Comparisons of ribonuclease HI homologs and mutants uncover a multi-state model for substrate recognition

James A. Martin,¹ Arthur G. Palmer, III^{2*}

ABSTRACT: Ribonuclease HI (RNHI) non-specifically cleaves the RNA strand in RNA:DNA hybrid duplexes in a myriad of biological processes, including retroviral reverse transcription. Several RNHI homologs contain an extended domain, termed the handle region, that is critical to substrate binding. Nuclear magnetic resonance (NMR) spectroscopy and molecular dynamics (MD) simulations have suggested a kinetic model in which the handle region can exist in open (substrate-binding competent) or closed (substrate-binding incompetent) states in homologs containing arginine or lysine at position 88 (using sequence numbering of *E. coli* RNHI), while the handle region populates a state intermediate between the open and closed conformers in homologs with asparagine at residue 88 [Stafford, K. A., et al., *PLoS Comput. Biol.* **2013**, 9, 1-10]. NMR parameters characterizing handle region dynamics are highly correlated with enzymatic activity for RNHI homologs with two-state (open/closed) handle regions [Martin, J. A., et al., *Biochemistry* **2020**, 59, 3201-3205]. The work presented herein shows that homologs with one-state (intermediate) handle regions display distinct structural features compared with their two-state counterparts. Comparisons of RNHI homologs and site-directed mutants with arginine at position 88 support a kinetic model for handle region dynamics that includes 12 unique transitions between eight conformations. Overall, these findings present an example of the structure-function relationships of enzymes and spotlight the use of NMR spectroscopy and MD simulations in uncovering fine details of conformational preferences.

INTRODUCTION

Ribonuclease HI (RNHI) is an omnipresent, non-sequence-specific, endonuclease that cleaves the RNA strand in RNA:DNA hybrid duplexes. The enzyme has roles in replication,¹ genome maintenance,² and is the C-terminal domain of retroviral reverse transcriptase (RT) proteins.^{3,4} Murine Leukemia Virus (MLV) and Human Immunodeficiency Virus (HIV) are two such retroviruses and their RNHI domains are necessary for viral replication.⁵⁻⁶ A number of RNHI homologs contain a “handle region”, an extended loop (residues 81-101, numbering based on *E. coli* RNHI is used herein) with a large cluster of positive residues, that is critical for substrate recognition (Figure 1A-B).⁸ Solution-state nuclear magnetic resonance (NMR) spectroscopy⁹ combined with Molecular Dynamics (MD) simulations¹⁰ have suggested a two-state kinetic model in which the handle region of the mesophile *E. coli* RNHI (EcRNHI) populates “open” (substrate-binding competent) and “closed” (substrate-binding incompetent) states, while the thermophile *Thermus thermophilus* RNHI (TtRNHI) predominantly populates the closed state at 300 K (Figure 1C).

This model has been extended by correlating NMR parameters describing the handle region conformation, including residue 101 χ_1 trans percentage and Trp 85 Nε1-Hε1 and Thr 92 N-H residual dipolar couplings (RDCs) with enzymatic activity for each homolog, in what is termed the Weighted Conformer Model.¹² However, this model was developed only for homologs with handle regions exhibiting two-state dynamics and able to populate open and closed conformations. Homologs that populate a single conformation intermediate between the open and closed conformations have been identified through MD simulations (Figure 1D).¹⁰ Two-state (open/closed) or one-state (intermediate) behavior of the handle region arises from the identity of

residue 88: long side-chain, positive amino acids (lysine and arginine) are associated with the former, while the shorter, neutral amino acid asparagine is associated with the latter. EcRNHI, TtRNHI, and the psychrophile *Shewanella oneidensis* RNHI (SoRNHI) have two-state handle regions. The moderate thermophile *Chlorobium tepidum* RNHI (CtRNHI) and the *Homo sapiens* RNHI domain (HsRNHI) have one-state handle regions. HsRNHI is the only RNHI homolog containing a handle region that has been co-crystallized with substrate;¹¹ apo-EcRNHI residues K86, K87 and R88 form crystal contacts with acidic active sites of neighboring molecules.¹³⁻¹⁴

The present work investigates whether homologs with one-state handle regions display dynamic and enzymatic properties that can be characterized similarly as for two-state counterparts. NMR experiments characterizing handle region dynamics were performed as described previously.¹² Reciprocal mutations were made in EcRNHI (R88N) and CtRNHI (N88R) to determine the effect of the residue type at position 88. Further comparisons between RNHI homologs were made using a combination of Circular Dichroism (CD) spectroscopy, NMR spectroscopy, MD simulations, and x-ray crystallographic structures. These data establish N88 as an important contributor in the thermostability of the hydrophobic core of RNHI, as well as amend the kinetic model for handle region dynamics to include 12 unique transitions between eight total conformations.

CtRNHI is a moderate thermophile with 57% sequence identity to TtRNHI and 53% sequence identity to EcRNHI. Its overall stability and melting temperature are similar to EcRNHI, but its change in heat capacity upon unfolding and stability at optimal growth temperature are more similar to TtRNHI. Finally, its optimal growth temperature is intermediate EcRNHI and TtRNHI.¹⁵ The x-ray crystal structure of

Figure 1. Ribonuclease HI Handle Region. (A) Multiple sequence alignment of RNHI handle regions from *E. coli* (Ec), *S. oneidensis* (So), *T. thermophilus* (Tt), *C. tepidum* (Ct), and *H. sapien* (Hs). Yellow highlights: DNA-binding residues.¹¹ Sequence numbering is according to EcRNHI. (B) Tertiary structure of EcRNHI with a focus on the handle region. Trp 85 (W85) and Thr 92 (T92) are shown in sticks. (C) Two-state kinetic model for handle region dynamics.¹⁰ Open and closed states are characterized by the distance from the Ca of W85 to the Ca of T92 at the tip of the dynamic handle loop; open distances are taken to be ≥ 7.8 Å and closed distances < 7.8 Å. Homologs with Lys or Arg at residue 88 (pink) follow this model; those with Asn (green) display a handle region conformation intermediate between the open and closed conformers (D).

CtRNHI has been solved at a resolution of 1.60 Å (PDB ID: 3H08) and shows two monomers in the unit cell, both with missing residues in the handle region (Figure S1).¹⁵ Superposition of CtRNHI and TtRNHI shows higher structural similarity of the handle regions (0.754 Å Ca-RMSD) compared to EcRNHI (0.975 Å Ca-RMSD). CtRNHI has a characteristic handle distance of 6.7 Å (see Fig.1 caption), which is more similar to TtRNHI (6.5 Å) than EcRNHI (7.8 Å). However, structures of CtRNHI observed in MD simulations have handle distances (~ 7.4 Å) that are intermediate and more like distances in EcRNHI.¹⁰ The disparity between the two measurements provides ambiguity about the handle region dynamics of CtRNHI.

RESULTS

Near complete backbone and sidechain NMR resonance assignments were made for CtRNHI (all NMR experiments herein used Cys-free mutant proteins) and handle region dynamics were characterized by NMR spectroscopy as described previously.¹² The Weighted Conformer Score calculated for CtRNHI from these NMR parameters (0.31) is intermediate between EcRNHI (0.21) and TtRNHI (0.63) (Table S1). If CtRNHI has handle region dynamics similar to homologs with two-state handle regions, then the weighted conformer score predicts a Michaelis constant (K_M) that is approximately 2.4 (± 0.4) fold greater than EcRNHI (Figure 2A). In contrast, enzymatic assays for CtRNHI and EcRNHI yield a relative Michaelis constant of 0.77 ± 0.29 (Figure 2B). These data suggest that the contributions of handle region dynamics to substrate recognition by CtRNHI differ in some regards to two-state handle region RNHI homologs.

Reciprocal mutations were made in CtRNHI (N88R) and EcRNHI (R88N) to better understand the features of one-

Figure 2. CtRNHI Enzyme kinetics. (A) **Open Diamond:** Michaelis constant (K_M) prediction for CtRNHI using the Weighted Conformer Model.¹² Experimental K_M ratios: K_M for each RNHI homolog normalized by the corresponding K_M for EcRNHI. **Square:** EcRNHI, **Circle:** SoRNHI, **Triangle:** V98A EcRNHI, **Inverted Triangle:** TtRNHI, **Closed Diamond:** CtRNHI (B) Initial reaction velocities are plotted versus substrate concentrations for EcRNHI (circle, solid line) and CtRNHI (square, dashed line). Lines are calculated from the Michaelis-Menten equation using values of K_M and V_{max} obtained from global fits of progress curves. Fitted values of K_M and V_{max} are 920 ± 240 nM and 13.5 ± 3.2 nM/s for EcRNHI and 710 ± 190 nM and 12.6 ± 2.9 nM/s for CtRNHI, respectively.

state handle regions. Backbone and sidechain NMR resonance assignments were obtained for each mutant and compared to its respective wildtype counterpart. Overlaid ^{15}N and ^{13}C heteronuclear single quantum coherence (HSQC) NMR spectra for CtRNHI show that the structural changes caused by the mutation are localized to the handle region (Figure S2). Specifically, resonances for K91 (the interaction partner of N88 in wild-type CtRNHI), V98 and I101 are the most perturbed, consistent with previous research suggesting that these residues have concerted movements.¹² K91 is also most perturbed in R88N EcRNHI, however, the overall differences are small and broadly distributed (Figure S3, consistent with studies showing that dynamics of the handle region are correlated with structural elements throughout EcRNHI.¹⁶⁻¹⁷ The NMR chemical shift perturbations are complemented by nearly identical CD spectra for the respective homolog/mutant pairs, suggesting the secondary structure of the enzyme remains unchanged upon mutation (Figure 3A-B). Thermal denaturation is more reversible for wild-type and N88R mutant CtRNHI than for wild-type and R88N EcRNHI; the thermal melts also show an 11% decrease in

Figure 3. Far-UV CD Spectra of wild-type and mutant RNHI enzymes. CD spectra recorded at 25°C for (A) **Orange:** CtRNHI, **Blue:** N88R CtRNHI and (B) **Orange:** EcRNHI, **Blue:** R88N EcRNHI proteins before (closed circles) and after (open circles) denaturation at 85 °C. CD signal from native (x-axis) vs renatured (y-axis) enzymes are respectively plotted for (C) CtRNHI and (D) EcRNHI. Slopes represent the recovery of CD signal (reversibility) over the spectra range.

reversibility for N88R CtRNHI (Figure 3C) and a 9% increase in reversibility for R88N EcRNHI (Figure 3D).

Experimental NH RDCs, an NMR parameter used for structure validation and refinement,¹⁸ agree poorly with values fitted to the crystal structure of CtRNHI (Figure S4A), with a Q factor of 0.50 and R = 0.81; agreement is slightly worse when only handle region residues are considered, with R = 0.76 (Figure S4B). Thus, the structure of the backbone and handle region in the crystal structure of CtRNHI may differ from the structure(s) in solution. MD simulations for CtRNHI were used to obtain a structural model more representative of the experimental RDCs. Briefly, experimental RDCs were fit to individual structural frames of a 100 ns trajectory at 300 K to identify the structure (BF CtRNHI) with the lowest Q factor.¹⁸ Agreement improved in comparison to the crystal structure for backbone NH RDCs (Figure S4C), with Q = 0.37 and R = 0.85, and for the handle region (Figure S4D) with R = 0.89. Additionally, the *in silico* structure BF CtRNHI has a better Molprobity score (metric evaluating the quality of a structure)¹⁹ than either of the monomers in the crystal structure (Table S2).

The handle distance for BF CtRNHI is 7.2 Å, intermediate between EcRNHI (7.8 Å) and TtRNHI (6.5 Å). Inspection of

the handle region of BF CtRNHI (Figure 4A) shows that the side chain of N88 makes two hydrogen bonds with the backbone of K91: N88 Oδ1:H K91 and N88 Hδ2:O K91 (as noted previously)¹⁰ and none with T92. HsRNHI is the other homolog that naturally contains asparagine at residue 88. Superposition of BF CtRNHI and HsRNHI (2QK9) show similar architecture, even though the HsRNHI structure was determined in complex with substrate (Figure 4B). The handle regions of HsRNHI and BF CtRNHI have a Cα-RMSD of 0.65 Å. The main perturbations occur at the end of the handle loop (after V98). When the end of the handle loop is excluded from alignment, the Cα-RMSD reduces to 0.47 Å. One of the hydrogen bonds (N88 Hδ2:O K91) is longer in HsRNHI (2.2 Å vs 2.0 Å), and the distance between N88 Nδ2-Hδ2 and T92 O is also larger (5.1 Å vs 4.9 Å). These slight differences combined with the steric constraint due to the presence of substrate may account for why HsRNHI is more open (7.8 Å) than BF CtRNHI (7.2 Å).

Structural analysis of a sextuple, thermostable mutant of SoRNHI (6X SoRNHI)²⁰ offers further insight into the interactions of N88 with other handle region residues. 6X SoRNHI has six mutations: N27K, D37G, M74V, K88N, R95G, and D134H that increase its thermostability by 28.8 K and decrease its relative activity to 43% of the wildtype enzyme. The effects of the mutations on stability are approximately additive.²⁰ Furthermore, the overall structure of 6X SoRNHI (2ZQB) is unaltered outside of the mutation sites. The four monomers in the crystallographic unit cell (Figure 5A-D) are structurally similar to each another (Figure 5E) and have an all Cα-RMSD to wild-type SoRNHI in the range of 0.653-1.062 Å (Table S3). The handle region for each monomer is also structurally similar to the wildtype structure, although the position of the handle loop relative to the protein core differs between monomers (vide infra).²⁰

Figure 4. N88 Interactions of HsRNHI and CtRNHI. (A) Closeup of the handle region of BF CtRNHI with a focus on W85, N88, K91 and T92. (B) Overlay of HsRNHI(cyan) and BF CtRNHI (green) handle regions. The overall architecture of the interactions of N88 with K91 and T92 remains the same. Sequence numbering is according to EcRNHI.

These observations suggest the mutations act independently of one another. This conclusion is consistent with a study of a quintuple thermostable EcRNHI mutant (5X EcRNHI: G23A, H62P, V74L, K95G, D134H) that showed the structures at the mutation sites are almost identical to the respective structures for the single mutant variants.²¹ Thus, with reasonable certainty, the interactions of N88 in 6X SoRNHI are primarily the effect of the K88N mutation.

Figure 5. N88 Interactions and structural comparisons of 6X SoRNHI monomers. Handle region of each monomer, A-D (labeled accordingly), of 6X SoRNHI (2ZQB). (E) Overlay of all 4 monomers of 6X SoRNHI: Monomer A (orange), B (cyan), C (red), and D (green). (F) Matrix of Cα-RMSDs (Å) for the overall structure (above diagonal) and for the handle region (below diagonal).

The interactions of N88 in the sextuplet mutant mirrors that of the other homologs and variants containing N88. The N88 sidechain interacts with the backbones of M91 and T92 in monomers A and C; according to the handle distance metrics (Table S3), their handle regions would be considered closed (6.8 and 6.7 Å respectively). The N88 side chain points away from the backbones of M91 and T92 in monomer B; its handle region metrics would classify monomer B as the most open structure (7.6 Å). The N88 side chain interacts with M91 O but not with T92 O in monomer D. This conformation is nearly identical to the other three monomers, with the main difference being the hydrogen bond with M91 O is longer and N88 Oδ1 does not form a hydrogen bond with M91 NH. Nonetheless, Monomer D appears to be in a transition state from closed to open; its handle metrics are characteristic of a closed structure, but its handle distance (7.4 Å) is closer to that of Monomer B (7.6 Å) than A and C (6.8 and 6.7 Å respectively).

A matrix of Cα-RMSDs for the four monomers show that monomers B and D are most alike overall (Figure 5F, upper triangle) and in their handle regions (Figure 5F, lower triangle). The handle conformations of monomers B and D are much more similar than to either monomers A and C (Fig. 5F, lower triangle). Due to the high overall structural similarity of the four monomers, each monomer may represent a step in the transition of the handle region from closed to open. The process may start with the most closed monomer (C: 6.7 Å) stabilized by hydrogen bonds between N88 Nδ2-Hδ2 and M91 O (2.4 Å) and T92 O (2.5 Å). Next, these hydrogen bonds lengthen (2.9 Å and 3.7 Å respectively) as the handle region becomes slightly less closed (Monomer A: 6.8 Å). The hydrogen bond between N88 Nδ2-Hδ2 and T92 O breaks (5.5 Å), transitioning the structure from closed to open (Monomer D: 7.4 Å). Lastly, the hydrogen bond between N88 Nδ2-Hδ2 and M91 O breaks (5.5 Å), allowing the sidechain of N88 to orient away from the backbone of

both K91 and T92 and enabling a more open state (Monomer B: 7.6 Å).

These observations suggest that the sidechain interactions of N88 with the backbones of residues 91 and 92 modulate handle region dynamics: the handle region conformations transition from the most closed structure (Monomer C: 6.7 Å) to the most open structure (Monomer B: 7.6 Å) and vice versa as the N88 sidechain hydrogen bonds break and reform. This mechanism is consistent with previous results showing that the interaction between R88 and T92 was characterized by the handle distance metric with shorter distances more prevalent in closed conformations.¹² A major difference between the RNHI homologs with Arg or Asn at position 88, however, is that the R88 sidechain is only involved in one interaction, probably water-mediated, with T92 O, whereas the N88 sidechain is involved in up to three hydrogen bond interactions with the handle loop backbone (N88 Oδ1: K91 N-H, N88 Nδ2-Hδ2: K91 O, and N88 Nδ2-Hδ2: T92 O). If all three interactions modulate handle distance, then eight (2³) possible handle region conformations can be characterized by the different hydrogen bonding patterns.

Table 1 shows the average handle distances observed when hydrogen bonds are present or absent in MD simulations of the N88 RNHI variants. The handle distances do not depend significantly on the presence or absence of N88 sidechain interactions with the K91 backbone. However, shorter average handle distances are observed when the hydrogen bond between N88 Nδ2-Hδ2 and T92 O is present than when the hydrogen bond is absent (paired t-test, $p = 0.0097$), consistent with data presented so far. Trends are also seen across the homologs for the eight possible combinations of the N88 sidechain interactions (Table 1). Ordering of the eight conformations can be done in similar manner as for the monomers of 6X SoRNHI, as sequential increases in handle distance result in either the loss or gain of a hydrogen bond. Interestingly, this ordering results in the first four combinations all containing a hydrogen bond between N88

Table 1. Handle Distance averages of N88 Hydrogen Bonding interactions.

Protein	88 Oδ1 : 91 N		88 Nδ2 : 91 O		88 Nδ2 : 92 O	
	Present	Absent	Present	Absent	Present	Absent
R88N EcRNHI	8.4(43.9%)	8.3(56.1%)	8.3(67.9%)	8.4(32.1%)	7.5(12.2%)	8.4(87.8%)
K88N SoRNHI	8.4(57.1%)	8.5(42.9%)	8.4(78.2%)	8.7(21.8%)	7.7(16.8%)	8.6(83.2%)
CtRNHI	7.6(95.3%)	7.2(4.7%)	7.6(97.1%)	7.9(2.9%)	6.9(2.9%)	7.3(93.4%)
HsRNHI	8.0(93.2%)	7.1(6.8%)	7.9(95.4%)	7.9(4.6%)	7.0(7.1%)	8.0(92.9%)

Protein	88 Oδ1 & 92 O	92 O	91 O & 92 O	All	88 Oδ1 & 91 O	91 O	None	88 Oδ1
R88N EcRNHI	(0%)	6.5(0.5%)	7.5(10.2%)	7.9(1.6%)	8.4(41.4%)	8.5(14.8%)	8.5(30.6%)	9.1(1.0%)
K88N SoRNHI	5.9(<0.1%)	6.5(0.8%)	7.7(11.5%)	7.9(4.5%)	8.4(50.6%)	8.9(11.6%)	8.7(19.1%)	9.2(1.9%)
CtRNHI	(0%)	6.1(<0.1%)	6.3(1.4%)	7.6(1.4%)	7.6(93.4%)	7.3(0.8%)	7.8(2.3%)	8.6(0.5%)
HsRNHI	(0%)	6.0(0.9%)	6.8(3.6%)	7.6(2.6%)	8.0(88.2%)	7.8(1.0%)	8.2(1.2%)	8.6(2.4%)

Top: 100 ns MD simulations at 300 K for 4 RNHI homologs and mutants (R88N EcRNHI, K88N SoRNHI, CtRNHI, and HsRNHI) were separated respectively by the presence or absence of the three hydrogen bonding interactions of the N88 sidechain (N88 Oδ1: K91 N-H, N88 Nδ2-Hδ2: K91 O, and N88 Nδ2-Hδ2: T92 O). The VMD H-Bond plugin (4.0 Å and 120° cutoff) was used to assess the presence of a hydrogen bond. The handle distance averages (Å) of each group and their respective percentages in parenthesis are reported. Bottom: Same as top but frames were separated by the eight possible combinations of interactions. The acceptor atom of each interaction is shown in the column name.

Nδ2-Hδ2 and T92 O (preferably closed), while the last four do not (preferably open). As such, the combinations can be grouped into proposed closed conformers that contain this interaction (Figure 6) and open conformers that do not (Figure 7).

DISCUSSION

CtRNHI has structural and thermodynamic features that are similar to its mesophilic (EcRNHI) and thermophilic (TtRNHI) counterparts. Additionally, MD simulations found that CtRNHI has a one-state handle region conformation that is intermediate between the two-state open and closed handle region conformations of EcRNHI and TtRNHI. To investigate these differences, NMR chemical shifts, scalar coupling constants and RDCs were measured to characterize the conformational properties of the CtRNHI handle region. The weighted conformer score for CtRNHI is intermediate between EcRNHI and TtRNHI, giving a *de novo* K_M prediction that is ~2.4 fold greater than EcRNHI; in contrast, activity assays yielded a K_M for CtRNHI that is statistically indistinguishable from EcRNHI. This discrepancy is consistent with MD simulations suggesting that CtRNHI populates a state of the handle region in between the open and closed states. RNHI homologs in which position 88 is an Arg residue populate both open and closed states of the handle region in solution and substrate affinity is correlated with the open state population. In contrast, RNHI homologs with Asn at position 88 populate one-state handle region conformations preferentially and behave differently than homologs with two-state handle region conformations. Thus one- and two-state RNHI homologs require different models to describe conformational dynamics and substrate affinity: for CtRNHI the open-state population is not a dominant predictor of substrate affinity.

Comparison of reciprocal mutants at sequence position 88 for EcRNHI and CtRNHI provides insight into the structural features of one-state handle regions. The effect of mutation

is predominantly local (confined to the handle region) as evident by ^{15}N and ^{13}C HSQC chemical shift perturbations for N88R CtRNHI, but small and broadly distributed in R88N EcRNHI. Complementary CD spectroscopy shows that the secondary structures of the homologs are barely changed by mutation. N88 is preferred in thermophilic homologs¹⁰ and is thermostabilizing with a low mutability score in the psychrophile, SoRNHI.²² Furthermore, previous studies of chimeras of RNHI homologs²³⁻²⁴ have shown that the hydrophobic core (residues 43–122) correlates with capacity for reversible thermal denaturation: the cores of CtRNHI and TtRNHI are sufficient for renaturation, while the core of EcRNHI is insufficient without the use of a chemical denaturant²⁵ or low pH.²⁶ HsRNHI, a homolog with N88, also unfolds reversibly upon thermal denaturation.²⁷ Additionally, N88 containing ancestral mutants of EcRNHI and TtRNHI, with the exception of AncB (last mesophilic ancestor with N88), unfolded reversibly upon thermal denaturation.²⁸ These data reinforce the notion of N88 as a key contributor to the thermostability of RNHI enzymes.

Further comparison of N88 RNHI homologs/mutants with R88 RNHI homologs/mutants amplify structural contributions to handle region dynamics.¹² Arginine has a long guanidino sidechain with only one of its protons in range to interact with the backbones of K91 or T92; a hydrogen bond is formed with K91, but its weaker interaction with T92 might be mediated through a potential water molecule. Asparagine on the other hand has a short carboxamide sidechain in range to form three hydrogen bonds with backbone atoms of K91 and T92; 2 with K91 and one with T92.

An Asn at position 88 in RNHI homologs can make three side chain hydrogen bonds with residues 91 and 92: N88 Oδ1: K91 NH, N88 Nδ2-Hδ2: K91 O, and N88 Nδ2-Hδ2: T92 O. The conformational states of the handle region for N88 RNHI homologs are modulated by the eight possible combinations of the N88 hydrogen bonds. Four closed states are associated with the presence of the N88 Nδ2-Hδ2: T92 O

hydrogen bond (Figure 6) and four open states are associated with the absence of this hydrogen bond (Figure 7). Seven of these eight conformational states are seen among the multiple monomers of 6X SoRNHI (2ZQB) and HsRNHI (2QKK). Moreover, MD simulations of N88 homologs and mutants show that structural transitions of the handle region between consecutive frames change are accompanied by changes in hydrogen bond interactions of ± 1 , for a total of 12 unique transitions between ordered states. In this model, the (rare in MD simulations) fully closed state is shown in Figure 6A, in which the N88 N82-H82: T92 O and N88 O81: K91 NH hydrogen bonds are formed. Upon losing the hydrogen bond with K91 NH, the enzyme transitions to the closed state that is most similar to that observed in R88 containing homologs (Figure 6B). Formation of subsequent hydrogen bonds with the backbone of K91, first N88 N82-H82: K91 O and then re-forming N88 O81: K91 NH, yield a pair of closed intermediates (Figure 6C & D). The loss of the hydrogen bond with T92 O yields a pair of quasi-open conformations with both N88 N82-H82: K91 O and N88 O81: K91 NH hydrogen bonds (Figure 7A) or only the N88 N82-H82: K91 O hydrogen bond (Figure 7B). All three hydrogen bonds are absent in the open state similar to that of R88 RNHI homologs (Figure 7C). Formation of the N88 O81: K91 NH hydrogen bond yields a fully open state (Figure 7D). The MD simulations of all four N88 RNHI homologs and mutants follow this kinetic model $\geq 99\%$ of the time (Table S4); an overwhelming 81-96% of these transitions feature no change in hydrogen bond interactions, consistent with the stable nature of one-state handle region conformations. An important consequence of this model is that intermediate open conformations of CtRNHI must have altered affinities for substrate compared with the open conformation to account for the difference between predicted and measured Michaelis constants using the Weighted Conformer Model developed for two-state RNHI homologs.

CONCLUSION

Residue 88 was previously shown to be a major player in RNHI handle region dynamics, as the choice of amino acid at this position affects conformational preferences: Arginine and lysine is present in homologs with two-state handle regions, while asparagine is present in homologs with one-state handle regions.¹⁰ A weighted conformer model was designed and shown to correlate open-state populations with the Michaelis constants for enzyme activity for homologs with two-state handle regions, but this model was not tested for homologs with one-state handle regions.¹² To better understand the mechanisms of substrate recognition by RNHI homologs with one-state handle regions, this study conducted near complete NMR backbone and sidechain resonance assignments of a representative homolog, CtRNHI, and showed that the structural features underlying conformational change in one-state handle regions differs from two-state counterparts. Structural comparisons of N88 homologs and mutants showed a direct correlation between the interactions of N88 with K91 and T92 and a handle distance metric defining a range of conformations between fully open and fully closed states. Using MD simulations of K88N SoRNHI and direct evidence from four monomers in the x-ray crystal structure of a mutant of SoRNHI containing six stabilizing mutations, including K88N, a kinetic model was developed that includes 12 unique transitions between eight confor-

mations of the handle region for N88 RNHI homologs. In contrast to RNHI homologs with two-state handle regions, the Michaelis constant of CtRNHI is not strongly determined by measures of the open state population, implying that other intermediate states make substantial contributions to substrate affinity. These findings exemplify the cohesiveness and power in the use of multiple biophysical tools to paint a more comprehensive picture of the linkage between protein dynamics and activity.

Figure 6. Proposed Closed Conformers for RNHI Homologs.

A randomly selected frame was taken from each of the first 4 interaction groups of Table 1B from K88N SoRNHI; K88N SoRNHI is the only RNHI variant that populates all eight interaction combinations in MD simulations. Distances of all 3 potential N88 sidechains are measured and handle distances are also displayed for interaction combinations: (A) N88 N82-H82: T92: O and N88 O81: M91 N-H (B) N88 N82-H82: T92 O (C) N88 N82-H82: T92: O and N88 N82-H82: M91 O (D) N88 N82-H82: T92: O, N88 N82-H82: M91 O, and N88 O81: M91 N-H. This group is defined by the presence of N88 N82-H82: T92 O hydrogen bond.

Figure 7. Proposed Open Conformers for RNHI Homologs.

Same as Figure 6 but for interaction groups: (A) N88 N82-H82: M91 O and N88 O81: M91 N-H (B) N88 N82-H82: M91 O (C) No sidechain interactions (D) N88 O81: M91 N-H. This group is defined by the absence of N88 N82-H82: T92 O hydrogen bond.

ASSOCIATED CONTENT

Supporting Information

Experimental procedures, one equation, five figures, and four tables are included in the Supporting Information.

Accession Codes

EcRNHI (UniProt P0A7Y4), CtRNHI (Uniprot Q93SU7), SoRNHI (UniProt A0A2W5DXL8), and HsRNHI (UniProt O60930). Backbone ^1H - ^{15}N backbone assignments are deposited at BioMagResBank: CtRNHI (BMRB 51160) N88R CtRNHI (BMRB 51162) and R88N EcRNHI (BMRB 51163).

AUTHOR INFORMATION

Corresponding Author

***Arthur G. Palmer III**- *Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York, U.S.A;* agp6@cumc.columbia.edu

Authors

James A. Martin- *Department of Biological Sciences, Columbia University, New York, New York, U.S.A*

Funding Sources

This work was supported by an NSF graduate research fellowship (DGE 1644869; J.A.M.) and the NIH (R35GM130398; A.G.P.). Some of the work presented here was conducted at the Center on Macromolecular Dynamics by NMR Spectroscopy located at the New York Structural Biology Center (NYSBC), supported by NIH grant P41 GM11830. The (Columbia) 600 and (NYSBC) 800 MHz NMR spectrometers were supported by NIH grants S10RR026540 and S10OD016432, respectively. CD spectrophotometer and fluorimeter were supported by NIH grant 1S10OD025102 and NSF MRI grant 1828491, respectively.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

This paper is dedicated to the memory of Arlene Davis. We acknowledge Kate Stafford (Columbia University) for providing MD simulations; Kate Stafford and Tassadite Dahmane (New York Structural Biology Center) for sample preparation; and Jia Ma (Columbia University) for assistance with CD and fluorescence spectroscopy.

REFERENCES

- Cozar, J. M. G. D.; Gerards, M.; Teeri, E.; George, J.; Dufour, E.; Jacobs, H. T.; Joers, P. RNase H1 promotes replication fork progression through oppositely transcribed regions of Drosophila mitochondrial DNA. *J. Biol. Chem.* **2019**, *294*(12), 4331–4344.
- Parajuli, S.; Teasley, D. C.; Murali, B.; Jackson, J.; Vindigni, A.; Stewart, S. A. Human ribonuclease H1 resolves R-loops and thereby enables progression of the DNA replication fork. *J. Biol. Chem.* **2017**, *292*(37), 15216–15224.
- Champoux, J. J.; Schultz, S. J. Ribonuclease H: Properties, substrate specificity and roles in retroviral reverse transcription. *FEBS J.* **2009**, *276*(6), 1506–1516.
- Beilhardt, G. L.; Gotte, M. HIV-1 Ribonuclease H: Structure, Catalytic Mechanism and Inhibitors. *Viruses* **2010**, *2*(4), 900–926.
- Hansen, J.; Schulze, T.; Mellert, W.; Moelling, K. Identification and characterization of HIV-specific RNase H by monoclonal antibody. *EMBO J.* **1988**, *7*(1), 239–243.
- Tisdale M.; Schulze T.; Larder B. A.; Moelling K. Mutations within the RNase H domain of human immunodeficiency virus type 1 reverse transcriptase abolish virus infectivity. *J. Gen. Virol.* **1991**, *72*, 59–66.
- Tramontano, E.; Corona, A.; Menéndez-Arias, L. Ribonuclease H, an unexploited target for antiviral intervention against HIV and hepatitis B virus. *Antiviral Res.* **2019**, *171*, 104613.
- Kanaya, S.; Katsuda-Nikai, C.; Ikehara, M. Importance of the positive charge cluster in *Escherichia coli* ribonuclease HI for the effective binding of the substrate. *J. Biol. Chem.* **1991**, *266*, 11621–11627.
- Butterwick, J. A.; Palmer, A. G. An inserted gly residue fine tunes dynamics between mesophilic and thermophilic ribonucleases H. *Protein Sci.* **2006**, *15*, 2697–2707.
- Stafford, K. A.; Robustelli, P.; Palmer, A. G. Thermal adaptation of conformational dynamics in ribonuclease H. *PLoS Comput. Biol.* **2013**, *9*, 1–10.
- Nowotny, M.; Gaidamakov, S. A.; Ghirlando, R.; Cerretti, S. M.; Crouch, R. J.; Yang, W. Structure of Human RNase H1 Complexed with an RNA/DNA Hybrid: Insight into HIV Reverse Transcription. *Mol. Cell* **2007**, *28*(3), 264–276.
- Martin, J. A.; Robustelli, P.; Palmer, A. G., 3rd, Quantifying the Relationship between Conformational Dynamics and Enzymatic Activity in Ribonuclease HI Homologues. *Biochemistry* **2020**, *59*(35), 3201–3205.
- Goedken, E. R.; Marqusee, S. Co-crystal of *Escherichia coli* RNase HI with Mn^{2+} ions reveals two divalent metals bound in the active site. *J. Biol. Chem.* **2001**, *276*(10), 7266–7271.
- Katayanagi, K.; Miyagawa, M.; Matsushima, M.; Ishikawa, M.; Kanaya, S.; Nakamura, H.; Ikehara, M.; Matsuzaki, T.; Morikawa, K. Structural details of ribonuclease H from *Escherichia coli* as refined to an atomic resolution. *J. Mol. Biol.* **1992**, *223*(4), 1029–52.
- Ratcliff, K.; Corn, J.; Marqusee, S. Structure, Stability, and Folding of Ribonuclease H1 from the Moderately Thermophilic *Chlorobium tepidum*: Comparison with Thermophilic and Mesophilic Homologues. *Biochemistry* **2009**, *48*(25), 5890–5898.
- Iwai, S.; Wakasa, M.; Ohtsuka, E.; Kanaya, S.; Kidera, A.; Nakamura, H. Interaction of the Basic Protrusion of *Escherichia coli* Ribonuclease HI with its Substrate. *J. Mol. Biol.* **1996**, *263*(5), 699–706.
- Livesay, D. R.; Jacobs, D. J. Conserved quantitative stability/flexibility relationships (QSFR) in an orthologous RNase H pair. *Proteins Struct. Funct. Bioinf.* **2005**, *62*(1), 130–143.
- Chen K.; Tjandra, N. The use of residual dipolar coupling in studying proteins by NMR. *Top. Curr. Chem.* **2012**, *326*, 47–67.
- Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. (MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*(1), 12–21.
- Rohman, M. S.; Tadokoro, T.; Angkawidjaja, C.; Abe, Y.; Matsumura, H.; Koga, Y.; Takano, K.; Kanaya, S. Destabilization of psychrotrophic RNase HI in a localized fashion as revealed by mutational and X-ray crystallographic analyses. *FEBS J.* **2009**, *276*(2), 603–613.
- Haruki, M.; Tanaka, M.; Motegi, T.; Tadokoro, T.; Koga, Y.; Takano, K.; Kanaya, S. Structural and thermodynamic analyses of *Escherichia coli* RNase HI variant with quintuple thermostabilizing mutations. *FEBS J.* **2007**, *274*, 5815–5825.
- Tadokoro, T.; Kazama, H.; Koga, Y.; Takano, K.; Kanaya, S. Investigating the structural dependence of pro-

- tein stabilization by amino acid substitution. *Biochemistry* **2013**, *52*(16), 2839-2847.
23. Robic, S.; Berger, J. M.; Marqusee, S. Contributions of folding cores to the thermostabilities of two ribonucleases H. *Protein Sci.* **2002**, *11*(2), 381-389.
24. Ratcliff, K.; Marqusee, S. Identification of residual structure in the unfolded state of ribonuclease H1 from the moderately thermophilic *Chlorobium tepidum*: comparison with thermophilic and mesophilic homologues. *Biochemistry* **2010**, *49*(25), 5167-5175.
25. Kanaya, S.; Katsuda, C.; Kimura, S.; Nakai, T.; Kitakuni, E.; Nakamura, H.; Katayanagi, K.; Morikawa, K.; Ikehara, M. Stabilization of *Escherichia coli* ribonuclease H by introduction of an artificial disulfide bond. *J. Biol. Chem.* **1991**, *266*(10), 6038-6044.
26. Kimura, S.; Oda, Y.; Nakai, T.; Katayanagi, K.; Kitakuni, E.; Nakai, C.; Nakamura, H.; Ikehara, M.; Kanaya S. Effect of cavity-modulating mutations on the stability of *Escherichia coli* ribonuclease H1. *Eur. J. Biochem.* **1992**, *206*(2), 337-43.
27. Wu, H.; Lima, W. F.; Crooke, S. T. Properties of cloned and expressed human RNase H1. *J. Biol. Chem.* **1999**, *274*(40), 28270-28278.
28. Hart, K. M.; Harms, M. J.; Schmidt, B. H.; Elya, C.; Thornton, J. W.; Marqusee, S. Thermodynamic system drift in protein evolution. *PLoS Biol.* **2014**, *12*(11), e1001994.