The role of conserved residues in the DEDDh motif: the proton transfer mechanism of HIV-1 RNase H

Simon L. Dürr^{1,2}[†], Olga Bohuszewicz¹, Dénes Berta³, Reynier Suardiaz¹[‡], Pablo G. Jambrina¹[¶], Christine Peter², Yihan Shao⁴, Edina Rosta^{1,3}*

¹Department of Chemistry, King's College London, London, SE1 1DB, United Kingdom

²Department of Chemistry, University of Konstanz, 78457 Konstanz, Germany

³Department of Physics and Astronomy, University College London; London WC1E 6BT, United Kingdom

⁴Department of Chemistry and Biochemistry, University of Oklahoma, 101 Stephenson Parkway, Norman, OK 73019–5251, USA

KEYWORDS: RNase H, QM/MM, nuclease, phosphodiester cleavage, HIV-1

ABSTRACT: RNase H is a prototypical example for two metal ion catalysis in enzymes. An RNase H activity cleaving the ribonucleic acid (RNA) backbone of a DNA/RNA hybrid is present in important drug targets, such as the HIV-1 reverse transcriptase but also in many other nucleases such as Homo sapiens (Hs), Escherichia coli (Ec) RNases H or, notably, in enzymes that are part of the CRISPR gene editing molecular machinery. Despite its importance, the reaction mechanism uncovering the proton transfer events are not yet understood. In particular, it is not known, which group is the proton donor for the leaving group. Moreover, several different proton acceptors were proposed, and the exact identity of the proton acceptor is also elusive. Here, we revisit the mechanism for RNAse H, whereby we find that the highly conserved Glu residue of the DDE motif acts as a proton donor via a mechanism further stabilized by the 2'O atom of the sugar. Additionally, we also describe an alternative proton transfer mechanism via a conserved catalytic His residue to deprotonate the attacking water molecule. Furthermore, our quantum mechanics/molecular mechanics (QM/MM) calculations combining Hamiltonian replica exchange with a finite-temperature string method provide an accurate free energy profile for the reaction catalyzed by the HIV-1 RNase H. Our reported pathway is consistent with kinetic data obtained for mutant HIV-1, Hs and Ec RNase H, with the calculated pK_a values of the DEDD residues and with crystallographic studies. The overall reaction barrier of ~19 kcal mol⁻¹, encountered in the phosphate cleavage step, matches the slow experimental rate of $\sim 1-100$ min⁻¹. Additionally, using molecular dynamics (MD) calculations, we sample the recently identified binding site for a third transient divalent metal ion in the vicinity of the scissile phosphate in the product complex. Our results account for the experimental observation of a third metal ion facilitating product release in an *Aquifex aeolicus* RNase III crystal structure and the Bh RNase H in crystallo reaction. Taken together, we provide a molecular mechanism of the nuclease catalytic reaction that is likely common for the broad family of two-metal ion catalytic phosphate cleaving enzymes with a DDE motif.

Introduction

The formation and cleavage of phosphodiester bonds are one of the most essential reactions carried out by enzymes to process the genetic material. The general mechanistic features of phosphodiester cleavage are universally conserved in ribozymes, polymerases, nucleases, and a wide variety of other nucleic acid processing enzymes such as integrases, transposases or topoisomerases.¹

The phosphodiester cleavage generally follows an associative mechanism due to the poor leaving group nature of the alkoxide group,² however, controversially, dissociative mechanisms were also reported occasionally.^{3,4} It utilizes the well-established two-metal ion mechanism.⁵ Two divalent ions are typically bound by one bridging aspartate coordinating both metal ions, another aspartate and a glutamate, referred to as the DDE motif. While a more acidic aspartate residue with a shorter sidechain can clearly provide a moiety for Mg^{2+} binding, the role for the universally present glutamate residue in the DDE motif that is easier to protonate than Asp is as of yet unexplained.

Enzymes of the nuclease family are found in all branches of life. Amongst many processes conferred by the nucleases are DNA replication, proofreading, RNA maturation and RNA interference.¹ One of the particularly well-studied nucleases is RNase H. The RNase H fold is conserved in many enzymes with nucleolytic activity including transposases, the argonaute proteins, and the Cas9 enzyme.^{1,6-8}

RNase H is also an essential part of the HIV-1 enzyme machinery as HIV-1 reverse transcriptase (RT) depends on both a polymerase and a RNase H active site localized on two domains for its full functionality. The RNase H subdomain (residues 1021–1144 of the gag polyprotein, Figure 1B) is located on the p66 subunit of RT where it catalyzes the removal of the RNA strand in the RNA/DNA hybrid generated by the polymerase.⁹ RNase H was found essential for proliferation of HIV-1.¹⁰ In higher eukaryotes RNase H1 is thought to be involved in generating RNA primers for replication and the prevention of genomic instability by removal of RNA nucleotides in R-Loops.^{11,12}

Despite its importance, several key details about the reaction mechanism are still unknown. These include proton transfer events that are often a disputed aspect of enzymatic phosphate cleavage or transfer studies.¹³ The identity of the general base deprotonating the nucleophilic water and the protonation of the leaving group leading to a stable product are still elusive. One possible pathway is the transfer of the proton from the attacking group to the leaving group, mediated by the phosphate undergoing the substitution.^{14,15} However, this scheme may not explain many of the experimental studies.^{16,17,18}



Figure 1: A) Superposition of *Bh* RNase H *in crystallo* reactant after 40ns (PDB 6DMV¹⁹, olive cartoon), HIV-1 QM/MM reactant model (gray cartoon, model based on PDB 1RTD²⁰/1ZBL⁶)) and Hs RNase H (PDB 2QKK²¹, light cyan cartoon). The inset shows positioning of the nucleophile and the active site geometry. Respective residues were colored in magenta (Asp), blue (Glu), cyan (His) and green (Mg²⁺). B) Comparison of the selected RNase H family members and their catalytic sites. Binding of DNA/RNA hybrid is supported by a hybrid binding domain (HBD) or basic protrusion (dark gray). Residues shown in italic, indicate reduced activity when substituted Residues shown in bold, indicate no activity for the $E \rightarrow Q$, $D \rightarrow N$ or $H \rightarrow A$ mutation. *partially active in Mn²⁺, HBD hybrid binding domain, GAG group specific antigen.

Kinetic experiments indicate a crucial role of a catalytic histidine in the HIV-1, human (*Hs*) and *E. coli* (*Ec*) RNase H reaction (H539, H264, H124 in Figure 1).^{17,18} Earlier computational studies of the RNase H catalytic reaction

were either conducted without the histidine included in the QM region,^{14,22} or using the *Bacillus halodurans (Bh*) sequence that lacks this crucial histidine.²³ In similar nucleases with nucleolytic activity, such as CRN-4, TREX1, 3'hExo, or RNase T^{1,24–27}, a conserved histidine residue (part of the conserved DEDDh motif²⁸) has already been suggested from experimental studies to deprotonate the water nucleophile. For the ϵ -subunit of Ec DNA pol III²⁹, QM/MM calculations support this claim.³⁰

More importantly, the identity of the general acid that protonates the leaving O3' is elusive for RNase H, similarly to other proton transfer steps involving the highenergy sugar alkoxide intermediates in other two-metal ion catalytic phosphate cleavage reactions, including polymerases.³¹ Our previous work suggested the protonation of Asp132 in *Bh* RNase H.²³ Water and/or the cleaved phosphate mediated proton transfer from the attacking nucleophile was also suggested.^{14,22} Interestingly, the E116 Glu from the DDE motif of the *Hs* Pol η was observed in two conformations, likely due to its proton acceptor role that would enable it to eventually obtain the proton from the 3'-OH group of the sugar.³²

The key role of the two essential divalent metal ions in the catalytic mechanism of *Bh* RNase H^{19,22,23,33} and *Hs* RNase H¹⁴ was previously elaborated. Interestingly, Samara and Yang¹⁹ recently suggested that a third transient Mg²⁺ ion participates in the catalytic reaction, similarly to polymerases during the reverse reaction.^{19,32,34} Additionally, Samara and Yang¹⁹ also implicate monovalent ions in the *Bh* RNase H catalysis.

To address several of the outstanding questions regarding the HIV-1 RNase H catalytic mechanism: the nature of the general base, the identity of the proton donor for the leaving group, and the role of the third metal ion, we employed classical molecular dynamics (MD) simulations and quantum mechanics/molecular mechanics (QM/MM) calculations utilizing the 6-31+G(d)/B3LYP density functional level of theory. We determined the multidimensional free energy surface underlying the catalvtic reaction of HIV-1 RNase H with two metal ions bound in the active site by combining Hamiltonian replica exchange (RE),³⁵ umbrella sampling (US),³⁶ and the finite-temperature string method.³⁷ In addition, we studied the cleavage product resulting from the action of RNase H using classical MD simulations to investigate the role of a third metal ion. Our simulations showed a reaction pathway that differs from earlier computational studies but are consistent with kinetic and thio-substitution experiments for *Ec* RNase H and *Hs* RNase H.¹⁶ We found that the fully conserved Glu residue in the DEDD motif acts as the proton donor for the protonation of the leaving group. Our suggested mechanism therefore provides a general functional explanation for the role of the Glu residue in two-metal ion catalytic enzymes. Additionally, in our mechanism we find that the conserved His in the DEDDh motif is the proton acceptor base that deprotonates the nucleophilic water molecule. Based on our MD simulations we propose that a third divalent cation, necessary for the displacement of the cleaved RNA strand, is recruited from the bulk potentially with the aid of the His. Using a structural comparison with the catalytic motifs of several nucleases, we suggest that a conserved histidine plays a dual role in many nucleolytic reactions, irrespective of the exact metal ion coordination of the two central Mg²⁺ ions in the active site.

Materials and Methods

Initial coordinates: The crystal structure of the catalytically trapped HIV-1 RT DNA/RNA complex by Huang et al,²⁰ (PDB: 1RTD) with 3.2 Å resolution was used as a starting model. The active site of RNase H was modeled with SWISS-MODEL³⁸ based on the higher resolution structure of D192N *Bh* RNase H catalytic domain by Nowotny et al.⁶ with 2.2 Å resolution (PDB: 1ZBL) to achieve appropriate coordination of the metal ions and proper orientation of the substrate. The obtained model was compared with the recently published catalysisready structure of RNase H in HIV-1 RT (PDB 6BSH).³⁹ Our template, thus our model, is in very good agreement with the 6BSH structure, especially considering the metal coordination (Figure S1).

The DNA/RNA hybrid of the crystal was replaced with the polypurine tract (PPT) sequence.⁴⁰ The RNase H reaction has a pH profile with maximum activity at pH 7.5.¹⁸ All carboxylate active site residues were therefore considered deprotonated at first and H539 was modelled as ε -protonated, while standard protonation states were used for all the remaining titrable residues.

MD simulations: MD simulations for the reactant state were setup using CHARMM-GUI⁴¹ and the model was subject to 20 ns equilibration with the NVT ensemble and the CHARMM27 force field.42 The system was solvated in a cubic water box with 138 Å in each dimension and 10 Å distance to the protein. The phosphate groups were modelled as completely deprotonated. Periodic boundary conditions were applied. The system was neutralized with 150 mM potassium chloride which has been shown to be optimal for catalysis in Hs RNase H.18 The Particle Mesh Ewald method was used to evaluate long range electrostatic interactions and the system was coupled to a Langevin thermostat at a temperature of 303.15 K. Non-bonded interactions were modeled with scaled 1-4 parameters using NAMD 2.943 and a cutoff of 12 Å with switching distance of 10 Å.

Subsequently, QM/MM minimizations were carried out to relax the active site before the reaction coordinate (RC) energy minimization scans to produce a product state. The product state with an added proton on the leaving group and deprotonated scissile phosphate was modelled based on the QM/MM minimized structure. Additionally, either 0 mM, 6 mM or 50 mM MgCl₂ was added. The concentration of 6 to 10 mM MgCl₂ has been shown to be optimal for catalysis in vitro^{18,44} and the concentration of 50 mM is considered inhibitory.⁶ After minimizing for 10 000 steps, the system was re-equilibrated for 5 ns with restrained protein backbone atoms to allow for the relaxation of side chains and to heat up the system in NVT conditions with identical simulation parameters as the reactant state, but using the CHARMM36 force field. Note that longer equilibrations already took place using both MD and QM/MM calculations, as described above. The restraints were then lifted and the trajectories for analysis were collected over 100 ns in NPT conditions. A time-step of 2 fs was used coupled to a Langevin barostat at standard pressure. All systems were run in triplicates.

QM/MM energy minimizations: To carry out the QM/MM calculations, we used the CHARMM 27 force field⁴⁵ and interfaced with Q-Chem 4.4.⁴⁶ The system was trimmed to a sphere of 20 Å centered on the scissile phosphate group. Atoms further than 15 Å away from that group were kept frozen throughout the simulations. We used full electrostatic coupling between the QM and MM regions⁴⁷ and standard link atom treatment for the bonds cut across the MM and QM regions⁴⁸ as described in previous work.^{23,49} The region treated with molecular mechanics was 4164 atoms in size. The QM region contained 116 atoms including two Mg²⁺ ions, the RNA backbone of the active site adenosine, the guanosine in 3' direction of the RNA strand and the side chains of residues Q475, D443, E478, D498, D549 and H539. The hydrogen link atoms were placed between the C_{β} and C_{α} bonds for the aspartates, glutamine and histidine or the C_{γ} and C_{β} bonds for the glutamates. Link hydrogens were restrained between the original bond partners and their bond length was scaled by a factor of 0.7261. Twelve water molecules were included in the QM region of which three coordinate Mg_A. To probe for the different proton transfers proposed in the literature,14,22,50 we also included all water molecules that make hydrogen bonds to the scissile phosphate, to the Bh D132 equivalent D498, to the O2' and the Q475 amide oxygen (Figure 1A). The charge of this QM region was -2. To probe for a possible new protonation pathway, a water molecule in the solvent channel egressing from the 02'/Q475 into the bulk solvent was protonated. The total number of atoms therefore was 117 and total charge of the QM region was -1 for the string calculation.

All calculations were carried out with the 6-31+G(d) basis set. Density functional theory calculations and the B3LYP⁵¹ exchange-correlation functional were used. The RC was defined by the bond distances between reactive atoms. The RC included the proton transfer (Q_p) and electron transfer (Q_e). Q_{ep} is taken as $Q_{ep} = Q_e + Q_p$. The Q_e coordinate includes r2-r1 (WAT_{nuc}:OH2 - RNA:P - RNA:O3'). The Q_p coordinate includes r4-r3 (H539:NE2 - WAT_{nuc}:H - WAT_{nuc}:O, and r6-r5 (RNA:O3' - RNA:H2' - E478:OE1).

String calculations: To determine the reaction free energies and the corresponding minimum free energy pathway of the mechanism, we performed finite temperature string calculations.³⁷ We selected 19 coordinates that define the space for the reaction mechanism (Table

S2), which include all bond breaking and forming distances for the proton transfers and phosphate cleavage. We performed 69 iterations during the string free energy calculations, running 100 steps of QM/MM dynamics at each iteration using Langevin dynamics with a timestep of 1 fs using CHARMM40b2 as the MD engine.⁵² Each coordinate was restrained using a force constant of 150 kcal mol⁻¹ Å².²³ The pathway was defined via 26 windows obtained from the QM/MM minimizations. After each iteration, a new pathway was calculated by fitting a polynomial of order 8 at each coordinate, and dividing the resulting 1 dimensional path within the 19 dimensional coordinate space into 26 equal arclengths.⁵⁰ For validation, further 500 steps of QM/MM dynamics were run in 12 windows equally spaced between the 18th and the 23rd windows to check if a finer resolution is required for capturing the profile during the second proton transfer. The initial coordinates were swapped using a Metropolis Monte Carlo Hamiltonian replica exchange algorithm after each iteration.⁵⁰ The free energy profiles and surfaces of reaction were obtained using the weighted histogram analysis method (WHAM).53,54 The set of strings were divided into four equal sections, and the projected free energies were calculated by WHAM. The standard error of these subset-based free energies represents the uncertainty of the calculation.

Results

General scheme of the reaction: We found, in accordance with previous works, that the reaction proceeds via an associative mechanism. The nucleophilic water (WAT_{nuc}) is bound to Mg_A: polarized to ease deprotonation and oriented for nucleophilic attack on the phosphate. The O3' leaving group is bound to Mg_B (Figure 2A). The reaction proceeds via the formation of a pentavalent phosphorane transition state (TS1, Figure 2B) after, in the same reaction step, WAT_{nuc} is deprotonated. The system then reaches a high energy intermediate state (INT, Figure 2C) followed by a second transition state (TS2, Figure 2D), where the leaving group is protonated to finally arrive at the stable product state (Figure 2E).

Deprotonation of WAT_{nuc} by H539: The order of events for the water deprotonation and the phosphodiester cleavage process was determined with extensive energy minimizations of the QM/MM system followed by finite temperature string free energy calculations. The geometry was optimized along the Qep RC driving the system from reactant to product (Figure 3A-B). $Q_{ep} = Q_e + Q_p$. The O_e coordinate corresponds to the phosphate cleavage, while the Q_p coordinate describes the proton transfer (Error! Reference source not found.A). The geometry minimized path was used in the first iteration to determine the minimum free energy pathway in subsequent string free energy calculations. Initially, the first coupled proton transfer and phosphate cleavage step until the formation of the intermediate was probed during minimizations. The Q_e reaction coordinate alone was used to drive the phosphate cleavage via the 03' leaving

oxygen, the phosphorous atom and the WAT_{nuc} oxygen. We did not specify *a priori* the WAT_{nuc} deprotonation mechanism that facilitates the nucleophilic attack. The QM region included the potential proton acceptors: the 3' downstream *pro*-R_p oxygen, the N_{ϵ} of H539 and the *pro*-R_p oxygen of the scissile phosphate.

We consistently found that the N_ϵ of H539 acted as a proton acceptor. Additionally, a hydrogen bond was formed between the 3' downstream *pro*-R_p oxygen and the other WAT_{nuc} proton.

To determine the effect of the 3' downstream pro-R_p oxvgen, H539 was excluded from the quantum mechanical treatment. In QM/MM minimizations with the smaller QM region, deprotonation was carried out via the 3' downstream pro-R_p oxygen. This is consistent with earlier simulations of *Bh* RNase H⁵⁰ that lacks the histidine residue. To verify that the histidine is the stronger general base the product state of the reaction was modelled by moving the WAT_{nuc} proton pointing towards H539 to the O3' leaving group yielding a product complex with cleaved phosphodiester bond and a monoprotonated phosphate with the hydrogen pointing towards the 3' downstream pro-R_p oxygen. In the subsequent unrestrained QM/MM energy minimization of this system, the WAT_{nuc} proton jumped to the N_{ϵ} of H539. This clearly demonstrates the higher basicity of H539 compared to the 3' downstream pro-R_p oxygen.

Protonation of the leaving O3' via E478: The product of the first reaction step is a cleaved phosphodiester bond with the O3' leaving oxygen tightly bound to Mg_B (Figure 2C). In earlier studies this intermediate has also been observed and found to be highly unstable,⁵⁰ and the protonation of the leaving oxygen was found to be essential for the stabilization of the cleavage product.⁵⁰ However, the source of the proton was not established. As we will show below, we found that protonation via the E478 is an energetically feasible pathway.

Initially, we tested whether the O3' leaving oxygen could be protonated via a file of water molecules in channel C1 passing under the substrate to the bulk solvent. Based on evidence from kinetic experiments with mutated O2',¹⁶ we assumed that the proton could be shuttled to the O3' via the O2' as there is evidence for a role of the O2' in the reaction.¹⁶



Figure 2: Calculated catalytic mechanism of the HIV-1 RNase H reaction. A) Reactant state B) Phosphorane transition state (TS1) after WAT_{nuc} deprotonation C) Intermediate state with the O3' leaving oxygen bound to Mg_B D) Proton transfer from E478 to the O3' at TS2 E) Product of the HIV-1 RNase H cleavage, the remaining proton of WAT_{nuc} can be abstracted by H539 after shuttling the first proton to the bulk solvent via sidechain flipping F) Hypothesized substrate release inferred from Aa-RNAse III crystal structure⁵⁵ (major conformation scissile phosphate in olive sticks, Mg²⁺ in dark grey spheres, PDB 2NUG) and MD simulation of HIV-1 RT/RNase H product complex at 50 mM Mg²⁺ at 30 ns in replicate 1 (light grey spheres, white sticks). The Mg²⁺ ion C bends the scissile phosphate and the *pro*-S_p oxygen is replaced with a solvent water (sphere in magenta). U is upstream nucleotide, S is the nucleotide directly in the active site containing the scissile phosphate, D is the 3' downstream nucleotide. *pro*-S_p and *pro*-R_p oxygen are abbreviated as S_p and R_p. An <u>interactive version</u> with 3D view of the reaction (Panels A-E) together with Figure 4 is available online.

QM/MM simulations with added proton: To test the hypothesis of protonation via solvent channel C1 a hydronium ion was placed in the channel. Energy minimizations showed that the proton is transferred to the terminal water molecule in the solvent channel close to the O2', where it is coordinated by a conserved serine (S499) and a conserved glutamine (Q475) (see Seq. alignment Figure S2). However, with using only energy minimizations, the proton does not directly transfer to the E478, suggesting some energy barrier for this transfer, which we estimated to be below 16 kcal mol⁻¹ using minimizations.

With restrained energy minimizations including the distances between the terminal water proton, the 02', the H2' and the O3' distance and the restraints from the above simulations we tried to obtain a potential energy estimate of this pathway. However, we consistently found that the proton was not directly transferred to the O3' but via E478, which was not included in the restraints. We therefore concluded that E478 could store the proton and then transfer it to the O3'. The proton is received from the water channel egressing from S499 to the bulk solvent. The restraints in the QM/MM minimizations to generate the initial path for the string calculations were modified to account for this pathway by including only the distances between E478:H, O3', P, Onuc, H_{nuc} and $H539:N_c$ in the RC. With this final reaction coordinate a good starting string was constructed conducting iterative QM/MM energy minimizations along the RC. Therefore, in our string calculations, a protonated E478 was modelled as the most likely state of the protein.

To support the possible role of E478 as a proton transfer, we calculated its pK_a using PROPKA 3.1^{56} obtaining a value of 7.80, significantly larger than its model pK_a of 4.5, which suggests that E478 could accept a proton throughout the catalytic cycle at no free energy cost (at pH ~7), similarly to our finding in the QM/MM energy minimizations (Table S1).

Therefore, considering the pK_a estimates, the protonation of E478 occurs with a negligible free energy penalty. Initially, we performed minimizations using the deprotonated form. Subsequently, we showed that there is a viable path for the protons to reach E478, with a free energy barrier less than 16 kcal/mol. This implies that the interconversion between the glutamate and the glutamic acid forms is significantly faster than the reaction itself. Hence whichever protonation state is more favorable for the reaction to proceed, it is readily available under normal conditions. Thus, our main QM/MM free energy calculations were carried out with the protonated E478, and in the MD simulations of the product the proton is already transferred to the leaving group.

Free energy profile of the reaction: The calculated free energy profile and key RCs are depicted in **Error! Reference source not found**.. To assess the convergence of the string calculations, we compared the constraint positions during the second half of the string iterations and found that they only exhibit a minor average variation of 0.0097 Å and a maximum variation of 0.017 Å, implying a good convergence of the final minimum free energy path. To additionally confirm the convergence of the free energies, we split all the QM/MM calculations into four sections and determined the statistical error for the free energy profile (Figure 3C, shaded area), demonstrating small uncertainties.



Figure 3: A) Scheme of the catalytic site of the HIV-1 RNase H in the reactant state. Residues in brackets are *Bh* equivalent residues. B) Distance plot of the distances specified in A. Average of distance in the last 1ps in each window is plotted. The shaded rectangles represent RS reactant state, TS1 transition state 1, INT intermediate, TS2 Transition state 2 and PS product state. C) Free energy reaction profile from string calculations projected onto Q_{ep} RC with fitted curve. Red shades depict the estimated variation of the profile along the energy axis.

We obtained two local maxima in the one-dimensional energy profile projected onto the Q_{ep} RC, corresponding to the two transition states (Figure 3C). The formation of the phosphorane TS (TS1) was the highest point in the free energy profile with ~19 kcal mol⁻¹. The intermediate state was ~4 kcal mol⁻¹ lower in energy compared to TS1. To form the product via TS2 another energy barrier with ~3 kcal mol⁻¹ had to be surmounted to protonate the leaving group. From Figure 3B and the plot of the free energy surface on the Q_e and Q_p coordinates (Figure 4) it can be seen that first a proton transfer, then an electron transfer and then again a proton transfer occurs.



Figure 4: Free energy profile of the reaction from the string simulations with RS reactant state, transition state 1 (TS1), intermediate (INT), transition state 2 (TS2) and product state (PS). Starting string in green, optimized final string in black. Colorbar indicates ΔG in kcal mol⁻¹.

Role of third ion for product release: A third divalent ion was also implicated in the mechanism by Samara and Yang.¹⁹ However, as they deduce importance of the third divalent ion from the in crystallo reaction further investigation into the dynamics of the ion recruitment for the real reaction in solution and in the presence of the histidine is still needed. Experimentally, it has been observed that at high Mg²⁺ concentrations (>50mM in Bh RNase H) the reaction stalls^{18,19,57-59}, which has led to the hypothesis that a third metal ion could bind to the product or disturb the nucleophilic attack even before it was detected in the in crystallo reaction. In their recent work Samara and Yang¹⁹ have shown that the third divalent metal ion coordinates to the scissile phosphate and to the 3' downstream phosphate and that product release can only take place after its departure. Interestingly, $H \rightarrow A$ mutation in Hs or Ec RNase H removes the attenuation of the reaction at high Mg²⁺ concentrations, suggesting that the catalytic histidine plays an important role in the recruitment and displacement of the third metal ion.^{18,57} We therefore investigated the product complex of the cleavage reaction in HIV-1 RT/RNase H at the catalytic (6mM), zero and high (50mM) MgCl₂ concentrations with classical MD simulations. In multiple 100 ns long MD simulations we found two prominent binding sites for cations in the product complex. The term "binding site" used in this context refers to the ion sampling a small region of space for some extended period of time close to the active site

and shall not be indicative of first shell binding via displacement of one of the first shell water ligands of the solvated K⁺ or Mg²⁺ ions. While it would be important to access longer timescales where first shell metal ion coordination changes can be observed, this is beyond the limits of current simulation times.

One of the sites (F in Figure 5, Figure S3) was located close to D498, Mg_A and Mg_B and was populated mostly by potassium ions. This binding site was also occupied during the reaction for Bh RNase H crystal structure and was probably populated because of the double negative charge of the scissile phosphate.



Figure 5: Alignment of MD (gray) to 600 s product structure of *Bh* RNase H (yellow, PDB 6DOG¹⁹). Water channel C1 with water molecules from MD (as sticks) and from crystal (red spheres) structures reaches E478. Mg²⁺ ion binding sites are also labelled from MD (magenta sphere, location C), from crystal structures (yellow spheres, locations A, B, C, F), and in QM/MM and MD calculations (gray spheres, locations A and B). S499 keeps the terminal water molecule in place with a hydrogen bond (0-0 distance in 6DOG, 2.7 Å)

The second binding site (ion C in Figure 2F, Figure 5, Figure S3) was located between the scissile phosphate and the 3' downstream phosphate, and was populated by Mg²⁺ and K⁺ ions alike. While no direct first-shell coordination was observed with the protein residues in our MD simulation, in a RNase III crystal structure the Mg²⁺ ion in site C directly coordinates the scissile phosphate (Figure 2F).55 Site C was also found to be occupied in the product state of the *in crystallo* reaction.¹⁹ The length of the binding events for the ions varied within the replicates of the MD simulations from very short binding $(\sim 5ns)$ to guite stable association with the scissile phosphate with retention times of up to 80 ns in binding site C. The relation of potassium vs. magnesium binding is biased towards more potassium ions as they were more abundant in the simulation (150 mM K⁺ vs 50 mM Mg²⁺). As is expected from the classical description of the ions, we do not observe loss of the first solvation shell during the relatively short simulation time. Instead, we observe coordination to the phosphate oxygens mediated via second shell interactions.

The simulations show that the binding of a metal ion in site C causes conformational changes in the RNA strand of the hybrid that disrupt the interaction of the RNA with the DNA strand (Figure S3). Complete dissociation could not be observed likely due to the length of the simulation, the length of the hybrid and/or due to the limitations of the nucleic acid force field. Nevertheless, a general trend for destabilization is visible (measured in terms of RMSD of the active site nucleotide and the two nucleotides in the 3' direction of the RNA strand, Figure S3).

As suggested for the $H \rightarrow A$ mutation in Hs RNase H,¹⁸ the catalytic histidine H539 placed on a flexible loop next to the active site seemed to participate in the binding of the third ion. We measured the position of the histidine using its N α -C α -C β -C γ dihedral angle, which varied from - 45° to 60° for the inwards facing position ready for the catalysis and from -170° to -180° for the outwards position ready for metal ion recruitment and shuttling of the abstracted proton to the solvent (Figure S4). When a third metal ion was in the binding site Mgc, we found the H539 to be almost exclusively facing inwards as a second shell ligand of the third ion (Figure S5). The metal ion dissociated from the binding site shortly after the histidine changed its conformation from in- to outwards in several of the replicas.

Discussion

Dual functional role of the histidine: Alla and Nicholson¹⁸ have proposed a dual functional role of H264 in Hs RNase H acting both as a proton acceptor and a facilitator of the product release congruent with the findings of our QM/MM and classical simulations. The mutation of the catalytic histidine in Hs RNase H (H264A) reduces the kcat 100-fold18 but has no effect on the Km. A similar effect of the H \rightarrow A mutation in RNase H can be seen in *Ec*, where the H124A substitution decreases the kcat 50-fold and increases K_m only by 4-fold.¹⁷ The H539F mutation in HIV-1 RNase H renders the enzyme inactive.⁶⁰ In addition, Tisdale et al.¹⁰ showed that even the conservative H539N mutation makes the virus non-infective, corroborating the crucial role of H539 for RNase H in the virus life cycle. It is likely that the replication of the viral genome is severely impacted by a slower RNase H catalytic rate.

The pH profile of the reaction^{18,61} with maximum activity at pH 7.5 and low activity at low pH is also consistent with the role of the histidine which is protonated at low pH, diminishing the catalytic activity.

Our QM/MM calculations explain all the aforementioned experimental evidence, evincing that H539 is likely to act as the proton acceptor in contrast to previous calculations for which the conserved histidine was not included in the QM region.¹⁴ This also underlines the importance of careful selection of the QM region.⁶²

As previously argued,^{22,50} the effect of the 3' downstream *pro*-R_p oxygen (with sulfur substitution 86% k_{cat} decrease⁶³) is most likely an indirect effect, where the 3' downstream *pro*-R_p oxygen increases the catalytic rate by correctly positioning the nucleophile. Another cause of the rate diminishing effect of the 3' downstream *pro*-R_p thio substitution is likely the hindered ability to stabilize a third metal ion.

As H539 was very flexible in the MD simulations and was shown to be able to also deprotonate the monoprotonated phosphate in the product complex in QM/MM minimizations, it is likely that H539 shuttles both protons from WAT_{nuc} to the solvent and acts as the general mediator for proton transfers to and from the solvent. It could also play a role in attracting the active site Mg^{2+} ions to the active site, similarly to the role of E188 in *Bh.*¹⁹

An alternate proton acceptor could be the *pro*-R_p oxygen of the scissile phosphate, however mutational studies showed that thio substitution of the *pro*-R_p oxygen does not decrease the catalytic rate¹⁶ which is why it was not considered as a possible proton acceptor,

The flexibility of H539 and the coupled occurrence of a metal ion in site C and the histidine pointing inwards (Figure S4) further support the role of the histidine in recruitment of a third ion similarly to that suggested by Alla and Nicholson¹⁸ experimentally. This effect is most likely electrostatic, as first shell coordination of a Mg^{2+} or K⁺ by the histidine is not observed in our simulations. This is congruent with findings of low affinity of nitrogens for earth alkali metals ⁶⁴ but might also be related to the timescale for exchange of water ligands on Mg^{2+} and the known deficiencies of the Mg^{2+} force field.

In the Bh *in crystallo* reaction, the residue K196 placed at the C-terminus on a flexible loop stabilizes the scissile phosphate. In HIV-1 RNase H a lysine (K550) is present as well, which, as the C-terminus of the HIV-1 protein extends further, is placed in an alpha helix of HIV-1 RNase H instead of the C-terminal loop of Bh RNase H. K550 in HIV-1 RNase H points away from the active site and is too far away to stabilize the cleaved phosphate (Figure S6).

A possible hypothesis for the function of the third divalent ion in the catalysis and product release can be inferred from Aa-RNase III⁵⁵ and from the *in crystallo* study of the Bh RNase H mechanism.¹⁹ Aa-RNase III performs a similar 3'-5' endonucleolytic cleavage like RNase H with the help of two metal ions but the architecture of its active site is slightly different (DDEE motif). In the crystal of the product complex of Aa-RNase III (res. 1.7 Å, PDB: 2NUG), two conformations of the scissile phosphate were observed with different occupancies. The minor conformation with occupancy 0.3 closely resembles the product state of the HIV-1 RNase H reaction from the OM/MM with the scissile phosphate pro-Sp oxygen coordinating both Mg_A and Mg_B. However, the major conformation in the crystal with an occupancy of 0.7 has the scissile phosphate out of the active site, bent away from Mg_A and Mg_B and closer to Mg_C (Figure 2E). In the major conformation, the scissile phosphate is a first shell ligand of Mg_C and the scissile phosphate O_{sp} between Mg_A and Mg_B is replaced by a water molecule. Gan et al.⁵⁵ have proposed that the minor conformation is a model of the complex directly after the reaction took place and the major conformation is the first step towards product release in Aa-RNase III. Thus, the product release is facilitated by Mgc.

Interestingly, the main binding site for the third ion (site C) and the scissile phosphate in the HIV-1 RNase H MD simulations are in the same position as in the minor conformation of the Aa-RNase III crystal, suggesting similarity between the Aa-RNase III and RNase H product release mechanism. The stabilization through the third ion likely has stabilized this state in Aa-RNase III enough in order to be detected by conventional crystallography.

Attempts to crystallize the *Bh* RNase H product complex with 100% occupation of a third Mg²⁺ have not been successful⁵⁹ and in the *in crystallo* reaction the third ion was only detectable in 500mM Mn²⁺ or in 5mM Mg²⁺/ 200mM Li⁺ buffer. It was found that the concentration of Mg²⁺ needed for optimal catalysis is higher than the concentration to fully occupy the A and B sites.¹⁹ In analogy to our observations, Samara and Yang found the third ion to bind to the product state.¹⁹ In the crystal it can be seen that with Mgc the 5' phosphate product is shifted by 1 Å.¹⁹

The suggested role of the histidine in the recruitment of the third ion is consistent with the observation that for $H \rightarrow A$ mutation the attenuation of RNase H activity at high Mg²⁺ concentration is lost.^{18,57} *Bh* E188A mutant, which has a lower activity, also displays no attenuation at high Mg²⁺ concentrations.⁶ Nowotny and Yang⁵⁹ have attributed the lower activity of *Bh* E188A to an abnormally high affinity to the cleavage product which supports the interpretation that a third metal ion is important for the product release. In *Hs* RNase H264A mutation increases the rate constant for the product release 12-fold which further supports the role of histidine in recruitment of the third ion.¹⁸

Role of monovalent ions: Samara and Yang¹⁹ implicate monovalent potassium ions in the catalysis based on the in crystallo analysis of the reaction mechanism in Bh RNase H. In Bh RNase H, K+w (site F in Figure 5) coordinates the scissile phosphate and K⁺U (site C in Figure 5) together with a transiently bound Mg²⁺c activates the nucleophile. Bh RNase H likely relies on the cation trafficking to activate the nucleophile because it lacks the histidine residue, which is present in other members of the RNase H family. Therefore, HIV-1, Hs and Ec RNase might be less reliant on monovalent ions as the histidine activates the nucleophile and therefore possibly replaces the function of the monovalent ions. This is also related to the slightly different active site architecture with Bh K197 able to point inwards stabilizing the intermediate, which is not possible for HIV-1 K550 (Figure S6).

Activation free energy of the reaction: The calculated barrier height of ~19 kcal mol⁻¹ is in good agreement with experimentally determined barriers calculated from kinetic experiments to 16.7 to 20.5 kcal mol⁻¹ in $Ec^{63,65}$ and Hs RNase H.^{18,66} The measured k_{cat} with modified substrate in HIV-1 RNase H is 1.01 to 0.11 min⁻¹ corresponding to a barrier of 19.8 to 21.2 kcal mol^{-1.67} The modified substrate is, however, likely to have negatively impacted the k_{cat} for the measurement of HIV-1 RNase H.²² The k_{cat} values measured with the human and

the *Ec* RNase H are therefore considered to be closer to the real value in HIV-1.

Glutamate as conserved proton donor in nucleases: Various experimental evidence supports the role of E478 in the proton transfer. Firstly, the homologous Ec E48D mutant is substantially less active with a 95% decrease in k_{cat}.¹⁷ This decrease is most likely connected with suboptimal positioning and change in pKa. Secondly, it was found that *Ec* E48 is less important for Mg²⁺ binding compared to the three other aspartates as the Mg²⁺ dissociation constant for the Ec E48Q mutant is 1.1 mM compared with 0.71 mM for the wild type but 18.8 mM for Ec D70N.68 As a result of the latter, Kanaya et al.68 proposed that the glutamate of the DEDD motif should play an additional role in the catalysis. Our results support that hypothesis, showing the role of E478 as a mediator in the proton-transfer reaction needed for the stabilization of the leaving group and competition of the cleavage reaction.

Uchiyama et al.¹⁶ have measured kinetics of *Ec* RNase H with modified substrates at the O2' position of the nucleotide immediately upstream of the scissile bond, which also supports the pathway via E478. Substitution of the O2' into an amino, fluoro or methoxy group or removal of the O2' did not affect the affinity for the substrate but greatly reduced the reaction rate.¹⁶ E478 interacts with the O2' through a hydrogen bond both in the crystal and in simulations. Altogether this supports a dual role of E478 in recognizing the RNA strand in the DNA/RNA hybrid via the hydrogen bond and then by shuttling the proton to the leaving O3'.

Also consistent with this hypothesis is the 50% decrease in catalytic activity in *Ec* by mutation of a serine close to the active site (*Ec* S71A, HIV-1 S499).¹⁷ This serine is conserved in *Bh*, *Hs*, *Ec* and HIV-1 RNase H (Figure S2). In our simulations, we found that the terminal water molecule in channel C1 is bound to S499 (Figure 5). The binding site for this terminal water is favorable as can be inferred from its presence in multiple crystal structures (e.g. 1ZBL, 2G8I, and 2R7Y, 6DOG) of *Bh* RNase H (Figure 5).⁵⁰ It is therefore likely that the water channel and the serine form part of a conserved replenishing mechanism to reconstitute the protonated state of the glutamate in the DEDD motif in RNases H.

Protonation of D498, which was previously considered as the most likely protonation pathway for *Bh* RNase H,⁵⁰ was not attempted as the pathway via E478 is more likely. Our pK_a calculations show that E478 (pK_a 7.80) is more susceptible to get protonated than D498 (pK_a 2.71), with the latter incurring an additional free energy cost at pH 7.0. Thus, E478 can better stabilize the proton ultimately donated to the O3' alkoxide after the intermediate has formed, because of its higher pK_a value. Another contributing factor is that the energetic cost of rotating the D498 necessary for protonation is not necessary if protonation occurs via the pre-positioned E478. The protonation pathway via the *pro*- R_p oxygen of the scissile phosphate is not consistent with available experimental findings therefore it was not considered in our calculations. Uchiyama et al.¹⁶ demonstrated that sulfur substitution of the *pro*- R_p oxygen of the scissile phosphate does not decrease activity. A decrease would however be expected for sulfur substitution of the *pro*- R_p oxygen because the phosphorothioate is a weaker base. It can therefore be concluded that protonation via E478 is the most likely route for the protonation of the leaving group in HIV-1 RNase H and possibly in the other members of the family (*Bh* E109, *Ec* E48, *Hs* E186) in contrast to earlier findings.

Role of catalytic histidine and glutamate in other nucleases: The RNase H topology is shared between many catalytic cores of enzymes that process nucleic acids. We therefore strategically mined the protein data bank for similar catalytic motifs to obtain a picture of the generality of the catalytic mechanism using the histidine and a third ion in cleaving nucleic acids. We used VMD to automatically select structures based on certain criteria.69 As criteria we used the presence of a histidine within 5 Å of two Mg²⁺, Ca²⁺ or Mn²⁺ ions in all structures with resolution better than 2.5 Å and manually inspected the resulting structures. From our analysis, we could determine that indeed catalytic sites of various nucleases are similarly composed of glutamates, aspartates and histidines as has been described before,^{1,70} but we were able to further reason the positioning of some of the residues based on the results presented in this article and the recently available in crystallo study of the reaction¹⁹ as well as the QM/MM dynamics of the Cas9 active site.⁷¹ All of the identified enzymes belong to the RNHL superfamilv.⁷⁰

While the DEDD motif is common among nucleases,^{1,72} we identified three different subclasses of this catalytic motif with histidine based on the PDB mining. Two with Glu bound to metal B (RNase H-like, RuvC-like, Figure 6AC) and the other with Glu bound to metal A (RNase T-like, Figure 6B). Both the RNase Hlike motif as well as the RNase T-like motif have been commonly referred to as DEDD motifs. The RuvC-like motif is referred to as DDE motif.71 For the RNase T-like structures, Zuo and Deutscher²⁸ established the DEDDh nomenclature to differentiate DEDD nucleases that have histidine from those with a tyrosine (DEDDy) in their active site. It is well known that the catalytic mechanism of these nucleases is strongly related to the coordination of the catalytic magnesium ions,^{21,50} therefore the RNase H-like and RNase T-like motifs are not the same DEDDh motifs even though the position of the histidine is very similar (Figure 6). In both types, the histidine is placed on a loop segment coordinating the nucleophile (Figure S4).27

Differences in the catalytic mechanism are mainly associated with Mg_B , which is penta-coordinated in the RNase H-type and hexa-coordinated in the RNase T-like, most likely due to different specificities for DNA and RNA substrates. In contrast, in the RNase T-like enzymes Mg_A is penta-coordinated.³⁰ For the first two types, the pK_a of the glutamate is high (pK_a 7.5-9 as calculated with PROPKA 3.1, see Table S1) due to the relatively nonpolar micro environment around the pocket with the protonated oxygen. The high pK_a for the active site Glu supports the role of the glutamate as proton donor (RNase H-like, RuvC like) or in stabilizing the WAT_{nuc} (RNase Tlike) as it is significantly more basic than a standard Glu ($pK_a \sim 4.5$). In QM/MM simulations of the mechanism of the ϵ -subunit of DNA pol III it was shown that the Glu in the RNase T-like motif can pick up the second proton from the WAT_{nuc} after cleavage of the scissile bond.³⁰

We found that in the RNase H-like enzymes protonation likely occurs via the Glu coordinated to Mg_B. The exception within the canonical DEDDh motif are the RNase Tlike enzymes due to the different position of the Glu (Figure 6). In QM/MM simulation of the RNase T-like ϵ -subunit of the pol III the protonation of the leaving O3' remained elusive, but pathways via the scissile phosphate were ruled out based on the calculated energy barriers of >30 kcal mol⁻¹.³⁰ In the RNase T-like enzymes, there is an Asp exclusively coordinated to Mg_B via a water molecule but its pK_a likely is much closer to a standard Asp $(pK_a \sim 3, see Table S1)$, making protonation via this residue less likely, but possible. The proton for the protonation of the leaving group thus possibly originates from the Asp or any of the up to three water molecules of the first solvation shell of Mg_B.⁷⁴ For example in the RNase T structure solved by Hsiao et. al.⁷⁴ Mg_B is solvated with 3 water ligands as opposed to zero water ligands in the RNase H-like and in the RuvC-like structures.⁷¹ This mechanism is also consistent with the lowering of pKa of Mg associated water molecules.75

In the RuvC-like enzymes the DEDD motif is not observed (only DDE) and the last Asp of the DEDD motif is possibly replaced by a His (e.g in S. pyogenes (Spy) Cas9 RuvC H983). This is supported by the fact that $H \rightarrow D$ mutation in *Tth* RuvC is shown to activate the enzyme,⁸ which would reconstitute the DEDD motif of RNase H. This indicates the strong similarity of the RuvC-type and the RNase H-type. In contrast to this finding, other studies report that the affinity of nitrogen (e.g., in histidine or imino side chains of nucleic acids) for Mg²⁺ is low⁶⁴ and in the QM/MM ab initio MD of the Cas9 active site it can be seen, that the stable distance of H983 to MgA is around 4.2 Å, which is not directly coordinating. We can reproduce the same distance if we align our QM/MM reactant structure to the inactive Spy Cas9 (PDB 4UN3) RuvC active site (Figure S7). H983 has recently been shown to be the general base in the RuvC reaction in Spy Cas9.71,76



Figure 6: Comparison of the three main types of catalytic motifs identified in the PDB search and from literature. A) In RNase Hlike, Glu coordinates a solvent excluded Mg_B and stores proton for leaving group protonation. His abstracts the proton and facilitates product release. Example PDBs: 1RTD, 2QKK, 6BSH. B) In RNase T-like enzymes, the Glu is shifted to Mg_A aiding in WAT_{nuc} positioning. Mg_B is coordinated by three water molecules⁷³ of which one is indirectly coordinated by an Asp. Example PDBs: 4KAZ, 1W0H, 1J53, 3NH1, 2O4G, 3CG7. C) In RuvClike enzymes, an additional Lys/Arg is present which we consider to be atypical for a general RNase H but present in *Bh* RNase H (K196). Mg_A is coordinated by a backbone (bb) oxygen.⁷¹ Example PDBs: 4LD0, 4EP4, 4UN3, 1HJR, 3HVR. See more details in Table S1.

For RuvC-type enzymes, neither H982 nor H983 is placed on a flexible loop as is the histidine in the RNase H-type or the RNase T-type.²⁷ These histidines therefore would likely be able to abstract the proton and act as general base but not as third ion recruiter or facilitator of product release. However, there is a conserved lysine (*Spy* K974, *Tth* K112) and a conserved arginine (*Spy* R976) in close vicinity of the active site. This lysine (2-4 residues in direction of the N-terminus from the arginine) and the arginine are conserved among various Cas9 family members such as from *Staphylococcus Aureus, Francisella Novicida, Actinomyces naeslundii* and *Campylobacter jejuni* (Figure S8). The lysine is very mobile (no sidechain density for *Tth* RuvC K112 PDB 4LD0) and most likely carries out a similar task such as K196 in *Bh*.¹⁹ In Gaussian Accelerated MD simulations of the activated Cas9 complex it has also been shown that R976 is more stable in a down conformation pointing in the active site.^{71,76} K196 in *Bh* stabilizes the negatively charged cleaved product until arrival of the third Mg²⁺ which then possibly facilitates the product release.¹⁹ We postulate that one positively charged residue (K or R) is needed to facilitate catalysis and removal of the product in the RuvC-like type.

Combinations of these general motifs are also possible such as in *Bh* RNase H. Here, the leaving group protonation most likely is RNase H-like, but the histidine, which is missing, is partly replaced by the substrate and transient metal ions^{19,50} and partly by the lysine as in the RuvClike motif.

Conclusion

HIV-1 RNase H catalyzes the cleavage of the DNA/RNA hybrid bound to the HIV-1 RT enzyme dependent on a non-canonical DEDDh motif for coordination of two catalytic Mg²⁺ ions. Our QM/MM calculations, in agreement with prior work, found that the reaction proceeds via two transition states that separate the two protonation events: (i) deprotonation of the nucleophilic water and (ii) protonation of the leaving group. Beyond this, here we identified a new proton transfer mechanism, both for the deprotonation of the nucleophilic water and for the protonation of the leaving group, that is able to account for the multitude of experimental results not only in RNase H but in many other nucleases. We identified a conserved histidine residue, which has a dual role as activator of the nucleophile and recruiter of a third divalent metal ion in the product complex. The histidine is structurally and functionally conserved across several other enzymes with a RNase or DNase function, and homologous mutational and computational studies in nucleases with a different positioning of the same active site residues (i.e the canonical DEDDh motif) support our mechanistic results. We highlight differences and similarities of the canonical DEDDh motif in 3'-5' exonucleases such as the ϵ subunit of polymerase III or RNase T with the non-canonical motif found in various RNases H in the discussion of our results.

Here, we revisited the protonation of the leaving group by assessing the possible path of the proton donor. We identified a conserved water channel and demonstrated that by carefully selecting the QM region,⁶² we are now able to observe that the proton preferably reaches the vicinity of the sugar leaving group.

We also identified a highly conserved glutamate residue as the most likely proton donor for the leaving group phosphate O3', providing evidence why a glutamate and not a fourth aspartate is present in the DEDD motif in RNase H. Our pK_a calculations are consistent with this mechanism showing that the pK_a of the Glu residue is strongly upshifted in several available crystallographic structures. We therefore suggest that this residue acts as the proton donor in RNase H and RuvC-like nuclease catalytic reactions, and likely participates as proton acceptor or donor for other two-metal ion catalytic enzymes such as polymerases.

Furthermore, to shed light on the post-reaction product release process, we performed MD simulations to identify the role of active-site metal ions and the conserved histidine residue. We sampled the possible metal ion binding sites of the cleaved product state for transiently bound mono- and divalent cations in multiple MD simulations. Our simulations show that a third metal cation is recruited from the bulk finding two potential cation binding sites that are consistent with those identified experimentally by Samara and Yang¹⁹ and Gan et. al.⁵⁵ Our results suggest that the third Mg_c binding site plays a key role in the product release based on the cleaved product stability in our MD simulations, aided also by the conserved histidine. Our results confirm that this mechanism is likely more general for other nucleases, based on structural similarity with Bh RNase H, HIV-1 RNase H and Aa-RNase II crystallographic data.

Through the novel understanding of structural and mechanistic roles of key conserved residues, we provide a number of key insights to help understanding the nucleic acid processing reaction mechanism in a wide array of biomedically important enzymes possessing a DDE motif, aiding structural and functional understanding into these ubiquitous biochemical processes.

ASSOCIATED CONTENT

Supporting Information including

- Figures S1-9, Alignments, geometries and data from the MD.
- Tables S1-2, PROPKA 3.1 calculation and string collective variables.
- Source code of interactive Figure of the free energy profile (<u>https://dev.simonduerr.eu/interactive/hiv/</u>) and the geometries from the QM/MM and from <u>https://doi.org/10.5281/ze-nodo.4775077</u>

AUTHOR INFORMATION

Corresponding Author

*e.rosta@ucl.ac.uk

Present Addresses

† Laboratory of computational biochemistry and chemistry, SB ISIC, Ecole polytechnique fédérale de Lausanne, CH-1015 Lausanne, Switzerland

‡ Universidad complutense de Madrid, Facultad de Ciencias Químicas, Departamento de Química Física I, Avd. Complutense s/n 28040, Madrid, Spain. ¶ Departamento de Química Física, Facultad de Ciencias Químicas, Universidad de Salamanca, 37008 Salamanca, Spain

Funding Sources

E.R acknowledges funding from EPSRC (EP/R013012/1), BBSRC (BB/N007700/1), and ERC (project 757850 BioNet). S.D received scholarships from Erasmus+ and Studienstiftung des deutschen Volkes. P. G. J Acknowledges funding by Fundacion Salamanca City of Culture and Knowledge ("Programme for attracting scientific talent to Salamanca"). E.R and R.S acknowledge funding from the EC Marie Curie IEF.

Notes

ORCID S.L.D 0000-0002-4304-8106 D.B 0000-0002-8299-3784 R.S 0000-0002-1035-9020 P.G.J 0000-0001-8846-3998 E.R 0000-0002-9823-4766

ACKNOWLEDGMENT

Computing time was provided on Rosalind HPC cluster by King's College London and the ARCHER UK National Supercomputing Service (http://www.archer.ac.uk) via HecBioSim (funded by EPSRC grant EP/R029407/1).

ABBREVIATIONS

Hs Homo sapiens, *Ec* Escherichia coli, *Bh* Bacillus halodurans, *Tth* Thermus thermophilus, *Spy* Streptococcus pyogenes, TS transition state, RS reactant state, INT intermediate, PS product state, HBD hybrid binding domain, GAG group specific antigen.

REFERENCES

- Yang, W. Nucleases: Diversity of Structure, Function and Mechanism. *Q. Rev. Biophys.* 2011, 44 (1), 1–93. https://doi.org/10.1017/S0033583510000181.
- (2) Klähn, M.; Rosta, E.; Warshel, A. On the Mechanism of Hydrolysis of Phosphate Monoesters Dianions in Solutions and Proteins. J. Am. Chem. Soc. 2006, 128 (47), 15310–15323. https://doi.org/10.1021/ja065470t.
- Casalino, L.; Palermo, G.; Rothlisberger, U.; Magistrato, A. Who Activates the Nucleophile in Ribozyme Catalysis? An Answer from the Splicing Mechanism of Group II Introns. J. Am. Chem. Soc. 2016, 138 (33), 10374–10377. https://doi.org/10.1021/jacs.6b01363.
- (4) Huang, W.; Huang, Y.; Xu, J.; Liao, J. Lou. How Does the Spliceosome Catalyze Intron Lariat Formation? Insights from Quantum Mechanics/Molecular Mechanics Free-Energy Simulations. J. Phys. Chem. B 2019, 123 (28), 6049–6055. https://doi.org/10.1021/acs.jpcb.9b04377.
- (5) Steitz, T. A.; Steitz, J. A. A General Two-Metal-Ion Mechanism for Catalytic RNA. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 6498– 6502. https://doi.org/10.1073/pnas.90.14.6498.
- 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 (7), 1005–1016. https://doi.org/10.1016/j.cell.2005.04.024.
- Jiang, F.; Doudna, J. A. CRISPR Cas9 Structures and Mechanisms. Annu.Rev.Biophys 2017, 46, 505–529. https://doi.org/DOI: 10.1146/annurev-biophys-062215-010822.
- (8) Górecka, K. M.; Komorowska, W.; Nowotny, M. Crystal Structure of RuvC Resolvase in Complex with Holliday Junction Substrate. *Nucleic Acids Res.* 2013, 41 (21), 9945– 9955. https://doi.org/10.1093/nar/gkt769.
- (9) Katz, R.; Skalka, A. M. The Retroviral Enzymes. Annu. Rev. Biochem. 1994, 63 (1), 133-173.

https://doi.org/10.1146/annurev.biochem.63.1.133.

- (10) Tisdale, M.; Schulze, T.; Larder, B. A.; Moelling, K. Mutations within the RNase H Domain of Human Immunodeficiency Virus Type 1 Reverse Transcriptase Abolish Virus Infectivity. *J Gen Virol* **1991**, *72* (Pt 1), 59–66.
- Chon, H.; Sparks, J. L.; Rychlik, M.; Nowotny, M.; Burgers, P. M.; Crouch, R. J.; Cerritelli, S. M. RNase H2 Roles in Genome Integrity Revealed by Unlinking Its Activities. *Nucleic Acids Res.* 2013, 41 (5), 3130–3143. https://doi.org/10.1093/nar/gkt027.
- (12) Cerritelli, S. M.; Crouch, R. J. Ribonuclease H: The Enzymes in Eukaryotes. *FEBS J.* **2009**, *276* (6), 1494–1505. https://doi.org/10.1111/j.1742-4658.2009.06908.x.
- (13) Berta, D.; Buigues, P. J.; Badaoui, M.; Rosta, E. Cations in Motion: QM/MM Studies of the Dynamic and Electrostatic Roles of H+ and Mg2+ Ions in Enzyme Reactions. *Curr. Opin. Struct. Biol.* **2020**, *61*, 198–206. https://doi.org/10.1016/j.sbi.2020.01.002.
- (14) Elsässer, B.; Fels, G. Atomistic Details of the Associative Phosphodiester Cleavage in Human Ribonuclease H. *Phys. Chem. Chem. Phys.* **2010**, *12* (36), 11081. https://doi.org/10.1039/c001097a.
- (15) Borišek, J.; Magistrato, A. All-Atom Simulations Decrypt the Molecular Terms of RNA Catalysis in the Exon-Ligation Step of the Spliceosome. ACS Catal. 2020, 10 (9), 5328–5334. https://doi.org/10.1021/acscatal.0c00390.
- Uchiyama, Y.; Miura, Y.; Inoue, H.; Ohtsuka, E.; Ueno, Y.; Ikehara, M.; Iwai, S. Studies of the Interactions between Escherichia Coli Ribonuclease HI and Its Substrate. *J. Mol. Biol.* **1994**, 243 (4), 782–791. https://doi.org/10.1016/0022-2836(94)90047-7.
- (17) Kanaya, S.; Kohara, A.; Miura, Y.; Sekiguchi, A.; Iwai, S.; Inoue, H.; Ohtsuka, E.; Ikehara, M. Identification of the Amino Acid Residues Involved in an Active Site of Escherichia Coli Ribonuclease H by Site-Directed Mutagenesis. *J. Biol. Chem.* **1990**, *265* (8), 4615–4621.
- (18) Alla, N. R.; Nicholson, A. W. Evidence for a Dual Functional Role of a Conserved Histidine in RNA·DNA Heteroduplex Cleavage by Human RNase H1. FEBS J. 2012, 279 (24), 4492– 4500. https://doi.org/10.1111/febs.12035.
- (19) Samara, N. L.; Yang, W. Cation Trafficking Propels RNA Hydrolysis. *Nat. Struct. Mol. Biol.* **2018**, *25* (8), 715–721. https://doi.org/10.1038/s41594-018-0099-4.
- Huang, H.; Chopra, R.; Verdine, G. L.; Harrison, S. C. Structure of a Covalently Trapped Catalytic Complex of HIV-1 Reverse Transcriptase: Implications for Drug Resistance. *Science* 1998, 282 (5394), 1669–1675. https://doi.org/10.1126/science.282.5394.1669.
- (21) Nowotny, M.; Gaidamakov, S. A.; Ghirlando, R.; Cerritelli, S. M.; Crouch, R. J.; Yang, W. Structure of Human RNase H1 Complexed with an RNA/DNA Hybrid: Insight into HIV Reverse Transcription. *Mol. Cell* 2007, 28 (2), 264–276. https://doi.org/10.1016/j.molcel.2007.08.015.
- (22) 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 (33), 10955–10962. https://doi.org/10.1021/ja8005786.
- (23) 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 (23), 8934–8941. https://doi.org/10.1021/ja200173a.
- (24) Hsiao, Y. Y.; Duh, Y.; Chen, Y. P.; Wang, Y. T.; Yuan, H. S. How an Exonuclease Decides Where to Stop in Trimming of Nucleic Acids: Crystal Structures of RNase T-Product Complexes. *Nucleic Acids Res.* **2012**, *40* (16), 8144–8154. https://doi.org/10.1093/nar/gks548.
- (25) Cheng, H.-L.; Lin, C.-T.; Huang, K.-W.; Wang, S.; Lin, Y.-T.; Toh, S.-I.; Hsiao, Y.-Y. Structural Insights into the Duplex DNA Processing of TREX2. *Nucleic Acids Res.* 2018, 46 (22), 12166– 12176. https://doi.org/10.1093/nar/gky970.
- (26) Cheng, Y.; Patel, D. J. Crystallographic Structure of the Nuclease Domain of 3'hExo, a DEDDh Family Member, Bound to RAMP. J. Mol. Biol. 2004, 343 (2), 305–312.

https://doi.org/10.1016/j.jmb.2004.08.055.

(27) Huang, K.-W.; Hsu, K.-C.; Chu, L.-Y.; Yang, J.-M.; Yuan, H. S.; Hsiao, Y.-Y. Identification of Inhibitors for the DEDDh Family of Exonucleases and a Unique Inhibition Mechanism by Crystal Structure Analysis of CRN-4 Bound with 2-Morpholin-4-Ylethanesulfonate (MES). *J. Med. Chem.* **2016**, *59* (17), 8019–8029.

https://doi.org/10.1021/acs.jmedchem.6b00794.

- (28) Zuo, Y.; Deutscher, M. P. Exoribonuclease Superfamilies: Structural Analysis and Phylogenetic Distribution. *Nucleic Acids Res.* 2001, 29 (5), 1017–1026. https://doi.org/10.1093/nar/29.5.1017.
- (29) Hamdan, S.; Carr, P. D.; Brown, S. E.; Ollis, D. L.; Dixon, N. E. Structural Basis for Proofreading during Replication of the Escherichia Coli Chromosome. *Structure* 2002, *10* (4), 535– 546. https://doi.org/10.1016/S0969-2126(02)00738-4.
- (30) Cisneros, G. A.; Perera, L.; Schaaper, R. M.; Pedersen, L. C.; London, R. E.; Pedersen, L. G.; Darden, T. A. Reaction Mechanism of the \$ɛ\$ Subunit of E. Coli DNA Polymerase III: Insights into Active Site Metal Coordination and Catalytically Significant Residues. J. Am. Chem. Soc. 2009, 131 (4), 1550– 1556. https://doi.org/10.1021/ja8082818.
- (31) Stevens, D. R.; Hammes-Schiffer, S. Exploring the Role of the Third Active Site Metal Ion in DNA Polymerase η with QM/MM Free Energy Simulations. J. Am. Chem. Soc. 2018, 140 (28), 8965–8969. https://doi.org/10.1021/jacs.8b05177.
- (32) Nakamura, T.; Zhao, Y.; Yamagata, Y.; Hua, Y.; Yang, W. Watching DNA Polymerase η Make a Phosphodiester Bond. *Nature* 2012, 487 (7406), 196–201. https://doi.org/10.1038/nature11181.
- Rosta, E.; Yang, W.; Hummer, G. Calcium Inhibition of Ribonuclease H1 Two-Metal Ion Catalysis. J. Am. Chem. Soc. 2014, 136 (8), 3137–3144. https://doi.org/10.1021/ja411408x.
- Yoon, H.; Warshel, A. Simulating the Fidelity and the Three Mg Mechanism of Pol η and Clarifying the Validity of Transition State Theory in Enzyme Catalysis. Proteins Struct. Funct. Bioinforma. 2017, 85 (8), 1446–1453. https://doi.org/10.1002/prot.25305.
- (35) Fukunishi, H.; Watanabe, O.; Takada, S. On the Hamiltonian Replica Exchange Method for Efficient Sampling of Biomolecular Systems: Application to Protein Structure Prediction. J. Chem. Phys. 2002, 116 (20), 9058–9067. https://doi.org/10.1063/1.1472510.
- Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. J. Comput. Phys. 1977, 23 (2), 187–199. https://doi.org/10.1016/0021-9991(77)90121-8.
- (37) Weinan, E.; Ren, W.; Vanden-Eijnden, E.; E, W.; Ren, W.; Vanden-Eijnden, E. Finite Temperature String Method for the Study of Rare Events. J. Phys. Chem. B 2005, 109 (14), 6688– 6693. https://doi.org/10.1021/jp0455430.
- (38) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; Beer, D. T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL : Homology Modelling of Protein Structures and Complexes. **2018**, *46* (May), 296–303. https://doi.org/10.1093/nar/gky427.
- (39) Tian, L.; Kim, M.-S.; Li, H.; Wang, J.; Yang, W. Structure of HIV-1 Reverse Transcriptase Cleaving RNA in an RNA/DNA Hybrid. Proc. Natl. Acad. Sci. U. S. A. 2018, 115 (3), 507–512. https://doi.org/10.1073/pnas.1719746115.
- (40) Abbondanzieri, E. A.; Bokinsky, G.; Rausch, J. W.; Zhang, J. X.; Le Grice, S. F. J.; Zhuang, X. Dynamic Binding Orientations Direct Activity of HIV Reverse Transcriptase. *Nature* 2008, 453, 184–189. https://doi.org/10.1038/nature06941.
- (41) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. J. Comput. Chem. 2008, 29 (11), 1859–1865. https://doi.org/10.1002/jcc.20945.
- Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone φ, ψ and Side-Chain χ₁ and χ₂ Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8* (9), 3257–3273. https://doi.org/10.1021/ct300400x.

- Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.;
 Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* 2005, *26* (16), 1781–1802. https://doi.org/10.1002/jcc.20289.
- (44) Goldschmidt, V.; Didierjean, J.; Ehresmann, B.; Ehresmann, C.; Isel, C.; Marquet, R. Mg2+ Dependency of HIV-1 Reverse Transcription, Inhibition by Nucleoside Analogues and Resistance. *Nucleic Acids Res.* 2006, 34 (1), 42–52. https://doi.org/10.1093/nar/gkj411.
- (45) Foloppe, N.; Mackerell, A. D. I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *Nucleic Acids Res.* 2000, 21 (2), 86–104.
- (46)Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio, R. A.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnava, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V; Levchenko, S. V; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J. W.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package. Mol. Phys. 2015, 113 184-215. (2).https://doi.org/10.1080/00268976.2014.952696.
- (47) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; DiStasio Jr, R. A.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Yeh Lin, C.; Van Voorhis, T.; Hung Chien, S.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Min Rhee, Y.; Ritchie, J.; Rosta, E.; David Sherrill, C.; Simmonett, A. C.; Subotnik, J. E.; Lee Woodcock III, H.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer III, H. F.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. Advances in Methods and Algorithms in a Modern Quantum Chemistry Program Package. Phys. Chem. Phys. 2006, 3172-3191. Chem. 8 (27), https://doi.org/10.1039/B517914A.
- Shao, Y.; Kong, J. YinYang Atom: A Simple Combined Ab Initio Quantum Mechanical Molecular Mechanical Model. J. Phys. Chem. A 2007, 111 (18), 3661–3671. https://doi.org/10.1021/jp067307q.
- (49) Rosta, E.; Woodcock, H. L.; Brooks, B. R.; Hummer, G. Artificial

ReactionCoordinate"Tunneling"inFree-EnergyCalculations:The Catalytic Reaction of RNaseH. J. Comput.Chem.2009,30(11),1634–1641.https://doi.org/10.1002/jcc.21312.1634–1641.

- (50) 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* (23), 8934–8941. https://doi.org/10.1021/ja200173a.
- (51) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. J. Chem. Phys. 1993, 98 (7), 5648– 5652. https://doi.org/10.1063/1.464913.
- Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* 2009, *30* (10), 1545–1614. https://doi.org/10.1002/jcc.21287.
- 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 (8), 1011–1021. https://doi.org/10.1002/jcc.540130812.
- (54) Rosta, E.; Hummer, G. Free Energies from Dynamic Weighted Histogram Analysis Using Unbiased Markov State Model. J. Chem. Theory Comput. 2015, 11 (1), 276–285. https://doi.org/10.1021/ct500719p.
- (55) Gan, J.; Shaw, G.; Tropea, J. E.; Waugh, D. S.; Court, D. L.; Ji, X. A Stepwise Model for Double-Stranded RNA Processing by Ribonuclease III. *Mol. Microbiol.* **2008**, *67* (1), 143–154. https://doi.org/10.1111/j.1365-2958.2007.06032.x.
- (56) Sondergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of p K a Values. J. Chem. Theory Comput. 2011, 7 (7), 2284–2295. https://doi.org/10.1021/ct200133y.
- (57) Keck, J. L.; Goedken, E. R.; Marqusee, S. Activation / Attenuation Model for RNase H. J. Biol. Chem. 1998, 273 (51), 34128–34133.
- (58) 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 (39), 13702–13712. https://doi.org/10.1021/ja102933y.
- (59) Nowotny, M.; Yang, W. Stepwise Analyses of Metal Ions in RNase H Catalysis from Substrate Destabilization to Product Release. *EMBO J.* 2006, 25 (9), 1924–1933. https://doi.org/10.1038/sj.emboj.7601076.
- (60) Schatz, O.; Cromme, V. F.; Grüninger-Leitch, F.; Grice, L. S. F. J. Point Mutations in Conserved Amino Acid Residues within the C-Terminal Domain of HIV-l Reverse Transcriptase Specifically Repress RNase H Function. *FEBS J.* **1989**, 257 (2), 311–314.
- (61) Kashiwagi, T.; Jeanteur, D.; Haruki, M.; Katayanagi, K.; Kanaya, S.; Morikawa, K. Proposal for New Catalytic Roles for Two Invariant Residues in Escherichia Coli Ribonuclease HI. *Protein Eng.* **1996**, *9* (10), 857–867.
- (62) Bennie, S. J.; Van Der Kamp, M. W.; Pennifold, R. C. R.; Stella, M.; Manby, F. R.; Mulholland, A. J. A Projector-Embedding Approach for Multiscale Coupled-Cluster Calculations Applied to Citrate Synthase. J. Chem. Theory Comput. 2016, 12 (6), 2689–2697. https://doi.org/10.1021/acs.jctc.6b00285.
- (63) Hamki, M.; Tsunaka, Y.; Morikawa, M.; Iwai, S.; Kanaya, S. Catalysis by Escherichia Coli Ribonuclease HI Is Facilitated by a Phosphate Group of the Substrate. *Biochemistry* 2000, *39* (45), 13939–13944. https://doi.org/10.1021/bi001469+.
- (64) Leonarski, F.; D'Ascenzo, L.; Auffinger, P. Mg2+ions: Do They Bind to Nucleobase Nitrogens? *Nucleic Acids Res.* 2017, 45 (2), 987–1004. https://doi.org/10.1093/nar/gkw1175.

- (65) Kanaya, E.; Kanaya, S. Kinetic Analysis of Escherichia Coli Ribonuclease HI Using Oligomeric DNA/RNA Substrates Suggests an Alternative Mechanism for the Interaction between the Enzyme and the Substrate. *Eur. J. Biochem.* **1995**, 231 (3), 557–562. https://doi.org/10.1111/j.1432-1033.1995.0557d.x.
- Lima, W. F.; Wu, H.; Nichols, J. G.; Prakash, T. P.; Ravikumar, V.; Crooke, S. T. Human RNase H1 Uses One Tryptophan and Two Lysines to Position the Enzyme at the 3???-DNA/5???-RNA Terminus of the Heteroduplex Substrate. J. Biol. Chem. 2003, 278 (50), 49860–49867. https://doi.org/10.1074/jbc.M306543200.
- (67) Shaw-Reid, C. A.; Feuston, B.; Munshi, V.; Getty, K.; Krueger, J.; Hazuda, D. J.; Parniak, M. A.; Miller, M. D.; Lewis, D. Dissecting the Effects of DNA Polymerase and Ribonuclease H Inhibitor Combinations on HIV-1 Reverse-Transcriptase Activities. *Biochemistry* 2005, 44 (5), 1595–1606. https://doi.org/10.1021/bi0486740.
- (68) Kanaya, S.; Oobatake, M.; Liu, Y. Thermal Stability of Escherichia Coli Ribonuclease HI and Its Active Site Mutants in the Presence and Absence of the Mg 2 Ion. *Biochemistry* 1996, 271 (51), 32729–32736.
- (69) Humphrey, W.; Dalke, A.; Schulten, K. VMD Visual Molecular Dynamics. J. Mol. Graph. **1996**, *14*, 33–38.
- (70) Majorek, K. A.; Dunin-Horkawicz, S.; Steczkiewicz, K.; Muszewska, A.; Nowotny, M.; Ginalski, K.; Bujnicki, J. M. The RNase H-like Superfamily: New Members, Comparative Structural Analysis and Evolutionary Classification. *Nucleic* Acids Res. **2014**, 42 (7), 4160–4179.

https://doi.org/10.1093/nar/gkt1414.

- (71) Palermo, G. Structure and Dynamics of the CRISPR-Cas9 Catalytic Complex. J. Chem. Inf. Model. 2019, 59 (5), 2394– 2406. https://doi.org/10.1021/acs.jcim.8b00988.
- Nowotny, M.; Figiel, M. The RNase H Domain: Structure, Function and Mechanism. In *Human Immunodeficiency Virus Reverse Transcriptase: A Bench-to-Bedside Success*; LeGrice, S. F. J., Goette, M., Eds.; Springer Science, 2013; pp 77–95. https://doi.org/10.1007/978-1-4614-7291-9.
- (73) Hsiao, Y. Y.; Yang, C. C.; Lin, C. L.; Lin, J. L.; Duh, Y.; Yuan, H. S. Structural Basis for RNA Trimming by RNase T in Stable RNA 3'-End Maturation. *Nat Chem Biol* **2011**, 7 (4), 236–243. https://doi.org/10.1038/nchembio.524.
- Hsiao, Y. Y.; Fang, W. H.; Lee, C. C.; Chen, Y. P.; Yuan, H. S. Structural Insights Into DNA Repair by RNase T-An Exonuclease Processing 3??? End of Structured DNA in Repair Pathways. *PLoS Biol.* 2014, 12 (3). https://doi.org/10.1371/journal.pbio.1001803.
- Pontius, B. W.; Lott, W. B.; von Hippel, P. H. Observations on Catalysis by Hammerhead Ribozymes Are Consistent with a Two-Divalent-Metal-Ion Mechanism. *Proc. Natl. Acad. Sci. USA* **1997**, 94 (6), 2290–2294. https://doi.org/10.1073/pnas.94.6.2290.
- (76) Casalino, L.; Nierzwicki, Ł.; Jinek, M.; Palermo, G. Catalytic Mechanism of Non-Target DNA Cleavage in CRISPR-Cas9 Revealed by Ab Initio Molecular Dynamics. ACS Catal. 2020, 10 (22), 13596–13605. https://doi.org/10.1021/acscatal.0c03566.



