BioTASQ v.2.

III. Chemical synthesis of BioTASQ v.2.

a. BioTASQ v.2 synthesis

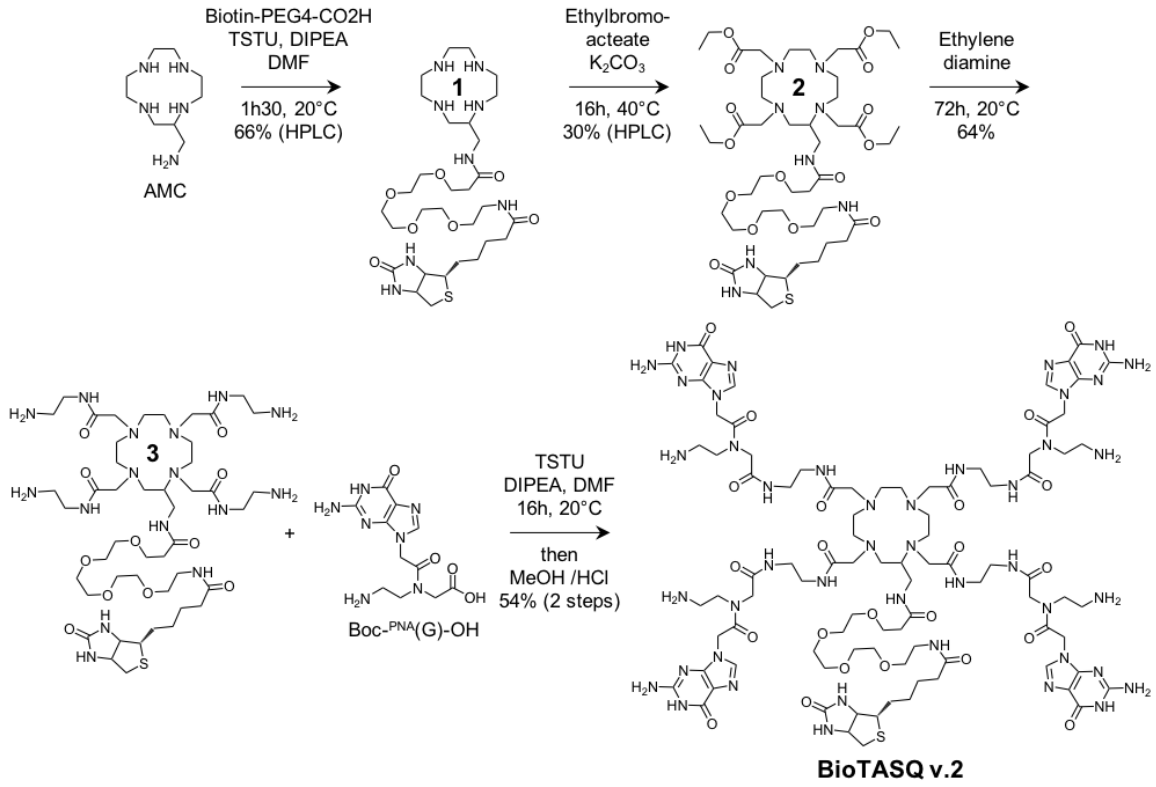

Figure S3. Synthetic scheme for the synthesis of **BioTASQ v.2**

b. **The AMC and Boc-^{PNA}(G)-OH precursors:** The aminomethylcyclen, or **AMC**, precursor was obtained according to the methodology described in Rousselin, Y. et al. *Eur. J. Org. Chem.* **2010**, 1688-1693. The **Boc-^{PNA}(G)-OH** precursor was obtained according to the methodology described in Haudecoeur, R. et al. *J. Am. Chem. Soc.* **2013**, *135*, 550-553.

c. **Compound 1:** TSTU (337 mg, 1.12 mmol), Biotin-PEG4-CO2H (500 mg, 1.02 mmol) and DIPEA (0.2 mL, 1.12 mmol) were dissolved in DMF (10 mL). The mixture was stirred at room temperature for 1 h. A solution of **AMC** (205 mg, 1.02 mmol) and DIPEA (0.35 mL, 2.03 mmol) in DMF (5 mL) was slowly added and the mixture was stirred at room temperature overnight. The solvent was evaporated under reduced pressure and the crude mixture was purified by flash chromatography to afford **1** as a white solid (456 mg, 66%). MS (ESI-MS): m/z $[C_{31}H_{61}N_8O_7S+H]^+$ = 675.30; $[M+Na]^+$ = 697.26. ^{1}H NMR (500 MHz, d6-DMSO, 300 K): δ (ppm) = 8.08 (m, 1 H), 7.84 (m, 1 H), 6.42 (br. s, 2 H), 4.30-4.10 (m, 2 H), 3.48-3.17 (m, 20 H), 2.98-2.83 (m, 14 H), 2.05 (t, 2 H), 1.61-1.28 (m, 8 H). **-> Figure S4**.

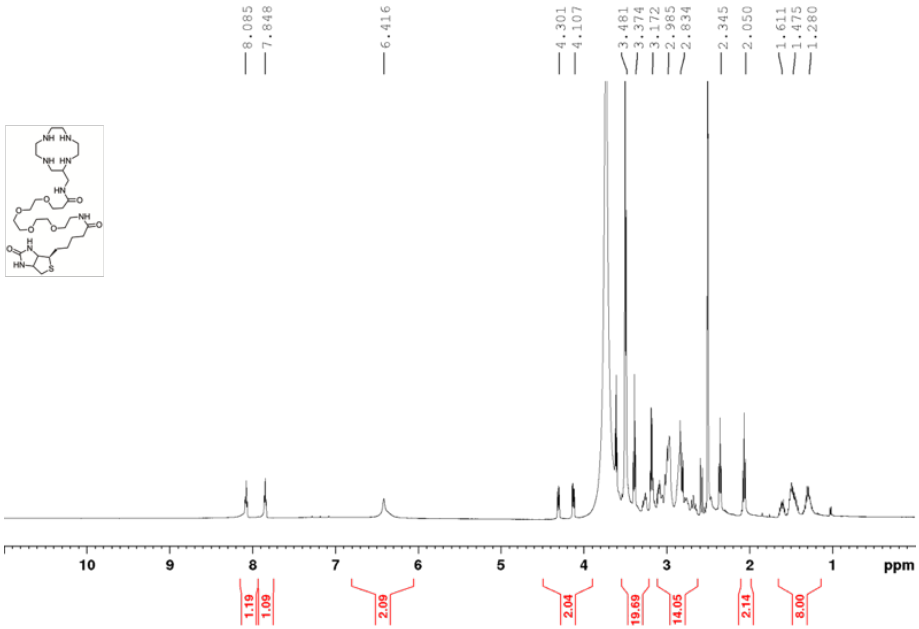

Figure S4: ^{1}H NMR (500 MHz, d6-DMSO, 300K) of **1**

d. **Compound 2:** To a solution of **1** (456 mg, 0.68 mmol) and K_2CO_3 (750 mg, 54.0 mmol) in acetonitrile (10 mL) was added ethylbromoacetate (306 µL, 2.7 mmol). The mixture was stirred overnight at 40 °C. It was filtered and washed with acetonitrile. The filtrate was evaporated under reduced pressure. The crude mixture was purified by semi-preparative HPLC to afford **2** as a yellow oil (220 mg, 30 %). MS (ESI-MS): m/z $[C_{46}H_{82}N_8O_{15}S+Na]^+$ = 1041.59. ^{1}H NMR (500 MHz, d6-DMSO, 300 K): δ (ppm) = 7.91-7.86 (m, 2 H), 6.41 (m, 1 H), 6.36 (m, 1 H), 4.33-4.23 (m, 8 H), 4.08 (m, 2 H), 3.61-3.58 (m, 20 H), 3.34-3.11 (m, 14 H), 2.08 (t, 2 H), 1.62-1.49 (m, 8 H), 1.22-1.15 (m, 12 H) **-> Figure S5**.

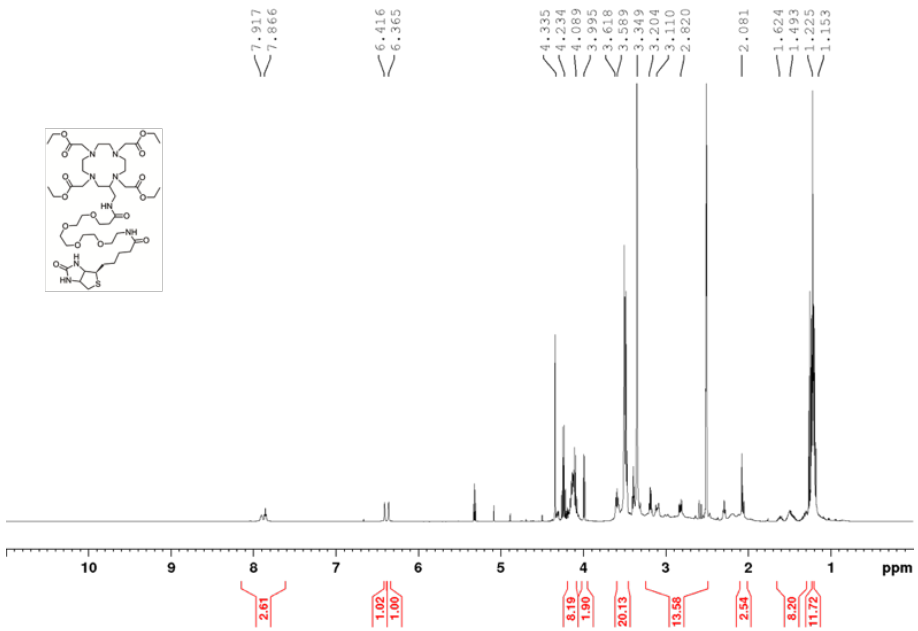

Figure S5: ^{1}H NMR (500 MHz, d6-DMSO, 300K) of **2**

e. **Compound 3:** A solution of **2** (203 mg, 0.17 mmol) in ethylene diamine (10 mL) was stirred at room temperature for 3 days. Ethylene diamine was evaporated under reduced pressure. The mixture was triturated in acetonitrile to afford **3** as a yellow solid (136 mg, 64%), which is directly engaged in the following step. ^{1}H NMR (500 MHz, d6-DMSO, 300 K): δ (ppm) = 7.86 (m, 4 H), 6.47 (s, 1 H), 6.37 (s, 1 H), 4.12 (m, 18 H), 3.59-3.50 (m, 20 H), 3.06-2.88 (m, 8 H), 2.88-2.70 (m, 14 H), 2.05 (m, 2 H), 1.61-1.26 (m, 8 H) **-> Figure S6**.

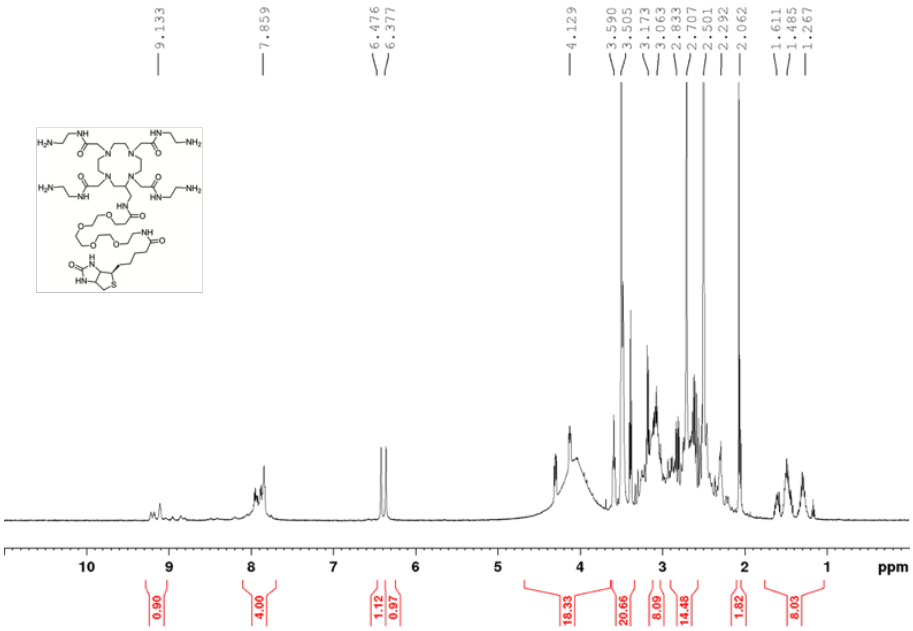

Figure S6: ^{1}H NMR (500 MHz, d6-DMSO, 300K) of **3**

f. **BioTASQ v.2:** Boc-^{PNA}G-OH (2, 31 mg, 0.57 mmol), TSTU (170 mg, 0.57 mmol) and DIPEA (98 µL, 0.57 mmol) were dissolved in DMF (5 mL). The mixture was stirred at room temperature for 30 min. A solution of **3** (129 mg, 0.11 mmol) and DIPEA (39 µL, 0.23 mmol) in DMF (2 mL) was then added dropwise at 0 °C and the mixture was stirred overnight at room temperature. The solvent was removed under reduced pressure. The crude mixture was purified by semi-preparative HPLC: the protected **BioTASQ v.2** was obtained as a yellow oil, diluted in TFA (10 mL) and stirred for 1h at room temperature. The final compound is then precipitated upon addition of a large excess of Et_2O; the solid is collected, triturated 3 times with a large excess of Et_2O. After removal of the solvent under reduced pressure, **BioTASQ v.2** is isolated as a white powder (13.7 mg, 54% over two steps). ESI-MS (HRMS): m/z $[C_{90}H_{142}N_{44}O_{23}S+H]^+$ = 2238.08703. ^{1}H NMR (500 MHz, d6-DMSO, 300 K): δ (ppm) = 8.51-8.33 (m, 4 H), 8.00 (m, 4 H), 7.76 (s, 8 H), 6.55 (m, 4 H), 5.08 (m, 4 H), 4.84 (m, 4 H), 4.17-3.89 (m, 20 H), 3.89-3.49 (m, 16 H), (m, 14 H), 3.18-2.98 (m, 14 H), 2.06 (m, 2 H), 1.61-1.26 (m, 8 H) **-> Figures S7-S9**.

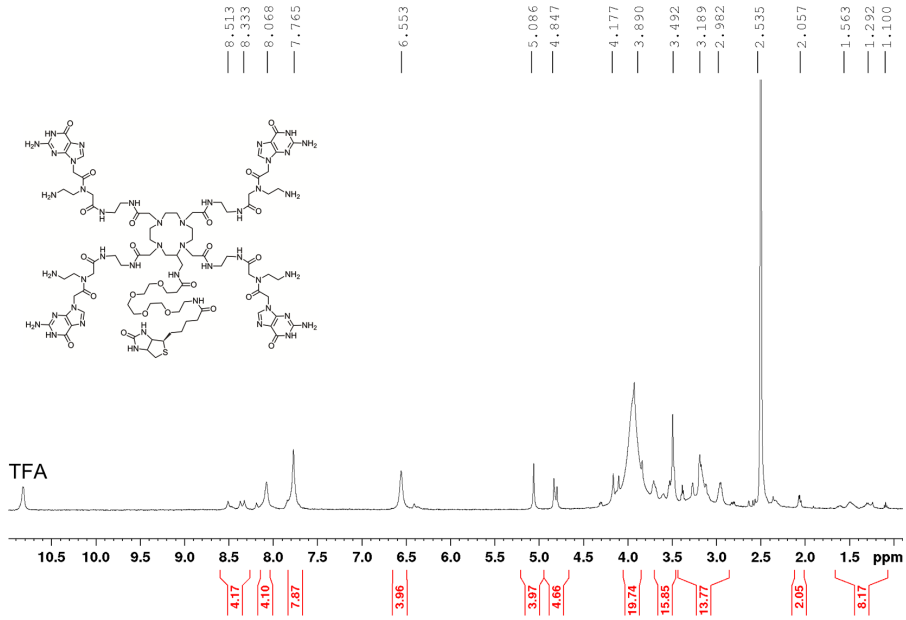

Figure S7: ^{1}H NMR (500 MHz, d6-DMSO, 300K) of BioTASQ v.2

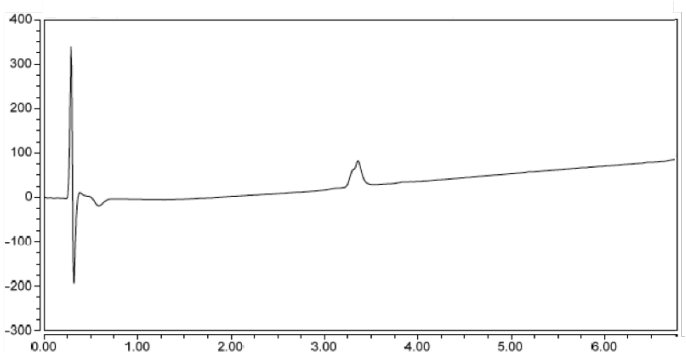

Figure S8: HPLC profile of BioTASQ v.2

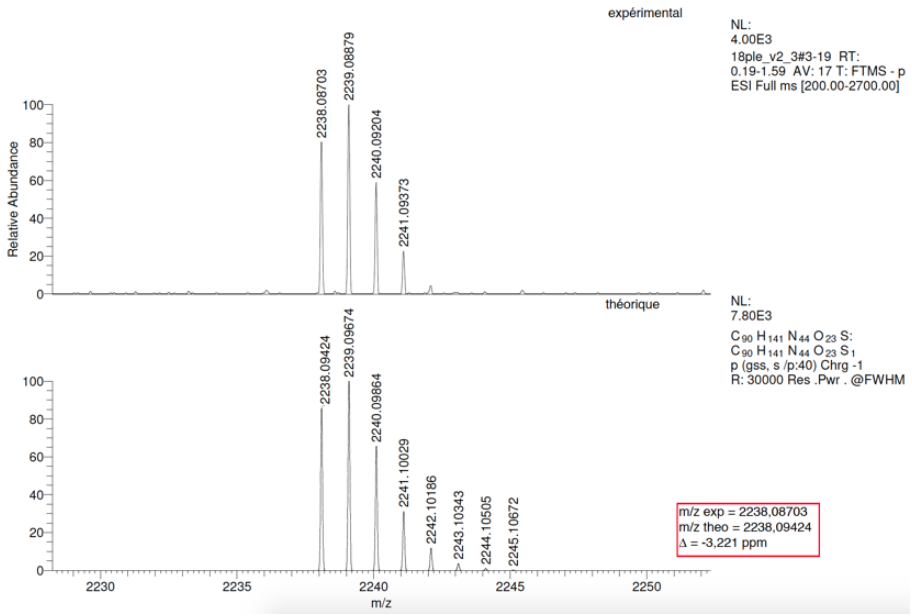

Figure S9: HRMS analysis of BioTASQ v.2.

IV. Preparation of the oligonucleotides

The sequences of non-labeled, labeled and doubly labeled oligonucleotides used herein were: F-Myc (FAM-d[$^{5'}$GAG$_3$TG$_4$AG$_3$TG$_4$A$_2$G$^{3'}$]), F-SRC (FAM-d[$^{5'}$G$_3$AG$_3$AG$_3$CTG$_5{}^{3'}$]), F-22AG (FAM-d[$^{5'}$AG$_3$(T$_2$AG$_3$)$_3{}^{3'}$]), F-duplex (FAM-d[$^{5'}$(TA)$_2$GC(TA)$_2$T$_6$(TA)$_2$GC(TA)$_2{}^{3'}$]), F-TERRA (FAM-r[$^{5'}$G$_3$(U$_2$AG$_3$)$_3{}^{3'}$]), F-TRF2 (FAM-r[$^{5'}$CG$_3$AG$_3$CG$_4$AG$_3$C$^{3'}$]), F-NRAS (FAM-r[$^{5'}$G$_3$AG$_4$CG$_3$UCUG$_3{}^{3'}$]), F-Myc-T (FAM-d[$^{5'}$GAG$_3$TG$_4$AG$_3$TG$_4$A$_2$G$^{3'}$]-TAMRA), ds17 (d[$^{5'}$C$_2$AGT$_2$CGTAGTA$_2$C$_3{}^{3'}$]/d[$^{5'}$G$_3$T$_2$ACTACGA$_2$CTG$_2{}^{3'}$]) and ds26 (d[$^{5'}$CA$_2$TCG$_2$ATCGA$_2$T$_2$CGATC$_2$GAT$_2$G$^{3'}$]). The lyophilized DNA strands purchased from Eurogentec (Seraing, Belgium) were firstly diluted at 500 µM in deionized water (18.2 MΩ.cm resistivity), expect calf thymus DNA (ctDNA) purchased form Sigma-Aldrich and slowly diluted at 3.1 mM in deionized water and used without further preparation. DNA/RNA structures were prepared in a Caco.K buffer, comprised of 10 mM lithium cacodylate buffer (pH 7.2) plus 10 mM KCl/90 mM LiCl. Quadruplex structures were prepared by mixing 40 µL of the constitutive strand (500 µM) with 8µL of a lithium cacodylate buffer solution (100 mM, pH 7.2), plus 8µL of a KCl/LiCl solution (100 mM/900 mM) and 24µL of water. The duplex structures were prepared by mixing 40 µL of each constitutive strand (500 µM) with 16µL of a lithium cacodylate buffer solution (100 mM, pH 7.2), plus 16µL of a KCl/LiCl solution (100 mM/900 mM) and 48 µL of water. The final concentrations were theoretically 250 and 125 µM, for mono- and bimolecular DNA structures, respectively; the actual concentration of each sample was determined through a dilution to 1 µM theoretical concentration *via* a UV spectral analysis at 260 nm (after 5 min at 90 °C) with the following molar extinction coefficient values: F-Myc (253000), F-SRC (204300), F-22AG (249500), F-duplex (279900), F-TERRA (257900), F-TRF2 (198700), F-NRAS (199300), 232000 (F-myc-T), 328300 (ds17) and 506400 $M^{-1}.cm^{-1}$ (ds26). The higher-order structures were folded according to two procedures: (a) for the monomolecular architectures (i.e., DNA/RNA quadruplexes and F-duplex), solutions were heated (90 °C, 5 min), cooled on ice (7 h) and then stored overnight (4 °C); (b) for the folding of bimolecular structures (ds17, ds26), the solutions were heated (90 °C, 5 min), gradually cooled (65, 60, 55, 50, 40 and 30 °C (30 min/step), 25 °C (2 h)) and then stored at 4 °C.

V. FRET-melting assay

Experiments were performed in a 96-well format using a Mx3005P qPCR machine (Agilent) equipped with FAM filters (λ_{ex} = 492 nm; λ_{em} = 516 nm) in 100 µL (final volume) of 10 mM lithium cacodylate buffer (pH 7.2) plus 1 mM KCl/99 mM LiCl with 0.2 µM of F-Myc-T and 0, 1 or 5 µM of both BioTASQ, ^{PNA}DOTASQ and commercially available biotin. After a first equilibration step (25 °C, 30 s), a stepwise increase of 1 °C every 30 s for 65 cycles to reach 90 °C was performed, and measurements were made after each cycle. Final data were analyzed with Excel (Microsoft Corp.) and OriginPro®9.1 (OriginLab Corp.). The emission of FAM was normalized (0 to 1), and $T_{1/2}$ was defined as the temperature for

which the normalized emission is 0.5; $\Delta T_{1/2}$ values, calculated as follows: $\Delta T_{1/2}$ = [$T_{1/2}$(F-Myc-T+ligand)-$T_{1/2}$(F-Myc-T)], are means of 3 experiments.

VI. in vitro pull-down assays

A. Two-step pull-down protocol. The *in vitro* quadruplex capture experiments were performed in 250 μL final volume as follows: BioTASQ (20-50 μM) was mixed with 5'-fluorescently labeled Myc oligonucleotide (F-Myc, FAM-d[$^{5'}$GAG$_3$TG$_4$AG$_3$TG$_4$A$_2$G$^{3'}$]) (1 μM) (or F-TERRA and F-SRC for control experiments, cf. Figure S12) in Caco.K buffer, comprised of 10 mM lithium cacodylate buffer (pH 7.2) plus 10 mM KCl/90 mM LiCl for 2 h at 25 °C. The mixtures were thus taken up for the first fluorescence analysis (fluorescence analysis 1), being distributed in 3 wells (80 μL each) of a 96-well plate, using a ClarioStar® machine (BMG Labtech) equipped with FAM filters (λ_{ex} = 492 nm; λ_{em} = 516 nm). The streptavidin MagneSphere® beads (Promega) were washed 3 times with Caco.K buffer: the supernatant of 200 μL of the commercial solution of beads (1 mg/mL) was removed and the beads were taken up in 200 μL of Caco.K, and this washing procedure was repeated 3 times. After the original solutions being reconstituted (as 250-μL mixtures), 25 μL (corresponding to 10 % of the final volume, 25 μg) of MagneSphere® beads were added to the BioTASQ/F-Myc mixtures; after 1-h incubation at 25 °C, beads were immobilized (attracted by a magnet) and the supernatant was taken up for analysis (fluorescent analysis 1), after being distributed in 3 wells (80 μL each) of a 96-well plate. The solid residue was resuspended in 240 μL of 10X TBE buffer, heated for 8 min at 90 °C (under gentle stirring 800 r.p.m.), centrifuged for 2 min (8,900 rpm) and the supernatant was taken up for analysis (fluorescent analysis 2), after being distributed in 3 wells (80 μL each) of a 96-well plate. Control experiments are performed without BioTASQ and/or without beads. Data were analyzed with Excel (Microsoft Corp.) and OriginPro®9.1 (OriginLab Corp.); normalized FAM emission values are means of 3 measurements, according to the following methodology: A/ Fluorescence analysis 1: each analysis originated in 4 different experiments, performed as triplicates: 1- 3 control wells in which the F-Myc was alone, whose FAM emission was normalized to 100; 2- 3 wells in which F-Myc was mixed with beads, in order to quantify the non-specific F-Myc/bead binding (compared to the previous control experiment); 3- 3 control wells in which the F-Myc was mixed with BioTASQ, whose FAM emission was normalized to 100, in order to quantify the influence of the presence of the ligand on the fluorescence of the F-Myc; and 4- 3 wells in which F-Myc and BioTASQ were mixed with beads, in order to quantify the specific F-Myc/BioTASQ/bead binding (compared to the previous control experiment). B/ Fluorescence analysis 2: each analysis originated in 2 different experiments, performed as triplicates: 1- 3 control wells comprising solutions that resulted from experiments performed with F-Myc and beads, in order to quantify the non-specific F-Myc/bead binding, the FAM emission of the solution was

normalized to 1; and 2- 3 wells comprising solutions that resulted from experiments performed with F-Myc, BioTASQ and beads, in order to quantify the actual BioTASQ capture capability when compared to the control experiments.

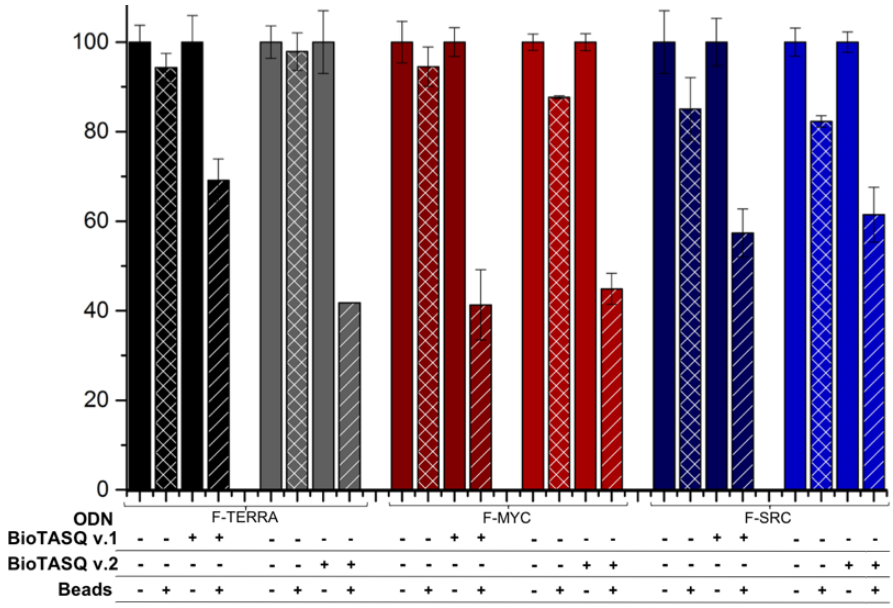

Figure S10: Control experiments performed with both BioTASQs and F-TERRA and F-SRC.

B. One-step pull-down protocol. The streptavidin MagneSphere® beads (Promega) were washed 3 times with Caco.K buffer as indicated above. BioTASQ (20 μM) was mixed with various 5'-labeled oligonucleotides (F-ON, 1 μM, here: F-Myc, F-SRC, F-22AG, F-duplex, F-TERRA, F-TRF2 and F-NRAS) and MagneSphere® beads (25 μg) in Caco.K buffer (250 μL final volume) and stirred for 16 h at 25 °C. The beads were immobilized (attracted by a magnet) and the supernatant was removed. The solid residue was resuspended in 240 μL of TBS 1X buffer, heated for 8 min at 90 °C (under gentle stirring 800 r.p.m.) and then centrifuged for 2 min (8,900 r.p.m.). The supernatant was taken up for analysis (the beads being immobilized by a magnet), after being distributed in 3 wells (80 μL each) of a 96-well plate, using a ClarioStar® machine (BMG Labtech) equipped with FAM filters (λ_{ex} = 492 nm; λ_{em} = 516 nm). Data were analyzed as above, with Excel and OriginPro®9.1. FAM emission is normalized as follows: the FAM emission of the 3 control wells (without BioTASQ) is collected and normalized to 1; then, the FAM emission of the 3 wells comprising F-ON, BioTASQ and beads is collected and compared to the control experiments. Competitive experiments: experiments were performed with BioTASQ (20 μM), F-ON (1 μM) and MagneSphere® beads (25 μg) in presence of an excess of nucleic acid competitor, either ds17 (20 μM), ds26 (20 μM) or calf thymus DNA (ctDNA, 100 μM). The protocol and data treatment was implemented as described above.

VII. G4RP-qPCR protocols

A. Dose-response Profiling of BRACO-19. MCF7 cells were seeded at 3000/well in a 96-well flat bottom plate. The cells were treated with a series of BRACO-19 concentrations, made from serial dilutions. The cells were then monitored in the EssenBioscience IncuCyte ZOOM live-cell monitoring system as described previously.[4] Three biological replicates were used to produce dose response curve. Doses between LD_{15} to LD_{25} were calculated from dose response curve profiles and used for subsequent G4RP-qPCR experiments.

B. G4 RNA-precipitation (G4RP) protocol. MCF7 cells were seeded at 3.5x10[5] cells per 10-cm dish before treatment with either vehicle (PBS) or BRACO 5 μg/mL (LD_{15}) for 72h. Cells were then crosslinked using 1% formaldehyde/PBS for 5 min at 25 °C and the crosslinking was then quenched with 0.125 M glycine for 5 min. Cells were scraped and resuspended in G4RP buffer (150 mM KCl, 25 mM Tris pH 7.4, 5 mM EDTA, 0.5 mM DTT, 0.5% NP40, RNase inhibitor (Roche), homebrew protease inhibitor cocktail). Cells were then sonicated using Covaris m220 Ultrasonicator using default settings at 10% duty for 2 min. The sonicated fractions were then incubated with 100 μM BioTASQ v.1 or BioTASQ v.2 (or 100 μM biotin for controls) overnight at 4 °C. 5% of the sonicate was collected as input control. 10 μg of streptavidin-magnetic beads (Promega) was added and the extract was incubated for 2h at 4 °C. Magnetic beads were then washed 4 times in G4RP buffer for 5 min. The beads were then incubated at 70 °C for 1 h to reverse crosslink. TRIZOL was then used to extract the RNA from the beads using manufacturer's instructions.

Reverse-transcription quantitative PCR (RT-qPCR). The primer sets used for RT-qPCR are:

mRNA	Forward	Reverse	Source
VEGFA	CCTTGCCTTGCTGCTCTACC	AGATGTCCACCAGGGTCTCG	Ref. *a*
NRAS	ATGACTGAGTACAAACTGGTGGT	CATGTATTGGTCTCTCATGGCAC	Ref. *a*
HPRT1	ACCAGTCAACAGGGGACATAA	CTTCGTGGGGTCCTTTTCACC	Ref. *a*
MALAT1	AAAGCAAGGTCTCCCCACAAG	GGTCTGTGCTAGATCAAAAGGCA	Ref. *b*
XIST	GGTCTGTGCTAGATCAAAAGGCA	AGCTCCTCGGACAGCTGTAA	Ref. *c*
RPPH1	GAGCTGAGTGCGTCCTGTC	TCAGGGAGAGCCCTGTTAGG	Ref. *d*

References:

a. Wang, X. et al. *Nucleic Acids Res.* **2011**, *40*, D1144-D1149; *b*. Ma, K. X. et al. *Tumor Biol*. **2015**, *36*, 3355-3359; *c*. Gilbert, S. L. et al. *J. Biol. Chem*. **2000**, *275*, 36491-36494; *d*. Cai, Y. et al. *Front. Mol. Neurosci*. **2017**, *10*, 27.

Extracted RNA was reverse transcribed with Superscript III (Thermos) and random hexamer primers using manufacturer's standard protocol to generate cDNA. cDNAs were quantified using 2X SYBR green mix (Bimake) with three technical replicates. C(t) values of pulldown samples were normalized to the input control. Three biological replicates were used for all qPCR-based quantifications. Exon-spanning primers for quantifying mRNA levels were derived from Primerbank (Ref. *a* above)).