Supporting Information

Multiple *in vitro* inhibition of HIV-1 proteins by 2,6-dipeptidyl-anthraquinone conjugates targeting the PBS RNA

Elia Gamba[a], Alice Sosic[a]*, Irene Saccone[b], Elisa Magli[b], Francesco Frecentese[b], Barbara Gatto[a]

[a]Dipartimento di Scienze del Farmaco, Università di Padova, via Marzolo 5, 35131 Padova, Italy

[b]Dipartimento di Farmacia, Università degli Studi di Napoli "Federico II", Via D. Montesano 49, 80131 Napoli, Italy

*alice.sosic@unipd.it

List of content:

EXPERIMENTAL DETAILS

Nucleic acid substrates and proteins. Synthetic RNA oligonucleotides were obtained in lyophilized form from commercial sources. The 18-nt RNA sequences were chemically synthesized by Metabion International AG (Martinsried, Germany). All oligonucleotides were used as received with no further purification. (+)PBS RNA and (-)PBS RNA sequences used were: 5'- UGG CGC CCG AAC AGG GAC -3' and 5'- GUC CCU GUU CGG GCG CCA -3', respectively. The samples were stored at -20°C in TE 1X buffer (Tris·HCl 10 mm, EDTA 1 mm, pH 8) prepared in DEPC-treated water. (+)PBS RNA and (-)PBS RNA were folded independently in Tris-HCl 10 mM (pH 7.5), NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM: RNA construct was heated to 95°C for 5 min and then ice-cooled in order to assume the proper hairpin structure. The dsRNA was also heated and slowly cool to room temperature in order to form the RNA-duplex. The full-length recombinant NC protein was obtained in house as previously described.[1] Tat(48-57) peptide was obtained from commercial sources. Sequence of the full-length NC protein: IQKGN FRNQR KTVKC FNCGK EGHIA KNCRA PRKKG CWKCG KEGHQ MKDCT ERQAN; sequence of the Tat(48-57) peptide: GRKKR RQRRR. Stock solutions were stored at -20°C and aliquots were freshly diluted to the desired concentration in Tris 10 mM (pH 7.5), NaCl 20 mM buffer.

Chemical reagents. 2,6-dipeptidyl-anthraquinone derivatives considered in this study were synthesized in house as previously described.[2-5] Aliquots were freshly prepared by diluting a 10 mm DMSO stock in MilliQ water. All the other chemical reagents, including salts and solvents, were purchased from Sigma–Aldrich (Milan, Italy).

Spectroscopic analysis. (+)PBS RNA construct was diluted in a 1 cm path quartz cuvette (Hellma) to a concentration of 1.6 μM in Tris-HCl 10 mM (pH 7.5), NaCl 20 mM, $Mg(ClO_4)_2$ 1mM buffer. UV-absorbance and circular dichroism spectra were acquired in the 240-350nm region at 20°C on a J-810 spectrophotometer (Jasco) connected to a Peltier temperature control system. Temperature was gradually increased from 20°C to 95°C at a rate of 50°C/h. Spectra were acquired at 20°C and 95°C.

RNA Annealer-Mediated Primer Annealing (RAMPA) assay. RNA Annealer-Mediated Primer Annealing (RAMPA) assay was developed to investigate the ability of NC and Tat proteins to enhance the annealing step involving the (+)PBS RNA and its complementary sequence (-)PBS RNA. Subsequently the RAMPA assay was employed to evaluate the ability of compounds to impair the biological activity of NC and Tat annealers, which involved monitoring the annealing of (+)PBS RNA with (-)PBS RNA. (+)PBS RNA, (-)PBS RNA and the dsRNA used as controls were folded as above described. Two different RAMPA assay formats were employed to evaluate the inhibition of NC annealer-mediated (+)/(-)PBS RNA formation. *RNA-preincubation mode*: folded (+)PBS RNA and (-)PBS RNA, each 1 μM, were incubated with increasing concentrations of compound (0, 3, 20, 50, 100, and 200 μM) for 15 min at room temperature. (+)PBS RNA and (-)PBS RNA samples were then mixed and incubated with NC 3 μM for 10 minutes at room temperature. Inhibition of the (+)/(-)PBS RNA NC- or Tat(48-57)-mediated annealing reaction by selected 2,6-dipeptidyl anthraquinone derivatives was then assessed by RAMPA assay following the RNA-preincubation mode. Inhibition effects were evaluated in the presence of increasing amounts of each compound (0, 3, 20, 50, 75, 100, 125, 150 and 200 μM). Finally, samples were added with gel loading buffer containing SDS (GLB_{SDS}: Tris-HCl 100 mM, EDTA 4 mM, 50 % w/v glycerol, 2% w/v SDS, 0.05% w/v bromophenol blue), kept on ice, and resolved on a 12% non-

denaturing polyacrylamide gel in TBE buffer (Tris-HCl 89 mM, boric acid 89 mM, and EDTA 2 mM, pH 8) for 2 h at 200 V. After electrophoresis, all gels were stained with SybrGreen II and analyzed on a Geliance 600 imaging system (PerkinElmer). Important to notice is that the bands corresponding to monomeric RNA sequences are not stained intensely enough to be unambiguously displayed. (+)PBS and (-)PBS RNA sequences always resulted barely detected in the gel system by SYBR Green II, probably due to the limited number of base-pairs characterizing their secondary structure. On the contrary, the RNA-duplex was clearly detected. Thus, reactions were evaluated by following the RNA-duplex species. Band quantification was always performed considering the band corresponding to the RNA-duplex. GeneTools software (PerkinElmer) was employed to quantify the percentage of RNA-duplex formation, which enabled the calculation of the sought-after IC_{50} value, namely the concentration of compound required to inhibit RNA-duplex formation by half.[3]

Direct binding of ligands to individual RNA construct. Electrospray-ionization mass spectrometry (ESI-MS) under non-denaturing conditions was employed to assess RNA/AQ interactions. Before mixing with compounds, RNA constructs were filtered by using 3K NMWL (Millipore Corporation, MA, USA) centrifugal filters to minimize the presence of magnesium salts that can adversely interfere with ESI performance. In a typical binding experiment, appropriate volume of folded RNA construct (final 2 μM) was mixed with each compound (10 μM final concentration) in 150 mM ammonium acetate (pH 7.5) and incubated for 30 min incubation at room temperature. Control experiments were performed on 2 μM solutions of (+)PBS and (-)PBS RNA in 150 mM ammonium acetate. Control experiments conducted in absence of compounds, provided an experimental mass of 5805.86 u and 5696.82 u for (+)PBS and (-)PBS RNA constructs, respectively, confirming the identity of the two RNAs having a calculated mass of 5805.85 u and 5696.76 u, based on the elemental composition of the two species (data not shown). All samples were analyzed in negative ion mode by direct infusion electrospray ionization (ESI) on a Waters Synapt G2 HRMS mass spectrometer equipped with a heated capillary source built in-house.[6] The analyses were performed in nanoflow ESI mode by using quartz emitters produced in-house by a Sutter Instruments Co. (Novato, CA) P2000 laser pipet puller. Up to 6 μL samples were loaded onto each emitter by using a gel loader pipet tip. A stainless steel wire was inserted in the back- end of the emitter and used to supply an ionizing voltage ranged around 0.8−1.0 kV, a desolvation voltage of 50 V, and a source temperature of 30°C. Data were analyzed by using MassLynx 4.1 software (Waters). Binding affinity was calculated from the full-scan ESI-MS spectra according to equation (1):

$$f_b = \frac{\sum\left(\frac{I_{RNA+L}}{z}\right)}{\sum\left(\frac{I_{RNA}+I_{RNA+L}}{z}\right)} \quad (1)$$

in which f_b represents the fraction of RNA bound by the ligand (L), z is the charge state of the specie, I_{RNA+L} and I_{RNA} are the areas under the curve of the signal corresponding to RNA-ligand complexes and free RNA, respectively. All the detected charged states were included in the calculation to minimize any quantification bias.[7] Tandem mass spectrometry, aimed at the investigation of the compound binding site on the oligonucleotide structure,[1] were conducted by using nitrogen as collision gas and activation of the precursor ion was achieved by increasing the voltage in the transfer ion trap of the Synapt G2 instrument.

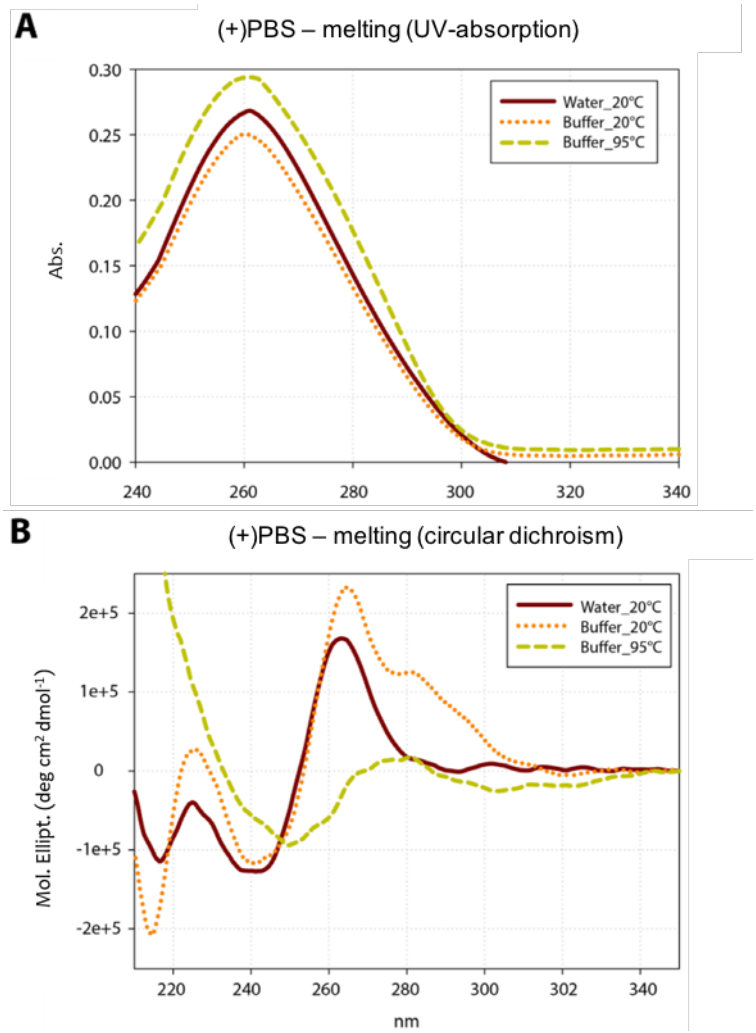

Figure S1. Melting/annealing analysis of (+)PBS RNA construct by UV-absorption (**A**) and circular dichroism (**B**).

(**A**) UV-absorption spectra of 1.6 µM (+)PBS RNA sequence in Tris-HCl 10 mM (pH 7.5), NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM buffer at 20°C (dotted orange line) or 95°C (dashed green line). The absorption peak at 260 nm at 95°C shows a reduction of 13.5 % when let to cool down to 20°C. When considering a 30 % loss in the absorbance upon complete base pairing (25), the hypochromic effect observed in (+)PBS RNA upon thermal melting indicates the formation of 4 base pairs.

(**B**) Circular dichroism (CD) spectra of 1.6 µM (+)PBS RNA sequence in Tris-HCl 10 mM (pH 7.5), NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM buffer at 20°C (dotted orange line) or 95°C (dashed green line). The far-UV CD spectrum recorded at 20°C shows the characteristic signals of a RNA helix (positive signal around 260 nm and a negative signal at 210 nm), suggesting the presence of a RNA stem in the (+)PBS RNA construct. The characteristic peaks at 260 nm and 210 nm disappear in the spectrum at 95°C due to unfolding of the hairpin structure.

Structural information obtained by UV and CD analyses confirm the predicted secondary structure of (+)PBS RNA construct consistent with the 4-base pair stem reported in the main text in **Figure 2**.

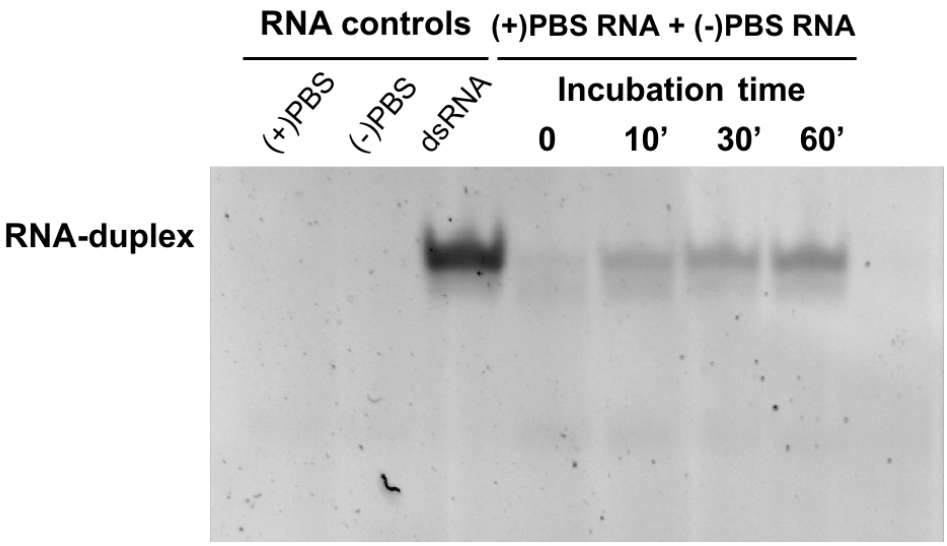

Figure S2. RNA annealer-mediated primer annealing (RAMPA) assay completed in the absence of the annealer proteins.

Folded (+)PBS and (-)PBS RNA, each 1 μM, were incubated for 0 min, 10 min, 30 min or 60 min at room temperature. Monomeric (+)PBS RNA, (-)PBS RNA and the RNA-duplex (dsRNA) annealed by thermal denaturation were used as controls. The samples were analyzed by electrophoresis on a 12% native polyacrylamide gel and stained with SybrGreen II (see ***Experimental Details***). In all the experiments, the bands corresponding to monomeric RNA sequences are not stained intensely enough to be unambiguously displayed. (+)PBS and (-)PBS RNA sequences always resulted barely detected in the gel system by SYBR Green II, probably due to the limited number of base-pairs characterizing their secondary structure. Reactions were therefore evaluated by following the RNA-duplex species. Band quantification was always performed considering the band corresponding to the RNA-duplex (see ***Experimental Details***).

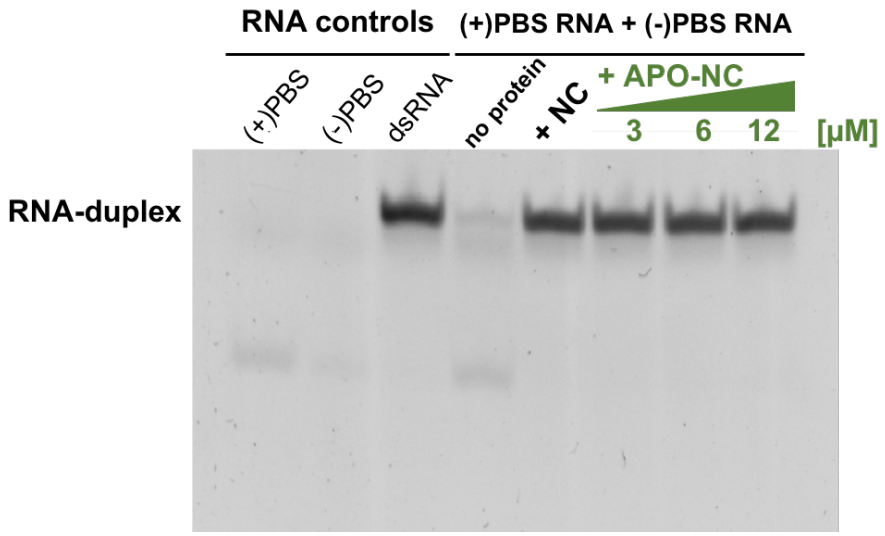

Figure S3. RNA annealer-mediated primer annealing (RAMPA) assay completed in the presence of the APO-NC protein.

APO-NC was obtained incubating full-length zinc-bound NC overnight in a mixture of EDTA 350 mM, methanol 17 % and acetic acid 0.7 %. The day after, the protein buffer was extensively changed against ammonium acetate using 3 kDa molecular weight cut-off centrifuge filters (Pearce Concentrators). The actual mass of the obtained protein was checked by mass spectroscopy (Mariner, PerSeptive Biosystems). No signals corresponding to the protein binding two zinc ions was detected after the treatment whereas little trace of protein bearing one zinc was observed, indicating that most of the protein was efficiently converted into APO-NC lacking both zinc ions.

Folded (+)PBS and (-)PBS RNA, each 1 μM, were incubated with increasing concentrations of APO-NC (3, 6 and 12 μM) for 10 min at room temperature. Monomeric (+)PBS RNA, (-)PBS RNA and the RNA-duplex (dsRNA) annealed by thermal denaturation were used as controls. (+)PBS and (-)PBS RNA were incubated also in the absence of the annealer proteins (no protein) and in presence of 3 μM NC for 10 min at room temperature as reference activity of the zinc-bound protein. The samples were analyzed by electrophoresis on a 12% native polyacrylamide gel and stained with SybrGreen II (see ***Experimental Details***).

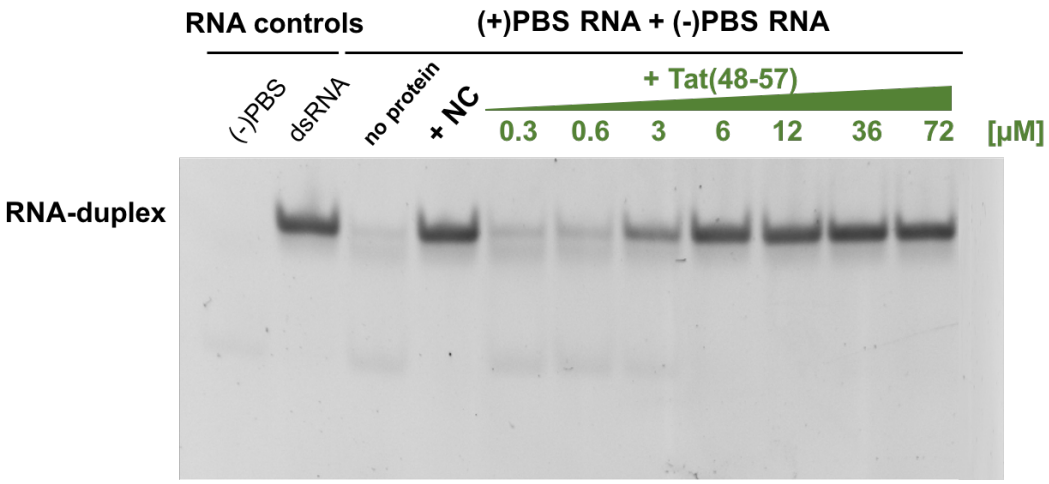

Figure S4. RNA annealer-mediated primer annealing (RAMPA) assay completed in the presence of the Tat(48-57) peptide.

Folded (+)PBS and (-)PBS RNA, each 1 μM, were incubated with increasing concentrations of Tat(48-57) peptide (0.3, 0.6, 3, 6, 12, 36 and 72 μM) for 10 min at room temperature. Monomeric (+)PBS RNA, (-)PBS RNA and the RNA-duplex (dsRNA) annealed by thermal denaturation were used as controls. (+)PBS and (-)PBS RNA were incubated also in the absence of the annealer proteins (no protein) and in presence of 3 μM NC for 10 min at room temperature as reference activity of NC protein. The samples were analyzed by electrophoresis on a 12% native polyacrylamide gel and stained with SybrGreen II (see ***Experimental Details***).

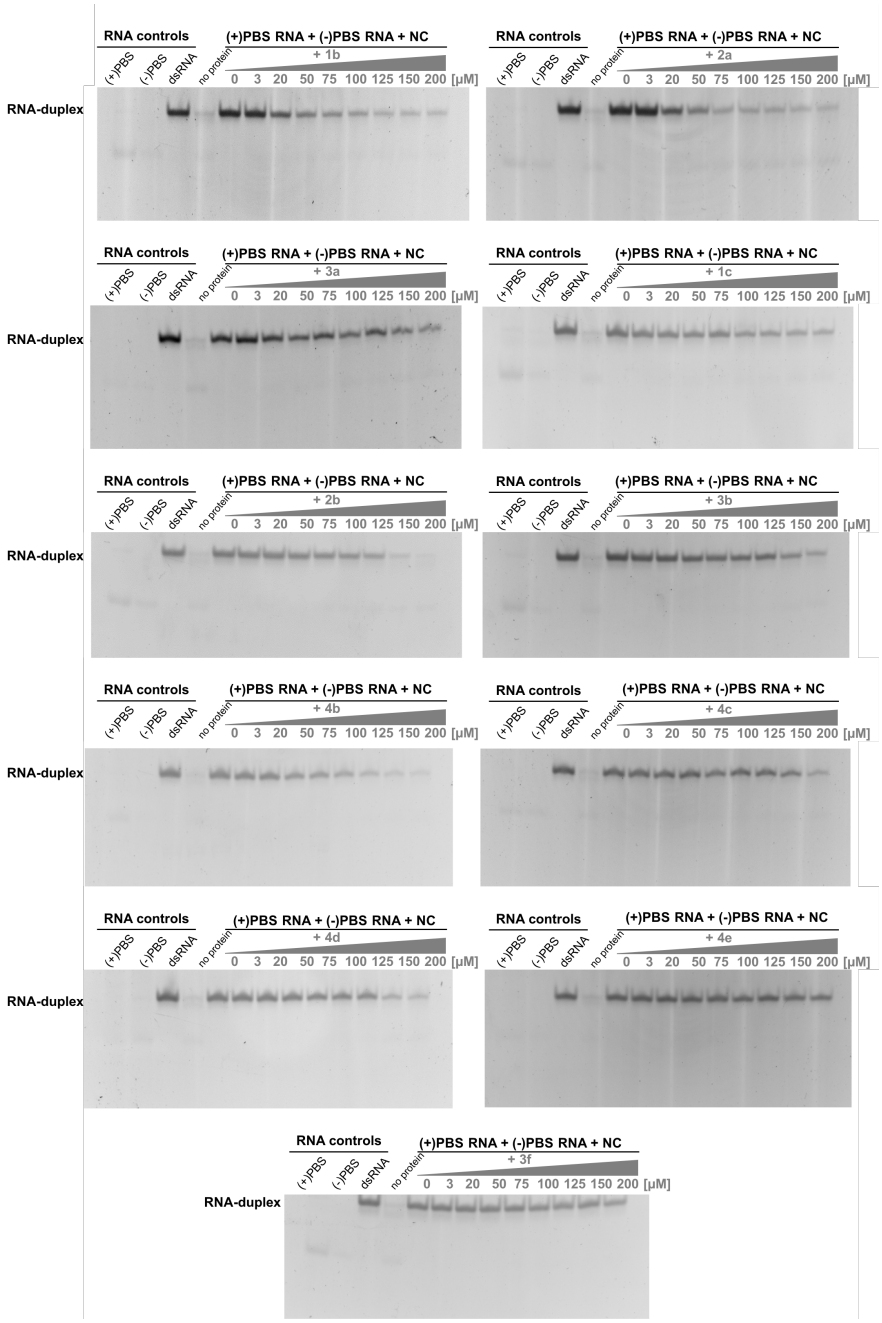

Figure S5. Inhibition of the (+)/(-)PBS RNA NC-mediated annealing reaction by selected 2,6-dipeptidyl anthraquinone derivatives.

RNA annealer-mediated primer annealing (RAMPA) assay was completed in the presence of the full-length NC protein following the RNA-preincubation mode. Inhibition effects were evaluated in the presence of increasing amounts of each compound. Folded (+)PBS and (-)PBS RNA, each 1 μM, were incubated with increasing concentrations of compound (0, 3, 20, 50, 75, 100, 125, 150 and 200 μM) for 15 min at room temperature. (+)PBS and (-)PBS RNA samples were then mixed and incubated with NC 3 μM for 10 minutes at room temperature. Monomeric (+)PBS RNA, (-) PBS RNA and the RNA-duplex (dsRNA) annealed by thermal denaturation were used as controls. (+)PBS and (-)PBS RNA were incubated also in the absence of the annealer protein (no protein). The samples were analyzed by electrophoresis on a 12% native polyacrylamide gel and stained with SybrGreen II (see ***Experimental Details***).

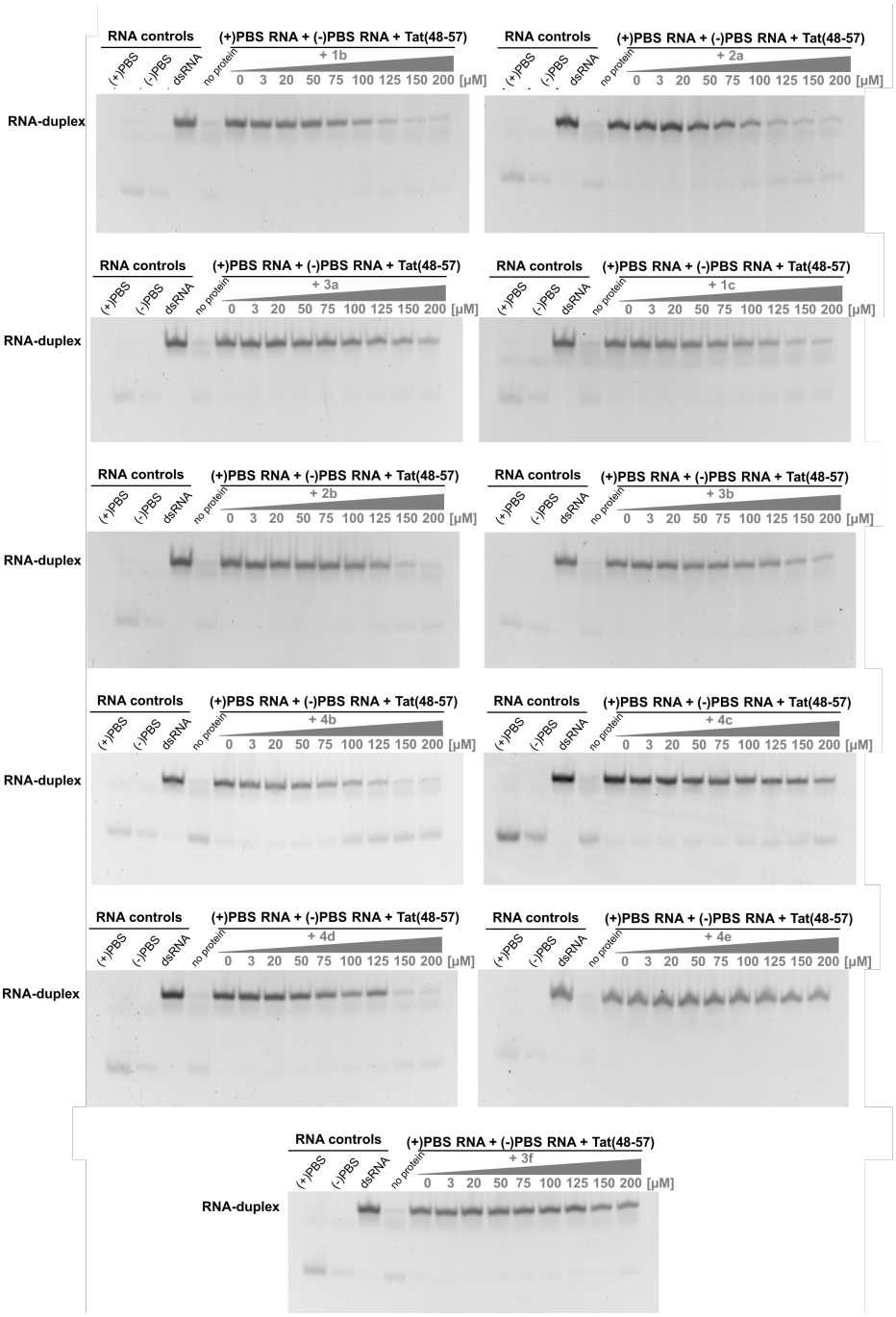

Figure S6. Inhibition of the (+)/(-)PBS RNA Tat-mediated annealing reaction by selected 2,6-dipeptidyl anthraquinone derivatives.

RNA annealer-mediated primer annealing (RAMPA) assay was completed in the presence of the Tat(48-57) peptide following the RNA-preincubation mode. Inhibition effects were evaluated in the presence of increasing amounts of each compound. Folded (+)PBS and (-)PBS RNA, each 1 μM, were incubated with increasing concentrations of compound (0, 3, 20, 50, 75, 100, 125, 150 and 200 μM) for 15 min at room temperature. (+)PBS and (-)PBS RNA samples were then mixed and incubated with Tat(48-57) 6 μM for 10 minutes at room temperature. Monomeric (+)PBS RNA, (-)PBS RNA and the RNA-duplex (dsRNA) annealed by thermal denaturation were used as controls. (+)PBS and (-)PBS RNA were incubated also in the absence of the annealer protein (no protein). The samples were analyzed by electrophoresis on a 12% native polyacrylamide gel and stained with SybrGreen II (see ***Experimental Details***).

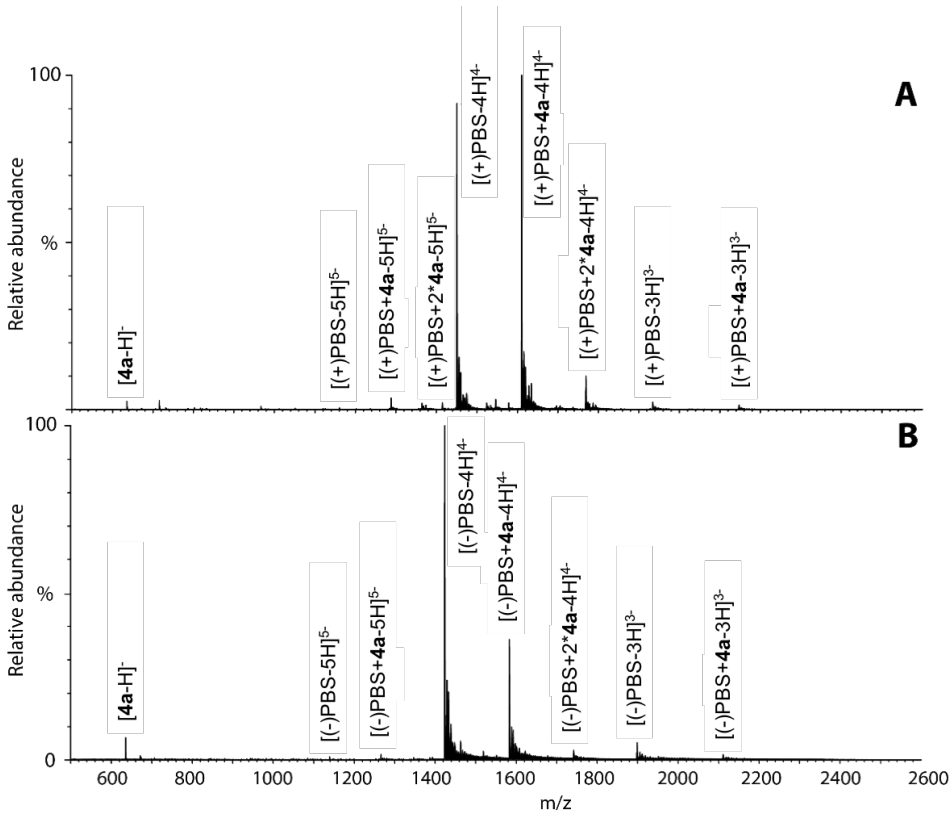

Figure S7. Representative ESI-MS spectra of samples containing a 1:5 molar ratio of either (+)PBS RNA (**A**) or (-)PBS RNA (**B**) and compound **4a**.

ESI-MS spectra were acquired after 30 minutes incubation in 150 mM ammonium acetate at room temperature. Spectra were acquired in negative ion mode in 150 mM ammonium acetate. Low-intensity signals observed near free/bound species consist of typical sodium, potassium and ammonium adducts. For both RNA substrates, the formation of complexes up to 2:1 compound:RNA stoichiometries was observed, confirming RNA as a putative target for tested AQs.

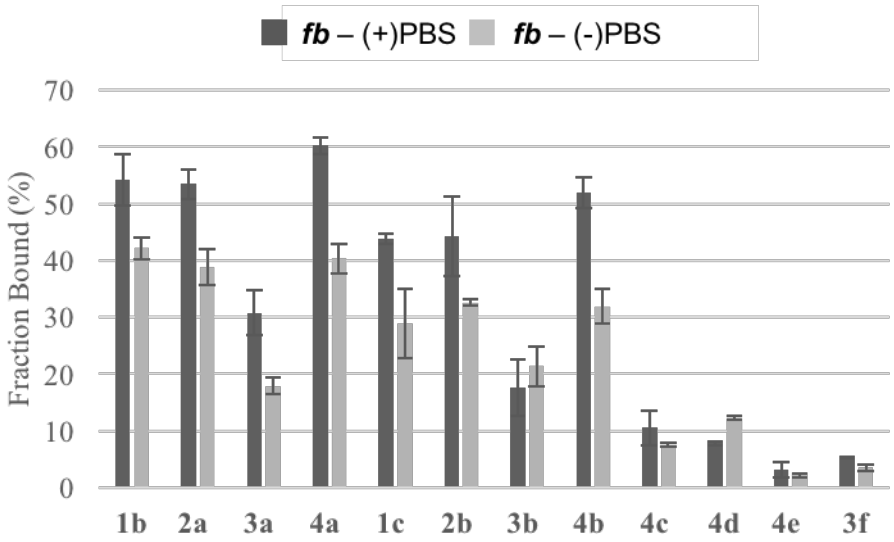

Figure S8. Relative affinity of 2,6-dipeptidyl-anthraquinone conjugates for (+)PBS and (-)PBS RNA substrates.

The histogram shows the percentages of bound substrate (fb, fraction bound) observed in the spectra of samples containing a 1:5 molar ratio of either (+)PBS and (-)PBS RNA and compound. The indicated percentages of bound substrate provide a measure of the relative affinities of each AQ for each RNA substrate.

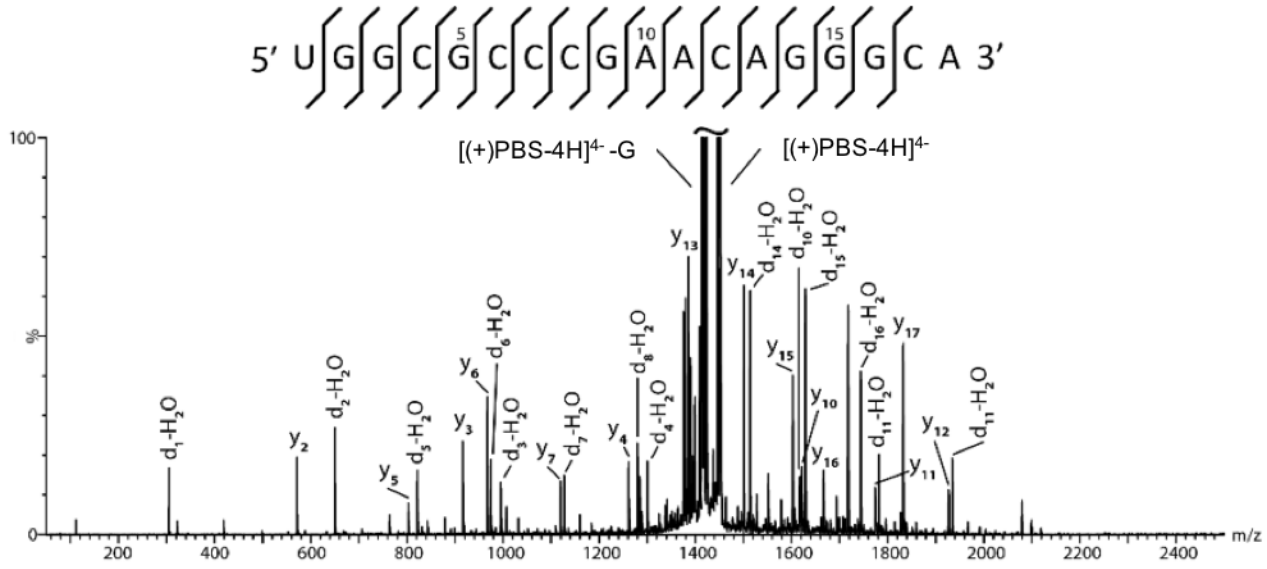

Figure S9. MS/MS spectrum of free (+)PBS RNA construct obtained in negative ion mode.

Characteristic ion series are labelled according to standard nomenclature. $[(+)PBS-4H]^{4-}$ indicates the precursor ion. Fragmentation pattern is summarized on the (+)PBS RNA sequence. Complete d-H_2O and y-series ions were observed.

References:

1. Turner, K. B.; Hagan, N. A.; Kohlway, A. S.; Fabris, D., Mapping noncovalent ligand binding to stemloop domains of the HIV-1 packaging signal by tandem mass spectrometry. *J Am Soc Mass Spectrom* **2006,** *17* (10), 1401-11.
2. Frecentese, F.; Sosic, A.; Saccone, I.; Gamba, E.; Link, K.; Miola, A.; Cappellini, M.; Cattelan, M. G.; Severino, B.; Fiorino, F.; Magli, E.; Corvino, A.; Perissutti, E.; Fabris, D.; Gatto, B.; Caliendo, G.; Santagada, V., Synthesis and in Vitro Screening of New Series of 2,6-Dipeptidyl-anthraquinones: Influence of Side Chain Length on HIV-1 Nucleocapsid Inhibitors. *J Med Chem* **2016,** *59* (5), 1914-24.
3. Sosic, A.; Frecentese, F.; Perissutti, E.; Sinigaglia, L.; Santagada, V.; Caliendo, G.; Magli, E.; Ciano, A.; Zagotto, G.; Parolin, C.; Gatto, B., Design, synthesis and biological evaluation of TAR and cTAR binders as HIV-1 nucleocapsid inhibitors. *MedChemComm* **2013,** *4* (10), 1388-1393.
4. Sosic, A.; Saccone, I.; Carraro, C.; Kenderdine, T.; Gamba, E.; Caliendo, G.; Corvino, A.; Di Vaio, P.; Fiorino, F.; Magli, E.; Perissutti, E.; Santagada, V.; Severino, B.; Spada, V.; Fabris, D.; Frecentese, F.; Gatto, B., Non-Natural Linker Configuration in 2,6-Dipeptidyl-Anthraquinones Enhances the Inhibition of TAR RNA Binding/Annealing Activities by HIV-1 NC and Tat Proteins. *Bioconjug Chem* **2018,** *29* (7), 2195-2207.
5. Zagotto, G.; Sissi, C.; Lucatello, L.; Pivetta, C.; Cadamuro, S. A.; Fox, K. R.; Neidle, S.; Palumbo, M., Aminoacyl-anthraquinone conjugates as telomerase inhibitors: synthesis, biophysical and biological evaluation. *Journal of medicinal chemistry* **2008,** *51* (18), 5566-74.
6. Lippens, J. L.; Mangrum, J. B.; McIntyre, W.; Redick, B.; Fabris, D., A simple heated-capillary modification improves the analysis of non-covalent complexes by Z-spray electrospray ionization. *Rapid Commun Mass Spectrom* **2016,** *30* (6), 773-83.
7. Hagan, N.; Fabris, D., Direct mass spectrometric determination of the stoichiometry and binding affinity of the complexes between nucleocapsid protein and RNA stem-loop hairpins of the HIV-1 Psi-recognition element. *Biochemistry* **2003,** *42* (36), 10736-45.