Supporting Information for

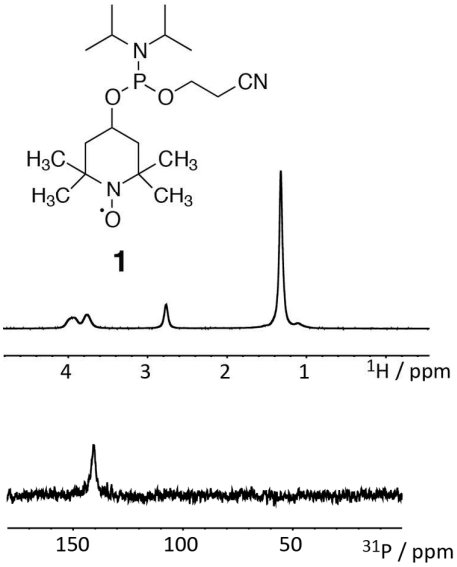
A novel paramagnetic relaxation enhancement tag for nucleic acids: A tool to study structure and dynamics of RNA

Christoph H. Wunderlich[‡1], Roland G. Huber[‡2], Romana Spitzer[1], Klaus R. Liedl[2], Karin Kloiber[*1], and Christoph Kreutz[*1]

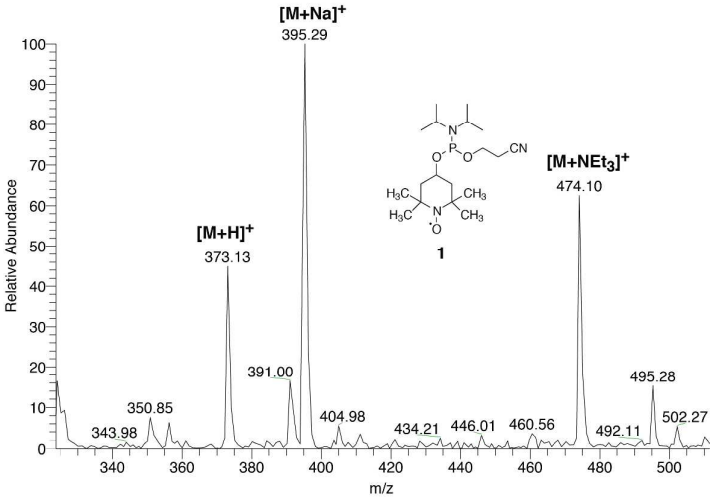
[1]Institute of Organic Chemistry and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innrain 80/82, 6020 Innsbruck, Austria

[2]Institute of General, Inorganic and Theoretical Chemistry and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innrain 80/82, 6020 Innsbruck, Austria

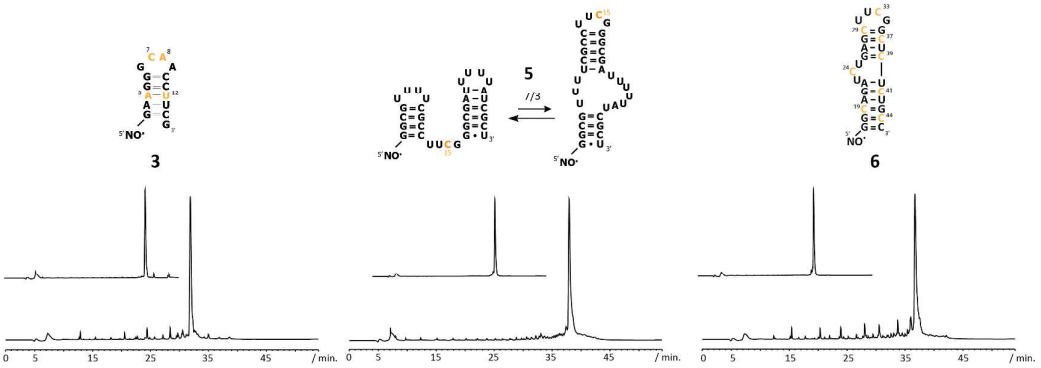
Spectral data of TEMPO phosphoramidite 1 and analytical data of TEMPO modified RNA sequences 3, 5 and 6



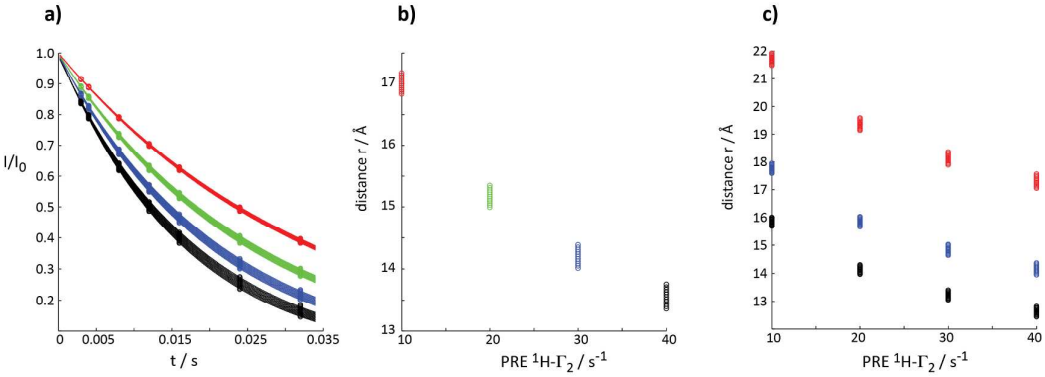
Supporting Figure 1a. ^{1}H- and ^{31}P- NMR spectra of the TEMPO phosphoramidite **1**. ^{1}H-NMR spectra acquired at 300 MHz in $CDCl_3$ at 298 K. ^{31}P-NMR spectra acquired at 121 MHz in $CDCl_3$ at 298 K.



Supporting Figure 1b. ESI-MS of TEMPO phosphoramidite **1**.

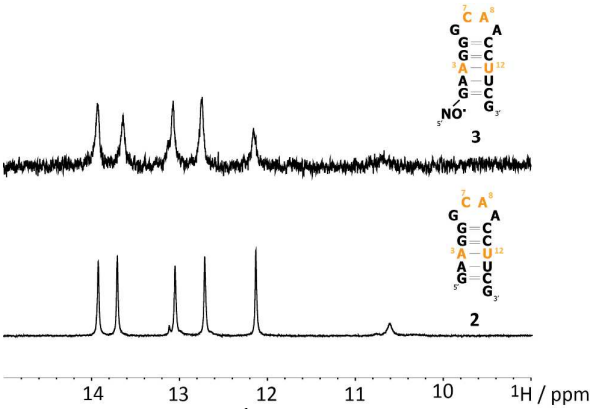


Supporting Figure 1c. Anion-exchange chromatograms of crude and purified (inset) products of RNA sequences **3**, **5** and **6**. As judged from the HPLC traces the 5'-TEMPO tag is attached with at least 98% efficiency.

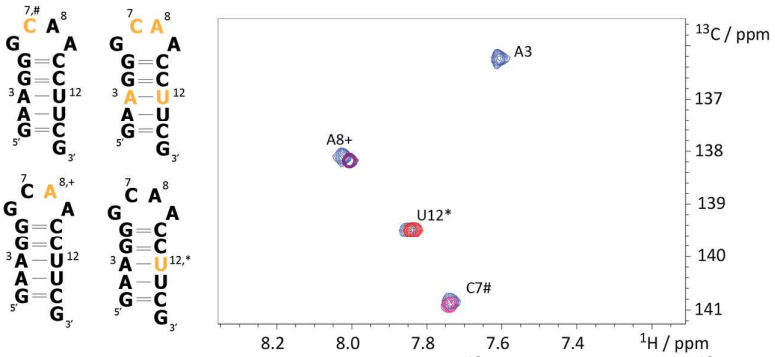


Supporting Figure 1d. Potential systematic errors arising from a partially degraded radical tag on PRE analysis. a) Effect of 'radical dilution' on the decays of transverse magnetization assuming that the unlabeled species decays at 20 s^{-1}. PREs were set to 10 s^{-1} (red lines), 20 s^{-1} (green lines), 30 s^{-1} (blue lines), and 40 s^{-1} (black lines). The most rapid decay represents pure TEMPO radical labeled species, the slowest decays (of the same color) represent the presence of 10% degraded radical species. b) The corresponding extracted distances for an assumed correlation time of 4.61 ns. Color coding is as in a) such that the smallest PRE corresponds to the largest distance. Dilution by non radical-tagged species leads to a slight overestimation of the 'exact' distance. c) Systematic errors depending on PRE (x-axis) and correlation time: black points correspond to 3 ns, blue lines to 6 ns, and the red circles are values obtained for correlation times of 20 ns. In all cases, systematic errors are small compared to experimental errors obtained in our study. Note also that the issue arises only if labeled and unlabeled species have overlapping resonances in the spectrum.

^{1}H-NMR and HSQC spectra of hairpins 2 and 3

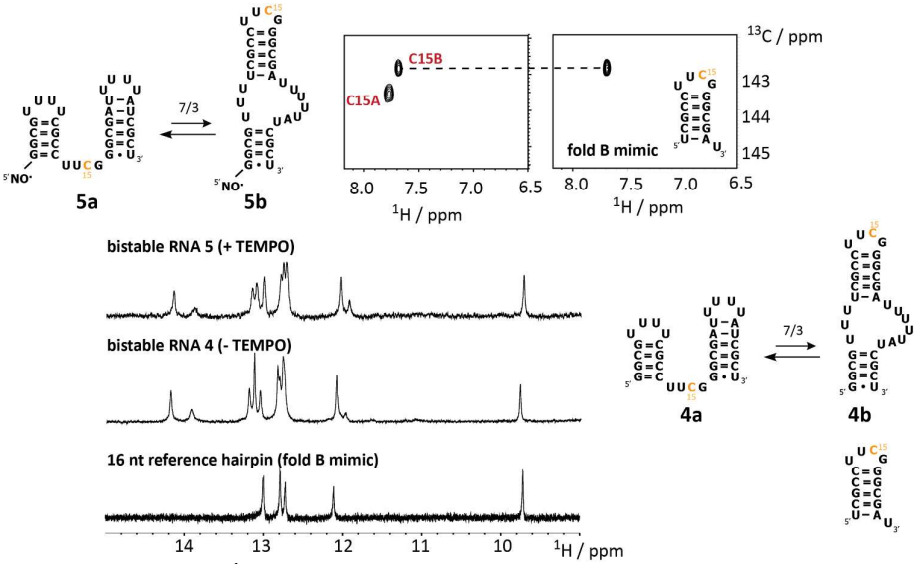


Supporting Figure 2a. ^{1}H-NMR spectra hairpin **3** and **2**. The imino proton spectra confirm that the overall fold is not perturbed by the attachment of the TEMPO radical.



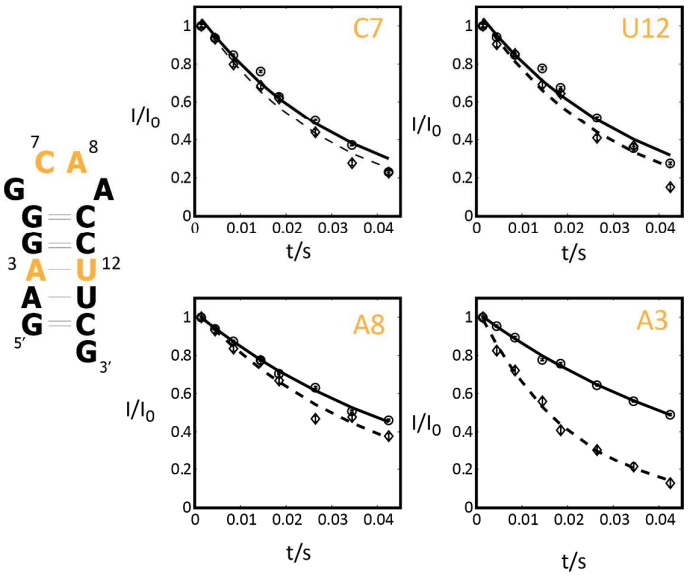
Supporting Figure 2b. HSQC spectra of fully ^{13}C-labeled and single ^{13}C-labeled hairpin RNAs used for resonance assignment.

^{1}H-NMR and HSQC spectra of bistable RNA 4 and 5 and references for fold assignment



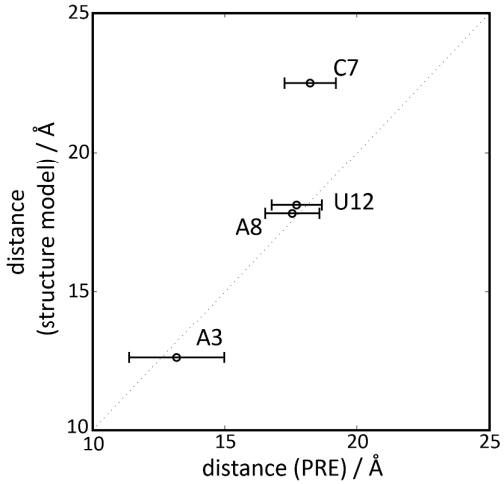
Supporting Figure 2c. ^{1}H-NMR and HSQC spectra of bistable RNA **4/5** (without/with TEMPO tag). A 16 nt hairpin fold B mimic was used for assignment purposes.

R_2 proton decays of hairpin RNAs 2 and 3



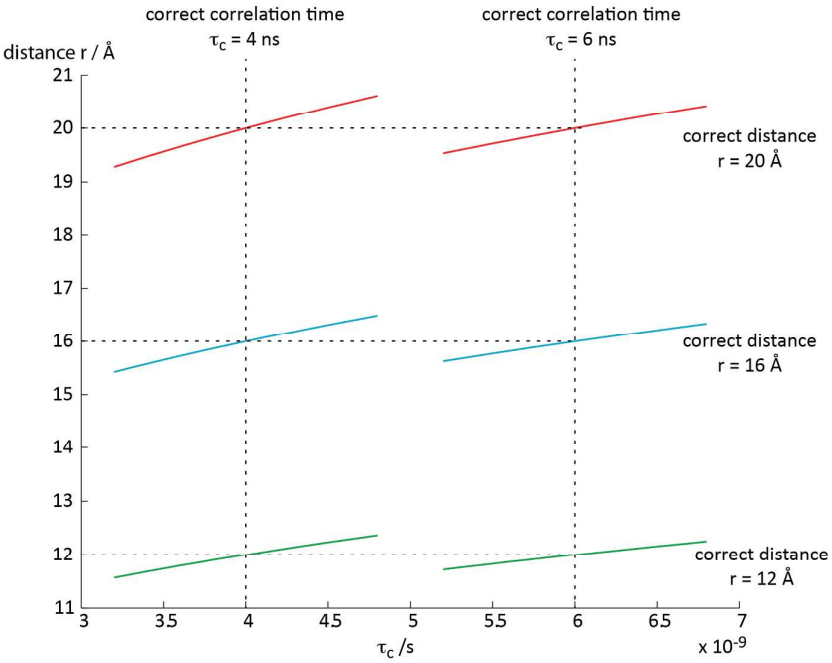
Supporting Figure 3a. Experimental decays of transverse proton H6/H8 magnetization in the absence and presence of the radical tag of selectively labeled nucleobases of hairpins **2**/**3**. Fits of experimental decays are shown as solid lines (- TEMPO) and dashed lines (+ TEMPO). Error bars are derived from MC analysis (500 runs) on the basis of experimental S/N ratios.

Correlation of PRE derived distances versus distances from structural model of hairpin 3



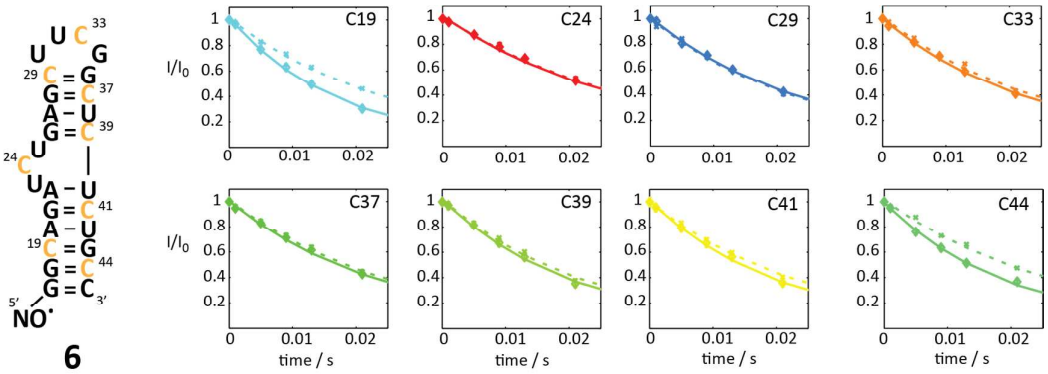
Supporting Figure 3b. Correlation of experimental (x-axis) and predicted (hairpin **3** structure model from MC-FOLD/MC-SYM pipeline) distances between the radical center and H6/H8 of the selectively labeled nucleotides.

PRE derived distance variations from fluctuations in local cytidine correlation times



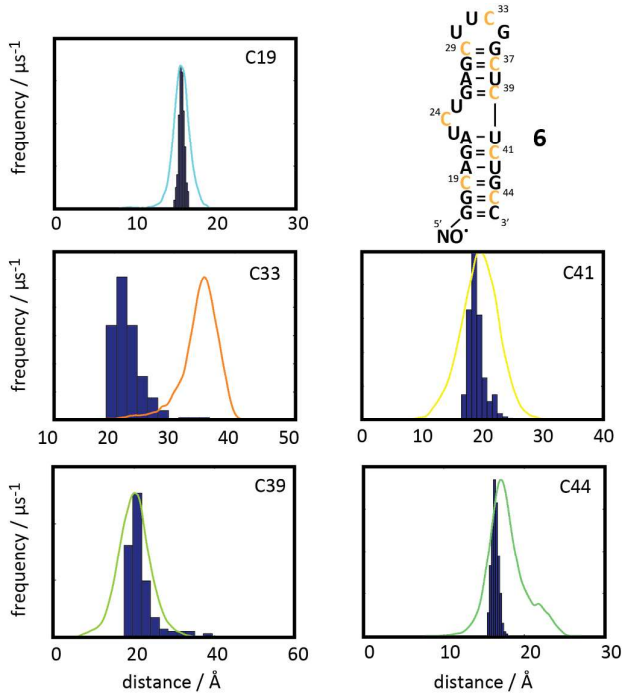
Supporting Figure 4. Propagation of experimental uncertainties in correlation times. Numerical simulations covering two 'correct correlation times' of 4 and 6 ns and three 'correct distances' 12, 16, and 20 Å. The dependency of PRE derived distances on the correlation time τ_c is shown for a range of 3.2 to 4.8 ns (around 4 ns) and 5.2-6.8 ns (around 6 ns). The extracted distances are within the range of experimental PRE errors.

R_2 proton decays of the free HIV-1 TAR RNA 6



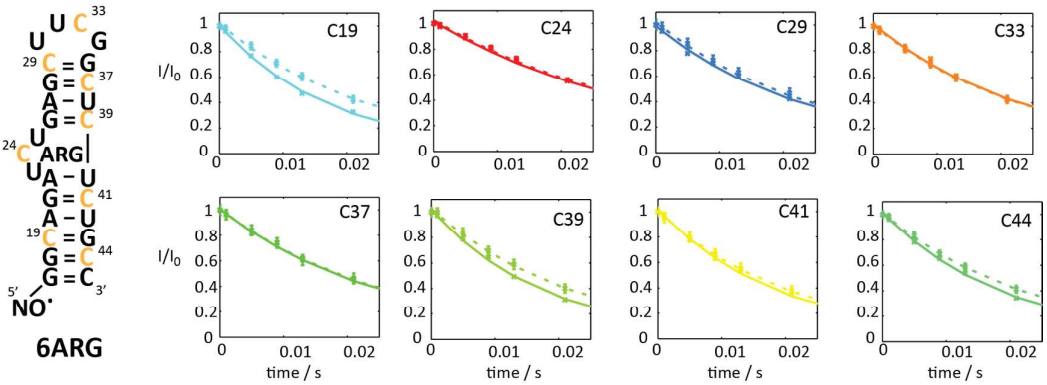
Supporting Figure 5a. Decays of ^{1}H transverse magnetization of H6 of selectively labeled cytidine residues and fits of free HIV-1 TAR RNA **6**. Transverse relaxation rates without (crosses) and with TEMPO (diamonds) are shown along with the respective mono-exponential fits (dashed/solid lines: without/with TEMPO). Largest effects are observed for C19, C44, C41, C39, and a small PRE was found for C33, which is located in the loop.

MD distance distributions and PRE derived distances of the free HIV-1 TAR RNA 6



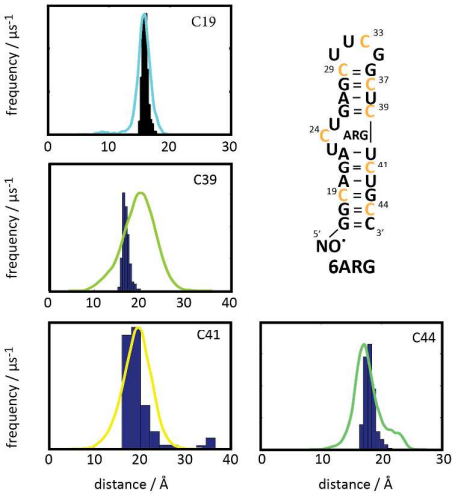
Supporting Figure 5b. Superposition of distance distributions from the MD simulation (solid lines) superimposed on results from Monte Carlo analysis of PRE-derived distances for free HIV-1 TAR RNA **6** (histograms). Measured PREs and determined rates represent a time-average over many conformations that are sampled in the rapid exchange process. The distributions represented by the histograms refer to the uncertainties in the experimentally derived mean values are not to be confused with the distributions yielded by the MD simulation.

R_2 proton decays of the binary HIV-1 TAR RNA argininamide complex 6ARG



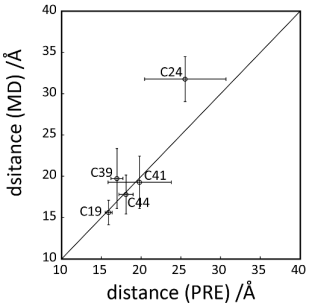
Supporting Figure 6a. Decays of ^{1}H transverse magnetization of H6 of selectively labeled cytidine residues and fits of argininamide-bound HIV-1 TAR RNA **6ARG**. Relaxation rates without (diamonds) and with TEMPO (crosses) are shown along with the respective mono-exponential fits (dashed/solid lines: without/with TEMPO). PREs are generally smaller than in the free form, with largest effects observed for C19, C44, C41, C39 (the PRE of C39 is increased with respect to the free form). The small PRE found for C33 in the free form is not found in the arginine bound form.

MD distance distributions and PRE derived distances of the binary HIV-1 TAR RNA argininamide complex 6ARG

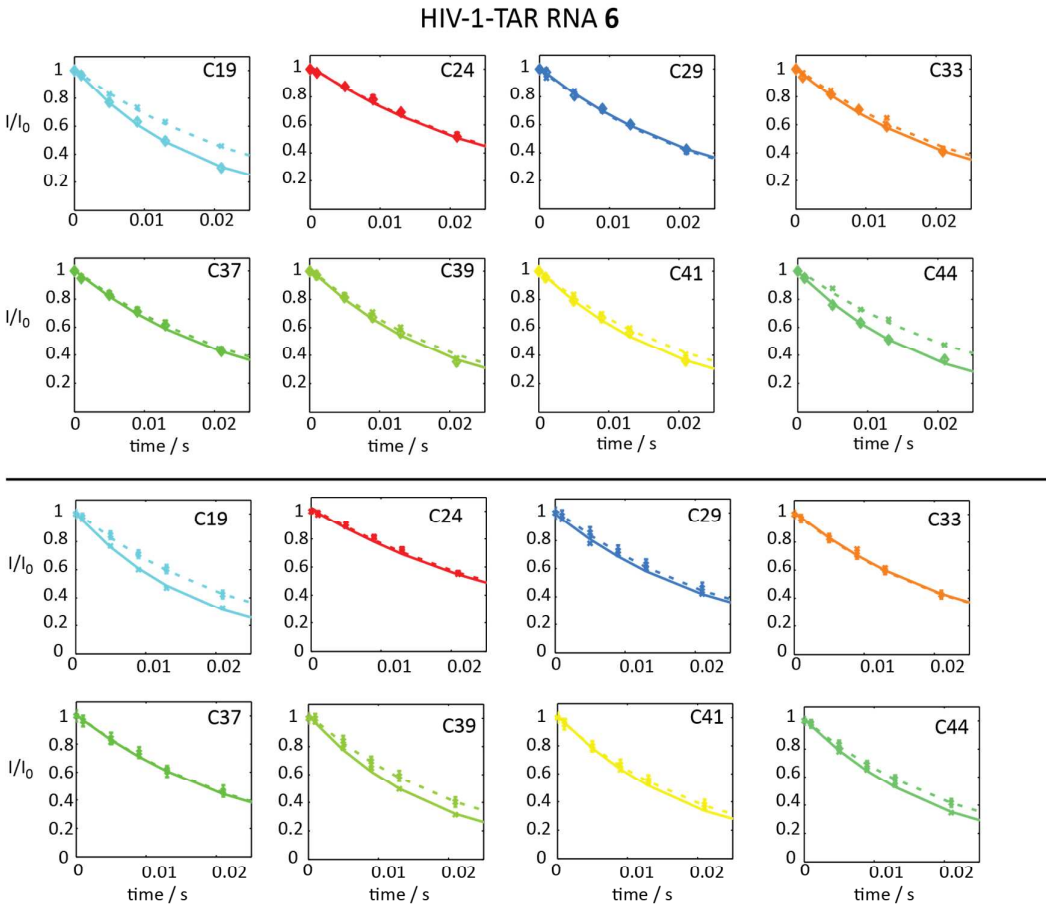


Supporting Figure 6b. Superposition of distance distributions from the MD simulation superimposed on results from Monte Carlo analysis of PRE-derived distances for argininamide bound HIV-1 TAR RNA **6ARG**. Measured PREs and determined rates represent a time-average over many conformations that are sampled in the rapid exchange process. The distributions represented by the histograms refer to the uncertainties in the experimentally derived mean values are not to be confused with the distributions yielded by the MD simulation. The small PRE detected for C33 in the free RNA is now absent and was not analyzed.

Correlation of PRE derived distances versus MD time averaged distances of the binary HIV-1 TAR RNA argininamide complex 6ARG

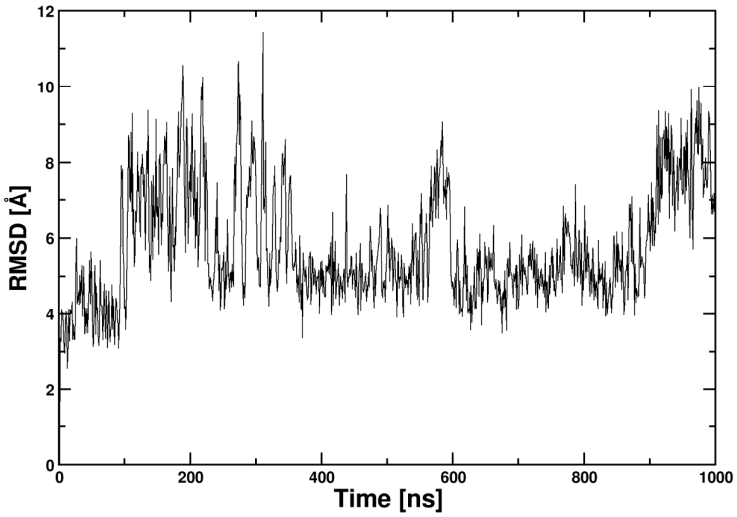


Supporting Figure 6c. Correlations between experimental distances of H6-^{13}C of cytidine residues from the radical center in the argininamide bound HIV-1 TAR RNA **6ARG** with the mean distance values obtained from an MD simulation. Error bars for the experimental values are derived from MC analysis (500 runs), whereas those for the structural ensemble and the MD simulation correspond to the standard deviation of the values.

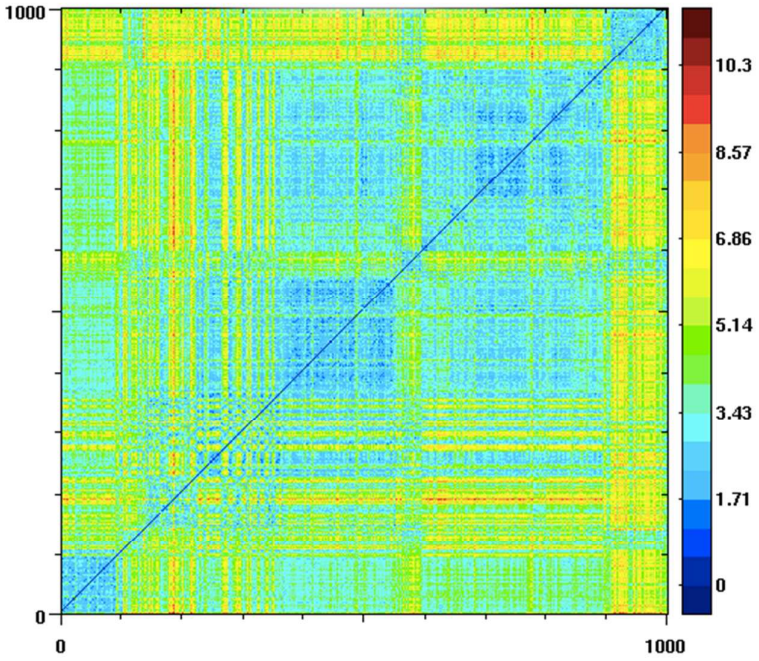


Supporting Figure 6d. Influence of argininamide binding on PRE. Decays of ^{1}H transverse magnetization of H6 of selectively labeled cytidine residues and fits of free HIV-1 TAR RNA **6** and argininamide-bound HIV-1 TAR RNA **6ARG**. Relaxation rates without and with TEMPO are shown along with the respective mono-exponential fits (dashed/solid lines: without/with TEMPO).

Analysis of the holo simulation of the TEMPO-tagged HIV-1 TAR RNA (6)

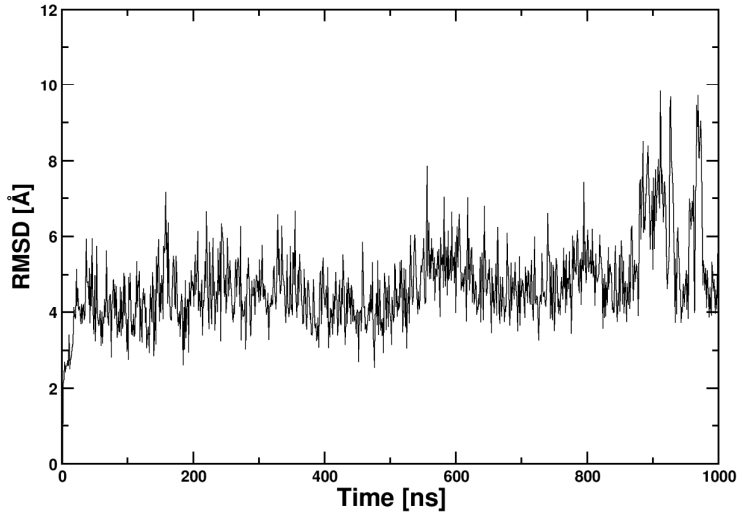


Supporting Figure 7a. RMSD-plot *vs.* 2L8H over simulation time (1 µs). After initial relaxation over the first 100 ns of the simulation, the system switches between two distinct states.

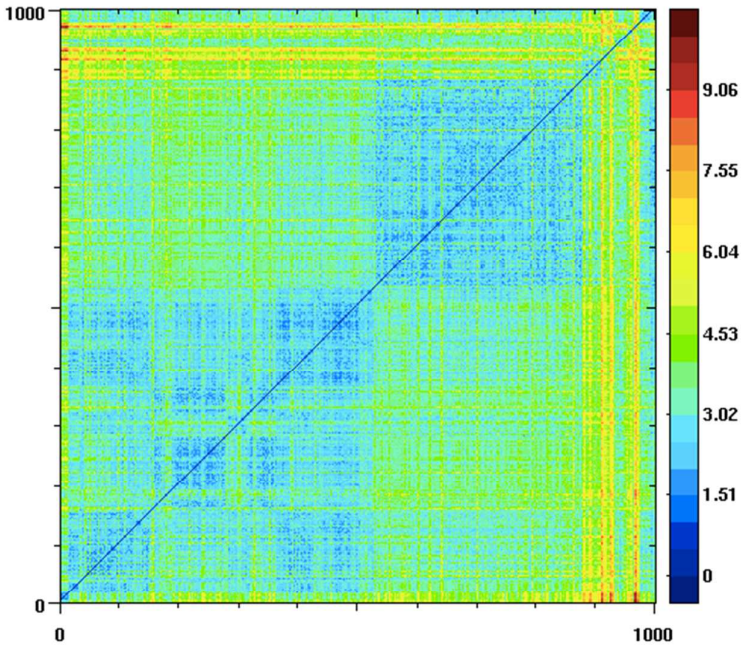


Supporting Figure 7b: The two-dimensional RMSD-Plot shows the similarity of the individual simulation frames against each other. The two conformations revealed in **Supporting Figure 7a** can be clearly identified. The dominant state is similar to the initial structure. A second state is regularly populated approximately 10% of the simulation time.

Analysis of the simulation of the TEMPO-tagged HIV-1 TAR RNA in complex with argininamide (6ARG)

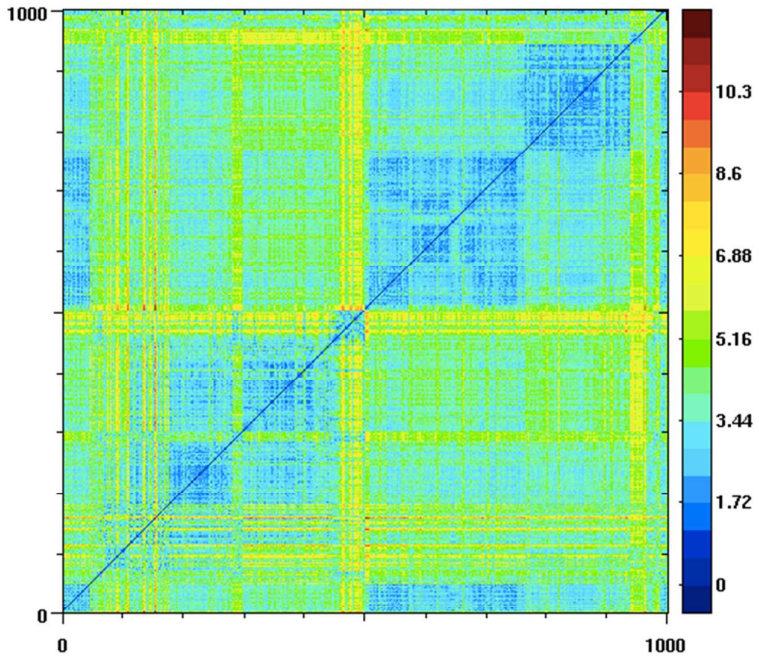


Supporting Figure 8a. The argininamide-complex simulation is much closer to the initial structure 2L8H. As 2L8H contains a ligand similar to argininamide, this indicates, that the complex is stabilized in this conformation by the presence of the ligand.



Supporting Figure 8b. Two-dimensional RMSD analysis shows most conformations are close to the initial structure. Two slightly different ligand-bounds conformation exist. A third conformation is present for approximately 5% of the simulation time. This third conformation relaxes quickly to one of the states close to the initial structure.

Comparison of the holo-HIV-1 TAR RNA (6) and the argininamide complex (6ARG) MD simulations



Supporting Figure 9. 2D-RMSD comparison between the two simulated systems. The first 500 frames correspond to the holo-simulation (construct **6**), the second 500 frames correspond to the HIV-1 TAR RNA argininamide complex (**6ARG**). The plot reveals that the holo-system relaxes to a defined stable conformational ensemble after approximately 100 ns. This ensemble resembles the bound state, hinting at a combined induced fit/conformational selection mechanism of the complex formation. The high-energy states of both simulations correspond in part.

Supporting Table 1. Distances between 5'-terminus (4.5 Å were added to take the radical tag into account) and the H6 of cytidine residues. HIV-1 TAR RNA structures (with various ligands) were taken from the PDB and IDs are indicated. In case of NMR structures the distance was averaged over all structures in the ensemble deposited in the PDB.

residue	**PDB ID**										
	1ARJ	1QD3	1UTS	1UUD	1UUI	2KDQ	2KX5	2L8H	397D	**average**	**PRE**
C19	15.3	13.7	14.1	14.5	14.0	15.2	16.3	14.5	14.1	**14.6**	**15.3±0.3**
C24	38.4	-	29.0	30.5	31.3	37.6	33.6	33.5	35.5	**33.7**	**-**
C29	31.8	32.9	31.5	24.9	25.9	35.4	26.7	24.5	31.2	**29.4**	**-**
C37	25.0	26.5	27.5	22.0	15.2	33.1	19.3	18.7	26.2	**23.7**	**22.2±4.7**
C39	22.6	24.5	20.8	13.3	22.9	24.1	16.6	14.6	20.5	**20.0**	**20.9±4.3**
C41	23.0	23.0	17.7	12.8	14.2	18.5	20.8	18.2	19.2	**28.6**	**18.7±1.6**
C44	19.1	-	19.6	18.2	18.0	16.8	22.0	20.2	19.0	**19.1**	**15.7±0.4**