

A Two Metal-Ion-Mediated Conformational Switching Pathway for HDV Ribozyme Activation

Tai-Sung Lee,[†] Brian K. Radak,^{†,‡} Michael E. Harris,[¶] and Darrin M. York^{*,†}

*Center for Integrative Proteomics Research and Department of Chemistry and Chemical
Biology, Rutgers University, Piscataway, NJ 08854, USA, Department of Chemistry,
University of Minnesota, Minneapolis, MN 55455, USA, and Department of Biochemistry,
Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA*

E-mail: Darrin.York@rutgers.edu

1 Simulation setup and protocols

This work utilizes the same protocol employed in our previous simulations.¹ In summary, more than thirty long time (≥ 300 ns each and roughly 10 μ s in total) molecular dynamics simulations at different stages along the HDVr reaction coordinate were performed with

*To whom correspondence should be addressed

[†]Center for Integrative Proteomics Research and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854, USA

[‡]Department of Chemistry, University of Minnesota, Minneapolis, MN 55455, USA

[¶]Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA

different starting conditions: the G25:U20 base pair in either its tWW or tWH forms² (Fig.1 in the main text), C75 protonated or neutral, and Mg^{2+} ion(s) at different binding modes.

1.1 Simulation protocols:

Each simulation was performed in a cubic cell of $60 \times 60 \times 120 \text{ \AA}^3$ filled with pre-equilibrated TIP3P waters³ with HDVr located at the center. The ion atmosphere consisted of Na^+ and Cl^- ions that were added to neutralize the system and reach the physiologic (extracellular) concentration of 0.14 M. The resulting system (the reactant state) contained 54,818 atoms: 16,914 water molecules, 2 Mg^{2+} , 102 Na^+ , 43 Cl^- , 1,582 protein atoms, and 2,347 RNA atoms.

Simulations were performed with the NAMD simulation package (version 2.7b3)⁴ using the AMBER^{5,6} parm99 force field with the α/γ corrections for nucleic acids.⁷ The parameters of Mg^{2+2+} optimized for RNA were utilized.⁸

Periodic boundary conditions were used along with the isothermal-isobaric ensemble (*NPT*) at 1 atm and 298 K using extended system pressure algorithm⁹ with effective mass of 500.0 *amu* and Nosé-Hoover thermostat^{10,11} with effective mass of 1000.0 kcal/mol-ps², respectively. The smooth particle mesh Ewald (PME) method^{12,13} was employed with a B-spline interpolation order of 6, the default κ value (0.258 \AA^{-1}), and 60, 60, and 120 FFT grid points were used for the x, y, and z cubic lattice directions, respectively. Non-bonded interactions were treated using an atom-based cutoff of 12 \AA with switching of the non-bonded potential beginning at 10 \AA . Numerical integration was performed using the leap-frog Verlet algorithm with 1 fs time step.¹⁴ Covalent bond lengths involving hydrogen were constrained using the SHAKE algorithm.¹⁵ For each simulation, 10 ns of water/ion equilibration followed by additional 10 ns of solute equilibration were performed before 350 ns of production simulation.

1.2 Initial structure:

The initial structure used in the reactant simulations was based on a 2.45 Å crystal structure with two resolved Sr^{2+} ions (PDB: 1VC7).¹⁶ The Sr^{2+} ions were replaced by native Mg^{2+} ions in the simulations. The C75U mutation was restored back to C75. The A-2 residue was removed.

1.3 Mg^{2+} binding sites:

It has been suggested that there are two possible Mg^{2+} binding sites, one near G1:N₇¹⁷⁻²¹ and another near the active site phosphate (possibly bound to U75:O₆ in the case of C75U mutants)¹⁷ or near G25:N₇ in other structures.^{17,22,23} In accordance with the previously reported notation,¹⁷ the binding site near G1:N₇ is defined as the “C-site” and the site near U75:O₆/G25:N₇ as the “B-site”. In all simulations reported here, the B-site initial position was either taken directly from the crystal structure (PDB: 1VC7) when G25:U20 was in the tWH form, or modeled by placing a Mg^{2+} ion 4.0 Å from G25:N₇ when G25:U20 was in its reverse wobble form. The C-site initial position was at 2.0 Å from G1:N₇, as suggested in other work.^{17,19}

1.4 G25:U20 base pairing modes:

There are two types of G25:U20 base pair conformations in the reported crystal structures (Fig.1 in the main text): *trans* Watson-Crick/Hoogsteen (tWH)^{16,17} and the unusual *trans* Watson-Crick/Watson-Crick wobble (tWW, also referred to as a reverse wobble).^{24-27,27} In order to explore these two possible G25:U20 base pairings, initial structures with different G25:U20 conformations were used. The initial starting positions of all other residues were the same as in our previous work,¹ which was derived from a C75U crystal structure (1VC7). To our knowledge, there is no fully resolved U-1 residue in any reported structure with G25:U20 in the tWW form. Hence, the following procedure was taken to build the initial structures

with G25:U20 in the tWW form. After 10 ns of solvent and ion equilibration, the structures with G25:U20 in the tWH form were modified by rotating the G25 glycosidic bond torsion ($O_{4'}-C_{1'}-N_9-C_4$) to 67.5° . In this *syn*-conformation two hydrogen bonds were nearly formed in the expected tWW pattern (G25:N₁-U20:O₄ and G25:O₆-U20:N₃). This was followed by an additional 1 ns equilibration of solvent and ions. The resulting structures underwent further relaxation with two hydrogen bonds (G25:N₁-U20:O₄ and G25:O₆-U20:N₃) restrained with 20 kcal/mol-Å² harmonic potentials for 10 ns to ensure the proper response and relaxation of the surrounding residues. The final structures then served as the starting structures for the reported MD simulations.

1.5 C75 protonation states and stages along the reaction coordinate:

Both C75 in its neutral and protonated forms have been used in the simulations as reported previously¹ and the same sets of parameters were employed here. Briefly, the parameters of deprotonated C75 were created by removing the O2' hydrogen followed by refitting of the charges of O2' oxygen and the nearby sugar ring atoms through RESP procedure.²⁸

1.6 Different stages along the reaction:

The initial structure of both the early (ETS) and late (LTS) transition state mimic were taken from the snapshot at 59.3 ns of the simulation of the deprotonated active reactant state. The initial structure of the product (Prod) simulation was taken from the snapshot at 130.0 ns of the simulation of the late transition state mimic. The residue U-1 was removed with additional 5 ns ion/water equilibration performed prior to production simulation.

2 Simulations performed

We use the following notations for Mg^{2+} binding sites: M_B stands for a Mg^{2+} placed at the B-site initially and M_C at the C-site. For G25:U20 base pairing modes, “tWH” stands for initial structures in the tWH pair form and “tWW” in the tWW (reverse wobble) form. For C75 protonation states, C75^0 stands for neutral C75 while C75^+ for protonated C75. Different stages of the reaction are marked as: reactant state (RT), activated precursor state (dRT), early transition state mimic (ETS), late transition state mimic (LTS), and product state (PD). All simulations are listed in the tables in the Results Section of the main text.

3 Result validation through repeated simulations:

To further validate our simulation results, four additional simulations were performed with the same protocol but re-generated water boxes and different initial velocity assignments through modifying the default initial random numbers, and were performed with newer version of NAMD (version 2.10) with GPU acceleration, which gives different force evaluation within the numerical error ranges.

The additional four simulations were

- tWw_RT_C75⁰: Shown in Table S1.
- tWW_RT_C75⁺M₂: Shown in Table S2.
- tWH_RT_C75⁺M_B: Shown in Table S2.
- tWH_dRT_C75⁺M_C: Shown in Table S3.

All simulations basically give the same results, compared to the simulations reported in the main text with the same conditions. The tWW_RT_C75⁺M₂ in Table S2 is a new simulation, where two Mg^{2+} ions were added at both B- and C- site initially. Its results support the conclusion in the main text that the B-site Mg^{2+} cannot stay when HDVr is in the tWW (reversed wobble) conformation.

Table S1: Key geometric parameters for B-site formation in two independent reactant state simulations (tWW_RT_C75⁰). The results were to verify the simulations in Table 1.

Simulation	Hydrogen Bond			G25: χ	Distance to C75:N ₃			Distance to Mg ²⁺ (B)				Distance to Mg ²⁺ (C)	
	G25:N ₁ U20:O ₄	U20:N ₃ G25:O ₆	U20:O ₂ G25:N ₁		G1:O _{1P}	G1:O _{2P}	$< \frac{1}{r_1} + \frac{1}{r_2} >$	U-1:O _{2'}	G1:O _{2P}	G25:N ₇	U20:O ₂	U-1:O _{2'}	G1:N ₇
tWW_RT_C75 ⁰ First run	19.9%	20.7%	16.0%	-169.2 (11.3)	5.54 (0.68)	4.84 (0.63)	0.39						
tWW_RT_C75 ⁰ Second run	27.3%	29.0%	18.3%	-169.4 (10.5)	5.39 (0.65)	4.79 (0.61)	0.39						

Notations used: Initial G25:U20 base pairing: tWW (reverse wobble) and tWH (Fig. 2); M_B : Mg²⁺ initially placed at the B-site (near G25:N₇); M_C: Mg²⁺ initially placed at the C-site (near G1:N₇); C75⁰ : C75 is neutral; RT: reactant state.

Entries are calculated from sampling of the range 50 to 300 ns from the 300 ns trajectory for each simulation condition with a sampling free frequency of 100 ps, resulting 2500 data points for every simulation. Distance units are in Å, and hydrogen bond occupations (listed as percentages) are defined as the percentage of the snapshots in which the distance between the proton donor and the proton acceptor $r \leq 3.5$ Å and the H-bond angle $\theta \geq 150^\circ$. The numbers in parenthesis are the standard deviation of sampling.

Table S2: Key geometric parameters for B-site formation in various redundant independent C75 protonated reactant state simulations. The results were to verify the simulations in Table 2.

Simulation	Hydrogen Bond			G25: χ	Distance to Mg ²⁺ (B)				Distance to Mg ²⁺ (C)	
	G25:N ₁ U20:O ₄	U20:N ₃ G25:O ₆	U20:O ₂ G25:N ₁		U-1:O _{2'}	G1:O _{2P}	G25:N ₇	U20:O ₂	U-1:O _{2'}	G1:N ₇
tWW_RT_C75 ⁺ M _B	0.3%	15.0%	8.4%	65.5 (34.3)	7.83 (0.28)	4.04 (0.26)	6.03 (1.31)	5.39 (0.93)		
tWW_RT_C75 ⁺ M _C	9.3%	6.5%	28.2%	120.7 (61.9)					4.09 (0.26)	2.23 (0.11)
tWW_RT_C75 ⁺ M ₂	5.9%	1.6%	31.6%	132.3 (88.1)					11.15 (4.46)	6.88 (2.35)
tWH_RT_C75 ⁺ M _B First run				-65.9 (17.0)	7.99 (0.21)	4.31 (0.26)	7.43 (1.05)	4.17 (0.36)		
tWH_RT_C75 ⁺ M _B Second run				-42.9 (8.5)	6.29 (0.75)	4.20 (0.21)	6.64 (0.28)	4.13 (0.20)		

Notations used: Initial G25:U20 base pairing: tWW (reverse wobble) and tWH (Fig. 2); M_B : Mg²⁺ initially placed at the B-site (near G25:N₇); M_C: Mg²⁺ initially placed at the C-site (near G1:N₇); M₂: two Mg²⁺ initially placed at both the B-site and C-site; C75⁺ : C75 is protonated and positively changed; C75⁰ : RT: reactant state.

Entries are calculated from sampling of the range 50 to 300 ns from the 300 ns trajectory for each simulation condition with a sampling free frequency of 100 ps, resulting 2500 data points for every simulation. Distance units are in Å, and hydrogen bond occupations (listed as percentages) are defined as as the percentage of the snapshots in which the distance between the proton donor and the proton acceptor $r \leq 3.5$ Å and the H-bond angle $\theta \geq 150^\circ$. The numbers in parenthesis are the standard deviation of sampling.

Table S3: Key geometric parameters in two independent activated precursor state simulations (tWH_dRT_C75⁺M_C). The results were to verify the simulations in Table 3.

Simulation	Hydrogen Bond			G25:χ	Distance to C75:N ₃			Inline angle		Distance to Mg ²⁺ (B)				Distance to Mg ²⁺ (C)	
	G25:N ₁ U20:O ₄	U20:N ₃ G25:O ₆	U20:O ₂ G25:N ₁		G1:O _{1P}	G1:O _{2P}	G1:O _{5'}	Average	%(>120°)	U-1:O _{2'}	G1:O _{2P}	G25:N ₇	U20:O ₂	U-1:O _{2'}	G1:N ₇
tWH_dRT_C75 ⁺ M _C	84.1%	72.9%		-84.7 (10.3)	3.71 (0.79)	3.23 (0.44)	5.12 (0.30)	129.1 (18.0)	78.3%					7.19 (0.67)	2.21 (0.10)
tWH_dRT_C75 ⁺ M _C	81.3%	73.8%		-71.7 (15.4)	3.90 (0.77)	3.52 (0.34)	5.17 (0.39)	124.5 (17.4)	74.2%					7.32 (0.75)	2.18 (0.12)

Notations used: Initial G25:U20 base pairing: tWW (reverse wobble) and tWH (Fig. 2); M_B : Mg²⁺ initially placed at the B-site (near G25:N₇); M_C: Mg²⁺ initially placed at the C-site (near G1:N₇); C75⁺ : C75 is protonated and positively charged; C75⁰ : dRT: activated reactant state (the nucleophile (U-1:O_{2'}) is deprotonated).

Entries are calculated from sampling of the range 50 to 300 ns from the 300 ns trajectory for each simulation condition with a sampling free frequency of 100 ps, resulting 2500 data points for every simulation. Distance units are in Å, and hydrogen bond occupations (listed as percentages) are defined as as the percentage of the snapshots in which the distance between the proton donor and the proton acceptor $r \leq 3.5$ Å and the H-bond angle $\theta \geq 150^\circ$. The numbers in parenthesis are the standard deviation of sampling.

References

- (1) Lee, T.-S.; Giambasu, G. M.; Harris, M. E.; York, D. M. *J. Phys. Chem. Lett.* **2011**, *2*, 2538–2543.
- (2) Leontis, N. B.; Westhof, E. *RNA* **2001**, *7*, 499–512.
- (3) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (4) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kaleé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (5) Case, D. A. et al. *AMBER 7*; University of California San Francisco: San Francisco, 2002.
- (6) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. R.; Cheatham, III, T.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1–41.
- (7) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham III, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817–3829.
- (8) Mayaan, E.; Moser, A.; MacKerell, Jr., A. D.; York, D. M. *J. Comput. Chem.* **2007**, *28*, 495–507.
- (9) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384–2393.
- (10) Nosé, S.; Klein, M. L. *Mol. Phys.* **1983**, *50*, 1055–1076.
- (11) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.
- (12) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Hsing, L.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

- (13) Sagui, C.; Darden, T. A. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155–179.
- (14) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, 1987.
- (15) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (16) Ke, A.; Zhou, K.; Ding, F.; Cate, J. H. D.; Doudna, J. A. *Nature* **2004**, *429*, 201–205.
- (17) Ke, A.; Ding, F.; Batchelor, J. D.; Doudna, J. A. *Structure* **2007**, *15*, 281–287.
- (18) Gong, B.; Chen, J.-H.; Bevilacqua, P. C.; Golden, B. L.; Carey, P. R. *Biochemistry* **2009**, *48*, 11961–11970.
- (19) Chen, J.-H.; Gong, B.; Bevilacqua, P. C.; Carey, P. R.; Golden, B. L. *Biochemistry* **2009**, *48*, 1498–1507.
- (20) Cerrone-Szakal, A. L.; Chadalavada, D. M.; Golden, B. L.; Bevilacqua, P. C. *RNA* **2008**, *14*, 1746–1760.
- (21) Golden, B. L. *Biochemistry* **2011**, *50*, 9424–9433.
- (22) Veeraraghavan, N.; Ganguly, A.; Chen, J.-H.; Bevilacqua, P. C.; Hammes-Schiffer, S.; Golden, B. L. *Biochemistry* **2011**, *50*, 2672–2682.
- (23) Chen, J.; Ganguly, A.; Miswan, Z.; Hammes-Schiffer, S.; Bevilacqua, P. C.; Golden, B. L. *Biochemistry* **2013**, *52*, 557–567.
- (24) Perrotta, A. T.; Been, M. D. *Nucleic Acids Res.* **1990**, *18*, 6821–6827.
- (25) Wadkins, T. S.; Perrotta, A. T.; Ferré-D’Amaré, A. R.; Doudna, J. A.; Been, M. D. *RNA* **1999**, *5*, 720–727.
- (26) Chen, J.-H.; Yajima, R.; Chadalavada, D. M.; Chase, E.; Bevilacqua, P. C.; Golden, B. L. *Biochemistry* **2010**, *49*, 6508–6518.

- (27) Kapral, G. J.; Jain, S.; Noeske, J.; Doudna, J. A.; Richardson, D. C.; Richardson, J. S. *Nucleic Acids Res.* **2014**, *42*, 12833–12846.
- (28) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollmann, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620–31.