SUPPORTING INFORMATION

Ligand Recognition in Viral RNA Necessitates Rare Conformational Transitions

Lev Levintov[‡], and Harish Vashisth[‡*]

[‡]*Department of Chemical Engineering, University of New Hampshire, 33 Academic Way, Durham, New Hampshire 03824, United States*

[*] Corresponding author; Email: harish.vashisth@unh.edu

I. MODELS AND METHODS

A. System Preparation and Simulation Details

Software and Force-Field: In this work, we focused on studying the unbinding process of acetylpromazine from HIV-1 TAR RNA using conventional molecular dynamics (MD) and steered molecular dynamics (SMD) simulations. The initial coordinates for our system were obtained from the NMR structure deposited in the Protein Data Bank (PDB code: 1LVJ).[1] A 2 μs long conventional MD simulation of the RNA/ligand complex was conducted using the AMBER force-field[2,3] in the AMBER software.[4,5] All SMD simulations were also carried out using the AMBER force-field (ff99OL3)[2,3] but using the software NAMD.[6] The analyses of all trajectories were carried out using the CPPTRAJ and VMD software.[6–8] For acetylpromazine, the "antechamber" program[9,10] in AMBER was used to develop the force field parameters with atomic charges using the AM1-BCC charge method.[11]

MD: The system was solvated in a 57×84×55 Å^3 periodic box of TIP3P water molecules and the total number of atoms was 22053. The overall charge of the system was neutralized with 29 Na^+ ions. No constraints were imposed on the RNA/ligand complex. The temperature was maintained using the Langevin thermostat at 310K, consistent with experimental conditions,[1] and the pressure was maintained at 1 atm using the Berendsen barostat. The steepest descent minimization was initially performed for 1000 steps followed by 100-500 steps of conjugate gradient minimization. The system was subjected to a 2 μs long MD simulation in the NPT ensemble with a 2 fs timestep. The configurations were saved every 20 ps. Data from this simulation are shown in Figs. S1 and S2.

SMD: The system was solvated in a 54×90×90 Å^3 periodic box of TIP3P water molecules and the total number of atoms was 33936. The overall system was charge neutralized with 29 Na^+ ions and was energy-minimized via 500 cycles of conjugate-gradient optimization. To equilibrate the box volume, a 500 ps MD simulation with a 2 fs timestep was initially conducted. The coordinates from the end of this MD simulation were used as initial conditions for subsequent 5 ns long SMD simulations in the NPT ensemble, conducted using a 2 fs timestep. Even though the ligand dissociated at 25 Å, we continued SMD simulations up to a distance of 60 Å. The temperature and pressure were maintained at 310 K and 1 atm using the Langevin thermostat and the Nosé-Hoover barostat. Periodic boundary conditions were used in all simulations, electrostatics were computed every time step using the particle mesh Ewald method, and van der Waals interactions were cut-off at 10 Å with switching initiated at 8 Å. In these simulations, phosphorus atoms in the RNA backbone were weakly restrained to prevent the overall rotation and translation of the RNA molecule. Configurations were saved every picosecond and SMD output was saved every 20 fs.

B. SMD Simulations and Potential of Mean Force (PMF) Calculation

In SMD simulations, an external force is applied to one atom or a group of atoms to enhance conformational sampling in biophysical processses (e.g. ligand dissociation) that are difficult to observe in conventional MD simulations. The constant velocity SMD (cv-SMD) simulations were implemented by applying a harmonic external force using a spring with a spring constant of k = 7 kcal mol^{-1} Å^{-2} that was attached to the center of mass of the ligand and was pulled at a constant velocity of 0.0125 Å/ps along the reaction coordinate ξ. The force constant value was chosen per stiff-spring approximation[12] to closely follow the reaction coordinate for ligand dissociation. The velocity was set relatively slower than is commonly used in SMD simulations, in order to capture the near-equilibrium dynamics of ligand unbinding.[13,14] The potential of mean force (PMF) is calculated as a function of the distance along the reaction coordinate (RC) λ, which increases at a constant velocity v such that $\lambda_t = \lambda_0 + vt$, where $\lambda_0 = 0$ initially. The external work performed for a trajectory is then given by the equation:

$$W_{0\to t} = -kv \int_0^t (\xi - (\lambda_0 + vt))dt$$

We followed the protocol developed by Jensen et al.[15] and used both the exponential averaging as well as the second-order cumulant expansion of the Jarzynski's equality[16] to estimate the PMF along the reaction coordinate from work distributions obtained using cv-SMD simulations.[12,17] The exponential averaging expression is:

$$\Delta F = -\beta^{-1} \ln \overline{\exp(-\beta W)},$$

where $\beta = 1/k_B T$ with k_B the Boltzmann's constant and T the temperature. W is the non-equilibrium work performed and ΔF is the equilibrium free energy difference. The expression for second-order cumulant expansion is:

$$\Delta F = \langle \overline{W} \rangle - \frac{1}{2}\beta(\langle \overline{W}^2 \rangle - \langle \overline{W} \rangle^2),$$

where overbars represent averages over time windows of 1 ps, and angle brackets denote ensemble averages over independent cv-SMD simulations.

C. Buried Surface Area (BSA)

We calculated the BSA for the ligand acetylpromazine from a 2 μs conventional MD simulation. The BSA was computed using the following equation:

$$\mathrm{BSA} = \mathrm{SASA}_{\mathrm{RNA}} + \mathrm{SASA}_{\mathrm{Ligand}} - \mathrm{SASA}_{\mathrm{Complex}}$$

where $\mathrm{SASA}_{\mathrm{RNA}}$ represents the solvent accessible surface area (SASA) of RNA, $\mathrm{SASA}_{\mathrm{Ligand}}$ represents the SASA of ligand, and $\mathrm{SASA}_{\mathrm{Complex}}$ represents the SASA of the RNA/ligand complex. The BSA value indicates the area of contact between the ligand and RNA. Data are shown in Fig. S2.

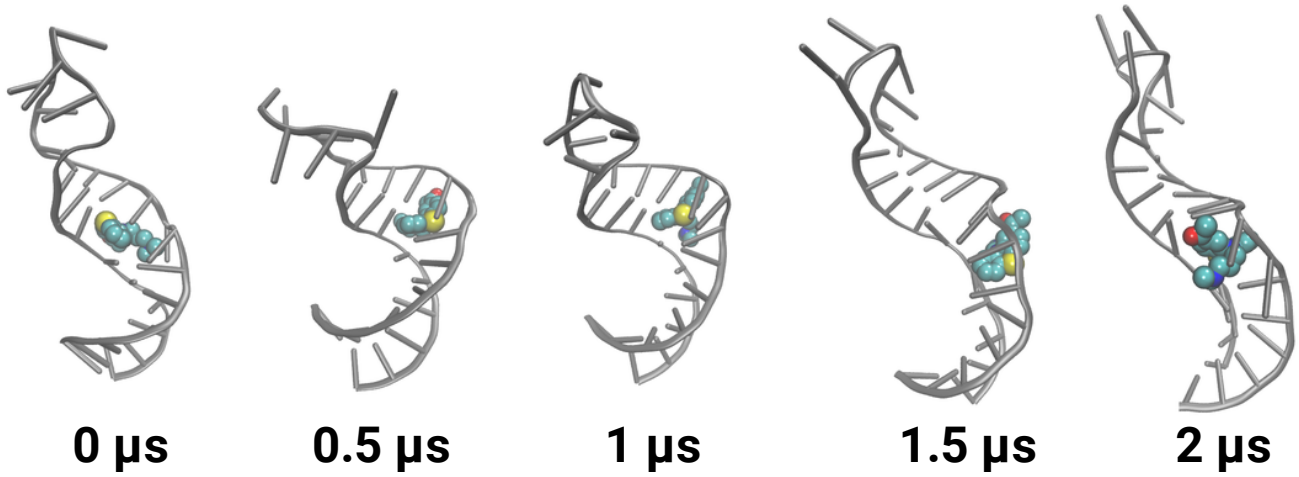

Figure S1. Long time-scale conventional MD simulation: Snapshots of RNA (gray cartoon) and ligand (space-filling) are shown at various timepoints from a 2 μs long conventional MD simulation, where ligand remains stably bound.

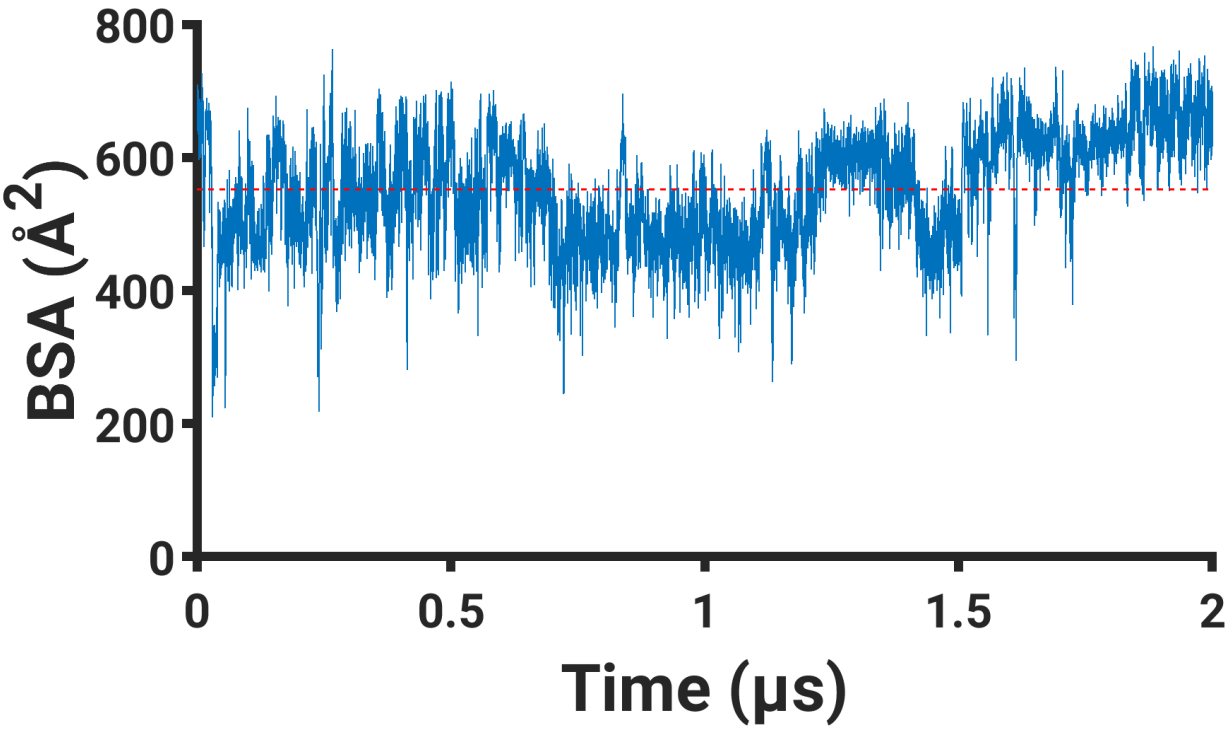

Figure S2. The buried surface area (BSA) trace *vs.* time is shown for the ligand from a 2 μs long conventional MD simulation (cf. Fig. S1). The dotted red line corresponds to the average BSA.

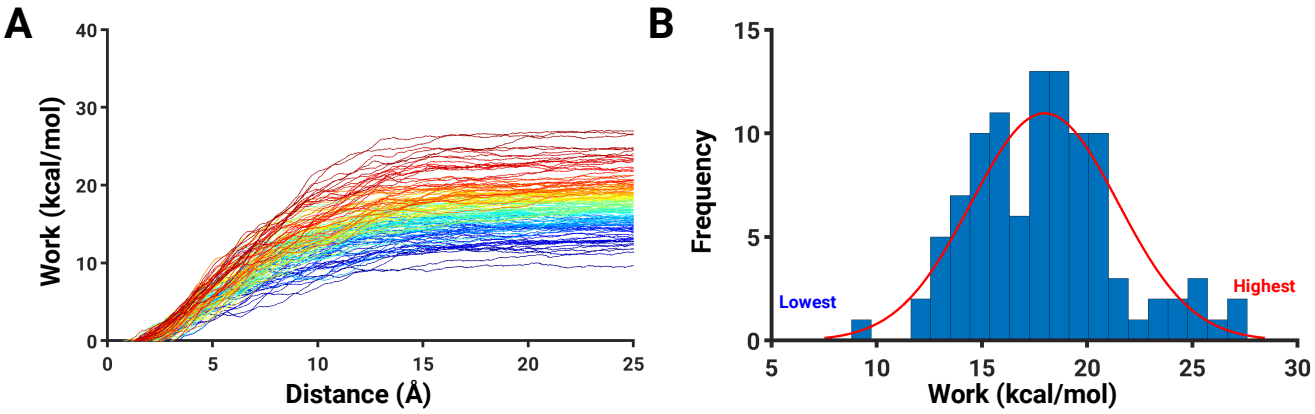

Figure S3. Ligand dissociation work from SMD simulations: (A) work values are plotted along the reaction coordinate (RC) from 102 independent cv-SMD simulations. Blue to red color palette indicates lower to higher values of work required for ligand dissociation. (B) A histogram of all work values (at 25 Å) is shown with a best-fit distribution line (red trace).

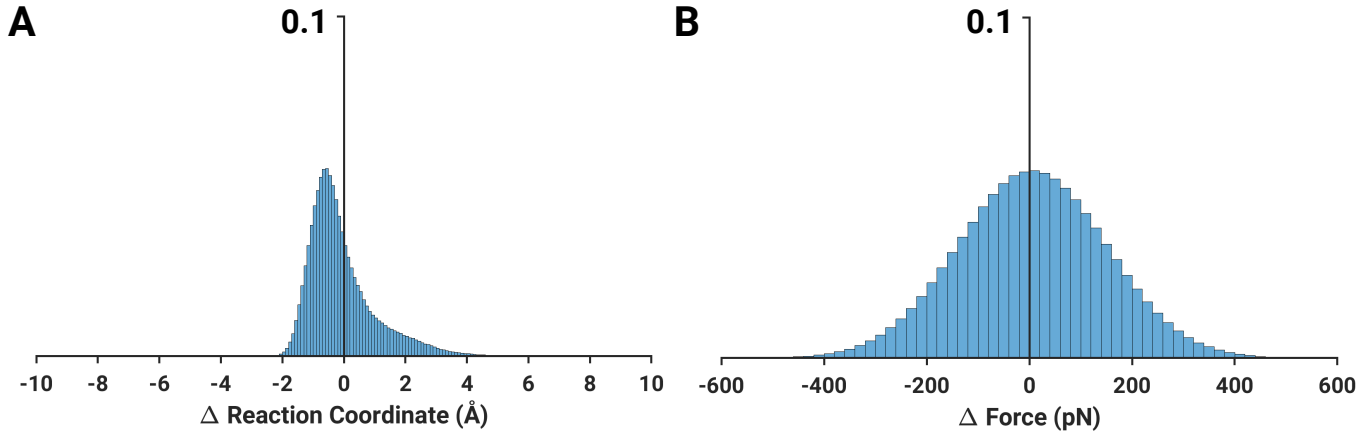

Figure S4. Reaction coordinate and force-convergence data: Shown are (A) distributions of ΔRC values computed from the deviations of pathways from cv-SMD simulations with respect to the actual reaction coordinate. (B) distributions of ΔForce values after the force on average converges to 0 (at ∼17.5 Å along the reaction coordinate; cf. Fig. 2B).

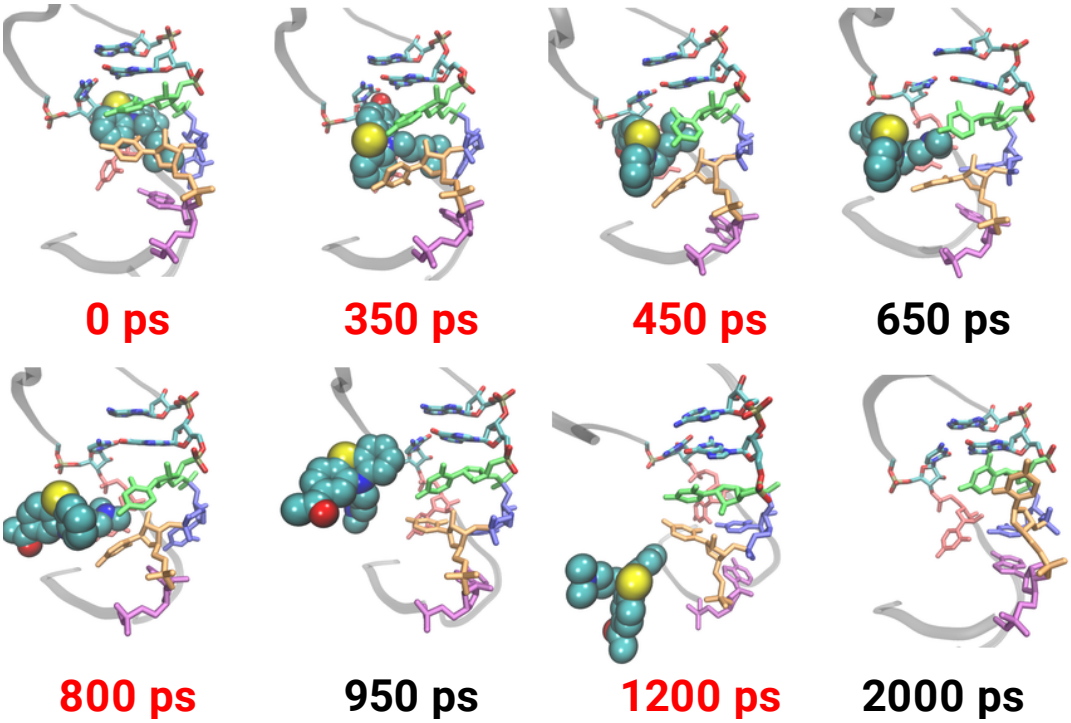

Figure S5. Ligand dissociation in the lowest work simulation: Snapshots of the ligand dissociation process at various timepoints are shown. The time values highlighted in red are those snapshots that are also shown in Fig. 3. Key nucleotides are highlighted: A22 (purple), U23 (orange), C24 (blue), U25 (green) and U40 (red). Ligand is shown in a space-filling representation.

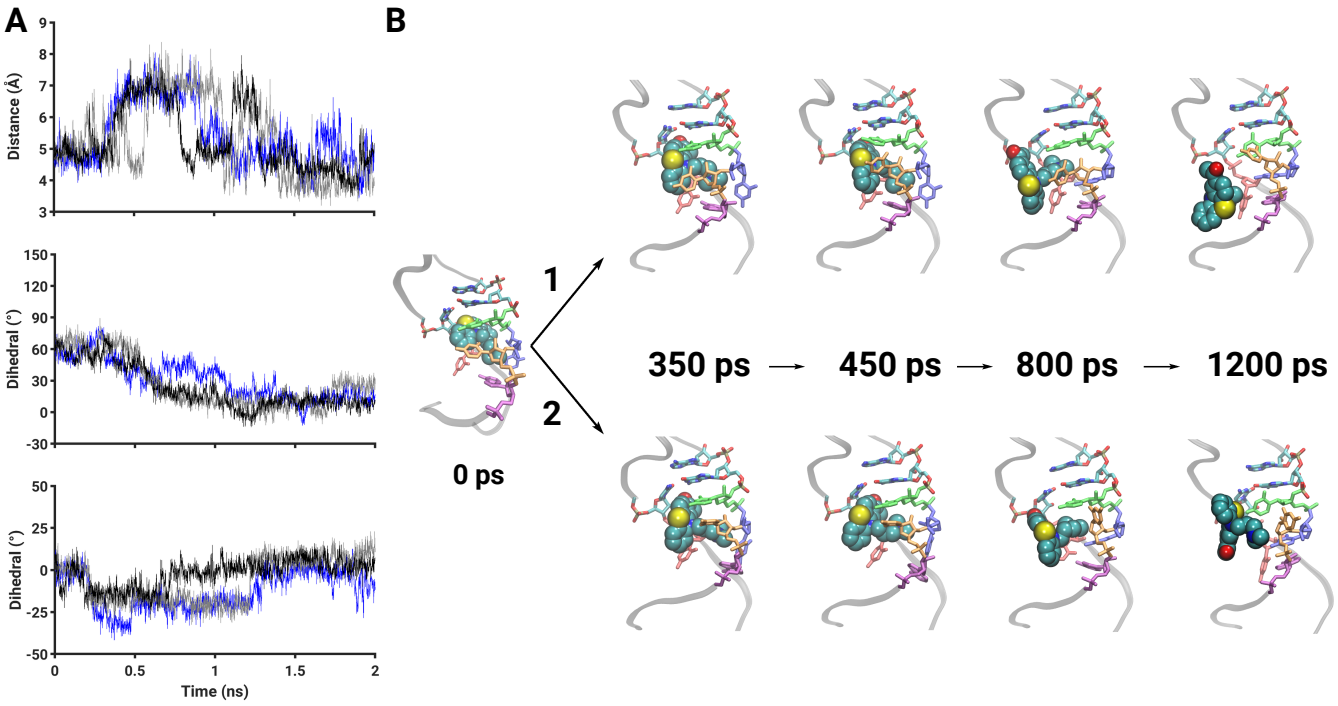

Figure S6. Conformational metrics and ligand dissociation mechanism from two additional simulations with lower work values: (A) Same metrics as in Figure 4 are presented. The traces correspond to the lowest work simulation (blue trace with W = 9.56 kcal/mol) and two additional simulations with the next lower values of work (black and gray traces with W = 12.59 and 12.76 kcal/mol, respectively). (B) Snapshots of ligand dissociation from two additional lower work simulations are shown. Color and labeling scheme is same as in Figs. 1 and 3.

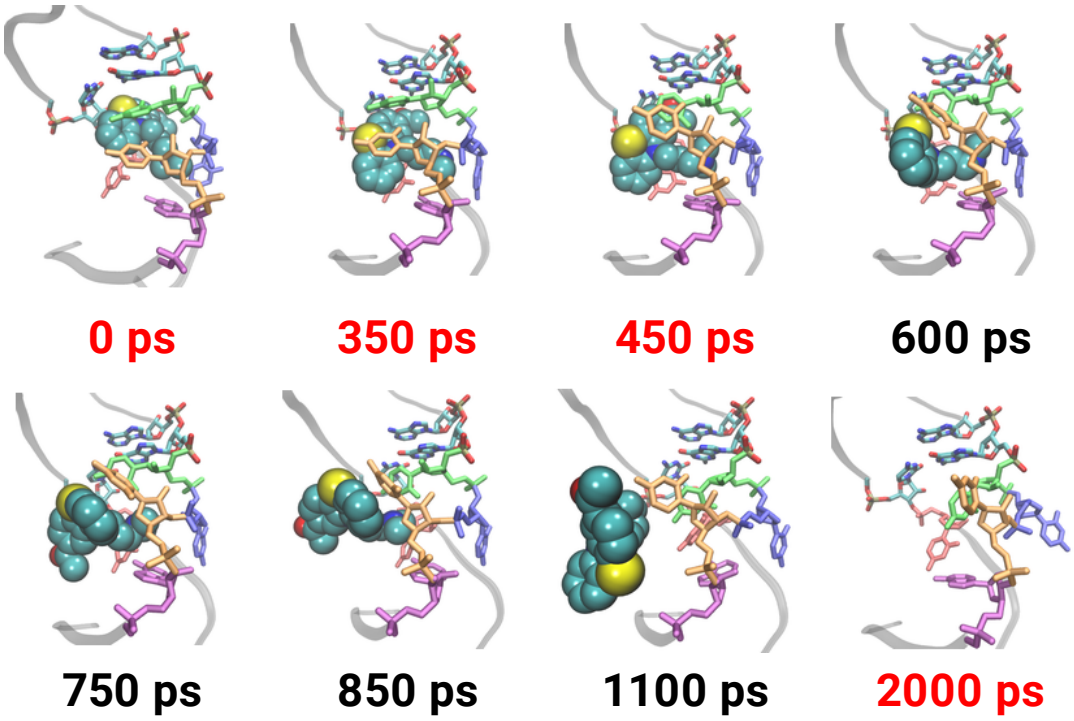

Figure S7. Ligand dissociation in the highest work simulation: Snapshots of the ligand dissociation process at various timepoints are shown. cf. Fig. 3 and Fig. S5 for other details.

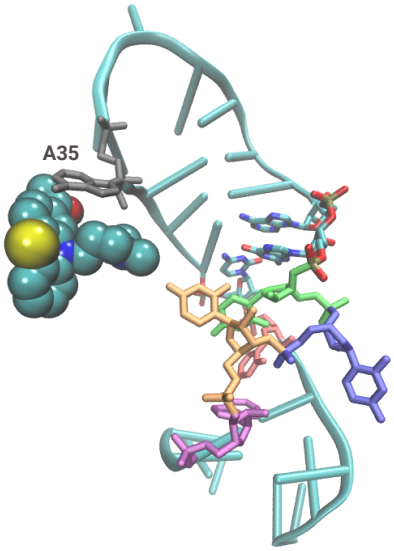

Figure S8. A side-view snapshot of the ligand interacting with the A35 nucleotide (gray sticks) is shown. Other nucleotides shown are the same as in Fig. 1C and Fig. 3.

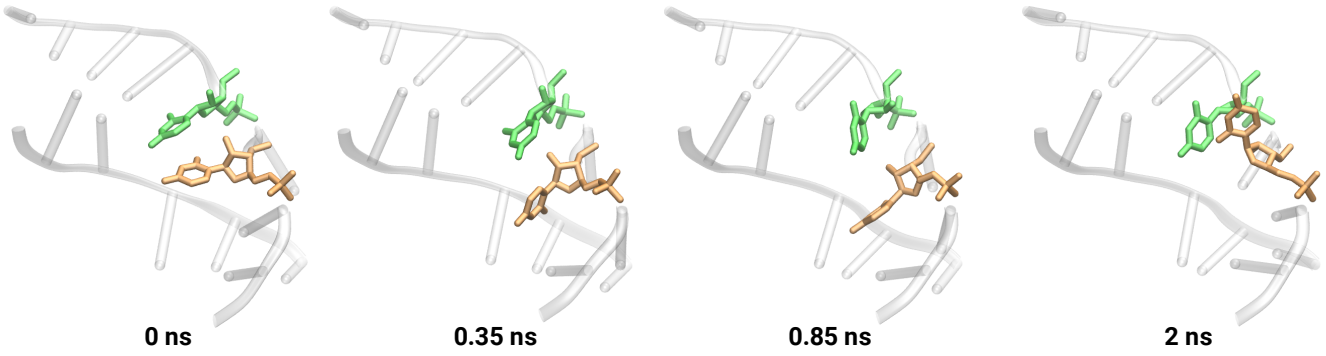

Figure S9. Shown are representative snapshots of TAR-RNA from the lowest work simulation highlighting nucleotides U23 (orange) and U25 (green).

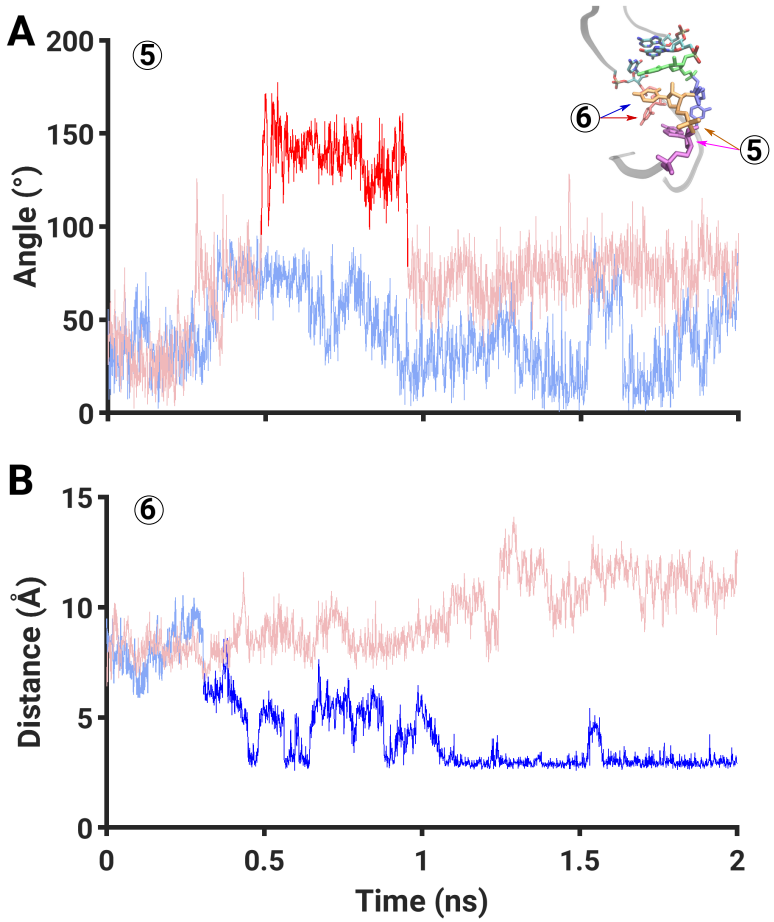

Figure S10. Additional conformational metrics: Shown are traces of additional conformational metrics from the lowest work (blue) and the highest work (red) simulations: (A) an interplane angle between A22 and U23 (marked as 5 in the *inset* in panel A and describing the relative position of U23); and (B) a hydrogen bond distance between C24 and U40 (marked as 6 in the *inset* in panel A). Darker colors signify transition regions of interest; cf. Fig. 4 for other details.

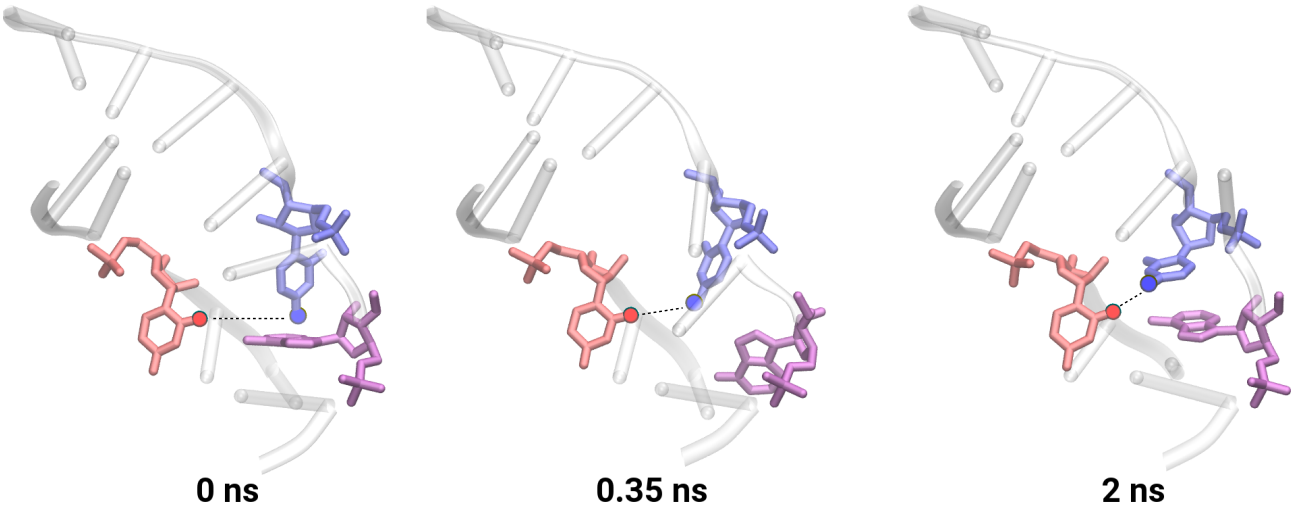

Figure S11. Representative snapshots from the lowest work simulation highlighting A22 (purple), C24 (blue), and U40 (red). The dotted line indicates a hydrogen bond between C24 and U40.

References

1. Du, Z.; Lind, K. E.; James, T. L. Structure of TAR RNA complexed with a Tat-TAR interaction nanomolar inhibitor that was identified by computational screening. *Chem. Biol.* **2002**, *9*, 707–712.
2. Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys. J.* **2007**, *92*, 3817–3829.
3. Zgarbova, M.; Otyepka, M.; Sponer, J.; Mladek, A.; Banas, P.; Cheatham, T. E.; Jurecka, P. Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J. Chem. Theory Comput.* **2011**, *7*, 2886–2902.
4. Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.* **2013**, *3*, 198–210.
5. Case, D. et al. AMBER 2018, University of California, San Francisco.
6. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
7. Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.
8. Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Molec. Graphics* **1996**, *14*, 33–38.
9. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247–260.
10. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
11. Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623–1641.
12. Park, S.; Schulten, K. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.* **2004**, *120*, 5946–5961.
13. Jin, H.; Zhu, J.; Dong, Y.; Han, W. Exploring the different ligand escape pathways in acylaminoacyl peptidase by random acceleration and steered molecular dynamics simulations. *RSC Adv.* **2016**, *6*, 10987–10996.
14. Yang, L.-J.; Zou, J.; Xie, H.-Z.; Li, L.-L.; Wei, Y.-Q.; Yang, S.-Y. Steered molecular dynamics simulations reveal the likelier dissociation pathway of imatinib from its targeting kinases c-Kit and Abl. *PLoS One* **2009**, *4*, e8470.
15. Jensen, M.; Park, S.; Tajkhorshid, E.; Schulten, K. Energetics of glycerol conduction through aquaglyceroporin GlpF. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6731–6736.
16. Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **1997**, *78*, 2690–2694.
17. Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J. Chem. Phys.* **2003**, *119*, 3559–3566.