Supporting information for:

Thermodynamic studies of a series of homologous HIV-1 TAR RNA ligands reveal that loose binders are stronger Tat competitors than tight ones.

Lise Pascale[1], Stéphane Azoulay[1], Audrey Di Giorgio[1], Laura Zenacker[1], Marc Gaysinski[1], Pascal Clayette[2], Nadia Patino[1,*]

[1] Institut de Chimie de Nice UMR7272, Université de Nice Sophia Antipolis, 06108 Nice Cedex, France

[2] Neurovirology Department, Bertin Pharma, CEA, 92265 Fontenay-aux-Roses Cedex, France

* To whom correspondence should be addressed. Tel: +33 (0)4 92 07 61 46; Fax: +33 (0)4 92 07 61 51; Email: patino@unice.fr

I. CHEMISTRY

I.1 Abbreviations:

NMP (N-methyl-2-pyrrolidinone), DIPEA (N,N-Diisopropylethylamine), DMF (Dimethylformamide), HBTU (2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluroniumhexafluorophosphate), NMM (N-methylmorpholine), Pyr (Pyridine), TEA (Triethylamine), TFA (Trifluoroacetic acid), TFMSA (Trifluoromethanesulfonic acid), Ac_2O (Acetic anhydride), Boc_2O (terbutyloxy anhydride), THF (Tetrahydrofurane), TIS (Triisopropylsilane), CH_2Cl_2 or DCM (Dichloromethane), EtOAc (Ethyl acetate), cyhex (Cyclohexane), DEA (Diethylamine), MeOH (Methanol), EtOH (Ethanol), DMSO (Dimethylsulfoxide), D_2O (Deuterium oxide), $CDCl_3$ (Deuterium Chloroform), Z (Benzyloxycarbonyl), Boc (Terbutyloxycarbonyl).

I.2 Materials and equipment

I.2.1 PAA monomers synthesis

Solvents and reagents were obtained from commercial sources and used without further purification. α-aminoacid and β-aminoacid residues were purchased from Novabiochem and from Chem-Impex International respectively. Analytical thin-layer chromatography (TLC) was conducted on Merck (VWR) precoated silica gel 60F254 plates and compounds were visualized with ninhydrin test and/or under ultraviolet light (254 nm). Column chromatographies were carried out on silica gel (Merck, SDS 60A, 63-200µm, VWR). Synthesis of α-PAA monomers was previously described (16-17). β-PAA (**13b–15b)** and C-α-PAA monomers (**13c–15c)** were synthesized as described below and analysed by reverse-phase HPLC on a Thermo RP-18 column (3.2 x 250 mm, 5µm, 300 Å) using a Waters apparatus (St Quentin en Yvelines, France) including HPLC Alliance 2695, 996 photodiode array detector. Data were monitored using a Waters Millenium software. All HPLC analyses were run at room temperature. Solvent A and solvent B respectively, 0.1% TFA in water and 0.1% TFA in acetonitrile were used for HPLC studies. For PAA monomers HPLC analyses, the A/B gradients employed at a flow rate of 1 mL/min are given independently for each compound. ^1H, ^{13}C and HSQC NMR spectra of PAA monomers were recorded with Brucker AC 200 (^1H: 200 MHz, ^{13}C: 50 MHz) or AV 500 (^1H: 500 MHz, ^{13}C: 125 MHz) spectrometers using deuterated DMSO or $CDCl_3$ purchased from Eurisotop. Chemical shift (δ) are reported in parts per million (ppm). ^1H NMR splitting patterns are designated as singlet (s), doublet (d), triplet (t), quartet (q). Splitting patterns that could not be interpreted or easily visualized were recorded as multiplet (m) or broad resonance (br). The NMR spectra of some compounds displayed split signals (or broad signals) because of the presence of an equilibrium mixture of two major isomers generated by the substituted amide bonds. Their chemical shifts are described in italics when they can be spotted on spectra. ESI mass spectra were recorded with a Bruker Esquire 3000 plus, equipped with an atmospheric pressure ionization source. This method used in positive mode and negative mode gives respectively either $(M+H)^+$ and/or $(M+Na)^+$ signals and $(M-H)^-$ signals.

I.2.2 Tri-PAA synthesis

Tri-PAA derivatives (**1b-c/7b-c**) were prepared following the solid-phase strategy previously described for tri-α-PAA (**1a/7a**) (16-17). All solvents and reagents for solid-phase synthesis were of peptide grade and purchased from Iris Biotech GmbH. TFMSA, TIS and Ac_2O were obtained from Alpha Aesar. Tri-PAA were analysed by reverse-phase HPLC with the same equipment and solvents as for PAA monomers, applying a A/B gradient from 90/10 to 0/100 for 30 min at a flow rate of 1 mL/min. Semi-preparative HPLC were

performed on a Waters system (600E system controller, 2487 dual wavelength absorbance detector) with a Thermo RP-18 column (10 x 250 mm, 5 μm, 300 Å) at a flow rate of 3 mL/min, with the same solvent gradient as for analytical separations. HRMS analyses were carried out on an LTQ Orbitrap hybrid mass spectrometer with an electrospray ionization probe (Thermo Scientific, San Jose, CA) by direct infusion from a pump syringe.

I.3 Synthesis

I.3.1 Synthesis of backbones

The synthesis of backbones **8** and **8'** was reported previously (16).

Methyl N-(3-(tertiobutyloxy)aminoethyl)-β-alaninate **9**

A solution of methylacrylate (44 mL, 0.5 mol) in 50 mL of acetonitrile was added to a solution of Boc-ethylenediamine (4 g, 25 mmol) in 100mL of acetonitrile, under an inert atmosphere of argon and in a dark place. The reaction was stirred under reflux for 9 h then the solvent was evaporated under reduced pressure. The crude residue was purified by column chromatography (from EtOAc/cyhex 8:2 to EtOAc/MeOH 9:1, v:v) to afford compound **9** (4.9 g, 79%) as a yellow oil.

R_f (EtOAc/MeOH 6:4) = 0.32

MS (ESI+) calcd for $C_{11}H_{23}N_2O_4Na$ $[M+Na]^+$ = 269.2; found : 269.0 $[M+Na]^+$

^{1}H NMR (200MHz, $CDCl_3$) δ= 4.99 (br s, 1H, NH(Boc)); 3.67 (s, 3H, OCH_3); 3.19 (q, J= 5.8Hz, 2H, CH_2); 2.87 (t, J = 6.4Hz, 2H, CH_2); 2.71 (t, J= 5.8Hz, 2H, CH_2); 2.49 (t, J= 6.4Hz, 2H, CH_2); 1.60 (s, 1H, NH); 1.42 (s, 9H, $3CH_3$).

^{13}C NMR (50MHz, $CDCl_3$) δ= 173.3 (1C, **C**OOMe); 156.2 (1C, **C**O(Boc)); 79.3 (1C, **C_{IV}**(Boc)); 51.7 (1C, O**C**H_3); 48.9 (1C, **C**H_2); 44.7 (1C, **C**H_2); 40.2 (1C, **C**H_2); 34.7 (1C, **C**H_2); 28.5 (3C, 3**C**H_3).

Allyl N-(3-(tertiobutyloxy)aminoethyl)-β-alaninate **9'**

Compound **9'** was obtained following the same procedure than above, starting from 4 g (25 mmol) of Boc-ethylenediamine and 30 mL (0.25 mol) of allylacrylate. Purification by column chromatography (from EtOAc/cyhex 8:2 to EtOAc/MeOH 8:2, v:v) afforded 5.45 g of compound **9'**(80% yield) as a yellow oil.

R_f (EtOAc/MeOH 8:2) = 0.35

MS (ESI+) calcd for $C_{13}H_{24}N_2O_4Na$ $[M+Na]^+$ = 295.2; found : 295.0 $[M+Na]^+$

^{1}H NMR (200MHz, $CDCl_3$) δ= 5.90 (ddt, J= 5.6, 10.20, 17.20Hz, 1H, CH); 5.36-5.19 (m, 2H, CH_2); 5.01 (br s, 1H, NH(Boc)); 4.58 (dt, J= 1.20, 5.80Hz, 2H, CH_2); 3.19 (q, J= 5.8Hz, 2H, CH_2); 2.88 (t, J= 6.4Hz, 2H, CH_2); 2.72 (t, J= 5.8Hz, 2H, CH_2); 2.52 (t, J= 6.4Hz, 2H, CH_2); 2.02 (s, 1H, NH); 1.47 (s, 9H, $3CH_3$).

^{13}C NMR (50MHz, $CDCl_3$) δ= 172.3 (1C, **C**OOAllyl; 156.1 (1C, **C**O(Boc)); 132.0 (1C, CH); 118.4 (1C, CH_2); 79.1 (1C, **C_{IV}**(Boc)); 65.15 (1C, **C**H_2); 48.7 (1C, **C**H_2); 44.5 (1C, **C**H_2); 40.0 (1C, **C**H_2); 34.6 (1C, **C**H_2); 28.4 (3C, 3**C**H_3).

I.3.2 General procedure for the synthesis of N-Boc β-PAA and C-α-PAA methyl and allyl ester monomers (**10b-c/12b-c**)

To a solution of a protected N-Z α- or β-aminoacid (5.16 mmol) in DMF (14 mL) was added DIPEA (17.2 mmol) and HBTU (5.16 mmol). The reaction was cooled to 0°C, stirred for 3 min and subsequently, a solution of backbone **8/8'** or **9/9'** (4.3 mmol) in DMF (2 mL) was added. The mixture was stirred at room temperature for 2 h, and the solvent was removed under reduced pressure. The residue was taken up in EtOAc (100 mL), washed with a 1M $KHSO_4$ solution (3 x 30 mL), a saturated $NaHCO_3$ solution (3 x 30 mL), H_2O (2 x 30 mL) and brine (50 mL). The organic layer was dried on $MgSO_4$ and subsequent removal of solid via filtration and concentration under reduced pressure afforded an oily residue which was purified by flash column chromatography (EtOAc/cyhex 2:8 to 100% EtOAc) to give the corresponding PAA monomer as a pale yellow oil in 80-95% yields.

I.3.2a N-BOC β-PAA methyl and allyl ester monomers

Boc-[Z-β-Phe]OMe **10b**

R_f (EtOAc/cyhex, 1:1, v:v) = 0.33

HPLC (A/B 90:10 to 0:100 over 30 min) Rt= 26.1 min

MS (ESI+) (m/z) calcd for $C_{28}H_{57}N_3O_7Na$ $[M+Na]^+$: 550.3, found: 550.5 $[M+Na]^+$

^{1}H NMR (200MHz, DMSO-d_6): δ (two major isomers)= 7.46-7.07 (m, 11H, $5CH_{Ar}$(Z), $5CH_{Ar}$(Ph), NH); 6.88, *6.72* (t, J= 5.0Hz, *J= 5.2Hz*, 1H, NH(Boc)); 4.94 (s, 2H, CH_2(Z)); 4.37-3.91 (m, 3H, CH, CH_2); *3.68*, 3.62 (s, 3H, OCH_3); 3.46-3.21 (m, 2H, CH_2); 3.16-2.93 (m, 2H, CH_2); 2.88-2.18 (m, 4H, $2CH_2$); 1.36 (s, 9H, $3CH_3$(Boc)).

^{13}C NMR (50MHz, DMSO-d_6) : δ (two major isomers) *170.81*, 170.61 (1C, N-**C**O); *170.21*, 169.89 (1C, **C**OOMe); 155.54, 155.20 (2C, **C**O(Z), **C**O(Boc)); 139.02, *138.95*, 137.12 (2C, $\mathbf{C}_{IV}$(Z), $\mathbf{C}_{IV}$(Ph)); *129.03*, 128.99, 128.19, 128.02, 127.56, 127.39, 125.95 (10C, 5**C**H_{Ar}(Z), 5**C**H_{Ar}(Ph)); 77.82, *77.59* (1C, $\mathbf{C}_{IV}$(Boc)); 64.84 (1C, **C**H_2(Z)); *52.07*, 51.60 (1C, O**C**H_3); 49.99, *49.80* (1C, **C**H); 47.70, 47.38, *46.22* (2C, 2**C**H_2); *40.52*, 40.10 (1C, **C**H_2); *39.69*, 39.27 (1C, **C**H_2); 37.75, *37.48* (1C, **C**H_2); 28.08 (3C, 3**C**H_3(Boc)).

Boc-[Z-β-Lys(Z)]OMe **11b**

R_f (EtOAc 100%) = 0.66

HPLC (A/B 90:10 to 0:100 over 30 min) Rt= 26.0 min

MS (ESI+) (m/z) calcd for $C_{33}H_{46}N_4O_9Na$ $[M+Na]^+$: 665.3, found: 665.6 $[M+Na]^+$

^{1}H NMR (200MHz, DMSO-d_6) δ (two major isomers) 7.59-7.00 (m, 12H, $10CH_{Ar}$(Z), 2NH); 6.83, *6.69* (t, J= 5.3Hz, *J= 4.3Hz*, 1H, NH(Boc)); 4.99 (s, 4H, $2CH_2$(Z)); 4.06-3.91 (m, 2H, CH_2); 3.89-3.72 (m, 1H, CH); *3.63*, 3.61 (s, 3H, OCH_3); 3.52-3.21 (m, 2H, CH_2); 3.15-2.83 (m, 4H, $2CH_2$); 2.62-2.12 (m, 2H, CH_2); 1.58-1.01 (m, 6H, $3CH_2$); 1.35 (s, 9H, $3CH_3$(Boc)).

^{13}C NMR (50MHz, DMSO-d_6) δ (two major isomers) *170.87*, 170.68, *170.20*, 169.88 (2C, N-**C**O, **C**OOMe); 155.99, 155.51, 155.45 (3C, 2**C**O(Z), **C**O(Boc)); 137.20, 137.13 (2C, 2$\mathbf{C}_{IV}$(Z)); 128.26, 127.62, 127.56 (10C, 10**C**H_{Ar}(Z)); 77.81, *77.59* (1C, $\mathbf{C}_{IV}$(Boc)); 65.02 (2C, 2**C**H_2(Z)); *51.73*, 51.56 (1C, O**C**H_3); 48.11, *48.00*, 47.65, 47.31 (3C, 2**C**H_2, **C**H); *40.53*, 40.12 (1C, **C**H_2); 39.70 (1C, **C**H_2); 37.39 (1C, **C**H_2); 33.39 (1C, **C**H_2); 29,22 (1C, **C**H_2); 28.09 (3C, 3**C**H_3(Boc)); 22.93, *22.86* (1C, **C**H_2).

Boc-[Z-β-Arg(Tos)]-OAll **12b**

R_f (EtOAc 100%) = 0.52

HPLC (A/B 90:10 to 0:100 over 30 min) Rt= 23.9 min

MS (ESI+) (m/z) calcd for $C_{34}H_{48}N_6O_9SNa$ $[M+Na]^+$: 739.3, found: 739.6 $[M+Na]^+$

^{1}H NMR (500MHz, DMSO-*d_6*) δ (two major isomers) 7.65-7.25 (m, 9H, $4CH_{Ar}$(Ts), $5CH_{Ar}$(Z); 7.12-7.05 (m, 1H, NH); 6.83-6.79 (m, 1H, NH(Boc)); 6.77-6.49 (m, 3H, 3NH); 5.96-5.84 (m, 1H, CH); *5.32,* 5.30 (dd, J= 17.5Hz and 1.5Hz, 1H, CH_2); *5.24*, 5.20 (m, 1H, CH_2); 5.00 (s, 2H, CH_2(Z)); *4.63*, 4.56 (m, 2H, CH_2); 3.85-3.75 (m, 1H, CH); 3.44-3.24 (m, 4H, $2CH_2$); 3.11-3.04 (m, 2H, CH_2); 3.04-2.91 (m, 2H, CH_2); 2.44-2.35 (m, 2H, CH_2); 2.33 (s, 3H, CH_3(Ts)); 2.24-2.18 (m, 2H, CH_2); 1.47-1.28 (m, 4H, $2CH_2$); *1.36,* 1.35 (s, 9H, $3CH_3$(Boc)).

^{13}C NMR (50MHz, DMSO-*d_6*) δ (two major isomers) 170.96, 169.89, 168.82 (3C, N-**C**O, **C**$_{IV}$(Ts)SO_2, **C**OOAllyl); 156.97 (1C, **C**$_{IV}$(Arg)); 155.67, 155.57 (2C, **C**O(Boc), **C**O(Z)); *141.53*, 141.31 (1C, **C**$_{IV}$(Ts)-CH_3); 137.66, *137.26* (1C, **C**$_{IV}$(Z)); 132.4 (1C, **C**H); 128.92, 128.25, 127.76, 127.59, 125.60 (9C, 4**C**H_{Ar}(Ts), 5**C**H_{Ar}(Z)); 118.13 (1C, **C**H_2); 78,57 (1C, **C**$_{IV}$(Boc)); *65.66*, 65.16 (1C, **C**H_2); 65.50 (1C, **C**H_2(Z)); 48.39, 48.06, 46.73 (3C, 2**C**H_2, **C**H); 40.76 (1C, **C**H_2); 38.76 (1C, **C**H_2); 37.77 (1C, **C**H_2); 31.96, *31.79* (1C, **C**H_2); 28.14 (3C, 3**C**H_3(Boc)); *26.31*, 26.15 (1C, **C**H_2); 20.83 (1C, **C**H_3(Ts)).

I.3.2b N-BOC C-α-PAA methyl and allyl ester monomers

Boc-[Z-Phe]-CH_2-OMe **10c**

R_f (EtOAc /cyclohex 1:1, v/v) = 0.53

HPLC (A/B 90:10 to 0:100 over 30 min) Rt= 26.4 min

MS (ESI+) (m/z): calcd for $C_{28}H_{37}N_3O_7Na$ $[M+Na]^+$= 550.3; found 550.3 $[M+Na]^+$

^{1}H NMR (200MHz, $CDCl_3$) δ (two major isomers) 7.49-7.24 (m, 10H, $5CH_{Ar}$(Z), $5CH_{Ar}$(Ph)); 5.90 (d, J= 8,8Hz, 1H, NH(Z)); 5.27-5.05 (m, 3H, CH_2(Z), NH(Boc)); 5.03-4.83 (m, 1H, CH); 3.97-3.54 (m, 2H, CH_2); 3.75, *3.73* (s, 3H, OCH_3); 3.50-2.95 (m, 6H, $3CH_2$); 2.75-2.49 (m, 1H, CH_2); 2.44-2.18 (m, 1H, CH_2); 1.52, *1.50* (s, 9H, $3CH_3$(Boc)).

^{13}C NMR (50MHz, $CDCl_3$) δ (two major isomers) *172.46*, 172.22, 171.99, *171.07* (2C, N-**C**O, **C**OOMe); *156.04*, 155.97, *155.83*, 155.77 (2C, **C**O(Z), **C**O(Boc)); *136.37*, 136.31, 136.16, *136.10* (2C, **C**$_{IV}$(Ph), **C**$_{IV}$(Z)); 129.54, 128.78, 128.68, 128.62, 128.24, 128.12, 128.05, 127.41, 127.22 (10C, 5**C**H_{Ar}(Z), 5**C**H_{Ar}(Ph)); *79.69*, 79.49 (1C, **C**$_{IV}$(Boc)); 67.07, *67.01* (1C, **C**H_2(Z)); 52.19 (1C, **C**H); 51.97, *51.80* (1C, O**C**H_3); *47.63*, 46.36 (1C, **C**H_2); *44.08*, 43.11 (1C, **C**H_2); *40.20*, 40.10 (1C, **C**H_2); *38.97*, 38.88 (1C, **C**H_2); *33.29*, 32.26 (1C, **C**H_2); *28.49*, 28.42 (3C, 3**C**H_3(Boc)).

Boc-[Z-Lys(Z)]-CH_2-OMe **11c**

R_f (EtOAc /cyhex 8:2, v/v) = 0.62

HPLC (A/B 90:10 to 0:100 over 30 min) Rt= 26.4 min

MS (ESI+) (m/z) calcd for $C_{33}H_{46}N_4O_9Na$ $[M+Na]^+$= 665.3; found 665.4 $[M+Na]^+$

^{1}H NMR (200MHz, $CDCl_3$) δ (two major isomers) 7.80-6.82 (m, 10H, $10CH_{Ar}$(Z)); 6.03-5.60 (m, 1H, NH); 5.52-4.83 (m, 5H, NH(Boc), $2CH_2$(Z)); 4.72-4.47 (m, 1H, CH); 3.92-2.88 (m, 8H, $4CH_2$); *3.66*, 3.63 (s, 3H, OCH_3); 2.85-2.31 (m, 2H, CH_2); 1.77-1.46 (m, 6H, $3CH_2$); 1.41, *1.40* (s, 9H, $3CH_3$(Boc)).

^{13}C NMR (50MHz, $CDCl_3$) δ (two major isomers) *173.24, 172.83 (*1C, N-**C**O);172.34, 171.28 (1C, **C**OOMe); 156.80, *156.74* (1C, **C**O(Boc)); *156.38*, 156.31, 156.21 (2C, 2**C**O(Z)); *136.68*, 136.64, *136.35*, 136.32 (2C, 2$\mathbf{C}_{IV}$(Z)); 128.60, 128.18, *128.07* (10C, 10$\mathbf{C}H_{Ar}$(Z)); 79.77 (1C, $\mathbf{C}_{IV}$(Boc)); 67.06, 66.80 (2C, 2$\mathbf{C}H_2$(Z)); *52.11*, 51.88 (1C, O$\mathbf{C}H_3$); *50.83*, 50.57 (1C, **C**H); *48.13*, 46.01 (1C, $\mathbf{C}H_2$); *44.06*, 43.12 (1C, $\mathbf{C}H_2$); *40.64*, 40.56 (1C, $\mathbf{C}H_2$); 38.99 (1C, $\mathbf{C}H_2$); *33.64*, 32.42 (1C, $\mathbf{C}H_2$); 32.80, *32.70* (1C, $\mathbf{C}H_2$); *29.39*, 29.31 (1C, $\mathbf{C}H_2$); *28.47*, 28.42 (3C, 3$\mathbf{C}H_3$(Boc)); 22.40, *22.20* (1C, $\mathbf{C}H_2$).

Boc-[Z-Arg(Z^2)]-CH_2-OAll **12c**

R_f (EtOAc/cyhex 1:1, v/v) = 0.56

HPLC (A/B 90:10 to 0:100 over 30 min) Rt= 29.9 min

MS (ESI+) (m/z) calcd for $C_{43}H_{54}N_6O_{11}Na$ $[M+Na]^+$= 853.4; found 853.6 $[M+Na]^+$

^{1}H NMR (200MHz, $CDCl_3$) δ (two major isomers) 9.17 (br s, 2H, 2NH); *7.59*, 7.51 (d, *J= 8.0Hz*, J= 7.6Hz, 1H, NH); 7.45-7.19 (m, 15H, 15CH_{Ar}(Z)); 6.90, *6.82* (t, J= 5.2Hz, *J= 4.0Hz*, 1H, NH(Boc)); 5.99-5.76 (m, 1H, CH); 5.26 (dd, J= 20.4 and 1.6Hz, 1H, CH_2); 5.21 (s, 2H, CH_2(Z)); 5.18 (dd, J= 10.4 and 1.4Hz, 1H, CH_2); 5.04 (s, 2H, CH_2(Z)); 4.99 (s, 2H, CH_2(Z)); 4.57-4.46 (m, 2H, CH_2); 4.43-4.28 (m, 1H, CH); 3.94-3.72 (m, 2H, CH_2); 3.68-2.88 (m, 6H, 3CH_2); 2.80-2.65 (m, 1H, CH_2); 2.58-2.43 (m, 1H, CH_2); 1.80-1.42 (m, 4H, 2CH_2); *1.34*, 1.33 (s, 9H, 3CH_3(Boc)).

^{13}C NMR (50MHz, $CDCl_3$) δ (two major isomers) 171.89, *171.74* (1C, N-**C**O); 170.51, *170.33* (1C, **C**OOAll); 162.93, 159.72, 154.99 (3C, 3**C**O(Z)); *156.04*, 155.94 (1C, **C**O(Boc)); 155.64 (1C, $\mathbf{C}_{IV}$(Arg)); 137.02, 136.88, *136.83*, 135.18 (3C, 3$\mathbf{C}_{IV}$(Z)); 132.46, *132.52* (1C, CH); 128.50, 128.30, *127.93*, 127.90, 127.84, *127.76*, 127.69, 127.65 (15C, 15$\mathbf{C}H_{Ar}$(Z)); *117.76*, 117.72 (1C, $\mathbf{C}H_2$); 77.83, *77.68* (1C, $\mathbf{C}_{IV}$(Boc)); 68.11, 66.07, *65.42*, 65.34 (3C, 3$\mathbf{C}H_2$(Z)); *64.50*, 64.39 (1C, $\mathbf{C}H_2$); *50.58*, 50.44 (1C, **C**H); 46.86, *45.52*, 44.29 (2C, 2$\mathbf{C}H_2$); 42.36 (1C, $\mathbf{C}H_2$); 37.83 (1C, $\mathbf{C}H_2$); *33.11*, 31.79 (1C, $\mathbf{C}H_2$); 28.87, *28.73* (1C, $\mathbf{C}H_2$); *28.09*, 28.02 (3C, 3$\mathbf{C}H_3$(Boc)); 24.87 (1C, $\mathbf{C}H_2$).

I.3.3 General procedure for the synthesis of N-Boc β-PAA and C-α-PAA acid monomers (**13b-15b**, **13c-15c**)

To a solution of a PAA methyl ester monomer (**10b-c/12b-c**) (2.5 mmol) in THF (7 mL) cooled to 0°C was slowly added an aqueous 1M lithium hydroxide solution (5.5 mmol). After stirring for 1 h at 0°C, the mixture was neutralised (pH = 7) with a 1M $KHSO_4$ solution and the solvent was concentrated under reduced pressure. The remaining aqueous layer was acidified with a 1M $KHSO_4$ solution until pH = 2, then extracted with EtOAc (2 x 40 mL). The combined organic layers were washed with H_2O and brine, dried on $MgSO_4$ and the solvent was evaporated under reduced pressure to afford without further purification the corresponding PAA acid monomer (**13b-15b, 13c-15c**) as a white solid, in 80-90% yields.

I.3.3a N-BOC β-PAA acid monomers

Boc-[Z-β-Phe]-OH **13b**

HPLC (A/B 90:10 to 0:100 over 30 min) Rt= 23.5 min

MS (ESI-) (m/z) calcd for $C_{27}H_{34}N_3O_7$ $[M-H]^-$ = 512.2; found 512.3 $[M-H]^-$

^{1}H NMR (200MHz, DMSO-d_6) δ (two major isomers) 7.47-7.03 (m, 11H, 5CH_{Ar}(Z), 5CH_{Ar}(Ph), NH); 6.99-6.68 (m, 1H, NH(Boc)); 5.05-4.78 (m, 2H, CH_2(Z)); 4.27-3.64 (m, 3H, CH, CH_2); 3.55-3.19 (m, 2H, CH_2); 3.17-2.94 (m, 2H, CH_2); 2.92-2.21 (m, 4H, 2CH_2); *1.36*, 1.34 (s, 9H, 3CH_3(Boc)).

^{13}C NMR (50MHz, DMSO-d_6) δ (two major isomers) 171.01, 170.10 (2C, **C**OOH, N-**C**O); 155.64, *155.59*, 155.36 (2C, **C**O(Z), **C**O(Boc)); 139.34, *139.31*, *137.24*, 137.20 (2C, **C**$_{IV}$(Z), **C**$_{IV}$(Ph)); 129.12, 128.27, 128.04, 127.61, 127.43, 125.94 (10C, 5**C**H_{Ar}(Z), 5**C**H_{Ar}(Ph)); *77.88*, 77.53 (1C, **C**$_{IV}$(Boc)); 64.90 (1C, **C**H_2(Z)); 50.09 (1C, **C**H); 47.97, 46.66 (2C, 2**C**H_2); 40.11 (1C, **C**H_2); 39.27 (1C, **C**H_2); 37.85, *37.40* (1C, **C**H_2); 28.21 (3C, 3**C**H_3(Boc)).

Boc-[Z-β-Lys(Z)]-OH **14b**

HPLC (A/B 90:10 to 0:100 over 30 min) Rt= 24.0 min

MS (ESI-) (m/z) calcd for $C_{32}H_{43}N_4O_9$ $[M-H]^-$ = 627.3; found 627.5 $[M-H]^-$

^{1}H NMR (200MHz, DMSO-d_6) δ (two major isomers) 7.42-7.20 (m, 11H, 10CH_{Ar}(Z), NH); 7.19-7.04 (m, 1H, NH); *6.99-6.89*, 6.87-6.60 (m, 1H, NH(Boc)); 5.01 (s, 4H, 2CH_2(Z)); 4.07-3.59 (m, 3H, CH, CH_2); 3.50-3.22 (m, 2H, CH_2); 3.18-2.83 (m, 4H, 2CH_2); 2.66-2.10 (m, 2H, CH_2); 1.65-1.07 (m, 6H, 3CH_2); 1.36 (s, 9H, 3CH_3(Boc)).

^{13}C NMR (50MHz, DMSO-d_6) δ (two major isomers) 171.10, 170.16 (2C, N-**C**O, **C**OOH); 156.14, 155.67, 155.59 (3C, 2**C**O(Z), **C**O(Boc)); 137.32, 137.25 (2C, 2**C**$_{IV}$(Z)); 128.35, *128.09*, *127.94*, 127.74, 127.69, *127.40* (10C, 10**C**H_{Ar}(Z)); *77.91*, 77.56 (1C, **C**$_{IV}$(Boc)); 65.15 (2C, 2**C**H_2(Z)); 48.24 (1C, **C**H); *46.88*, 46.78, 46.68 (2C, 2**C**H_2); 37.96 (1C, **C**H_2); 33.92, *33.81* (1C, **C**H_2); 29.36 (1C, **C**H_2); 28.26 (3C, 3**C**H_3(Boc)); 23.14 (1C, **C**H_2); NB; two CH_2 signals are hidden under the DMSO signal.

Boc-[Z-β-Arg(Tos)]-OH **15b**

HPLC (A/B 90:10 to 0:100 over 30 min) Rt= 20.8 min

MS (ESI-) (m/z) calcd for $C_{31}H_{43}N_6O_9SNa$ $[M-H]^-$ = 675.3; found 675.4 $[M-H]^-$

^{1}H NMR (200MHz, DMSO-d_6) δ (two major isomers) 7.70-6.89 (m, 14H, 4CH_{Ar}(Ts), 5CH_{Ar}(Z), 5NH); 4.98 (s, 2H, CH_2(Z)); 3.94-3.19 (m, 5H, CH, 2CH_2); 3.12-2.91 (m, 4H, 2CH_2); 2.46-2.02 (m, 2H, CH_2); 2.32 (s, 3H, CH_3(Tos)); 1.61-1.25 (m, 4H, 2CH_2); 1.34 (s, 9H, 3CH_3(Boc)).

^{13}C NMR (50MHz, DMSO-d_6) δ (two major isomers) *173.49*, 173.23 (1C, **C**OOH); 171.02 (1C, N-**C**O); 169.96 (1C, **C**$_{IV}$(Ts)-SO_2); 156.97 (1C, **C**$_{IV}$(Arg)); 155.67, 155.57 (2C, **C**O(Boc), **C**O(Z)); *141.91*, 140.99 (1C, **C**$_{IV}$(Ts)-CH_3); 137.31, *137.27* (1C, **C**$_{IV}$(Z)); 131.65, 131.46, 129.06, 128.93, 128.70, 128.37, 127.69, 127.65, 125.65 (9C, 4**C**H_{Ar}(Ts), 5**C**H_{Ar}(Z)); *77.78*, 77.50 (1C, **C**$_{IV}$(Boc)); 65.12 (1C, **C**H_2(Z)); 49.62, 48.08, 47.02 (3C, **C**H, 2**C**H_2); 40.20 (1C, **C**H_2); 37.92 (1C, **C**H_2); 31.68 (1C, **C**H_2); 28.28 (3C, 3**C**H_3(Boc)); *25.77*, 25.71 (1C, **C**H_2); 20.92 (1C, **C**H_3(Ts)).

I.3.3b N-BOC C-α-PAA acid monomers

Boc-[Z-Phe]-CH_2-OH **13c**

HPLC (A/B 90:10 to 0:100 over 30 min) Rt= 23.5 min

MS (ESI-) (m/z) calcd for $C_{27}H_{34}N_3O_7$ $[M-H]^-$ = 512.2; found 512.2 $[M-H]^-$

^{1}H NMR (200MHz, $CDCl_3$) δ (two major isomers) 7.52-7.21 (m, 10H, $5CH_{Ar}$(Z), $5CH_{Ar}$(Ph)); *6.07*, 5.96 (d, *J= 8.6Hz*, J= 8.6Hz, 1H, NH(Z)); 5.33-5.83 (m, 4H, NH(Boc), CH_2(Z), CH); 3.95-2.93 (m, 8H, $4CH_2$); 2.76-2.53 (m, 1H, CH_2); 2.48-2.25 (m, 1H, CH_2); 1.52, *1.50* (s, 9H, $3CH_3$(Boc)).

^{13}C NMR (50MHz, $CDCl_3$): δ (two major isomers) 175.13 (1C, **C**OOH); 173.79 (1C, N-**C**O); 156.31, *156.29*, *156.20*, 156.04 (2C, **C**O(Z), **C**O(Boc)); *136.31*, 136.25, 136.02 (2C, $\mathbf{C}_{IV}$(Ph), $\mathbf{C}_{IV}$(Z)); 129.62, 129.55, 128.81, 128.78, 128.66, 128.30, 128.22, 128.15, 128.05, 127.43, 127.36 (10C, 5**C**H_{Ar}(Z), 5**C**H_{Ar}(Ph)); 79.92, *79.78* (1C, $\mathbf{C}_{IV}$(Boc)); 67.20, *67.14* (1C, **C**H_2(Z)); *52.13*, 52.10 (1C, **C**H); 47.61, *46.41* (1C, **C**H_2); 44.16, *43.12* (1C, **C**H_2); 40.12, *40.00* (1C, **C**H_2); *38.79*, 38.74 (1C, **C**H_2); *33.14*, 32.26 (1C, **C**H_2); *28.51*, 28.43 (3C, 3**C**H_3(Boc)).

Boc-[Z-Lys(Z)]-CH_2-OH **14c**

HPLC (A/B 90:10 to 0:100 over 30 min) Rt= 24.1 min

MS (ESI-) (m/z) calcd for $C_{32}H_{43}N_4O_9$ $[M-H]^-$ = 627.3; found 627.4 $[M-H]^-$

^{1}H NMR (200MHz, $CDCl_3$) δ (two major isomers) 7.78-6.81 (m, 10H, $10CH_{Ar}$(Z)); 6.08-5.86, *5.52-5.39* (m, 1H, NH); 5.26-4.94 (m, 5H, NH(Boc), $2CH_2$(Z)); 4.83-4.49, *4.34-4.08* (m, 1H, CH); 4.00-2.97 (m, 8H, $4CH_2$); 2.89-2.39 (m, 2H, CH_2); 1.82-1.34 (m, 15H, $3CH_2$, $3CH_3$(Boc)).

^{13}C NMR (50MHz, $CDCl_3$) δ (two major isomers) 174.53 (1C, **C**OOH); 173.34, *173.04* (1C, N-**C**O); 156.91, 156.46, 156.36 (3C, 2**C**O(Z), **C**O(Boc)); 136.57, 136.31, *136.27* (2C, 2$\mathbf{C}_{IV}$(Z)); 128.58, 128.20, 128.08, 128.02 (10C, 10**C**H_{Ar}(Z)); 79.75, *79.64* (1C, $\mathbf{C}_{IV}$(Boc)); *67.08*, 67.04, 66.80 (2C, 2**C**H_2(Z)); *50.82*, 50.62 (1C, **C**H); *48.24*, 46.01 (1C, **C**H_2); 44.19, *43.11* (1C, **C**H_2); 40.56 (1C, **C**H_2); *38.90*, 38.73 (1C, **C**H_2); *33.59*, 32.29 (1C, **C**H_2); 32.75, *32.62* (1C, **C**H_2); 29.34 (1C, **C**H_2); *28.45*, 28.40 (3C, 3**C**H_3(Boc)); 22.38, *22.15* (1C, **C**H_2).

Boc-[Z-Arg(Z^2)]-CH_2-OH **15c**

HPLC (A/B 90:10 to 0:100 over 30 min) Rt= 21.0 min

MS (ESI-) (m/z) calcd for $C_{40}H_{49}N_6O_{11}$ $[M-H]^-$ = 789.4; found 789.6 $[M-H]^-$

^{1}H NMR (200MHz, DMSO-d_6) δ (two major isomers) = 9.22 (br s, 2H, 2NH); 7.90-7.08 (m, 16H, NH, $15CH_{Ar}$(Z)); 7.04-6.90, *6.89-6.76* (m, 1H, NH(Boc)); 5.23 (s, 2H, CH_2(Z)); 5.07 (s, 2H, CH_2(Z)); 5.02 (s, 2H, CH_2(Z)); 4.41 (br s, 1H, CH); 4.09-3.72 (m, 2H, CH_2); 3.69-2.85 (m, 6H, $3CH_2$); 2.75-2.53 (m, 1H, CH_2); 2.47-2.25 (m, 1H, CH_2); 1.88-1.47 (m, 4H, $2CH_2$); 1.36 (s, 9H, $3CH_3$(Boc)).

^{13}C NMR (50MHz, DMSO-d_6) δ (two major isomers) = *173.87*, 173.83 (1C, **C**OOH); *171.92*, 171.88 (1C, N-**C**O); 163.14, 159.92 (2C, 2**C**O(Z); *156.22*, 156.14, 155.84 (2C, **C**O(Boc), $\mathbf{C}_{IV}$(Arg)); 155.20, *155.18* (1C, **C**O(Z)); 137.26, *137.11*, 137.05, 135.38 (3C, 3$\mathbf{C}_{IV}$(Z)); *128.95*, 128.66, 128.44, 128.12, 128.02, 127.82 (15C, 15**C**H_{Ar}(Z)); 77.97, *77.80* (1C, $\mathbf{C}_{IV}$(Boc)); 68.40, 66.37, 65.73, *65.64* (3C, 3**C**H_2(Z)); *50.88*, 50.74 (1C, **C**H); *47.10*, 45.79, 44.53 (2C, 2**C**H_2); 42.92 (1C, **C**H_2); 37.88 (1C, **C**H_2); *34.30*, 32.85 (1C, **C**H_2); 29.18 (1C, **C**H_2); *28.33*, 28.25 (3C, 3**C**H_3(Boc)); 25,15 (1C, **C**H_2).

I.3.4 Characterization of C-α and β-PAA trimers

Table S1: Characterization of C-α and β-PAA trimers

Tri-PAA	HPLC retention time (min)	HPLC purity of crude compound (%)	Molecular Formula	HRMS (ESI+) m/z	
				$[M+H]^+$ calculated	$[M+H]^+$ found
C-α-FFR	10.5	87	$C_{44}H_{71}N_{14}O_8$	923.55738	923.55731
C-α-FRF	9.6	85	$C_{44}H_{71}N_{14}O_8$	923.55738	923,55884
C-α-RFF	10.1	90	$C_{44}H_{71}N_{14}O_8$	923.55738	923.55669
C-α-FRR	8.9	93	$C_{41}H_{74}N_{17}O_8$	932.59008	932.58913
C-α-RRF	8.5	86	$C_{41}H_{74}N_{17}O_8$	932.59008	932.58893
C-α-FFK	10.4	88	$C_{44}H_{71}N_{12}O_8$	895.55123	895.55115
C-α-FKR	8.6	85	$C_{41}H_{74}N_{15}O_8$	904.58393	904.58350
β-FFR	10.3	96	$C_{44}H_{71}N_{14}O_8$	923.55738	923.55713
β-FRF	10.2	83	$C_{44}H_{71}N_{14}O_8$	923.55738	923.55737
β-RFF	10.4	78	$C_{44}H_{71}N_{14}O_8$	923.55738	923.55682
β-FRR	9.3	98	$C_{41}H_{74}N_{17}O_8$	932.59008	932.59052
β-RRF	9.5	98	$C_{41}H_{74}N_{17}O_8$	932.59008	932.58867
β-FFK	10.9	92	$C_{44}H_{71}N_{12}O_8$	895.55123	895.55063
β-FKR	8.9	83	$C_{41}H_{74}N_{15}O_8$	904.58393	904.58390

HPLC conditions: Solvent A: 0.1% TFA/H_2O, solvent B (0.1% TFA/CH_3CN); gradient (A/B) from 90/10 to 0/100 for 30 min; flow rate: 1 mL/min.

II. TAR BINDING STUDIES

II.1 NMR experiments

II.1.1 Materials and equipment

High resolution NMR experiments were recorder on a BRUKER AVANCE Ultra shield DRX 500 spectrometer operating at 500.13 MHz for ^1H, equipped with a temperature control unit (BCU 6.0, BVT 3000), and an inverse probe head (5mm PHTXI 1H-13C/15N Z-GRD). Proton chemical shift was referenced internally by setting the carrier frequency on water at the center of the spectrum (4.71ppm at 13°C and 4.70ppm at 35°C). Chemical shifts (δ) are expressed in parts per million (ppm). All NMR experiments were carried out using standard pulse sequences supplied by the spectrometer manufacturer (BRUKER). 1D and 2D spectra were processed using TOPSPIN 2.1 NMR Software (BRUKER).

For the preparation of all NMR samples, a folding of TAR RNA (100 µM) was performed in the appropriate buffer (*vide infra*) as described in material and methods section of the paper. After refolding, the NMR sample (alone or with the appropriate amount of PAA) is incorporated into a Shigemi NMR tube.

^{1}H NMR imino proton spectra were recorded in a H_2O/D_20 (90/10) buffer containing 20 mM phosphate and 50 mM NaCl at 286°K (13°C) by using a WATERGATE 3-9-19 water suppression. Each proton NMR spectrum was acquired using 10.964 KHz Spectral Width (SW), 64K complex data point, acquisition time (aq) of 2.98 s, relaxation delay (D1) of 1s, number of scan (ns) between 1000 and 2000, number of dummy scan (ds) 4 and a 90° flip angle pulse width. Water suppression was achieved using WATERGATE pulse sequence. Gradient pulse were sine shape (SINE.100), 1.5 ms long (P16) with 100 μs gradient recovery delay (D16) and strengths set to 8.44 Gauss.cm^{-1} (20%). A 45.6 μs delay (D19) was used for binomial water suppression. Prior to Fourrier transformation, the fids were multiplied by an exponential line broadening function of 3 Hz.

gs-TOCSY Phase sensitive (States – TPPI mode) experiments using MLEV 17 pulse sequence for spin lock were recorded in a D_20 buffer (50 mM NaCl, 20 mM phosphate, pH 7.4) at 308K (35°C) by using a WATERGATE 3-9-19 water suppression. Each TOCSY 2D NMR spectrum was acquired with a spectral width of 5 KHz in both dimension, 2K complex data point in F2, 256 t1 increments (between 32 and 64 scans by increment) in F1, 0.20 s for aq and D1 of 2 s. MLEV 17 pulse sequence for spin lock was set to 60 ms. Water suppression was achieved using WATERGATE pulse sequence. Gradient pulse were sine shape (SINE.100), 1.5 ms long (P16) with 100 μs gradient recovery delay (D16) and strengths set to 8.44 Gauss.cm^{-1} (20%). A 100 μs delay (D19) was used for binomial water suppression. Prior to Fourier transformation a QSINE window function (SSB =2) was applied in both dimension and the data were zero filled and linear predicted (NC=32) to 1K data points in F1.

II.1.2 2D TOCSY

Table S2: Chemical shift variation of TAR RNA pyrimidine H5-H6 cross-peaks upon addition of 5 eq. of β-PAA **(β-FRR** and **β-FRF**).

Nucleic base	TAR / **β-FRR** (1/5)		TAR / **β-FRF** (1/5)	
	H5	H6	H5	H6
C18	-0.03	0.01	-0.04	0
C19	-0.01	0.01	-0.02	0
U23	0.09	-0.10	0.05	-0.13
C24	0.17	0.09	0.1	0.06
U25	0.12	0.07	0.1	0.05
C29	-0.03	0.1	-0.03	0,10
C30	-0.02	0.04	-0.04	0.03
U31	-0.01	-0.01	-0.01	-0.02
C37	-0.03	0.07	-0.01	0.05
U38	0	0.02	0.01	0.02
C39	-0.09	-0.27	-0.07	-0.22
U40	-0.15	0.18	-0.12	0.16
C41	-0.18	0.01	-0.17	0
U42	-0.13	-0.03	-0.15	-0.02

II.2 Comment on statistical validity for EEC phenomena.

It has been argued that the apparent compensation may be in many cases a statistical artifact resulting from experimental uncertainties in the measured thermodynamic data collected using the van't Hoff analysis. Krug proposed a simple statistical test to determine whether the observed compensation is relevant or not by determining the confidence interval for the compensation temperature (Tc) (50). If the experimental temperature Texp lies outside this confidence interval, the correlation would be significant at the 95% confidence level. In our case, Texp falls always outside the confidence limits, which could indicate that the observed phenomenon has a biological relevance.

II.3 UV melting studies

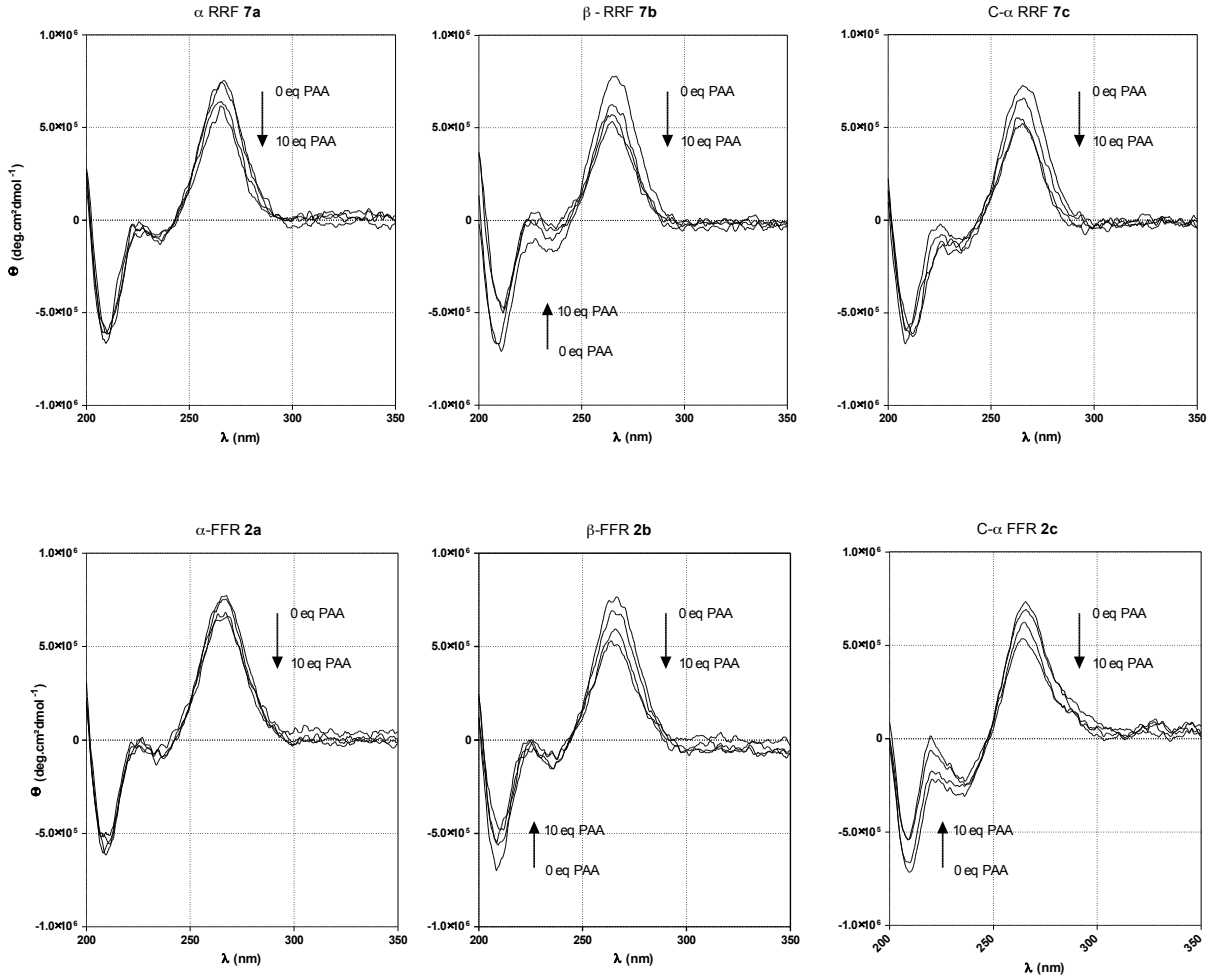

Figure S1: Circular Dichroism spectra of PAA-TAR complexes at 20°C.

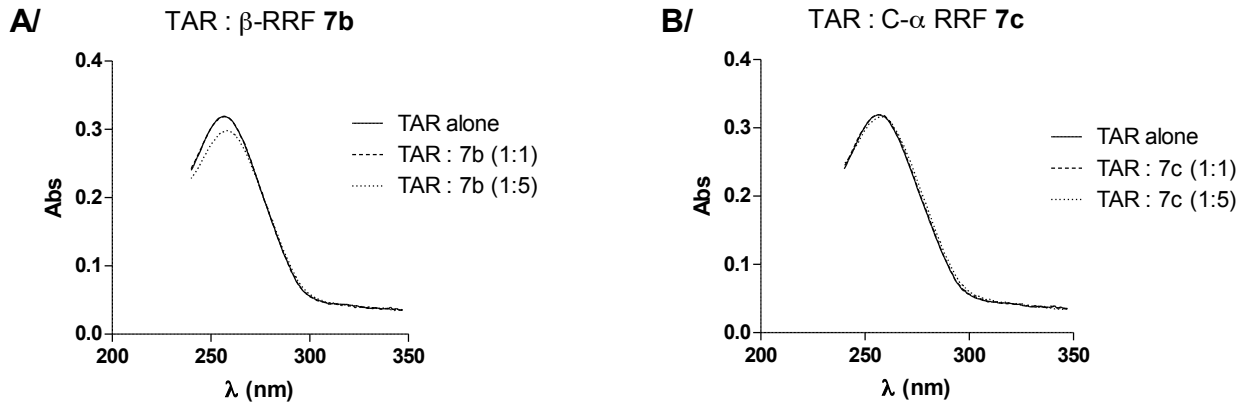

Figure S2: Absorption spectra of TAR and its complex with β-PAA 7b (A) or C-α-PAA 7c (B)

Table S3: melting temperature of TAR and TAR-PAA complexes

	Tm (°C)	***ΔTm (°C)***
TAR alone	57.6 ± 0.1	-
TAR : 7b (1:1)	61.0 ± 0.1	+ 3.4
TAR : 7b (1:5)	64.3 ± 0.1	+ 6.7
TAR : 7c (1:1)	60.1 ± 0.2	+ 2.5
TAR : 7c (1:5)	61.7 ± 0.1	+ 4.1

II.4 Biological assay: Antiviral activity of tri-PAA compounds

Antiviral activity of PAA was evaluated on phytohemaglutinin-P (PHA-P)-activated PBMC experimentally infected with the X4-tropic HIV-1-LAI strain (51). PBMC were pre-treated for 30 min by six concentrations of each drug (1:10 dilutions between 100 µM and 1 nM) and infected with one hundred 50% tissue culture infectious doses (TCID50) per 100,000 cells of the HIV-1-LAI strain.

This virus was amplified *in vitro* on PHA-P-activated PBMC and the viral stock titrated using PHA-P-activated PBMC and the Kärber's formula. In our antiviral assay, molecules were maintained throughout the culture, and cell supernatants were collected at day 7 post-infection and stored at -20°C. Azidothymidine (AZT) at 10 nM was used in these experiments as internal control; this AZT concentration corresponds to the ED99 in our experimental model (52). Viral replication was measured by quantifying reverse transcriptase (RT) activity in cell culture supernatants using the RetroSys HIV RT kit (Innovagen). In parallel, cytotoxicity of the samples was evaluated in uninfected PHA-P-activated PBMC by methyltetrazolium salt (MTT) assay on day 7. Experiments were performed in triplicate and percentages of inhibition were calculated to quantify the cell viability and the anti-HIV activity of each PAA compound.

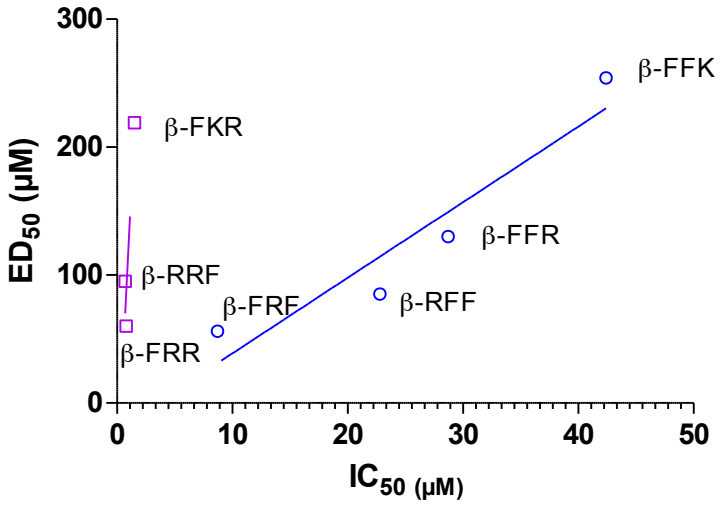

Figure S3: Correlation between IC_{50} and ED_{50}

- **Supplementary references**

16. Bonnard,V., Azoulay,S., Di Giorgio,A. and Patino,N. (2009) Polyamide amino acids: a new class of RNA ligands. *Chem. Commun. (Camb.)*, 2302-2304.

17. Bonnard,V., Pascale,L., Azoulay,S., Di Giorgio,A., Rogez-Kreuz,C., Storck,K., Clayette,P. and Patino,N. (2010) Polyamide amino acids trimers as TAR RNA ligands and anti-HIV agents. *Bioorg. Med. Chem.*, **18**, 7432-7438.

50. Sharp,K. (2001) Entropy-enthalpy compensation: fact or artifact? *Protein Sci.*, **10**, 661-667.

51. Baleux,F., Loureiro-Morais,L., Hersant,Y., Clayette,P., renzana-Seisdedos,F., Bonnaffe,D. and Lortat-Jacob,H. (2009) A synthetic CD4-heparan sulfate glycoconjugate inhibits CCR5 and CXCR4 HIV-1 attachment and entry. *Nat. Chem. Biol.*, **5**, 743-748.

52. Barre-Sinoussi,F., Chermann,J.C., Rey,F., Nugeyre,M.T., Chamaret,S., Gruest,J., Dauguet,C., xler-Blin,C., Vezinet-Brun,F., Rouzioux,C. *et al.* (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, **220**, 868-871.