

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/259649823>

Magnesium-Dependent RNA Binding to the PA Endonuclease Domain of the Avian Influenza Polymerase

ARTICLE in THE JOURNAL OF PHYSICAL CHEMISTRY B · JANUARY 2014

Impact Factor: 3.3 · DOI: 10.1021/jp408383g · Source: PubMed

CITATIONS

7

READS

25

7 AUTHORS, INCLUDING:



Shiyan Xiao

University of Science and Technology of China

16 PUBLICATIONS 61 CITATIONS

[SEE PROFILE](#)



Haojun Liang

University of Science and Technology of China

152 PUBLICATIONS 2,249 CITATIONS

[SEE PROFILE](#)



Christopher M Macdermaid

Temple University

11 PUBLICATIONS 154 CITATIONS

[SEE PROFILE](#)



Mercedes Alfonso-Prieto

University of Barcelona

30 PUBLICATIONS 413 CITATIONS

[SEE PROFILE](#)

Magnesium-Dependent RNA Binding to the PA Endonuclease Domain of the Avian Influenza Polymerase

Shiyan Xiao,[†] Michael L. Klein,[‡] David N. LeBard,[§] Benjamin G. Levine,^{||} Haojun Liang,[†] Christopher M. MacDermaid,[‡] and Mercedes Alfonso-Prieto*,[‡]

[†]CAS Key Laboratory of Soft Matter Chemistry, Department of Polymer Science and Engineering, University of Science and Technology of China, Hefei, Anhui 230026, P. R. China

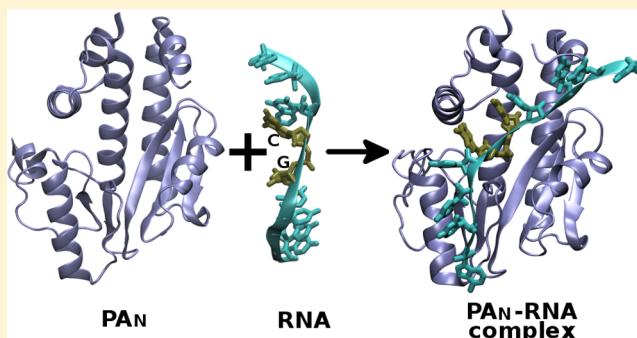
[‡]Institute for Computational Molecular Science, Temple University, Philadelphia, Pennsylvania 19122, United States

[§]Department of Chemistry, Yeshiva University, New York, New York 10033, United States

^{||}Department of Chemistry, Michigan State University, East Lansing, Michigan 48824-1322, United States

S Supporting Information

ABSTRACT: Influenza A viruses are highly pathogenic and pose an unpredictable public health danger to humans. An attractive target for developing new antiviral drugs is the PA N-terminal domain (PA_N) of influenza polymerase, which is responsible for the endonuclease activity and essential for viral replication. Recently, the crystal structures of the holo form of PA_N as well as PA_N bound to different inhibitors have been reported, but the potency and selectivity of these inhibitors still need to be improved. New drug design can be guided by a better understanding of the endonuclease activity of PA_N . However, this requires the structure of PA_N in complex with the host mRNA, which has not been determined yet. In particular, divalent metal ions are known to be essential for RNA cleavage, but it is not clear whether there is either one or two Mg ions in the PA_N active site. In the present work, we have modeled the complex of the PA_N endonuclease domain with the host mRNA in the presence of either one or two Mg^{2+} by using all-atom molecular dynamics. These simulations identify crucial interactions between the enzyme and the nucleic acid. Moreover, they validate a previous hypothesis that a second metal ion binds in the presence of the RNA substrate and therefore support a two-metal ion mechanism, in which K134 decreases the pK_a of the nucleophilic water. Nevertheless, at low Mg concentrations an alternative, one-metal ion mechanism is possible, with K137 as the catalytic lysine and H41 as the general base, rationalizing previous unexpected mutagenesis results. The RNA–enzyme interactions determined here could likely be used to design more specific endonuclease inhibitors to fight influenza viral infections.



INTRODUCTION

The highly pathogenic influenza A virus poses a growing and unpredictable threat to world public health, with severe pandemics occurring in 1918, 1957, and 1968.¹ Recently, the avian originated H5N1 influenza virus turned out to be highly pathogenic in humans and is thought to be a candidate for the next pandemic upon adaptation for human-to-human transmission.^{2,3} Current antiviral inhibitors are directed against the influenza M2 ion channel (target of adamantanes^{4,5}) and neuraminidase (target of flu drugs Oseltamivir⁶ and Zanamivir⁷). However, the lack of proofreading function in the influenza virus polymerase results in a high rate of gene mutation and thus the rapid development of antiviral drug resistance,⁸ making it urgent to develop alternative antiviral drugs. The influenza polymerase, which is involved in both the transcription and replication of the viral RNA (vRNA), plays an essential role in infection and evolution of influenza⁹ and thus

has been identified as a novel and attractive target for new antiviral drugs.^{10–14}

The influenza polymerase is a ~250 kDa RNA-dependent RNA polymerase (RdRP) that synthesizes the viral RNA. Transcription initiation requires short capped primers that are obtained from the host cell pre-mRNA through a “cap-snatching” mechanism.¹⁵ In this process, the polymerase cuts the 5' 7-methyl-guanosine cap-containing oligonucleotide of the host cell pre-mRNA, then it attaches the primer onto the viral mRNA, and finally it polyadenylates the 3' end.^{16,17} Structurally, the enzyme is a heterotrimer formed by subunits PA, PB1, and PB2. Crystallographic studies show that PB1 works as the central structural component of the polymerase, with no direct interaction between PA and PB2.^{18–20} The PA

Received: August 21, 2013

Revised: January 9, 2014

Published: January 9, 2014

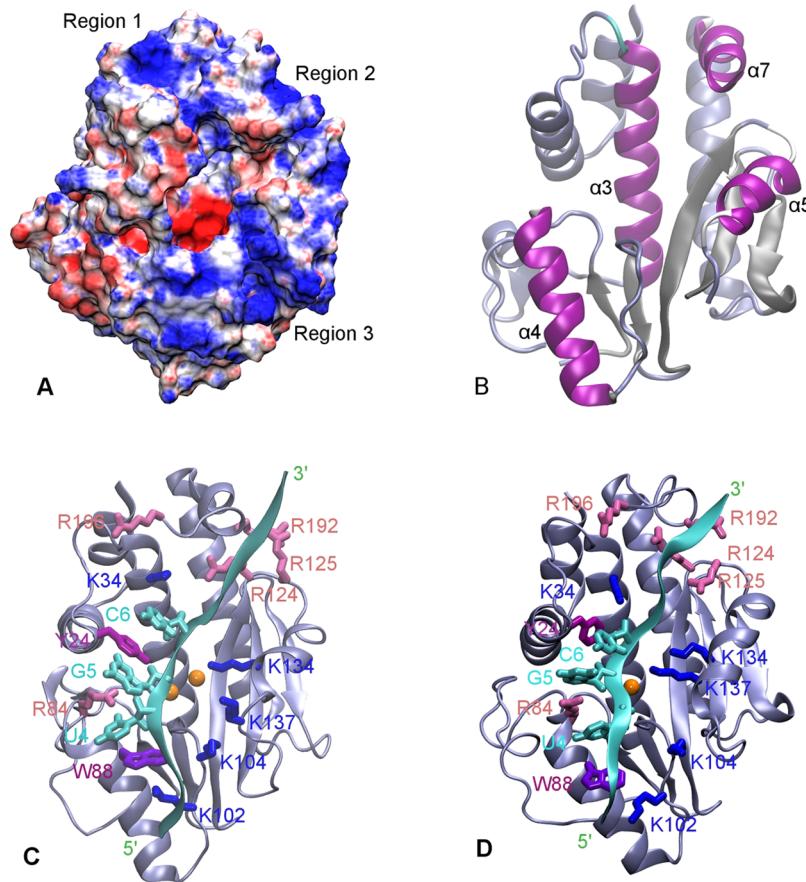


Figure 1. (A) Electrostatic surface potential of PA_N , ranging from $-10.0k_{\text{B}}T/e$ (red) to $10.0k_{\text{B}}T/e$ (blue). (B) Cartoon representation of PA_N , showing the main α helices involved in RNA binding. (C) Cartoon representation of the two- Mg^{2+} PA_N –RNA complex. PA_N is in gray, RNA is in cyan, and the Mg^{2+} ions are in orange. A licorice representation is used for the protein residues important for RNA binding (in blue, pink, and maroon) and the three central nucleotides of the RNA strand containing the cleavage site and the adjacent 3' phosphodiester (in cyan). (D) Cartoon representation of the one- Mg^{2+} PA_N –RNA complex. The color code is the same as in (C). In both cases we have chosen the complexes obtained in run 1 as representatives.

subunit can be further divided into the N-terminal (PA_N) and the C-terminal (PA_C) domains, which are linked by a long flexible peptide. PA_N is thought to be responsible for the endonuclease activity crucial for the cap-snatching mechanism^{21,22} and is an attractive target for developing new antiviral drugs.^{10–14} Inhibitors binding to the highly conserved endonuclease active site are less prone to generate viral resistance, and they can stop viral replication at the initial phase of the infection.

Recently, the crystallographic structure of the holo form of PA_N has been reported by several groups.^{21–24} PA_N is a ~ 25 kDa domain with α/β architecture (Figure 1A,B). It consists of five β -strands surrounded by seven α -helices, forming a highly negatively charged active pocket that binds divalent metal ions. However, it is not clear the number and type of divalent metal ions required for the endonuclease activity. The structure by Yuan et al.²² contains a single Mg^{2+} ion (hereafter Mg2) that is coordinated by E80, D108, and three water molecules stabilized by H41, E119, L106, and L107. In contrast, Dias et al. reported two Mn^{2+} ions in the active pocket.²¹ Mn2 is equivalent to the Mg ion in the structure by Yuan et al. and is coordinated by E80, D108, and two water molecules. On the other hand, Mn1 is coordinated by D108, E119, H41, and the carbonyl oxygen of I120, and interestingly, D108 and E119 (along with P107 and K134/K137) can constitute a (P)D(X)_N(D/E)XK motif,

present in many type II endonucleases. Similarly, two metal ions are also present in the structures by Kowalinski et al.²³ and DuBois et al.,²⁴ which were crystallized in the presence of both MnCl_2 and MgCl_2 . Both structures show a Mn ion at site 1, whereas site 2 is occupied by either Mg²³ or Mn.²⁴

The discrepancy in the number and type of divalent ions among these crystallographic structures can be explained in view of the results of the metal binding experiments by Crépin and co-workers.²⁵ They found that PA_N contains two divalent metal ion binding sites, one with a higher affinity than the other, and that the active pocket has a higher affinity for Mn than for Mg ions. Taken together with the crystallographic data,²³ it has been suggested that, in the absence of the substrate, Mg binds preferentially to site 2, whereas Mn is more favored at site 1, most likely due to the presence of H41. In other words, site 1 is a low affinity site for Mg, but a high affinity site for Mn, because Mg ions have lower affinity for histidine than Mn.^{26–29}

However, the situation may change upon binding of the RNA substrate. Indeed, the activity dependence on the metal ion concentration shows a Hill coefficient of 2,³⁰ suggesting that the endonuclease cavity is loaded with two metal ions in the presence of RNA substrate. Although the PA_N –RNA reactive complex has not been crystallized yet, there are several structures of PA_N in complex with different nucleoside

monophosphates (NMP) available, mimicking the product of the endonuclease reaction.^{23,31} The structure by Zhao et al.³¹ shows a single Mg^{2+} ion located at site 2, whereas site 1 is occupied by a water molecule. In contrast, Kowalinski et al.²³ found two Mn^{2+} ions in the active site, and a different conformation of the bound NMP. These discrepancies have been attributed to the different crystallization methods used.²³

In any case, the structure of the PA_N domain in complex with the host mRNA substrate has not yet been determined, and thus the question of the number and nature of the metal ions involved in the endonuclease reaction remains to be fully clarified. Knowledge of the PA_N -RNA complex structure is crucial for understanding the endonuclease reaction mechanism and can assist the design of more specific flu inhibitors targeting the endonuclease activity.

In the present work, we have investigated the interaction between PA_N and the host mRNA in the presence of Mg^{2+} . Despite the higher affinity and endonuclease activity of Mn compared to those of Mg,²⁵ it is more likely that Mg ions are the natural cofactors of PA_N ³¹ because the cellular free Mg concentration (approximately millimolar) is about 1000 times higher than that of Mn (approximately micromolar).^{32–35} In particular, we use molecular dynamics (MD) simulations to (i) construct structural models of the PA_N -RNA complex in the presence of either one or two Mg ions and (ii) investigate the stability of the two metal ions in the active site of the holo form of PA_N . Simulations are a powerful tool to understand biological processes at the atomic level and have been successfully applied in the field of nucleases,³⁶ e.g., predicting protein–nucleic acid complexes,^{37–45} investigating the role of metal ions in nucleic acid recognition,^{46–50} unraveling endonuclease catalytic mechanisms,^{51–56} or explaining the activity dependence on the Mg ion concentration.⁵⁷ The simulations reported here identify the structural determinants of the RNA binding to PA_N and also provide information about the conformation and dynamics of the endonuclease active site in the presence of either one or two Mg^{2+} ions. Namely, we have investigated the coordination between the phosphodiester bond to be cleaved during the reaction, the metal ion(s) and the nucleophilic water molecule, as well as the position of the catalytic lysine. On the basis of these results, we suggest a possible mechanism for the endonuclease activity of PA_N in the presence of Mg^{2+} . Unraveling the molecular basis of the Mn^{2+} dependent endonuclease activity of PA_N will be the subject of future studies.

SIMULATION DETAILS

Modeled Systems. *Two- Mg^{2+} Bound PA_N System.* The initial structure of the two- Mg^{2+} holo form of PA_N is based on the two- Mn^{2+} containing PA_N crystal structure (PDB code 2W69, 2.05 Å resolution)³¹ by replacing the Mn^{2+} ions present in the active site with Mg^{2+} . Mn is often used in crystallography as a replacement of Mg because it has a higher electron density, while maintaining a similar coordination geometry. Molecular dynamics (MD) was carried out on this system to characterize the Mg binding to PA_N . Three independent simulations were performed by assigning different initial velocity distributions to the minimized system (hereafter runs 1, 2, and 3, respectively) and each run was extended for ~100 ns. We note here that we have not attempted to model PA_N bound to Mn^{2+} (either the two-Mn state^{21,24} or the hybrid Mn–Mg state²³). Although a nonbonded force field for Mn has been recently published,⁵⁸ this model was not available at the time this investigation was

started, and testing its transferability to endonucleases is out of the scope of the present work. Moreover, computing the charge transfer from the ligands to the metal ion (see below) for the paramagnetic Mn^{2+} requires knowledge of the spin state of the Mn ion. Although Mn^{2+} has been found to be in a high spin state in most Mn-containing proteins, EPR spectroscopy will be needed to confirm this assignment in PA_N .

One- Mg^{2+} Bound PA_N System. The initial structure of the one- Mg^{2+} holo form of PA_N is also based on the crystal structure of the two- Mn^{2+} containing PA_N crystal structure by removing Mn1 and replacing the Mn^{2+} ion at site 2 by Mg^{2+} . MD was used to explore the stability of the metal ion bound to PA_N . As for the other holo system, three independent simulations were performed by assigning different initial velocity distributions to the minimized system (hereafter runs 1, 2, and 3, respectively) and each run was extended for ~100 ns.

Two- Mg^{2+} PA_N -RNA Complex System. To construct a model of the PA_N -RNA complex, we applied a homology modeling-like approach, using as a template the crystal structures of the two- Mn^{2+} bound form of PA_N ²¹ and *Bacillus halodurans* RNase H in complex with a hybrid RNA:DNA substrate⁵⁹ (PDB code 1ZBL, 2.20 Å resolution). Both endoribonucleases share a similar structure (with the strands $\beta 1-\beta 5$ and nearby helices αA and αD of RNase H resembling the central twisted β -strands and the helices $\alpha 2$ and $\alpha 3$ of PA_N) and their active sites are comparable (Figure S1A, Supporting Information). After superimposing the two structures, we merged the RNA strand in the RNase H-substrate complex onto PA_N and mutated its sequence to 5'-ACUUGCUUUU-3' using the VMD software package.⁶⁰ A U-rich oligonucleotide with a central GC cleavage site is commonly used to measure the endonuclease activity of the influenza polymerase.²² Moreover, this 10-mer substrate is twice longer than the 5-mer minimal substrate cleaved efficiently by PA.⁶¹ Therefore, by using this substrate, we can obtain information not only about the catalytic mechanism but also about the binding of the flanking regions, while still maintaining affordability in terms of computer time. Finally, we replaced the two Mn^{2+} ions in the original structure with two Mg^{2+} ions.

A similar modeling strategy was applied by Zhao et al.,³¹ using their crystal structure of PA_N in complex with an NMP and the X-ray structure of EcoRV in complex with an 11-mer DNA substrate.⁶² The resulting model showed that the catalytic residues of both type II endonucleases are superimposable (Figure S1B, Supporting Information), but the NMP and DNA substrates are not. This is most likely due to the lack of flexibility of their modeling approach. To overcome this limitation, we refine our two- Mg^{2+} PA_N -RNA complex by means of steered MD simulations, to optimize the binding of the RNA substrate onto PA_N (Supporting Information). RNA binding is a dynamical process that occurs in the microsecond–millisecond time scale,⁶³ much longer than the nanosecond time scale that can be easily achieved in MD simulations, and thus an enhanced sampling technique such as steered MD is required to accelerate the RNA binding. Three structures were selected from the end of the steered MD trajectory and then relaxed for 20 ns using unrestrained MD (hereafter runs 1, 2, and 3, respectively). For each run, an additional ~80 ns MD was performed to test the stability of the proposed binding mode and to analyze the structure of the active site in relation to the endonuclease reaction.

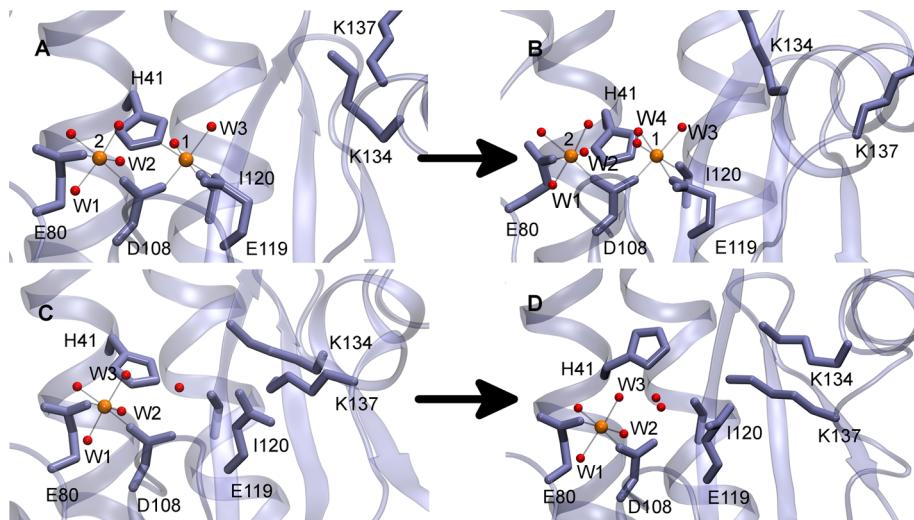


Figure 2. Active site structure of two- Mg^{2+} (A, B) and one- Mg^{2+} (C, D) bound PA_N . The key protein residues are shown as light blue licorice, the Mg ions are in orange, and the water molecules coordinated to the metal ions are in red. (A) Two- Mg^{2+} apo configuration where H41 is coordinated to Mg1. (B) Two- Mg^{2+} apo configuration where a water molecule replaces H41 in the coordination shell of Mg1. (C) One- Mg^{2+} apo configuration where site 1 is occupied by one water molecule. (D) One- Mg^{2+} apo configuration where two water molecules are present in site 1 and H41 is rotated.

One- Mg^{2+} PA_N -RNA Complex System. The one- Mg^{2+} PA_N -RNA complex was constructed from three fully equilibrated snapshots of the two- Mg^{2+} PA_N -RNA complex (at 10, 20, and 30 ns, respectively) by first removing Mg1, followed by a 2 ns MD simulation to further relax the system. A 100 ns MD simulation was then performed on each of these initial structures (hereafter runs 1, 2, and 3, respectively) to test the stability of the complex.

Molecular Dynamics. The modeled systems were immersed in a $\sim 106 \text{ \AA} \times \sim 100 \text{ \AA} \times \sim 94 \text{ \AA}$ water box composed of $\sim 31\,400$ water molecules plus 150 mM NaCl. The resulting systems contain a total of $\sim 98\,000$ atoms each. Protein and RNA were described using the Cornell et al.⁶⁴ and the parm99bsc0⁶⁵ force fields, respectively, and the TIP3P model⁶⁶ was used for the water molecules.

The metal-containing active site was treated using a flexible nonbonded approach based on the “atoms in molecules” partitioning scheme⁶⁷ of the DFT-BLYP level electronic density⁶⁸ (Supporting Information). This model accounts for the metal–ligand charge transfer effects, while allowing us to explore possible structural rearrangements (such as ligand exchange) during the MD simulations. In other words, this nonbonded model does not impose any coordination geometry, thus permitting us to investigate the flexibility of metal ion in the active site, which is one of the aims of the present study. The Bader-corrected charges for the metal and the ligands are listed in Tables S1–S4 (Supporting Information). The Mg^{2+} Lennard-Jones interactions were calculated using the Åqvist's parameters⁶⁹ for the metal ion.

All simulations were performed with the NAMD program.⁷⁰ The SHAKE/RATTLE algorithm⁷¹ was used to constrain the bonds involving hydrogen atoms in both the solvent and the solute. Multiple time step integration was carried out using r-RESPA,⁷² with a base time step of 2 fs and a secondary time step of 4 fs for long-range interactions. Periodic boundary conditions were applied and the particle mesh Ewald (PME) method⁷³ was employed to evaluate the long-range electrostatic interactions. The systems were simulated in the *NPT* ensemble, with a Nosé–Hoover thermostat⁷⁴ and a Langevin piston

Nosé–Hoover barostat to maintain the temperature at 300 K and the pressure at 1 atm, respectively. Each system was simulated for 100 ns and the last 80 ns were used for analysis.

RESULTS

Holo Form of PA_N . Two- Mg^{2+} Bound PA_N . The available crystal structures of the holo form of PA_N ^{21–24} show either Mg^{2+} ^{22,23} or Mn^{2+} ^{21,24} bound to site 2, whereas site 1 is empty²² or occupied by Mn.^{21,23,24} However, the electron density on site 2 is much weaker than for site 1, complicating the crystallographic refinement and the divalent metal ion identification. Moreover, as mentioned in the Introduction, Mg is expected to be the main cofactor of PA_N in vivo.³¹ Therefore, MD simulations can be useful to obtain a structural model of the two- Mg^{2+} holo form to complement the crystallographic studies. Besides, because the type of metal ions has been shown to affect both the cleavage pattern⁶¹ and inhibitor potency,⁷⁵ obtaining a model for the two- Mg^{2+} bound PA_N , in contrast to the two- Mn^{2+} ^{21,24} or Mn– Mg^{2+} ²³ containing PA_N crystal structures, can also be useful to predict reactivity and to design drugs.⁶¹

The overall structure of the two- Mg^{2+} containing PA_N model is very similar to that of the two- Mn^{2+} apo crystal structure by Dias and co-workers,²¹ with a backbone RMSD of 1.88 (run 1) and 2.11 (run 2) Å (Figure S2 and Table S5, Supporting Information). Only run 3 exhibits a somewhat larger RMSD (3.05 Å), which can be explained by the tilting of helix α 5 and the movement of the loop following helix α 4. Most likely, the flexibility of these secondary structural elements is overestimated in our model due to the absence of the other domains of the influenza polymerase. The most variable region in all three runs is the loop spanning residues 53–73, as evidenced by the high per-residue RMSD values in Figure S3 (Supporting Information). The flexibility of this solvent exposed loop is not unexpected, because the available crystal structures show either a disordered region²² or different orientations of the loop.^{21,23} Indeed, this region explores several conformations during the 100 ns time scale of our simulations (Figure S4, Supporting Information). Interestingly,

it has been suggested that closing of the loop may interfere with binding of inhibitors to the active site.^{23,24} To overcome this problem, the recently solved crystal structures of PA_N in complex with different inhibitors^{23,24} were generated using loop-deleted forms of PA_N. Therefore, the structural model of the full-length two-Mg²⁺ PA_N presented here, exploring different (open and closed) loop conformations, could potentially be used in computational docking studies, along with the experimental (loop-deleted) PA_N-inhibitor complex structures.^{23,24}

Regarding the endonuclease active site, Mg2 is coordinated by E80, D108 and four water molecules, whereas Mg1 is interacting with D108, E119, I120, and three water molecules (Figure 2A,B). These metal coordination spheres are slightly different from the crystallographic structures. In the two-Mn²⁺ containing PA_N crystal structure by Dias et al.²¹ (used as a starting point of our simulations), the Mn ion at site 2 is also coordinated by protein residues E80 and D108, but only two water molecules. However, in that X-ray structure the metal coordination is most likely affected by the intrusion of a loop from a neighbor crystal molecule. In our simulation, the distortion caused by crystal packing is relieved by relaxing the protein in a solution environment, allowing the metal ion to achieve the expected 6-fold coordination, with four water molecules bound to Mg2. This is exactly the same metal coordination observed at site 2 in the Mg2–Mn1 bound PA_N crystal structure by Kowalinski and co-workers²³ and in the Mn2–Mn1 bound PA_N X-ray structure by DuBois et al.,²⁴ where the crystal packing artifact was avoided by expressing the PA_N protein without the problematic loop. This shows that the MD-based refinement of the PA_N models used here alleviates the possible bias due to the initial structure chosen for building the models.

The Mg2–carboxylate distances in our model are ~1.90–1.91 Å (Table S6 and Figure S9, Supporting Information), only slightly shorter than the typical Mg–Asp/Glu distances (2.07 ± 0.10 Å) observed in all Mg-containing structures deposited in the Protein Data Bank,²⁶ confirming the reliability of the classical nonbonded model used here to describe the Mg coordination. Nevertheless, the MD bond lengths are shorter (by ~0.2–0.5 Å) than the Mn2–ligand distances in the crystal structure used as a starting point of the simulations (PDB code 2W69).²¹ The different divalent ion present in the active site can account for this difference, because typical Mg–O distances are ~0.1 Å shorter than Mn–O distances^{26,76} due to the different ionic radius of the two metals. Moreover, the distances observed in the two-Mn²⁺ containing PA_N crystal should be taken with caution because, as mentioned above, metal coordination is slightly distorted by intrusion of a loop from a neighbor crystal molecule. As a consequence, the *B*-factor of the crystallographic Mn ions is much higher than the protein (Table S8, Supporting Information) and the Mn2–ligand experimental distances vary by ~0.1–0.3 Å (Table S6, Supporting Information) depending on which of the three molecules present in the asymmetric unit is considered. Altogether, our simulation shows that Mg binds at site 2 in holo PA_N through E80, D108, and four water molecules, in agreement with the most recent crystallographic data, and that the Mg2 binding mode is well maintained within the time scale of our simulation.

In contrast, Mg1 is found to be more flexible than Mg2. Initially, Mg1 is coordinated by D108, E119, I120, H41, and two water molecules (Figure 2A), similar to Mn1 in the crystal

structures.^{21,23,24} However, after a few tenths of nanoseconds of MD simulation (56, 32, or 20 ns for run 1, 2, or 3, respectively, see Figure S10, Supporting Information), the Ne@H41–Mg1 distance increases from ~2.24 to ~4.08 Å (Table S6, Supporting Information), such that Mg1 loses the interaction with H41 and the imidazole ligand is replaced by a water molecule (Figure 2B). The changes in Mg1 coordination can be traced back to the movement of helix α5, which shows a larger RMSD compared to those of other secondary structures (Table S5, Supporting Information). Helix α5 contains two conserved lysine residues, K134 and K137, whose positive charge most likely affects the orientation of the nearby negatively charged E119. Movement of helix α5 results in the rotation of E119, and the moving E119 is able to drag concomitantly Mg1, breaking the Mg–H41 interaction, because Mg has higher affinity for O-ligands than N-ligands.^{26–29} The loss of the H41 bond observed in our Mg1-containing model compared to the crystal structures with Mn1^{21,23,24} is most likely due to the change of divalent metal. Crépin et al.²⁵ have shown that the dissociation constants of Mg bound to PA_N are higher than for Mn. Besides, Mn ions are known to have higher affinity for histidine than Mg.^{26–29} Therefore, it is reasonable to expect some differences between our Mg1-containing model and the crystallographic data with Mn1. Moreover, even the crystal structures with Mn occupying site 1^{21,23,24} show some variability among the different molecules present in the asymmetric unit. For example, in the structure by Dias et al.,²¹ one out of the three molecules present in the asymmetric unit has an empty site 1, and in the other two the occupancy of the metal ion is less than 1, indicating a weak electron density. This suggests that the crystallographic structure most likely represents an average of the two-Mn²⁺ and one-Mn²⁺ structures. In addition, the X-ray structure shows different Mn1–Ne@H41 distances (2.47–2.91 Å, Table S6, Supporting Information), and they are significantly longer than the typical Mn–N bond length (2.21 Å),²⁶ most likely distorted due to the interaction with the Glu59 loop residue from a neighbor crystal molecule. If such a variability exists in the presence of the more strongly bound Mn²⁺ ion, it is reasonable that the flexibility at site 1 increases further when Mn is substituted by the less tightly bound Mg²⁺. It is also worth mentioning that a similar insertion of a water molecule in the active site due to a helix sliding has been observed in simulations of the de novo designed dimanganese DF1 protein.⁷⁷ As mentioned above, the use of a nonbonded model allows us to explore whether small structural rearrangements of the metal-containing active site are possible at room temperature.⁶⁸ Indeed, the simulations on the di-Mn DF1 protein provided a dynamic picture of how the wetting transition between the open and close active site conformations present in the crystal structure occurs. Therefore, it is tempting to suggest that a future crystal structure of the two-Mg containing PA_N might display the two different active site conformations described above, one with Mg1 coordinated to H41 and another with H41 replaced by a water molecule.

One-Mg²⁺ Bound PA_N. The fact that crystallographic studies could only trap Mg at site 2,^{22,23} the low affinity of Mg for site 1,²⁵ and the flexibility of the Mg1 coordination sphere in our two-Mg²⁺ apo simulation suggest that in the absence of the RNA substrate Mg ions may occupy only site 2. Therefore, we also simulated the one-Mg²⁺ holo form of PA_N.

The overall structure of our one-Mg²⁺ model is very similar to the one-Mg²⁺ apo crystal structure by Yuan et al.,²² with a

backbone RMSD of 1.42–1.77 Å (Figure S2 and Table S5, Supporting Information), again confirming that relaxation by MD removes any possible bias due to the crystal structure used to build the model (i.e., the X-ray structure by Dias et al.²¹). In the endonuclease active site, Mg2 is coordinated by E80, D108, and four water molecules (Figure 2C,D), similar to the two-Mg²⁺ apo model, and this octahedral coordination is stable for the 100 ns MD simulation. The average Mg–O bond lengths are ~1.92–1.93 Å for Asp/Glu, and ~2.05–2.07 Å for water (Table S7 and Figure S8, Supporting Information). These values are ~0.1–0.3 Å shorter than in the corresponding crystal structure.²² Nevertheless, the metal–ligand distances obtained here are closer to the average Mg–O bond lengths (2.07 and 2.15 Å, respectively) observed in the Mg-containing structures deposited in the Protein Data Bank.^{26,70}

The empty site 1 is occupied by either one or two water molecules. Initially, there is a single water molecule (W3) H-bonded to I120, E119, and H41 (Figure 2C) that is occupying the position of the missing metal ion. However, the interaction between H41 and W3 is weak, such that the distance between the two oscillates significantly and it can even break (i.e., the H41–W3 distance increases from 3.44 to 6.62 Å, Table S7 and Figure S8, Supporting Information). The rotation of H41 leaves a gap in the active site that is readily filled by a second water molecule coming from the solution (Figure 2D). Similar to the two-Mg²⁺ apo simulation, this active site conformational change is caused by the movement of helix α 5, which exhibits the largest RMSD value among all the secondary structure elements of PA_N (Table S5, Supporting Information). The electrostatic attraction of K134 and K137 in helix α 5 causes the rotation of E119, resulting in the formation of a salt bridge between E119 and K134. This leaves an empty space near site 1 that is easily filled by a water molecule from the solvent. In turn, insertion of this additional water molecule pushes H41 away from the active site. The discrepancy between our one-Mg²⁺ holo PA_N model and the corresponding X-ray structure²² can be attributed to the higher temperature and water content of the simulated system compared to that of the crystal. In this regard, it is noteworthy that the number of water molecules and their position, as well as the orientation of K134, varies among the different molecules present in the asymmetric unit of the crystal. In summary, our simulations show that site 1 is very flexible in the absence of the RNA substrate, in line with the crystallographic data for Mg.^{21–24}

Host mRNA-PA_N Complex. RNA Binding to PA_N. One of the principal features of RNA is the presence of phosphate groups resulting in a highly negatively charged backbone at physiological pH. Therefore, the possible binding regions of RNA on PA_N are likely positively charged. Hence we calculated the electrostatic potential surface of PA_N using the implementation of APBS⁷⁸ in VMD.⁶⁰ Figure 1A shows that there are three strongly positively charged regions (displayed in blue) on the surface of PA_N that may be involved in RNA binding, and Figure 1B displays the corresponding tertiary and secondary structures, in particular, the α helices possibly interacting with RNA. Region 1 consists of helices α 3 (containing K34) and α 7 (R196), region 2 is formed by helix α 7 (R196, R192) and the α 5- β 3 loop (R124, R125), and region 3 comprises helix α 4 (R84), the α 4- β 2 loop (K102, K104), and the α 5- β 4 loop (K139). It has been suggested³¹ that the 3' end of RNA binds to the C terminus of PA_N, where region 1 and region 2 are located, whereas the 5' end is predicted to interact with the broad cleft corresponding to region 3. This proposal

was based on the comparison of the crystal structures of the one-Mg²⁺ PA_N-NMP complex and the two-Mn²⁺ EcoRV–DNA complex, i.e., the product of the endonuclease reactions of PA_N and EcoRV, respectively. Both PA_N and EcoRV contain a (P)DX_N(D/E)XK endonuclease motif and share a highly similar active site (Figure S1B, Supporting Information), such that the binding mode of the 5'-3' strand of the DNA complexed to EcoRV can be used to predict the conformation of the single stranded RNA bound to the active site of PA_N. Therefore, during the refinement of the PA_N–RNA complex, the 3' end of RNA was pulled first onto region 1 and then onto region 2, followed by pulling of the 5' end onto region 3, using steered MD (Supporting Information). Unrestrained MD simulations were then performed on the resulting models to check the validity of the proposed binding mode³¹ (Supporting Information).

The two-Mg²⁺ and one-Mg²⁺ PA_N–RNA complexes obtained in this work are shown in Figure 1C,D. Though the time scale of our simulations is limited, the overall RNA binding mode assayed here is observed to be stable, suggesting that the proposal by Zhao et al. used as a pathfinder in the steered MD refinement is correct. Moreover, the subsequent unrestrained MD simulations reveal further details about the specific interactions between RNA and PA_N. For both complexes region 1 of PA_N is unable to maintain a tight RNA binding. For example, for run 1 of the two-Mg²⁺ complex the 3' end of RNA moved away from the α 3- α 7 cleft after only a few nanoseconds of unrestrained MD, as shown by the rapid increase of the RMSD of the RNA nucleotides compared to that of the initial complex model (Figure S6, bottom, Supporting Information). The lack of binding to helix α 3 can be explained by the presence of two nearby glutamates (E26 and E31) on the α 2- α 3 loop that compensate the positive charge of K34 and thus decrease the attractive force between region 1 and RNA. In contrast, the binding of the 3' end of RNA to region 2 was stable during the 100 ns MD simulation due to the high positive charge of this region, which contains four arginines. Namely, R124 and R125 (on the α 5- β 3 loop) together with R192 (on helix α 7) are found to be the key residues for RNA binding, forming stable salt bridges with RNA during the whole simulation, whereas the contribution of R196 (on helix α 7) appears to be weaker. Most likely R196 is not able to establish strong stable interactions with the RNA because of its flexibility (it is the C-terminal residue of PA_N) and the neutralization of its positive charge by the nearby E195. Therefore, our simulations indicate that R124, R125, and R192, located in region 2, are the main residues responsible for the tight binding of the 3' end of the RNA onto the C-terminus of PA_N.

It is noteworthy that the 3' RNA tail displays a different behavior in the two- and one-Mg²⁺ complexes. The 3' end of the RNA strand is more flexible in the two-Mg than in the one-Mg complex, as evidenced by the larger fluctuations of the RMSD values of U9 and U10 in the two-Mg complex (Figure S6, bottom, Supporting Information) compared to the one-Mg complex (Figure S7, bottom, Supporting Information). The rigidity of the 3' tail in the one-Mg complex is due to a more stable interaction of the phosphodiester group of U9 with the arginine residues in region 2, which in turn is the result of the RNA strand being slightly displaced along the Mg1–Mg2 direction upon loss of the Mg1 ion (see below). Therefore, it seems that, in the one-Mg²⁺ complex, the enzyme strengthens the interaction between the 3' RNA tail and the protein surface,

to compensate for the weaker binding of the RNA central part to the endonuclease active site.

On the other hand, binding of the 5' end of RNA onto region 3 turned out to be much more flexible than the 3' tail. In the initial two-Mg²⁺ complex obtained from the steered MD simulations, the first three residues (A1, C2, and U3) were interacting with region 3 ($t = 0$ ns in Figure S6, top, Supporting Information); however, the RMSD increase after a few nanoseconds indicates that three nucleotides moved away from region 3. It seems that the attractive electrostatic forces from the arginines in region 3 are not strong enough to keep the 5' end of RNA substrate bound tightly onto region 3. Although it is difficult to predict the precise effect of the protein–RNA contacts on the endonuclease reaction, it is tempting to suggest that the flexibility of the 5' end may help to release the product RNA after the phosphodiester cleavage. Interestingly, R84 and the nearby W88 (both located at helix α 4) are found to interact with the phosphodiester groups U3 and C4 nucleotides, contributing to bend the RNA strand near the cleavage site (see below).

Besides the binding of the RNA tails on the protein surface, we also checked the interactions of the central part of the RNA chain with the active site. The central nucleotides (G5, C6, U7) are found to be the most stable part of the RNA oligomer (Figures S6 and S7, middle, Supporting Information), maintaining their interactions with the active site in all three 100 ns simulation runs. In the two-Mg²⁺ PA_N–RNA complex, the scissile phosphodiester of the nucleic acid (O2P) is accommodated between the two metal ions, and the O3' atom of the sugar is coordinated apically to Mg2, as proposed for the two-metal mediated nucleic acid cleavage in EcoRV, another type II endonuclease.^{36,79} Besides, the scissile phosphodiester is forming a salt bridge with K134, which has been proposed to be the catalytic lysine based on structural²¹ and mutagenesis data.²⁵ On the other hand, the guanine and cytosine bases of the G5 and C6 nucleotides forming the cleavage site are sandwich π stacked against Y24 (Figure 1C), further contributing to maintain the stability of the active site of the PA_N–RNA complex. Differently from the two-Mg²⁺ complex, in the one-Mg²⁺ PA_N–RNA complex the single Mg ion is only coordinated to the scissile phosphodiester, similar to the one-metal ion mechanism proposed in EcoRI, also a type II endonuclease.^{80,81} This new metal coordination upon loss of Mg1 results in the RNA strand moving ~ 2.0 Å along the Mg1–Mg2 direction compared to the case for the two-Mg complex. As a consequence, only the C6 base is π – π interacting with Y24 (Figure 1D). The scissile phosphodiester is again forming a salt bridge with a lysine residue, but in the one-Mg complex it is K137 instead of K134. More details about the endonuclease active site structure in the two complexes are given in the next subsection.

Bending of the complexed nucleic acid at the active site is critical for the catalytic activity of other type II endonucleases, such as EcoRV.^{47,48,82} As a measure of the bending of the single-stranded RNA, we have used the angle formed by the phosphorus atom of three consecutive nucleotides, G5, C6, and U7, similar to the helical roll angle of the double-stranded DNA characterizing substrate bending in EcoRV.^{47,48} The RNA bend angle is $\sim 123 \pm 11^\circ$ in the two-Mg²⁺ complex and $\sim 107 \pm 9^\circ$ in the one-Mg²⁺ complex. The smaller angle in the one-Mg complex indicates that the loss of Mg1 induces a more severe RNA bending. Unlike in the two-Mg complex, where only one lysine of the protein (K134) is forming a salt bridge with the

cleavage site (i.e., with the scissile phosphodiester between C5 and G6), in the one-Mg complex both K137 and K134 interact with the central part of the RNA strand (i.e., with the scissile phosphodiester and the adjacent 3' phosphodiester group of U7, respectively), to compensate for the missing metal positive charge, resulting in a smaller P@G5–P@C6–P@U7 angle. Therefore, K134 and K137 have two roles in the endonuclease activity: helping the complexed RNA to achieve the bent conformation needed for cleavage, and participating in the endonuclease chemical reaction (see below).

PA_N has been found to cleave RNA specifically at the 3' side of a guanine base.⁸³ Because the substrate we used to model the PA_N–RNA complex has a central GC cleavage, we have also checked whether these two bases can form any interaction with the protein that can explain this specificity at a molecular level. In the two-Mg²⁺ complex, G5 forms two H-bonds with the protein. On one hand, the amino group (N2) interacts either with the backbone carbonyl of A20 in run 1 (2.96 ± 0.24 Å) or with the carboxylate group of E23 in runs 1–2 (2.89 ± 0.17 Å and 3.13 ± 0.22 Å, respectively). On the other, the carbonyl group (O6) interacts with the guanidinium group of R84 in all three runs (3.24 ± 0.38 Å). The observed interactions, although tentative due to the enhanced flexibility of the single-stranded RNA in our model devoid of the other polymerase domains, are in line with the general model proposed for single-stranded RNA recognition,^{84–86} in which one or two hydrogen bonds can be enough to achieve specificity and, in particular, guanine discrimination is mostly determined by the O6, N7, and N2 atoms interacting with backbone carbonyl groups or arginine residues.^{84,86} Additionally, the amino group of C6 (N4) may form a transient H-bond with E26 in runs 1 and 3 (i.e., the carboxylate of E26 is oscillating between C6 and K34). In contrast, in the one-Mg²⁺ complex these specific H-bonds are lost. Instead, G5 is π stacked against the guanidinium group of R84, whereas C6 now forms a H-bond with either K137 in runs 1–2 (through its O2 carbonyl group) or E26 in run 3 (through its N4 amino group). Hence, one could imagine that the guanine specificity is not so stringent at low Mg concentrations. Loss of sequence discrimination at low Mg concentrations has also been observed for EcoRV.⁸⁷

Overall, our results show that the strong interaction between the 3' end of the RNA and region 2 (R124, R125, and R192) is the main factor determining the RNA direction on the protein surface and keeping the RNA bound, whereas the interaction between the 5' end of RNA with region 3 is weaker, probably to facilitate product release. In addition, the central nucleotides of the RNA strand bind to the active site through the metal ion(s), K134 (in the two-Mg complex), or K134 and K137 (in the one-Mg complex) in a bent conformation that promotes catalysis. Nevertheless, it remains to be seen whether this RNA binding mode changes for the full-length viral RNA and the intact polymerase. Additional interactions between the RNA flanking regions and the other domains of the polymerase might influence the structure of the complex and the endonuclease active site.

Protein Conformational Changes upon RNA Binding. To get insight into the possible RNA-induced conformational changes of the protein, we analyzed the RMSD of the protein backbone in the complexes, with respect to the two-Mg²⁺ apo²¹ and the one-Mg²⁺ apo²² crystal structures, respectively. The protein backbone RMSD as a function of time is shown in Figure S2 (Supporting Information). After equilibration, the RMSD remains stable during the 100 ns MD simulation, with

an average RMSD of 2.07–2.50 Å for the two-Mg²⁺ complex and 1.65–1.92 Å for the one-Mg²⁺ complex (Table S5 and Figure S2, Supporting Information). The small fluctuations about the converged RMSD (~0.2 Å) for both complexes indicate that the obtained PA_N–RNA models are stable within the submicrosecond time scale of our simulations and most likely represent realistic models of the actual RNA binding mode.

The protein backbone RMSD can be decomposed into the secondary structure elements of PA_N (Table S5, Supporting Information). The most significant changes in RMSD occur in helices α 4 and α 7, where RMSD values are ~0.2–0.4 Å larger in the complexes than in the apo systems, and in helix α 5, whose RMSD value decreases by ~1 Å upon RNA binding. As discussed in the previous subsection, the 5' end of the RNA substrate interacts with helix α 4 and the 3' end binds to helix α 7; these RNA–protein interactions induce some structural adjustment in these two helices, explaining their larger RMSD values compared to those of other secondary structure elements. In contrast, binding of the RNA substrate involves the formation of salt bridges between the central part of RNA (bound to the active site) and K134 and K137 in helix α 5, stabilizing the conformation of this helix and precluding the helix displacement away from the active site observed in the apo simulations. Because K134 and K137 are involved in catalysis (see below), stabilization of helix α 5 upon RNA binding is crucial for the endonuclease reaction.

Protein flexibility can also be expressed in terms of per-residue RMSD (Figure S3, Supporting Information). In the apo state (Figure S3, top, Supporting Information), loss of Mg1 increases the flexibility of PA_N. In contrast, in the RNA-bound state (Figure S3, bottom, Supporting Information) the protein becomes more rigid in the absence of Mg1. As mentioned above, the weaker binding to the active site due to the loss of Mg1 seems to be compensated by stronger interactions between PA_N and the RNA tails, resulting in a more rigid protein in the one-Mg complex.

Finally, we investigated further the flexibility of the active site using principal component analysis (PCA)^{88,89} (Figure S5, Supporting Information). The eigenvalues of the two-Mg²⁺ bound PA_N model are lower than the one-Mg²⁺ bound PA_N model, and a similar behavior can also be seen for the PA_N–RNA complexes. The increase in the active site rigidity upon binding of Mg1 reflects the decrease in the conformational fluctuations of E119 and H41 (see above) and is in line with the increase in the melting temperature of PA_N in the presence of metals.²⁵ In addition, both complexes show lower eigenvalues than their apo counterparts, indicating that the active site becomes more rigid upon RNA binding. Stabilization of the protein active site by the RNA substrate is in line with the higher melting temperatures observed in the presence of DPBA,²⁵ an inhibitor that binds to the metal-containing endonuclease active site in a manner similar to RNA.^{23,24}

Endonuclease Active Site Structure. Two-Mg²⁺ PA_N–RNA Complex. Figure 3A shows the endonuclease active site structure obtained for the two-Mg²⁺ PA_N–RNA complex. In contrast to the two-Mg²⁺ holo form, the coordination shells of the metal ions in the complex (Table 1) are well preserved during the MD simulation, and the fluctuations of the metal–ligand distances are small, in agreement with the previous suggestion that the presence of the RNA substrate can stabilize a second metal ion in the active site.^{25,30} In detail, Mg2 is coordinated by E80, D108, two water molecules (W1 and W2),

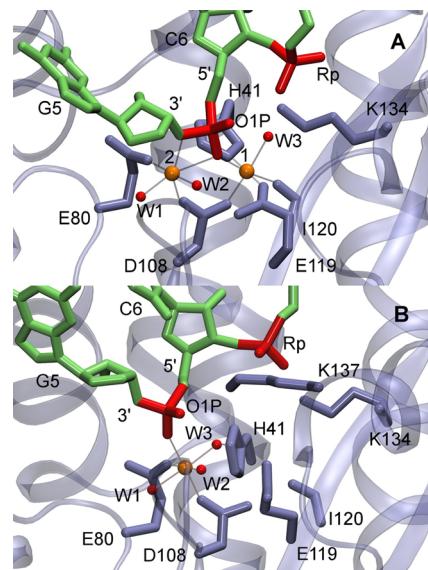


Figure 3. Active site structure of the two-Mg²⁺ (A) and one-Mg²⁺ (B) PA_N–RNA complexes. The key protein residues are shown as light blue licorice and the RNA strand is in green, except for the scissile phosphodiester and the adjacent 3' phosphodiester in red. The Mg ions are displayed in orange, and the water molecules coordinated to the metal ions are in red.

Table 1. Active Site Structural Parameters for the Two-Mg²⁺ PA_N–RNA Complex Simulations in Comparison with the X-ray Structure of Two-Mn²⁺ PA_N in Complex with a Nucleoside Monophosphate, NMP (PDB code 4AWH)²³

distance (Å)	simulation RNA complex ^a	X-ray NMP complex ^b
Mg1–Mg2	3.85(0.12)	3.82, 3.81, 3.80, 3.76
Mg1–W3 ^c	2.02(0.06)	2.09, 2.29, 2.45, 2.27
Mg1–H41	2.18(0.09)	2.31, 2.31, 2.36, 2.27
Mg1–D108	1.89(0.05)	2.26, 2.20, 2.13, 2.10
Mg1–E119	1.89(0.05)	2.20, 2.14, 2.17, 2.15
Mg1–I120	1.98(0.06)	2.05, 2.09, 2.11, 2.02
Mg1–C6 ^d	1.99(0.07)	2.25, 2.26, 2.30, 2.36
Mg2–W1	2.00(0.06)	2.19, 2.37, 2.40, 2.30
Mg2–W2	1.99(0.05)	2.28, 2.27, 2.12, 2.29
Mg2–E80	1.88(0.04)	2.19, 2.14, 2.35, 2.43
Mg2–D108	1.86(0.04)	2.14, 2.05, 2.23, 2.22
Mg2–G5 ^e	2.11(0.09)	2.09, 2.09, 2.14, 2.13
Mg2–C6 ^f	2.30(0.21)	2.15, 2.21, 2.22, 2.13

^aEnsemble average distances (with the standard deviation in parentheses) of the three runs. ^bValues corresponding to each of the four chains present in the asymmetric unit of the crystal. ^cMg1–W3 distance in the RNA complex or Mg1–phosphate oxygen distance in the NMP complex (Figure 4A). ^dMg1–O2P distance, where O2P is a phosphate oxygen atom of C6 (RNA complex) or NMP (NMP complex). ^eMg2–O3' distance in the RNA complex or Mg2–water distance in the NMP complex. ^fMg2–O2P distance, where O2P is a phosphate oxygen atom of C6 (RNA complex) or NMP (NMP complex).

and the bridging (O3') and nonbridging oxygen (O2P) atoms of the scissile phosphodiester of the RNA strand. The Mg2 coordination geometry deviates slightly from an ideal octahedral configuration with average Mg2–O distances of 1.86 Å (D108), 1.88 Å (E80), 2.00 Å (W1), 1.99 Å (W2), 2.11 Å (O3'), and 2.30 Å (O2P) (Table 1 and Figure S13, Supporting Information). On the other hand, Mg1 is

coordinated by H41, D108, E119, the carbonyl oxygen of I120, the O2P atom of the RNA strand, and one water molecule (W3). The Mg1 coordination geometry is close to an ideal octahedron, with average Mg1–ligand distances of 1.89 Å (D108 and E119), 1.99 Å (O2P), 2.02 Å (W3), 2.18 Å (H41), and 1.98 Å (I120) (Table 1 and Figure S14, Supporting Information). The Mg1–Ne@H41 distance is slightly larger than the Mg1–O distances because Mg coordination to histidines is not as favorable as to acidic residues.^{26–29,76} Compared to the two-Mg²⁺ apo simulation (Figure 2A,B), the coordination sphere of Mg1 is much more stable in the complex, in line with previous proposals that RNA binding stabilizes a second metal ion in the active site.^{25,30} Besides, the Mg1–Mg2 distance decreases from 4.91 to 3.85 Å (Table S6, Supporting Information, and Table 1), due to the presence of the bridging scissile phosphodiester that counteracts the repulsion between the two cations. Interestingly, the water molecule coordinated to Mg1 that is to act as the nucleophile group during the endonuclease reaction (W3) forms a hydrogen bond with the pro-Rp oxygen atom (O@W3–O1P@C6 distance = 3.36 ± 0.26 Å, see Figure S15, Supporting Information), and hence it is well positioned to attack the scissile phosphodiester. Further details about the role of the active site residues in the endonuclease reaction are given in the Discussion.

Recently, PA_N has been crystallized bound to a mononucleotide (UMP)²³ with two Mn ions in the active site bridged by one of the oxygen atoms of the UMP phosphate group. The metal coordination spheres in the crystal structure are very similar to the ones found here for the two-Mg²⁺ PA_N–RNA complex (Table 1 and Figure 4A). The small differences between the two structures can be easily explained by taking into account that our model represents the reactants, whereas the crystal structure mimics the products of the two-metal mediated endonuclease reaction. The nucleophilic water W3 coordinated by Mg1 in our complex is replaced in the X-ray structure by a phosphate oxygen atom mimicking the 5' protonated phosphate oxygen resulting from the nucleophile attack. Additionally, the ribose O3' atom coordinated by Mg2 is substituted by a water molecule, equivalent to the leaving 3'-hydroxyl group. Moreover, both our model and the crystal structure show the nitrogenous base of the nucleotide stacked on Y24, and K134 H-bonded to the phosphodiester group. Therefore, the two-Mg²⁺ PA_N–RNA complex obtained here seems to be a good candidate for the reactive conformation of the two-metal mediated endonuclease reaction.

One-Mg²⁺ PA_N–RNA Complex. Comparison of the two-Mg²⁺ (Figure 3A) and the one-Mg²⁺ (Figure 3B) PA_N–RNA complexes shows that removal of the Mg1 ion results in the occupancy of site 1 by a water molecule and a ~0.5 Å displacement of Mg2 toward site 1. The RNA strand concomitantly moves ~2.0 Å toward Mg2 along the Mg1–Mg2 axis, such that the O2P atom of the C6 phosphodiester group is only coordinated by the single Mg ion and the O3' atom is no longer bound to Mg2. Namely, Mg2 is coordinated by E80 and D108 in a monodentate way, as well as the O2P atom of the scissile phosphodiester group, and three water molecules (W1, W2, and W3 in Figure 3B). The corresponding Mg²⁺–ligand bond lengths (Table 2 and Figure S11, Supporting Information) are similar to the average Mg–ligand distances observed in the Protein Data Bank.^{26,76} Interestingly, the nucleophilic water W3 is not H-bonded to the pro-Rp oxygen atom of the RNA strand directly, as in the two-Mg²⁺

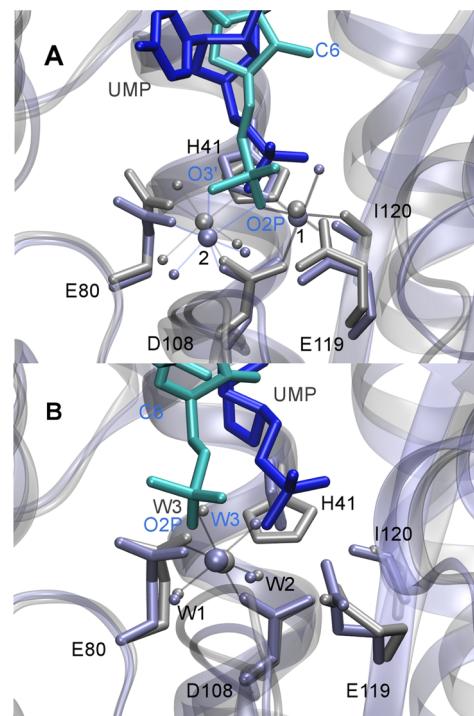


Figure 4. Comparison of the PA_N–RNA complexes obtained in this work with the available crystal structures of PA_N in complex with a nucleoside monophosphate (NMP). (A) Superposition of the simulated two-Mg²⁺ PA_N–RNA complex with the experimental two-Mn²⁺ PA_N–NMP (PDB code 4AWH).²³ (B) Superposition of the simulated one-Mg²⁺ PA_N–RNA complex with the experimental one-Mg²⁺ PA_N–NMP (PDB code 3HW3).³¹ Only the residues coordinated to the metal ion are shown, in light blue for PA_N–RNA and silver for PA_N–NMP. The C6 ribonucleoside is colored in cyan, and the NMP is in dark blue.

complex, but through another water molecule, and this water chain is broken and re-formed several times during the 100 ns MD simulation. Therefore, the pro-Rp oxygen is not very efficient at properly orienting W3 to attack the phosphorus atom of the scissile phosphodiester. Instead, H41 may play this role because this residue is also connected to W3, with a Ne@H41–O@W3 distance of ~3.34 ± 0.35 Å (Table 2).

The active site structure of the one-Mg²⁺ PA_N complex obtained here is slightly different from the crystallographic structure with one Mg²⁺ ion in complex with NMP³¹ (Figure 4B and Table 2). In both structures, there is a third water molecule coordinated to the Mg ion, but the position of W3 and the phosphate group are interchanged. In our PA_N–RNA structure, W3 occupies one of the equatorial positions of the Mg2 coordination shell and the phosphodiester group is coordinated axially, on the opposite side of D108. In contrast, in the PA_N–NMP complex, W3 is the axial ligand and the phosphate group occupies a position similar to that of the nucleophilic water in the PA_N–RNA model. This interchange is due to the formation of a salt bridge between O2P of the NMP and the ζ -amino group of K134 in the X-ray structure. In our model K134 is interacting with the adjacent 3' phosphodiester group, whereas the scissile phosphodiester is forming a salt bridge with K137. The difference between the two structures may reflect a different conformation of the nucleic acid in the reactants (PA_N–RNA complex) and the products (PA_N–NMP complex) of the one-metal mediated endonuclease reaction. Indeed, the authors of the PA_N–NMP crystallographic study³¹

Table 2. Active Site Structural Parameters for the One-Mg²⁺ PA_N–RNA Complex Simulations in Comparison with the X-ray Structure of One-Mg²⁺ PA_N in Complex with a Nucleoside Monophosphate, NMP (PDB code 3HW3)³¹

distance (Å)	simulation RNA complex ^a	X-ray NMP complex
Mg–W1	2.07(0.07)	2.03
Mg–W2	2.09(0.09)	2.12
Mg–W3 ^b	2.08(0.10)	2.15
Mg–E80	1.92(0.05)	2.15
Mg–D108	1.92(0.05)	1.92
Mg–C6 ^c	1.93(0.06)	2.02
Mg–E119	4.58(0.39)	4.35
Mg–H41	5.38(0.95)	4.46
Mg–I120	6.84(0.19)	5.81
E119–W2	2.79(0.25)	2.71
H41–W3 ^d	3.34(0.35)	3.35
I120–W3 ^e	4.93(0.30)	4.36

^aEnsemble average distances (with the standard deviation in parentheses) of the three runs. ^bNote that the position of the W3 water molecule and the phosphate coordinated to Mg are interchanged in the RNA and NMP complexes (Figure 4B). ^cMg–O2P distance, where O2P is a phosphate oxygen atom of C6 (RNA complex) or NMP (NMP complex). ^dH41–W3 distance in the RNA complex or H41–phosphate oxygen distance in the NMP complex. The MD ensemble average is calculated by excluding the part of the trajectories of runs 2–3 where an additional water molecule is mediating the interaction between H41 and W3. ^eI120–W3 distance in the RNA complex or I120–phosphate oxygen distance in the NMP complex.

have already pointed out that the conformation of the bound NMP seems to be rotated compared to the position of the substrate in the homologous EcoRV–DNA product complex. Moreover, the conformation of NMP in the one-Mg²⁺ crystal structure³¹ is different from the X-ray structure of two-Mn²⁺ PA_N in complex with the same nucleotide.²³ Indeed, the authors of the latter structure have suggested that the structure by Zhao et al. represents an active site configuration that has not fully adapted to the ligand, because it was obtained by soaking pregrown PA_N crystals in an NMP solution.²³ Therefore, our one-Mg²⁺ PA_N complex seems to be a reasonable model for a possible one-metal endonuclease catalysis.

■ DISCUSSION

Number of Metal Ions in the Active Site. The endonuclease reaction of PA_N has been proposed to follow a two-metal ion mechanism,³⁰ on the basis of the cooperative effect of metal ion binding on the RNA cleavage activity. Nevertheless, the crystal structures of holo PA_N only show one Mg occupying site 2, with site 1 either empty²² or containing a Mn ion^{23,24} when MnCl₂ concentrations higher than physiological conditions are used. Binding of a second metal ion has been proposed to be stabilized by the presence of the RNA substrate.^{23,30} To investigate the stability of the Mg ions bound in the active pocket, we have analyzed the MD trajectories of PA_N in the holo form and in complex with RNA in terms of the calculated B-factor of the metal ions (Table S8, Supporting Information).

For holo PA_N, Mg1 shows a higher B-factor (\sim 13.8 Å²) than Mg2 (\sim 11.2 Å²), and the flexibility of Mg2 is not affected by the presence or absence of Mg1 (Table S8, Supporting Information). The larger fluctuations of Mg1 around its time-averaged position suggest that in the holo state either Mg1

dissociates easily from the active site or PA_N binds Mg only at site 2. This is in line with Mg not being captured at site 1 in the holo PA_N crystal structures,^{22–24} and with the low affinity of site 1 for Mg.²⁵ Therefore, both the crystallographic and simulation data show the same trend, higher flexibility for Mg bound at site 1 compared to site 2, indicating a weaker binding for Mg1 than Mg2.

Upon RNA binding, the B-factor of both metal ions decreases, indicating that the stability of both Mg2 and Mg1 improves in the presence of the substrate. Therefore, our simulations confirm previous proposals that in the PA_N–RNA complex the active site can be occupied by two metal ions.^{25,30} This is further reinforced by the change in the Mg2–Mg1 distance. For the two-Mg²⁺ bound PA_N system, the average Mg1–Mg2 distance is 4.91 Å (Table S6, Supporting Information), whereas in the PA_N–RNA complex it decreases to 3.85 Å (Table 1). The bridging phosphodiester group alleviates the electrostatic repulsion between the two positively charged cations, resulting in a shorter Mg2–Mg1 distance and thus a more stable metal binding. Binding of a second metal ion in the presence of the nucleic acid substrate is also observed for other endonucleases. RNase H shows only one Mg²⁺ in the apo state (unless Mg concentrations much higher than the physiologically relevant are used⁹⁰), but it contains two Mg²⁺ ions when complexed with its RNA:DNA hybrid substrate, even at a Mg concentrations below 2.5 mM.⁹¹ Accordingly, binding of two metal ions to BamHI (another type II endonuclease) occurs only in the presence of the cognate DNA and the second metal ion has been shown to be required to strengthen the nuclease–substrate interaction.^{92–94} Therefore, both experimental and computational data indicate a close relationship between binding of the nucleic acid substrate and a second metal ion. Nevertheless, our simulations cannot discriminate whether RNA binding stabilizes the two-Mg holo form (i.e., conformational selection) or that binding of RNA and the second metal ion is a cooperative process. Binding of the second metal ion and RNA may be indeed cooperative, because the cleavage pattern observed with Mg and Mn is different.⁶¹ RNA in solution is bound to divalent metal ions, which upon binding to the enzyme are replaced by protein residues or shifted to the endonuclease active center. These metal ions bound to the substrate in solution act as sequence-specific markers (by modulating the structure or electrostatic properties of the nucleic acid). As a consequence, different metals can yield different substrate selectivities.⁹⁵ For instance, EcoRV shows specific cleavage in the presence of Mg, but a less strict sequence specificity with Mn.⁹⁶

The combination of two unequal metal binding sites, 1 and 2, in the active site of the PA_N–RNA complex is similar to other endonucleases, such as RNaseH and EcoRV.^{52,59,91,97} Similarly, we speculate that in PA_N the Mg²⁺ ion bound to site 1 would help to activate the nucleophile water molecule by reducing its pK_a. Indeed, the change in the pH dependency of the PA_N endonuclease activity with the type of divalent metal supports the idea that the metal ion could be involved in the activation of the nucleophilic water.³⁰ Moreover, Mg1 may assist the nucleophile attack by moving toward Mg2 and bringing the nucleophile close to the scissile phosphodiester. On the other hand, the Mg²⁺ ion bound to site 2 may contribute to increase the efficiency of the catalytic process. By analogy with the two-metal ion mechanism proposed for EcoRV and RNase H,^{52,59,91,97} Mg2 in PA_N may stabilize the pentacovalent intermediate formed during the phosphodiester bond cleavage

and likely destabilize the substrate–enzyme complex. More quantitative investigations of the catalytic roles of the two metal ions need to be assessed by QM/MM simulations.

It has been suggested that endonucleases containing a basic residue in the active site, such as H41 in PA_N, may only require one metal ion to cleave the nucleic acid substrate, with the second metal ion only slightly increasing the catalytic efficiency.^{81,98} Moreover, the high dissociation constants of Mg bound to holo PA_N²⁵ compared to the 1–5 mM Mg cellular concentration suggest that in vivo both the two- and the one-Mg²⁺ PA_N–RNA complexes could coexist in equilibrium. In this putative one-metal ion mechanism of PA_N, the role of the single Mg would be to decrease the pK_a of the bound nucleophilic water molecule and to stabilize the transition state of the nucleophilic attack, by analogy with EcoRI, another type II endonuclease that operates via a one-metal mechanism.⁸⁰ Nevertheless, one-metal mediated catalysis is at odds with the Hill coefficient of 2 found for the Mg-mediated endonuclease reaction.³⁰ Further kinetic experiments as well as theoretical investigations are needed to characterize the catalytic efficiency of the two- and one-metal states and thus unravel the actual number of ions required for the endonuclease activity. Provisionally, it is tempting to suggest that the one-metal ion mechanism might at least be operative at low Mg concentrations.

It is also interesting that the endonuclease activity of PA_N increases with the MgCl₂ concentration up to a maximum value around 1 mM and then decreases at higher MgCl₂ concentrations.³⁰ A similar “attenuation” effect has been observed for RNase H^{90,99} and has been attributed to the binding of a third Mg ion in the vicinity of the endonuclease active site, modulating its structure and thus the catalytic efficiency.⁵⁷ Interestingly, isothermal titration calorimetry (ITC) experiments²⁵ have detected binding of two Mg ions to PA_N, with dissociation constants of 148 μM and 4 mM. The high affinity site has been assigned to site 2 and the low affinity site to site 1,^{23,25} on the basis of the crystal structures of holo PA_N not being able to capture Mg bound to site 1.^{22,23} At first sight, a second dissociation constant for Mg of 4 mM, which is in the inhibitory concentration range, might seem to indicate that binding of a second Mg ion to site 1 impairs catalysis, and hence that PA_N operates more efficiently via a one-metal ion mechanism. However, two alternative interpretations are also possible. On one hand, one should consider that the ITC experiments were performed in the absence of the substrate and that RNA binding is expected to stabilize the second metal ion, bringing the K_d of site 1 closer to the Mg concentration optimal for endonuclease activity (i.e., 1 mM). For example, the affinity for Mg of the Klenow exonuclease site increases by a factor of 2 in the presence of the substrate analog TMP.¹⁰⁰ Metal binding ITC experiments for PA_N in the presence of substrate analogs (NMP or DPBA) will be needed to confirm this RNA-mediated enhancement of the Mg²⁺ affinity. On the other hand, one could propose that the attenuation effect at millimolar Mg concentrations is due to the binding of a third Mg ion, similar to RNase H,⁵⁷ and thus that the 4 mM K_d corresponds to the affinity of this putative site 3, instead of site 1. In RNase H, site 3 is formed by one of the endonuclease active site residues (D192) and an additional second-shell acidic residue (E188). In PA_N, the acidic residues closer to the endonuclease active site (and thus possible candidates for the second-shell ligand) are either E23 or D83, and interestingly, a sequence alignment using 13 000 sequences of PA_N from influenza A, B, and C²⁴

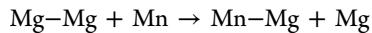
shows that E23 is fully conserved among the three types of influenza viruses and D83 is conserved only in influenza A. The magnesium dependency of the endonuclease activity for the mutants E23A and D83A needs to be measured to confirm the existence of a third Mg binding site in PA_N and identify unambiguously the residue equivalent to RNase H E188.

Notwithstanding, the latter interpretation implies that the first dissociation constant measured for Mg of 148 μM would correspond to site 1, and that Mg would bind to site 2 so tightly that it is not lost during protein purification or dialysis. Interestingly, full-length PA shows endonuclease activity even in the absence of added salt,⁶¹ suggesting that a copurified Mg ion (bound to site 2) might indeed be possible. Nevertheless, dialysis in the presence of excess EDTA followed by ITC measurements are needed to confirm this assumption. Moreover, our alternative interpretation of the ITC data, with the Mg²⁺ K_d values of 148 μM and 4 mM corresponding to binding to sites 1 and 3, respectively, might explain why the binding enthalpy measured for Mg is endothermic.²⁵ Site 1 contains an imidazole ligand (H41) and coordination of Mg to His is not as favorable as for Mn,^{26–29,76} which shows an exothermic binding enthalpy in the equivalent ITC experiments.²⁵ Besides, the orientation of E23 or D83 in the metal-bound PA_N structures^{21–24} is not optimal to coordinate a third Mg close to the endonuclease active site and thus a conformational rearrangement of E23 or D83 would be needed to form site 3, explaining the heat absorption associated with Mg binding. Nevertheless, a copurified Mg ion bound to site 2 seems to contradict the crystal structures of holo PA_N showing two Mn ions in the active site.^{21,24} The coordination of two Mn ions in the structure by Dias and et al.²¹ is likely to be affected by the insertion of E59 from a neighbor molecule and thus is not fully conclusive. Moreover, as Noble and co-workers pointed out, the addition of MnCl₂ could compete out the copurified Mg ion.⁶¹ As for the two Mn ions in the structure by DuBois et al.,²⁴ the assignment of a Mn ion in site 2 needs to be confirmed by anomalous scattering measurements. Indeed, Kowalski and co-workers, who have also crystallized holo PA_N in the presence of both MgCl₂ and MnCl₂, have shown that site 2 has no significant anomalous signal, indicating that site 2 is occupied by Mg. Furthermore, our new interpretation of the ITC data might explain why the species trapped in the latter crystal structure²³ is the “hybrid” species Mn–Mg (bound to site 1 and site 2, respectively), instead of the two-Mn. Within the original interpretation,²⁵ the dissociation constant for Mg bound to site 2 (148 μM) is ~20 times higher than for Mn (6.5 μM), which seems difficult to reconcile with the observation of Mg bound to site 2, taking into account that the crystallization buffer contained identical (2.5 mM) concentrations of both MgCl₂ and MnCl₂.

In this regard, it has been suggested that the catalytically active two-metal ion species is the “hybrid” Mn–Mg (bound to site 1 and site 2, respectively),^{23,25} instead of the two-Mg studied here. This proposal is based on (i) the endonuclease activity with Mn (or a mixture of Mn and Mg) being higher than with Mg,^{30,61} (ii) the crystal structures of holo PA_N not being able to capture Mg bound to site 1,^{22–24} and (iii) the dissociation constants for Mn being lower than for Mg.²⁵ Nonetheless, the bioavailability of the Mg and Mn also needs to be factored in.^{34,100} The cellular Mn concentration is in the ~1 μM range,^{32–35} whereas the optimal Mn concentration for the endonuclease activity is ~100 μM.³⁰ Moreover, the higher activity with MnCl₂ cannot be interpreted directly as a larger

population of the hybrid Mn–Mg species, because the Mn²⁺ ion is intrinsically a better Lewis acid than the Mg²⁺ ion. Even at the same concentration of both two-metal species, the amount of RNA cleaved by the hybrid Mn–Mg species is larger than the two-Mg species, because the pK_a of a water molecule coordinated to Mn is lower than for Mg,³⁰ and thus Mn is more efficient at activating the nucleophile. For instance, QM/MM calculations on the ϵ subunit of DNA polymerase III have shown that the exonuclease reaction barrier is lower with Mn²⁺ than with Mg²⁺.¹⁰¹ As for the crystallographic studies, metal concentrations much higher than physiological were used to populate the metal binding sites in a kinetically reasonable time scale. In particular, the crystallization buffers contained similar concentrations of Mg²⁺ and Mn²⁺ (10 mM Mg(acetate)₂ and 2.5 mM MnCl₂ in ref 21, 10 mM MgCl₂ and 5 mM MnCl₂ in ref 24, and 2.5 mM for both MgCl₂ and MnCl₂ in ref 23), whereas *in vivo* the concentration of Mg is 1000 times higher than for Mn. Finally, although the dissociation constants for Mn (0.3 and 6.5 μ M for sites 1 and 2, respectively) have been proposed to be 4 or 3 orders of magnitude lower than for Mg (4 mM and 148 μ M, according to the original interpretation in ref 25), it should be noted that these K_d values were measured in the absence of the substrate.²⁵ RNA binding is expected to increase the affinity of PA_N for both metals, as discussed above. One could wonder whether the change in the dissociation constants is the same for both Mg²⁺ and Mn²⁺. On one hand, Pearson's HSAB theory¹⁰² seems to indicate that the harder Mg acid binds more tightly to the hard phosphate base than the softer Mn ion, and thus RNA-mediated stabilization is larger for Mg²⁺ than for Mn²⁺. On the other hand, the recent crystal structure of PA_N bound to UMP shows two Mn²⁺ ions in the active site, as confirmed by anomalous scattering measurements. Mn²⁺ ions are probably better able than Mg²⁺ to coordinate the two negatively charged oxygen atoms of the UMP phosphate group, because the transition metal can withdraw more negative charge from the ligands than the alkaline earth metal, allowing the formation of a bidentate complex with the (−2) charged phosphate. Whether the Mn²⁺ preference is maintained when binding the actual RNA substrate, where the scissile phosphodiester only has a single negative charge and forms a monodentate complex, remains to be determined, e.g., by performing ITC measurements of metal binding to PA_N in the presence of a dinucleotide. Altogether, whether the Mn–Mg or the two-Mg species is more abundant *in vivo* still seems to be a matter of debate.

Furthermore, the exchange reaction between the two possible two-metal species



is governed by the equilibrium constant

$$K_{\text{exch}} = \frac{[\text{Mn–Mg}][\text{Mg}]}{[\text{Mg–Mg}][\text{Mn}]}$$

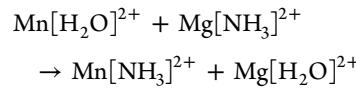
and thus the concentration ratio between the two species is

$$\frac{[\text{Mn–Mg}]}{[\text{Mg–Mg}]} = K_{\text{exch}} \frac{[\text{Mn}]}{[\text{Mg}]}$$

Because the cellular concentration of Mg is 1000 times higher than for Mn,^{32–35} then

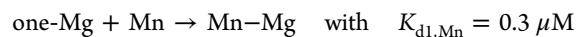
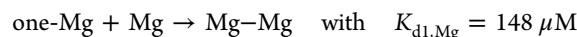
$$\frac{[\text{Mn–Mg}]}{[\text{Mg–Mg}]} = 10^{-3} K_{\text{exch}}$$

and thus the population of the hybrid Mn–Mg species is larger than the Mg–Mg only if $K_{\text{exch}} > 1000$ (or in other words if the exchange free energy is lower than −4.1 kcal/mol at 298 K). Because there is no ITC data to calculate K_{exch} , one could use an alternative, very simplified model of the metal exchange reaction above:



where the ammonia complex represents the divalent metal ion bound to site 1 (containing H41) and the water complex mimics Mg²⁺ or Mn²⁺ in solution. MP2 calculations²⁹ give an exchange free energy for this model of −2.8 kcal/mol, which is of the same order of magnitude as the threshold value of −4.1 kcal/mol. Therefore, despite the approximations of the model and the possible errors in the calculated value, both the Mn–Mg and the Mg–Mg species are likely to coexist *in vivo*.

Finally, we comment on our new interpretation of the ITC data presented above. If a Mg ion bound to site 2 is indeed copurified with PA_N and is not lost during dialysis (i.e., a one-Mg species is titrated in the ITC experiments), the measured first metal dissociation constants correspond to the binding reactions:



The difference between these Mn and Mg binding reactions yields the exchange reaction presented above, and thus the exchange equilibrium constant can be calculated as

$$K_{\text{exch}} = \frac{K_{d1,\text{Mg}}}{K_{d1,\text{Mn}}}$$

Then, the concentration ratio between the two two-metal species is

$$\frac{[\text{Mn–Mg}]}{[\text{Mg–Mg}]} = 10^{-3} \frac{K_{d1,\text{Mg}}}{K_{d1,\text{Mn}}} \sim 0.5$$

indicating that *in vivo* the Mg–Mg species studied here might be 2 times more abundant than the hybrid Mn–Mg trapped in the crystal. Indeed, coexistence of a (major but less active) two-Mg species with a (minor but more active) Mn–Mg species explains why the Mg-dependent endonuclease activity of PA_N is increased synergistically in the presence of low concentrations of MnCl₂ compared to the activity with either Mg or Mn alone, while maintaining the endonuclease inhibition at MgCl₂ concentrations higher than 1 mM.

General Base. The first step in phosphodiester hydrolysis is activation of the water nucleophile.⁸⁰ Similar to many type II restriction endonucleases, PA_N contains a (P)D(X)_N(D/E)XK motif,²² in which the conserved lysine residue is essential for catalysis and is proposed to act as the general base. However, the position of the catalytic lysine in the protein sequence varies among endonucleases, making difficult the assignment of this role to a specific lysine. Mutagenesis experiments on PA_N have shown that K134A is completely inactive for the endonuclease reaction in the presence of both Mg²⁺ and Mn²⁺, pointing to K134 as the catalytic lysine.²⁵ Further support comes from the crystallographic data. In the two-Mn²⁺ holo form of PA_N,²¹ K134 is the lysine closest to the active site (with a N ζ –Mn1 distance of 4.17–5.18 Å) and superimposes well with the

equivalent catalytic lysine of EcoRV (K92). Similarly, the most recent crystal structures of the PA_N holo form with two metal ions in the active site,^{23,24} show K134 H-bonded to the putative nucleophilic water molecule coordinated to Mg1. Nevertheless, in the crystal structure of the one-Mg²⁺ bound PA_N²² the distance between the N ζ atom of K134 and the Mg ion appears to be too long (5.32–6.78 Å) for K134 to stabilize the nucleophile. The authors speculated that either the α 5 helix containing K134 changes position in the heterotrimetric polymerase, bringing the lysine closer to the active site, or H41 can replace K134 as the general base. Considering all these previous proposals, we have investigated the basic residues close to the active site in the PA_N–RNA complexes with two or one Mg²⁺ ions.

In our two-Mg²⁺ PA_N–RNA complex (Figure 3A), the water molecule W3 coordinated by Mg1 is the attacking nucleophile, with a distance from the phosphorus atom of the scissile phosphodiester of 3.18 ± 0.19 Å. The lysine closest to the active site is K134, showing N ζ @K134–Mg1 distances of 5.10 ± 0.42 Å (run 1) and 5.86 ± 0.81 Å (run 3).¹⁰³ Therefore, our structure of the two-Mg²⁺ complex suggests that K134 is the catalytic lysine, in line with previous proposals based on crystallographic and mutagenesis data.^{21,23–25} However, K134 does not form a H-bond with W3 (N ζ @K134–O@W3 = 3.89 ± 0.35 Å in run 1 and 4.30 ± 0.64 Å in run 3), as observed in the crystal structures of the holo form.^{23,24} Instead, K134 interacts with the scissile phosphodiester, with a distance between its N ζ and the nonbridging phosphodiester oxygen (O1P) of 2.86 ± 0.31 Å (run 1) or 3.66 ± 0.67 Å (run 3) (Figure S15, Supporting Information). Therefore, the catalytic role for K134 does not seem to be as a general base, which is in agreement with the PA_N endonuclease activity being observed at pH 6.5, well below the pK_a of the lysine side chain.³⁰ Rather K134 may increase the electrophilicity of the P atom and/or lower the pK_a of the nearby nucleophilic water by neutralizing the negative charge of the phosphodiester group. The different H-bonding of K134 in our reactive complex compared to the crystal structures of the apo state^{21,23,24} suggests the possibility of induced fit conformational changes in the active site during the reaction.¹⁰⁴ Upon RNA binding, the K134–W3 H-bond seems to be replaced by two H-bonds (K134–O1P and W3–proRp oxygen, see below). This may contribute to properly orient both the scissile phosphodiester and the nucleophilic water by decreasing the electrostatic repulsion (and thus the distance) between the scissile and the 3'-adjacent phosphodiester groups.

Instead of K134, an alternative general base may be the phosphodiester group 3' of the scissile bond (i.e., substrate-assisted catalysis). Indeed, in our two-Mg²⁺ complex W3 is connected to the pro-Rp oxygen through a hydrogen bond (O@W3–proRp oxygen = 2.63 ± 0.10 Å in run 1 and 2.69 ± 0.13 Å in run 3). Nevertheless, previous studies on EcoRV^{55,104,105} and RNase H^{52,91} have shown that the essential role of the pro-Rp oxygen in catalysis is most likely not as a general base, but in maintaining the near-attack orientation of W3 and decreasing the pK_a of the nucleophile. Additionally, the pK_a of W3 can be further lowered by Mg1, because the metal positive charge stabilizes the resulting hydroxide anion. As mentioned above, the change in the pH dependency of the endonuclease activity with the type of metal supports that the metal ion could be involved in the activation of the nucleophilic water.³⁰ Altogether, our results on the two-Mg²⁺ PA_N–RNA complex suggest a combined effect of the metal (Mg1), K134,

and the adjacent 3' phosphodiester on lowering the pK_a of W3 and thus on nucleophile activation. This is in line with recent pK_a calculations on several endonucleases containing the (P)D(X)_N(D/E)XK motif,¹⁰⁶ including EcoRV, which showed that both the metal ion and the catalytic Lys contribute to lower the pK_a of the nucleophilic water. Moreover, because W3 is connected to the bulk solvent through a water network, deprotonation of the nucleophilic water could occur by proton transfer to the solution, without the need of a specific general base.

On the other hand, in our one-Mg²⁺ PA_N–RNA complex (Figure 3B), the nucleophilic water molecule W3 is coordinated to Mg2. As in the two-Mg²⁺ complex, K134 is the closest lysine to the active site, with a Mg2–N ζ distance of 6.58 ± 0.89 Å. However, K134 does not form a salt bridge with the scissile phosphodiester, but with the phosphodiester group 3' of the cleavage site (Figure S12, Supporting Information), and its interaction with the nucleophile water W3 is indirect, through a bridging water molecule. Moreover, this water network is very flexible, and replacement of the bridging water by one or more solvent molecules happens several times during the 100 ns time scale of each of the runs. Hence, it is unlikely that K134 acts as the general base, though it may still play an indirect role in nucleophile activation, as previously hypothesized by Yuan et al.²² In view of the structure obtained here, we suggest that the precise effect of K134 is most likely electrostatic, compensating the loss of the positive charge of Mg1 and thus stabilizing the attacking hydroxide.

Compared to K134, K137 (i.e., the other lysine in helix α 5) is slightly further away (6.77 ± 0.80 Å) from the Mg ion,¹⁰⁷ but interestingly it forms a salt bridge with the scissile phosphodiester group, with an average N ζ @K137–P distance of 3.27 ± 0.88 Å (Figure S12, Supporting Information). In other words, the location of K137 in the one-Mg²⁺ complex is equivalent to the position of K134 in the two-Mg²⁺ complex. Therefore, K137 appears to be the catalytic lysine in the one-Mg²⁺ complex. Its role in catalysis may be to increase the electrophilicity of the P atom and/or decrease the pK_a of the water nucleophile through electrostatic effects. This might explain in part why the K137A PA_N mutant is inactive in the presence of MgCl₂.²⁵ According to our simulations, in the one-Mg²⁺ state populated at low Mg concentrations, K137 acts as the catalytic lysine and thus mutation of K137 hinders the endonuclease activity.

Alternatively, it has also been proposed that H41 may act as the general base.²² In our simulation of the one-Mg²⁺ complex, H41 is H-bonded to the nucleophilic water (Figure S12, Supporting Information), and thus H41 not only may contribute to the proper orientation of W3 for the nucleophilic attack but also may be a good candidate for the general base in the one-metal catalytic mechanism. This might explain in part why the Mg²⁺ mediated endonuclease reaction is only observed above pH 6.5.³⁰ Taken together, our results on the one-Mg²⁺ PA_N–RNA complex suggest that at low Mg concentrations K137 (instead of K134) is the lysine involved in catalysis by lowering the pK_a of the nucleophilic water, and H41 is the general base that deprotonates the attacking water molecule.

In summary, our simulations confirm that K134 is the catalytic lysine in the two-metal mediated mechanism, in agreement with crystallographic and mutagenesis data, and suggest that its role in the endonuclease reaction is to decrease the pK_a of the nucleophilic water, along with the metal ion and the phosphodiester group 3' of the scissile bond. Moreover, we

have found that at low Mg concentrations, where the one-metal mediated mechanism may be operative, K134 may be replaced by K137 as the catalytic lysine and H41 may act as the general base, explaining the reduced activity of the K137A PA_N mutant.²⁵

Comparison with Other Endonucleases. We have proposed that the endonuclease reaction in PA_N follows a two-metal ion mechanism, with the nucleophilic water being activated by a combined effect of the Mg1 ion, the adjacent 3' phosphodiester and the catalytic lysine (K134). The two-metal ion mechanism in PA_N seems to be an intermediate case between those of EcoRV and RNaseH,^{52,91,104,105} two other endonucleases that also use two metal ions for catalysis. Both PA_N and EcoRV contain a (P)D(X)_N(D/E)XK motif and use a lysine residue (K134 or K92, respectively) to promote phosphodiester cleavage. However, the position of the catalytic lysine is different in the two endonucleases, either H-bonded to the scissile phosphodiester in PA_N or to the nucleophilic water in EcoRV. Nevertheless, in both cases the catalytic lysine seems to lower the barrier of the nucleophilic attack, either by increasing the electrophilicity of the phosphorus atom in PA_N or by stabilizing the nucleophilic hydroxide anion in EcoRV.^{104,105} On the other hand, both PA_N and RNaseH are endoribonucleases and probably use the phosphodiester group 3' of the scissile bond to assist catalysis. Both ribonucleases show a H-bond between the nucleophilic water and the adjacent 3' phosphodiester group that may help to properly orient the attacking water. The subtle differences between PA_N and the other two endonucleases are probably due to the different nucleic acid substrate (single-stranded RNA for PA_N but double-stranded DNA or an RNA:DNA hybrid for EcoRV and RNaseH, respectively) and the different protein residues coordinated to the two metal ions, specially the additional His in PA_N compared to a water molecule in the other two enzymes.

We have also suggested that a one-metal ion mechanism may be operative in PA_N at low Mg concentrations, with K137 being the catalytic lysine and H41 as the general base. This putative one-metal ion catalysis resembles that of EcoRI,^{80,81} another type II restriction endonuclease that uses a single metal ion for catalysis. In both enzymes the nucleophilic water is H-bonded to the adjacent 3' phosphodiester group; however, the substrate-assisted water deprotonation in EcoRI seems less likely in PA_N due to the presence of a histidine residue in the PA_N active site that can act as a general base.

CONCLUSIONS

In this work, we have studied the interaction between the PA N-terminal domain of the influenza polymerase (PA_N) and the host mRNA in the presence of Mg²⁺ using computational methods. We have obtained a structural model for the PA_N–RNA complex with either one- or two-Mg²⁺ ions, which could likely be used to understand the RNA specificity of PA_N and design more potent and efficient endonuclease inhibitors. We have also proposed a two-metal ion mechanism for the endonuclease reaction, where K134 acts as the catalytic lysine by decreasing the pK_a of the nucleophilic water. Nevertheless, we cannot discard that a one-metal ion mechanism could be operative at low Mg concentrations, with K137 replacing K134 and H41 as the general base. QM/MM simulations are needed to further investigate the mechanism of the endonuclease reaction, and to test the role of the catalytic residues, as suggested here.

ASSOCIATED CONTENT

Supporting Information

Additional computational and analysis details, as well as figures and tables regarding the system setup and the molecular dynamics simulations: Bader-corrected charges; average protein backbone RMSD; active site structural parameters; B factors; comparison with crystal structures; loop conformational analysis; graphs of time evolution of the protein backbone RMSD, the RNA RMSD, Mg2–ligand and Mg1–ligand distances, and active site hydrogen bond network; per-residue Cα RMSD values; principle component analysis. This material is available free of charge via the Internet at <http://pubs.acs.org/>.

AUTHOR INFORMATION

Corresponding Author

*M. Alfonso-Prieto: e-mail, malfonsoprieto@temple.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

S.X. acknowledges the financial support from Chinese Scholarship Council (CSC) and Temple University for his visit to the Institute for Computational Molecular Science at Temple University. S.X. and H.L. are grateful for the financial support provided by the National Natural Science Foundation of China (NNSFC) (Grant Nos. 20934004 and 91127046), the National Basic Research Program of China (NBRPC) (Grant Nos. 2012CB821500 and 2010CB934500), and the “Bairen” fund of the Chinese Academy of Sciences. M.A.-P. acknowledges the financial support from the Government of Catalonia (Commissioner for Universities and Research; Department of Innovation, Universities and Enterprise) and the European Union through a Beatriu de Pinós fellowship (BP-A 2009). The authors thank Dr. Vincenzo Carnevale for insightful discussions and Prof. Allen W. Nicholson for careful reading and comments on the manuscript. The computations were performed on the Kraken (Cray XT5) supercomputer at the National Institute for Computational Sciences (<http://www.nics.tennessee.edu/>) through the XSEDE allocation TG-MCA93S020, on the Eos cluster at the Oak Ridge Leadership Computing Facility through the INCITE allocation CHM04S, and on the Owlsnest HPC cluster at Temple University, which is supported in part by the National Science Foundation through a major research instrumentation grant (CNS-09-58854).

REFERENCES

- (1) Reid, A. H.; Taubenberger, J. K.; Fanning, T. G. The 1918 Spanish Influenza: Integrating History and Biology. *Microbes Infect.* **2001**, *2*, 81–87.
- (2) Horimoto, T.; Kawaoka, Y. Influenza: Lessons from Past Pandemics, Warnings from Current Incidents. *Nat. Rev. Microbiol.* **2005**, *3*, 591–600.
- (3) Noah, D. L.; Krug, R. M. Influenza Virus Virulence and its Molecular Determinants. *Adv. Virus Res.* **2005**, *65*, 121–145.
- (4) Hay, A. J.; Wolstenholme, A. J.; Skehel, J. J.; Smith, M. H. The Molecular Basis of the Specific Anti-Influenza Action of Amantadine. *EMBO J.* **1985**, *4*, 3021–3024.
- (5) Lamb, R. A.; Holsinger, L. J.; Pinto, L. H. The Influenza A Virus M2 Ion Channel Protein and its Role in the Influenza Virus Life Cycle. In *Cellular receptors for animal viruses*; Wimmer, E., Ed.; Cold Spring Harbor Laboratory Press: Woodbury, NY, 1994; pp 303–321.
- (6) Kim, C. U.; Lew, W.; Williams, M. A.; Liu, H.; Zhang, L.; Swaminathan, S.; Bischofberger, N.; Chen, M. S.; Mendel, D. B.; Tai,

- C. Y.; Laver, W. G.; Stevens, R. C. Influenza Neuraminidase Inhibitors Possessing a Novel Hydrophobic Interaction in the Enzyme Active Site: Design, Synthesis, and Structural Analysis of Carbocyclic Sialic Acid Analogues with Potent Anti-Influenza Activity. *J. Am. Chem. Soc.* **1997**, *119*, 681–690.
- (7) von Itzstein, M.; et al. Rational Design of Potent Sialidase-Based Inhibitors of Influenza Virus Replication. *Nature* **1993**, *363*, 418–423.
- (8) Drake, J. W. Rates of Spontaneous Mutation Among RNA Viruses. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 4171–4175.
- (9) Liu, Y.; Lou, Z.; Bartlam, M.; Rao, Z. Structure-Function Studies of the Influenza Virus RNA Polymerase PA Subunit. *Sci. China, Ser. C: Life Sci.* **2009**, *52*, 450–458.
- (10) Krug, R. M.; Aramini, J. M. Emerging Antiviral Targets for Influenza A Virus. *Trends Pharmacol. Sci.* **2009**, *30*, 269–277.
- (11) Das, K.; Aramini, J. M.; Ma, L.-C.; Krug, R. M.; Arnold, E. Structures of Influenza A Proteins and Insights into Antiviral Drug Targets. *Nat. Struct. Mol. Biol.* **2010**, *17*, 530–538.
- (12) Boivin, S.; Cusack, S.; Ruigrok, R. W. H.; Hart, D. J. Influenza A Virus Polymerase: Structural Insights into Replication and Host Adaptation Mechanisms. *J. Biol. Chem.* **2010**, *285*, 28411–28417.
- (13) Ruigrok, R. W.; Crépin, T.; Hart, D. J.; Cusack, S. Towards an Atomic Resolution Understanding of the Influenza Virus Replication Machinery. *Curr. Opin. Struct. Biol.* **2010**, *20*, 104–113.
- (14) Rogolino, D.; Carcelli, M.; Sechi, M.; Neamati, N. Viral Enzymes Containing Magnesium: Metal Binding as a Successful Strategy in Drug Design. *Coord. Chem. Rev.* **2012**, *256*, 3063–3086.
- (15) Plotch, S. J.; Bouloy, M.; Ulmanen, I.; Krug, R. M. A Unique Cap(M^7G P_nX)_m-Dependent Influenza Virion Endonuclease Cleaves Capped RNAs to Generate the Primers that Initiate Viral RNA Transcription. *Cell* **1981**, *23*, 847–858.
- (16) Poon, L. L. M.; Pritchard, D. C.; Fodor, E.; Brownlee, G. G. Direct Evidence that the Poly(A) Tail of Influenza A Virus mRNA is Synthesized by Reiterative Copying of a U Track in the Virion RNA Template. *J. Virol.* **1999**, *73*, 3473–3476.
- (17) Zheng, H.; Lee, H. A.; Palese, P.; Garcia-Sastre, A. Influenza A Virus RNA Polymerase Has the Ability to Stutter at the Polyadenylation Site of a Viral RNA Template During RNA Replication. *J. Virol.* **1999**, *73*, 5240–5243.
- (18) Sugiyama, K.; Obayashi, E.; Kawaguchi, A.; Suzuki, Y.; Tame, J. R. H.; Nagata, K.; Park, S.-Y. Structural Insight into the Essential PB1-PB2 Subunit Contact of the Influenza Virus RNA Polymerase. *EMBO J.* **2009**, *28*, 1803–1811.
- (19) He, X.; Zhou, J.; Bartlam, M.; Zhang, R.; Ma, J.; Lou, Z.; Li, X.; Li, J.; Joachimiak, A.; Zeng, Z.; Ge, R.; Rao, Z.; Liu, Y. Crystal Structure of the Polymerase PA_C-PB1_N Complex from an Avian Influenza H5N1 Virus. *Nature* **2008**, *454*, 1123–1126.
- (20) Obayashi, E.; Yoshida, H.; Kawai, F.; Shibayama, N.; Kawaguchi, A.; Nagata, K.; Tamel, J. R. H.; Park, S.-Y. The Structural Basis for an Essential Subunit Interaction in Influenza Virus RNA Polymerase. *Nature* **2008**, *454*, 1127–1131.
- (21) Dias, A.; Bouvier, D.; Crépin, T.; McCarthy, A. A.; Hart, D. J.; Baudin, F.; Cusack, S.; Ruigrok, R. W. H. The Cap-Snatching Endonuclease of Influenza Virus Polymerase Resides in the PA Subunit. *Nature* **2009**, *458*, 914–918.
- (22) Yuan, P.; Bartlam, M.; Lou, Z.; Chen, S.; Zhou, J.; He, X.; Lv, Z.; Ge, R.; Li, X.; Deng, T.; Fodor, E.; Rao, Z.; Liu, Y. Crystal Structure of an Avian Influenza Polymerase PA_N Reveals an Endonuclease Active Site. *Nature* **2009**, *458*, 909–913.
- (23) Kowalinski, E.; Zubietta, C.; Wolkerstorfer, A.; Szolar, O. H. J.; Ruigrok, R. W.; Cusack, S. Structural Studies of Specific Metal Chelating Inhibitor Binding to the Endonuclease Domain of Influenza PH1N1 (2009) Polymerase. *PLoS Pathog.* **2012**, *8*, e1002831.
- (24) DuBois, R. M.; Slavish, P. J.; Baughman, B. M.; Yun, M.-K.; Bao, J.; Webb, R. J.; Webb, T. R.; White, S. W. Structural and Biochemical Basis for Development of Influenza Virus Inhibitors Targeting the PA Endonuclease. *PLoS Pathog.* **2012**, *8*, e1002830.
- (25) Crépin, T.; Dias, A.; Andrés Palencia, C. S.; Cusack, S.; Ruigrok, R. W. H. Mutational and Metal Binding Analysis of the Endonuclease Domain of the Influenza Virus Polymerase PA Subunit. *J. Virol.* **2010**, *84*, 9096–9104.
- (26) Harding, M. M.; Nowicki, M. W.; Walkinshaw, M. D. Metals in Protein Structures: a Review of Their Principal Features. *Crystallogr. Rev.* **2010**, *16*, 247–302.
- (27) Dudev, T.; Lin, Y.; Dudev, M.; Lim, C. First-Second Shell Interactions in Metal Binding Sites in Proteins: A PDB Survey and DFT/CDM Calculations. *J. Am. Chem. Soc.* **2003**, *125*, 3168–3180.
- (28) Yang, T.-Y.; Dudev, T.; Lim, C. Mononuclear versus Binuclear Metal-Binding Sites: A Metal-Binding Affinity and Selectivity from PDB Survey and DFT/CDM Calculations. *J. Am. Chem. Soc.* **2008**, *130*, 3844–3852.
- (29) Bock, C. W.; Katz, A. K.; Markham, G. D.; Glusker, J. P. Manganese as a Replacement for Magnesium and Zinc: Functional Comparison of the Divalent Ions. *J. Am. Chem. Soc.* **1999**, *121*, 7360–7372.
- (30) Doan, L.; Handa, B.; Roberts, N. A.; Klumpp, K. Metal Ion Catalysis of RNA Cleavage by the Influenza Virus Endonuclease. *Biochemistry* **1999**, *38*, 5612–5619.
- (31) Zhao, C.; Lou, Z.; Guo, Y.; Ma, M.; Chen, Y.; Liang, S.; Zhang, L.; Chen, S.; Li, X.; Liu, Y.; Bartlam, M.; Rao, Z. Nucleoside Monophosphate Complex Structures of the Endonuclease Domain from the Influenza Polymerase PA Subunit Reveal the Substrate Binding Site inside the Catalytic Center. *J. Virol.* **2009**, *83*, 9024–9030.
- (32) Vettori, M. V.; Gatti, R.; Orlandini, G.; Belletti, S.; Alinovi, R.; Smargiassi, A.; Mutti, A. An *in vitro* Model for the Assessment of Manganese Neurotoxicity. *Toxicol. In Vitro* **1999**, *13*, 931–938.
- (33) Dobson, A. W.; Erikson, K. M.; Aschner, M. Manganese Neurotoxicity. *Ann. N. Y. Acad. Sci.* **2004**, *1012*, 115–128.
- (34) Maret, W. Metalloproteomics, Metalloproteomes, and the Annotation of Metalloproteins. *Metalomics* **2010**, *2*, 117–125.
- (35) Jähnen-Dechent, W.; Ketteler, M. Magnesium Basics. *Clin. Kidney J.* **2012**, *5*, i3–i14.
- (36) Dupureur, C. M. Roles of Metal Ions in Nucleases. *Curr. Opin. Chem. Biol.* **2008**, *12*, 250–255.
- (37) MacKerell, A. D.; Nilsson, L. Molecular Dynamics Simulations of Nucleic Acid-Protein Complexes. *Curr. Opin. Struct. Biol.* **2008**, *18*, 194–199.
- (38) Abyzov, A.; Uzun, A.; Strauss, P. R.; Ilyin, V. A. An AP Endonuclease 1-DNA Polymerase β Complex: Theoretical Prediction of Interacting Surfaces. *PLoS Comput. Biol.* **2008**, *4*, e1000066.
- (39) Allawi, H. T.; Kaiser, M. W.; Onufriev, A. V.; Ma, W.-P.; Brogaard, A. E.; Case, D. A.; Neri, B. P.; Lyamichev, V. I. Modeling of Flap Endonuclease Interactions with DNA Substrate. *J. Mol. Biol.* **2003**, *328*, 537–554.
- (40) Swift, R. V.; Durrant, J.; Amaro, R. E.; McCammon, J. A. Toward Understanding the Conformational Dynamics of RNA Ligation. *Biochemistry* **2009**, *48*, 709–719.
- (41) Kormos, B. L.; Baranger, A. M.; Beveridge, D. L. Do Collective Atomic Fluctuations Account for Cooperative Effects? Molecular Dynamics Studies of the U1A-RNA Complex. *J. Am. Chem. Soc.* **2006**, *128*, 8992–8993.
- (42) Ditzler, M. A.; Otyepka, M.; Sponer, J.; Walter, N. G. Molecular Dynamics and Quantum Mechanics of RNA: Conformational and Chemical Change We Can Believe In. *Acc. Chem. Res.* **2010**, *43*, 40–47.
- (43) Formoso, E.; Matxain, J. M.; López, X. Molecular Dynamics Simulation of Bovine Pancreatic Ribonuclease A-CpA and Transition State-Like Complexes. *J. Phys. Chem. B* **2010**, *114*, 7371–7382.
- (44) Feig, M.; Burton, Z. F. RNA Polymerase II with Open and Closed Trigger Loops: Active Site Dynamics and Nucleic Acid Translocation. *Biophys. J.* **2010**, *99*, 2577–2586.
- (45) Palermo, G.; Stenta, M.; Cavalli, A.; Dal-Peraro, M.; De-Vivo, M. Molecular Simulations Highlight the Role of Metals in Catalysis and Inhibition of Type II Topoisomerase. *J. Chem. Theory Comput.* **2013**, *9*, 857–862.
- (46) Guérout, M.; Picot, D.; Abi-Ghanem, J.; Hartmann, B.; Baaden, M. How Cations Can Assist DNase I in DNA Binding and Hydrolysis. *PLoS Comput. Biol.* **2010**, *6*, e1001000.

- (47) Zahran, M.; Daidone, I.; Smith, J. C.; Imhof, P. Mechanism of DNA Recognition by the Restriction Enzyme EcoRV. *J. Mol. Biol.* **2010**, *401*, 415–432.
- (48) Zahran, M.; Bereznak, T.; Imhof, P.; Smith, J. C. Role of Magnesium Ions in DNA Recognition by the EcoRV Restriction Endonuclease. *FEBS Lett.* **2011**, *585*, 2739–2743.
- (49) Chen, C.; Krause, K.; Pettitt, B. M. Advantage of Being a Dimer for *Serratia Marcescens* Endonuclease. *J. Phys. Chem. B* **2009**, *113*, 511–521.
- (50) Davis, B. C.; Thorpe, I. F. Molecular Simulations Illuminate the Role of Regulatory Components of the RNA Polymerase from Hepatitis C Virus in Influencing Protein Structure and Dynamics. *Biochemistry* **2013**, *52*, 4541–4552.
- (51) Fuxreiter, M.; Osman, R. Probing the General Base Catalysis in the First Step of BamHI Action by Computer Simulation. *Biochemistry* **2001**, *40*, 15017–15023.
- (52) 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.
- (53) De-Vivo, M.; Ensing, B.; Dal-Peraro, M.; Gomez, G. A.; Christianson, D. W.; Klein, M. L. Proton Shuttles and Phosphatase Activity in Soluble Epoxide Hydrolase. *J. Am. Chem. Soc.* **2007**, *129*, 387–394.
- (54) 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.
- (55) Imhof, P.; Fischer, S.; Smith, J. C. Catalytic Mechanism of DNA Backbone Cleavage by the Restriction Enzyme EcoRV: a Quantum Mechanical/Molecular Mechanical Analysis. *Biochemistry* **2009**, *48*, 9061–9075.
- (56) Sgrignani, J.; Magistrato, A. The Structural Role of Mg²⁺ Ions in a Class I RNA Polymerase Ribozyme: a Molecular Simulation Study. *J. Phys. Chem. B* **2012**, *116*, 2259–2268.
- (57) 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.
- (58) Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, K. M., Jr. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J. Chem. Theory Comput.* **2013**, *9*, 2733–2748.
- (59) 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.
- (60) Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.
- (61) Noble, E.; Cox, A.; Deval, J.; Kim, B. Endonuclease Substrate Selectivity Characterized with Full-Length PA of Influenza A Virus Polymerase. *Virology* **2012**, *433*, 27–34.
- (62) Horton, N. C.; Perona, J. J. Recognition of Flanking DNA Sequences by EcoRV Endonuclease Involves Alternative Patterns of Water-Mediated Contacts. *J. Biol. Chem.* **1998**, *273*, 21721–21729.
- (63) Deka, P.; Rajan, P. K.; Perez-Canadillas, J. M.; Varani, G. Protein and RNA Dynamics Play Key Roles in Determining the Specific Recognition of GU-Rich Polyadenylation Regulatory Elements by Human Cstf-64 Protein. *J. Mol. Biol.* **2005**, *347*, 719–733.
- (64) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
- (65) 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 α/γ Conformers. *Biophys. J.* **2007**, *92*, 3817–3829.
- (66) 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.
- (67) Bader, R. F. *Atoms In Molecules - A Quantum Theory*; Oxford University Press: Oxford, U.K., 1990.
- (68) 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.
- (69) Aqvist, J. Ion-Water Interaction Potentials Derived from Free Energy Perturbation Simulations. *J. Phys. Chem.* **1990**, *94*, 8021–8024.
- (70) 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*, 1781–1802.
- (71) Ryckaert, J.-P.; Cicotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (72) Jesus A. Izaguirre, S. R.; Skeel, R. D. Longer Time Steps for Molecular Dynamics. *J. Chem. Phys.* **1999**, *110*, 9853–9864.
- (73) 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.
- (74) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, *31*, 1695–1697.
- (75) Parkes, K. E. B.; Ermert, P.; Fässler, J.; Ives, J.; Martin, J. A.; Merrett, J. H.; Obrecht, D.; Williams, G.; Klumpp, K. Use of a Pharmacophore Model to Discover a New Class of Influenza Endonuclease Inhibitors. *J. Med. Chem.* **2003**, *46*, 1153–1164.
- (76) Harding, M. M. Small Revisions to Predicted Distances Around Metal Sites in Proteins. *Acta Crystallogr. D* **2006**, *62*, 678–682.
- (77) Spiegel, K.; DeGrado, W. F.; Klein, M. L. Structural and Dynamical Properties of Manganese Catalase and the Synthetic Protein DF1 and Their Implication for Reactivity from Classical Molecular Dynamics Calculations. *Proteins* **2006**, *65*, 317–330.
- (78) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of Nanosystems: Application to Microtubules and the Ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10037–10041.
- (79) Yang, W. Nucleases: Diversity of Structure, Function and Mechanism. *Q. Rev. Biophys.* **2011**, *44*, 1–93.
- (80) Pingoud, A.; Fuxreiter, M.; Pingoud, V.; Wende, W. Type II Restriction Endonucleases: Structure and Mechanism. *Cell. Mol. Life Sci.* **2005**, *62*, 685–707.
- (81) Dupureur, C. M. One is Enough: Insights into the Two-Metal Ion Nuclease Mechanism from Global Analysis and Computational Studies. *Metalomics* **2010**, *2*, 609–620.
- (82) Horton, N. C.; Perona, J. J. Crystallographic Snapshots along a Protein-Induced DNA-Bending Pathway. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5729–5734.
- (83) Datta, K.; Wolkerstorfer, A.; Szolar, O. H.; Cusack, S.; Klumpp, K. Characterization of PA-N Domain of Influenza A Polymerase Reveals Sequence Specific RNA Cleavage. *Nucleic Acids Res.* **2013**, *41*, 8289–8299.
- (84) Allers, J.; Shamoo, Y. Structure-Based Analysis of Protein-RNA Interactions Using the Program ENTANGLE. *J. Mol. Biol.* **2001**, *311*, 75–86.
- (85) Messias, A. C.; Sattler, M. Structural Basis of Single-Stranded RNA Recognition. *Acc. Chem. Res.* **2004**, *37*, 279–287.
- (86) Morozova, N.; Allers, J.; Myers, J.; Shamoo, Y. Protein-RNA Interactions: Exploring Binding Patterns with a Three-Dimensional Superposition Analysis of High Resolution Structures. *Bioinformatics* **2006**, *22*, 2746–2752.
- (87) Thielking, V.; Selent, U.; Koehler, E.; Landgraf, A.; Wolfe, H.; Alves, J.; Pingoud, A. Mg²⁺ Confers DNA Binding Specificity to the EcoRV Restriction Endonuclease. *Biochemistry* **1992**, *31*, 3727–3732.
- (88) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. Essential Dynamics of Proteins. *Proteins* **1993**, *17*, 412–425.

- (89) Pérez, A.; Blas, J. R.; Rueda, M.; López-Bes, J. M.; de la Cruz, X.; Orozco, M. Exploring the Essential Dynamics of B-DNA. *J. Chem. Theory Comput.* **2005**, *1*, 790–800.
- (90) Goedken, E. R.; Marqusee, S. Co-Crystal of *Escherichia coli* RNase HI with MN²⁺ Ions Reveals Two Divalent Metals Bound in the Active Site. *J. Biol. Chem.* **2001**, *276*, 7266–7271.
- (91) 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.
- (92) Vipond, I. B.; Halford, S. E. Specific DNA Recognition by EcoRV Restriction Endonuclease Induced by Calcium Ions. *Biochemistry* **1995**, *34*, 1113–1119.
- (93) Viadiu, H.; Aggarwal, A. K. Structure of BamHI Bound to Nonspecific DNA: a Model for DNA Sliding. *Mol. Cell* **2000**, *5*, 889–895.
- (94) Lee, J. Y.; Chang, J.; Joseph, N.; Ghirlando, R.; Rao, D. N.; Yang, W. MutH Complexed with Hemi- and Unmethylated DNAs: Coupling Base Recognition and DNA Cleavage. *Mol. Cell* **2005**, *20*, 155–166.
- (95) Solt, I.; Simon, I.; Csaszar, A. G.; Fuxreiter, M. Electrostatic versus Nonelectrostatic Effects in DNA Sequence Discrimination by Divalent Ions Mg²⁺ and Mn²⁺. *J. Phys. Chem. B* **2007**, *111*, 6272–6279.
- (96) Bowen, L. M.; Dupureur, C. M. Investigation of Restriction Enzyme Cofactor Requirements: a Relationship between Metal Ion Properties and Sequence Specificity. *Biochemistry* **2003**, *42*, 12643–12653.
- (97) 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.
- (98) Pingoud, V.; Wende, W.; Friedhoff, P.; Reuter, M.; Alves, J.; Jeltsch, A.; Mones, L.; Fuxreiter, M.; Pingoud, A. On the Divalent Metal Ion Dependence of DNA Cleavage by Restriction Endonucleases of the EcoRI Family. *J. Mol. Biol.* **2009**, *393*, 140–160.
- (99) Keck, J. L.; Goedken, E. R.; Marqusee, S. Activation/Attenuation Model for RNase H: a One-Metal Mechanism with a Second Metal Inhibition. *J. Biol. Chem.* **1998**, *273*, 34128–34133.
- (100) Cowan, J. A. Metal Activation of Enzymes in Nucleic Acid Biochemistry. *Chem. Rev.* **1998**, *98*, 1067–1087.
- (101) Cisneros, G. A.; Perera, L.; Schaaper, R. M.; Pedersen, L. C.; London, R. E.; Pedersen, L. G.; Darden, T. A. Reaction Mechanism of the ϵ Subunit of *E. coli* DNA Polymerase III: Insights into Active Site Metal Coordination and Catalytically Significant Residues. *J. Am. Chem. Soc.* **2009**, *131*, 1550–1556.
- (102) Pearson, R. G. Hard and Soft Acids and Bases. *J. Am. Chem. Soc.* **1963**, *85*, 3533–3539.
- (103) We excluded run 2, because of the presence of additional water molecules, mediating the interaction of K134 with the catalytic residues. Most likely, the absence of the other domains of the polymerase in our model increases the solvent exposure of the active site and enhances the flexibility of the α 5 helix, where K134 is located, leading to the system being trapped in this metastable state.
- (104) Horton, N. C.; Perona, J. J. DNA Cleavage by EcoRV Endonuclease: Two Metal Ions in Three Metal Ion Binding Sites. *Biochemistry* **2004**, *43*, 6841–6857.
- (105) Horton, N. C.; Newberry, K. J.; Perona, J. J. Metal Ion-Mediated Substrate-Assisted Catalysis in Type II Restriction Endonucleases. *Proc. Natl. Acad. Sci.* **1998**, *95*, 13489–13404.
- (106) Xie, F.; Briggs, J. M.; Dupureur, C. M. Nucleophile Activation in PD...D/E/XK Metallonucleases: an Experimental and Computational Study. *J. Inorg. Biochem.* **2010**, *104*, 665–672.
- (107) Hereafter we exclude run 3 from the calculation of the ensemble average distances involving K137, because in this simulation the interaction of K137 with the catalytic residues is mediated by extra water molecules (see reference 103).