Electronic Supplementary Information

to the manuscript:

Identification of novel interactors of human telomeric G-quadruplex DNA

Bruno Pagano,*[a] Luigi Margarucci,[b] Pasquale Zizza,[c] Jussara Amato,[a] Nunzia Iaccarino,[a] Chiara Cassiano,[b] Erica Salvati,[c] Ettore Novellino,[a] Annamaria Biroccio,[c] Agostino Casapullo[b] and Antonio Randazzo[a]

[a] *Department of Pharmacy, University of Naples "Federico II", Via D. Montesano, 49, 80131 Napoli, Italy. E-mail: bruno.pagano@unina.it*

[b] *Department of Pharmacy, University of Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano (SA), Italy.*

[c] *Experimental Chemotherapy Laboratory, Regina Elena National Cancer Institute, 00158 Rome, Italy.*

Contents **Page**

Experimental Section

Preparation of oligonucleotide samples. All synthetic oligonucleotides have been purchased by Biomers (Germany), purified employing standard HPLC protocols and checked for their integrity by MALDI mass spectrometry. In particular, the following DNA sequences have been used for the experiments:

$^{5'}$TTAGGGTTAGGGTTAGGGTTAGGGTT$^{3'}$ (tel_{26});

biotin-TEG-$^{5'}$TTAGGGTTAGGGTTAGGGTTAGGGTT$^{3'}$ (biotin-tel_{26});

biotin-TEG-$^{5'}$GAAGTGTGTGTGTGTGTGTGTGTGAA$^{3'}$ (biotin-scr_{26}).

G-quadruplex (G4) formation was carried out by dissolving the oligonucleotides in a buffer containing 10 mM KH_2PO_4, 70 mM KCl, and 0.2 mM EDTA, pH 7.0, at 10 mM single strand concentration. Formation of secondary structures was achieved by heating the DNA solution at 90 °C for 5 min, and slowly cooling it to room temperature. The solutions were then equilibrated at 4 °C overnight. The concentration of all oligonucleotides was measured at 260 nm by UV measurements at 90 °C using the molar extinction coefficients calculated by the nearest neighbor model.[1] After dilution necessary for the experiments, the concentration of a sample was further refined. In addition, to verify that the dilution did not alter the G-quadruplex species in solution, CD spectral changes with time were checked, without any appreciable change observed over the period of time required to complete the experiments.

Circular Dichroism (CD) experiments. CD spectra and CD melting curves of the oligonucleotides were recorded on a Jasco J-715 spectropolarimeter equipped with a Jasco PTC-423S Peltier temperature controller. CD spectra were recorded using 1 mm path-length cuvettes, in the wavelength range 200–320 nm at 20 °C, with a scan rate of 100 nm/min, a response time of 1 s and a 0.5 nm bandwidth. All the spectra were averaged over 3 scans. The DNA concentration was 10 μM as single strand. CD melting experiments were performed in the temperature range 20–90 °C, at the heating rate of 0.5 °C/min by following changes of the CD signal at the wavelengths of

maximum variation upon oligonucleotide folding. The melting temperatures were determined from fit of melting curves using two state transition model implemented in Origin 8.0 program. Each melting experiment was performed at least two times. CD spectra of G-quadruplex/protein complexes were recorded using 1 cm path-length cuvettes, in the wavelength range 190–320 nm at 20 °C, with a scan rate of 20 nm/min, a response time of 2 s and a 0.5 nm bandwidth. All the spectra were averaged over 10 scans. Buffer baseline was subtracted from each spectrum.

Nuclear protein extraction. Nuclear extracts were obtained from HCT116 human colorectal carcinoma cell line purchased by ATCC and maintenance in Dulbecco Modified Eagle Medium (D-MEM, Invitrogen Carlsbad, CA, USA) supplemented with 10% fetal calf serum, 2 mM L-glutamin and antibiotics. Briefly, cells from subconfluent cultures were washed and scraped into phosphate-buffered saline (PBS) and pelleted by centrifugation at 300x*g* for 10 min. Then, cells were lysed in NE-PER extraction reagent (Pierce, IL, USA) according to the manufacturer's protocol and protein concentration was determined by BCA protein assay (Pierce).

G-quadruplex partners identification. Biotin-tel_{26} (50 nmol) and the same amount of biotin and biotin-$scr_{26,}$ as controls, were separately incubated for 1 h at 4 °C with 1 mg of HCT116 nuclear extract and maintained under continuous shaking. Following the incubation, 100 µl of streptavidin beads (Streptavidin agarose resin, Pierce, Thermo) were added to the lysate under continuous shaking (1 h, 4 °C). The beads were collected by centrifugation (900 x*g*, 1 min, 4 °C) and the supernatant, containing the unbound proteins, was removed. Since buffers with high salt concentration and high ionic strength are commonly used to remove the proteins slightly bounds to the beads reducing the unspecific binding, the beads we washed 3 times with 1 mL of PBS (50 mM Sodium phosphate, 150 mM pH 7.4 with 2.7 mM KCl and 137 mM NaCl). After washing, the bound proteins were eluted by boiling the beads in SDS-PAGE loading buffer (60 mM Tris/HCl pH 6.8, 2% SDS, 0.001% bromophenol blue, 10% glycerol, 2% 2-mercaptoethanol). The eluted

proteins were separated on 12% SDS-PAGE (Fig. S1), and stained with Coomassie G-250 (Bio-Rad, Hercules, CA). SDS-PAGE gel lanes were cut and digested. Briefly, gel pieces were washed with MilliQ water and AcN, reduced with DTT (10 mM) and alkylated with iodoacetamide reagent (54 mM). After a few washings, pieces were incubated with 50 ml of trypsin (10 ng/ml) on ice for 1 h. After the addition of 30 μl ammonium bicarbonate (10 mM, pH 7.5) samples were digested overnight at 37 °C. The supernatant was collected and the peptides were extracted from the gel pieces by incubation with 100 μl of pure AcN. Finally, the supernatant was collected and both were combined. Each peptide mixture was dried out and dissolved in 10 ml of 10% formic acid.[2] The experiment has been repeated twice using an opportune control matrix bearing the linker without any were combined. The peptide sample was dried and dissolved in formic acid (FA, 10%) before MS analysis. The peptide mixture (5 μL) was injected into a nano-ACQUITY UPLC system (Waters). Peptides were separated on a 1.7 mm BEH C18 column (Waters) at a flow rate of 1000 nL/min. Peptide elution was achieved with a linear gradient (solution A: 95 % H_2O, 5 % CH_3CN, 0.1 % FA; solution B: 95 % CH_3CN, 5 % H_2O, 0.1 % FA); 15–50 % B over 50 min). MS and MS/MS data were acquired on a LTQ-Orbitrap XL high-resolution mass spectrometry system (ThermoScientific). The ten most intense doubly and triply charged peptide ions were chosen and fragmented. The resulting MS data were processed to generate peak lists for protein identifications. Database searches were carried out on the Mascot server. The SwissProt database (release 2014_07 of 09-Jul-14 of UniProtKB/Swiss-Prot contains 546000 sequence entries, comprising 194259968 amino acids abstracted from 229917 references) was employed (settings: two missed cleavages; carbamidomethyl (C) as fixed modification and oxidation (M) and phosphorylation (ST) as variable modifications; peptide tolerance 20 ppm; MS/MS tolerance 0.8 Da).

Western blotting analysis on protein eluates. To further confirm the chemoproteomic protein identification, a Western blotting experiment was performed on high mobility group protein B1 (HMGB1) and far upstream element-binding protein 2 (KHSRP, FUBP2), using the eluates from

both the control and modified beads. 5 μL of each sample were resolved on a 10% SDS-PAGE and then transferred onto a nitrocellulose membrane. After the treatment with a blocking solution (25 mM Tris pH 8, 125 mM NaCl, 0.1% Tween-20, 5% nonfat dried milk) for 1 h, membranes were separately incubated with primary monoclonal antibodies raised against HMGB1 (OriGene, 1:1000 dilution) and KHSRP (OriGene, 1:1000 dilution). After incubation with anti-rabbit peroxidase-conjugated secondary antibodies, ECL Western Blotting substrates (Thermo, Pierce) and LAS 4000 (GE Healthcare, Waukesha, WI, USA) digital imaging system was used for detection.

Binding affinity evaluation by surface plasmon resonance. The recombinant KHSRP, HMGB1 and LMNB1 proteins were purchased from OriGene (MD, USA) and separately immobilized onto a CM-5 sensor chip using standard amine coupling procedure and HBS-P as running buffer (GE Healtcare). The activation of sensor chip surface was done injecting a 1:1 mixture of 100 mM EDC and 100 mM NHS at flow rate of 10 μL min^{-1}. 100 μL of each protein were injected onto the activated chip surface at a flow rate of 5 μL min^{-1} at a final concentration of 30 ng μL^{-1} in 10 mM potassium acetate at pH 4.5. The protein concentration was selected to obtain an optimal response (around 10000 ΔRU). The residual active groups were blocked by injection of 100 μL ethanolamine–HCl 1.0 M, pH 8.5, at 10 μL min^{-1} over the chip surface. Tel_{26} solutions (0.1–25 μM) were prepared in 10 mM KH_2PO_4, 70 mM KCl and 0.2 mM EDTA (pH = 7.0) and injected over the protein-functionalized chip surface at a flow rate of 10 μL min^{-1} using the same solution as running buffer. Each injection was at least duplicated, and since the dissociation curves reached the baseline within a reasonable time frame, no regeneration was required. The dissociation time was set at 400 s. The rate constants for association (ka), dissociation (kd) and the dissociation constant (KD) were obtained by globally fitting the data from all of the injections of different concentrations of each compound, using the BIAevaluation software, using the simple 1:1 Langmuir binding model.

Immunofluorescence. Immunofluorescence (IF) was performed in BJ fibroblasts expressing hTERT plus SV40 early region (BJ-EHLT) as previously described.[3] Briefly, cells were fixed in 2% formaldehyde and permeabilized in PBS plus 0.25% Triton X-100 for 5 min at room temperature. For immunolabeling, cells were incubated with primary antibodies for 2 h at room temperature, washed twice in PBS and finally incubated with the secondary antibodies for 1 hr. The following antibodies were used: mouse monoclonal anti-TRF1 antibody (Abcam Ltd.; Cambridge UK); rabbit polyclonal anti-TRF1 antibody and goat polyclonal anti-LMNB1 antibody (M20) (San Diego, CA, USA); rabbit polyclonal anti-HMGB1 antibody and rabbit polyclonal anti-KHSRP antibody (OriGene Technologies, MD, USA); TRITC-conjugated Chicken anti Goat (Agrisera AB, Vännäs, Sweden); TRITC-conjugated Goat anti-Rabbit, FITC-conjugated Goat anti-Mouse, FITC-conjugated Goat anti-Rabbit (Jackson Immunoresearch, Suffolk, UK). Nuclei were immunostained with DAPI. Fluorescence signals were recorded by using a Leica DMIRE2 microscope equipped with a Leica DFC 350FX camera and elaborated by Leica FW4000 deconvolution software (Leica, Solms, Germany).

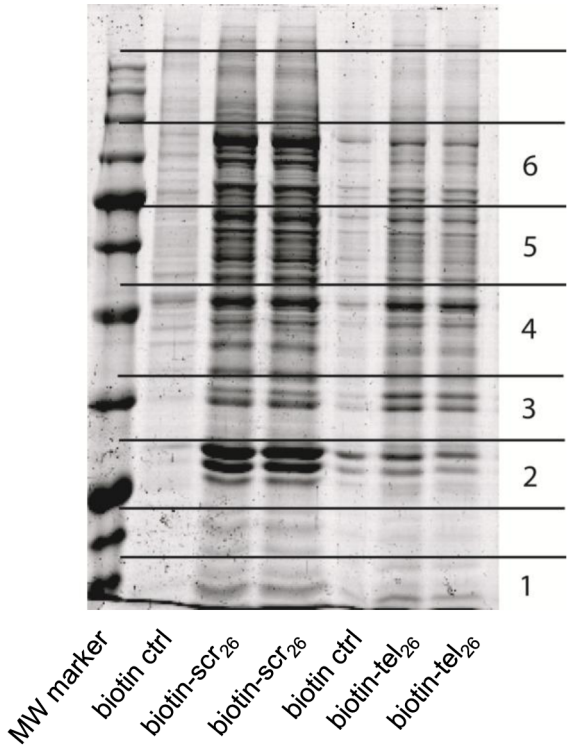

Fig. S1. SDS-PAGE of the eluted proteins from biotin-tel_{26}, biotin-scr_{26} and biotin control (ctrl). Gel regions of interest are delimited by lines and numbered.

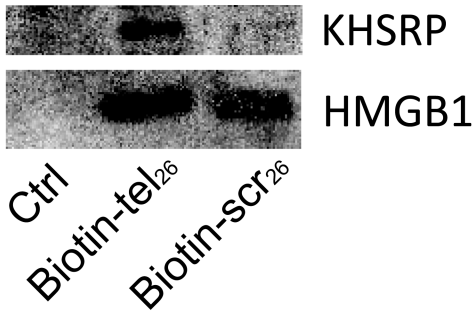

Fig. S2. Western blot analysis of biotin-tel_{26}, biotin-scr_{26} and biotin control (Ctrl) showing tel_{26} ability to pull down KHSRP and HMGB1 in the nuclear lysate. A positive response for biotin-scr_{26} towards HMGB1 is consistent with a relevant enrichment of the isoform HMGB2 in the biotin-scr_{26} fishing (see the Table in Supporting Information for the complete list of proteins identified).

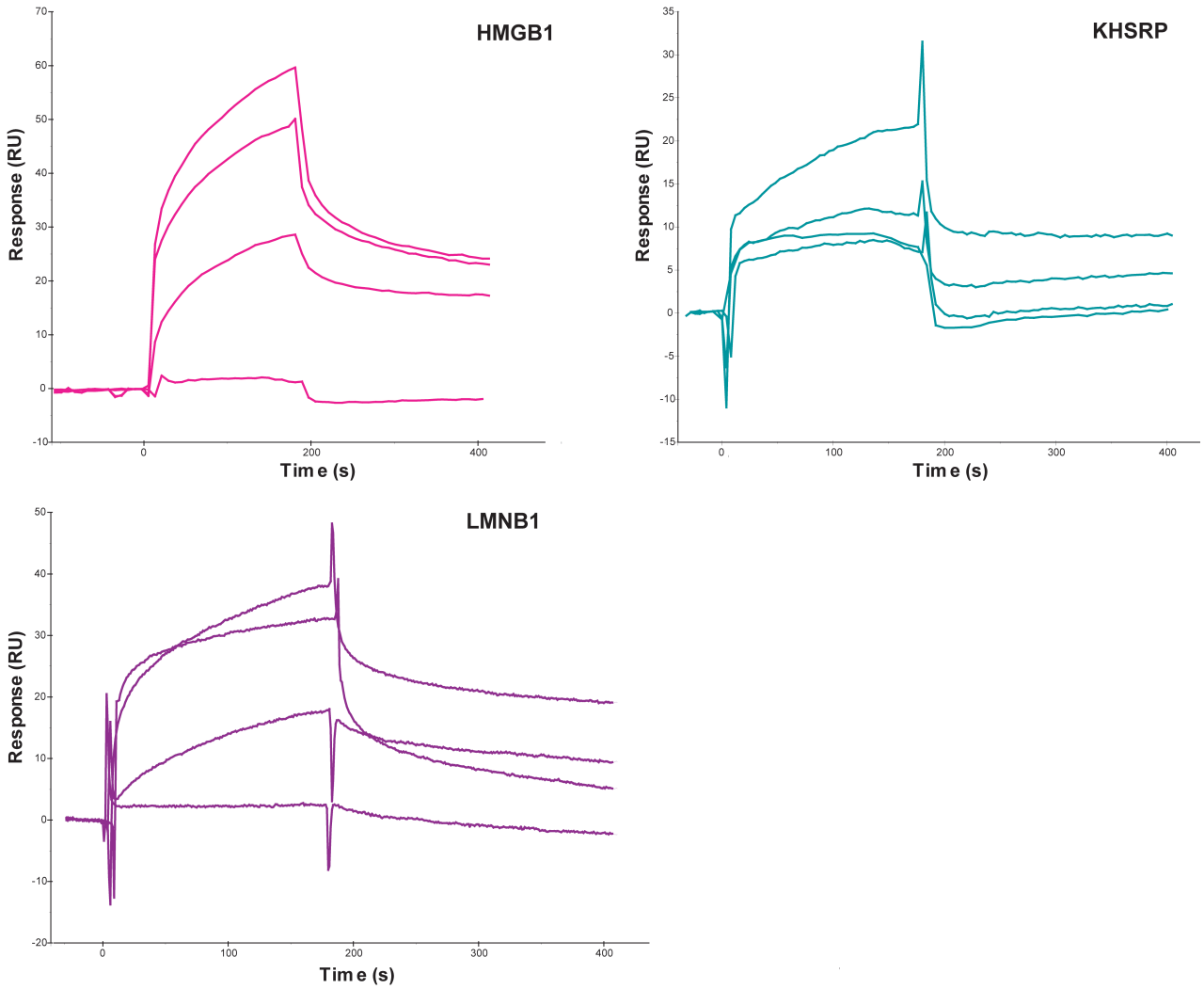

Fig. S3. SPR sensorgrams obtained by injections of tel_{26} (0.1, 1, 10 and 25 μM from bottom to up) on the chip-immobilized HMGB1, KHSRP and LMNB1.

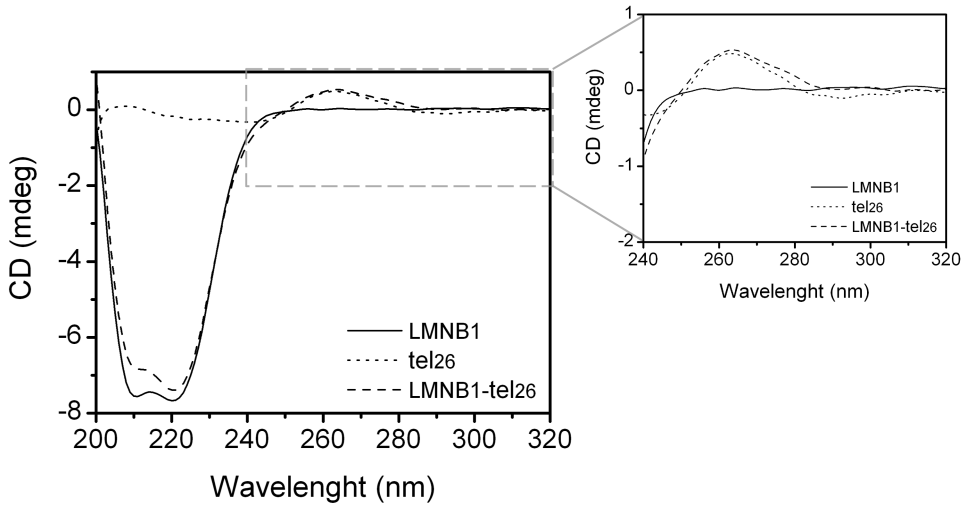

Fig. S4. CD spectrum of LMNB1-tel_{26} complex recorded at 1:1 molar ratio (50 nM) at 20 °C. CD spectra of the protein and tel_{26} alone (50 nM) are also shown.

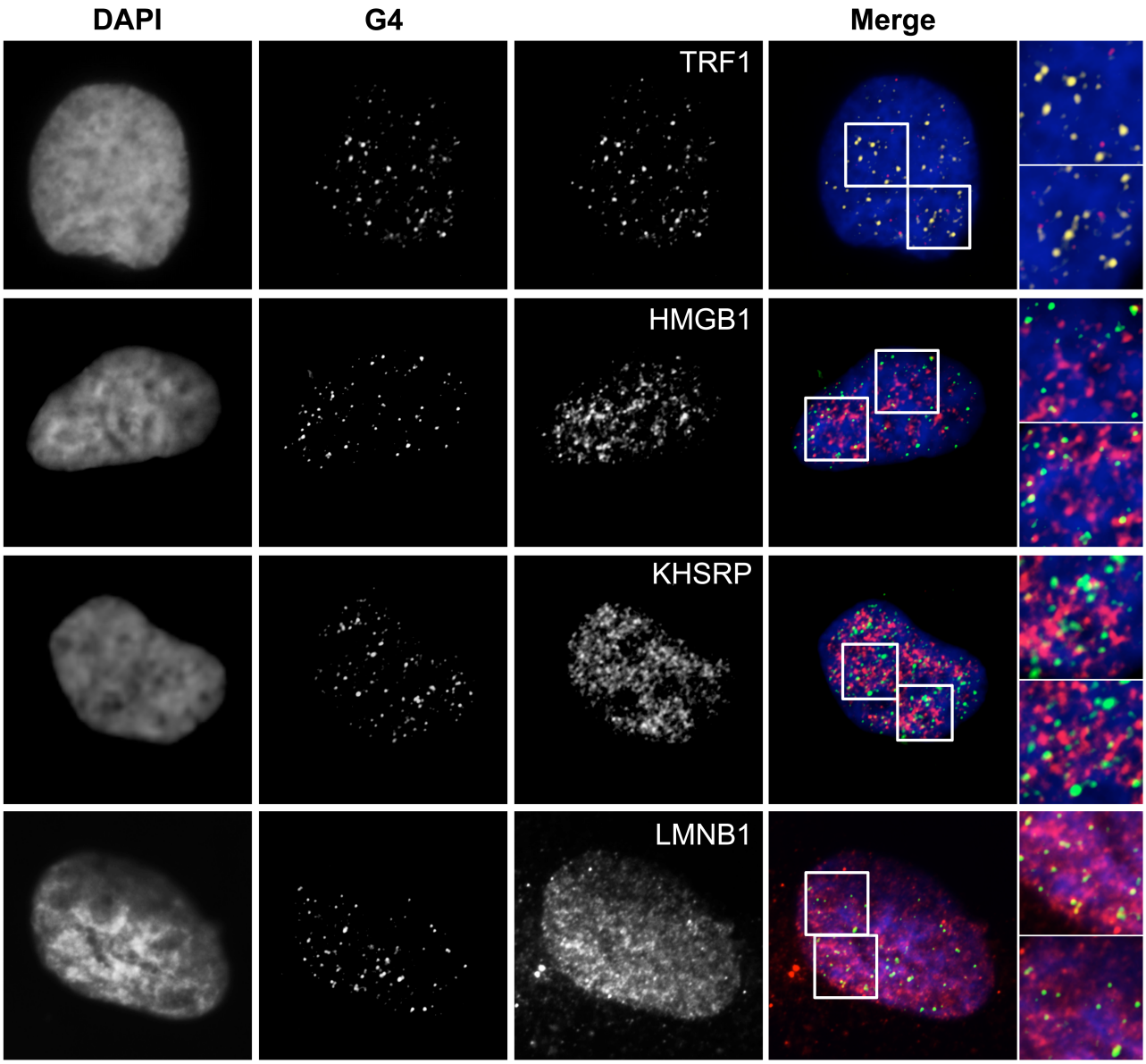

Fig. S5. HMGB1, KHSRP and LMNB1 colocalize with G-quadruplex structures *in vivo*. Actively growing BJ-EHLT fibroblasts were fixed, permeabilized and immunostained using a conventional immunofluorescence protocol (see Experimental Section). Representative IF images acquired by using a Leica Deconvolution microscope (magnification 100x) are shown. Foci deriving from the colocalization of TRF1, HMGB1, KHSRP and LMNB1 (red) with a monoclonal anti-G4 antibody (green, kindly provided by Prof. P. M. Landscorp) are clearly visible into the nuclei (blue). Enlarged views of colocalization foci are reported on the right of each merged picture.

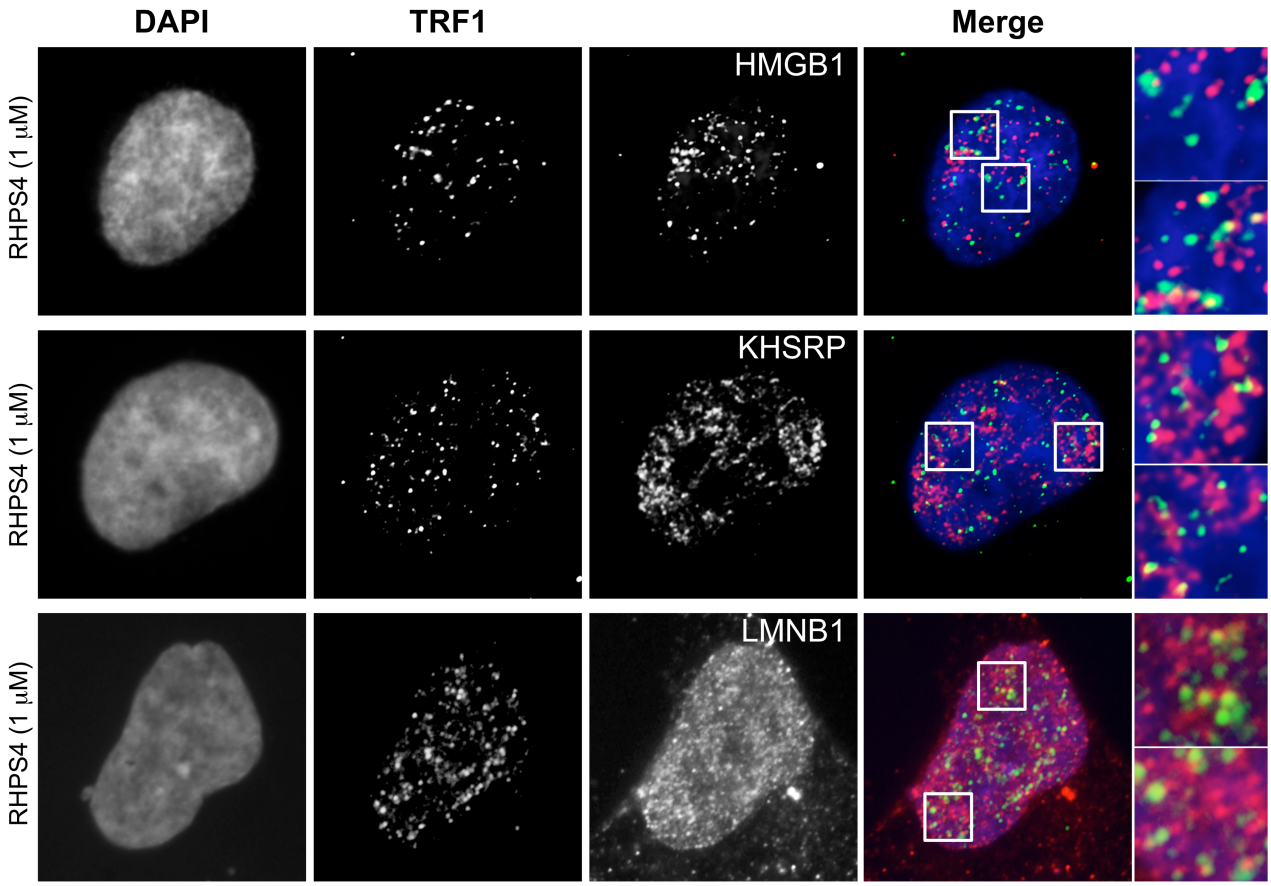

Fig. S6. BJ-EHLT fibroblasts were grown for 24 h in presence of 1 μM of RHPS4,[3] and the telomeric localization of HMGB1, KHSRP and LMNB1 was analyzed. Representative IF images acquired by using a Leica Deconvolution microscope (magnification 100x) are shown. Discrete foci deriving from the colocalization of HMGB1, KHSRP or LMNB1 (reds) with TRF1 (green) are visible into the nucleus (blue). Enlarged views of colocalization spots are reported on the right of each merged picture.

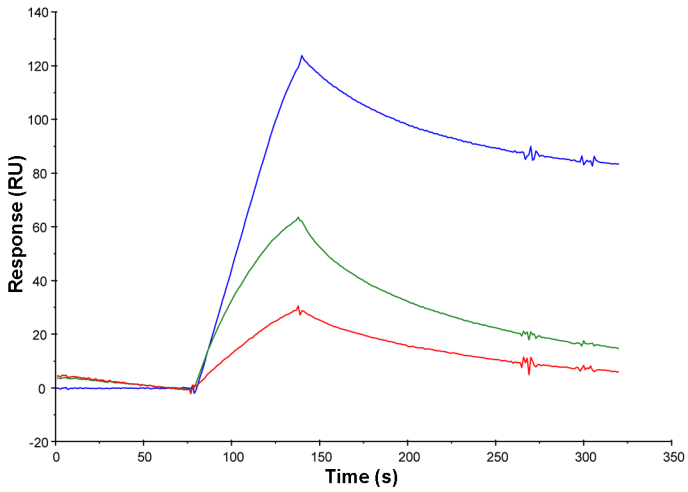

Fig. S7. SPR sensorgrams obtained by injections of 50 μM G-quadruplex-forming oligonucleotides tel_{26} (blue), c-myc ($^{5'}$TGAGGGTGGGGAGGGTGGGGAAGG$^{3'}$, green),[4] and c-kit1 ($^{5'}$AGGGAGGGCGCTGGGAGGAGGG$^{3'}$, red)[5] on the chip-immobilized HMGB1 protein. Tel_{26} displays better binding profile compared to the other two samples for which a weak interaction is observed.

References

1 C. R. Cantor, M. M. Warshaw and H. Shapiro, *Biopolymers,* 1970, **9**, 1059.

2 A. Shevchenko, H. Tomas, J. Havlis, J. V. Olsen and M. Mann, *Nat. Protoc.,* 2006**, 1**, 2856.

3 E. Salvati, C. Leonetti, A. Rizzo, M. Scarsella, M. Mottolese, R. Galati, I. Sperduti, M. F. G. Stevens, M. D'Incalci, M. Blasco, G. Chiorino, S. Bauwens, B. Horard, E. Gilson, A. Stoppacciaro, G. Zupi and A. Biroccio, *J. Clin. Invest.,* 2007, **117**, 3236.

4 A. T. Phan, V. Kuryavyi, H. Y. Gaw, D. J. Patel, *Nat. Chem. Biol.*, 2005, **1**, 167.

5 A. T. Phan, V. Kuryavyi, S. Burge, S. Neidle, D. J. Patel, *J. Am. Chem. Soc.*, 2007, **129**, 4386.