Supporting Information for:

Deciphering structure-activity relationships in a series of Tat/TAR inhibitors

Lise Pascale,^[a] Alejandro López González,^[b] Audrey Di Giorgio,^[a] Marc Gaysinski,^[a] Jordi Teixido Closa,^[b] Roger Estrada Tejedor,^[b] Stéphane Azoulay,^[a] and Nadia Patino*^[a]

a] Institut de Chimie de Nice UMR7272, Université Nice Sophia Antipolis, 06108 Nice Cedex, France. [b] Molecular Design Lab., IQS School of Engineering, Universitat Ramon Llull, 08017 Barcelona, Spain.

***Corresponding author:** patino@unice.fr; Tel: +33 (0)4 92 07 61 46; Fax: +33 (0)4 92 07 61 51

1 - CHEMISTRY

Characterization of PAA pentamers

Table S1: HPLC and HRMS data for individual penta-PAA

Penta-AA	HPLC		HRMS m/z [M+H] ⁺		
	Retention time (min)*	Purity* * (%)	Molecular Formula	Calculated	experimental
FFFRF Ia	11.6	68	C ₇₂ H ₁₀₈ N ₂₀ O ₁ 2	1445.85284	1445.85280
KFFRF Ib	10.6	92	C ₆₉ H ₁₁₂ N ₂₁ O ₁ 2	1426.87938	1426.87876
RFFRF Ic	10.7	76	C ₆₉ H ₁₁₂ N ₂₃ O ₁ 2	1454.88553	1454.88476
FKFRF IIa	10.5	89	C ₆₉ H ₁₁₂ N ₂₁ O ₁ 2	1426.87938	1426.87775
KKFRF IIb	9.7	82	C ₆₆ H ₁₁₅ N ₂₂ O ₁ 2	1407.90593	1407.89865
RKFRF IIc	9.8	90	C ₆₆ H ₁₁₅ N ₂₄ O ₁ 2	1435.91208	1435.90468
FRFRF IIIa	10.7	88	C ₆₉ H ₁₁₂ N ₂₃ O ₁ 2	1454.88553	1454.88222
KRFRF IIIb	9.9	86	C ₆₆ H ₁₁₅ N ₂₄ O ₁ 2	1435.91208	1435.91166
RRFRF IIIc	9.8	89	C ₆₆ H ₁₁₄ N ₂₆ O ₁ 2	1463.91823	1463.91788

*HPLC conditions: C₁₈, H₂O (0.1% TFA)/Acetonitrile (0.1% TFA): 90/10 to 0/100 for 30 min ; flow rate: 1 mL/min.

**HPLC purity of the crude product. All compounds were purified before testing.

HPLC analyses after semi-preparative HPLC purification

- PAA Ia

Sample Information

SampleName	-C-PentaL-FRFFF n 22 somr	Sample Type	Unknown
Vial	1	Date Acquired	17/02/11 13:43:13
Injection	2	Acq Method Set	Analytique Monomere
Injection Volume	20,00 ul	Processing Method	Analytique Monomere
Channel	996	Date Processed	21/02/11 18:35:47
Run Time	47,0 Minutes

	Retention Time (min)	Area	% Area	Height
1	11,556	9816948	100,00	1651733

Instrument Method: Analytique Monomere

Stored: 19/10/10 16:09:01

W600 Gradient Table

	Time	Flow	%A	%B	%C	%D	Curve
1		1,00	0,0	0,0	90,0	10,0	
2	30,00	1,00	0,0	0,0	0,0	100,0	6

	Time	Flow	%A	%B	%C	%D	Curve
3	35,00	1,00	0,0	0,0	0,0	100,0	6
4	36,00	1,00	0,0	0,0	90,0	10,0	6

- PAA Ib

Sample Information

SampleName L-Pental-FRFKF n 10 somm
 Vial 1
 Injection 3
 Injection Volume 20,00 ul
 Channel 996
 Run Time 47,0 Minutes

Sample Type Unknown
 Date Acquired 18/03/11 16:40:28
 Acq Method Set Analytique Monomere
 Processing Method Analytique Monomere
 Date Processed 18/03/11 18:22:15

	Retention Time (min)	Area	% Area	Height
1	10,507	8923493	100,00	1316224

Instrument Method: Analytique Monomere

Stored: 19/10/10 16:09:01

W600 Gradient Table

	Time	Flow	%A	%B	%C	%D	Curve
1		1,00	0,0	0,0	90,0	10,0	
2	30,00	1,00	0,0	0,0	0,0	100,0	6

	Time	Flow	%A	%B	%C	%D	Curve
3	35,00	1,00	0,0	0,0	0,0	100,0	6
4	36,00	1,00	0,0	0,0	90,0	10,0	6

- PAA Ic

Sample Information

SampleName	-C-PentaL-FRFRF n 34 somr	Sample Type	Unknown
Vial	1	Date Acquired	11/05/11 14:59:24
Injection	2	Acq Method Set	Analytique Monomere
Injection Volume	20,00 ul	Processing Method	Analytique Monomere
Channel	996	Date Processed	11/05/11 15:59:57
Run Time	47,0 Minutes		

Chromatogram

	Retention Time (min)	Area	% Area	Height
1	10,681	13779051	100,00	1988507

Instrument Method: Analytique Monomere

Stored: 02/05/11 11:58:06

W600 Gradient Table

	Time	Flow	%A	%B	%C	%D	Curve
1		1,00	0,0	0,0	90,0	10,0	
2	30,00	1,00	0,0	0,0	0,0	100,0	6

	Time	Flow	%A	%B	%C	%D	Curve
3	35,00	1,00	0,0	0,0	0,0	100,0	6
4	36,00	1,00	0,0	0,0	90,0	10,0	6

- PAA IIa

Sample Information

SampleName -C-PentaL-FRFFK n 23 somr
 Vial 1
 Injection 9
 Injection Volume 20,00 ul
 Channel 996
 Run Time 47,0 Minutes

Sample Type Unknown
 Date Acquired 21/02/11 17:40:59
 Acq Method Set Analytique Monomere
 Processing Method Analytique Monomere
 Date Processed 21/02/11 18:34:51

	Retention Time (min)	Area	% Area	Height
1	10,613	12069331	100,00	1350270

Instrument Method: Analytique Monomere

Stored: 19/10/10 16:09:01

W600 Gradient Table

	Time	Flow	%A	%B	%C	%D	Curve
1		1,00	0,0	0,0	90,0	10,0	
2	30,00	1,00	0,0	0,0	0,0	100,0	6

	Time	Flow	%A	%B	%C	%D	Curve
3	35,00	1,00	0,0	0,0	0,0	100,0	6
4	36,00	1,00	0,0	0,0	90,0	10,0	6

- PAA IIb

Sample Information

SampleName L-PentaC-FRFFK n 8 somm
 Vial 1
 Injection 1
 Injection Volume 20,00 ul
 Channel 996
 Run Time 47,0 Minutes

Sample Type Unknown
 Date Acquired 18/03/11 11:21:41
 Acq Method Set Analytique Monomere
 Processing Method Analytique Monomere
 Date Processed 18/03/11 12:15:42

	Retention Time (min)	Area	% Area	Height
1	9,902	10812600	100,00	874692

Instrument Method: Analytique Monomere

Stored: 19/10/10 16:09:01

W600 Gradient Table

	Time	Flow	%A	%B	%C	%D	Curve
1		1,00	0,0	0,0	90,0	10,0	
2	30,00	1,00	0,0	0,0	0,0	100,0	6

	Time	Flow	%A	%B	%C	%D	Curve
3	35,00	1,00	0,0	0,0	0,0	100,0	6
4	36,00	1,00	0,0	0,0	90,0	10,0	6

- PAA IIc

Sample Information

SampleName	-C-PentaC-FRFRK n 32 som	Sample Type	Unknown
Vial	1	Date Acquired	10/05/11 16:20:09
Injection	4	Acq Method Set	Analytique Monomere
Injection Volume	20,00 ul	Processing Method	Analytique Monomere
Channel	996	Date Processed	10/05/11 17:22:09
Run Time	47,0 Minutes		

Chromatogram

	Retention Time (min)	Area	% Area	Height
1	9,749	13550520	100,00	1609957

Instrument Method: Analytique Monomere

Stored: 02/05/11 11:58:06

W600 Gradient Table

	Time	Flow	%A	%B	%C	%D	Curve
1		1,00	0,0	0,0	90,0	10,0	
2	30,00	1,00	0,0	0,0	0,0	100,0	6
3	35,00	1,00	0,0	0,0	0,0	100,0	6
4	36,00	1,00	0,0	0,0	90,0	10,0	6

- PAA IIIa

Sample Information

SampleName	-C-PentaL-FRFFR n 24 somr	Sample Type	Unknown
Vial	1	Date Acquired	21/02/11 20:02:38
Injection	12	Acq Method Set	Analytique Monomere
Injection Volume	20,00 ul	Processing Method	Analytique Monomere
Channel	996	Date Processed	21/02/11 21:04:28
Run Time	47,0 Minutes

	Retention Time (min)	Area	% Area	Height
1	10,735	10433468	100,00	1786648

Instrument Method: Analytique Monomere

Stored: 19/10/10 16:09:01

W600 Gradient Table

	Time	Flow	%A	%B	%C	%D	Curve
1		1,00	0,0	0,0	90,0	10,0	
2	30,00	1,00	0,0	0,0	0,0	100,0	6

	Time	Flow	%A	%B	%C	%D	Curve
3	35,00	1,00	0,0	0,0	0,0	100,0	6
4	36,00	1,00	0,0	0,0	90,0	10,0	6

- PAA IIIb

Sample Information

SampleName	L-PentaC-FRFR n 9 somm	Sample Type	Unknown
Vial	1	Date Acquired	18/03/11 15:42:49
Injection	2	Acq Method Set	Analytique Monomere
Injection Volume	20,00 ul	Processing Method	Analytique Monomere
Channel	996	Date Processed	18/03/11 16:38:22
Run Time	47,0 Minutes		

Chromatogram

	Retention Time (min)	Area	% Area	Height
1	9,700	10932898	100,00	662611

Instrument Method: Analytique Monomere

Stored: 19/10/10 16:09:01

W600 Gradient Table

	Time	Flow	%A	%B	%C	%D	Curve
1		1,00	0,0	0,0	90,0	10,0	
2	30,00	1,00	0,0	0,0	0,0	100,0	6

	Time	Flow	%A	%B	%C	%D	Curve
3	35,00	1,00	0,0	0,0	0,0	100,0	6
4	38,00	1,00	0,0	0,0	90,0	10,0	6

- PAA IIIc

Sample Information

SampleName	-C-PentaL-FRFR n 30 somi	Sample Type	Unknown
Vial	1	Date Acquired	10/05/11 15:27:07
Injection	3	Acq Method Set	Analytique Monomere
Injection Volume	20,00 ul	Processing Method	Analytique Monomere
Channel	996	Date Processed	10/05/11 17:36:04
Run Time	47,0 Minutes

	Retention Time (min)	Area	% Area	Height
1	9,759	12102873	100,00	1752375

Instrument Method: Analytique Monomere

Stored: 02/05/11 11:58:06

W600 Gradient Table

	Time	Flow	%A	%B	%C	%D	Curve
1		1,00	0,0	0,0	90,0	10,0	
2	30,00	1,00	0,0	0,0	0,0	100,0	6

	Time	Flow	%A	%B	%C	%D	Curve
3	35,00	1,00	0,0	0,0	0,0	100,0	6
4	36,00	1,00	0,0	0,0	90,0	10,0	6

HRMS analyses

• PAA Ia

C:\xcalibur\data\310511NP22

5/30/2011 4:40:18 PM

310511NP22 #48 RT: 1.11 AV: 1 NL: 3.19E5

F: FTMS + p ESI d Full ms2 723.43@cid21.00 [185.00-1460.00]

310511NP22_XT_00001_MHp_#1 RT: 1.00 AV: 1 NL: 3.84E7

T: FTMS + p ESI Full ms [50.00-2000.00]

- **PAA Ib**

C:\Xcalibur\data\310511NP23

5/30/2011 4:44:26 PM

310511NP23 #61 RT: 1.39 AV: 1 NL: 1.74E5

F: FTMS + p ESI d Full ms2 713.94@cid21.00 [185.00-1440.00]

310511NP23_XT_00001_MHp_#1 RT: 1.00 AV: 1 NL: 3.86E7

T: FTMS + p ESI Full ms [50.00-2000.00]

- **PAA Ic**

C:\Xcalibur\data\310511NP24

5/30/2011 4:48:32 PM

310511NP24 #62 RT: 1.43 AV: 1 NL: 1.68E5

F: FTMS + p ESI d Full ms2 727.94@cid21.00 [190.00-1470.00]

310511NP24_XT_00001_MHp_#1 RT: 1.00 AV: 1 NL: 3.83E7

T: FTMS + p ESI Full ms [50.00-2000.00]

- **PAA IIa**

C:\Xcalibur\data\310511NP28

5/30/2011 5:05:08 PM

310511NP28 #61 RT: 1.41 AV: 1 NL: 1.11E5

F: FTMS + p ESI d Full ms2 713.94@cid21.00 [185.00-1440.00]

310511NP28_XT_00001_MHp_#1 RT: 1.00 AV: 1 NL: 3.98E7

T: FTMS + p ESI Full ms [50.00-2000.00]

- **PAA IIb**

C:\Xcalibur\data\310511NP29

5/30/2011 5:09:15 PM

310511NP29 #58 RT: 1.36 AV: 1 NL: 1.75E5

F: FTMS + p ESI d Full ms2 704.45@cid21.00 [180.00-1420.00]

310511NP29_XT_00001_MHp_110530180639 #1 RT: 1.00 AV: 1 NL: 1.32E7

T: FTMS + p ESI Full ms [50.00-2000.00]

- **PAA IIc**

C:\Xcalibur\data\310511NP30

5/30/2011 5:13:24 PM

310511NP30 #59 RT: 1.35 AV: 1 NL: 1.92E5

F: FTMS + p ESI d Full ms2 718.46@cid21.00 [185.00-1450.00]

310511NP30_XT_00001_MHp__110530181103 #1 RT: 1.00 AV: 1 NL: 1.58E7

T: FTMS + p ESI Full ms [50.00-2000.00]

- **PAA IIIa-MS2 on 4+ m/z 364.48**

C:\Xcalibur\data\310511NP34

5/30/2011 5:29:49 PM

310511NP34 #8 RT: 0.16 AV: 1 NL: 3.78E5

F: FTMS + p ESI d Full ms2 364.48@cid21.00 [90.00-1470.00]

310511NP34_XT_00001_MHp_#1 RT: 1.00 AV: 1 NL: 4.89E4

T: FTMS + p ESI Full ms [50.00-2000.00]

- **PAA IIIb**

C:\Xcalibur\data\310511NP35

5/30/2011 5:33:56 PM

310511NP35 #64 RT: 1.46 AV: 1 NL: 1.20E5

F: FTMS + p ESI d Full ms2 718.46@cid21.00 [185.00-1450.00]

310511NP35_XT_00001_MHp_#1 RT: 1.00 AV: 1 NL: 2.06E7

T: FTMS + p ESI Full ms [50.00-2000.00]

- **PAA IIIc**

C:\Xcalibur\data\310511NP36

5/30/2011 5:38:03 PM

310511NP36 #32 RT: 0.70 AV: 1 NL: 9.77E4

F: FTMS + p ESI d Full ms2 732.46@cid21.00 [190.00-1475.00]

310511NP36_XT_00001_MHp_#1 RT: 1.00 AV: 1 NL: 2.95E7

T: FTMS + p ESI Full ms [50.00-2000.00]

2 - TAR BINDING STUDIES

2.1 - NMR experiments : material and methods

High resolution NMR experiments were recorded 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 (4.71 ppm at 13°C and 4.70 ppm 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 μ M) was performed in the appropriate buffer 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.

¹H NMR imino proton spectra were recorded in a H₂O/D₂O (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 μ s gradient recovery delay (D16) and strengths set to 8.44 Gauss.cm⁻¹ (20%). A 45.6 μ s delay (D19) was used for binomial water suppression. Prior to Fourier 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 D₂O 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 μ s gradient recovery delay (D16) and strengths set to 8.44 Gauss.cm⁻¹ (20%). A 100 μ s 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.

2.2 - Circular Dichroism and 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 **Ic** (A) or **IIIb** (B) – Tm TAR alone = $56.2 \pm 0.2^\circ\text{C}$, Tm TAR-**Ic** = $58.5 \pm 0.1^\circ\text{C}$, Tm TAR-**IIIb** = $60.2 \pm 0.1^\circ\text{C}$

2.3 - Thermodynamic analyses

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 (T_c).¹ If the experimental temperature T_{exp} lies outside this confidence interval, the correlation would be significant at the 95% confidence level. In our case, T_{exp} falls always outside the confidence limits, which could indicate that the observed phenomenon has a biological relevance.

3 - MOLECULAR MODELING

3.1 - Protocols and inputs

AutoDock Vina docking protocol

¹ Sharp, K. (2001) Entropy-enthalpy compensation: fact or artifact? *Protein Sci*, **10**, 661.

Docking of the PAAs on HIV-1 TAR X-ray structure (PDB ID 2kx5(16)) was performed with AutoDock Vina software². First, the TAR structure was prepared by trimming the peptide from the original pdb file. Then, the RNA was prepared for docking by using the automated preparation protocol of AutoDock Tools (ADT), which includes protonation of the structure and Gasteiger partial charges assignment. The following input file was used for running all docking calculations:

```
-----  
AutoDock Vina input  
-----
```

```
receptor = tar.pdbqt  
ligand = paa.pdbqt  
center_x = -28.779  
center_y = 63.327  
center_z = -53.986  
size_x = 36.75  
size_y = 36.75  
size_z = 28.5  
out = vina.pdbqt  
-----
```

Autodock Vina was run 10 times for each PAA and only the 40 most energetically favorable conformations were selected and clustered with ADT. A representative conformation from the best scoring docking was selected and subjected to molecular dynamics.

Molecular dynamics protocol

The molecular dynamics simulations were run with pmemd.MPI. The results were visualized and analyzed with VMD³ and Pymol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.). The *closest* command from *cpptraj* module was used for stripping all except the closest *N* waters from the trajectory. All the MM/GBSA calculations were run in a standard basis.

² Trott,O. and Olson,A.J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.*, **31**, 455.

³ Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: visual molecular dynamics. *J Mol. Graph.*, **14**, 33.