Supporting Information to:

The Role of an Active Site Mg^{2+} in the HDV Ribozyme Self-Cleavage: Insights from QM/MM Calculations

Vojtěch Mlýnský,[1] Nils G. Walter,[,3] Jiří Šponer,[2,4] Michal Otyepka,[*,1,2] and Pavel Banáš[*,1,2]*

[1]Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacky University, tr. 17 listopadu 12, 771 46, Olomouc, Czech Republic,

[2]Institute of Biophysics, Academy of Sciences of the Czech Republic, Kralovopolska 135, 612 65 Brno, Czech Republic,

[3]Department of Chemistry, Single Molecule Analysis Group, University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109-1055

[4] CEITEC – Central European Institute of Technology, Masaryk University, Campus Bohunice, Kamenice 5, 625 00 Brno, Czech Republic

Setup for MD simulations

Molecular dynamics (MD) simulations were used to prepare starting structures for our QM/MM calculations. Starting structures were based on the most recent crystal structure of the *trans*-acting genomic HDV ribozyme (PDB ID 3NKB; resolution 1.9 Å).[1] We performed three different simulations containing zero, one (the active site Mg^{2+} ion) or four Mg^{2+} ions (all Mg^{2+} ions presumed to have important structural and/or functional roles, including the active site Mg^{2+} ion; see Figure S1). The well depth ε of Mg^{2+} ion was 0.8947 kcal/mol and the radii R = 0.7926 Å.[2] The initial positions of these Mg^{2+} ions were taken from the crystal structure. Note that the divalent ions are poorly described by the approximate non-polarizable force field so that it is prudent to minimize their usage in classical MD simulations whenever they do not play a significant structural role and can be substituted by monovalent ions.[3,4] All three simulated systems were further neutralized with Na^{+} counter ions (well-depth ε = 0.0028 kcal/mol and radii R = 1.8680 Å)[2] and immersed in a rectangular water box with an at least 10.0-Å thick layer of TIP3P water molecules all around the RNA solute.

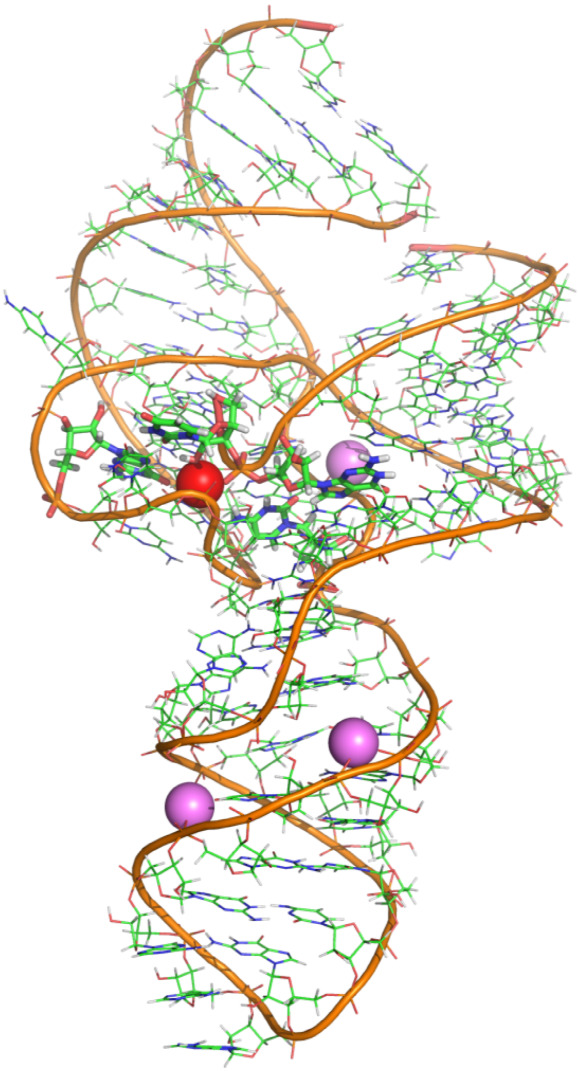

Figure S1: *Initial positions of the Mg^{2+} ions in our MD simulations. The active site Mg^{2+} ion is depicted in red, while the remaining three Mg^{2+} ions with presumably important structural roles are in violet. The snapshot has the same orientation and representation as Fig. 1B in the main text.*

The solute-solvent system was minimized prior to MD simulation as follows. Minimization of the ribozyme hydrogen atoms was followed by minimization of counter ions and water molecules. Subsequently, the ribozyme was frozen and solvent molecules with counter ions were allowed to move during a 10 ps-long MD run to relax the density in the box. The nucleobases were allowed to relax in several minimization runs with decreasing force constants applied to the backbone phosphate atoms. After full relaxation, the system was slowly heated to 298.15 K over 80 ps using 2-fs time steps and NpT conditions. The simulations were performed under periodic boundary conditions in the NpT ensemble (298.15 K, 1 atm) with 2-fs time steps. The particle-mesh Ewald method was used to calculate electrostatic interactions and a 10.0-Å cutoff was applied for Lennard-Jones interactions. The SHAKE algorithm was applied to all bonds containing hydrogen atoms. The length of all three simulations (with none, one or four Mg^{2+} ions) was 80 ns.

Selection of starting geometries for QM/MM calculations

Our three simulations revealed a total of 22 varying conformational states with a population of at least 0.1% during at least one MD simulation, which differed in their coordination and position of the ion within the active site (Table S1). Based on the probability of occurrence of the hydrogen bond between one of water molecule (WAT) from the first coordination shell of the active site Mg^{2+} ion and the U-1(2'-OH) group, only 16 positions of the ion in the active site were considered. Starting snapshots for QM/MM calculations were selected from geometries achieving the best orientation of the key residues in the active site, where two catalytically important hydrogen bond distances (U-1(2'-OH)…WAT(O) and $C75H^{+}$(N3H)…G1(O5')) were shorter than 2.8 Å and the in-line attack angle of (U-1(O2')…G1(P)…G1(O5')) was at its highest (most favorable) value (typically greater than 160°). This set of 16 selected snapshots contained two geometries with triple-inner-shell coordination of the active site ion to [U-1(O2'), G1(*pro*-R_P), U23(*pro*-S_P)], which were chosen from the MD simulations with either one or four divalent Mg^{2+} ions. Seven snapshots corresponded to a double-inner-shell coordination to [G1(*pro*-R_P), G23(*pro*-S_P)], [U-1(O2), G25(N7)], [U-1(O2), G25(O6)], [G1(*pro*-R_P), U20(O2)], [G25(O6), G25(N7)], [U-1(O2), U23(*pro*-S_P)], and [G1(*pro*-R_P), G25(O6)]. Six other snapshots had single-inner-shell coordination to [U-1(O2)], [G25(N7)], [G25(O6)], [G1(*pro*-R_P)], [U23(*pro*-S_P)], or [U20(O2)]. The remaining snapshot contained an active site ion with a complete inner shell of six water molecules.

Preliminary QM/MM minimizations

One of the coordinated water molecules with a hydrogen bond to the U-1(2'-OH) group was deprotonated prior to minimization and, if necessary, the monovalent Na^{+} ion in the active site was replaced with the divalent Mg^{2+} (a total of 13 snapshots originated from the MD simulation in the presence of monovalent ions only). QM/MM minimizations reduced the total number of testing snapshots from 16 to 13 since in three minimizations the coordination of the active site Mg^{2+} ion changed to another coordination that was already considered. Notably, only one snapshot with triple-inner-shell coordination to [U-1(O2'), G1(*pro*-R_P), U23(*pro*-S_P)] passed the criteria enumerated above. In addition, the snapshot with outer-shell coordination of the Mg^{2+} ion converted to a single-inner-shell coordination to [G25(N7)], whereas the snapshot with single-inner-shell coordination to [U20(O2)] converted to a double-inner-shell coordination to [G1(*pro*-R_P), U20(O2)]. These newly established coordinations of the active site ion were redundant with the initial set.

QM/MM calculations of reaction pathways

The set of 13 different position/coordinations of the Mg^{2+} ion was probed by QM/MM calculations. QM/MM calculations revealed that two snapshots (with double-inner-shell coordination to [U-1(O2), G25(O6)] and single-inner-shell coordination to [G25(N7)]) changed the active site geometry and established another double-inner-shell coordination to [U-1(O2), U-1(O2')]. We further observed that the double-inner-shell coordination to [G1(*pro*-R_P), G25(O6)] changed to the newly formed triple-inner-shell coordination to [G1(*pro*-R_P), U20(O2), G25(O6)]. The remaining coordinations of the Mg^{2+} ion without complete reaction profiles (i.e., double-inner-shell to [U-1(O2), U23(*pro*-S_P)] and single-inner-shell to [G25(O6)], [G1(*pro*-R_P)], and [U23(*pro*-S_P)]) also converted to a different coordination of the Mg^{2+} ion (typically one of double-inner-shell coordinations with an already localized reaction profile, Table 1 in the main text). In summary, we succeeded in obtaining 8 complete reaction pathways with two triple-inner-shell coordinations, five double-inner-shell coordinations, and one single-inner-shell coordination of the active site Mg^{2+} ion (Table 1 in the main text).

Table S1: *Behavior of the active site ion during MD simulations, where we analyzed the position and coordination of the ion towards functional groups in the active site of the HDV ribozyme.*

MD simulation	specific coordination number [a]	functional groups involved in coordination	population during MD (%) [b]	population, where X…WAT…U-1(O2') favorable (%) [c]
$4Mg^{2+}$ ions	3	[U-1(O2'), G1(*pro*-R_P), U23(*pro*-S_P)]	30.1	99.4
	2	[G1(*pro*-R_P), G23(*pro*-S_P)]	69.9	87.7
	1	-	0.0	-
	0	-	0.0	-
$1Mg^{2+}$ ion	3	[U-1(O2'), G1(*pro*-R_P), U23(*pro*-S_P)]	99.3	99.8
	2	[G1(*pro*-R_P), G23(*pro*-S_P)]	0.7	91.9
	1	-	0.0	-
	0	-	0.0	-
Na^+ ion	3	[U-1(O2), G1(*pro*-R_P), U23(*pro*-S_P)]	1.1	6.7
		[U-1(O2), U23(O2'), G25(N7)]	0.1	18.2
		[G1(*pro*-R_P), U20(O2), G25(O6)]	0.1	61.7
		[U-1(O2), G25(O6), G25(N7)]	0.1	60.0
	2	[G1(*pro*-R_P), G23(*pro*-S_P)]	4.3	8.3
		[U-1(O2), G25(N7)]	4.1	40.1
		[U-1(O2), G25(O6)]	3.5	39.6
		[G1(*pro*-R_P), U20(O2)]	1.8	38.3
		[G25(O6), G25(N7)]	1.3	55.5
		[U-1(O2), U23(*pro*-S_P)]	1.0	7.1
		[G1(*pro*-R_P), G25(O6)]	0.6	57.9
		[U-1(O2), G1(*pro*-R_P)]	0.2	5.6
		[U-1(O2), U23(O2')]	0.1	16.3
		[U20(O2), G25(O6)]	0.1	46.2
	1	[U-1(O2)]	18.3	40.4
		[G25(N7)]	13.9	47.9
		[G25(O6)]	6.6	50.1
		[G1(*pro*-R_P)]	4.4	42.2
		[U23(*pro*-S_P)]	1.4	11.1
		[U20(O2)]	0.4	50.0
	0	-	36.4	49.6

[a] number of the active site groups involved in the coordination of the ion excluding water molecules form the first coordination shell (3 – triple-inner-shell coordination, 2 – double-inner-shell coordination, 1 – single-inner-shell coordination, and 0 – outer-shell coordination).

[b] we considered all atoms as possible ligands to the active site ion with distance from the ion < 2.5 Å.

[c] number (in %) of those states from first population (previous column), where any water molecule from the first coordination shell (< 2.5 Å) of the active site ion (either Mg^{2+} or Na^+, labeled as X) was within strong hydrogen bond distance (< 3.0 Å) from the U-1(O2') group.

Model of the uncatalyzed reaction to estimate the Gibbs energy corrections

To estimate the necessary Gibbs energy corrections, we modeled the self-cleavage reaction using as sugar-phosphate model 3'-(1'-amino-4'-methylribose)-5'-methylphosphodiester, accompanied by a H1-protonated N1-methylcytosine and a hydrated Mg^{2+} ion with one of the inner-shell waters deprotonated while forming an outer-shell coordination to the non-bridging oxygen of the scissile phosphate. The same approach was used in our recent QM/MM studies of the HDV and hairpin ribozymes.[5,6] Starting geometries of the pre-cleavage, transition, intermediate and product states of the sugar-phosphate backbone (i.e., without cytosine and hydrated Mg^{2+} ion, in total containing 27 atoms) were adopted from our previous work.[5] All

structures were optimized using the MPW1K/6-31+G(d,p) functional in implicit water, represented by a polarizable conductor calculation model (CPCM) as implemented in Gaussian 09.[7] Frequencies (under harmonic approximation) were calculated at the same level for each optimized structure to estimate corrections to the Gibbs energy at 298 K and 1 atm. The differences between the gas phase MPW1K/6-31+G(d,p) energies and the CPCM (ε_r = 78.4)/MPW1K/6-31+G(d,p) energies were used to estimate the solvation contributions to the energy profile of the reference reaction.

Table S2: *The MPW1K/6-31+G(d,p) gas phase energies, solvation and Gibbs energy corrections (calculated at the CPCM(ε_r=78.4)/MPW1K/6-31+G(d,p) level), as well as the total Gibbs energy profiles of the self-cleavage reaction of the endo/exo-3'-(1'-amino-4'-methylribose)-5'-methylphosphodiester sugar-phosphate backbone model extended by an N3-protonated N1-methylcytosine and a deprotonated hydroxide anion coordinated to a partially hydrated Mg^{2+} ion ($[Mg(H_2O)_5OH]^+$). The hydrated Mg^{2+} ion acts as the general base, while the protonated cytosine acts as the general acid. All energies and energy corrections are in kcal/mol and are relative to the pre-cleavage (reactant) state. No pK_a corrections for the $[Mg(H_2O)_5OH]^+$ ion and the N3-protonated N1-methylcytosine were included in calculating the total Gibbs energies.*

	R	TS	P
MPW1K/6-31+G(d,p)	0.0	13.6	22.0
Solvation energy	0.0	-1.2	-24.0
Gibbs energy correction	0.0	-0.7	-5.0
MPW1K Gibbs energy in water	0.0	11.7	-7.0

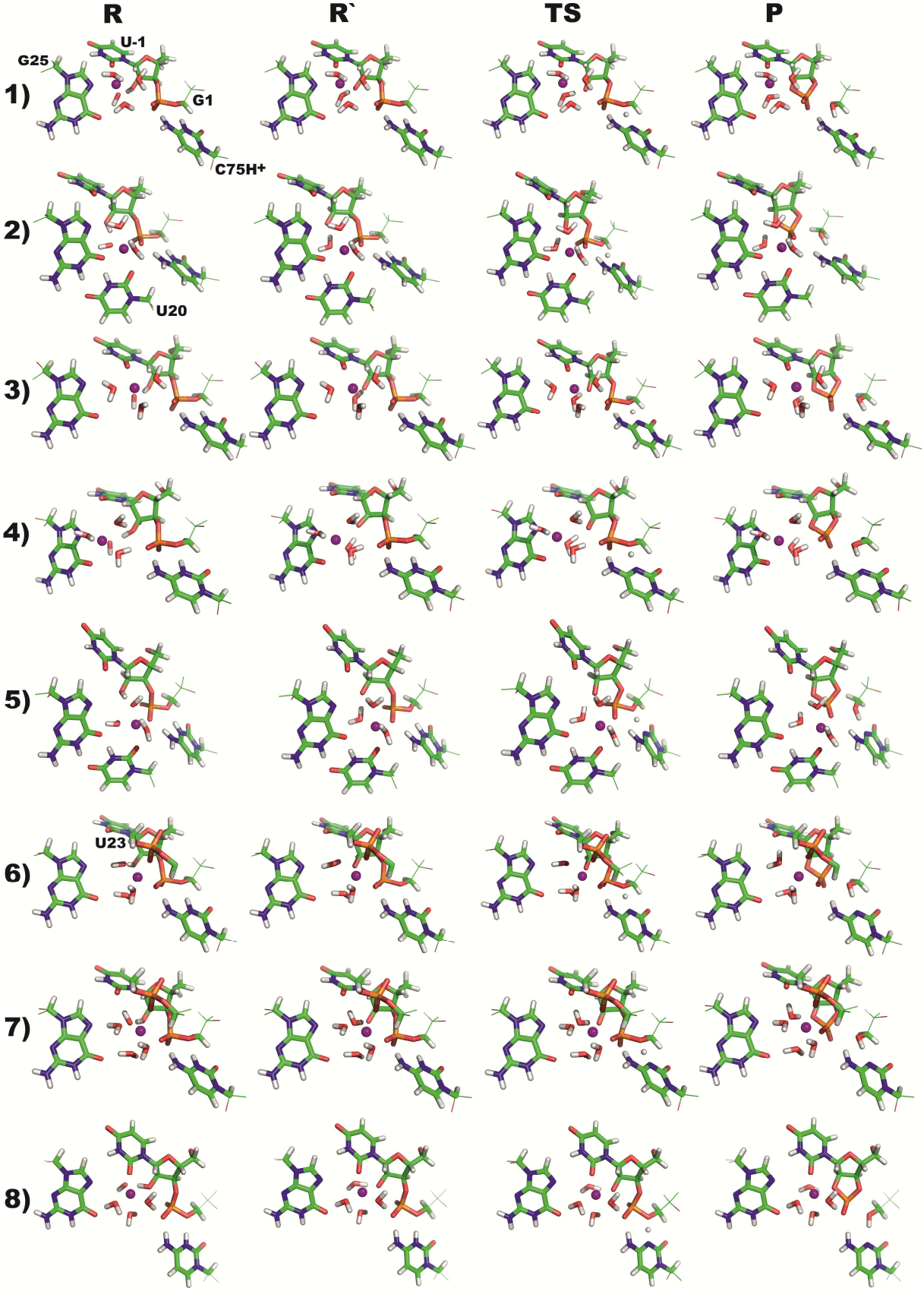

Figure S2: *Detailed QM core geometries obtained by QM/MM calculations along reaction pathways in the HDV ribozyme. Geometries are oriented in the same way as in Fig. 2 in the main text. Calculated reaction pathways with different position and coordination of the active*

site Mg^{2+} ion are numbered in the same order as they are displayed in Table 1 in the main text, i.e., 1) double-inner-shell coordination to [U-1(O2), G25(N7)], 2) triple-inner-shell coordination to [G1(pro-R_P), U20(O2), G25(O6)], 3) single-inner-shell coordination to [U-1(O2)], 4) double-inner-shell coordination to [G25(O6), G25(N7)], 5) double-inner-shell coordination to [G1(pro-R_P), U20(O2)], 6) triple-inner-shell coordination to [U-1(O2'), G1(pro-R_P), U23(pro-S_P)], 7) double-inner-shell coordination to [G1(pro-R_P), U23(pro-S_P)], and 8) double-inner-shell coordination to [U-1(O2), U-1(O2')].

References:

(1) Chen, J. H.; Yajima, R.; Chadalavada, D. M.; Chase, E.; Bevilacqua, P. C.; Golden, B. L. *Biochemistry* **2010**, *49*, 6508-6518.

(2) Aqvist, J. *J. Phys. Chem.* **1990**, *94*, 8021-8024.

(3) Banas, P.; Jurecka, P.; Walter, N. G.; Sponer, J.; Otyepka, M. *Methods* **2009**, *49*, 202-216.

(4) Ditzler, M. A.; Otyepka, M.; Sponer, J.; Walter, N. G. *Acc. Chem. Res.* **2010**, *43*, 40-47.

(5) Banas, P.; Rulisek, L.; Hanosova, V.; Svozil, D.; Walter, N. G.; Sponer, J.; Otyepka, M. *J. Phys. Chem. B* **2008**, *112*, 11177-11187.

(6) Mlynsky, V.; Banas, P.; Walter, N. G.; Sponer, J.; Otyepka, M. *J. Phys. Chem. B* **2011**, *115*, 13911-13924.

(7) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A. et al. Gaussian 09; Gaussian, Inc., Wallingford CT, 2009.