Evidence from molecular dynamics simulations of conformational preorganization in the ribonuclease H active site
[version 2; peer review: 2 approved]

Kate A. Stafford, Arthur G. Palmer III
Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA

Abstract
Ribonuclease H1 (RNase H) enzymes are well-conserved endonucleases that are present in all domains of life and are particularly important in the life cycle of retroviruses as domains within reverse transcriptase. Despite extensive study, especially of the E. coli homolog, the interaction of the highly negatively charged active site with catalytically required magnesium ions remains poorly understood. In this work, we describe molecular dynamics simulations of the E. coli homolog in complex with magnesium ions, as well as simulations of other homologs in their apo states. Collectively, these results suggest that the active site is highly rigid in the apo state of all homologs studied and is conformationally preorganized to favor the binding of a magnesium ion. Notably, representatives of bacterial, eukaryotic, and retroviral RNases H all exhibit similar active-site rigidity, suggesting that this dynamic feature is only subtly modulated by amino acid sequence and may primarily be imposed by the distinctive RNase H protein fold.

Corresponding author: Arthur G. Palmer III (agp6@columbia.edu)
Competing interests: No competing interests were disclosed.
Grant information: This work was funded by an NSF graduate research fellowship (KAS) and NIH grant GM50291 (AGP).
Copyright: © 2014 Stafford KA and Palmer III AG. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).
How to cite this article: Stafford KA and Palmer III AG. Evidence from molecular dynamics simulations of conformational preorganization in the ribonuclease H active site [version 2; peer review: 2 approved] F1000Research 2014, 3:67 (https://doi.org/10.12688/f1000research.3605.2)
Amendments from Version 1

Version 2 contains a new Figure 8 showing a structure-based multiple sequence alignment for all RNase H homologs used to initiate simulations as part of this work. The locations of the active site residues and substrate-binding regions are highlighted. Additionally, a comment has been added to the Materials and methods section discussing the choice of pH in this work, the text in the Conclusions section has been edited for clarity, and a new reference has been added to a recent paper on the evolution of the RNase H superfamily.

See referee reports

Introduction

Ribonuclease H1 (RNase H) proteins are well-conserved endonucleases that are found in all domains of life and cleave the RNA strand of an RNA-DNA duplex substrate. The RNase H active site canonically consists of a highly conserved DED(D) motif (Figure 1), three to four carboxylate-containing residues collectively participating in the binding of catalytically required divalent cations, Mg$^{2+}$ under physiological conditions. This active-site sequence motif and requirement for Mg$^{2+}$ is widely shared with other nucleases, suggesting a common catalytic mechanism.

The best-studied member of the RNase H family is the homolog from Escherichia coli (ecRNH), in which this active-site motif is represented as D10, E48, D70, and D134$^\text{1}$ (Figure 2A). Measurements of the pK$a$ values of the active-site residues indicate perturbed pK$a$ values for D10 and D70 which normalize upon Mg$^{2+}$ binding, clearly establishing these residues as critical for interaction with ions. The pH optimum for the RNase H reaction in vitro is approximately 7.5–8.5$^\text{1}$, a value at which all active-site residues should be deprotonated and therefore accessible for ion binding.

Despite extensive study, the interaction of metal ions with the ecRNH active site is poorly understood. Activity has been reported in the presence of Mn$^{2+}$ as well as the physiologically relevant Mg$^{2+}$.$^\text{1}$ Significant differences have been observed between the protein’s interactions with Mg$^{2+}$ and Mn$^{2+}$. Co-crystallization studies of ecRNH with high concentrations of Mg$^{2+}$ find a single bound metal ion (Figure 2B). By contrast, co-crystallization with Mn$^{2+}$ reveals two bound ions, one associated with residues D10 and D134 (denoted the A site), and one associated with D10, E48, and D70 (denoted the B site) (Figure 2C); the B site is similar but not identical to the previously identified Mg$^{2+}$ site. Single Mn$^{2+}$ sites have been identified in the presence of mutations of E48 and/or D134$^\text{7}$, both of which are dispensable for Mn$^{2+}$-dependent activity$^\text{1}$. Crystallographic studies of related RNases H from the archaeal extremophile Bacillus halodurans$^\text{8}$ and from Homo sapiens$^\text{9}$ in complex with substrate find two bound ions in the active site (Figure 2D).

Experimental evidence from nuclear magnetic resonance (NMR) studies locates the area surrounding the active site as the region most susceptible to perturbation upon interaction with ions (Figure 3). Titration of Mg$^{2+}$ with ecRNH, monitored independently by $^1$H and $^{13}$C NMR, suggests that only a single ion binds to the protein in the absence of substrate$^\text{10}$. The identified binding site has relatively weak affinity: $K_b$ has been reported in the micromolar to low millimolar range$^\text{11}$. The second site may be occupied only upon binding of substrate$^\text{11}$, possibly due to the presence of high local concentration in the ion cloud of the highly negatively charged nucleic acid molecule. Conformational changes in the active site upon binding the first ion have been suggested as well, with the second site proposed as being responsible for the attenuation of activity at high ion concentrations$^\text{12}$. Collectively, these results have been used to propose both a one-metal$\text{10,14}$ and a two-metal$\text{13,15,16}$ catalytic mechanism. Computational work using the quantum mechanics/molecular mechanics (QM/MM) method applied to the Bacillus halodurans$\text{17,18}$ and Homo sapiens$\text{8}$ complexes generally supports the two-metal mechanism.

RNase H domains are widely distributed in cellular organisms, but also occur as a component of the reverse transcriptase protein found in retroviruses, in which they are required for viral proliferation$\text{31}$.

Figure 1. Conservation of the ribonuclease H active site. (A) Residue conservation among bacterial RNase H homologs mapped onto the structure of ecRNH (PDB ID 2RN2). Highly conserved residues are shown in red, highly variable residues in green, and sites with insufficient data in yellow. Image produced using ConSurf. (B) Electrostatic map of the solvent-accessible surface of ecRNH produced using APBS. Red represents regions of negative charge and blue represents regions of positive charge. The active-site residues are represented as sticks in both cases.
Figure 2. Conformational diversity of metal-ion interactions with ecRNH as determined by crystallography. In all cases the backbone and active-site sidechains from ecRNH in the absence of ion (PDB ID 2RN2) are shown in light blue for comparison. (A) Structural superposition of the four active-site residues in two structures of ecRNH in the absence of metal ions: 2RN2 (light blue) and 1RNH (dark blue). (B) Structural diversity of RNases H in complex with a single Mg$^{2+}$ ion: ecRNH (1RDD), green; XMRV WT (4E89), dark cyan; XMRV ΔC (3P1G), maroon; MoMLV ΔC (2HB5), purple. The two deletion mutants (indicated as ΔC) bind Mg$^{2+}$ in slightly different positions than do their corresponding full-length proteins and both contain two alternate conformations for E48 and D134. (C) Comparison of Mg$^{2+}$ and Mn$^{2+}$ complexes: ecRNH with Mg$^{2+}$ (1RDD), green; ecRNH with Mn$^{2+}$ (1G15), orange; HIV RNase H domain with ecRNH helix C insertion with Mn$^{2+}$ (3HYF), brown. (D) Structural diversity of RNases H in complex with substrate: Homo sapiens RNase H with Ca$^{2+}$ ions (2QKK), brown; Bacillus halodurans RNase H with Mn$^{2+}$ ions (1ZBI), dark green.

Figure 3. Experimental NMR measurements of the effects of Mg$^{2+}$ binding mapped onto the structure of ecRNH. (A) Active site residues (yellow) and other DENQ residues (green) that experience perturbation upon Mg$^{2+}$ binding. (B) Chemical shift perturbation values for sidechain C$\gamma$δ reflecting the effects of Mg$^{2+}$ binding. White corresponds to no chemical shift change, red corresponds to a large change, and non-DENQ residues are shown in light blue. (C) Residues previously shown to experience backbone $^{15}$N or $^1$H chemical shift changes upon binding Mg$^{2+}$.

For this reason, inhibitors of retroviral RNase H domains, particularly that of HIV, have been widely reported, although none to date have reached clinical use. Most such inhibitors interact with the active site in the metal-bound state and therefore must be selective for retroviral RNase H domains to find clinical utility. The HIV RNase H domain (hivRNH) has been reported to bind two metal ions even in the absence of substrate, although the reason for this difference in behavior between hivRNH and ecRNH is not clear. The Homo sapiens RNase H domain (hsRNH) has not been structurally characterized in the absence of substrate and its binding behavior is less well understood. However, it has higher sequence identity and is more structurally similar to ecRNH than hivRNH. Combined nuclear magnetic resonance (NMR) and molecular dynamics (MD) studies of the behavior of the ecRNH active site residues suggest that the residues of the ecRNH active site are pre-organized in the apo state for the binding of a single Mg$^{2+}$ ion. However, experimental constraints prevent the detailed observation of the protein’s dynamic behavior in the presence of a bound ion at ps-ns timescale. The present work aims to more fully understand the dynamics of ecRNH in the Mg$^{2+}$-bound state through molecular dynamics simulations of ecRNH in the presence of single Mg$^{2+}$ ions in various positions in the active site as suggested by crystallographic studies. In addition, the dynamic behavior of the active site in the apo state is compared with homologs from other organisms.
Methods
For each initial protein structure, protonation states for titratable residues were assigned either by experimental measurement (for ecRNH) or by prediction using the H++ pKa predictor. Unless otherwise specified, all simulations were performed at a pH of 5.5 to recapitulate the conditions used in prior NMR experiments on ecRNH. It should be noted that the experimental pH optima for the RNase H reaction in vitro are approximately 7.5-8.5 for ecRNH and 8.5-9.5 for ttRNH; however, NMR data are not available at these pH ranges due to sample precipitation. The active site residue D10 has a pKa of 6.1 in the absence of Mg2+; however, its pKa normalizes in the presence of Mg2+ and therefore is not expected to be protonated at either the experimental or optimal pH. Crystallographic water molecules were removed from all structures prior to solvation using Schrodinger’s Maestro tool, version 8.5 or 9.1, as distributed in the Desmond software package. Simulations were performed as described using Desmond academic release 3 or source release 2.4.2.1. Proteins were described with the Amber99SB force field, solvated with TIP3P water in a cubic box with a 10Å buffer region from solute to box boundary, and neutralized with Cl- ions. Bonds to hydrogen atoms were constrained using the M-SHAKE algorithm. Simulations containing Mg2+ ions used the Aqvist parameter set. Electrostatics were calculated with the PME method using a 9Å cut-off. All simulations used a 2.5fs inner timestep on a 1-1-3 RESPA scheme to calculate long-range electrostatics. The crystallographic Mg2+ site is unstable in simulation. A simulation was initiated from the crystal structure of ecRNH in the Mg2+-bound state (PDB ID 1RDD) and then run for 100ns unless otherwise noted. These simulation conditions applied to the apo state of RNase H homologs have previously been shown to reproduce NMR data well.

Results and discussion
The crystallographic Mg2+ site is unstable in simulation
A simulation was initiated from the crystal structure of ecRNH in the Mg2+-bound state (PDB ID 1RDD). However, the position of the ion identified in this structure is not stable in simulation and exits the binding site immediately upon initiation of the trajectory. The ion transiently interacts with the protein at a variety of sites on the protein surface over the course of the 89ns trajectory but never returns to its original position in the active site (Figure 4A).

Historically, simulation of the behavior of multivalent ions using standard molecular mechanics force fields has been a long-standing challenge. It is therefore possible that the instability of this position in simulation is an artifact of force field errors. However, given

<table>
<thead>
<tr>
<th>PDB ID</th>
<th>Protein</th>
<th>Source organism</th>
<th>Resolution (Å)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>2RN2</td>
<td>ecRNH</td>
<td><em>Escherichia coli</em></td>
<td>1.48</td>
<td>—</td>
</tr>
<tr>
<td>1RDD</td>
<td>ecRNH</td>
<td><em>Escherichia coli</em></td>
<td>2.80</td>
<td>E. coli WT with single bound Mg2+ ion</td>
</tr>
<tr>
<td>1RIL</td>
<td>ttRNH</td>
<td><em>Thermus thermophilus</em></td>
<td>2.80</td>
<td>—</td>
</tr>
<tr>
<td>2E4L</td>
<td>soRNH</td>
<td><em>Shewanella oneidensis</em></td>
<td>2.00</td>
<td>—</td>
</tr>
<tr>
<td>3H08</td>
<td>ctRNH</td>
<td><em>Chlorobium tepidum</em></td>
<td>1.60</td>
<td>Missing residues in handle and active-site loop modeled in from 1RIL</td>
</tr>
<tr>
<td>2OK9</td>
<td>hsRNH</td>
<td><em>Homo sapiens</em></td>
<td>2.55</td>
<td>Substrate removed and catalytically inactivating D210N mutation reversed in Maestro 8.5</td>
</tr>
<tr>
<td>3KP2</td>
<td>hivRNH</td>
<td>HIV</td>
<td>2.04</td>
<td>Inhibitor and bound metal ions removed in Maestro 9.1; chosen as the HIV structure with lowest RMSD to the unbound state (PDB ID 1HRH) with the active-site loop resolved</td>
</tr>
<tr>
<td>3V1O</td>
<td>xmrvRNH</td>
<td>XMRV</td>
<td>1.88</td>
<td>Full length XMRV RNase H domain with no bound ion</td>
</tr>
<tr>
<td>3P1G</td>
<td>xmrvRNH</td>
<td>XMRV</td>
<td>1.60</td>
<td>Helix C and handle region deletion mutant of XMRV RNase H domain with single bound Mg2+ ion</td>
</tr>
</tbody>
</table>

Summary of crystal structures used as starting points for molecular dynamics simulations of RNase H homologs.
that ions in this position are not observed in the substrate-bound structures of RNase H homologs (Figure 2D), and that the B-factor of the Mg\textsuperscript{2+} ion in the 1RDD structure is much higher than those of the surrounding residues (Figure 4B), it is likely that this position does not reflect the most stable conformation of the protein-ion complex in solution.

Both the crystallographic A and B sites are stable in simulation

Additional simulations were carried out under the same conditions for single Mg\textsuperscript{2+} ions in each of the two Mn\textsuperscript{2+} binding sites identified for ecRNH. Because the crystal structure of ecRNH solved in the presence of Mn\textsuperscript{2+} (PDB ID 1G15) exhibits disorder in both the active-site and handle loops\textsuperscript{7}, the ion positions were instead modeled into the apo ecRNH structure (PDB ID 2RN2) by superposition. For the model of the B-site Mg\textsuperscript{2+} ion, the rotamer of E48 was also corrected to match the orientation observed in the 1G15 structure. For comparison to an alternative homolog, the Mg\textsuperscript{2+} ion in the B site was also modeled into the RNAse H structure from the thermophilic bacterium Thermus thermophilus (ttRNH, PDB ID 1RIL), whose structure was also solved in the absence of divalent ions\textsuperscript{8}.

Mg\textsuperscript{2+} ions were found to be stably associated with the ecRNH active site in both simulations, despite the fact that the ions were modeled into a structure that did not originally contain them (Figure 5). This observation clearly supports the hypothesis that conformational preorganization in the active site promotes ion binding. It is possible that the effectiveness of this modeling procedure was facilitated by a well-documented feature of crystal packing in ecRNH, in which the amino group of a lysine sidechain in a neighboring molecule inserts into the negatively charged active site in a position approximating the B site. However, a short simulation of ttRNH, whose structure does not contain this contact, with Mg\textsuperscript{2+} modeled into the B site was also stable, suggesting that crystal contacts in ecRNH are not responsible for the observation of preorganization in its active site.

The active site is conformationally preorganized for ion binding in the B site

The presence of Mg\textsuperscript{2+} located in either the A or the B site did not substantially affect the dynamics of the active-site residues as determined by S\textsuperscript{2}. All four residues remain highly rigid in the presence of a Mg\textsuperscript{2+} ion in either position (Figure 6). The major difference between the

---

**Figure 4.** Mg\textsuperscript{2+} ion dynamics in a simulation initiated from the ecRNH Mg\textsuperscript{2+}-bound crystal structure. (A) Occupancy map from an 89ns simulation initiated from the ecRNH structure solved in the presence of Mg\textsuperscript{2+} (PDB ID 1RDD), contoured to 0.05% occupancy (corresponding to at least 45ps total residence time). The ion exits the active site and interacts with a variety of regions on the protein surface. (B) The active-site region of the 1RDD structure, colored by atomic B-factor. The B-factor of the ion is substantially larger than the surrounding residues, and is in fact larger than the B-factor of any other atom in the structure save crystallographic waters.

**Figure 5.** Mg\textsuperscript{2+} ion dynamics in an ecRNH simulation initiated with single ions in the A and B sites. (A) Occupancy maps contoured at 10% occupancy for the A site (blue) and B site (green). Neither ion leaves the active site over the 100ns trajectory length. (B) Closer view of the active site in which the four active-site residues in apo ecRNH and the positions of the two Mn\textsuperscript{2+} from which the trajectories were initiated (derived from PDB ID 1G15) are shown for comparison.
unbound, A site, and B site trajectories’ sidechain dynamics was observed in a short loop between helix D and β-sheet 5. Experiments have demonstrated dynamics in this region on the ps-ns timescale, suggesting that the loop is simply incompletely sampled in 100ns simulations rather than significantly perturbed by ion binding. No significant differences in the behavior of these residues upon introduction of ions are observed experimentally.

Of the four conserved catalytic residues, D134 is known to be somewhat dispensable; catalytic activity is retained, though reduced, by substitutions with N or H, which also increase thermostability. In conjunction with crystallographic evidence, this suggests that the B site is occupied in the absence of substrate. Because measurements of the sidechain $^{13}$C$_{\gamma\delta}$ resonances by NMR could not clearly distinguish the behavior of D134 (the unique participant in the A site) from E48 (the unique participant in the B site), comparisons of the two trajectories provide an additional opportunity to distinguish between these two sites.

Although single metal ions in both sites were stably bound to the protein, the RMSD over the course of each 100ns trajectory was larger for the ion in the A site (1.2Å) compared to the B site (0.6Å), which in turn is similar to the RMSD of a 30ns control simulation of trRNH with an ion modeled into the B site (0.6Å). Additionally, a small amount of motion in the direction of the B site was observed for the ion in the A site; the initial and final positions differ by 1.7Å (Figure 5). (By comparison, the A and B sites are about 4Å apart.)

Distinct conformations were also observed for several neighboring residues, reflecting reorganization of local hydrogen bonding networks to accommodate ion binding in each of the two sites. N45 does not differ significantly in sidechain rigidity between the two trajectories, but it does differ in conformation: in the A site trajectory, it is oriented away from the substrate-binding site and participates in a network of interactions that also includes the conserved site T43, while in the B site trajectory N45 is primarily oriented into solvent and occupies the rotamer found in the hsRNH-substrate complex.

The hydrogen-bonding network surrounding D134 unsurprisingly differs considerably between the A and B site trajectories. Occupancy of inter-sidechain hydrogen bonds in this region is summarized in Table 2. H124, which interacts with substrate in the hsRNH complex and is known to be associated with product release, forms hydrogen bonds with D134 in the B site trajectory, partially displacing one of the hydrogen bonds.

<table>
<thead>
<tr>
<th>H-bond</th>
<th>Apo</th>
<th>A Site</th>
<th>B Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>H124-E131</td>
<td>2.2%</td>
<td>47.0%</td>
<td>16.7%</td>
</tr>
<tr>
<td>H124-D134</td>
<td>0.8%</td>
<td>1.5%</td>
<td>23.8%</td>
</tr>
<tr>
<td>H124-E135</td>
<td>0%</td>
<td>0%</td>
<td>0.3%</td>
</tr>
<tr>
<td>R138-E131</td>
<td>0%</td>
<td>11.3%</td>
<td>0%</td>
</tr>
<tr>
<td>R138-D134</td>
<td>74.1%</td>
<td>48.4%</td>
<td>65.5%</td>
</tr>
<tr>
<td>R138-E135</td>
<td>43.4%</td>
<td>29.8%</td>
<td>50.1%</td>
</tr>
</tbody>
</table>

Comparison of hydrogen-bonding environments in the network surrounding the active-site residue D134 in the apo trajectory of ecRNH compared to trajectories containing an Mg$^{2+}$ ion in either the A or the B site. Hydrogen bonds were considered formed if the donor-acceptor distance was less than 3.1Å and the donor-hydrogen-acceptor angle was less than 25°.
the hydrogen bonds formed between D134 and R138 in the apo trajectory. By contrast, H124 interacts primarily with E131 in the A site trajectory, while D134 coordinates Mg\(^{2+}\) in a monodentate manner, partially displacing the R138-D134 interaction. This conformation too is at odds with experimental evidence, since E131 experiences minimal chemical shift perturbation upon Mg\(^{2+}\) binding\(^\text{31}\).

These results collectively add to prior experimental evidence that the B site is the primary site for metal ion binding in the absence of substrate. Furthermore, the presence of a metal ion in the B site may induce reorganization of the surrounding sidechains into conformations conducive to subsequent substrate binding.

**Rigidity of active-site sidechains is conserved within the RNase H family**

Given that all known RNase H homologs have extremely similar active-site structures, it is likely that measurements made on the ecRNH protein can be generalized to other RNase H homologs. \(S_{ap}^2\) values were therefore calculated from previous simulations of the four handle-region-containing bacterial RNase H homologs of known structure, as well as for hsRNH in the absence of substrate\(^\text{29}\).

As might be expected from the high level of structural conservation in the active-site region, the five handle-region-containing RNase H homologs compared differ very little in the dynamics of their active site residues (Figure 9). Notably, the trajectory initiated from the hsRNH structure, which was solved in the presence of substrate and which contained a Na\(^+\) ion in a position similar to the B site in ecRNH, differs very little from trajectories initiated from any other RNase H structure lacking these additional components. This observation provides strong support for the interpretation that the rigid active-site residues are conformationally preorganized for metal-ion interactions even in the unbound state.

In order to better understand the relationships between dynamic processes in RNase H domains of retroviral origin compared to those from cellular organisms, additional simulations in the absence of divalent ions were performed on a set of retroviral RNase H homologs, whose sequences are shown in Figure 8. In brief, no significant differences are observed between simulations initiated from the XMRV full-length structure compared to its \(\Delta C\) mutant (in which helix C and the handle region are removed), between the XMRV \(\Delta C\) mutant compared to the HIV homolog (which naturally lacks the handle sequence), or between any of the retroviral domains compared to ecRNH (Figure 9). This result suggests that the preorganization of the active site on the ps-ns timescale is not significantly altered by differences in amino acid sequence among the subset of family members examined.

**Conclusions**

In this work we aimed to use molecular dynamics simulations to understand the dynamic behavior of the RNase H family in complex with catalytically required Mg\(^{2+}\) ions. We observe that the well-studied RNase H homolog from \(E.\ coli\) contains a conformationally preorganized active site that is highly rigid on the ps-ns timescale in the presence of a single Mg\(^{2+}\) ion, which is likely located at the B site crystallographically identified by examining the Mn\(^{2+}\) complex. Additionally, we examined the apo state dynamics of the active site—previously validated by comparison to NMR data in the case of the \(E.\ coli\) homolog\(^\text{34}\)—and found that similar patterns of active-site rigidity are present in all homologs examined, including representatives of bacterial, eukaryotic, and retroviral RNases H. This result suggests that the dynamics of the active site residues are only subtly modulated by the amino acid sequence variations present among structurally characterized family members. Instead, we hypothesize that this behavior is imposed by the overall protein fold, the topology of which serves to force the negatively charged active-site residues into proximity with one another. Although it has long been recognized that RNases H share similar topologies and active-site conformations with other endonucleases\(^\text{4}\), the present work extends this observation from static crystal structures to dynamics on the ps-ns timescale.

Simulations of the \(Homo sapiens\) RNase H homolog and the \(Thermus thermophilus\) argonaute protein (a distant RNase H homolog with similar active-site architecture) in complex with two Mg\(^{2+}\) ions and substrate analogs also observe high active-site rigidity\(^\text{4}\); this is consistent with the present data and implies that active-site preorganization is a general property of this larger family of nucleases. Recent analysis of the larger RNase H-like superfamiliy identifies a common catalytic core in which the relative orientation of the active-site residues correlates with function, distinguishing endonucleases such as RNase H1 from exonuclease family members\(^\text{15}\). Although selective inhibitors of the HIV RNase H domain have been developed based on the hypothesis that the metal ion dependence of the HIV domain’s catalytic mechanism differs from that of the human homolog\(^\text{15}\), it is...
Figure 8. Multiple sequence alignment of RNase H homologs. The Promals3D<sup>46</sup> structural alignment of the PDB structures used to initiate simulations (see Table 1) is shown with active-site residues highlighted in yellow. Residues that form contacts with substrate in the hsRNH complex (PDB ID 2QK9) are highlighted in orange. The ΔC deletion of XMRV is shown in green. The catalytically inactivating D70N mutation in 2QK9 has been reversed and missing residues from the ctRNH structure (PDB ID 3H08) have been filled in for clarity. Residues are numbered relative to ecRNH. Figure prepared with UCSF Chimera<sup>47</sup>.

Figure 9. Active-site dynamics in RNase H homologs. Calculated S<sup>2</sup> values are shown for the four active-site residues in RNase H homologs. Left: soRNH (dark blue), ecRNH (light blue), ctRNH (magenta), ttRNH (red), hsRNH (purple). Right: ecRNH (light blue), XMRV WT (green), XMRV ΔC (yellow), HIV (brown). All simulations were carried out at 300K in the AMBER99SB force field with TIP3P water with structures protonated to reflect a pH of 5.5.
likely that the physical origin of this selectivity is not dependent on the active-site conformations sampled on the ps-ns timescale.

**Data availability**

ZENODO: Molecular dynamics derived side chain order parameters for Asp, Gln, Asn, and Gln residues in ribonucleases H, and molecular dynamics trajectories for E. coli ribonuclease H, doi:10.5281/zenodo.8431

**Author contributions**

KAS: Conceived and designed experiments, performed experiments, analyzed data, wrote the paper. AGP: Conceived and designed experiments, performed experiments, wrote the paper.

**Author contributions**

KAS: Conceived and designed experiments, performed experiments, analyzed data, wrote the paper. AGP: Conceived and designed experiments, performed experiments, wrote the paper.

**Author contributions**

KAS: Conceived and designed experiments, performed experiments, analyzed data, wrote the paper. AGP: Conceived and designed experiments, performed experiments, wrote the paper.

**Author contributions**

KAS: Conceived and designed experiments, performed experiments, analyzed data, wrote the paper. AGP: Conceived and designed experiments, performed experiments, wrote the paper.

**References**


Open Peer Review

Current Peer Review Status: ✔ ✔

Version 1

Reviewer Report 20 May 2014

https://doi.org/10.5256/f1000research.3859.r4492

© 2014 Montelione G. This is an open access peer review report distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Gaetano Montelione
Department of Molecular Biology and Biochemistry, Rutgers University, Piscataway, NJ, USA

This is an exciting paper which beautifully demonstrates the value of MD simulations coupled with experimental data for understanding structure-dynamic-function mechanisms. The research appears to be carefully executed, with proper controls, and the paper is well written. Together with related papers (e.g, Stafford et al., JACS 2014 - ref 24) it would be an excellent topic paper in the area of protein structure and function for an undergraduate or graduate seminar course.

The results indicate that the apo state of Ec RNase H is conformationally preorganized to favor binding of magnesium ion. MD simulations suggest the primary Mg+2-binding site is actually the observed Mn+2 –binding site B. This is an interesting application of MD to assess the significance of various metal-binding sites previously observed in X-ray crystal structures. 13C nuclear relaxation measurements reveal no statistically significant differences in the fast-time-scale side-chain dynamics of DENQ residues for complexes with Mg+2 modelled into metal binding sites compared with apo Ec RNase H. However, these modelling results suggest some minor structural changes in the active site that may enhance subsequent substrate binding.

MD simulations from a phylogenetically diverse set of RNase H proteins, including representatives of bacterial, eukaryotic, and retroviral RNAses H, suggest that active sites that are conformationally-preorganized for metal binding are a common feature of the RNase H domain family.

The work is suitable for publication in F1000Research but I do suggest minor revisions to address the following points:

1. Fig. 1 shows strong conservation of active site residues among bacterial homologues, and Fig. 8 documents similarity in active site dynamics across a broad phylogenetic distribution, as characterized by MD simulations. The authors conclude that these dynamic properties are only subtly modulated by amino acid sequence and primarily imposed by the characteristic protein fold. As suggested by Reviewer 1 (Rafael Najmanovich), it would be helpful to document the active-site sequence variation across these homologs to support this conclusion.
2. The authors state “all simulations were performed at a pH of 5.5 to recapitulate the conditions used in prior NMR experiments on ecRNH”. What is the pH optimum of enzyme function? Is there any information about “physiological pH”. Based on the pKa’s of active site ionizable groups, how might the pH selected for this study impact the conclusions.

**Competing Interests:** No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 20 Jun 2014

Kate Stafford, Columbia University, New York, USA

Thank you for your thoughtful comments on our paper. In response to your suggestions, we have published a version 2 of the paper which addresses both of your points, specifically:

1. A new Figure 8 has been added to the paper, showing a structure-based sequence alignment for all of the bacterial (soRNH, ecRNH, ctRNH, ttRNH), eukaryotic (hsRNH), and retroviral (XMRV, HIV) RNase H structures used to initiate simulations as part of this work.

2. The pH optimum of the RNase H reaction was previously mentioned only in the Introduction (refs 3 and 4). We have added a comment in the Methods section on the motivation for the choice of pH 5.5 in the NMR experiments and the impact of this on the simulations. In brief, the active site residues are expected to be unprotonated in the presence of Mg2+ at both the experimental pH of 5.5 and the reaction optima of 7.5-9.5.

**Competing Interests:** No competing interests were disclosed.

Reviewer Report 21 March 2014

https://doi.org/10.5256/f1000research.3859.r4005

© 2014 Najmanovich R. This is an open access peer review report distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Rafael Najmanovich

Department of Biochemistry, Université de Sherbrooke, Sherbrooke, Canada

The authors set out to use molecular dynamics simulations to the binding of a required magnesium ion on Ribonuclease H1, utilizing a number of distinct structures including those crystallized in different conditions or forms (Apo or different types of Holo forms) and orthologs. The main result found by the authors is that the binding site is very rigid and pre-organized in the presence of a single magnesium ion. The authors show that this observation holds for all orthologs studied. Based on this, the authors suggest that 'the active site dynamics are only subtly modulated by amino acid sequence and are primarily imposed by the characteristic protein fold'. It is indeed interesting that most RNases not only conserve the structure of the binding site but also dynamic properties in the ns-ps range. However, I feel that showing a
multiple sequence alignment of the RNase H family highlighting the members used in this study as well as the positions that constitute the substrate and metal binding sites would considerably help understand that indeed this pre-organization comes about despite divergent binding-site sequences. Without seeing such an alignment is hard to accept that the rigidity of the binding-site is not affected by changes in binding-site sequence. If the binding-site sequences turn out to be highly conserved upon analysis of the multiple sequence alignment, it may be necessary to perform mutations in silico in the binding-site to support the conclusion above.

**Competing Interests:** No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

---

**Author Response 20 Jun 2014**

**Kate Stafford, Columbia University, New York, USA**

Thank you for your thoughtful comments on our paper. In response to your suggestion, we have published a version 2 of the paper containing a new Figure 8, which shows a structure-based multiple sequence alignment for all RNases H examined (see reviewer response above). The positions of the active-site residues and the substrate-binding regions have been highlighted on the MSA to facilitate comparison of the sequences within these regions. The point you raise can be subtle - since some sequence similarity is required for maintenance of the overall fold, rather than for specific substrate interactions - so we have edited the text of the Conclusions section for clarity and changed the final sentence of the Abstract to read "...may primarily be imposed...". We have also updated the Conclusions section with a new reference to a recent paper on the evolution of the RNase H superfamily (ref 50).

**Competing Interests:** No competing interests were disclosed.