Supplementary Information

Identification of Biologically Active, HIV TAR RNA-Binding Small Molecules Using Small Molecule Microarrays

*Joanna Sztuba-Solinska,[†¶] Shilpa R. Shenoy,[⌘¶] Peter Gareiss,[§] Lauren H. Krumpe,[‡⌘] Stuart F.J. Le Grice,[†] Barry R. O'Keefe,[‡] and John S. Schneekloth, Jr.[#]**

[†]HIV Drug Resistance Program, National Cancer Institute, Frederick, MD, USA,
[‡]Molecular Targets Laboratory, National Cancer Institute, Frederick, MD, USA
[⌘]Leidos Biomedical Research, Inc., Frederick National Laboratory, Frederick, MD, USA
[§]Center For Molecular Discovery, Yale University, New Haven CT, USA
[#]Chemical Biology Laboratory, National Cancer Institute, Frederick, MD, USA

*e-mail: schneeklothjs@mail.nih.gov

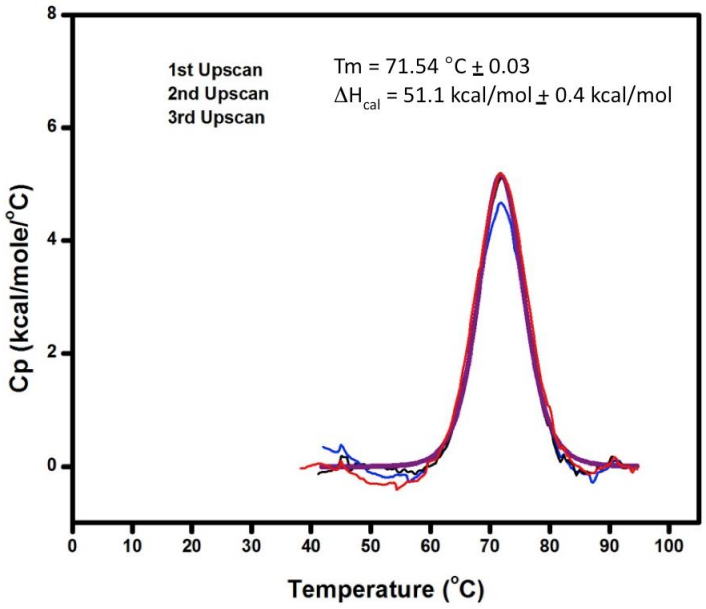
General Information:

Nucleic acids were purchased from Dharmacon (ThermoFisher), with the following sequences:

TAR RNA: 5'-GGCAGAUCUGAGCCUGGGAGCUCUCUGCC-3'
Cy5-labeled TAR RNA: 5'-Cy5-GCAGAUCUGAGCCUGGGAGCUCUCUGCC-3'
2-AP-labeled TAR RNA: 5'-GGCAGAUC(2AP)GAGCCUGGGAGCUCUCUGCC-3'
Cy5-labeled miR-21 RNA:5'-Cy5-GGGUUGACUGUUGAAUCUCAUGGCAACCC-3'

Small molecules were purchased from commercial vendors, assessed for purity with LC/MS analysis and ^{1}H NMR, dissolved in DMSO at a concentration of 10 mM, and used without further purification. See Figure SI-4 for compound supplier/product ID information.

Figure SI-1: Reversibility of TAR RNA hairpin folding by differential scanning calorimetry.



Differential Scanning Calorimetry (DSC) – DSC experiments were carried out on a Microcal VP-DSC microcalorimeter (GE Healthcare, originally Microcal, Northampton, MA). Prior to running the experiments, all surfaces were treated with RNase Zap Wipes (Ambion®) to minimize RNase contamination. The calorimeter, syringes, degassing tubes and stir bars were cleaned with 70% ethanol, followed by DEPC treated water (Ambion®). A 50 μM solution of TAR RNA in PBS pH 7.4 was heated to 95 °C in a hot water bath for 5 minutes and was allowed to cool to room temperature (over a 2 hour period). Annealed RNA was frozen in 1 mL aliquots at -20 °C. Per routine protocol, buffer (1X PBS, pH 7.4) was introduced to both the reference and sample cells and the calorimeter was allowed to ramp through one heat-cool cycle (10 °C to 95 °C) at a rate of 60 °C/hr. During the down scan at 25° C, the buffer solution from the sample cell was quickly replaced with a degassed solution of 50 μM TAR RNA. The entire system was re-pressurized to approximately 30 psi of positive pressure to prevent evaporation at higher temperatures, and the experiment was allowed to continue. A total of 8 alternating up-down scans (10 C to 95 C) was performed to measure reversibility of folding/unfolding of the RNA. According to Origin DSC Anaylsis software, buffer effects were corrected and integration of the unfolding transition was performed. Thermograms were fitted to a two-state melting model and the calorimetric transition enthalpy (ΔH_{unf}) was obtained from the area under the excess heat capacity peak, the midpoint of the transition calculated as the melting temperature (T_m).

Minus strand strong stop assay performed in the presence of 4

RNA-dependent DNA polymerase activity was determined on a viral RNA corresponding to nt 1–365 of the HIV-1 NL4-3 (+) strand genome. A fluorescently labeled DNA primer (5'-Cy5 GTC CCT GTT CGG GCG CCA-3') was combined with template RNA at a ratio of 1:1.2 in 10 mM Tris/HCl, pH 7.6, 25 mM KCl and heated in a thermal cycler at 85 °C for 3 min and then cooled to 4 °C at 0.2 °C/s. **4** was added at final concentrations of 0, 1, 5, 10 μM and incubated for 10 min at room temperature. DNA polymerase reactions were performed at 37 °C and contained 10 mM Tris/HCl, pH 7.8, 50 mM $MgCl_2$, 80 mM KCl, 1 mM DTT, 0.2 mM dNTPs, 25 nM template/ primer, and 12.5 nM RT. Aliquots were removed at 0, 5, and 20 min time points and combined with an equal volume of 8 M urea in 1x Tris/borate/EDTA. Before loading, samples were heated to 95 °C for 3 min and immediately placed on ice. Nucleic acids were fractionated by denaturing 8% polyacrylamide gel electrophoresis. Gels were scanned with a GE Health- care Typhoon Trio + and analyzed with Image Quant Total Lab software.

Figure SI-2.

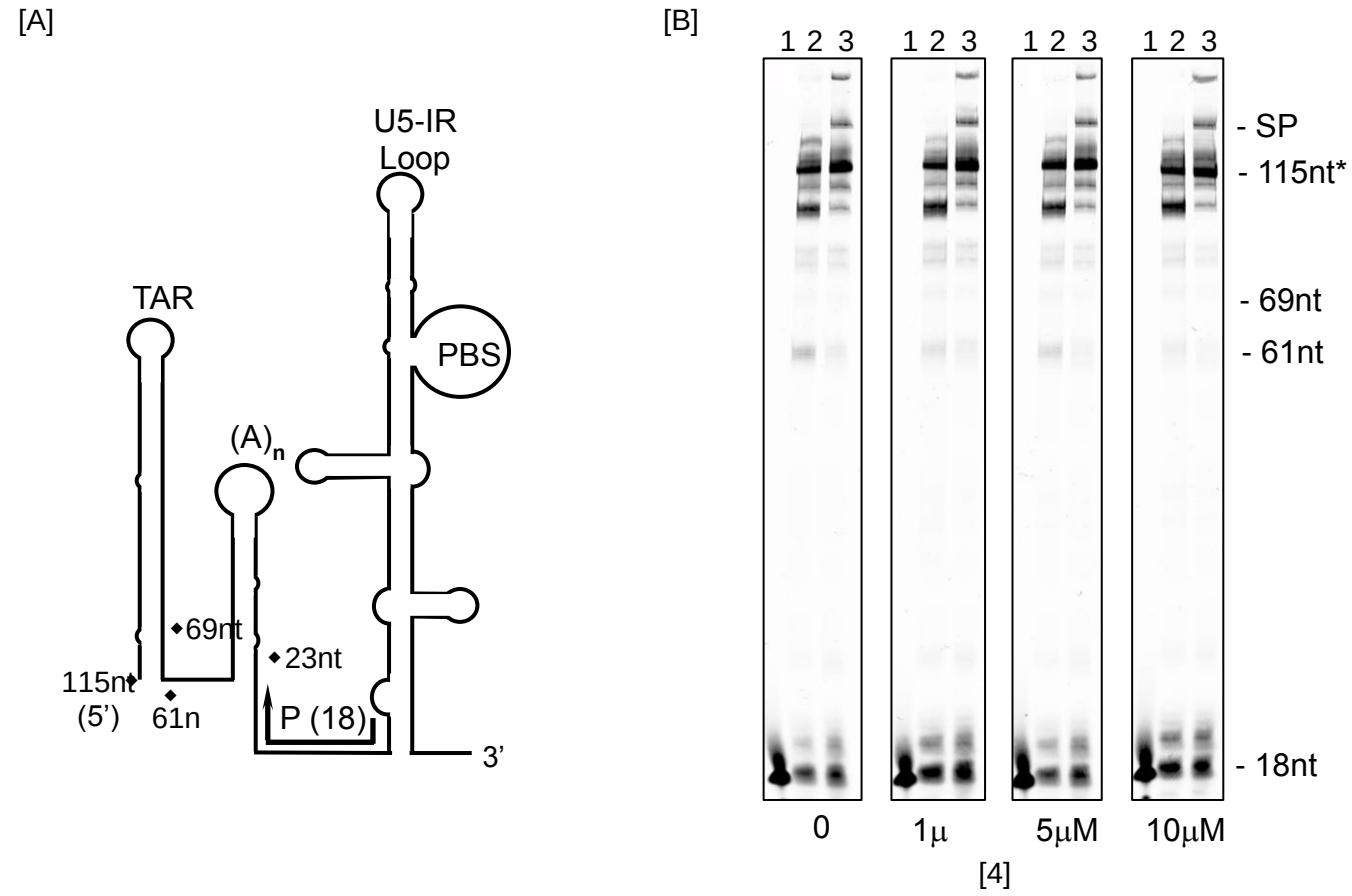


Figure SI-2. Minus strand strong stop assay performed in the presence of **4**.

A schematic representation of the HIV-1 genomic RNA substrate is presented in panel [A]. PBS, primer binding site; Poly(A), poly(A) hairpin; TAR, transactivation hairpin. The position to which the synthetic primer was hybridized is indicated via arrow. In the RNA-dependent DNA synthesis assay of panel [B], the asterisk indicates the expected DNA synthesis product (~115 nt), and SP reflects the results of a self-priming event. Lanes 1, 2 and 3 represent 0, 5 and 20 min time points, respectively, performed in the presence of 4 compound at different final concentrations: 0, 1, 5, 10 uM.

Figure SI-3.

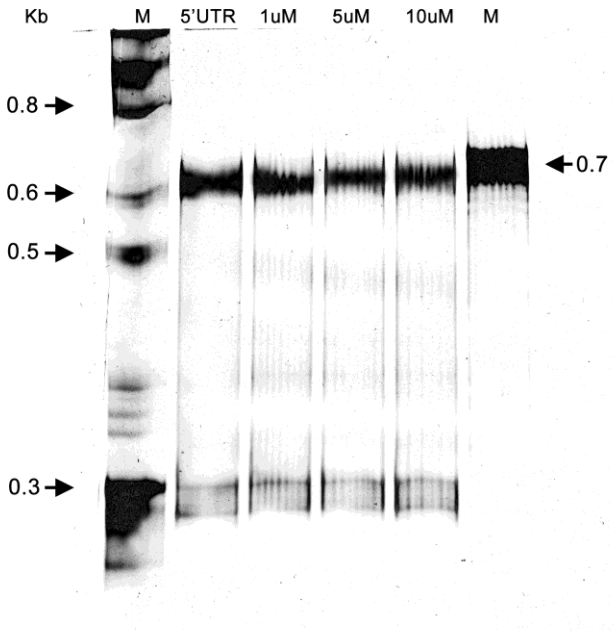
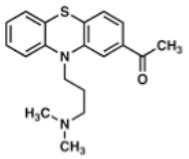


Figure SI-3. Verification of homogenous conformation of the HIV-1 RNA 5' UTR region. Increasing concentrations of the compound **4** (indicated on top) incubated with HIV-1 RNA 5' UTR region were added prior the non-denaturing polyacrylamide gel electrophoresis. The leftmost lane provides molecular weight markers.

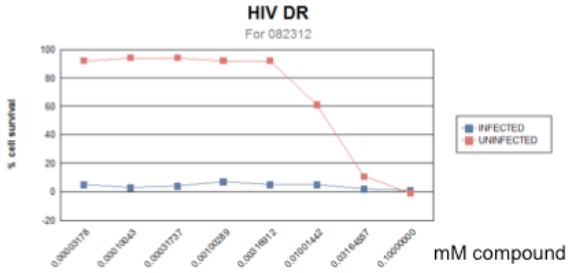
Figure SI-4. Cell-based Anti-HIV activity of **1, 4**, and related compounds. Compounds were purchased from commercial vendors, analyzed for purity by LC/MS, and used without further purification. For each compound, vendor name (below, in parenthesis) and product ID are indicated. Assays were performed as described in the main text.

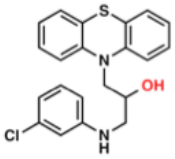
Compound	EC_{50} (μM)	CC_{50} (μM)	Comments
ST4133609 (Compound **4**)	28	ND	
4478-7480 (Compound **1**)	123	ND	
ST50055849	> 100	ND	
AKOS001656555 (Compound **2**)	> 10	9	
5251219	> 20	40	
4340-1594 (Compound **3**)	> 30	ND	
9233834	> 63	ND	
5277323	> 63	ND	
7878578	> 63	ND	
7928037	> 63	ND	
7852383	> 63	ND	
7911696	> 63	ND	
Acepromazine maleate	> 16	18	
ST4119563	> 330	190	
ST4133739	> 316	ND	
7746407	> 3	ND	Abnormal curves
ST014353 (Compound **5**)	ND	ND	Abnormal curves

ND = Not Determined.

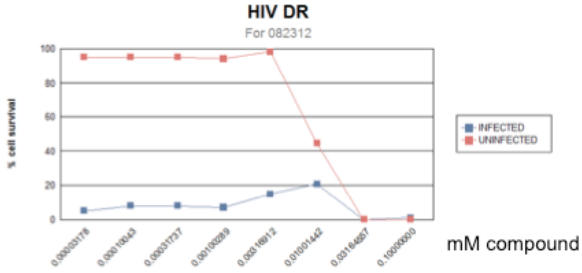


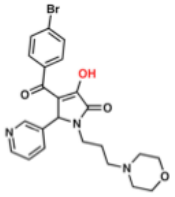
Acepromazine maleate



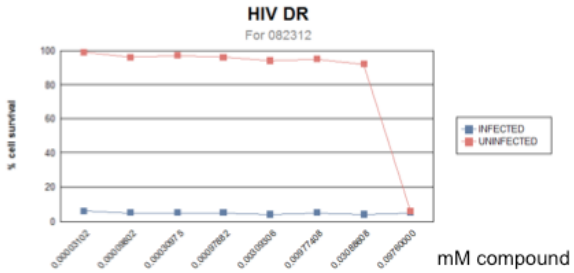


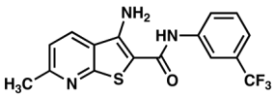
AKOS001656555 (AKos)



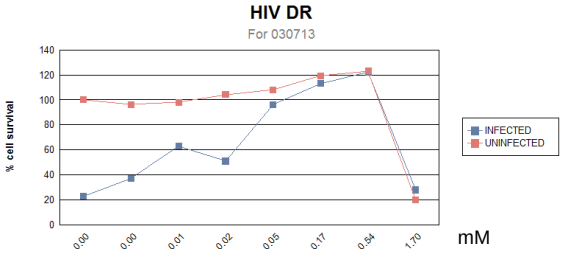


4340-1594 (ChemDiv)

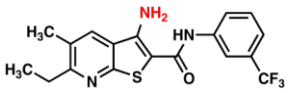




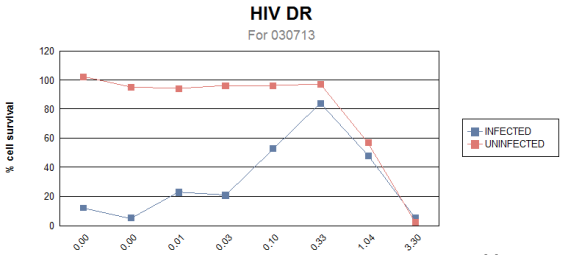
ST4133609
(TimTec)



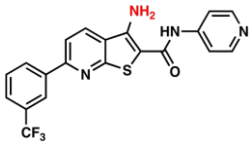
mM



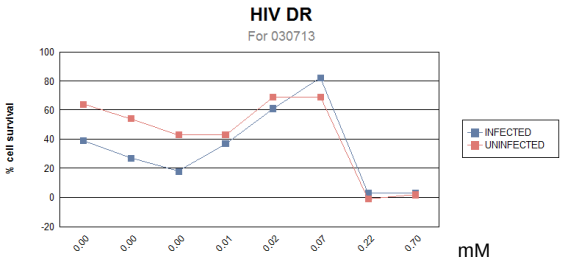
4478-7480
(ChemDiv)



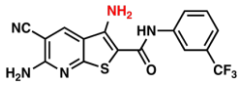
mM



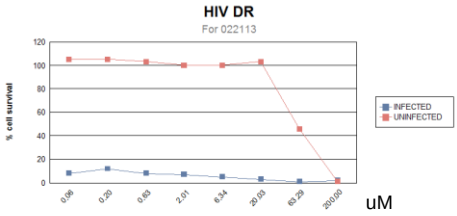
ST014353
(TimTec)



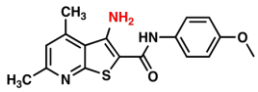
mM



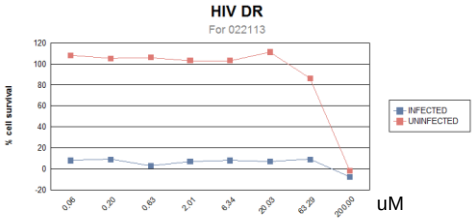
9233834
(ChemBridge)



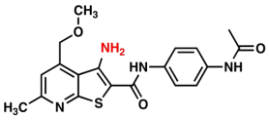
uM



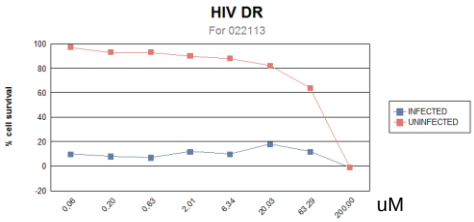
5277323
(ChemBridge)



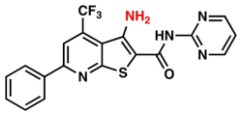
uM



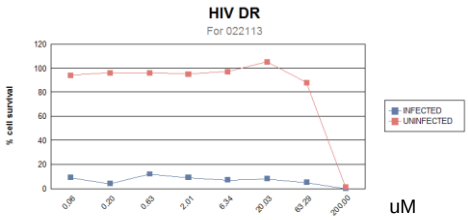
7878578
(ChemBridge)



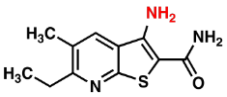
uM



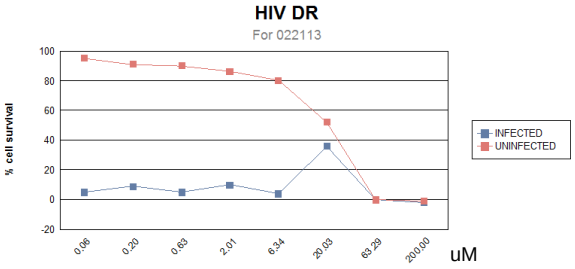
7928037
(ChemBridge)

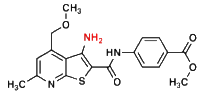


uM

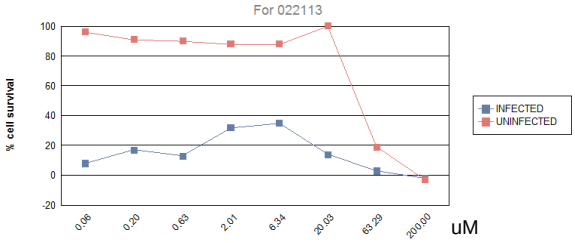


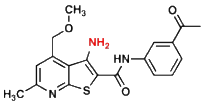
5251219
(ChemBridge)



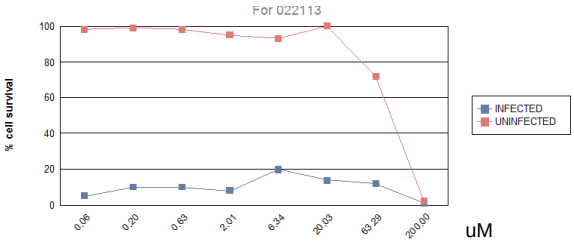


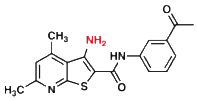
7746407
(ChemBridge)



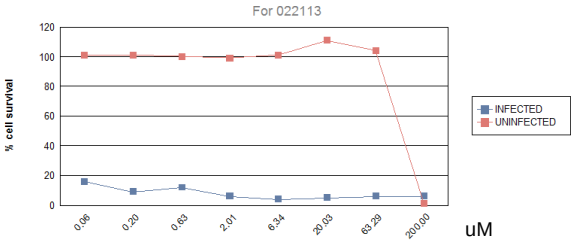


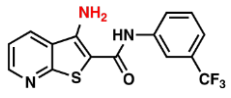
7852383
(ChemBridge)



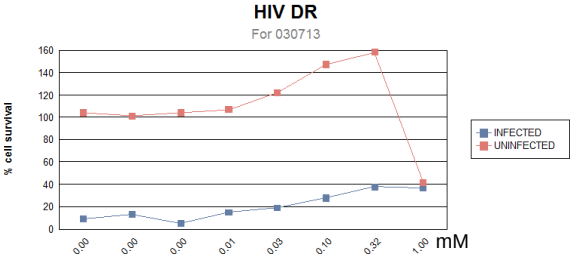


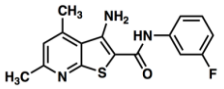
7911696
(ChemBridge)





ST4133739
(TimTec)





ST50055849
(TimTec)

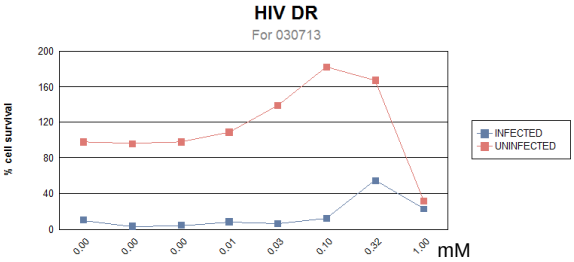


Figure SI-5. K_d measurement with compound **5** using a 2-aminopurine titration. The titration was carried out as described in the main text.

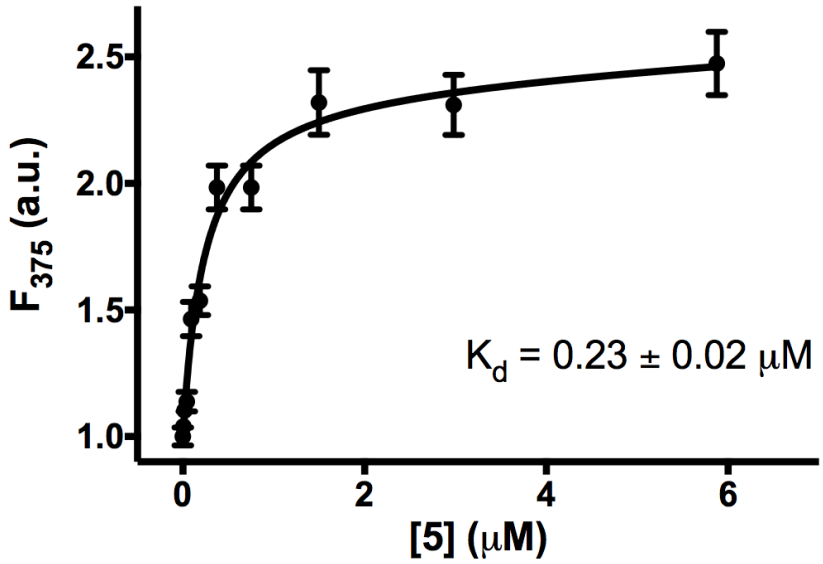
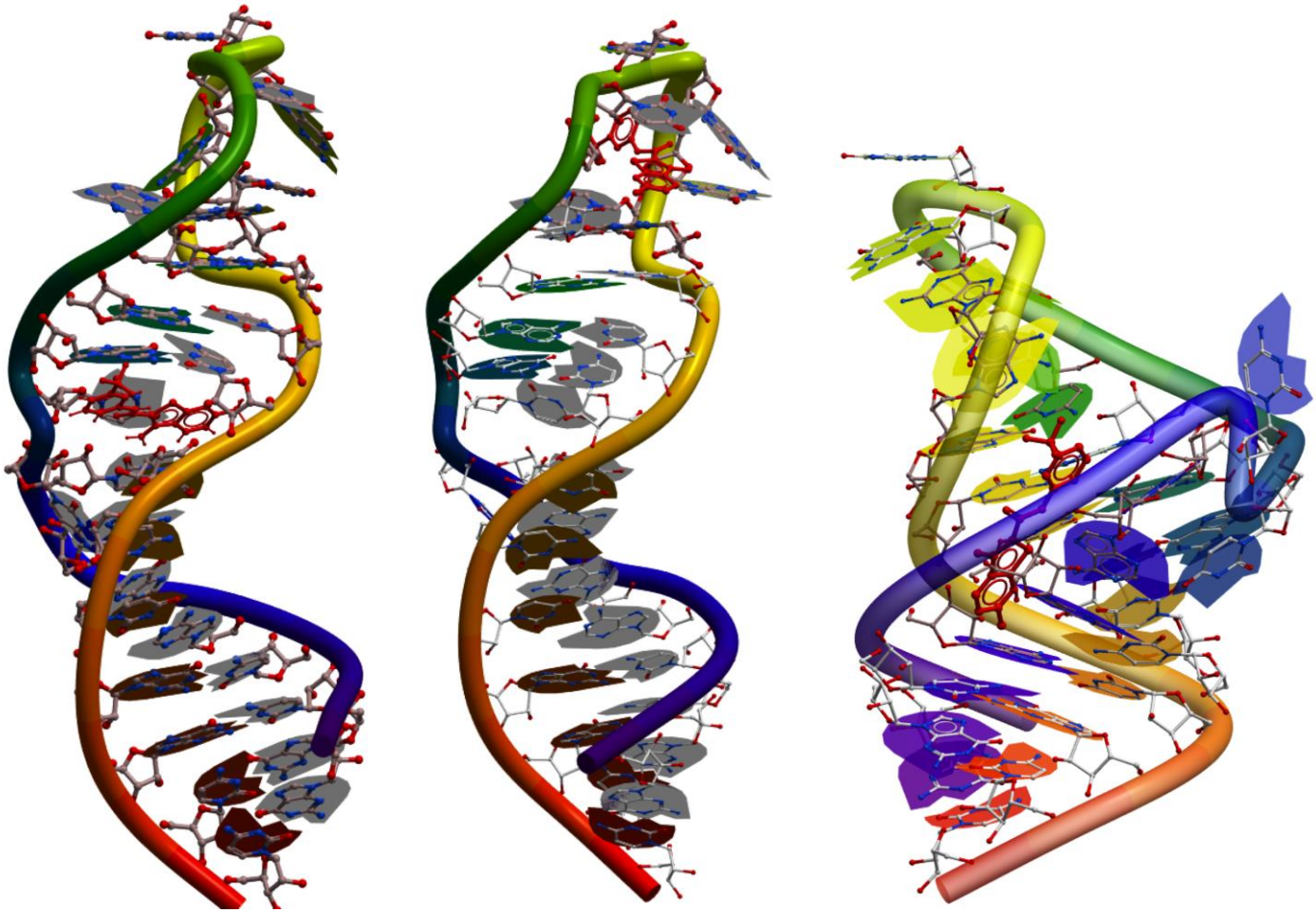


Figure SI-6. Docking compound **4** to multiple known TAR conformations



	1LVJ (Bulge)	1LVJ (Loop)	2L8H (Bulge)
ICM Score:	**-28.46**	**-22.02**	**-24.18**
Other forces: (van der Waals, hydrogen bonding, hydrophobic forces)	**-39.06**	**-41.33**	**-38.31**
Electrostatics:	**6.21**	**14.21**	**14.6**

Docking studies were performed with ICM (Molsoft) using TAR structures from the PDB (listed below) and compound **4**. Binding pockets on TAR were defined both using the ICM PocketFInder module as well as by docking **4** to known small molecule binding sites. For each of the structures, compound **4** was docked both to all identified binding sites. Docking simulations were carried out at a thoroughness of 10 (the maximum number of Monte Carlo iterations). Each simulation provided ~15-20 poses, each of which was visually inspected after completion of the simulations. The three best scoring poses are illustrated above, with compound **4** drawn in red. In each case, the relative contribution of non-electrostatic forces (van der Waals, hydrophobic interactions, hydrogen bonding interactions) and electrostatics to the binding score are indicated.

Structures Used in Docking Simulations:

1LVJ (acetylpromazine-bound)
1QD3 (neomycin-bound)
1UTS (synthetic compound-bound)
1UUD (synthetic compound-bound)
1UUI (synthetic compound-bound)
2L8H (synthetic compound-bound)