# Molecular mechanism of rRNA maturation by M5: interplay between conformational flexibility and reactivity

Julie Puyo-Fourtine, Laetitia Kantin, Milorad Andelkovic, Carine Tisné, Elisa Frezza, Inaki Tunon, Elise Duboué-Dijon

# Molecular mechanism of rRNA maturation by M5: interplay between conformational flexibility and reactivity

Julie Puyo-Fourtine,[*,†] Laetitia Kantin,[†] Milorad Andelkovic,[‡] Carine Tisné,[¶] Elisa Frezza,[§] Iñaki Tuñón,[‡] and Elise Duboué-Dijon[*,†]

†*Université Paris Cité, CNRS, Laboratoire de Biochimie Théorique, 13 rue Pierre et Marie Curie, 75005, Paris, France*

‡*Departament de Química Física, Universitat de València, 46100 Burjassot, Spain*

¶*Université Paris Cité, CNRS, Institut de Biologie Physico-Chimique, IBPC, Expression Génétique Microbienne, Paris 75005, France*

§*Université Paris Cité, CiTCoM, Paris, France*

E-mail: puyo@ibpc.fr; elise.duboue-dijon@cnrs.fr

**Abstract**

Combining enhanced sampling molecular dynamics simulations with dynamic QM/MM-MD exploration of the mechanisms, we shed light at the molecular level on the maturation process of 5S rRNA in *B. subtilis*, performed by the M5 enzyme. Providing a small displacement of the scissile phosphate and the use of adequate force fields, the structure of the experimentally reconstructed complex is stable, with small conformational fluctuations of the interface. Enhanced sampling is critical to reveal the full extent of the conformational variability of the rRNA:M5 interface, which we show critically impacts the active site structure by modulating the distance and orientation between key reactive residues. This rich conformational landscape thus suggests a variety of possible reaction mechanisms, that depend on the active site pre-organization. Comprehensive characterization of the different possible mechanisms from each conformational basin suggests that the most favorable reaction route would involve deprotonation of the nucleophilic water by the active site Glu96 and reprotonation of the 3′ leaving group by a $Mg^{2+}$-bound water. Contrary to other RNases, M5 activity does not rely on a bifurcated aspartate between the 2 active site cations, nor on the release of initially frustrated hydration shells.

# Introduction

Ribosomes are macromolecular machines composed of proteins and RNA, that translate the genetic information contained into messenger RNA (mRNA) and perform protein synthesis. Their synthesis and assembly dominates energy consumption in living organisms and are subjected to strict quality control mechanisms. In bacteria, the three ribosomal RNAs (rRNAs)—16S, 23S, and 5S— are initially produced as a single long polynucleotide (pre-30S), which then goes through a series of maturation steps to form a fully functional ribosome ribonucleoprotein particle.[1] These maturation processes are the subject of increasing interest as they could be involved in the regulation of gene expression and the response and adaptation

adaptation to stress.

While the initial cleavage of the long pre-30S precursor is performed by Ribonuclease (RNase) III,[1] the following maturation steps, occuring directly on the pre-assembled ribosomal particle, are found to be highly dependent on the studied prokaryotic model.[2] In Firmicutes, such as *Bacillus subtilis*, final maturation of the 5S rRNA—shortening of its 3′ and 5′ ends—is performed by the RNase M5,[3–5] a ∼190 residue enzyme composed of 2 domains separated by a flexible linker (Figure 1). While the C-terminal domain serves to anchor the enzyme on the rRNA, catalysis of the cleavage reaction is performed by the Toprim N-terminal domain, shared by other endonucleases, such as DNA type II topoisomerases and DNA repair recombination proteins RecR.[6] M5, which acts as a monomer,[7,8] is however the only Toprim domain-containing enzyme known to cleave its substrate on both sides of an RNA double stranded stem, which makes it of special interest. Previous works[5,7,9] have suggested that the cleavage occurs in two steps, starting with the removal of the 3′ extension before reorganization of the complex to position the catalytic domain for cleavage of the 5′ end (Figure 1a).

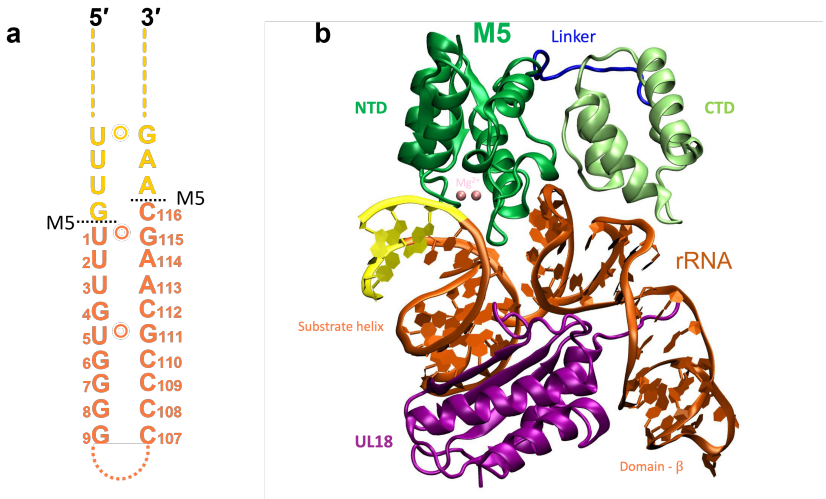


Figure 1: a) Scheme of the pre-5S rRNA with the two identified locations for cleavage by M5. b) Simulated complex between M5, a 104-nucleotide rRNA fragment and the uL18 protein cofactor.

C. Tisné and coworkers recently provided molecular insights[9,10] into this maturation

process, with a cryo-EM structure of the complex between M5 and the pre-5S ribosome particle, complemented by X-ray structures of the two domains of M5. They showed that M5 binds the ribosomal particle as a monomer and that its Toprim active site is poised to coordinate two $Mg^{2+}$ ions, evidenced in a cristallographic structure of unbound M5 N-terminal domain obtained at a high $Mg^{2+}$ concentration (Figure 2a).

Nucleases—enzymes that catalyze the cleavage of RNA or DNA phosphodiester bonds—play essential roles in a large number of key biological processes, e.g. DNA replication, recombination and repair, DNA topoisomerization and RNA processing, maturation and splicing.[11] While they all cleave phosphodiester bonds, they present a broad diversity of structures and of mechanisms, that can either be ion-independent or rely on 1 or 2 ions—typically $Mg^{2+}$—for catalysis. Given the importance of these processes, there is a pronounced interest in better understanding them at a molecular level, which is where molecular simulations can bring valuable additional insights.

Only few computational studies of ion-dependant RNases mechanisms are found in the literature, almost exclusively on RNase H,[12–16] and several points of its mechanism remain debated, including for instance the initial state—deprotonated or not—of the nucleophilic water molecule, the proton shuttle pathway(s), and the identity of a possible general acid to reprotonate the 3′OH leaving group. The recently obtained structures of the rRNA:M5 complex thus appeared as an opportunity to examine the molecular mechanism of a different 2-ion dependent RNase.

In addition, earlier studies suggested that the presence of a bifurcated aspartate bridging the two active site $Mg^{2+}$ ions was essential for RNase activity[17] and that release of the frustration of an initially undercoordinated $Mg^{2+}$ in site B played a key role in catalysis.[18,19] As these two features are absent from the experimentally reconstructed structure of the M5:rRNA complex, this was an opportunity to re-examine their role in RNase catalysis. Examining the role of structural features in reactivity requires to ensure proper sampling of active site conformations in simulations, as recent works increasingly suggest the interplay

between conformational dynamics and reactivity to be central to understand enzyme catalysis with, for instance, conformational selection now being used as a guiding principle to enzyme design.[20]

In this work, we thus investigated using various molecular simulation approaches the structural stability of the M5:rRNA complex and characterized the conformational dynamics of this interface. We show that fluctuations of the interface impact key structural features of the active site and we identify distinct conformations, whose structures appear pre-organized to favor distinct mechanisms. Finally, we characterize using dynamic QM/MM-MD simulations the reaction mechanism and free energy profile in each of the identified conformational basins, identify the most favorable route and provide a comprehensive picture of the interplay between conformational fluctuations and reactivity of this RNase.

# Computational Methods

## Molecular Dynamics simulations

**System Preparation.** The simulated complex was prepared with Pymol[21] by superimposing the cryo-EM structure of the complex formed between *B. Subtilis* 50S (pre-5S) ribosomal RNA subunit and the M5 enzyme[9] (PDB: 6TPQ), with the X-ray structure[10] (code PDB: 6Z2B) of uncomplexed M5 Toprim domain (chain B)—which contains two $Mg^{2+}$ ions in the active site. All water molecules resolved in the X-ray structure (in particular in the $Mg^{2+}$ hydration shell) were retained, and the mutation of Asp58 to alanine was reversed.

Focusing on the rRNA:M5 interface, our simulated model comprised the M5 enzyme, the UL18 cofactor (PDB chain h) and part of the rRNA chain (PDB chain V), selecting a 104-nucleotide portion (5′ strand up to residue U72, and 3′ strand up to A99) relevant for the interaction with M5 (see Figure 1b). A loop (sequence G-G-U-A-U-G-A-G from 5′ to 3′) was then added using RNA Composer[22,23] to avoid fraying of the RNA ends. Using Gromacs tools, this rRNA:protein complex was then immersed in a 130 Å box of water molecules with

a background NaCl concentration of 150 mM.

**Force fields.** Proteins were described with the ff14SB force field,[24] and RNA with the amberff99-bsc0-$\chi_{OL3}$ force field.[25,26] These force fields were combined with the TIP3P water model[27] and Joung-Cheatham parameters for monovalent ions.[28] The stability of the active site structure was tested using 3 different force field for $Mg^{2+}$: the non-polarizable full charge model proposed by Merz and coworkers[29] "$Mg_m$" (TIP3P CM set version); the parametrization proposed by Schwierz and coworkers[30] (TIP3P version) with pair-specific Lennard-Jones parameters to capture the $Mg^{2+}$–phosphate interaction strength; and our scaled charge "Electronic Continuum Correction" (ECC) force field, with a charge reduced to +1.6 on both $Mg^{2+}$ atoms.[31] The ECC approach is a mean-field strategy[32,33] that takes into account electronic polarization as a dielectric continuum, which is formerly equivalent to scaling ionic charges by a factor $1/\sqrt{\epsilon_\infty} \sim 0.75$ in water. This approach has been found very promising to describe electrolyte solutions[34] and capture ion pairing to phosphate moieties, where a scaling factor of 0.8 was found more appropriate.[31] We used this force field in combinaison with the SPC/E water model,[35] that it was developed to be compatible with. An additional simulation was performed using also an ECC scaled-charge description for the protein and RNA residues involved in $Mg^{2+}$ coordination. The detailed parameters for each force fields are provided in SI (Tables S1, S2, S3, S4, S5, S6).

All the parameter files, as well as typical input files for classical and QM/MM-MD simulations and selected trajectories are provided in an online repository.[36]

**MD Simulation protocol.** All classical molecular dynamics simulations were performed using Gromacs2022 software.[37] Visualization of the trajectories and evolution of key collective variables was performed in VMD[38] augmented by the Colvars dashboard.[39]

The system was first minimized for 2000 steps, using flat-bottom harmonic restraints ($k$ = 20000 kJ/mol/nm$^2$) that act when the distance between the $Mg^{2+}$ ions and all the oxygen atoms initially present in their first hydration shells (4 water molecules, the bound

oxygen atom of Asp56, Glu10 and Glu96) exceed 2.2 Å. In addition—and after initial tests that all immediately resulted in the active site falling apart—the scissile phosphate group of A117 was moved to adopt a bridge localization in between the two $Mg^{2+}$ ions, thus replacing the water molecule that bridges the two ions in the X-ray structure (see Figure 2. This was done using flat-bottom harmonic restraints on both $Mg^{2+}$–O2P(pro-Sp)(A117) distances, with a harmonic wall at 2.2 Å. In addition, position restraints were applied to all heavy atoms of the complex.

The system was equilibrated in the NPT ensemble for a total of 100 ns, using the Parrinello-Rahman barostat[40] and a stochastic velocity rescaling thermostat[41] at 300 K with a 1 ps coupling time. Hydrogen-containing bonds were constrained using the LINCS algorithm[42] and the SETTLE algorithm for water molecules,[43] which allowed the use of a 2 fs propagation time step. Particle mesh Ewald[44] with a cutoff distance of 12 Å was used to treat long-range electrostatic interactions, and the same cutoff was used for Van der Waals parameters. Position restraints on heavy atoms were gradually decreasing until complete removal (Table S7 for the detailed procedure), while the restraints on the $Mg^{2+}$ coordination sphere were kept during the entire equilibration. From the end of this equilibration protocol, in order to test the performance of different force fields to describe the active site structure, an additional 20 ns NPT equilibration was performed with each of the tested force fields, followed by a 500 ns equilibration keeping only restraints (k = 20000 kcal/mol/nm$^2$) on the distance between $Mg^{2+}$ ions and the non-water oxygen atoms in their hydration sphere. These restraints were then removed and an additional 500 ns NPT production run was performed with each force field in 4 replicates, with different initial velocities.

Once the force field impact on the active site characterized, we generated 4 long 500 ns NPT simulations starting from the structure obtained at the end of the equilibration protocol, with restraints on the distance between $Mg^{2+}$ ions and the non-water oxygen atoms in their hydration sphere. We used here the pair-specific Lennard Jones force field for $Mg^{2+}$ ions, but note that the specific choice of the force field is here irrelevant, since restraints

maintain the active site structure around the ions. In addition, extensive exploration of the complex conformational space was performed using a Hamiltonian replica exchange scheme to accelerate sampling, namely Replica Exchange with Solute Tempering (REST2),[45] scaling being applied to a subset of atoms comprising the RNA and M5, as well as the 2 $Mg^{2+}$ ions and 102 $Na^{+}$ ions to keep a neutral charge in the rescaled region. 32 replicas were used with scaling factors $\lambda$ ranging from 1 to 0.731707, thus covering an effective temperature range for rescaled atoms of 300 K to 410 K. Exchange was attempted every 2 ps, which yielded an average exchange rate of $\sim$20 %. This simulation was replicated using random initial velocities.

## Quantum Mechanims(QM)/Molecular Mechanics(MM)-MD Simulations

Hybrid QM/MM-MD simulations were used to investigate the phosphate cleavage reaction mechanism by M5. All QM/MM-MD calculations were carried out with the sander module of the Amber23 software,[46,47] using a 1 fs time step. First explorations of different possible mechanisms were conducted using the semi-empirical DFTB3[48] method for the QM region with OPhyd[49] parameters, specifically developed to describe phosphorus chemistry with improved accuracy. The most promising mechanism was then re-examined *(in progress)* with a higher DFT level of theory, B3LYP/6-31G*,[50–52] using the Gaussian16/sander QM/MM interface. This functional and basis set were chosen as a good compromise between computational cost and proven accuracy in similar phosphoryl transfer or phosphate hydrolysis reactions.[53,54] Dispersion corrections were recently shown to be detrimental to the description of the interaction of divalent cations with their surrounding,[55] and were thus not used here, where $Mg^{2+}$ ions are central for the reactivity. A 15 Å cutoff was set for QM/MM interactions.

The QM region comprised the ribose phosphate backbone of A117 and C116, the two active site $Mg^{2+}$ ions, the water molecules in their first hydration shell as well as protein

residue side chains involved in $Mg^{2+}$ inner-sphere coordination (Glu96, Glu10, Asp56), for a total of 94 atoms (see Figure S1). Starting structures with active site geometries appropriate for each mechanism were selected from the above described MD simulations and further equilibrated for $\sim$ 60 ps at the QM(DFTB3)/MM-MD[48] level after conversion of the input and parameter files from Gromacs to Amber format using the Parmed tool.[56] The masses of the hydrogens involved in proton transfer were adjusted to 2 amu to increase the weight of the proton transfer events in the overall path mass-weighted reaction coordinate $s$. A 12.0 Å cutoff radius was used for electostatic and Lennard-Jones QM/MM interactions in B3LYP calculations while DFTB3 simulations employed Ewald for long-range electrostatic interactions. The RNase M5 reaction free energy landscape was explored using a dynamic path sampling approach, the adaptive string method (ASM),[57] developed by one of us. The ASM approach allows to dynamically find the minimum free energy path (MFES) for the reaction, describe it with a single path coordinate $s$, and straightforwardly obtain the associated reaction free energy profile. In this method, the path is discretized into $N$ replicas of the system called "string nodes" ($N = 128$ in this work) defined along an assumed reaction path in a $M$-dimensional space, where $M$ is the number of selected collective variables (CVs). These CVs are the key degrees of freedom (typically, distances or angles) that describe the reaction. Here, we used typically a set 7 CVs to describe the reaction, comprising the distances of the bonds broken or formed during the chemical reaction (Tables S8, S9, S10 for full list).

Starting from an initially chemically-educated guess of the reaction path, the positions of the string nodes are evolved according to the free energy gradient while being kept equidistant in the CV space, until the MFEP is found. Typically the string was propagated until the root-mean-square deviation of the CVs remained below 0.2 $amu^{1/2}$ Å during 1.5 to 2 ps. The free energy profile along the path-CV ($s$) was then computed with Umbrella Sampling[58] with Replica Exchange,[59,60] using the location of string nodes and biasing force constant along the path CV obtained at convergence of the string as parameters for the US windows. At

least 120 ps of data were accumulated in each node to compute the free energy profile at the DFTB3/MM level. The free energy profile was finally reconstructed using the Weighted Histogram Analysis Method (WHAM) algorithm.[61]

## Analysis of interfacial contacts

The rRNA:M5 interface structure and dynamics was characterized by analysis of the contacts between the protein and rRNA, and their fluctuations. On a given structure or snapshot, two residues across the interface were considered in contact if at least one pair of their heavy atoms are distant by less than 5 Å. The list of interfacial contacts was computed on both the starting structure (after phosphate repositioning) and every 80 ps of simulation.

**Per-residue change in contacts.** For a given simulation frame (time $t$), the change in contacts with respect to the initial experimentally-derived structure for a given residue $i$ was quantified through the Manhattan distance $d_i(t)$ defined as:

$$d_i(t) = \sum_j (\Delta C_{ij}(t))^2 = \sum_j (C_{ij}(t) - C_{ij}^{ref})^2 \tag{1}$$

where $C_{ij}$ is the 2D contact matrix between protein and rRNA residues ($C_{ij} = 1$ if residues $i$ and $j$ are in contact, 0 else). How much the contacts established by a residue $i$ differ in the simulations compared to the initial structure was then quantified through the averaged quantity

$$\langle d_i(t) \rangle = \left\langle \sum_j (\Delta C_{ij}(t))^2 \right\rangle \tag{2}$$

Per-residue fluctuations $f_i$ in contacts along the MD trajectories can be quantified through the standard deviation of the distance in contact space $d_i$ :

$$f_i = \sqrt{\langle d_i^2 \rangle - \langle d_i \rangle^2} \tag{3}$$

To assess whether a residue $i$ gains or looses contacts, we assigned a sign based on the scalar product $\alpha_i$ between the sum vector $\sigma_i$ and the difference vector $\delta_i$:

$$\sigma_i(t) = \sum_j (C_{ij}(t) + C_{ij}^{pdb}) \tag{4}$$

$$\delta_i(t) = \sum_j (C_{ij}(t) - C_{ij}^{pdb}) \tag{5}$$

so that

$$\alpha_i(t) = \sigma_i(t) \cdot \delta_i(t) \tag{6}$$

If $\alpha_i > 0$, the residue at time $t$ has gained contacts compared to the starting structure, while a negative sign indicates a loss of contacts. Overall, we thus quantify for each residue the average change in its interfacial contacts with respect to the reference structure with the scalar per-residue quantity:

$$\Delta_i = \text{sign}\left(\langle \alpha_i(t) \rangle\right) \langle d_i(t) \rangle \tag{7}$$

**Clustering**

Clustering analysis was performed following the methodology and tools developed by Sabei et al..[62] Clustering was performed after concatenation of all the replicas, keeping one frame every 200 ps. The interface similarity between two structures S1 and S2 was quantified with the Jaccard index $J$ defined as

$$J(C_1, C_2) = \frac{|C_1 \cap C_2|}{|C_1 \cup C_2|} \tag{8}$$

where $C_1$ and $C_2$ represent the two sets of contacts. The Jaccard matrix is of dimension $\mathrm{N}_{frames}$ x $\mathrm{N}_{frames}$. $J = 1$ if S1 and S2 have identical contacts. On the contrary, $J = 0$ if S1 and S2 have no common contacts. Hierarchical clustering using the Ward.D2 method[63] as implemented in R was performed on the dissimilarity matrix $1 - J$ to identify clusters. The choice of the optimal number of clusters was guided by the analysis of the topology of the

clustering dendograms and the assessment of cluster size and stability in time when varying the number of clusters as proposed in ref.[62]

# Results

## Structure and stability of the active site

The initial structure of the active site[10]—obtained from superimposition of the cryo-EM structure of the M5:rRNA complex and that of the solution Xray structure of the N-terminal domain of M5—shows two $Mg^{2+}$ separated by only 4.23 Å. They are coordinated by 5 carboxylate groups from protein side chain—3 directly (Glu96, Glu10 and Asp56) and 2 through a water molecule (Asp14 and Asp58). The scissile phosphate is positioned to be in direct contact with only one of the two $Mg^{2+}$ ions ($Mg_A^{2+}$), and the oxygen of a water molecule is located in between the two cations (Figure 2a). This arrangement differs from the typical organization of 2-ion dependant nucleases, where the scissile phosphate usually occupies a bridging position in between the two active site ions.[19,64]

Simulations of a reduced model of the system—comprising M5, the protein cofactor uL18 and a fragment of rRNA terminated by a loop (Figure 1), all explicitly solvated in water with 150 mM background KCl— showed that, when starting from the structure with phosphate coordinated to a single ion, the active site structure is unstable and the distance between the two ions increases quickly (Figure S2).

By analogy with other nucleases active sites, the scissile phosphate was thus moved to a bridging position between the two ions, so that it replaces the bridging water molecule resolved in the X-ray structure of the isolated enzyme. This small rearrangement implies minimal distorsion of the overall structure (Figure 2), which remains compatible with the cryo-EM structure of the complex—of limited resolution in this area and obtained with a mutated active site. It increased the stability of the active site in simulations performed with the standard Amber99ff_OL15 force field (Figure 3a). However, the two ions still moved

apart, thus breaking apart the active site structure, in all four 500ns-replicas.

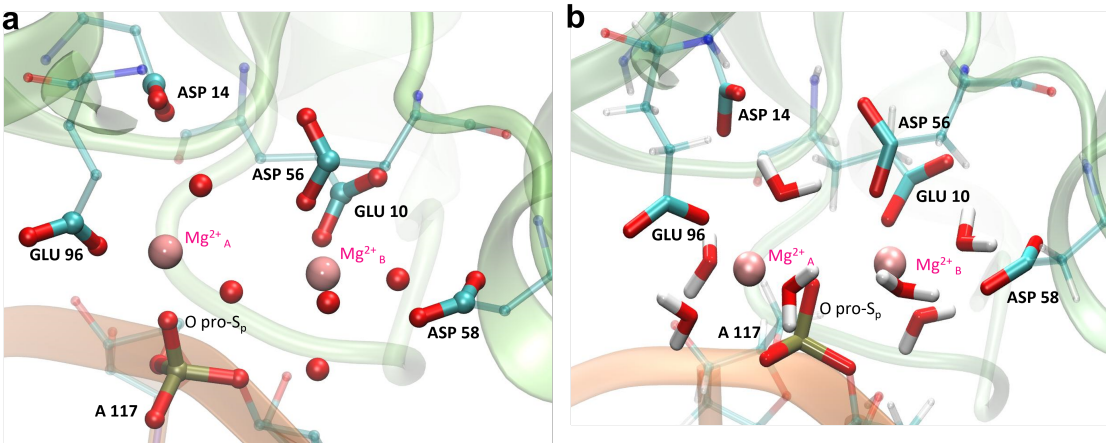


Figure 2: a) Initial structure of the active obtained from superposition of the cryo-EM structure of the complex and X-ray structure of unbound M5 at high $Mg^{2+}$ concentration. Resolved water oxygen atoms in the $Mg^{2+}$ hydration shell are drawn as red balls. b) Structure of M5 active site after repositioning of the scissile phosphate in between the two magnesium ions.

Such active sites with two divalent ions very close in space are notoriously difficult to describe with standard force fields, due to the large electrostatic repulsion between the two ions, which is overestimated in the absence of electronic polarization and charge transfer.[65] We then explored how different computational strategies could stabilize the active site structure. First, we tested a recently-proposed $Mg^{2+}$ force field[66] that includes pair-specific Lennard-Jones interactions with RNA phosphate backbone oxygen atoms, specifically tuned to properly capture $Mg^{2+}$–phosphate binding free energy. This modification did not increase the stability of the active site, which fell apart in all 4 replicas within 150 ns.

Another strategy to account for the missing electronic polarization in fixed charge force fields is to use scaled charges, which is formally equivalent to including the electronic polarization as a dielectric continuum. This mean-field theory, called Electronic Continuum Correction (ECC),[32,33] prescribes scaling all ionic charges by a factor $1/\sqrt{\epsilon_\infty} \simeq 0.75$, and has been shown to improve the description of ion–protein and ion–phosphate interactions.[31,33,67] Using ECC force fields for the active site $Mg^{2+}$ ions considerably improved the active site stability, with a stable structure over 500 ns in 2 of the 4 replicas. Using it for all charged

residues in the active site even further stabilizes the structure (see Figure 3 c,d).

These results demonstrate that advanced force fields with improved electrostatics can stabilize the active site structure reconstructed from the low-resolution cryo-EM structure of the complex and X-ray structure of the isolated protein, after a small displacement of the reactive phosphate. To complement these observations, it would be interesting in future work to assess the stability of this active structure when using explicitly polarizable force fields, that, in principle, should provide an even finer electrostatic description. The structure of the active site is then very similar to that of previously described nucleases,[11] even if it lacks the typical aspartate residue bridging the two ions, previously thought to be essential for reactivity.[11]

## Interface structural dynamics

We quantified the location and dynamics of contacts between M5 and the rRNA, on long (4 replicas of 500 ns) standard MD simulations, which we compared with two 500 ns replicas using the REST2 Hamiltonian Replica Exchange enhanced sampling methodology.[45] This technique allows to accelerate sampling of biomolecular conformations by using replicas with scaled interactions—intramolecular and intermolecular—for M5 and rRNA, which mimic a local temperature increase. To avoid spurious disruption of the active site during these simulations, the distance between both $Mg^{2+}$ and their non-water ligands was restrained (see Methods).

Overall, the interface between M5 and rRNA is stable, even in simulations with enhanced sampling, with a significant fraction of native contacts ($> 70$ %, Figure S5) at the interface preserved all along the simulations. Unsurprisingly the main structural fluctuations can be assigned to flexible RNA and protein loops, often distant from the interface (Figure S4). We quantified at the residue-level the change in interfacial contacts with respect to the initial structure using the average value of the signed distance in the space of contacts $\Delta_i$ relative to the initial structure (see Methods) (Figure 4a), and characterized the fluctuation in contacts

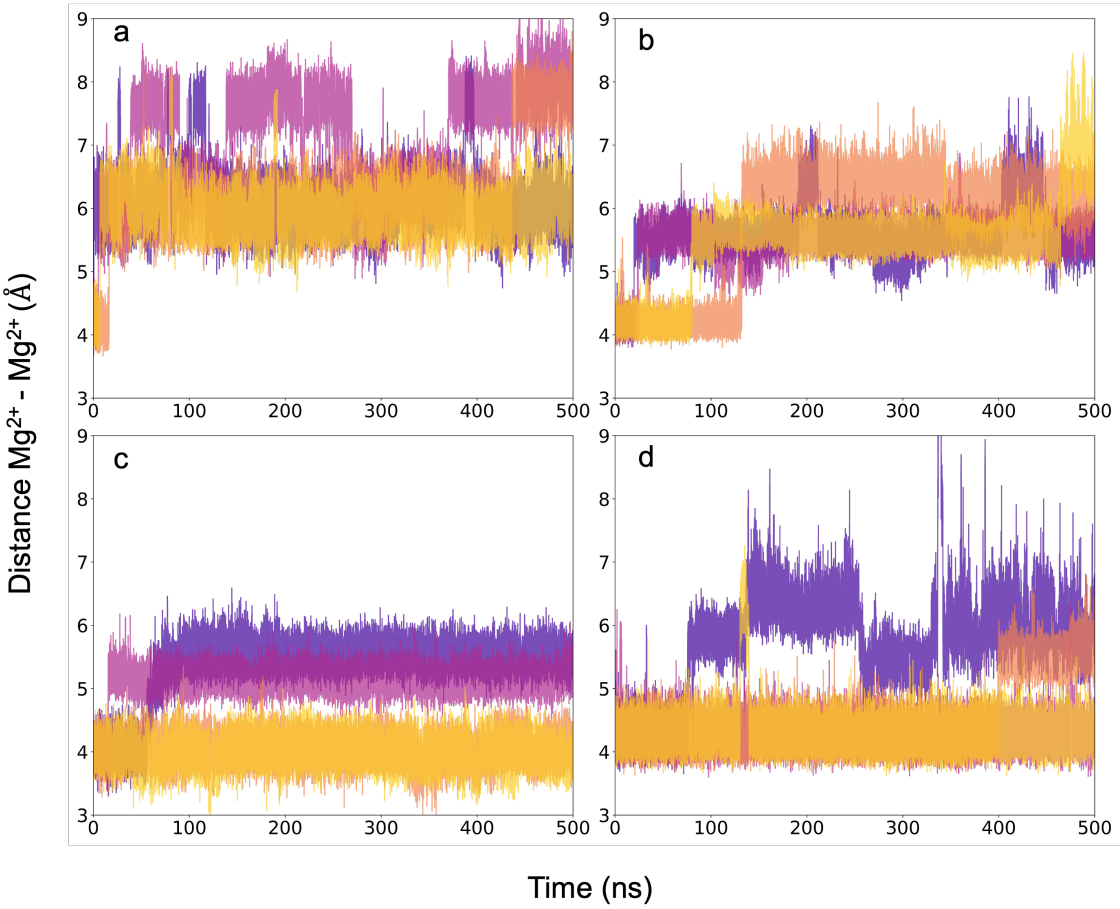


Figure 3: Time evolution of the inter-magnesium distance over 4 replicas of 500 ns with a) the standard full-charge force field, b) pair-specific Lennard-Jones parameters for $Mg^{2+}$, c) our ECC force field for active site $Mg^{2+}$ ions or d) our ECC force field for all active site residues.

along the simulations with the per-residue variability of contacts $f_i$ (see Methods). In non-accelerated simulations, the largest changes and flexibility are observed around the U23 RNA base—which is inserted in between the two M5 protein domains— and for the C116 nucleotide, which interacts with the residues around Lys91, in an unstructured loop of M5 N-terminal domain. Large changes in contacts are also found for A55 in the domain $\beta$, which changes interaction patterns with the residues around Met136, at the foot of M5 C-terminal domain $\alpha$7-helix. Three distinct conformations of the interface were identified by performing

clustering in the contact space (Figure S6), without large conformational changes from the initial experimentally-derived structure. Rather, we observe small side chain rearrangements —e.g. formation of additional salt bridge interactions between protein side chains and RNA phosphate backbone or insertion of the U23 RNA base deeper in the groove between the two protein domains.

The same conformations are visited in REST2 simulations (Figure 4c), which share similar clusters, but enhanced sampling reveals alternative conformations (additional clusters) with fraying of the initially paired rRNA 3′ and 5′ ends (Figure 4e). This allows new interactions between the now unpaired rRNA ends and the protein $\alpha 2$ helix, that were not sampled with standard MD.

Overall, our simulations thus picture a stable rRNA:M5 interface, with significant fluctuations observed mostly at the RNA ends, where possible fraying was uncovered thanks to enhanced sampling simulations. At this stage, it is however unclear how much these fluctuations could impact the RNase M5 activity, which we thus proceed to investigate.

Monitoring selected distances across the active site along our enhanced sampling simulations, we revealed a rich conformational landscape in the active site, with multiple basins corresponding to various orientations of the active site Glu96 and distinct distances between the backbone phosphate A118 and the $Mg^{2+}$ ion ($Mg_A^{2+}$) coordinating the attacking water molecule (see Figure 5a). The center of the basin corresponding to our starting structure is associated with long O2P–$Mg_A$ and OE1(Glu96)–P(A117) distances ($\sim 6.5 - 7\,$Å), which means that both pairs of groups are separated by more than a single water molecule. However, we identify secondary conformations on the side of this broad basin, with a free energy only 2.0 kcal/mol higher, with Glu96 OE1 oxygen much closer ($< 6.0\,$Å) to the phosphate, which reflects the presence of a single molecule in between those two groups (see Figure 5c).

In addition to this main broad basin, we evidence a secondary basin with a much shorter distance ($\sim 4\,$Å) beetween the neighbor backbone phosphate O2P(A118) and $Mg_A^{2+}$. This corresponds to a conformation with a single water molecule bridging those two groups (see

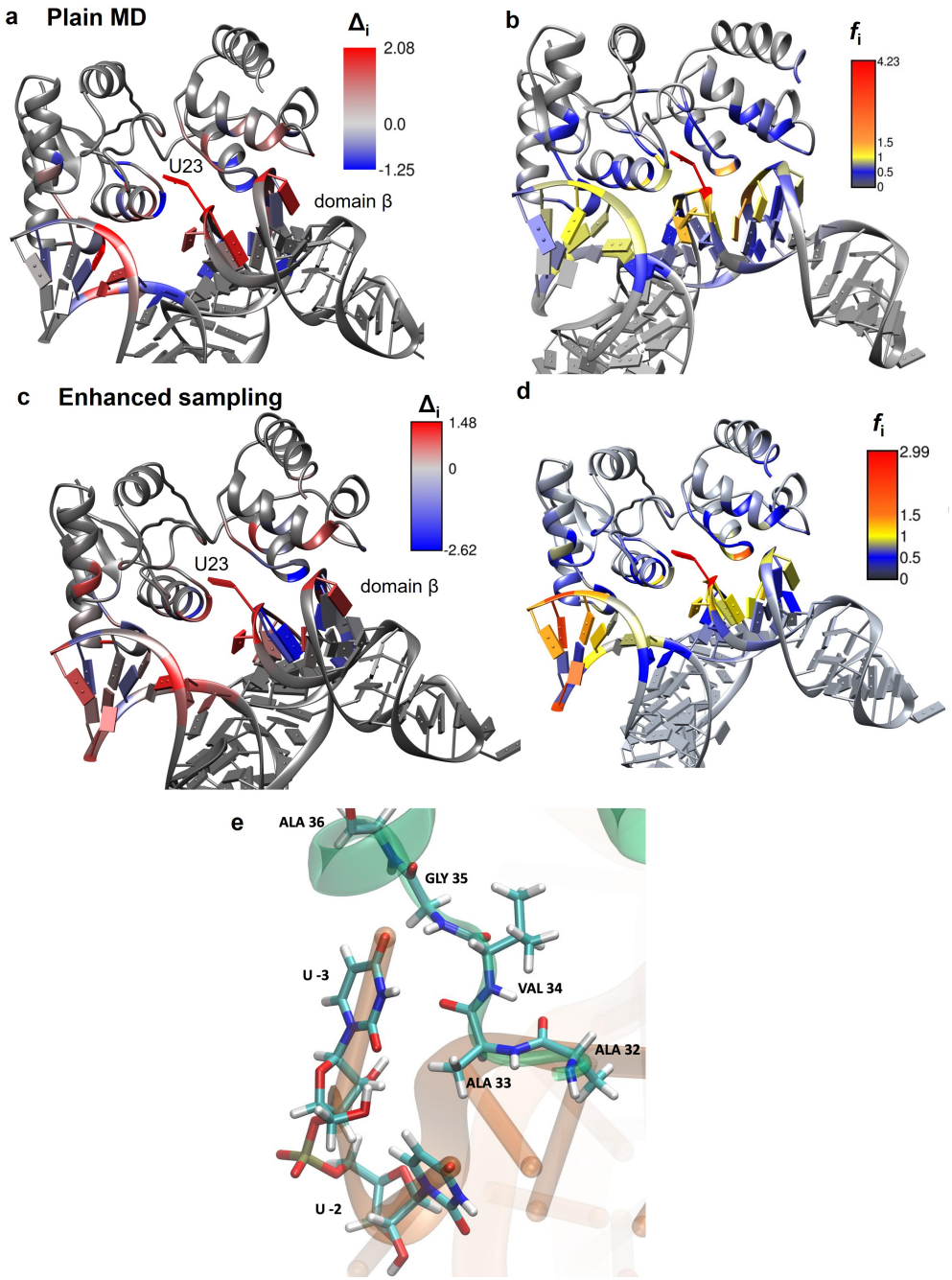


Figure 4: a) Per-residue change $\Delta_i$ ($> 0$ if gain, $< 0$ if loss) in interfacial contacts along 4x500 ns of MD simulation with respect to the initial structure of the complex, and b) per-residue fluctuation $f_i$ in contacts. c) Per-residue change $\Delta_i$ in interfacial contacts and d) per-residue fluctuation $f_i$ in contacts along 2x500 ns of REST2 enhanced sampling simulations. e) Example of conformation with fraying at the RNA ends as sampled in our REST2 simulation.

Figure 5c). From our REST2 simulations, we estimate that this conformation could be almost as populated as the main one (see SI). Further, we found the existence of these distinct

conformations of the active site structure to be correlated with the changes in interfacial contacts examined earlier (Figure S10), as it is a change in the conformation of the RNA ends that brings A118 phosphate closer to the active site (Figure S8). These distinct active site conformations are thus associated with different relative positioning of the residues (in particular Glu96 and A118) in the active site, which may favor distinct reaction mechanisms, as we will now further investigate.

## QM/MM-MD investigation of the reaction mechanism

The cleavage reaction of rRNA by M5 occurs through the nucleophilic attack of a water molecule to the scissile phosphate, followed by departure of the 3′ leaving group (see Figure 6), similarly to other RNases.[11] However, it is a priori unclear whether the attacking water molecule is deprotonated before the attack—and in that case, where the proton is transfered, either to Glu96 or a backbone phosphate—or whether its proton is shuttled to the leaving group through the scissile phosphate oxygen atoms (see Figure 6). In addition, the two proton transfers can happen in a more or less concerted fashion with the nucleophilic attack and departure of the 3′ leaving group. Only few computational studies of RNases are reported in the literature, almost exclusively on RNase H,[12–16] where these points are still debated, including the identity of a possible general acid to reprotonate the leaving group. We thus setup to investigate the different possible mechanistic routes M5 from the different identified conformations, exploring the interplay between conformational dynamics and reactivity.

Given the multiplicity of possible reaction mechanisms, the number of chemical steps and the need to capture active site reorganization along the reaction path, we chose to use a path-based method—The Adaptive Finite Temperature String Method[57] (ASM) developed by one of us—to dynamically explore the reaction free energy landscape in a multi-dimensional space of collective variables (see Methods). Given the computational cost of such QM/MM-MD

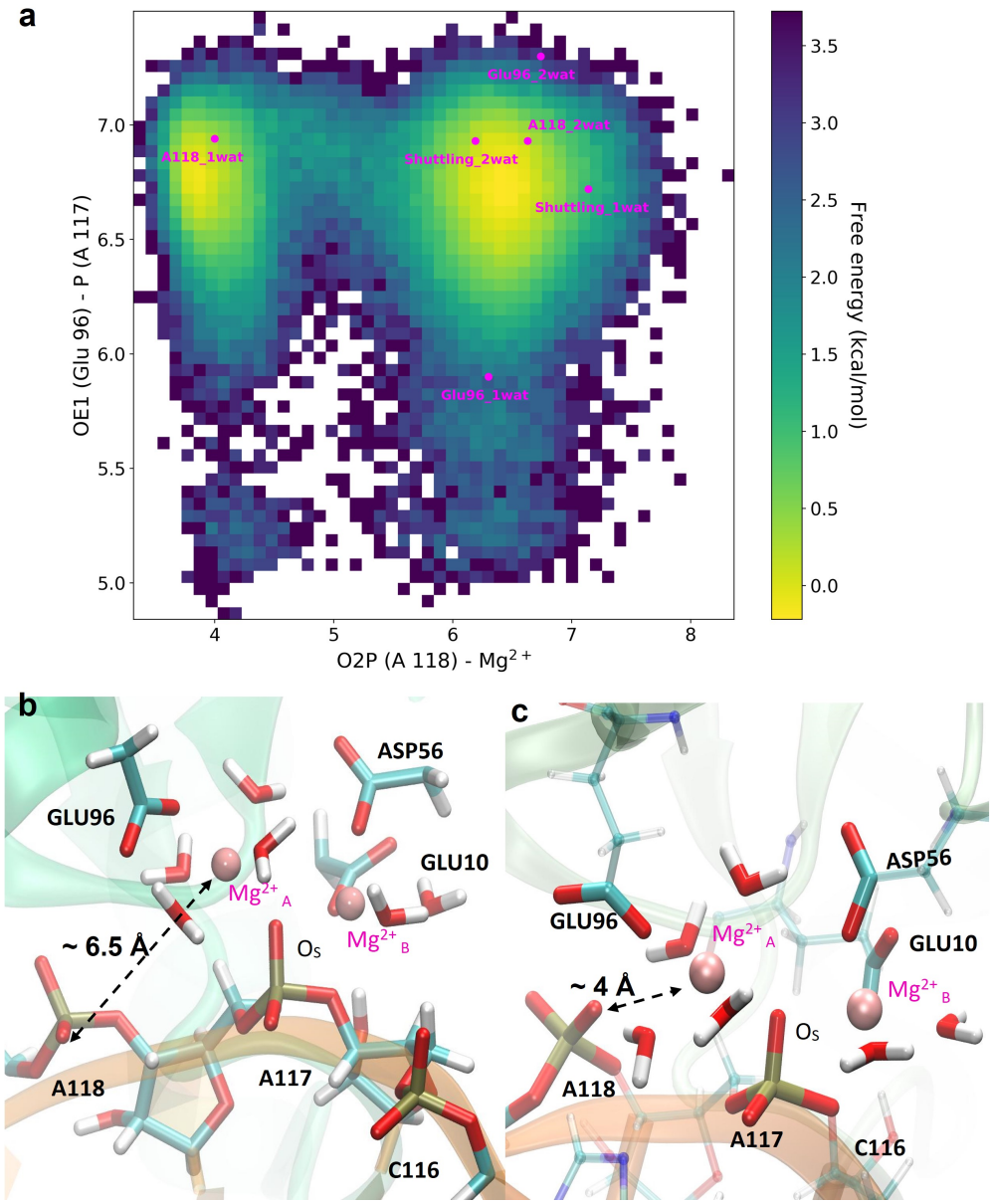


Figure 5: a) Free energy lanscape along two key active site distances computed from the 2D probability distributions obtained in our enhanced sampling simulations. b) Representative structure of a typical conformations of the M5 active site with Glu96 poised to deprotonate the attacking water molecule and b) with the A118 phosphate positioned to deprotonate the nucleophilic water molecule, in a bridge position between A118 and a $Mg^{2+}$ ion. Prepared with VMD[38]

simulations and to allow for sufficient sampling—typically, around 500 ps are needed to converge the path—initial explorations of the different mechanistic hypotheses were performed at the DFTB3[48]/MM level of theory. The most promising mechanism was then be

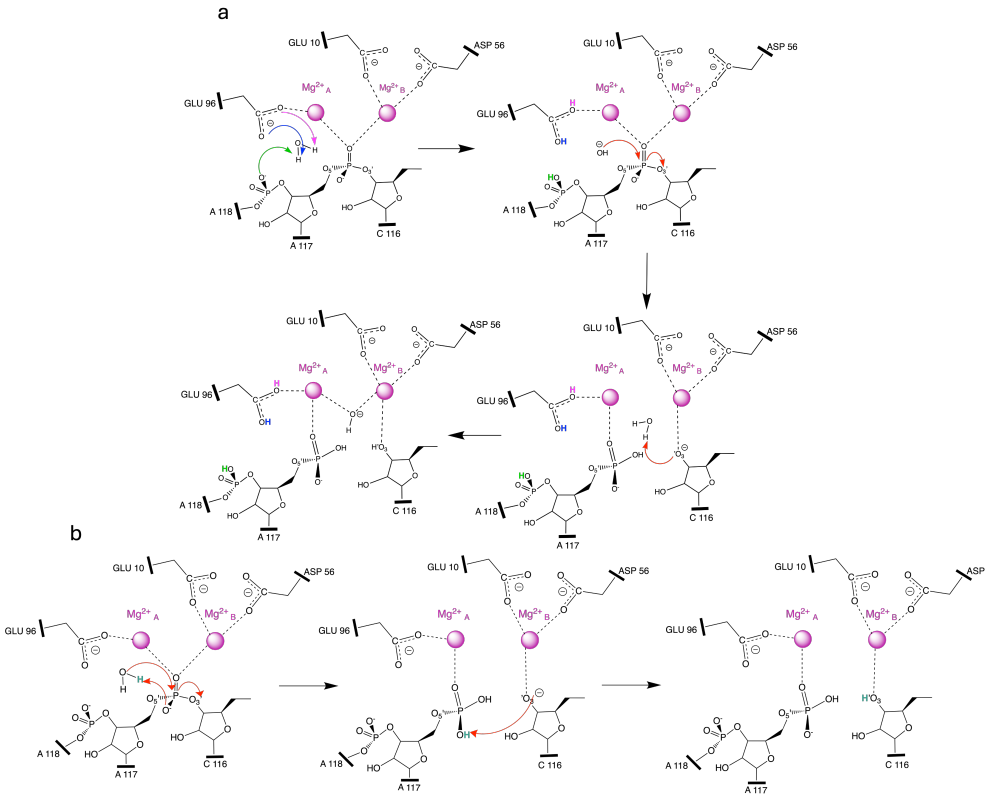


Figure 6: a)Scheme of three possible mechanisms (represented with the different colored arrows for PT1) for M5 rRNA cleavage b) Scheme of the proton shuttling mechanism.

recalculated at the B3LYP/MM level.

The different conformations of the active site previously identified seem prone to favor distinct mechanisms. In particular, the conformation with a single water molecule—in the $Mg^{2+}$ hydration shell—bridging the carboxylate group of Glu96 and the reactive phosphate (see starting point "Glu96_1wat" on Figure 5a and Figure 5b) seems ideally pre-organized for deprotonation of the nucleophilic water molecule directly by Glu96. Structures from the secondary basin ("A118_1wat" on Figure 5a and Figure 5c) with a single water molecule bridging the $Mg_A^{2+}$ and the neighbor A118 phosphate pro-$R_p$ ($O_R$) would in contrast be better arranged for deprotonation by the phosphate backbone. Both deprotonation routes could also take place from conformations with increased distances between the base and scissile phosphate ("A118_2wat" and "Glu96_2wat" in Figure 5a), with indirect proton transfer through an additional water molecule (Figure S11). A final possible mechanistic

route is the attack of a non-deprotonated water molecule, with proton shuttling towards the leaving group through the scissile A117 phosphate. This proton shuttling can also involve a string of several water molecules, as suggested in RNase H.[12] The different possible reaction mechanisms are depicted in Figure 6 and Figure S11.

Using the ASM methodology, we characterized the reaction pathway associated to all these possible mechanisms, starting in each case from the a priori most favorable conformation. The computed free-energy barriers are summarized in Table 1. Note that we could not converge any reaction pathway starting with deprotonation by the $Mg^{2+}$-bound Glu96 oxygen (purple arrow in Fig 6) which thus seems highly unfavored. Of all the explored mechanisms, direct deprotonation by the Glu96 oxygen non involved in $Mg^{2+}$ coordination appears by far the most favorable, with a free energy barrier $\Delta G^{\ddagger} \simeq 18.2$ kcal mol$^{-1}$, while all the other mechanisms are associated with barriers over 30 kcal mol$^{-1}$ (Figures S12, S13, S14).

Table 1: Activation free energies and name of the starting point (localized on Fig. 5a) for all tested reaction mechanisms.

| **Mechanism** | **Name** | $\Delta G^{\ddagger}$ **(kcal/mol)** |
|---|---|---|
| Deprotonation by Glu96 (direct) | Glu96_1wat | 18.2 |
| Deprotonation by Glu96 (via 2 $H_2O$) | Glu96_2wat | 32.1 |
| Deprotonation by A118 (direct) | A118_1wat | 46.0 |
| Deprotonation by A118 (via 2 $H_2O$ ) | A118_2wat | 44.2 |
| Proton shuttling (via 1 $H_2O$) | Shuttling_1wat | 36.9 |
| Proton shuttling (via 2 $H_2O$) | Shuttling_2wat | 35.5 |

The reaction starts with a proton transfer from a water bound to $Mg_A^{2+}$ to the carboxylate oxygen of Glu96 not involved in $Mg_A^{2+}$ coordination. This first deprotonation event is associated with a very small barrier and the deprotonated state has a very similar free energy as the initial state, the hydroxide ion being stabilized by $Mg_A^{2+}$. In a second step, the hydroxide ion attacks the scissile phosphate, to reach an associative transition state with a pentavalent phosphate geometry (Figure 7c). Departure of the leaving group is then associated with only small free energy barriers, as is its reprotonation by a water molecule located in between

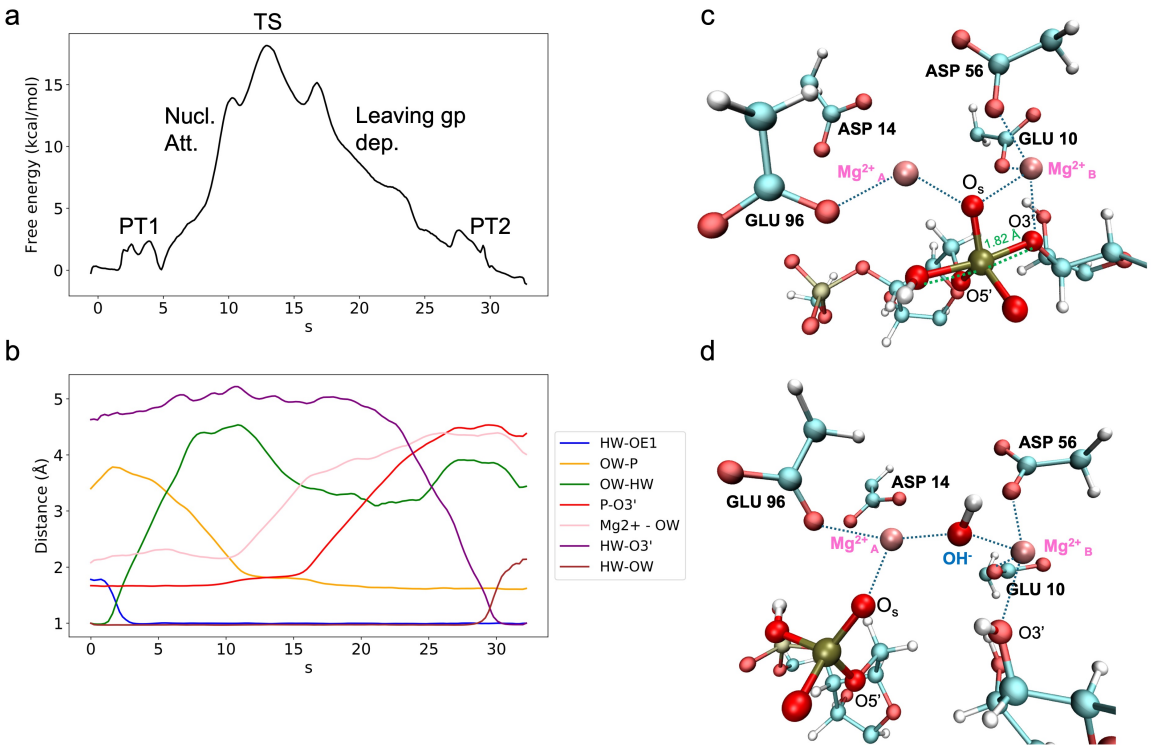


Figure 7: a) Free energy profile at the DFTB3/MM-MD level (*will be updated with DFT/MM data*) along the reaction path coordinate $s$. b) Evolution of the defined CVs along the reaction pathway. c) Representative structure of the phosphorane transition state with P-O3′ and P-nucleophile distances of 1.82 Å. d) Representative structure of the product, with the hydroxide ion between the two $Mg^{2+}$ ions.

the two $Mg^{2+}$ ions. While our initial mechanistic guess did not impose any reprotonation of the leaving group, a $Mg^{2+}$-bound water molecule spontaneously oriented itself to reprotonate the leaving group during our initial explorations of the reaction free energy landscape. In the course of the reaction, the scissile phosphate, initially coordinating both $Mg^{2+}$ ions, moves towards $Mg^{2+}_A$ and is replaced by a water molecule that then donates its proton to reprotonate the 3′OH leaving group. In the final state, a hydroxide ion is thus located in between the two active site cations (Figure 7d).

We note that the selected starting structure for the mechanism is not in the center of the main population basin, as Glu96 is closer to the scissile phosphate than in the dominant geometry (see Figure 5a), which allows the nucleophilic water molecule to be directly deprotonated by Glu96. In the dominant conformation of the active site, direct proton transfer is

not possible, but could instead happen through a string of two water molecules (Figure S11). Characterization of the free energy profile associated with such a mechanism however shows that it is associated with a much higher free energy barrier (see Table 1 and Figures S12, S13, S14). These results thus suggest that from the dominant conformation of the active site, the most favorable route would be first a conformational fluctuation to bring Glu96 in an appropriate geometry (with an estimated free energy cost of about 2 kcal $mol^{-1}$), followed by direct deprotonation of the nucleophilic water molecule by Glu96 from this specific geometry.

# Discussion

**Active site structure and fluctuations.** The experimentally determined M5 active site structure[10] is similar to that of other RNases reported in the literature,[64] with 2 $Mg^{2+}$ ions coordinated by several negatively charged protein side chains, either directly or through water molecules. While the initial superimposition of the cryoEM structure of the complex—mutated to prevent reactivity—and the X-ray structures of M5 isolated domains suggested an unusual position of the scissile phosphate, coordinating only one of the two ions, this structure is unstable in simulations. We propose that the scissile phosphate should instead, as in other RNases, be positioned in between the two active site ions, thus effectively replacing the water molecule resolved in the crystal structure of the isolated N-terminal domain. This (small) change being done, the active site structure now closely resembles that of other RNases, in particular the well-characterized RNase H.[18,68]

Yet, substantial differences can be noticed. First, the binding of two active site aspartate residues is mediated by water molecules in M5—in contrast with direct inner sphere coordination in RNase H. In addition, while the presence of an aspartate that simultaneously binds both $Mg^{2+}$ ions is often reported to be a key element for 2-ion catalysis,[11,69] M5 is missing such a feature. Our present work suggests that the active site is stable—provided an

adequate force field is used in simulations—and that a reasonable reaction free energy path exist without this feature, which may not be as essential as previously thought.

Also, RNase H crystal structures have suggested[64] that the cation occupying the A site has a perfect octahedral hydration shell, while site B coordination is initially frustrated, with a distorted hydration geometry. Release of this frustration was thought[64] to play a key role in catalysis. The available crystal structure of M5 N-terminal domain[10] paints a very different picture, with a perfect octahedral coordination for site B $Mg^{2+}$ and a distorted square-planar coordination in site A. Note that this square planar coordination may be due to limited resolution of structures, difficulty to identify, and that in our molecular simulations both cations have a octahedral hydration shell, with 6 ligands coordinated. Since a thermodynamically favorable reaction pathway, with a low free-energy barrier could still be found in simulations, release of an initial geometric frustration does not seem to be an essential feature for these reactions.

Finally, RNase H structures[18] suggest that the nucleophilic water molecule, coordinated to the $Mg^{2+}{}_{A}$ ion, could donate a H-bond to the downstream phosphate. This feature led computational studies[12,14] to consider the downstream phosphate as a potential general base for the initial proton transfer. The starting structure of the M5:rRNA complex does not exhibit such a feature, with a downstream phosphate (A118) too far from the active site to allow formation of a single-water bridge. However, our enhanced sampling simulations revealed the conformational flexibility of the active site, and the existence of an alternative conformation where the A118 phosphate group has moved close enough to be bridged to one of the active site cations by a single water molecule, similarly to what was observed in RNase H. Our simulations further suggest that the two conformations could have similar populations. While not explicitly discussed and quantified in the case of RNase H, such flexibility of the active site may be also relevant there, as earlier studies[14] mention that the interaction between the nucleophilic water molecule and downstream phosphate may be either direct or mediated by another water molecule.

This suggests that the conformational fluctuations of the M5 active site, which we quantified with enhanced sampling simulations, may have a broader relevance for RNases. We found that these distinct conformations of the active sites where correlated with larger scale conformational fluctuations of the interface, and notably with fraying motions at the RNA ends. Here, we have to pause and note that *in vivo*, the RNA 3′ end is much longer than in our simulations (42 nucleotides), but this part of the RNA is unstructured and not resolved in the cryoEM structure of the complex. This suggests that the RNA 3′ end is indeed very flexible and prone to adopt varied conformations. While the exact populations of different conformational states may differ in the full-length system, we can thus reasonably assume that large conformational fluctuations may occur at the RNA ends and lead to distinct populations of active site structure.

**RNA cleavage mechanism** Previous computational studies of 2-ion RNases mostly focused on RNase H,[12–16] and suggested different, possibly co-existing, mechanisms for the cleavage. A first study[12] singled out two possible mechanisms: a first one where the proton of the attacking water molecule is shuttled to the leaving group through several water molecules (similar to our "Shuttling_2wat" mechanism), and a second where the reaction, starting with the attack of a hydroxide anion (deprotonated water), ends with stabilization of the deprotonated 3′OH by the site B $Mg^{2+}$ ion. Another study[14]—starting from a slightly different active site conformation with an additional Glutamate Glu188 bound to site A $Mg^{2+}$, which again echos our observation of a rich conformational landscape in such active sites—suggests that the water nucleophile donates its proton to the downstream phosphate (similar to our "A118_1wat" mechanism) and that the leaving group is reprotonated by an aspartate residue, initially protonated. It however could not exclude alternative reprotonation routes that were not explicitly considered.

In this work, we evidenced an interplay between reactivity and conformational fluctuations in the active site, as key aspects of the reaction (e.g. nature of the general base)

are intimately coupled to conformational fluctuations of the active site, that modulate the distance between reactive partners. Distinct conformations of the active site appear pre-organized to favor different mechanisms. This echos recent works that increasingly suggest an interplay between conformational fluctuations and catalytic activity in enzymes.[70,71] This work thus underlines the importance to perform extensive conformational exploration of the reactant state of a reactive complex before mechanistic investigations.

Our mechanistic investigations evidenced one favorable reaction route, that first requires a small conformational fluctuation from the dominant active site conformation to reach a suitable organization of the active site. This RNA cleavage mechanism differs from that suggested for RNase H in several aspects: the most favorable route for deprotonation of the nucleophilic water molecule is here a $Mg^{2+}$-bound glutamate, and the leaving group is then reprotonated by a water molecule that moves along the reaction path to adopt a bridging position between the two active site $Mg^{2+}$ ions, thus taking the place of the scissile phosphate in the reactant state. Note that a water molecule is visible in the X-ray structure of isolated M5 at this specific position, with ill-determined protonation state, so that a hydroxide ion may well bridge the two $Mg^{2+}$ ions in the absence of the RNA subtrate. Our product structure could thus be reminiscent of that of the uncomplexed active site. Experimentally testing the nature of the base for the initial deprotonation will prove challenging: mutating Glu96 would presumably prevent the initial proton transfer, but it may also be detrimental for $Mg^{2+}$ binding in the active site, and the two effects will be difficult to deconvolute.

While we evidenced here a clearly favored reaction pathway from one of the conformational basins, it would be interesting to investigate in future work whether such a rich conformational landscape, relevant for reactivity, is pertinent for other RNases. The existence of distinct reaction pathways associated with different conformations paints a picture of enzyme catalysis much more complex than the traditional view. In cases of competing reaction routes with similar barriers, such effects may be key to rationalize reactivity, understand mutation effects, and predict activity changes, for instance in the context of enzyme

design.

# Conclusion

Molecular simulations of the complex between M5 and its rRNA substrate have thus demonstrated the plausibility of the experimentally suggested complex and inferred active site structure, which are—provided a small displacement of the scissile phosphate and the use of specialized force fields—stable and with a reasonable reactive pathway. Combining advanced enhanced sampling techniques and dynamic QM/MM-MD reaction path exploration, we have uncovered a rich conformational landscape in the active site and characterized different mechanistic routes associated with various orientations of key active site residues. The most favorable mechanism features deprotonation of the nucleophilic water molecule by a $Mg^{2+}$-bound glutamate (Glu96). The formed hydroxide ion, stabilized by the $Mg^{2+}$, then attacks the scissile phosphate in an associative mechanism with a pentavalent phosphate at the transition state. The leaving group is later reprotonated by a water molecule, whose acidity is exacerbated by coordination to both $Mg^{2+}$ ions. Such a mechanism is not possible from the main conformational basin, so that the system must wait for a conformational fluctuation to bring it in a reactive conformation.

From a methodological point of view, we stress that enhanced sampling techniques were key to uncover these distinct conformational states of the active sites and characterize the associated free energy landscape. This work thus underlines the importance to perform extensive conformational exploration of the reactant state of a reactive complex before mechanistic investigations, echoing in that recent work on different RNA processing systems, ribozymes.[71]

Finally, while the present study focused on the cleavage of the 3′ end by M5, experiments suggest[9] that the complex then undergoes reorganization to position the same M5 enzyme on the 5′ end to perform the second cleavage step. While M5 presumably remains anchored

to the rRNA with its C-terminal domain, significant reorganization of the complex has to occur after removal of the 3′ end to allow for the needed repositionning. Characterization of this reorganization is a challenge both for experiments and computations that will be tackled in future work.

## Acknowledgement

This work was supported by the "Initiative d'Excellence" program from the French State (Grants "DYNAMO", ANR-11-LABX-0011, and "CACSICE", ANR-11-EQPX-0008). JPF acknowledges a PhD grant from the DYNAMO Labex. EDD and LK acknowledge support from the ANR JCJC MUSIRICAT (ANR-22-CE29-0004-01). EDD acknowledge an IEA grant from CNRS. This project was provided with computing HPC and storage resources by GENCI at CINES, TGCC and IDRIS (project A0150711021).

## Supporting Information Available

Supporting Information is available, with additional computational details and complementary analyses. Input and parameter files for all simulations *will be* shared on a Zenodo public folder.[36]

# References


(1) Srivastava, A. K.; Schlessinger, D. Mechanism and Regulation of Bacterial Ribosomal RNA Processing. *Annu. Rev. Microbiol.* **1990**, *44*, 105–129.

(2) Deutscher, M. P. Twenty Years of Bacterial RNases and RNA Processing: How We've Matured. *RNA* **2015**, *21*, 597–600.

(3) Sogin, M. L.; Pace, B.; Pace, N. R. Partial Purification and Properties of a Ribosomal RNA Maturation Endonuclease from Bacillus Subtilis. *J. Biol. Chem.* **1977**, *252*, 1350–1357.

(4) Stahl, D. A.; Pace, B.; Marsh, T.; Pace, N. R. The Ribonucleoprotein Substrate for a Ribosomal RNA-processing Nuclease. *J. Biol. Chem.* **1984**, *259*, 11448–11453.

(5) Condon, C.; Brechemier-Baey, D.; Beltchev, B.; Grunberg-Manago, M.; Putzer, H. Identification of the Gene Encoding the 5S Ribosomal RNA Maturase in Bacillus Subtilis: Mature 5S rRNA Is Dispensable for Ribosome Function. *RNA* **2001**, *7*, 242–253.

(6) Aravind, L.; Leipe, D. D.; Koonin, E. V. Toprim—a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. *Nucleic Acids Res.* **1998**, *26*, 4205–4213.

(7) Allemand, F.; Mathy, N.; Brechemier-Baey, D.; Condon, C. The 5S rRNA Maturase, Ribonuclease M5, Is a Toprim Domain Family Member. *Nucleic Acids Res.* **2005**, *33*, 4368–4376.

(8) Oerum, S.; Dendooven, T.; Catala, M.; Gilet, L.; Dégut, C.; Trinquier, A.; Bourguet, M.; Barraud, P.; Cianferani, S.; Luisi, B. F.; Condon, C.; Tisné, C. Structures of B. Subtilis Maturation RNases Captured on 50S Ribosome with Pre-rRNAs. *Mol. Cell* **2020**, *80*, 1–10.

(9) Oerum, S.; Dendooven, T.; Catala, M.; Gilet, L.; Dégut, C.; Trinquier, A.; Bourguet, M.; Barraud, P.; Cianferani, S.; Luisi, B. F.; Others, Structures of B. Subtilis Maturation RNases Captured on 50S Ribosome with Pre-rRNAs. *Mol. Cell* **2020**, *80*, 227–236.

(10) Oerum, S.; Catala, M.; Bourguet, M.; Gilet, L.; Barraud, P.; Cianférani, S.; Condon, C.;

Tisné, C. Structural Studies of RNase M5 Reveal Two-Metal-Ion Supported Two-Step dsRNA Cleavage for 5S rRNA Maturation. *RNA Biol.* **2021**, *18*, 1996–2006.

(11) Yang, W. Nucleases: Diversity of Structure, Function and Mechanism. *Q. Rev. Biophys.* **2011**, *44*, 1–93.

(12) De Vivo, M.; Dal Peraro, M.; Klein, M. L. Phosphodiester Cleavage in Ribonuclease H Occurs via an Associative Two-Metal-Aided Catalytic Mechanism. *J. Am. Chem. Soc.* **2008**, *130*, 10955–10962.

(13) Ho, M. H.; De Vivo, M.; Dal Peraro, M.; Klein, M. L. Understanding the Effect of Magnesium Ion Concentration on the Catalytic Activity of Ribonuclease H through Computation: Does a Third Metal Binding Site Modulate Endonuclease Catalysis? *J. Am. Chem. Soc.* **2010**, *132*, 13702–13712.

(14) Rosta, E.; Nowotny, M.; Yang, W.; Hummer, G. Catalytic Mechanism of RNA Backbone Cleavage by Ribonuclease H from Quantum Mechanics/Molecular Mechanics Simulations. *J. Am. Chem. Soc.* **2011**, *133*, 8934–8941.

(15) Rosta, E.; Yang, W.; Hummer, G. Calcium Inhibition of Ribonuclease H1 Two-Metal Ion Catalysis. *J. Am. Chem. Soc.* **2014**, *136*, 3137–3144.

(16) Sgrignani, J.; Magistrato, A. QM/MM MD Simulations on the Enzymatic Pathway of the Human Flap Endonuclease (hFEN1) Elucidating Common Cleavage Pathways to RNase H Enzymes. *ACS Catal.* **2015**, *5*, 3864–3875.

(17) Yang, W. *Nucleases: Diversity of Structure, Function and Mechanism*; 2011; Vol. 44.

(18) Nowotny, M.; Gaidamakov, S. A.; Crouch, R. J.; Yang, W. Crystal Structures of RNase H Bound to an RNA/DNA Hybrid: Substrate Specificity and Metal-Dependent Catalysis. *Cell* **2005**, *121*, 1005–1016.

(19) Nowotny, M.; Yang, W. Stepwise Analyses of Metal Ions in RNase H Catalysis from Substrate Destabilization to Product Release. *EMBO J.* **2006**, *25*, 1924–1933.

(20) Maria-Solano, M. A.; Serrano-Hervás, E.; Romero-Rivera, A.; Iglesias-Fernández, J.; Osuna, S. Role of conformational dynamics in the evolution of novel enzyme function. *Chem. Commun.* **2018**, *54*, 6622–6634.

(21) DeLano, W. L.; Others, Pymol: An Open-Source Molecular Graphics Tool. *CCP4 Newsl, Protein Crystallogr* **2002**, *40*, 82–92.

(22) Sarzynska, J.; Popenda, M.; Antczak, M.; Szachniuk, M. RNA Tertiary Structure Prediction Using RNAComposer in CASP15. *Proteins Struct. Funct. Bioinf.* **2023**, *91*, 1790–1799.

(23) Popenda, M.; Szachniuk, M.; Antczak, M.; Purzycka, K. J.; Lukasiak, P.; Bartol, N.; Blazewicz, J.; Adamiak, R. W. Automated 3D Structure Composition for Large RNAs. *Nucleic Acids Res.* **2012**, *40*, e112–e112.

(24) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

(25) Pérez, A.; Marchán, 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.

(26) Zgarbová, M.; Otyepka, M.; Šponer, J.; Mládek, A.; Banáš, P.; Cheatham, T. E.; Jurečka, 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.

(27) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(28) Joung, I. S.; Cheatham, T. E. I. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112*, 9020–9041.

(29) Li, P.; Song, L. F.; Merz, K. M. Parameterization of Highly Charged Metal Ions Using the 12-6-4 LJ-type Nonbonded Model in Explicit Water. *J. Phys. Chem. B* **2015**, *119*, 883–895.

(30) Grotz, K. K.; Cruz-León, S.; Schwierz, N. Optimized Magnesium Force Field Parameters for Biomolecular Simulations with Accurate Solvation, Ion-Binding, and Water-Exchange Properties. *J. Chem. Theory Comput.* **2021**, *17*, 2530–2540.

(31) Puyo-Fourtine, J.; Juillé, M.; Hénin, J.; Clavaguéra, C.; Duboué-Dijon, E. A Consistent Picture of Phosphate-Divalent Cation Binding from Models with Implicit and Explicit Electronic Polarization. *J. Phys. Chem. B* **2022**, *126*, 4022–4034.

(32) Leontyev, I. V.; Stuchebrukhov, A. A. Electronic Continuum Model for Molecular Dynamics Simulations. *J. Chem. Phys.* **2009**, *130*, 85 102.

(33) Kirby, B. J.; Jungwirth, P. Charge Scaling Manifesto: A Way of Reconciling the Inherently Macroscopic and Microscopic Natures of Molecular Simulations. *J. Phys. Chem. Lett.* **2019**, *10*, 7531–7536.

(34) Duboué-Dijon, E.; Javanainen, M.; Delcroix, P.; Jungwirth, P.; Martinez-Seara, H. A Practical Guide to Biologically Relevant Molecular Simulations with Charge Scaling for Electronic Polarization. *J. Chem. Phys.* **2020**, *153*.

(35) Berendsen, H. J.; Grigera, J.-R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

(36) Zenodo repository.

(37) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

(38) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.

(39) Fiorin, G.; Klein, M. L.; Hénin, J. Using Collective Variables to Drive Molecular Dynamics Simulations. *Mol. Phys.* **2013**, *111*, 3345–3362.

(40) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

(41) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 14101.

(42) Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.

(43) Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13*, 952–962.

(44) 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.

(45) Wang, L.; Friesner, R. A.; Berne, B. Correction to "Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2)". *J. Phys. Chem. B* **2011**, *115*, 11305–11305.

(46) Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

(47) Case, D. A.; Aktulga, H. M.; Belfon, K.; Cerutti, D. S.; Cisneros, G. A.; Cruzeiro, V. W. D.; Forouzesh, N.; Giese, T. J.; Götz, A. W.; Gohlke, H.; Others, AmberTools. *J. Chem. Inf. Model.* **2023**, *63*, 6183–6191.

(48) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*, 931–948.

(49) Gaus, M.; Lu, X.; Elstner, M.; Cui, Q. Parameterization of DFTB3/3OB for Sulfur and Phosphorus for Chemical and Biological Applications. *J. Chem. Theory Comput.* **2014**, *10*, 1518–1537.

(50) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(51) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785–789.

(52) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A: At. Mol. Opt. Phys.* **1988**, *38*, 3098–3100.

(53) Kamerlin, S. C.; Florián, J.; Warshel, A. Associative versus dissociative mechanisms of phosphate monoester hydrolysis: on the interpretation of activation entropies. *ChemPhysChem* **2008**, *9*, 1767–1773.

(54) Lu, X.; Gaus, M.; Elstner, M.; Cui, Q. Parametrization of DFTB3/3OB for magnesium and zinc for chemical and biological applications. *J. Phys. Chem. B* **2015**, *119*, 1062–1082.

(55) Kostal, V.; Mason, P. E.; Martinez-Seara, H.; Jungwirth, P. Common Cations Are Not Polarizable: Effects of Dispersion Correction on Hydration Structures from Ab Initio Molecular Dynamics. *J. Phys. Chem. Lett.* **2023**, *14*, 4403–4408.

(56) Shirts, M. R.; Klein, C.; Swails, J. M.; Yin, J.; Gilson, M. K.; Mobley, D. L.; Case, D. A.; Zhong, E. D. Lessons Learned from Comparing Molecular Dynamics Engines on the SAMPL5 Dataset. *J. Comput.-Aided Mol. Des.* **2017**, *31*, 147–161.

(57) Zinovjev, K.; Tuñón, I. Adaptive Finite Temperature String Method in Collective Variables. *J. Phys. Chem. A* **2017**, *121*, 9764–9772.

(58) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.

(59) Sabri Dashti, D.; Roitberg, A. E. Optimization of Umbrella Sampling Replica Exchange Molecular Dynamics by Replica Positioning. *J. Chem. Theory Comput.* **2013**, *9*, 4692–4699.

(60) Sugita, Y.; Kitao, A.; Okamoto, Y. Multidimensional Replica-Exchange Method for Free-Energy Calculations. *J. Chem. Phys.* **2000**, *113*, 6042–6051.

(61) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. THE Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(62) Sabei, A.; Hognon, C.; Martin, J.; Frezza, E. Dynamics of Protein–RNA Interfaces Using All-Atom Molecular Dynamics Simulations. *J. Phys. Chem. B* **2024**, *128*, 4865–4886.

(63) Ward Jr., J. H. Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* **1963**, *58*, 236–244.

(64) Yang, W.; Lee, J. Y.; Nowotny, M. Making and Breaking Nucleic Acids: Two-Mg2+-Ion Catalysis and Substrate Specificity. *Mol. Cell* **2006**, *22*, 5–13.

(65) Dal Peraro, M.; Spiegel, K.; Lamoureux, G.; De Vivo, M.; DeGrado, W. F.; Klein, M. L. Modeling the Charge Distribution at Metal Sites in Proteins for Molecular Dynamics Simulations. *J. Struct. Biol.* **2007**, *157*, 444–453.

(66) Cruz-León, S.; Grotz, K. K.; Schwierz, N. Extended Magnesium and Calcium Force Field Parameters for Accurate Ion–Nucleic Acid Interactions in Biomolecular Simulations. *J. Chem. Phys.* **2021**, *154*.

(67) Duboue-Dijon, E.; Javanainen, M.; Delcroix, P.; Jungwirth, P.; Martinez-Seara, H. A Practical Guide to Biologically Relevant Molecular Simulations with Charge Scaling for Electronic Polarization. *J. Chem. Phys.* **2020**, *153*.

(68) Katayanagi, K.; Miyagawa, M.; Matsushima, M.; Ishikawa, M.; Kanaya, S.; Ikehara, M.; Matsuzaki, T.; Morikawa, K. Three-Dimensional Structure of Ribonuclease H from E. Coli. *Nature* **1990**, *347*, 306–309.

(69) Yang, W. An Equivalent Metal Ion in One- and Two-Metal-Ion Catalysis. *Nat. Struct. Mol. Biol.* **2008**, *15*, 1228–1231.

(70) Maffucci, I.; Laage, D.; Sterpone, F.; Stirnemann, G. Thermal adaptation of enzymes: impacts of conformational shifts on catalytic activation energy and optimum temperature. *Chem. Eur. J.* **2020**, *26*, 10045–10056.

(71) Forget, S.; Juillé, M.; Duboué-Dijon, E.; Stirnemann, G. Simulation-Guided Conformational Space Exploration to Assess Reactive Conformations of a Ribozyme. *J. Chem. Theory Comput.* **2024**, *20*, 6263–6277.

## TOC Graphic