

Molecular mechanism of processive 3' to 5' RNA translocation in the active subunit of the RNA exosome complex

Lela Vuković^{1,2,3,4*}, Christophe Chipot^{1,3,5}, Debora L. Makino^{6#}, Elena Conti⁶, and Klaus Schulten^{1,2,3*}

¹Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States

²Center for the Physics of Living Cells, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States

³Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States

⁴Department of Chemistry, University of Texas at El Paso, El Paso, TX 79968, United States

⁵Laboratoire International Associé CNRS-University of Illinois, Université de Lorraine, Vandoeuvre-lès-Nancy, France

⁶Department of Structural Cell Biology, Max Planck Institute of Biochemistry, Martinsried, Germany

* Authors to whom all correspondence should be addressed: kschulte@ks.uiuc.edu, Lvukovic@utep.edu

Present address: CRELUX GmbH, 82152 Martinsried, Germany

Electronic Supporting Information (ESI)

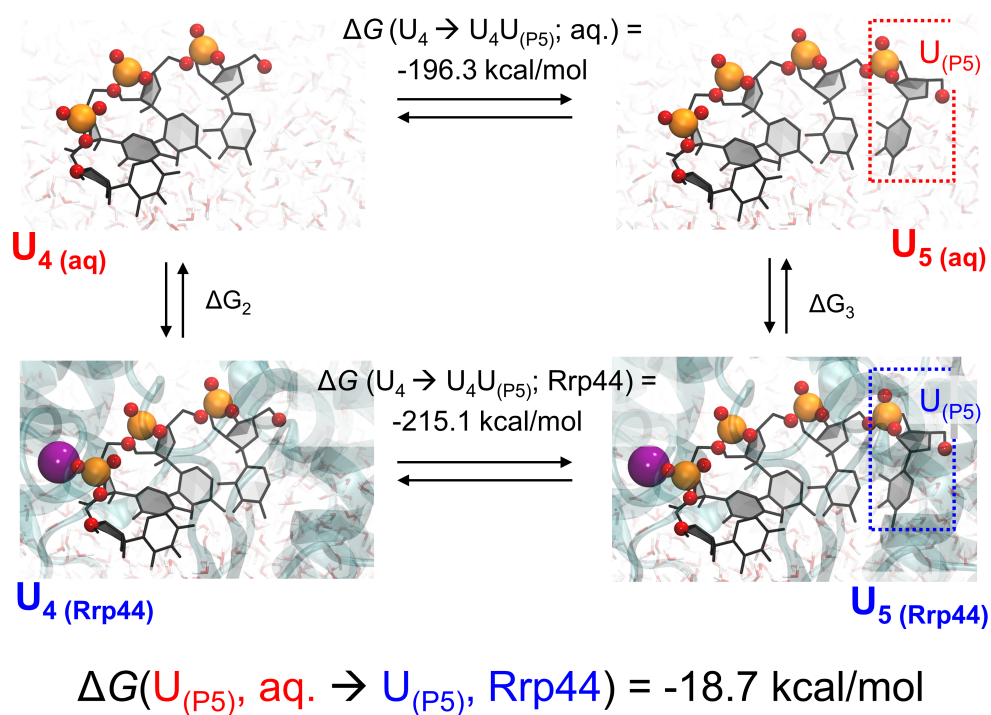


Figure S1: Thermodynamic cycle used to determine transfer free energies of single nucleotides. The representative cycle used in FEP calculations is shown for the addition of the fifth nucleotide, labeled $U_{(P5)}$, to a 4-nt RNA segment, U_4 . The heavy atoms of RNA are shown in licorice in black, and P, O and Mg atoms are shown in orange, red and purple, respectively, as vdW spheres. Hydrogen atoms on RNA are not shown for clarity.

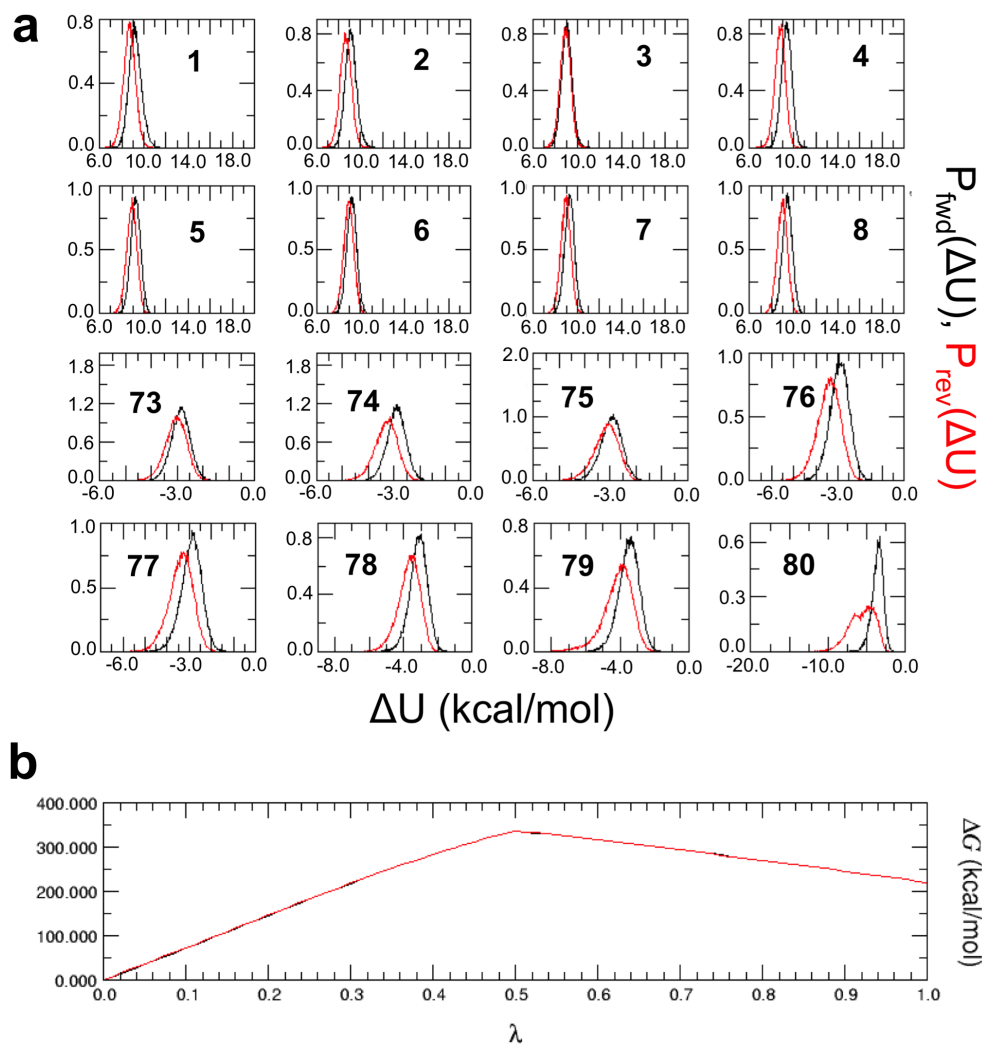


Figure S2: (a) Selected probability distribution functions for representative FEP transformations of a single RNA nucleotide to a terminal hydrogen. The transformation involves the 5'-end nucleotide in position P₅ of the Rrp44 exonuclease active site tunnel, as shown in the two bottom insets in Fig. S1. The figure shows distributions in the first eight and the last eight intermediate states of the 80-window FEP stratification strategy, described in the Methods section. The forward distributions are shown in red, and the backwards distributions are shown in black. (b) Free energy change as a function of λ -window value in FEP calculations (as defined in Methods) in forward (red) and backward (black) transformations.

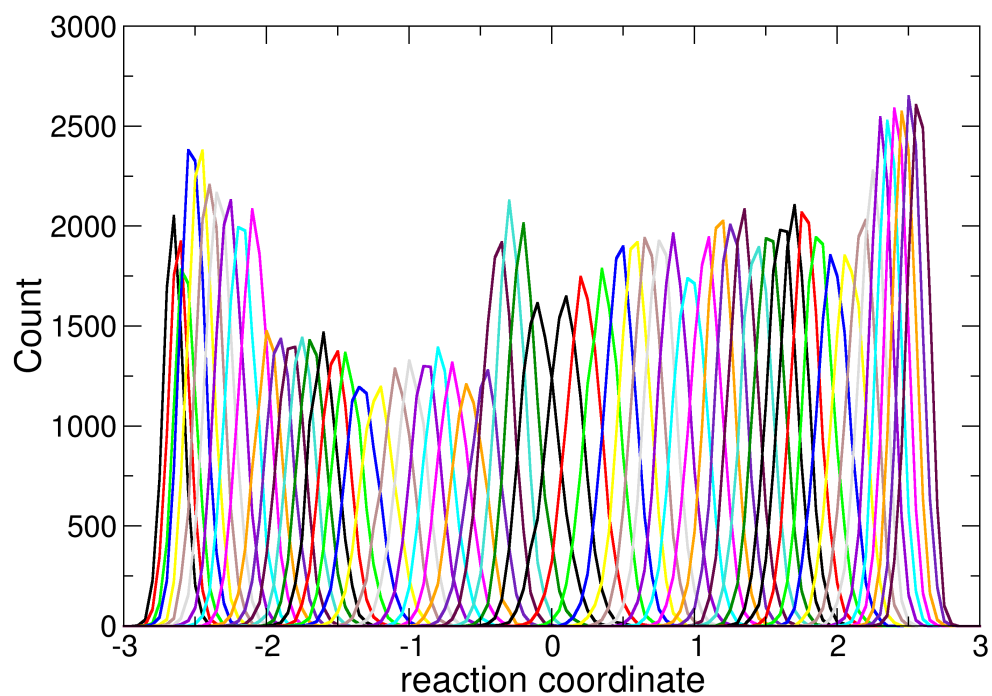


Figure S3: Histograms for 60 umbrella sampling windows, used to reconstruct the potential of mean force for a representative case of RNA translocation. The reaction coordinate is defined in the Methods section, and histograms are obtained from complete REMD-US trajectories.

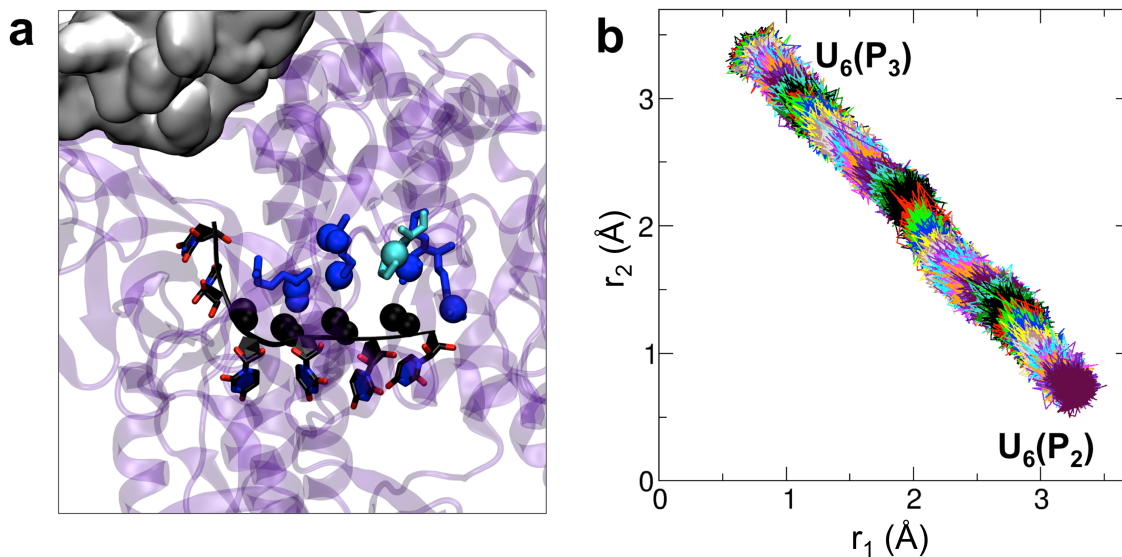


Figure S4: Reaction coordinate in REMD-US simulations. a) The atoms of the system which define the RMSD reaction coordinate for REMD-US simulations are shown as van der Waals spheres. The remaining heavy atoms of residues that form a reaction coordinate are shown in licorice representation for clarity. The environment proteins are shown in transparent purple (Rrp44), and in grey surface (exosome core). b) Trajectories from 60 REMD-US windows, projected onto two RMSD coordinates. The coordinate r_1 is defined as $r_1 = \langle (\vec{r}_i - \vec{r}_{i,initial})^2 \rangle^{1/2}$, where \vec{r}_i is the coordinate vector of atom i , $\vec{r}_{i,initial}$ is the coordinate vector of atom i in the initial state (when ssRNA has its 3'-end at P_3), and the averages are taken over the set of atoms shown in panel a. The coordinate r_2 is defined as $r_2 = \langle (\vec{r}_i - \vec{r}_{i,final})^2 \rangle^{1/2}$, where \vec{r}_i is the coordinate vector of atom i , $\vec{r}_{i,final}$ is the coordinate vector of atom i in the final state (when ssRNA has its 3'-end at P_2), and the averages are taken over the set of atoms shown in panel a. Each of the trajectories from 60 REMD-US windows is shown in different color; some of the colors are repeated due to a large number of windows, however, these trajectories do not overlap in the plot and can be distinguished.

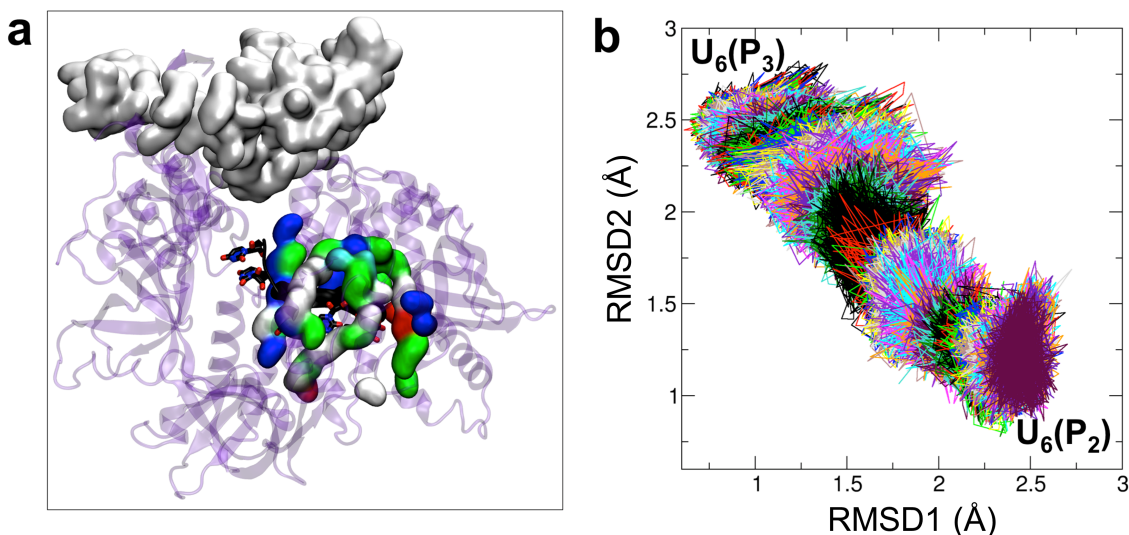


Figure S5: Translocation reaction in REMD-US simulations. a) The atoms of the system that define the RMSD coordinates examined in panel (b). The atoms are selected to test whether translocation occurs for the reaction coordinate chosen in REMD-US simulations. The selected atoms include heavy atoms of RNA segment U₆, and the Rrp44 active site tunnel atoms shown in surface representation. Different colors mark atoms of basic (blue), acidic (red), polar (green) and nonpolar (white) protein residues. The environment proteins are shown in transparent purple (Rrp44), and in grey surface (exosome core). b) Trajectories from 60 REMD-US windows, projected onto two RMSD coordinates, RMSD1 and RMSD2. The coordinate RMSD1 is defined as $\text{RMSD1} = \langle (\vec{r}_i - \vec{r}_{i,\text{initial}})^2 \rangle^{1/2}$, where \vec{r}_i is the coordinate vector of atom i , $\vec{r}_{i,\text{initial}}$ is the coordinate vector of atom i in the initial state (when ssRNA has its 3'-end at P₃), and the averages are taken over the set of atoms shown in panel a. The coordinate RMSD2 is defined as $\text{RMSD2} = \langle (\vec{r}_i - \vec{r}_{i,\text{final}})^2 \rangle^{1/2}$, where \vec{r}_i is the coordinate vector of atom i , $\vec{r}_{i,\text{final}}$ is the coordinate vector of atom i in the final state (when ssRNA has its 3'-end at P₂), and the averages are taken over the set of atoms shown in panel a. The trajectories are distinguished by color as in case of Fig. S4 (b).

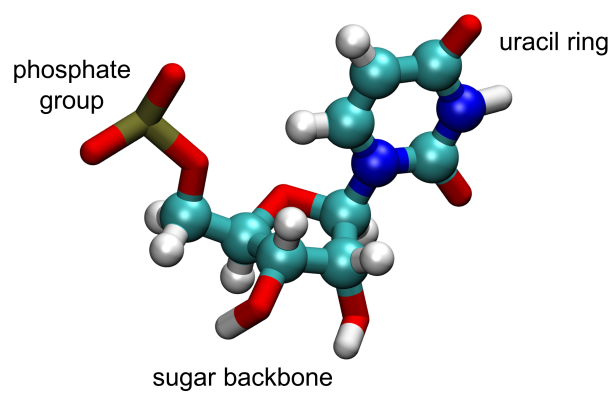


Figure S6: Definition of nonpolar atoms (van der Waals spheres) of a uracil nucleotide (licorice representation) in calculations of contact areas reported in Fig. 6 (c-d).

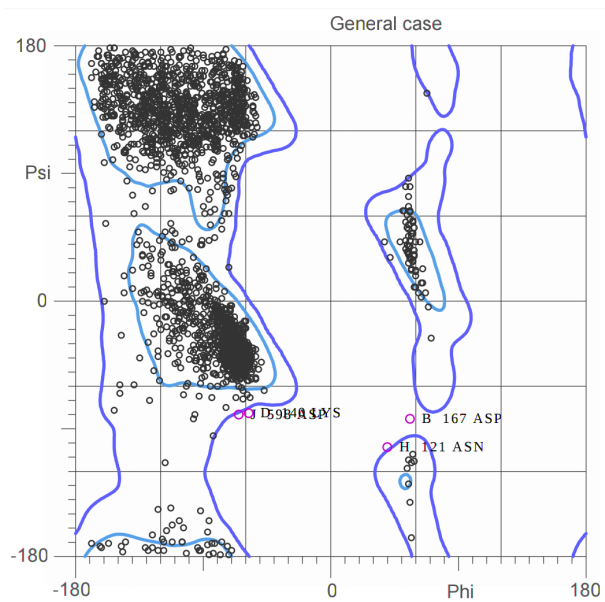


Figure S7: MolProbity [1] Ramachandran analysis for the modeled structure (pdbID: 4IFD). The plot is the output of the analysis performed on the MolProbity web server.

Table S1: Summary of results of FEP calculations.

Nucleotide/System	$G_{forward}$	G_{back}	G_{BAR}	$G_{BAR,corr}^a$	$G_{BAR,error}$	hysteresis	$G_{transfer}^b$
P ₅ , 5'-end, (U ₅ → U ₄) in water	205.07	-203.69	204.25	196.30	0.11	1.38	n/a
P ₁ , 3'-end, (U ₆ → U ₅) in Rrp44	215.12	-214.92	215.16	211.75	0.21	0.20	-15.00
P ₂ , 3'-end, (U ₆ → U ₅) in Rrp44	211.29	-207.88	208.93	205.52	0.17	3.41	-8.77
P ₃ , 3'-end, (U ₆ → U ₅) in Rrp44	214.74	-213.61	213.64	210.23	0.17	1.13	-13.48
P ₄ , 3'-end, (U ₆ → U ₅) in Rrp44	217.29	-216.54	216.75	213.34	0.16	0.75	-16.59
P ₅ , 5'-end, (U ₅ → U ₄) in Rrp44	219.67	-218.00	218.46	215.05	0.19	1.67	-18.74
P ₆ , 5'-end, (U ₆ → U ₅) in Rrp44	214.72	-213.11	213.22	209.81	0.19	1.61	-13.06
P ₇ , 5'-end, (U ₇ → U ₆) in Rrp44	204.69	-204.77	205.22	201.81	0.29	0.08	-5.50
P ₈ , 5'-end, (U ₈ → U ₇) in Rrp44	205.57	-205.62	205.19	202.19	0.14	-0.05	-5.89
P ₉ , 5'-end, (U ₉ → U ₈) in Rrp44	207.33	-206.16	206.21	202.80	0.16	1.17	-6.05

^a $G_{BAR,corr}$ is obtained by subtracting from G_{BAR} the analytical correction due to presence of a net charge in systems examined with FEP calculations [2] after creation or annihilation of single nucleotides. The correction is 7.95 kcal/mol for the system containing ssRNA in the water box and 3.41 kcal/mol for systems of ssRNA within Rrp44.

$$^b G_{transfer} = G_{BAR,corr}(water) - G_{BAR,corr}(Rrp44)$$

References

- [1] V. B. Chen, W. B. Arendall, III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, and D. C. Richardson. *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D*, 66:12–21, 2010.
- [2] J. Gumbart and B. Roux. Determination of membrane-insertion free energies by molecular dynamics simulations. *Biophys. J.*, 102:795–801, 2012.