Supporting Information

Evidence for a Catalytic Strategy to Promote Nucleophile Activation in Metal-Dependent RNA-Cleaving Ribozymes and 8-17 DNAzyme

Abir Ganguly[†,‡], Benjamin P. Weissman[‖], Joseph A. Piccirilli[‖,§], and Darrin M. York[†,‡,¶,*]

[†]*Laboratory for Biomolecular Simulation Research, Rutgers, The State University of New Jersey, Piscataway, NJ 08854-8087, USA* [‡]*Institute for Quantitative Biomedicine, Rutgers, The State University of New Jersey, Piscataway, NJ 08854-8087, USA* [‖]*Department of Chemistry, The University of Chicago, Chicago, IL 60637, USA* [§]*Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL 60637, USA* [¶]*Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, NJ 08854-8087, USA*

* E-mail: Darrin.York@rutgers.edu

Methods

MD simulations on VSr, HHr, Psr, and 8-17dz performed in this study are based on crystal structures with PDB IDs 5V3I,[1] 2OEU,[2] 5K7C,[3] and 5XM8,[4] respectively. The simulations were performed in truncated octahedral boxes containing TIP4P-Ew water molecules,[5] having a buffer distance of 12Åwith the solute. Monovalent ions (Na+ and Cl-) were added to neutralize the boxes and provide them with ionic strengths of 150 mM. Simulations were run with a time step of 1 fs with full periodic boundary conditions. The smooth particle mesh Ewald (PME) method[6] with a 12Åcutoff was used to calculate long range electrostatic interactions and the SHAKE algorithm[7] was used to constrain all bonds involving hydrogen. Langevin dynamics with a collision frequency (γ) of 5.0 was employed to maintain constant temperature during simulations. For constant pressure and temperature simulations, the Berendsen barostat[8] with a pressure relaxation time of 1 ps and isotropic pressure scaling was used along with Langevin dynamics (γ=5.0). The putative general base guanines were kept deprotonated at their N1 positions. Parameters for nonstandard nucleobases were derived from Restrained Electrostatic Potential (RESP)[9,10] calculations and were adopted from previous studies.[11] For each system, the crystal coordinates are subjected to a series of equilibration simulations (described below) to obtain the respective active states in solution. The active state of VSr is reported in Ganguly et al., HHr is reported in Lee et al.,[12] Psr is reported in Kostenbader et al.,[13] and 8-17dz is reported in Ekesan et al.[14]

For each system, six independent sets of one dimensional umbrella sampling (US) simulations were performed; three sets of US simulations were performed in presence of the catalytic metal ion along the reaction coordinates $d_{HO2'-G}$, $d_{HO2'-proR_P}$, and $d_{HO2'-proS_P}$, respectively, where $d_{HO2'-G}$ coordinate corresponds to the Cartesian distance of the HO2′ from the WC edge of the catalytic guanine, and $d_{HO2'-proR_P}$ and $d_{HO2'-proS_P}$ coordinates correspond to the Cartesian distance of the HO2′ from the pro-R_P and pro-S_P NPOs of the scissile phosphate, respectively. Three additional sets of analogous US simulations were performed in absence of the catalytic metal ion. Each US simulation set was divided into 6

windows; in each window the given reaction coordinate were kept restrained to 1.8Å, while the strength of harmonic potentials along the windows were varied by changing the force constants according to (20,16,12,8,4,0). In some case additional windows with force constants (5,4,3,2,1) were inserted to ensure convergence. Force constants are expressed in units of kcal/mol. For each window, 1 ns of initial equilibration is performed followed by 20-40 ns of production MD. Thus, in each set of US simulations 120-240 ns of MD was performed, and consequently for each system typically 700 ns - 1 μs of MD was performed.

In order to construct the unbiased 2D and 1D PDFs for each system, the US simulation sets were jointly unbiased using the multistate Bennett acceptance ratio (MBAR) method[15] to obtain the relative free energies of the various windows (ΔA_ξ). The biasing potential associated with each biased configuration, $w(\xi)$, is calculated according to $\frac{1}{2}\sum_i k(\xi_i - \xi^0)^2$. Using the biasing potential and associated free energy, the unbiased configuration is obtained from the corresponding biased configuration, according to the equation

$$P_u(\xi) = P_b(\xi)e^{-\beta(w(\xi)-\Delta A_\xi)} \tag{1}$$

The theory underlying this approach is described in a later section.

Equilibration protocol for MD simulations

Before initiating production trajectories, the MD boxes were subjected to a series of rigorous equilibration steps that are summarized below:

1. During the first part of equilibration, the solute (RNA) is kept fixed to its crystal coordinates by imposing harmonic restraints with force constants of 100 kcal/mol-^{2} on solute heavy atoms, and following steps were performed -

   (a) Energy minimization of the solvent (water molecules + ions)

   (b) Simulated annealing (NVT ensemble) of solvent according to

(c) Increase temperature from 0 K to 298 K in 298 ps, MD at 298 K for 500 ps

(d) Increase temperature from 298 K to 600 K in 302 ps, MD at 600 K for 500 ps

(e) Decrease temperature from 600 K to 298 K in 302 ps, MD at 298 K for 1500 ps

(f) MD at 298 K for 3000 ps (NPT ensemble) of solvent

(g) Repeat of Step b

(h) Repeat of Step c

2. The equilibrated solvent configuration is kept fixed and energy minimization of the solute is performed.

3. With the solvent free to move, the solute is relaxed in stages by slowly releasing the harmonic restraints on the solute heavy atoms according to –

(a) MD at 298 K for 100 ps (NPT ensemble) with a force constant of 50 kcal/mol-Å^2

(b) MD at 298 K for 100 ps (NPT ensemble) with a force constant of 25 kcal/mol-Å^2

(c) MD at 298 K for 100 ps (NPT ensemble) with a force constant of 10 kcal/mol-Å^2

(d) MD at 298 K for 100 ps (NPT ensemble) with a force constant of 5 kcal/mol-Å^2

(e) MD at 298 K for 100 ps (NPT ensemble) with a force constant of 2 kcal/mol-Å^2

4. With both solute and solvent free to move, MD at 298 K for 5000 ps (NPT ensemble).

Theory

Consider a constant NVT ensemble generated by applying a biasing potential, $w(\xi)$, along the reaction coordinate ξ, such that the energy term governing the ensemble is altered according to

$$U_b(\mathbf{q}) = U_u(\mathbf{q}) + w(\xi) \tag{2}$$

where superscripts b and u correspond to the biased and unbiased states, respectively, and $\mathbf{q}$ is a generalized position coordinate.The unbiased probability distribution of the ensemble can be expressed as

$$P_u(\xi) = \langle \delta[\xi(\mathbf{q}) - \xi] \rangle_u \tag{3}$$

, where $\delta(.)$ is the Dirac delta function and $\langle\rangle$ represent ensemble averages. Equation (3) can be rewritten as

$$\begin{aligned}
P_u(\xi) &= \frac{\int d\mathbf{q} e^{-\beta U_u(\mathbf{q})} \delta\left[\xi(\mathbf{q}) - \xi)\right]}{\int d\mathbf{q} e^{-\beta U_u(\mathbf{q})}} \\
&= \frac{\int d\mathbf{q} e^{-\beta[U_b(\mathbf{q}) - w(\xi)]} \delta[\xi(\mathbf{q}) - \xi]}{\int d\mathbf{q} e^{-\beta[U_b(\mathbf{q}) - w(\xi)]}} \\
&= \frac{\int d\mathbf{q} e^{-\beta U_b(\mathbf{q})} e^{\beta w(\xi)} \delta[\xi(\mathbf{q}) - \xi]}{\int d\mathbf{q} e^{-\beta U_b(\mathbf{q})} e^{\beta w(\xi)}} \\
&= \frac{\int d\mathbf{q} e^{-\beta U_b(\mathbf{q})} e^{\beta w(\xi)} \delta[\xi(\mathbf{q}) - \xi] \cdot \int d\mathbf{q} e^{-\beta U_b(\mathbf{q})}}{\int d\mathbf{q} e^{-\beta U_b(\mathbf{q})} e^{\beta w(\xi)} \cdot \int d\mathbf{q} e^{-\beta U_b(\mathbf{q})}} \\
&= \frac{\int d\mathbf{q} e^{-\beta U_b(\mathbf{q})} e^{\beta w(\xi)} \delta[\xi(\mathbf{q}) - \xi] \cdot \int d\mathbf{q} e^{-\beta U_b(\mathbf{q})}}{\int d\mathbf{q} e^{-\beta U_b(\mathbf{q})} \cdot \int d\mathbf{q} e^{-\beta U_b(\mathbf{q})} e^{\beta w(\xi)}} \\
&= \langle e^{\beta w(\xi)} \delta[\xi(\mathbf{q}) - \xi] \rangle_b \frac{\int d\mathbf{q} e^{-\beta U_b(\mathbf{q})}}{\int d\mathbf{q} e^{-\beta U_u(\mathbf{q})}} \\
&= \langle e^{\beta w(\xi)} \delta[\xi(\mathbf{q}) - \xi] \rangle_b \frac{Z_b}{Z_u}
\end{aligned} \tag{4}$$

, where Z is the configurational partition function of a given state, and is related to its free energy, A, as

$$A = -(1/\beta) \ln Z \tag{5}$$

The free energy of the biased state relative to the unbiased state can thus be expressed as

$$\Delta A_\xi = A_b - A_u = -(1/\beta)\ln(Z_b/Z_u), or, equivalently, Z_u/Z_b = e^{\beta \Delta A_\xi} \quad (6)$$

Combining equations (4) and (6),

$$P_u(\xi) = e^{-\beta \Delta A_\xi} \langle \delta[\xi(\mathbf{q}) - \xi] e^{\beta w(\xi)} \rangle_b \quad (7)$$

Since $\delta[\xi(\mathbf{q}) - \xi]$ operates on all degrees of freedom except ξ,

$$P_u(\xi) = e^{-\beta \Delta A_\xi} e^{\beta w(\xi)} \langle \delta[\xi(\mathbf{q}) - \xi] \rangle_b \quad (8)$$

In equation (8), $\langle \delta[\xi(\mathbf{q}) - \xi] \rangle_b$ corresponds to the biased probability distribution of the ensemble, $P_b(\xi)$. Thus,

$$P_u(\xi) = P_b(\xi) e^{-\beta(w(\xi) - \Delta A_\xi)} \quad (9)$$

According to equation (9), the unbiased probability distribution of the ensemble along a given reaction coordinate can be calculated from a biased probability distribution by knowing the biasing potential and the relative free energy between the biased and unbiased states.

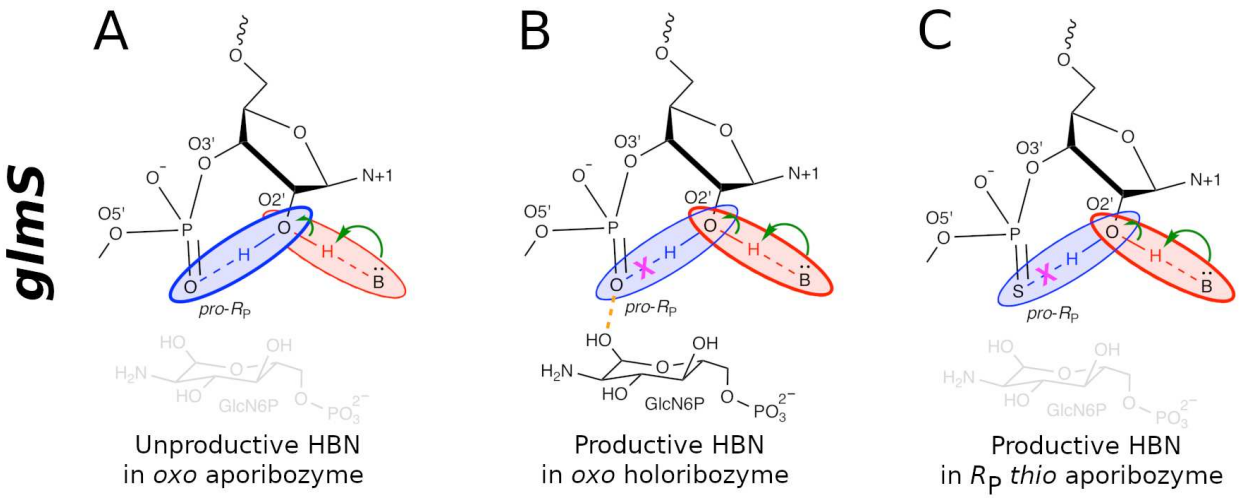

Figure S1: **3°γ (HBN) effect observed in glmS ribozyme.** In absence of the GlcN6P cofactor (aporibozyme), the HO2′ forms non-productive interaction with the pro-R_P NPO in the wild type (oxo) substrate (A). In the holoribozyme, the cofactor donates a hydrogen bond to the pro-R_P NPO, disrupting the HO2′- pro-R_P NPO interaction and thus promoting productive interactions of the O2′ nucleophile (B). In the R_P thio substrate of the aporibozyme, the sulfur atom at the R_P position disrupts the HO2′- pro-R_P NPO interaction, thus also promoting productive interactions of the O2′ nucleophile (C). The 3°γ (HBN) effect in glmS ribozyme was reported by the Bevilacqua group.[16,17]

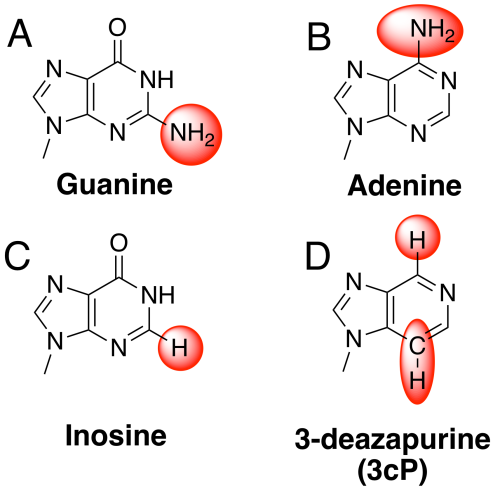

Figure S2: **G638I and A756(3cP) mutation in VSr** In the double mutant of VSr, the guanine G638 (A) was mutated to Inosine (C) and the adenine A756 (B) was mutated to 3-deazapurine (3cP) (D).

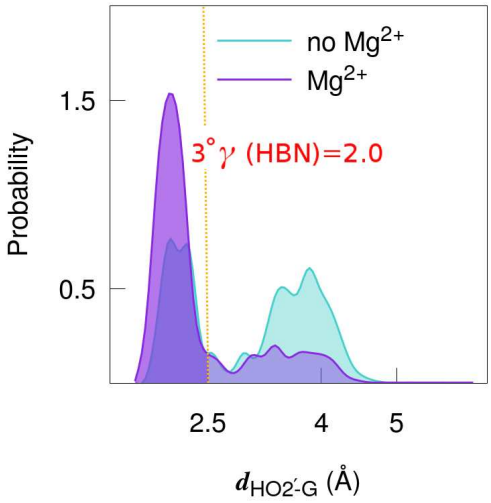

Figure S3: **3°γ (HBN) effect in Psr induced by inner sphere metal binding.** One dimensional probability distribution functions (1D PDFs) constructed from the unbiased and reweighted umbrella sampling simulations are shown for Psr with the active site Mg^{2+} ion making an inner sphere contact with the pro-R_P NPO. 1D PDFs obtained in presence of the metal ion are shown in purple and 1D PDFs obtained in absence of metal ion are shown in cyan. The 3°γ (HBN) factor is defined as the ratio of the areas under the 1D PDFs obtained in presence and absence of the metal ion, that reside below a threshold value (2.5Å) of the HO2′-guanine distance (shaded region). These results are denoted as PSr^* is Table 1 of the main text.

References

(1) DasGupta, S.; Suslov, N. B.; Piccirilli, J. A. Structural Basis for Substrate Helix Remodeling and Cleavage Loop Activation in the Varkud Satellite Ribozyme. *J. Am. Chem. Soc.* **2017**, *139*, 9591–9597.

(2) Martick, M.; Lee, T.-S.; York, D. M.; Scott, W. G. Solvent Structure and Hammerhead Ribozyme Catalysis. *Chem. Biol.* **2008**, *15*, 332–342.

(3) Ren, A.; Vusurovic, N.; Gebetsberger, J.; Gao, P.; Juen, M.; Kreutz, C.; Micura, R.; Patel, D. Pistol Ribozyme Adopts a Pseudoknot Fold Facilitating Site-specific In-line Cleavage. *Nat. Chem. Biol.* **2016**, *12*, 702–708.

(4) Liu, H.; Yu, X.; Chen, Y.; Zhang, J.; Wu, B.; Zheng, L.; Haruehanroengra, P.; Wang, R.; Li, S.; Lin, J.; Li, J.; Sheng, J.; Huang, Z.; Ma, J.; Gan, J. Crystal Structure of an RNA-Cleaving DNAzyme. *Nat. Commun.* **2017**, *8*, 2006–2015.

(5) Horn, H. W.; Swope, W. C.; Pitera, J. W. Characterization of the TIP4P-Ew Water Model: Vapor Pressure and Boiling Point. *J. Chem. Phys.* **2005**, *123*, 194504.

(6) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(7) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(8) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(9) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic

Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(10) Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. Application of the Multimolecule and Multiconformational RESP Methodology to Biopolymers - Charge Derivation for DNA, RNA, and Proteins. *J. Comput. Chem.* **1995**, *16*, 1357–1377.

(11) Lee, T.-S.; York, D. M. Origin of Mutational Effects at the C3 and G8 Positions on Hammerhead Ribozyme Catalysis from Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2008**, *130*, 7168–7169.

(12) Lee, T.-S.; Silva Lopez, C.; Giambaşu, G. M.; Martick, M.; Scott, W. G.; York, D. M. Role of Mg^{2+} in Hammerhead Ribozyme Catalysis from Molecular Simulation. *J. Am. Chem. Soc.* **2008**, *130*, 3053–3064.

(13) Kostenbader, K.; York, D. M. Molecular Simulations of the Pistol Ribozyme: Unifying the Interpretation of Experimental Data and Establishing Functional Links with the Hammerhead Ribozyme. *RNA* **2019**, *in press, doi 10.1261/rna.071944.119.*

(14) Ekesan, Ş.; York, D. M. Dynamical Ensemble of the Active State and Transition State Mimic for the RNA-cleaving 8-17 DNAzyme in Solution. *Nucleic Acids Res.* **2019**, *in press, doi 10.1093/nar/gkz773.*

(15) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129*, 124105.

(16) Bingaman, J. L.; Zhang, S.; Stevens, D. R.; Yennawar, N. H.; Hammes-Schiffer, S.; Bevilacqua, P. C. The GlcN6P Cofactor Plays Multiple Catalytic Roles in the glmS Ribozyme. *Nat. Chem. Biol.* **2017**, *13*, 439.

(17) Bingaman, J. L.; Gonzalez, I. Y.; Wang, B.; Bevilacqua, P. C. Activation of the glmS

Ribozyme Nucleophile via Overdetermined Hydrogen Bonding. *Biochemistry* **2017**, *56*, 4313–4317.