Electronic Supplementary Material (ESI) for Chemical Communications

New anti-HIV aptamers based on tetra-end-linked DNA G-quadruplexes: effect of the base sequence on anti-HIV activity

Valentina D'Atri,[a] Giorgia Oliviero,[a] Jussara Amato,[a] Nicola Borbone,*[a] Stefano D'Errico,[a] Luciano Mayol,[a] Vincenzo Piccialli,[b] Shozeb Haider,[c] Bart Hoorelbeke,[d] Jan Balzarini[d] and Gennaro Piccialli[a]

[a] *Dipartimento di Chimica delle Sostanze Naturali, Università degli Studi di Napoli Federico II, Via D. Montesano 49, I-80131 Napoli, Italy. E-mail: Nicola.borbone@unina.it.*
[b] *Dipartimento di Chimica, Università degli Studi di Napoli Federico II, Via Cynthia 4, I-80126 Napoli, Italy.*
[c] *Centre for Cancer Research and Cell Biology, Queen's University of Belfast, 97 Lisburn Road, BT9 7BL, Belfast, UK*
[d] *Rega Institute for Medical Research, KU Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium*

Contents **Page Number**

General instrumentations

Automated solid-phase oligonucleotide synthesis was performed on a 8909 DNA-Synthesizer (Applied Biosystems). Mass spectra of TEL-ODNs were performed on a Bruker Autoflex I MALDI-TOF spectrometer using a picolinic/3-hydroxypicolinic acid mixture as the matrix. CD measurements were performed on a JASCO J-715 spectropolarimeter equipped with a Peltier Thermostat JASCO ETC-505T using 0.1 cm path length cuvettes. NMR spectra were acquired on Varian ^{UNITY}INOVA 500 MHz, 700 MHz and Mercury Plus 400 MHz spectrometers and processed using the Varian VNMR software package. For the experiments in H_2O, water suppression was achieved by including a double pulsed-field gradient spin-echo (DPFGSE) module[1,2] in the pulse sequence prior to acquisition. Chemical shifts are reported in parts per million (δ) relative to the residual solvent signals.

Synthesis of 3'-O-tert-butyldiphenylsilyl-propyl-1'-O-(2-cyanoethyl)-N,N'-diisopropylphosphoramidite

1,3-propandiol (1 g, 13.1 mmol) dissolved in 5 mL of dimethylformammide (DMF) was treated with imidazole (1.07 g, 15.7 mmol) and TBDPS-chloride (3.24 g, 11.8 mmol) at 0°C under stirring at room temperature. After 2 h the reaction was quenched by addition of water and the resulting mixture was taken to dryness, dissolved in ethyl acetate (30 mL) and washed with water. The organic phase was dried over sodium sulphate, evaporated to an oil and then purified by column chromatography (50 g Kieselgel 60H) eluted with esane/ethyl acetate 80:20 (v/v), giving 3.3 g (10.5 mmol, 80% yield) of the monosilylated 1,3-(O-TBDPS)-propandiol.

^{1}H-NMR (400 MHz; CD_3OD): δ 7.67-7.66 (m, 4H, Ph), 7.45-7.37 (m, 6H, Ph), 3.78 (t, J = 6.2 Hz, 2H, CH_2O), 3.70 (t, J = 6.4 Hz, 2H, CH_2O), 1.79 (quintuplet, J = 6.3 Hz, 2H, CH_2), 1.03 (s, 9H, *tert*-Bu).

To the above product, dissolved in 20 mL of dry CH_2Cl_2, N,N-diisopropylethylammine (2.9 mL, 12.6 mmol) and 2-cyanoethyl-N,N'-diisopropylchlorophosphoramidite (2.81 mL, 12.6 mmol) were added and the mixture was stirred at r.t. for 1h. The mixture was diluted with ethyl acetate (100 mL) and

[1] T. L. Hwang and A. J. Shaka, *J. Magn. Reson. Ser. A*, 1995, **A112**, 275.

[2] C. Dalvit, *J. Biomol. NMR*, 1998, **11**, 437.

washed with saturated aq. sodium chloride (3 x 50 mL). The organic phase, dried over sodium sulphate and evaporated to an oil, was purified by column chromatography (10 g Kieselgel 60H) eluted with *n*-esane/ethyl acetate/lutidine 90:9.5:0.5 (v/v) to give 3'-O-tert-butyldiphenylsilyl-propyl-1'-O-(2-cyanoethyl)-N,N'-diisopropylphosphoramidite as a white powder (9.4 mmol, 90% yield).

^{1}H-NMR (400 MHz; acetone-d_6): δ 7.71-7.70 (m, 4H, Ph), 7.48-7.42 (m, 6H, Ph), 3.88-3.76 (complex signal, 6H, 3 × CH_2O), 3.67-3.58 (complex signal, 2H, 2 × CH), 2.71 (t, J = 6.0 Hz, 2H, CH_2CN), 1.88 (quintuplet, J = 6.1 Hz, 2H, CH_2), 1.18 (d, J = 6.8 Hz, 6H, *iso*Pr), 1.15 (d, J = 6.8 Hz, 6H, ^{i}Pr), 1.04 (s, 9H, ^{t}Bu); ^{31}P-NMR (202 MHz; acetone-d_6): δ 150.4 (s); HRESIMS: *m/z* 537.2672 $[M+Na]^+$, $C_{28}H_{43}N_2NaO_3PSi$, requires 537.2678.

Synthesis of TEL-ODNs

The syntheses of TEL-ODNs **I-V** were performed using a solid-phase automated DNA synthesizer according to the previously described procedure (G. Oliviero, J. Amato, N. Borbone, S. D'Errico, A. Galeone, L. Mayol, S. Haider, O. Olubiyi, B. Hoorelbeke, J. Balzarini, and G. Piccialli, *Chem. Commun.*, 2010, **46**, 8971–8973). In the case of TEL-ODN **VI**, in the last coupling step the 3'-O-tert-butyldiphenylsilyl-propyl-1'-O-(2-cyanoethyl)-N,N'-diisopropylphosphoramidite was used at 50 mg/mL in acetonitrile solution. After automated synthesis, the oligonucleotides were detached from the support and deprotected by using concentrated aqueous ammonia at 55 °C for 17 h. The combined filtrates and washings were dried, redissolved in water, analyzed and purified by HPLC on an anion exchange column (Macherey-Nagel, 1000-8/46, 4.4x50 mm, 5 μm) using a linear gradient from 0 to 100% B in 30 min, flow rate = 1 mL/min and detection at 260 nm (buffer A: 20 mM NaH_2PO_4 aq. solution pH 7.0, containing 20% (v/v) CH_3CN; buffer B: 20 mM NaH_2PO_4 aq. solution pH 7.0, containing 1M NaCl and 20% (v/v) CH_3CN.

After HPLC purification, the oligonucleotide samples were desalted on a biogel column (BIORAD) eluted with H_2O/CH_3CH_2OH (9:1 v/v). The isolated oligomers were characterized by MALDI TOF mass:

I calculated: [M] = 9765.3; found $[M+H]^+$ = 9764

II calculated: [M] = 9829.3; found $[M+H]^+$ = 9830

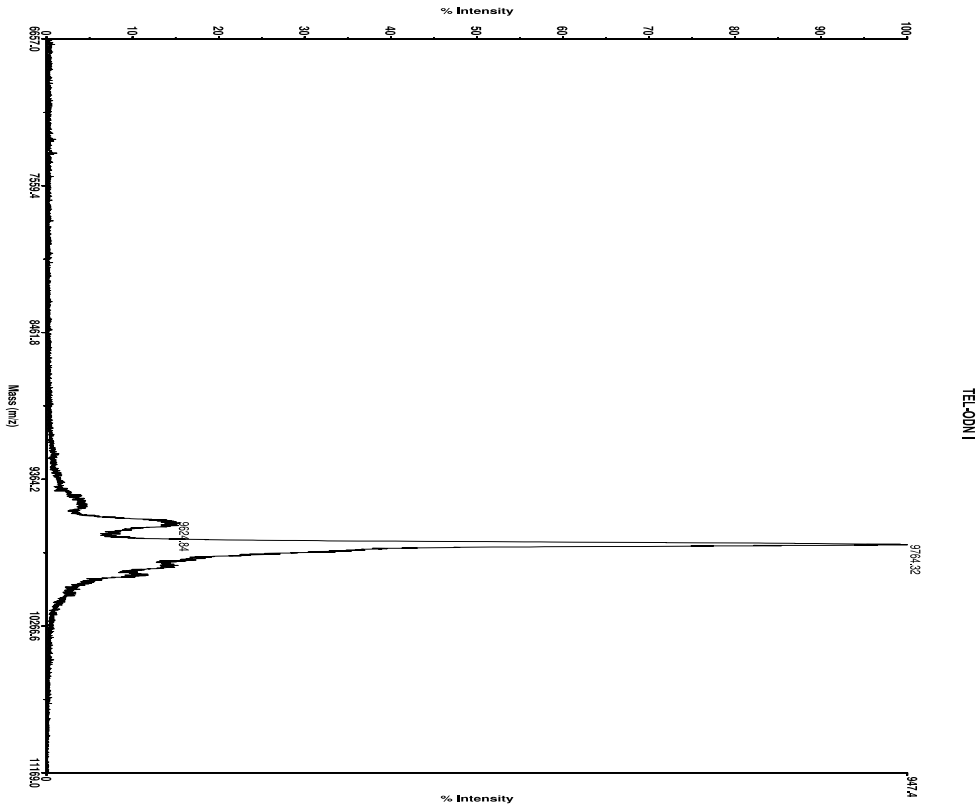
III calculated: [M] = 9729.3; found $[M+H]^+$ = 9730

IV calculated: [M] = 9669.3; found $[M+H]^+$ = 9669

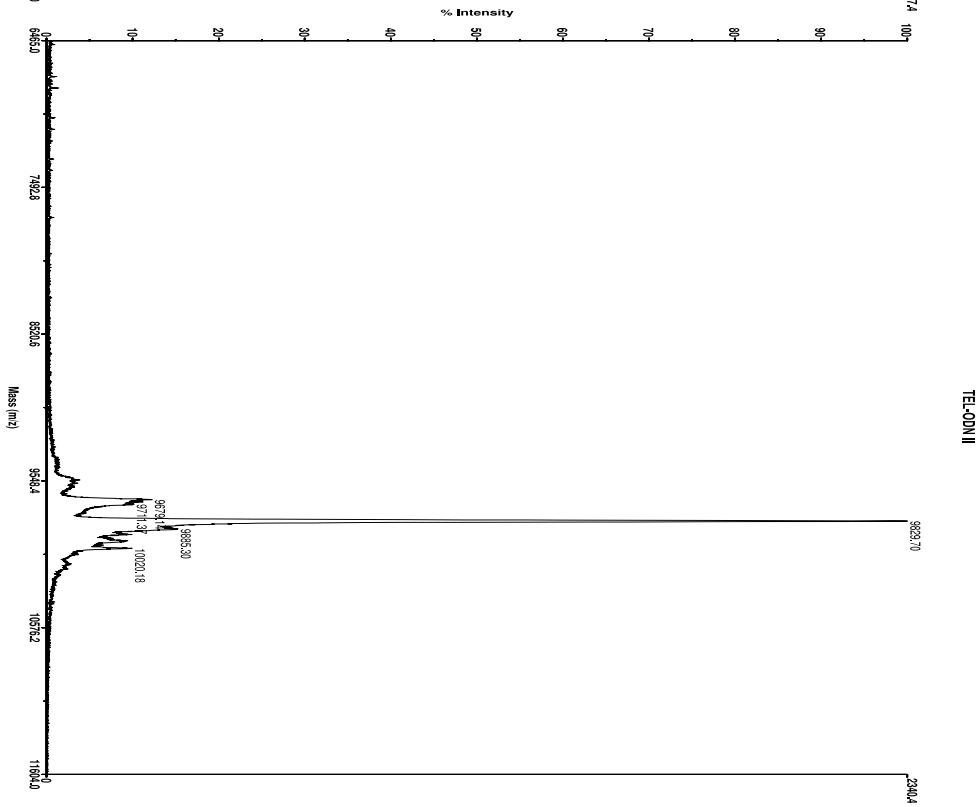
V calculated: [M] = 9729.3; found $[M+H]^+$ = 9729

VI calculated: [M] = 10316.1; found $[M+H]^+$ = 10318

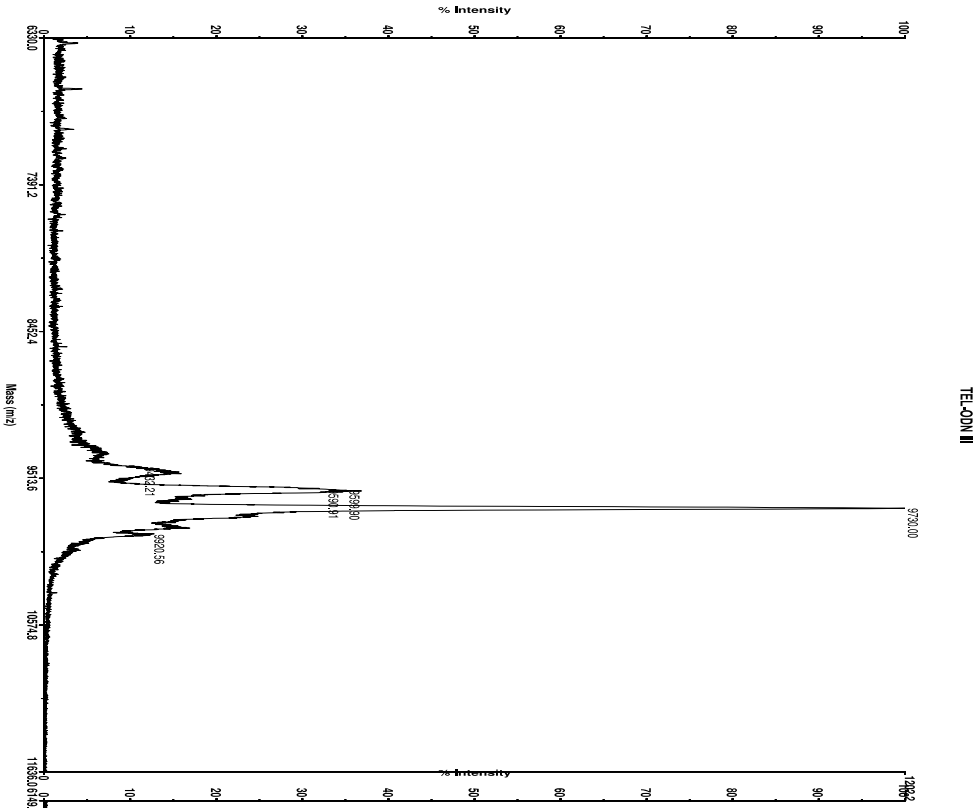
A


% Intensity
Mass (m/z)
TEL-ODN I
9624.84
9764.32
947.4

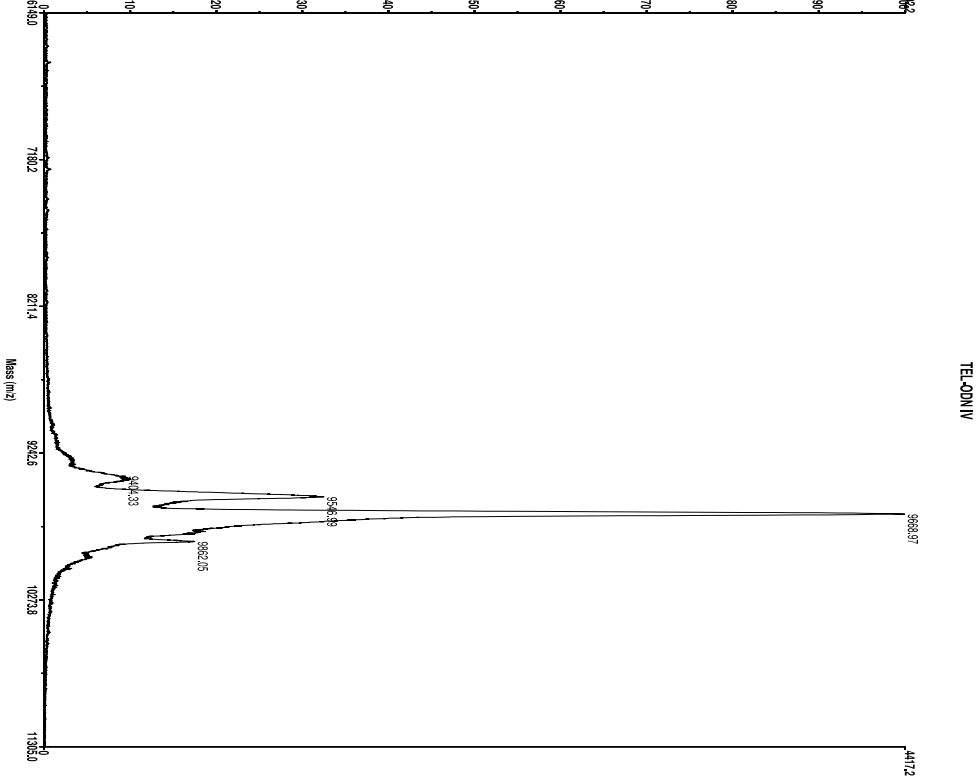
B


% Intensity
Mass (m/z)
TEL-ODN II
9673.14
9711.37
9885.20
10020.18
9823.70
2340.4

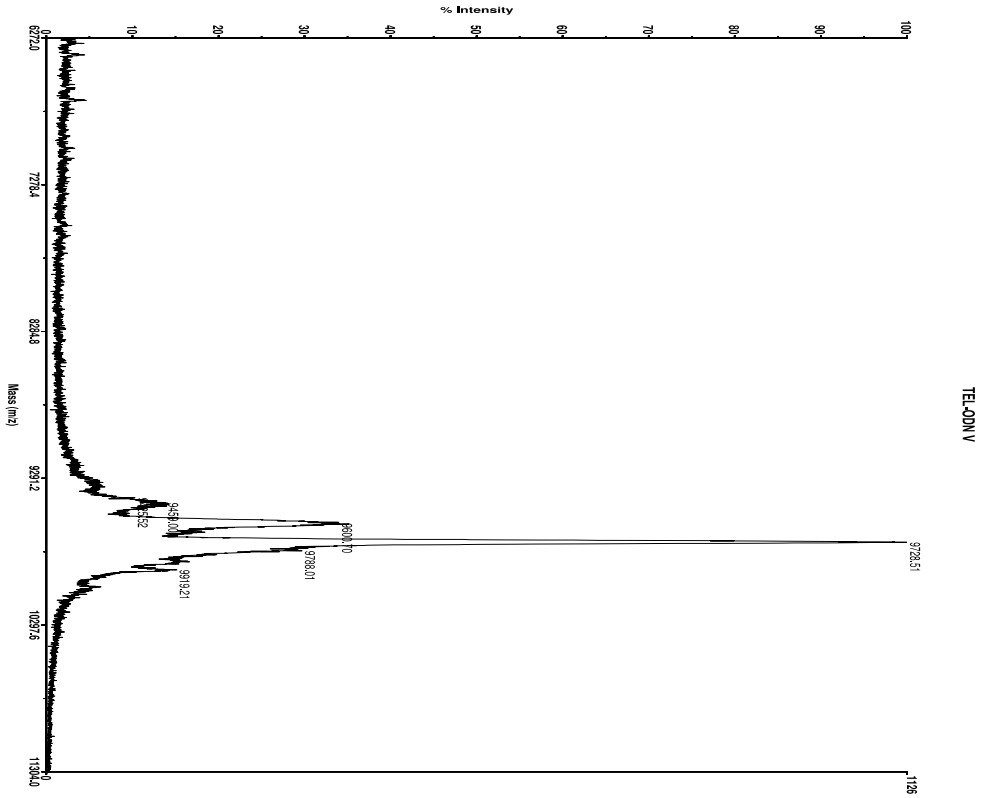
C


% Intensity
Mass (m/z)
TEL-ODN III
9730.00
9589.80
9590.01
9020.56
9433.21

D


% Intensity
Mass (m/z)
TEL-ODN IV
9668.97
9566.89
9420.33
9862.05

E



F

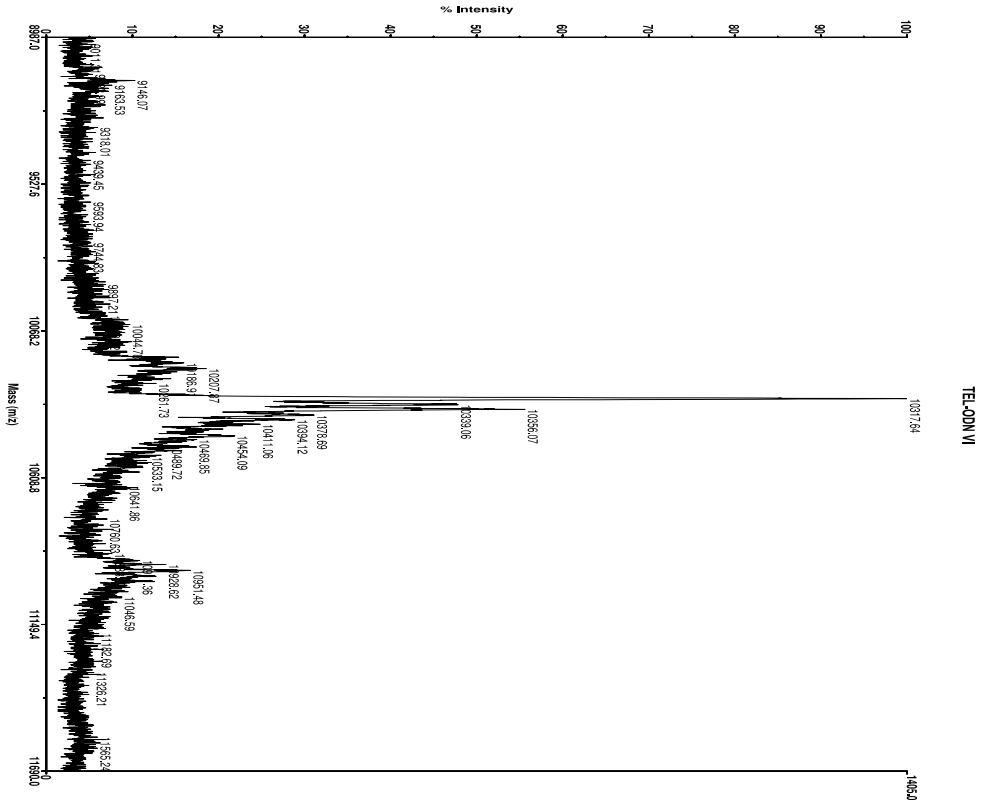


Fig. S1: MALDI-TOF spectra (positive mode) of TEL-ODNs **I-VI** (A-F).

Preparation of quadruple helices (annealing).

TEL-ODN samples were dissolved in 100 mM potassium buffer (10 mM KH_2PO_4 aq. solution containing 90 mM KCl, pH 7.0) and the corresponding quadruplexes were formed by heating at 90 °C for 5 min and slowly cooling at room temperature for 12 h. Samples were stored at 4°C for 24 h before measurements. TEL-ODNs concentrations were determined in water by measuring the absorbance at 260 nm at 90 °C, using the nearest-neighbour calculated molar extinction coefficient of d[$^{5'}$TGGGXG$^{3'}$] (for **I-IV** and **VI)** or d[$^{5'}$TGGGGT$^{3'}$] (for **V**) multiplied by 4 (the number of strands in each TEL-ODN complex). A contribute of 8658 $cm^{-1}M^{-1}$ was added for each TBDPS group.[3]

I: calculated ε = 284632 $cm^{-1}M^{-1}$

II: calculated ε = 272232 $cm^{-1}M^{-1}$

III: calculated ε = 266232 $cm^{-1}M^{-1}$

IV: calculated ε = 258232 $cm^{-1}M^{-1}$

V: calculated ε = 265832 $cm^{-1}M^{-1}$

VI: calculated ε = 284632 $cm^{-1}M^{-1}$

[3] J. D'Onofrio, L. Petraccone, E. Erra, L. Martino, G. Di Fabio, L. De Napoli, C. Giancola and D. Montesarchio, *Bioconjugate Chem.*, 2007, **18**, 1194-1204

CD spectroscopy measurements

TEL-ODN samples (2×10^{-5} M) were annealed as described above and stored at 4°C overnight before measurements. CD spectra were recorded in the 360-200 nm wavelength range as an average of 5 scans (100 nm/min, 1 s response time, 1 nm bandwidth) and normalized by subtraction of the background scan containing only the buffer. Thermal denaturation experiments (Fig. S3) were carried out in the temperature range of 5-90 °C by monitoring CD values at 264 nm at a heating rate of 0.5 °C/min.

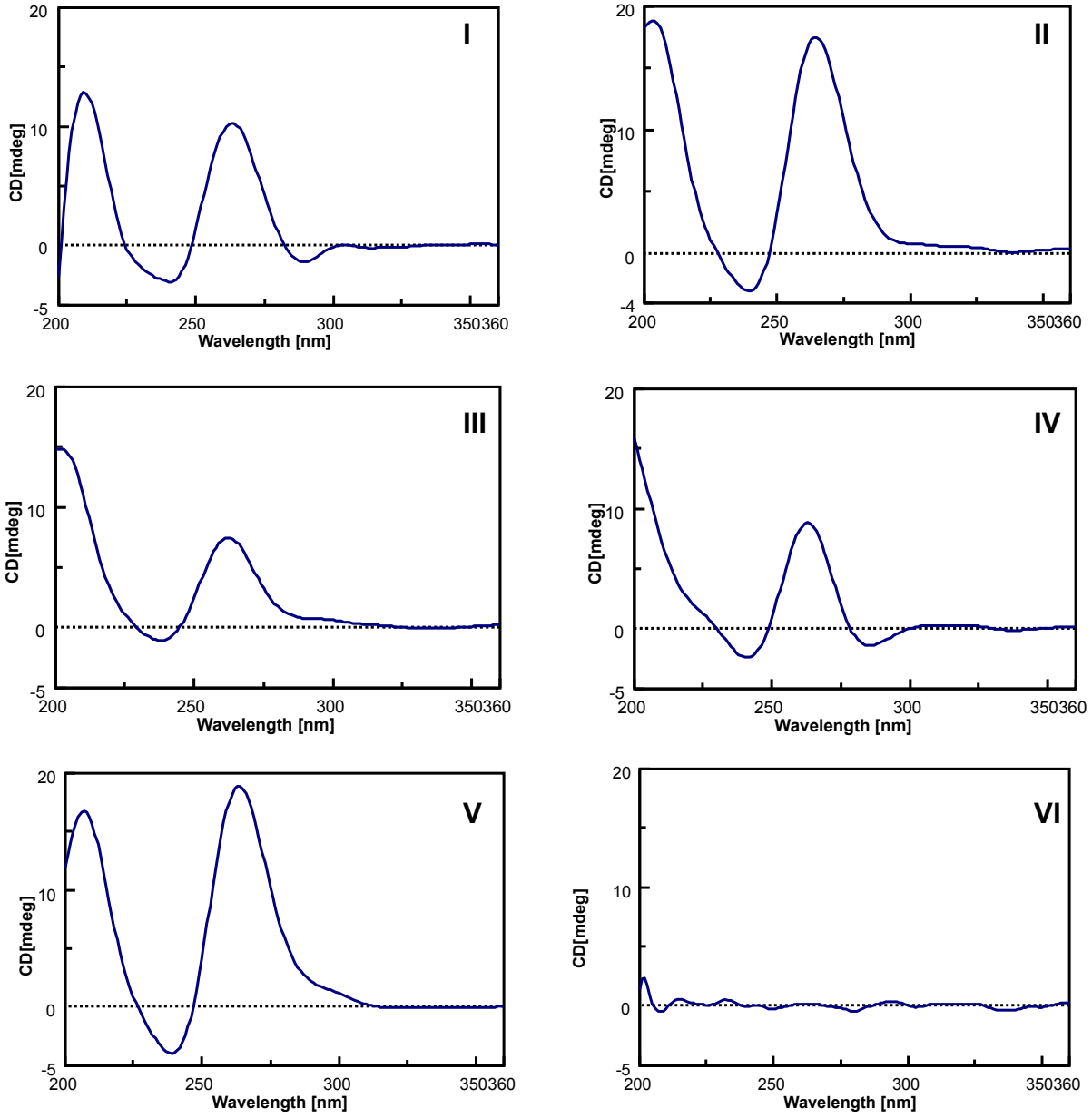


Fig. S2: CD profiles of **I-VI** in 100 mM K^+ (left column) buffer at 25°C.

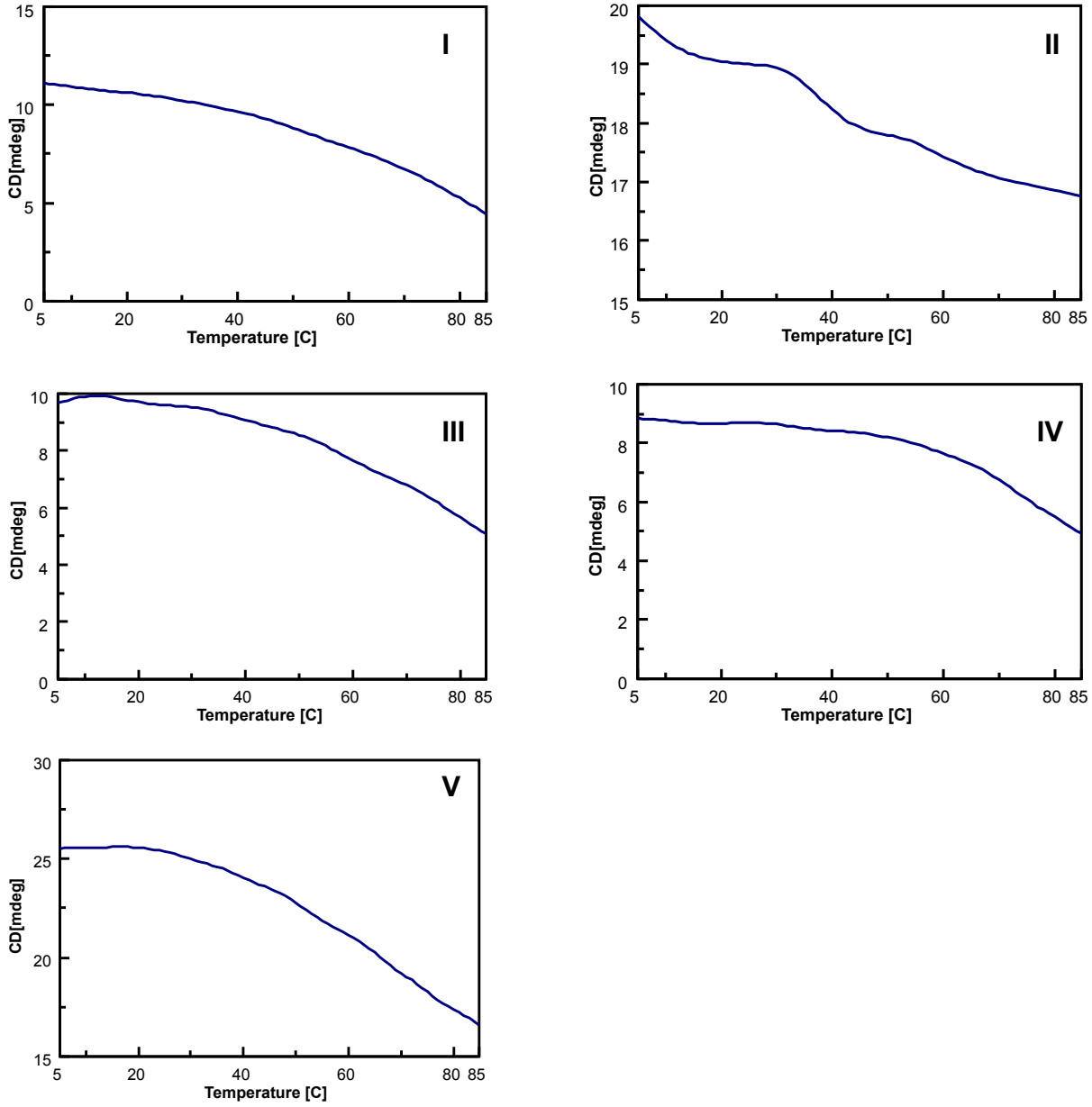


Fig. S3: CD-melting curves of **I-V** recorded at 264 nm in 100 mM K^+ buffer.

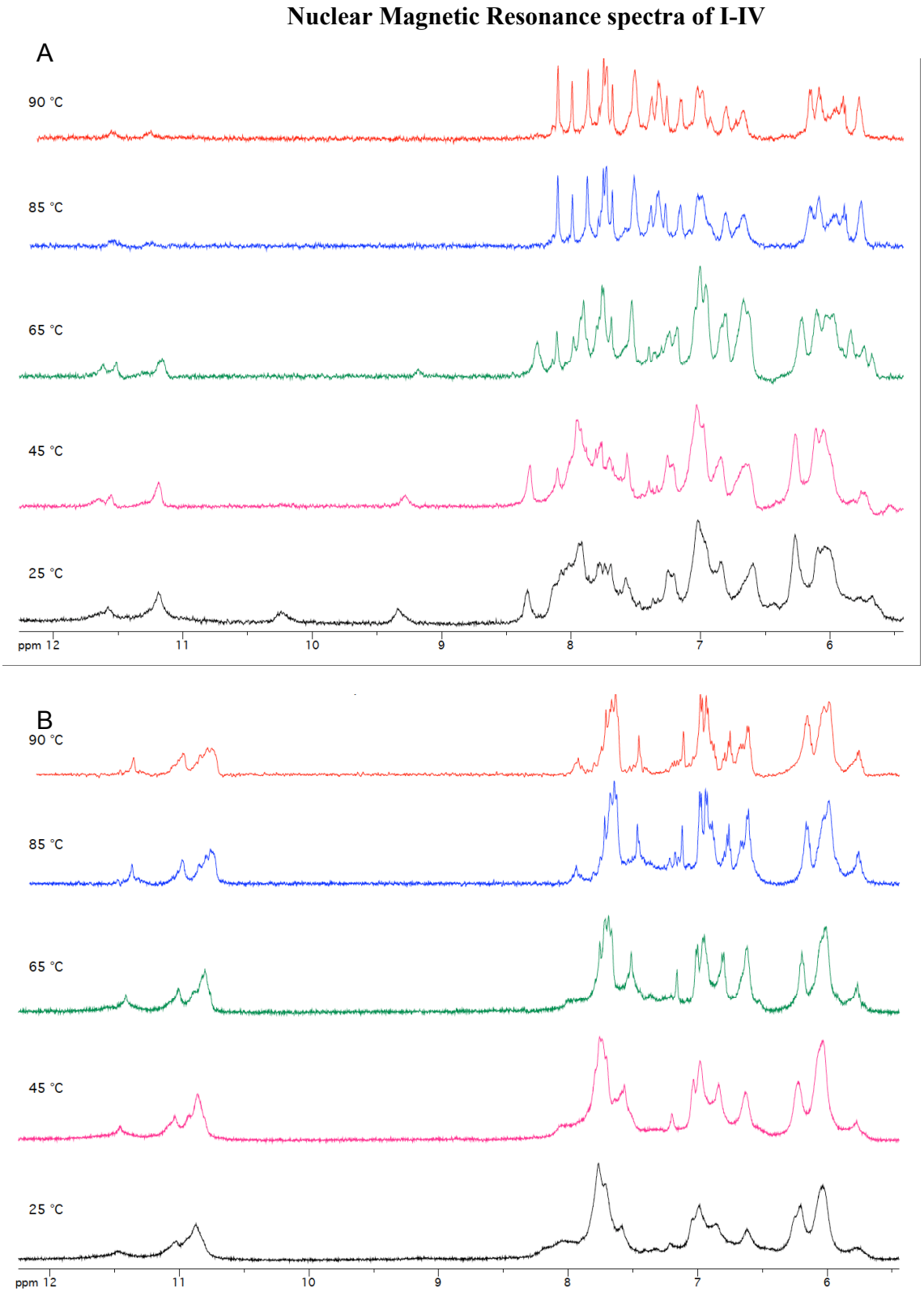


Fig. S4 ^{1}H NMR spectra of **I** (A) and **II** (B) recorded in the temperature range 25-90 °C (100 mM K^{+} buffer; H_2O/D_2O 9:1).

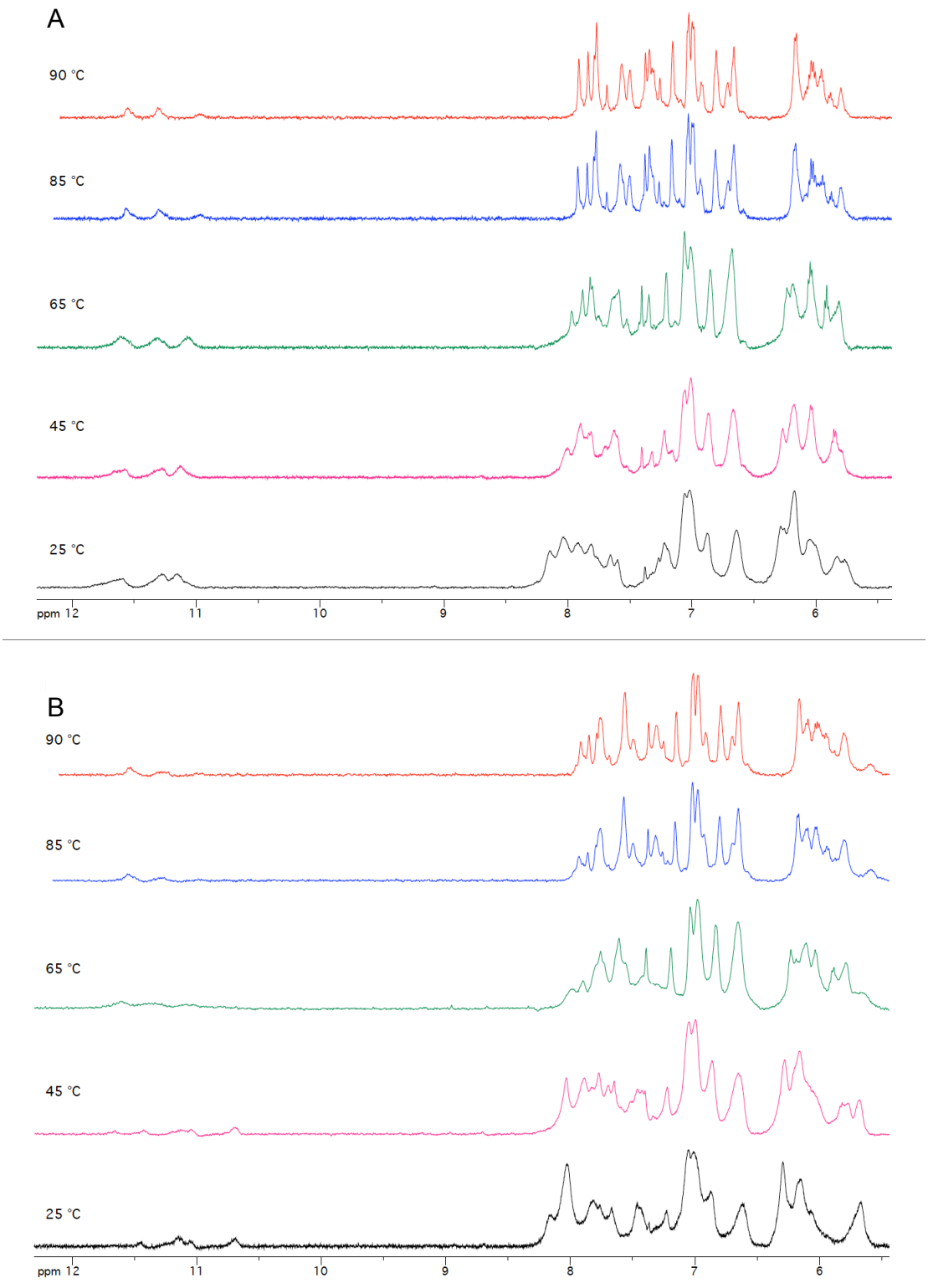


Fig. S5 ^{1}H NMR spectra of **III** (A) and **IV** (B) recorded in the temperature range 25-90 °C (100 mM K^{+} buffer; H_2O/D_2O 9:1).

Biological evaluation assays

The methodology of the anti-HIV assays was as follows: ~ 3 x 10^5 human CD_4 T-lymphocyte (CEM) cells were infected with 100 $CCID_{50}$ of HIV(III_B) or HIV-2(ROD)/ml and seeded in 200 µL wells of a 96-well microtiter plate containing appropriate dilutions of the test compounds. After 4 days of incubation at 37°C, HIV-induced syncytia formation in the CEM cell cultures was examined microscopically. Data represent the mean (± SD) of 3 independent experiments.

Surface plasmon resonance (SPR) analysis

Recombinant gp120 protein from HIV-1(III_B) (ImmunoDiagnostics Inc., Woburn, MA) (produced by CHO cell cultures) was covalently immobilized on a CM5 sensor chip in 10 mM sodium acetate, pH 4.0, using standard amine coupling chemistry. The chip density was 2345 RU (~19.5 fmol of gp120). A reference flow cell was used as a control for non-specific binding and refractive index changes. The experiment was performed at 25°C on a Biacore T200 instrument (GE Healthcare, Uppsala, Sweden). The test compounds were serially diluted in HBS-P (10 mM HEPES, 150 mM NaCl and 0.05% surfactant P20; pH 7.4), covering a concentration range between 1.17 and 4.69 µM. Samples were injected for 120 sec at a flow rate of 30 µl/min and the dissociation was followed for 180 sec. One duplicate sample and several buffer blanks were used as a control and as double referencing, respectively. The CM5 sensor chip surface was regenerated with 1 injection of 50 mM NaOH. The experimental data were fit using the 1:1 binding model (Biacore T200 Evaluation software 1.0) to determine the binding kinetics.

Surface plasmon resonance (SPR) analysis

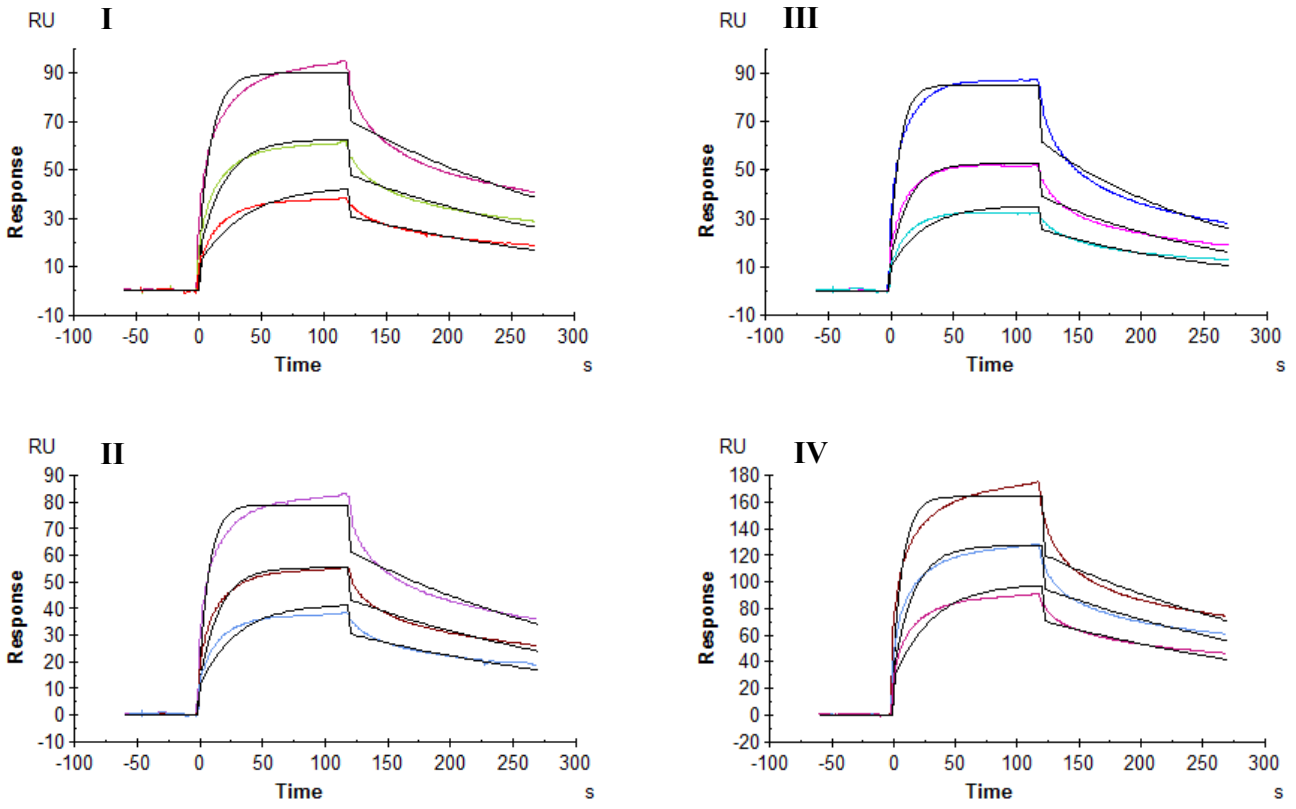


Fig. S6 Kinetic analysis of interactions of compounds **I-IV** with recombinant HIV-1(IIIB) gp120 expressed in CHO cell cultures using SPR technology. The experimental data (coloured curves) were fit using the 1 : 1 binding model (black lines) to determine the kinetic parameters. The data are a representative example of two independent experiments.

Docking protocol

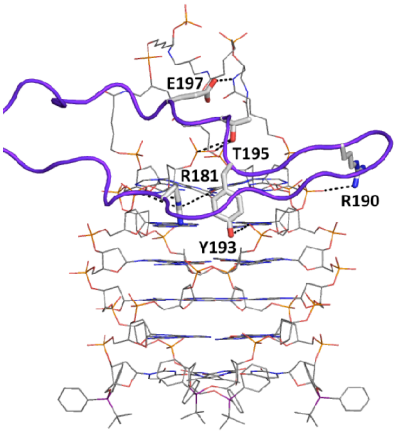
The molecular models of TEL-ODNs were generated using the InsightII suite program (Accelrys, Inc), by employing the Biopolymer module, to make the quadruplexes, and the Builder module, to add the modified linkers at 3'- and 5'-ends. The molecular models were parameterized according to the AMBER force field and then brought to their energetically minimized structures with Discovery3 module. The minimizations were performed using 1000 cycles of steepest descent method followed by conjugate gradient calculation until convergence at 0.1 rms was reached. A distance-dependent macroscopic dielectric constant of 4.0 and an infinite cut-off for non-bonded interactions were used to partially compensate for the lack of solvent. The Protein Data Bank was used to download the HIV gp120 protein (PDB ID 2B4C).

All dockings between the V3 loop in gp120 and TEL-ODN were performed using the ZDOCK server (zdock.umassmed.edu), that evaluates Pairwise shape complementarity, desolvation and electrostatic energies using Fast Fourier Transform algorithm to search multiple binding possibilities for receptor and ligand. The output from the program is a list of transformation values for the docks, from which models can be generated. For each docking, a total of 2000 transformations were calculated, ranked based on internal scoring functions and presented through a global energy term. The best solution from each dock was chosen and further minimised to relieve any steric clashes that might have arisen in docking. No atoms were restrained during the minimization process. The minimized models were then visualized using the Pymol software and the electrostatic interactions between protein and aptamer were defined.

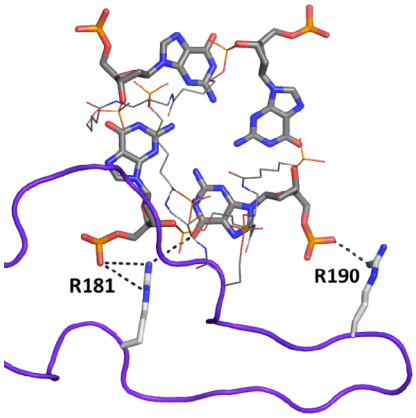
Table S1 Table of interactions among the residues of V3 loop and **I** (◆), **II** (⊙), **III** (❖), **IV** (O) and **V** (■).

ǂ	G6-O6	X5-N3	G6-OP3'	G6-O3'/O4'/O5'	T6-OP3'	X5-OP3'	X5-O3'/O4'/O5'	TEL-O	TEL-N
Arg181/NH	⊙		O◆ǂ	⊙Oǂ	ǂ	ǂ	⊙❖ǂ	ǂ	ǂ
Arg181/NH2	O◆❖	⊙	◆ǂ	ǂ	■■ǂ	ǂ	⊙O❖❖ǂ		
Lys182/N	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	■■ǂ	ǂ
Lys182/O	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	■ǂ
Ser183/OH	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	■ǂ	ǂ
Ile184/O	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	■ǂ
Ile186/O	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	■ǂ
Arg190/NH2			◆ǂ	ǂ	ǂ	ǂ	ǂ	⊙O❖ǂ	ǂ
Tyr193/N	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	⊙O❖ǂ	ǂ
Tyr193/OH	O		◆	ǂ	ǂ	⊙ǂ	ǂ	ǂ	ǂ
Thr194/N	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	⊙Oǂ	ǂ
Thr195/N	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	⊙ǂ	ǂ
Thr195/OH	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	⊙⊙O◆◆❖ǂ	ǂ
Glu197/OH	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	⊙⊙OO❖❖ǂ	◆ǂ
Ile198/N	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	⊙ǂ	ǂ

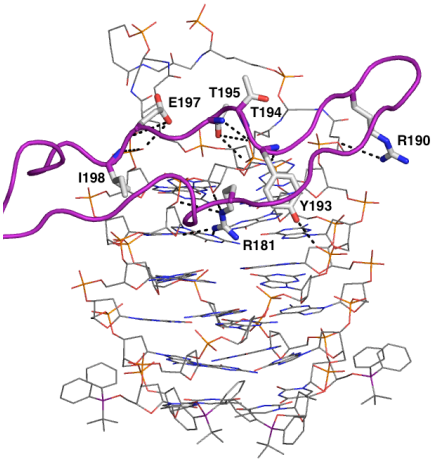
ǂ



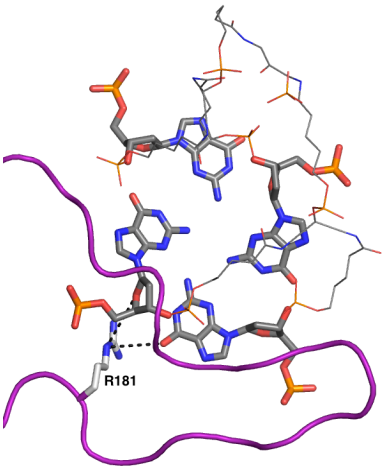
(TBDPS-5'TGGGAG3')$_4$-TEL (**I**)



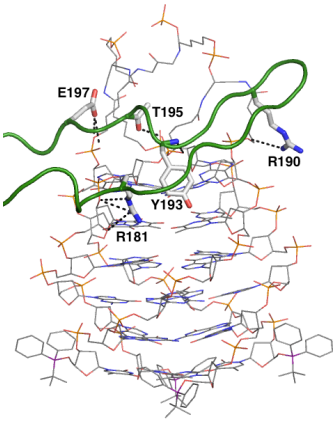
G-tetrad at 3' terminus interactions



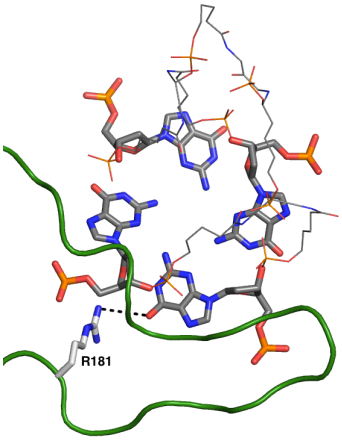
(TBDPS-5'TGGGGG3')$_4$-TEL (**II**)



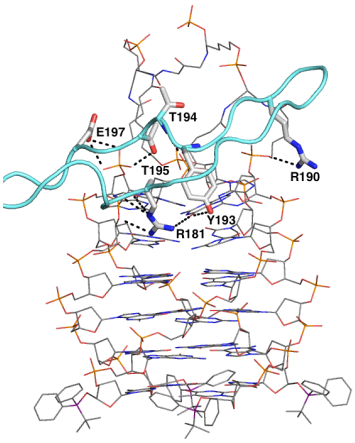
G-tetrad at 3' terminus interactions



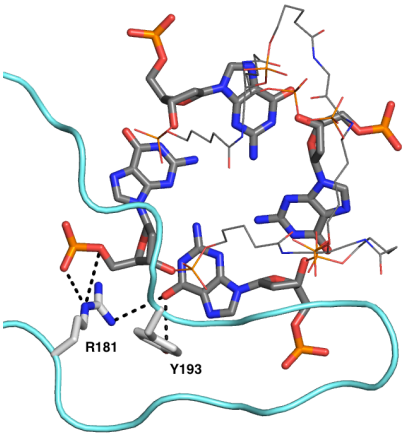
(TBDPS-5'TGGGTG3')$_4$-TEL (**III**)



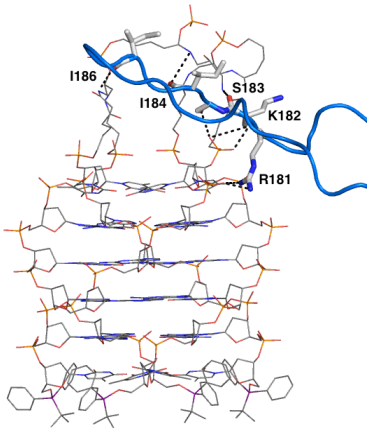
G-tetrad at 3' terminus interactions



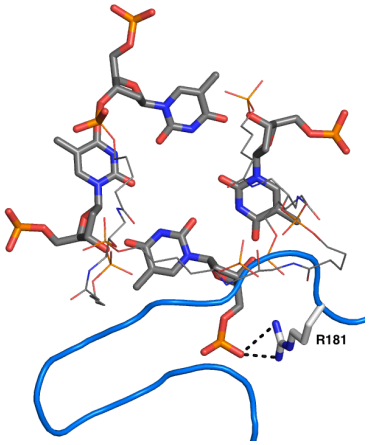
(TBDPS-5'TGGGCG3')$_4$-TEL (**IV**)



G-tetrad at 3' terminus interactions



(TBDPS-5'TGGGGT3')$_4$-TEL (**V**)



T-tetrad at 3' terminus interactions

Fig. S7 Molecular models of interactions between aptamers **I-V** and the V3 loop of gp120. The aptamers are drawn as lines, the V3 loop as cartoon and hydrogen bonding interactions between the V3 loop and the aptamers are drawn as dashed lines.