Conformational Sub-states and Dynamics in Human Ribonuclease Family

Khushboo Bafna
University of Tennessee, kbafna@vols.utk.edu

Follow this and additional works at: https://trace.tennessee.edu/utk_graddiss

Recommended Citation
https://trace.tennessee.edu/utk_graddiss/5683

This Dissertation is brought to you for free and open access by the Graduate School at Trace: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of Trace: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.
Conformational Sub-states and Dynamics in Human Ribonuclease Family

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Khushboo Bafna
August 2019
Acknowledgments

I am deeply grateful to the Genome Science and Technology program for providing me the opportunity to pursue graduate studies at the University of Tennessee, Knoxville. I would like to extend my sincere gratitude toward Dr. Pratul Agarwal for giving me the opportunity to obtain my Ph.D. under his mentorship. His constant support and guidance have helped me complete my dissertation research successfully. Despite the changes at his workplace, he made sure that I had access to all the necessary resources I need throughout these years, first at Oak Ridge National Laboratory and later at the University of Tennessee Knoxville. I appreciate his efforts to help me develop critical thinking and research ethics, which are extremely essential to thrive in the scientific arena. His attention to details and meticulous nature has brought the best out of me for conducting scientific research and presenting it to the broader scientific community.

I would also like to extend my most sincere gratitude to our collaborator and my committee member late Dr. Elizabeth Howell. A marvelous teacher and an amazing human being, she will be dearly missed! I admire her love for science and her desire to learn new skills needed to complement her own expertise. She spent time with me to learn and eventually perform computational studies and adding new skillsets her research techniques. Her questions always challenged my thinking and helped me to take a more in-depth look at a research problem from multiple perspectives. I will always look up to her for inspiration. Furthermore, I would also like to thank members of Howell lab, Dr. Michael Duff Jr., Dr. Deepika Nambiar, Gabriel Fuentes and Seda Kocaman (from Serpersu Lab) for their support and feedback. I am especially thankful to Dr. Purva Bhojane for teaching me steady-state kinetics experiments. Regular discussion with Dr. Howell and her group helped me developed a profound appreciation for the complex and exciting field of enzyme catalysis.

Next, I would like to thank our collaborator Dr. Nicolas Doucet, for INRS in Canada, for offering valuable suggestions and critical insights into my research. He was kind enough to allow me to visit his lab in Canada and to work with his students on CPMG NMR data analysis. Multiple discussions with Doucet lab members, Dr. Chitra Narayana on computational aspect of research and with David Bernard regarding detailed analysis of NMR data, have proved very insightful and significant for successful completion of my doctoral dissertation.
I am also grateful to my committee members Dr. Hong Guo, Dr. Engin Serpersu, and Dr. Xiaolin Cheng for lending their expertise and time. Their suggestions have been helpful in the progress of my dissertation research.

A special note of thanks to Shantanu Shukla, for recognizing my talent and encouraging me to pursue a scientific career. With his passion and enthusiasm for science, he motivated me to cross traditional boundaries. He has been the most supportive and at the same time the most critical person of my work. I thank him for always being in the background and supporting me through the highs and lows of graduate school. He continues to encourage me in my pursuits to build a successful career.

This journey would not have been possible without the support of my family and friends despite the long distance between us. I am especially grateful to my parents for supporting my education and letting me follow my dreams. I am thankful to my brother for believing in me and always wanting the best for me. A special thanks to my friends in Knoxville and those scattered around the world, thank you for your well-wishes, phone calls, texts, visits, editing advice, and for just being there whenever I needed a friend. Lastly, I would like to thank all my school and college teachers for instilling in me the knowledge required to thrive in life.
Abstract

The enzymes of the ribonuclease (RNase) family catalyze the hydrolysis of ribonucleic acid (RNA). There is a wide interest in therapeutic interventions of human RNases (hRNases) due to their critical role in host defense, cancer cell proliferation and neurodegenerative diseases. Development of structure-based therapeutics targeting individual members of a closely related enzyme family is difficult. The structural conservation among hRNases cannot explain a million-fold difference in the catalytic efficiency of these enzymes. We hypothesize, this ambiguity in the structure activity relationship of hRNases can be explained by their dynamical behavior. The rate of substrate turnover in RNases correlates strongly with the rates of conformational dynamics. Moreover, catalytic efficiency is also linked to the ability of an enzyme to sample conformational sub-states that promote specific interaction between reactants at various stages of catalysis. Detailed characterization of the structure-function-dynamics relationship in hRNases can help in determining similarities and differences in the motions associated to catalysis thereby enhancing the possible druggability of this enzyme family.

This study investigates the role of functionally relevant conformational sub-states and dynamics in each step of the RNase members catalytic cycle; apo, substrate-bound, the chemical step, and product release. Computer simulations and nuclear magnetic resonance (NMR) relaxation dispersion experiments indicated an increased dynamics in distal loop regions of RNases in their apo state. Similar dynamical patterns were observed within the members of an individual sub-family, while dynamics were different across sub-families. Nucleotide binding properties probed using computer simulations indicated diverse binding preferences in hRNases. Additionally, structural and dynamical properties of hRNases varied significantly with subtle change in temperature. Quasi anharmonic analysis revealed sampling of separate conformational sub-states in product release step of RNases. Further, NMR chemical shift titrations along with the chemical shift projection analysis revealed distinct conformational rearrangements in hRNases upon binding of ligands that mimic cleaved products. Steady-state kinetics experiments showed million-fold difference in the catalytic efficiency of hRNases. Finally, hybrid quantum mechanical/molecular mechanics calculations identified conformational sub-states associated with the chemical step in bovine RNase A. Broadly, these results strongly support our hypothesis that dynamics modulates catalytic efficiencies of structurally related enzyme super-families like RNases.
Table of Contents

Chapter 1 Introduction ........................................................................................................ 1
  1.1 Enzymes and ribonuclease A superfamily ................................................................. 2
  1.2 Human RNases and their role in health ................................................................. 6
  1.3 Biochemical and biophysical characterization of RNases ........................................ 10
     1.3.A Substrate binding and specificity ...................................................................... 10
     1.3.B Catalytic mechanism of phosphodiester bond cleavage .................................... 14
     1.3.C Catalytic efficiency ......................................................................................... 18
  1.4 Role of dynamics in enzyme function ................................................................. 18
     1.4.A Dynamics in ribonucleases ............................................................................. 23
  1.5 Methods for studying protein dynamics ............................................................... 24
     1.5.A Molecular dynamics simulations ..................................................................... 24
     1.5.B Hybrid QM/MM ............................................................................................. 26
     1.5.C Quasi Anharmonic Analysis (QAA) ................................................................ 27
     1.5.D Carr-Purcell-Meiboom-Gill (CPMG) - NMR relaxation dispersion .................... 28
  1.6 Proposed study ........................................................................................................ 31

Chapter 2 Conservation of Dynamics Associated with Biological Function in an Enzyme Superfamily ........................................................................................................... 33
  2.1 Abstract .................................................................................................................. 35
  2.2 Introduction ............................................................................................................. 35
  2.3 Materials and methods ........................................................................................... 37
  2.4 Results .................................................................................................................... 41
     2.4.A Phylogenetic analysis of ribonuclease superfamily .......................................... 41
     2.4.B Conformational dynamics on the microsecond-millisecond (µs - ms) time-scale .... 42
     2.4.C Conformational dynamics on the microsecond (µs) time-scale ................................ 48
     2.4.D Quantitative characterization of dynamical conservation .................................... 49
     2.4.E Faster time-scale conformational dynamics in functionally distinct RNase subfamilies ......................................................................................................................... 56
## Chapter 3  Nucleotide Substrate Binding Characterization in Human Pancreatic-type Ribonucleases

3.1 Abstract ..............................................................................................................................70
3.2 Introduction ........................................................................................................................70
3.3 Materials and methods .......................................................................................................74
3.4 Results ...............................................................................................................................77
   3.4.A Stability of the substrate-bound RNase complexes ................................................77
   3.4.B Structural interactions ..............................................................................................79
   3.4.C Electrostatic interactions .........................................................................................88
   3.4.D Quantitative characterization of interaction energies .............................................89
   3.4.E Dynamical characterization of the enzyme-substrate complexes .........................98
3.5 Discussion .......................................................................................................................104
3.6 Summary of observations for each RNases investigated in this study ......................106
3.7 Appendix ........................................................................................................................110

## Chapter 4  Conformational Sub-states Associated with Product Release in Ribonucleases

4.1 Abstract ............................................................................................................................123
4.2 Introduction ......................................................................................................................123
4.3 Materials and methods .....................................................................................................126
4.4 Results .............................................................................................................................131
   4.4.A Evaluating the conformational coverage of umbrella sampling simulations ..........131
   4.4.B Computational Analysis of enzyme-product interactions ......................................132
   4.4.C Conformational dynamics associated product release ..........................................135
   4.4.D Conformational sub-states associated with product release ...............................140
   4.4.E Potential of Mean Force (PMF) .............................................................................140
4.5 Discussion .......................................................................................................................143
Chapter 5  Specialized Techniques to Identify Conformational Fluctuations Associated with Enzyme Catalysis ................................. 147

5.1 Abstract .............................................................................................................................................. 149
5.2 Introduction ......................................................................................................................................... 149
5.3 Materials and methods ................................................................................................................... 151
5.4 Results ................................................................................................................................................ 159
  5.4.A Elucidating the differences in kinetics parameters for human ribonucleases.............. 159
  5.4.B Ligand-induced variations in structural and dynamical properties of RNases........ 160
  5.4.C Conformational dynamics in the chemical step of bRNaseA ................................. 165
  5.4.D Conformational sub-states in the chemical step of bRNaseA function ............ 167
5.5 Discussion ......................................................................................................................................... 167

Chapter 6  Conclusions and Future Directions ................................................................. 171

6.1 Conclusions ......................................................................................................................................... 172
6.2 Future directions .................................................................................................................................. 177
6.3 Assessing the quality of computer simulations ......................................................................... 180

List of References .................................................................................................................................. 183

Vita ......................................................................................................................................................... 204
List of Tables

Table 1.1: Sub-site base preference in RNases. .................................................................12
Table 1.2: Summary of reported substrate binding affinities and enzyme activity. .............19
Table 2.1: PDB IDs of RNase homologs used for dynamical characterization .....................40
Table 2.2: Amino acid substitutions and their effect on the catalytic and biological activities of select RNase homologs ........................................................................................................61
Table 3.1: Stability of nucleotide substrates ACAC and AUAU in the active-site of RNases. ....80
Table 3.2: H-bonding properties of bovine and human pancreatic-type RNases at 300 K. ......86
Table 3.3: H-bonding properties of bovine and human pancreatic-type RNases at 310 K. ......87
Table 3.4: Averaged total interaction energy between nucleotide substrates and RNases. ....97
Table 3.5: Residues showing significant favorable interactions with the substrates at 300 K...110
Table 3.6: Residues showing significant favorable interactions with the substrates at 310 K...112
Table 4.1: H-bonding interaction of pancreatic-type RNases with product nucleotides. ........136
Table 4.2: Interaction of product nucleotides with hydrophobic amino acids .........................138
Table 5.1: Enzyme concentrations and range of substrate (UpA) concentrations used in steady-state kinetics experiments. ........................................................................................................152
Table 5.2: Comparing the steady-state kinetic parameters for RNases. ...............................159
Table 5.3: Chemical shift perturbations upon ligand binding ..............................................162
List of Figures

Figure 1.1: Schematic representation of Ribonuclease A function .................................................. 3
Figure 1.2: Multiple sequence alignment of RNases ....................................................................... 4
Figure 1.3: Structure and structural alignment of RNases family members. ................................. 5
Figure 1.4: Location of hRNases on human chromosome 14. ....................................................... 6
Figure 1.5: Nucleotide binding sub-sites in ribonucleases. ........................................................... 13
Figure 1.6: Mechanism of Ribonuclease function. ......................................................................... 15
Figure 1.7: Widely accepted general acid/base mechanism for transesterification step ............... 16
Figure 1.8: Calculated low energy path for tri-ester like mechanism ........................................ 17
Figure 1.9: Schematic representation of conformational landscape of protein ............................. 22
Figure 1.10: QAA describes conformational sub-states leading to transition state during catalysis in cyclophilin A .......................................................... 30
Figure 1.11: Schematic representation of the RNase catalytic cycle and proposed work ............. 32
Figure 2.1: Structural and phylogenetic analysis of RNase homologs ......................................... 43
Figure 2.2: Millisecond conformational exchange in Hominidae and bovine RNases .................. 45
Figure 2.3: Conformational exchange in HsR5 residues ............................................................... 46
Figure 2.4: Dynamical effects of loop 1 swapping on BtRA .......................................................... 47
Figure 2.5: Microsecond conformational dynamics of RNase homologs ..................................... 50
Figure 2.6: Dynamical properties of BtRAHsR3 chimera ............................................................. 52
Figure 2.7: Quantitative characterization of dynamical similarities ............................................. 54
Figure 2.8: Faster time-scale motions in representative RNases .................................................. 57
Figure 2.9: Conformational exchange rates of selected RNases ................................................... 64
Figure 2.10: Consensus sequence corresponding to the multiple sequence alignment .................. 65
Figure 2.11: μs dynamics of all 23 RNase homologs ................................................................. 66
Figure 2.12: Dynamic cross-correlation maps (DCCMs) for the 23 RNase homologs ............... 67
Figure 3.1: RNase-substrate binding sub-sites ............................................................................. 72
Figure 3.2: Substrate behavior over the course of 0.5 μs MD simulations ................................... 81
Figure 3.3: Substrate behavior over the course of 0.5 μs MD simulations at 310 K ..................... 82
Figure 3.4: Hydrogen-bonding interactions between RNases and model substrates at 300 K .... 84
Figure 3.5: Hydrogen-bonding interactions between RNases and model substrates at 310 K .... 85
Figure 3.6: Electrostatic surface of RNases with the two model substrates at 300 K ............... 90
Figure 3.7: Electrostatic surface of RNases with the two model substrates at 310 K ............... 91
Figure 3.8: Enzyme-substrate interaction energy at 300 K ....................................................... 93
Figure 3.9: Enzyme-substrate interaction energy at 310 K ........................................................95
Figure 3.10: Dynamical behavior of RNases bound to model substrates at 300 K .................. 100
Figure 3.11: Dynamical behavior of RNases bound to model substrates at 310 K .................. 102
Figure 3.12: Important H-bond behavior over time for bRNaseA-substrate complexes ......... 114
Figure 3.13: Important H-bond behavior over time for hRNase1-substrate complexes .......... 115
Figure 3.14: Important H-bond behavior over time for hRNase2-substrate complexes .......... 116
Figure 3.15: Important H-bond behavior over time for hRNase3-substrate complexes .......... 117
Figure 3.16: Important H-bond behavior over time for hRNase4-substrate complexes .......... 118
Figure 3.17: Important H-bond behavior over time for hRNase5-substrate complexes .......... 119
Figure 3.18: Important H-bond behavior over time for hRNase6-substrate complexes .......... 120
Figure 3.19: Important H-bond behavior over time for hRNase7-substrate complexes .......... 121
Figure 4.1: Schematic representation of ribonuclease reaction .............................................. 128
Figure 4.2: Representation of geometric center of mass (COM) of catalytic residues and products. .................................................................................................................................... 129
Figure 4.3: Schematics of product release umbrella sampling simulations .............................. 130
Figure 4.4: Overlap of histograms between umbrella sampling windows ............................... 133
Figure 4.5: Root mean square deviation of protein structures along the reaction coordinate..... 134
Figure 4.6: Conformational dynamics along the reaction coordinate ...................................... 139
Figure 4.7: Conformational sub-states associated with release of products in ribonucleases .. 141
Figure 4.8: Free energy profile of product release ................................................................. 142
Figure 4.9: Comparison of computational conformational sub-states and NMR conformational exchange .......................................................................................................................... 146
Figure 5.1: Illustration of the CHESPA approach ................................................................. 155
Figure 5.2: The NMR chemical shift projection analysis (CHESPA) ...................................... 155
Figure 5.3: Starting structures for enzyme-substrate and enzyme-products complexes ........ 156
Figure 5.4: Reactive center for reactant and product states .................................................... 158
Figure 5.5: Effect of ligand binding on functionally distinct RNases ................................... 161
Figure 5.6: Chemical shift projection analysis of ligand binding to RNases ......................... 164
Figure 5.7: Dynamical behaviour of bRNaseA in the chemical step ...................................... 166
Figure 5.8: Conformational sub-states in the chemical step of bRNaseA function ............... 168
Figure 6.1: Conformational dynamics in the catalytic cycle of bRNaseA ............................. 176
List of Abbreviations

3'-CMP  Cytidine 3'-Monophosphate
3'-UMP  Uridine 3'-Monophosphate
3'U    3'Uracil
5'-AMP  Adenosine 5'-Monophosphate
5'A    5'Adenine
ALS    Amyotrophic Lateral Sclerosis
AMBER Assisted Model Building and Energy Refinement
APBS   Adaptive Poisson-Boltzman Solver
BMRB   Biological Magnetic Resonance Bank
bRNaseA bovine ribonuclease A
CEST   Chemical Exchange Saturation Transfer
CHARMM Chemistry at HARvard Macromolecular Mechanics
CHESPA Chemical Shifts Projection Analysis
COM    center of mass
CPMG   Carr Purcell Meiboom Gill
CPU    Central Processing Unit
CypA   Cyclophilin A
DFT    Density Function Theory
DHFR   dihydrofolate reductase
DNA    deoxyribonucleic acid
dsRNA  double stranded ribonucleic acid
EAR    eosinophil associated ribonucleases
ECP    eosinophil cationic protein
EDN    eosinophil-derived neurotoxin
$E_{el}$ electrostatic energy
$E_{tot}$ total interaction energy
$E_{vdw}$ van der Waal’s energy
fs     femtoseconds
GPU    Graphical Processing Unit
GROMOS GROningen MOlecular Simulation
H-bonding hydrogen bonding
HIV    Human Immunodeficiency Virus
hRNase  human ribonuclease
hRNase1  human ribonuclease 1
hRNase2  human ribonuclease 2
hRNase3  human ribonuclease 3
hRNase4  human ribonuclease 4
hRNase5  human ribonuclease 5
hRNase6  human ribonuclease 6
hRNase7  human ribonuclease 7
HSQC  Heteronuclear Single Quantum Coherence
$k_{cat}$  turnover number
$k_{cat}/K_M$  catalytic efficiency
kDa  kilodalton
$k_{ex}$  rate constant of chemical exchange
$K_M$  Michaelis constant
$k_{off}$  rate constant for dissociation
MD  Molecular Dynamics
ms  milliseconds
NEB  nudged elastic band
NMR  nuclear magnetic resonance
ns  nanoseconds
OPLS  Optimized Potential for Liquid Simulations
ORF  open reading frame
PDB  protein data bank
pI  isoelectric point
PME  Particle Mesh Ewald
PMF  Potential of Mean Force
ppm  parts per million
ps  picoseconds
QAA  Quasi Anharmonic Analysis
QHA  Quasi Harmonic Analysis
QM/MM  Quantum Mechanics / Molecular Mechanics
RMSF  Root Mean Square Fluctuation
RNA  ribonucleic acid
RNase  ribonuclease
RSV  Respiratory Syncytial Virus
SCC-DFTB  Self-Consistent Charge Density Functional Tight-Binding
ssRNA  single stranded ribonucleic acid
TS  transition state
WHAM  Weighted Histogram Analysis Method
$\Delta\delta_{\text{obs}}$  compounded chemical shift changes
$\mu$s  microseconds
Chapter 1  Introduction
1.1 Enzymes and ribonuclease A superfamily

Enzymes are biological macromolecules that catalyze cellular reactions essential to sustain life. Enzyme have the ability to enhance the rate of a chemical reaction by more than $10^{20}$ folds, making them the most efficient catalysts occurring naturally. The molecules upon which enzymes act are called as substrates; with each enzyme recognizing and acting upon a defined list of substrates. Upon completion of an enzymatic reaction, substrate(s) are converted into product(s). Enzymatic reactions occur in a cycle of multistep process, called as the catalytic cycle, which includes substrate(s) binding, chemical step and product(s) release. All these steps can occur within a few milliseconds or even less. During an enzymatic reaction the smallest fraction of time is spent in the most important part of the catalytic cycle, called the transition state (TS). TS is the highest energy transitory structure in a reaction which is no longer a substrate but not a product as yet. The energy required to reach the TS is the minimum energy required to convert a reactant into product, in this case also called as the activation energy. Compared to reaction in water or gas-phase, enzymes decrease activation energy by shaping their active site such that it binds to the transition state even better than the substrate. Most enzymes are tailored to bind a specific substrate(s) and perform a specific function (reaction). Enzymes' specificity comes from their unique three-dimensional (3D) structures. Enzymes are proteins, and their structure comprises of unbranched chain of amino acids that adopts a specific three-dimensional fold. Enzymes that share a common fold and catalyze same/similar reaction belong to a same protein family. A number of protein families together form a super family. In other words, protein super family refer to a group of structurally or functionally related proteins that may or may not have a common evolutionary origin.

Ribonuclease A superfamily is a group of enzymes that catalyze the hydrolysis/degradation of ribonucleic acid (RNA) molecules into smaller components. Ribonucleases A enzymes specific to vertebrates are found in high quantity in the pancreas of certain mammals and some reptiles, and therefore are also called as pancreatic-type ribonuclease superfamily, hence forth called as RNase family. RNases have a marked preference for cleavage of RNA at the P-O bond on the 3’ side of the pyrimidine residue and do not require any metal ion or cofactor (Figure 1.1). The sequence identity among the RNase family members is sometime as low as ~20%. However, two histidine and one lysine residues essential for the RNase function are conserved (Figure 1.2). Structural information for several RNase homologs has been obtained by X-ray crystallography. The members of RNase superfamily are small proteins with ~130 amino acids.
(~13-15 kilodaltons, kDa). In addition, they generally also have a short signal peptide of ~25 amino acids. Signal peptides are short segments usually at the N-terminal end of a newly synthesized protein, with a function to prompt the cell to translocate the synthesized protein usually to cell membranes. Even with a low primary sequence identity RNases adopt a common $\alpha+\beta$ kidney shaped fold (Figure 1.3 A). The protein secondary structure elements ($\alpha$ helix and $\beta$ sheets) are conserved, and variations can be observed only in loop regions (Figure 1.3 B). RNases have 6 to 8 cysteine residues in their primary sequence that form three to four disulfide bonds in their three-dimensional structure (Figure 1.3 A). Two antiparallel $\beta$-sheets define the $V_1$ and $V_2$ domains linked by a hinge region in their kidney shaped structure. They have a well-defined binding cleft between the two $V_1$ and $V_2$ domains.

**Figure 1.1: Schematic representation of Ribonuclease A function.** RNA substrate is shown on left. The phosphodiester bond cleaved during RNase reaction is marked with an arrow. The cleave products formed after the RNase reaction are shown on right.
**Figure 1.2: Multiple sequence alignment of RNases.** Sequence for human RNases and bovine RNase A is included. Fully conserved residues are highlighted in red boxes. Catalytic residues are marked with blue stars. Gaps in the alignment are represented using dotted lines. Consensus sequence number is shown at the top and bovine RNase A sequence numbering is shown at the bottom. Sequence alignment was performed in *Clustal Omega* \(^\text{10}\) and rendered in *ESPript* \(^\text{11}\)
Figure 1.3: Structure and structural alignment of RNases family members. (A) Ribbon diagram of the crystal structure of bovine RNase A (PDB ID 7RSA); cysteine residues are represented as purple sticks and the disulfide bonds are shown in yellow. V₁ and V₂ domain is represented in dashed line. (B) Structural alignment of representative RNases. Structures include 1QMT (violet), 2K11(orange); 1ANG (cyan) and 1RNF (green). Catalytic residues are displayed as sticks and labeled. Loops L1, L3, L4, L6 and L7 are marked.
1.2 Human RNases and their role in health

RNase superfamily has remarkably high rates of gene duplication and gene loss, resulting in variable numbers of genes in different species.\textsuperscript{12} In humans, there are thirteen RNase type genes located on chromosome 14 within cluster 14q11.2 (Figure 1.4).\textsuperscript{13} Eight of these human RNases (hRNases 1-8) have the RNases function (i.e. can hydrolyze RNA molecule) and are referred to as canonical RNases. The other five open reading frames (ORFs), i.e. RNase 9-13 encode proteins that lacks the catalytic residues (two histidine and one lysine) essential for RNase function.\textsuperscript{5} The canonical hRNases are secretory protein, i.e. they are released from the cells where they were produced and translocated to target cells when they perform their designated function. These proteins are secreted by various innate immune cells, from blood cells to epithelial cells.\textsuperscript{14} They exhibit diverse expression pattern and are widely distributed in various organs, tissues and body fluids. The levels RNases in our body fluids correlate with infection and inflammation processes ensuring tissue health and body homeostasis.\textsuperscript{15} Apart from hydrolysis of dietary RNA, these enzymes participate in a wide variety of biological functions, including neurotoxicity, angiogenesis (blood vessel formation), immuno-modulatory, anti-pathogen and anti-tumor activities.\textsuperscript{8, 16, 17} Described below are the highlights of common features of the eight canonical hRNases with regard to their respective sequences, structure, biochemical characterization, regulation and role in human health and disease.

![Figure 1.4: Location of hRNases on human chromosome 14. hRNase1-8 perform ribonuclease function and degrade RNA molecules. hRNases 9-13 lack the residues required for enzymatic function and therefore do not catalyze RNase reaction. hRNase10 is towards the centromere; hRNase8 is towards the telomere; ps (pseudogene). (Adapted from \textsuperscript{5})](image-url)
**Human RNase 1 (hRNase1)** is expressed in a wide variety of human organs and tissues. Apart from hydrolysis of RNA, hRNase1 has significant physiological implications in the regulation of vascular homeostasis due to its wide extracellular distribution. In addition to cleavage of single-stranded RNA (ssRNA) this enzyme has been reported to hydrolyze double-stranded RNA (dsRNA) molecule as well as DNA (deoxyribonucleic acid):RNA hybrids. Due to its catalytic versatility, hRNase1 is reported to act as an extracellular RNA scavenger and inactivates human immunodeficiency virus (HIV), participate in inflammatory and immune response by promoting maturation and activation of dendritic cells as well as subsequent production of different cytokines. Evidence indicate that hRNase1 important for normalization of serum viscosity and clearance of perivascular pathogenic polynucleotides. Further, it is reported that hRNase1 enters human cells and catalyze the cleavage of RNA, leading to cell death. Utilizing this ability, Raines and co-workers have demonstrated the potential of hRNase1 as a chemotherapeutic agent that is toxic to tumor cells.

**Human RNase 2 (hRNase2),** also known as eosinophil-derived neurotoxin (EDN) is found in spleen, liver, kidney, placenta, and urine. It has also been found in the secondary granules of eosinophil granulocytes. hRNase2 shares ~32% primary structure identity to hRNase1. It is reported to be inactive towards poly (A) and dsRNA and does not hydrolyze 2':3'-cyclic nucleotides. hRNase2/EDN show neurotoxic potency and has a specific antiviral action against respiratory syncytial virus (RSV), HIV-I and hepatitis B DNA virus. Its anti-viral activity is reported to be dependent on its ribonucleolytic activity. However, it does not display any bactericidal activity. Studies also suggest that hRNase2 might trigger an inflammatory response that activates killer cells or cell death pathways. Alternatively, hRNase2 can participate in further modulating the immune response by degrading the viral RNA in the cytoplasm of infected cells.

**Human RNase 3 (hRNase3),** also known as eosinophil cationic protein (ECP) is the other major secretory protein of the human eosinophil granule and is released upon cell activation/stimulation. hRNase3 shares ~26% primary structure identity to hRNase1 while it is ~67 % similar to hRNase2. It is about two orders of magnitude less active on yeast tRNA as compared to hRNase1 and hRNase2. hRNase3 displays anthelminthic, bactericidal, antiviral and antiparasitic activities. The ribonucleolytic activity of hRNase3 does not appear to be necessary for cytotoxicity or bactericidal activity, while it is essential of the antiviral activity. hRNase3’s high toxic activity has been demonstrated in vitro against Schistosoma mansoni, Plasmodium falciparum, Brugia pahangi, and Trichinella spiralis. The bactericidal action of
hRNase3/ECP is reported to correlate with its membrane disruption capability and is linked to aromatic and cationic residues located on the surface of the protein. Although not completely understood, a stepwise mechanism of hRNase3 antibacterial activity is proposed. hRNase3 binds to the surface of bacteria promoting conformational changes in the protein. All rearranged hRNase3 molecules bind to each other leading to agglutination of bacteria. The protein then forms amyloid-like aggregation that disrupts the lipopolysaccharide bilayer, resulting in possible membrane disruption in bacteria.

**Human RNase 4 (hRNase4),** is detected in several somatic tissues including liver, pancreas, lung, heart, kidney, and placenta. Cytoplasmic granules of monocytes also express this enzyme. This enzyme has the shortest amino acid sequence (119 residues) in this family and shares about 90% sequence identity with bovine ribonuclease A (bRNaseA), the prototypical member of the RNase superfamily. It is structurally more like hRNase1 (43% identity) than to hRNase2 (31% identity) but shares some catalytic properties with both RNase types. Like hRNase1 it shows hydrolytic activity towards 2’:3’-cyclic nucleotides. Although considerable work has been undertaken to interpret its biological properties, its specific biological function remains unknown. Nearly ubiquitous distribution of hRNase4 suggests a housekeeping role for this enzyme. Interestingly, hRNase4 has also been found in the soluble factors secreted by T-cells showing anti-HIV activity further suggesting its involvement in host defense activities. Additionally, it is reported that hRNase4 prevent the degeneration of neuron by promoting angiogenesis, neurogenesis, and promotes the neuronal survival under stress. Expression of hRNase4 is reduced in amyotrophic lateral sclerosis (ALS) affected patients, however a direct link between ALS and hRNase4 is yet to be established.

**Human RNase 5 (hRNase5),** was first isolated from tumor cells and later found in a wide variety of tissues and organs. This enzyme is a potent inducer angiogenesis, and therefore it is also called as angiogenin. hRNase5 is structurally similar to bRNaseA and hRNase1 however it has a million-fold lower ribonucleolytic activity than that of bRNaseA. Although hRNase5 exhibits extremely weak ribonucleolytic activity, it is still essential for angiogenesis. Studies report that hRNase5 undergoes nuclear translocation to stimulate blood vessel growth and plays an essential role in cell proliferation. A recent work correlated elevated hRNase5 expression in some tumor cell lines with the promotion of cell proliferation and development of malignant cancer. An association between ANG and cancer has been observed in more than 20 clinical studies to date. hRNase5 is also reported to be involved in motor neuron physiology, and genetic deficiencies in
hRNase5 can lead to immune-related diseases, such as ALS. Recent studies report the involvement of hRNase5 in cell defense against stress and microbicidal activities.

**Human RNase 6 (hRNase6),** also referred to as RNase K6, is found in neutrophils and monocytes and therefore proposed to have host defense functions. The amino acid sequence of hRNase6 is ~43% identical to hRNase2 and ~38% identical to hRNase1. Its ribonucleolytic activity towards yeast tRNA is 40 folds lower than hRNase2. RNase6 has been reported to display a high antimicrobial activity in vivo and it showed very low cytotoxicity levels against mammalian cells. It is also suggested to play an essential role in bacterial clearance at the urinary tract. Boix et al. recently proposed the antimicrobial mechanism of action. Highly cationic hRNase6 first interacts with the negatively charged envelopes of the bacterial cell, subsequently leading membrane destabilization and bacterial agglutination. The investigators also found that the N-terminal domain of the protein retains the antimicrobial properties of the whole protein. This suggests the possibility of using the scaffold of hRNase6 for developing peptide-based antimicrobial agents in the future.

**Human RNase 7 (hRNase7),** is expressed in various epithelial tissues and organs. It was detected in skin as well as in the respiratory tract, genitourinary tract, liver, kidney, skeletal muscle, and heart. The catalytic activity of hRNase7 is 50 fold higher hRNase3 and 8 fold lower than hRNase1. hRNase7 is reported to have potent bactericidal activity against several Gram-positive and Gram-negative bacteria. It is therefore called as epithelial-derived antimicrobial protein. Recent studies have reported the ability of hRNase7 to permeate the bacterial membranes. Specific lysine residues (K1, K3, K111, and K112) were identified to be crucial for its antimicrobial mechanism of action, and ribonucleolytic activity was not essential for its antibacterial action.

**Human RNase 8 (hRNase8),** is the most recently discovered active human RNase and is expressed uniquely in the placenta. It is reported to have antimicrobial activity against various bacteria including *Candida albicans* and does not affect the human cells, suggesting that it plays a role in protecting the placenta against infections. Its reported ribonucleolytic activity against yeast tRNA is two folds lower than that of hRNase7. At present, no information regarding structure or catalytic properties are available for this enzyme.
1.3 Biochemical and biophysical characterization of RNases

The physical, chemical and enzymatic properties of the bovine RNase (bRNaseA) have been extensively characterized. However, comparatively lesser information is available for recently discovered human RNases. In this section, the biochemical and biophysical properties bRNaseA and hRNases is summarized. bRNaseA is a catanionic protein (pI = 9.63) at physiological pH. It has several positively charged residues (10 lysine and 4 arginine residues) that enable its binding to RNA (a polyanion). In other words, bRNaseA has smaller number of hydrophobic groups, especially aliphatic ones. Its low hydrophobic content may also serve to reduce the physical repulsion between highly charged groups (its own and those of its substrate RNA) and regions of low dielectric constant (the nonpolar residues). Similar to bRNaseA, hRNases also have high isoelectric points and net positive charges (with pI values between 8.69 – 10.12). These properties are associated with antibacterial activity and strong interactions with their negatively charged polynucleotide substrates. However, each hRNases varies in its nucleotide substrate preference and ribonucleolytic activity.

1.3.A Substrate binding and specificity

Binding of single-stranded nucleotides to bRNaseA involves up to seven coulombic interactions suggesting that the enzyme-substrate interaction extends well beyond the scissile bond. Structural and kinetic studies, using different oligonucleotides substrates and polynucleotides (as homologous to the RNA substrate) reveal the existence of several substrate binding sub-sites that make up the active site in the structure of RNases as shown in Figure 1.5. The amino acid residues that form the sub-sites for bRNaseA have been well defined. The B1 sub-site consisting of residues Thr45 and Asp83 appear to bind only pyrimidine bases. Thr45 sterically hinders the binding of purine base in B1 sub-site and mediates pyrimidine specificity by forming a hydrogen bond. Thr45 can act as a donor or acceptor depending upon the binding nucleotide, and the conformation of Asp83 is responsible for switching this role of Thr45. Additionally, Phe120 contributes to pyrimidine binding in B1 sub-site by forming a van der Waals contact and acts as a hydrophobic mattress for the base. Phe120 also plays an important role in stabilizing the TS intermediate in catalysis by forming a hydrogen bond with non-bridging oxygen of the reactive phosphoryl group. The B2 sub-site, comprising of residues Gln69, Asn71 and Glu111, has a preference of adenine base. The difference in the specificities of B1 and B2 sub-sites makes RNases a distributive enzyme, that binds a polymeric substrate, catalyzes a chemical reaction and releases a polymeric product to the solvent.
Three other enzymatic sub-sites $P_0$, $P_1$, and $P_2$ interacts with the phosphoryl groups of the bound RNA substrate. The $P_1$ sub-site comprising of residues Gln11, His12, Lys41, His119, and Asp121 forms the active site in bRNaseA (Figure 1.5). The phosphoryl group bound at the $P_1$ site participates in the reaction mechanism. Site-directed mutagenesis has shown that residues Lys7, Arg10, Gln11, His12, Lys41, Thr45, Lys66, Asn71, Asp83, Glu111, His119, Phe120, and Asp121 contribute towards substrate binding or turnover in bRNaseA. Substitution of these residues has resulted in reduced catalytic activity by $\sim 10^4$ fold.7

Sequence alignment for canonical hRNases and bRNaseA show that the residues of main phosphate binding $P_1$ sub-sites sites are conserved (Figure 1.2) in all members. The threonine residue of $B_1$ sub-site and asparagine of $B_2$ sub-site are also conserved. However, variability is observed in other residues of base binding sub-sites which attributes to substrate preference and cleavage pattern in RNase homologs. Further, nucleotide binding recognition patterns for hRNases has been deduced based on the available protein-ligand complexes in the Protein Data Bank (PDB). The base preference of select hRNases and bRNaseA is reported in Table 1.1. bRNaseA and hRNase1, has a higher preference for cytosine containing substrates. It is reported that bRNaseA can hydrolyze cytosine containing substrates at-least 2 folds faster than uridylyl substrates.7 Table 1.1 reports that hRNases 2 - 4 have a clear preference for uracil containing substrates. Asn71 (bRNaseA numbering) can contribute as both acceptor and donor while bonding to adenine base, and all hRNases share this purine recognition trait. However, variations are observed in pyrimidine binding. In hRNase4 the aspartate residue of $B_1$ sub-site is not flexible to move and therefore locks the threonine residue as a hydrogen bond acceptor facilitating interaction mainly with uracil nucleotide. This substitution of Arg101 (in place of Lys104 in bRNaseA) prevents the binding of cytosine base in $B_1$ sub-site of hRNase4. Substitution of this lysine residue is also found in hRNase2 and 3 increasing their preference for uracil base.71 In hRNase5, Gln117 is oriented towards the active site in the solved crystal structure and hinders the access of the substrate to the pyrimidine binding site suggesting the requirement of conformational rearrangement to allow the main base binding. Moreover, specific orientation of catalytic residues is not favorable for cleavage of RNA resulting in million-fold lower catalytic efficiency in hRNase5. Therefore, non-conserved substitutions at binding sub-sites might explain the distinct substrate specificities in these homologs reported in Table 1.1. These observations should, however, be considered with caution; as many complex structures were solved in non-physiological conditions and limited data was available for pyrimidine bases and some homologs.
Table 1.1: Sub-site base preference in RNases. Preference of B₁ and B₂ sub-sites is deduced from crystal structure and kinetic studies (Adapted from ⁹)

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrates investigated</th>
<th>Preference at B₁ sub-site</th>
<th>Preference at B₂ sub-site</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>hRNase1</td>
<td>Poly(C), Poly(U)</td>
<td>C &gt; U</td>
<td>A</td>
<td>⁷⁴</td>
</tr>
<tr>
<td>hRNase2</td>
<td>UpA, UpG Poly(C), Poly(U)</td>
<td>U &gt; C</td>
<td>A &gt; G</td>
<td>⁷⁴, ⁷⁵</td>
</tr>
<tr>
<td>hRNase3</td>
<td>CpA, UpA, UpG Poly(C), Poly(U)</td>
<td>U &gt; C</td>
<td>A</td>
<td>⁷⁴, ⁷⁶</td>
</tr>
<tr>
<td>hRNase4</td>
<td>UpA, CpA Poly(C), Poly(U)</td>
<td>U &gt;&gt;&gt; C  (10:1)</td>
<td>A</td>
<td>⁴⁶, ⁷⁴</td>
</tr>
</tbody>
</table>
**Figure 1.5: Nucleotide binding sub-sites in ribonucleases.** Either uracil or cytosine can bind to B₁ sub-site. B₂ sub-site has a preference for adenosine in RNase A family. The catalytic residues form the phosphate binding sub-site P₁. The phosphodiester bond cleave during RNase reaction is marked with an arrow. bRNaseA residues that interact with the substrate RNA molecule at distinct sub-sites are shown. (Adapted from 7).
1.3.B Catalytic mechanism of phosphodiester bond cleavage

RNase catalyze the cleavage of phosphodiester bond (P-O5') found in the RNA backbone. The mechanism of catalysis by RNases has been studied by a number of biochemical and biophysical approaches.\(^7\), \(^79\)-\(^84\) The overall RNase reaction have been proposed to involve two separate steps. The first (transphosphorylation) step, 3',5'-phosphodiester bond of RNA molecule is cleaved to give an oligonucleotide terminating in a pyrimidine 2',3'-cyclic nucleotide and a 5' hydroxyl terminated fragment; in the second (hydrolysis) step this 2',3'-cyclic nucleotide is hydrolyzed, by water attack, and a terminal 3'-nucleotide is formed (Figure 1.6).\(^7\) Two types of mechanisms for the transphosphorylation step have been extensively discussed in the literature.\(^7\), \(^85\)-\(^89\) The first is the widely accepted classical mechanism that follows a minimal acid/base mechanism with a di-anionic TS (Figure 1.7).\(^88\), \(^89\) The second type of mechanism that has been proposed is based on a tri-ester mechanism leading to a monoanionic TS (Figure 1.8).\(^86\) This mechanism has received less acceptance. The difference in the two proposed mechanisms of transphosphorylation is the destination of protons as well as the sequence in which they are transferred.\(^85\)

The widely accepted mechanism of transphosphorylation step proposes that the reaction is catalyzed by concerted general acid/base catalysis by His12 and His119 (in bRNaseA).\(^90\), \(^91\) The imidazole group of His12 acts as a base that abstracts a proton from the 2'-O of a substrate molecule. This nucleophilic oxygen attacks the electrophilic phosphorous and forms a pentavalent intermediate. Lys41 directly interacts with the phosphate moiety and stabilizes the pentavalent TS intermediate.\(^7\), \(^92\) The imidazole group of His119 acts as an acid for the cleavage reaction and protonates the 5'O of the leaving group.

An alternative tri-ester mechanism has also been proposed.\(^93\) Evidence from experimental observations and gas-phase quantum chemistry calculations support this method.\(^94\) In this mechanism, a proton originating via migration from the O2' hydroxide coordinates with one of the nonbridging phosphate oxygen rendering the phosphate neutral, which facilitates nucleophilic attack by the 2'-alkoxide to form a phosphate tri-ester like intermediate structure. After the formation of the trigonal bipyramidal phosphorane structure, the catalytic proton is abstracted from the nonbridging oxygen by the catalytic base His119, leading to the expulsion of the leaving group. Such a mechanism has also been proposed as the reaction path for a similar Ribonuclease T1 enzyme.\(^95\)
Figure 1.6: Mechanism of Ribonuclease function. bRNaseA catalyzed (A) Transphosphorylation of 3',5'-phosphodiester bond of RNA molecule leading to a 2',3'-cyclic nucleotide and 5' hydroxyl terminated fragment. (B) Hydrolysis of 2',3'-cyclic nucleotide to a terminal 3'-nucleotide is shown. (Adapted from 7)
Figure 1.7: Widely accepted general acid/base mechanism for transesterification step.

The reactant, product and transitions state (TS) are shown. His12 acts as a general base and His119 acts as a general acid. Catalytic residues are numbered according to bRNaseA (This figure is adapted from 85)

After the transphosphorylation step each histidine residues is protonated appropriately to catalyze the hydrolysis (Figure 1.6 B) of the bound cyclic intermediate. In the second hydrolysis step, the 2',3' cyclic phosphate is hydrolyzed to a 3' terminal nucleotide.96 The catalytic roles of His12 and His119 are now reversed. His119 activates the attack of water by acting as a general base while His12 protonates the leaving group acting as a general acid. His119 abstracts the proton from a water molecule, activating it for nucleophilic attack. The activated water molecule then attacks the cyclic phosphate resulting in the cleavage of the 2',3' cyclic phosphate intermediate. Simultaneously, His12 donates a proton to the leaving group, which is the 3' oxygen of the cyclic intermediate. Finally, the truncated nucleotide with a 3' terminal phosphate group is released. After the process of hydrolysis, the histidine residues return to their initial protonation states completing the catalytic cycle. bRNaseA, however, can release the cyclic intermediate rather than hydrolyzing it.96, 97 Under such instances, the protonation states of the two histidine residues are interconvert by an iso mechanism that does not involve substrate molecule.
Figure 1.8: Calculated low energy path for tri-ester like mechanism. This pathway is calculated based on a first principle quantum mechanics/molecule mechanics (QM/MM) with density function theory (DFT/B3LYP) level of electronic structure calculations. Catalytic residues are numbered according to bRNaseA (This figure is adapted from 85).
1.3.C Catalytic efficiency

The rate of enzymatic reaction has been measured for RNases using steady-state kinetics experiments. The kinetic parameters; the rate of catalytic turnover ($k_{cat}$), the amount of substrate needed for the enzyme to obtain half of its maximum rate of reaction ($K_M$) and catalytic efficiency (ratio of $k_{cat}/K_M$) for bRNaseA and hRNases is reported in Table 1.2. The data in Table 1.2 indicates that binding affinities differ over several orders of magnitude between these RNases. Further, catalytic efficiencies for these homologs differ by as much as $10^5$-fold. However, these kinetic assays have been performed under different experimental conditions (pH, temperature, buffer) and using variety of substrates with different nucleotide sequences with variable lengths resulting in diverse binding affinities and catalytic efficiencies. Therefore, these differences should be carefully considered when comparing the efficiency in these homologs.

1.4 Role of dynamics in enzyme function

The role of enzyme structure as a scaffold to facilitate the catalyzed reaction is well established. The active site of enzyme provides a complementary structural and electrostatic environment to preferentially bind specific substrates and/or cofactors and subsequently stabilize the TS allowing the reactions to proceed at a much faster rate than in solution. This perspective, however, fails to explain many phenomena observed in enzyme structure and function; First, the allosteric and cooperative effects of residues outside the active site; many studies report significant changes in catalytic efficiency of enzymes pertaining to mutations in distal residues. Second, the impact of bulk and hydration shell in enzyme activity. Third, large rate-enhancement achieved by structurally similar enzymes, for example, the members of the RNase family share conserved structural scaffold and catalytic machinery, yet their catalytic efficiency differs by a million folds (Table 1.2). And finally, efforts towards designing proteins with active sites environment similar to natural enzymes have not been very successful. Therefore, the rate enhancements by enzymes and their reaction mechanisms cannot be fully explained by the structural effects alone. In addition to the role of direct structural interactions, several other factors contribute to catalysis for example networks of hydrogen bond and long-range effects.

The emergence of new information has allowed enzymes to be viewed beyond a single rigid structural scaffold. Evidence from multiple experimental and computational techniques show that enzymes are flexible and exist as an ensemble of inter-converting structures called as
Table 1.2: Summary of reported substrate binding affinities and enzyme activity. In addition to hRNases 1-6, the parameters for bRNaseA are also shown for comparison. tRNA denotes substrate of non-specific sequence obtained from cellular tRNA.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Substrate</th>
<th>$K_M$ (mM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_M$ (s$^{-1}$M$^{-1}$)</th>
<th>Temp. (K)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>hRNase1</td>
<td>Poly(C)</td>
<td>0.10 ± 0.013</td>
<td>2416.67 ± 33.3</td>
<td>1.48 ± 0.11</td>
<td>310</td>
<td>105</td>
</tr>
<tr>
<td>hRNase1</td>
<td>tRNA</td>
<td>4.0</td>
<td>64.8</td>
<td>0.0.162</td>
<td></td>
<td>106</td>
</tr>
<tr>
<td>hRNase2</td>
<td>tRNA</td>
<td>0.0007</td>
<td>0.91</td>
<td>0.13</td>
<td></td>
<td>107</td>
</tr>
<tr>
<td>hRNase2</td>
<td>tRNA</td>
<td>0.00336 ± 0.00015</td>
<td>31.37 ± 0.96</td>
<td>9.34 ± 0.17</td>
<td>310</td>
<td>108</td>
</tr>
<tr>
<td>hRNase2</td>
<td>tRNA</td>
<td>0.0128</td>
<td>3.7</td>
<td>3.2</td>
<td></td>
<td>109</td>
</tr>
<tr>
<td>hRNase2</td>
<td>Poly(U)</td>
<td>80 ± 7</td>
<td>190 ± 11</td>
<td></td>
<td></td>
<td>108</td>
</tr>
<tr>
<td>hRNase2</td>
<td>tRNA</td>
<td>2.5 ± 0.21</td>
<td>20 ± 0.65</td>
<td></td>
<td></td>
<td>108</td>
</tr>
<tr>
<td>hRNase2</td>
<td>tRNA</td>
<td>0.0036 ± 0.00015</td>
<td>31.37 ± 0.96</td>
<td>9.34 ± 0.17</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>hRNase2</td>
<td>Poly(U)</td>
<td>53.05 ± 4.04</td>
<td>1248.93 ± 12.10</td>
<td>4.79 ± 0.52</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>hRNase3</td>
<td>CpA</td>
<td>1.7±0.3</td>
<td>0.55±0.06</td>
<td>0.000323</td>
<td>298</td>
<td>57</td>
</tr>
<tr>
<td>hRNase3</td>
<td>UpA</td>
<td>2.7±0.66</td>
<td>1.22±0.12</td>
<td>0.000447</td>
<td>298</td>
<td>57</td>
</tr>
<tr>
<td>hRNase3</td>
<td>C &gt; p</td>
<td>3±0.53</td>
<td>0.0032±0.00051</td>
<td>1.07×10$^{-6}$</td>
<td>298</td>
<td>57</td>
</tr>
<tr>
<td>hRNase3</td>
<td>C &gt; p</td>
<td>1.5</td>
<td>0.014</td>
<td>0.000011</td>
<td>298</td>
<td>110</td>
</tr>
<tr>
<td>hRNase3</td>
<td>U &gt; p</td>
<td>1.0</td>
<td>0.0043</td>
<td>0.000004</td>
<td>298</td>
<td>110</td>
</tr>
<tr>
<td>hRNase3</td>
<td>CpA</td>
<td>2.4</td>
<td>4.2</td>
<td>0.001750</td>
<td>298</td>
<td>110</td>
</tr>
<tr>
<td>hRNase3</td>
<td>UpA</td>
<td>5.4</td>
<td>6.2</td>
<td>0.001150</td>
<td>298</td>
<td>110</td>
</tr>
<tr>
<td>hRNase3</td>
<td>(Up)2 U &gt; p</td>
<td>1.4</td>
<td>0.56</td>
<td>0.000400</td>
<td>298</td>
<td>110</td>
</tr>
<tr>
<td>hRNase3</td>
<td>(Up)3 U &gt; p</td>
<td>0.7</td>
<td>1.2</td>
<td>0.001714</td>
<td>298</td>
<td>110</td>
</tr>
<tr>
<td>hRNase3</td>
<td>(Up)4 U &gt; p</td>
<td>0.17</td>
<td>1.4</td>
<td>0.008235</td>
<td>298</td>
<td>110</td>
</tr>
<tr>
<td>hRNase3</td>
<td>ytRNA</td>
<td>0.0041</td>
<td>0.0024</td>
<td>0.000590</td>
<td>298</td>
<td>110</td>
</tr>
<tr>
<td>hRNase3</td>
<td>Poly(U)</td>
<td>253 ± 10</td>
<td>36 ± 0.4</td>
<td>0.00014 ± 0.000005</td>
<td>111</td>
<td></td>
</tr>
<tr>
<td>hRNase4</td>
<td>UpA</td>
<td></td>
<td>0.25</td>
<td></td>
<td>298</td>
<td>112</td>
</tr>
<tr>
<td>hRNase4</td>
<td>CpA</td>
<td>0.00066</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hRNase5</td>
<td>CpA</td>
<td>0.000122 ± 0.0000002</td>
<td></td>
<td>298</td>
<td>69, 113</td>
<td></td>
</tr>
<tr>
<td>hRNase5</td>
<td>CpG</td>
<td>0.000004 ± 0.0000001</td>
<td></td>
<td>298</td>
<td>69, 113</td>
<td></td>
</tr>
<tr>
<td>hRNase5</td>
<td>CpC</td>
<td>0.000013 ± 0.0000001</td>
<td></td>
<td>298</td>
<td>69, 113</td>
<td></td>
</tr>
<tr>
<td>hRNase5</td>
<td>CpU</td>
<td>0.000006 ± 0.0000001</td>
<td></td>
<td>298</td>
<td>69, 113</td>
<td></td>
</tr>
<tr>
<td>hRNase5</td>
<td>UpA</td>
<td>0.000007 ± 0.0000001</td>
<td></td>
<td>298</td>
<td>69, 113</td>
<td></td>
</tr>
<tr>
<td>hRNase5</td>
<td>UpG</td>
<td>0.0000013 ± 0.0000002</td>
<td></td>
<td>298</td>
<td>69, 113</td>
<td></td>
</tr>
<tr>
<td>hRNase5</td>
<td>UpC</td>
<td>0.0000006 ± 0.0000001</td>
<td></td>
<td>298</td>
<td>69, 113</td>
<td></td>
</tr>
<tr>
<td>hRNase5</td>
<td>UpU</td>
<td>0.00000031 ± 0.000000007</td>
<td></td>
<td>298</td>
<td>69, 113</td>
<td></td>
</tr>
</tbody>
</table>
Table 1.2 continued

<table>
<thead>
<tr>
<th>Protein</th>
<th>Substrate</th>
<th>$K_M$ (mM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_M$ (s$^{-1}$ M$^{-1}$) x 10$^6$</th>
<th>Temp. (K)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>hRNase5</td>
<td>pCpA</td>
<td>0.005</td>
<td>0.13</td>
<td>0.000026</td>
<td>298</td>
<td>69, 113</td>
</tr>
<tr>
<td>hRNase5</td>
<td>CpAp</td>
<td>2.63±0.3</td>
<td>12.9±1.1</td>
<td>0.00049</td>
<td>298</td>
<td>57</td>
</tr>
<tr>
<td>hRNase5</td>
<td>CpApG</td>
<td>1.22±0.2</td>
<td>1.08±0.1</td>
<td>0.000885</td>
<td>298</td>
<td>57</td>
</tr>
<tr>
<td>hRNase5</td>
<td>CpApA</td>
<td>2.06±0.3</td>
<td>0.0325 ± 0.06</td>
<td>1.6 x 10$^{-6}$</td>
<td>298</td>
<td>57</td>
</tr>
<tr>
<td>hRNase6</td>
<td>tRNA</td>
<td>0.0022</td>
<td>5.1</td>
<td>2.3</td>
<td>310</td>
<td>58</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>UpA</td>
<td>0.20 ± 0.06</td>
<td>347 ± 40</td>
<td>1.7 ± 0.3</td>
<td>298</td>
<td>114</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>UpA</td>
<td>0.33</td>
<td>1920</td>
<td>5.8</td>
<td>298</td>
<td>115</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>UpA</td>
<td>0.62 ± 0.09</td>
<td>1400 ± 150</td>
<td>2.3 ± 0.4</td>
<td>298</td>
<td>116</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>UpA</td>
<td>0.7</td>
<td>2690</td>
<td>3.0 ± 0.1</td>
<td>298</td>
<td>117</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>CpA</td>
<td>0.5</td>
<td>2300</td>
<td>6.4 ± 0.7</td>
<td>298</td>
<td>117</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>C &gt; p</td>
<td>1.06±0.1</td>
<td>2.28±0.18</td>
<td>0.0215</td>
<td>298</td>
<td>57</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>Poly(C)</td>
<td>0.047 ± 0.011</td>
<td>190 ± 11</td>
<td>4.0 ± 1.0</td>
<td>283</td>
<td>118</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>Poly(C)</td>
<td>0.089 ± 0.009</td>
<td>507 ± 15</td>
<td>5.7 ± 0.5</td>
<td>298</td>
<td>114</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>Poly(C)</td>
<td>0.0331 ± 2.6</td>
<td>368.0 ± 11.0</td>
<td>11.2 ± 0.6</td>
<td>298</td>
<td>119</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>Poly(U)</td>
<td>0.06</td>
<td>24</td>
<td>0.4</td>
<td>298</td>
<td>115</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>Poly(C)</td>
<td>0.034</td>
<td>510</td>
<td>15</td>
<td>298</td>
<td>115</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>Poly(U)</td>
<td>0.06 ± 0.01</td>
<td>24 ± 15</td>
<td>0.4 ± 0.3</td>
<td>298</td>
<td>116</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>Poly(C)</td>
<td>0.034 ± 0.002</td>
<td>510 ± 10</td>
<td>15 ± 1</td>
<td>298</td>
<td>116</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>Poly(A)</td>
<td>0.080 ± 0.009</td>
<td>0.023 ± 0.001</td>
<td>0.00028 ± 0.0004</td>
<td>298</td>
<td>116</td>
</tr>
</tbody>
</table>
Alternate conformations are sampled as a result of internal motions in the enzymes driven by the surrounding environment and temperature. Internal protein motions occurring at all time scales is known as protein dynamics. These motions associated with protein dynamics occur on a wide range of time-scales ranging from femtosecond to millisecond (fs - ms). Fast motions governing side chain and backbone rearrangements occur at picosecond to nanosecond (ps - ns) time scales and allow sampling of conformations in the near vicinity. Large conformational changes are slow motions that occur mainly at longer time-scales microsecond to millisecond (μs - ms) and are associated with overcoming large energy barriers allowing interconversion between different group of populations (Figure 1.9).

Studies report that enzyme landscape is full of conformational states, which are sampled over different time-scales. The conformational fluctuations occurring at the time scale of reaction allow the enzyme to access the region of conformational landscape that contains the correct structural and dynamical features to promote specific interactions between reactants. In other words, the rate of the various events in the enzyme cycle is connected to the rate of sampling of correct motions/conformations. Only a small fraction of the conformations sampled by an enzyme facilitate in substrate (and/or cofactor) binding, reactant ground state stabilization, the chemical step of conversion, and product release along the catalytic cycle. These function promoting conformations are called as conformational sub-states. Conformational sub-states promote specific interactions between reactants in various stages of the catalytic cycle. Accessing such conformational sub-states is not a random event but form an integral part of enzymes function.

A number of studies summarizing the contribution of dynamical motions and conformational fluctuations in enzyme catalysis have reported the presences of complex residue networks in enzyme structure that connects the surface loop regions to the active site by a series of hydrogen bonds and hydrophobic interactions. Conformational fluctuations in these network residues introduce favorable changes in the active site allowing the enzyme to become more productive. Another vital role of these network pathways is to couple the surrounding solvent to enzyme function. The motions on the surface of the enzyme harness energy from the surrounding solvent which is eventually fed into the global conformational fluctuation of the enzyme providing energy for overcoming the activation energy barrier. The presence of such networks and motions is an integral part of enzymes structure and function that has been conserved during evolution.
**Figure 1.9: Schematic representation of conformational landscape of protein.** The arrows correspond to the conformational fluctuations within and between the distinct sub-states. Faster time-scales (ps - ns) allow conformational sampling within a well. Conformational fluctuations at slower time-scales (µs - ms) allow interconversion between two conformational states A (lower energy ground state) and B (excited state). \( \rho_A \) and \( \rho_B \) represent respective populations and \( \rho_A \gg \rho_B \).
1.4.A Dynamics in ribonucleases

The effect of protein motion on different aspects of enzyme function and protein complex formation in a variety of systems has become accessible through NMR relaxation experiments. Based on NMR dispersion relaxation experiments Loria and coworkers found that bRNaseA was rigid (absence of internal motions) at lower time-scales (ps - ns). However, residues in bRNaseA were dynamic at the µs to ms time-scale and showed chemical exchange in $^{15}$N-CPMG NMR experiments, i.e., displayed dynamic behavior. These residues also included the residues that form the B$_1$, B$_2$ and P$_0$ sub-sites in bRNaseA. The average chemical exchange rate ($k_{ex}$) quantified for dynamic residues in CPMG experiments was 1640 s$^{-1}$ which is similar to the rate of catalytic turnover in bRNaseA (1900 s$^{-1}$), suggesting a possible correlation between enzyme dynamics and catalysis in bRNaseA.

Further Loria and coworkers identified that mutations in distal residues His48 and loop1 (Figure 1.3) residues resulted in the loss of ms motions and 10-fold decrease in catalytic activity of bRNaseA suggesting that motions in RNase A were functionally relevant and more widespread than the active site. Additionally, their studies highlight the importance of H-bonding interactions in the propagation of functional motions in bRNaseA. His48 was eventually shown to be an important conformational switch that controls distal motions that are essential for optimal product release. Doucet et al. in their work demonstrated that altering one methyl $\sim$10 Å away from the active site significantly alters conformational dynamics of bRNaseA on ns - ms time-scale. This change in the dynamics of the enzyme affects ligand binding and specificity. Further studies on bRNaseA motion using NMR and computer simulations reveal that RNase does not undergo large conformational changes as observed in highly flexible proteins/enzymes. Flexibility in bRNaseA propels subtle movements of the loops and secondary structure elements. The movements play a crucial role in important steps of enzymatic cycle which is substrate recognition, catalysis, and/or product release. Human homologs of bRNaseA have also been subjected to dynamic investigation.

Global and local motions in hRNases are closely related to dynamics observed in bRNaseA, and similar motions are observed for the conserved catalytic residues in some members. In hRNase1, catalytic residue His119 is flexible and undergoes chemical exchange. NMR studies revealed that loop 4 in hRNase1 is highly flexible in the absence of ligand consistent with motions observed in bRNaseA; however, it becomes rigid upon inhibitor binding. NMR studies of hRNase3 show that loop4 residues undergo conformational exchange and MD simulations show
higher-than average residue motion of this loop consistent with bRNaseA.\textsuperscript{142, 144} However, their rate of conformational exchange varies significantly from bRNaseA. Further, experimental evidence also illustrates that the millisecond motions essential to product release in bRNaseA are absent from hRNase3. In hRNase5 the V\textsubscript{1} and V\textsubscript{2} domains undergo large breathing motions on the ns time-scale, like the hinge-bending motions observed in the bRNaseA\textsuperscript{143}. Like bRNaseA His114 (bRNaseA His119) and Lys40 (bRNaseA Lys41) residues were dynamic, while His13 was found to be very rigid on NMR time-scales.\textsuperscript{145} Considerable flexibility is observed in loop4 on the ns time-scale which contributes towards purine selection in hRNase5. However, at ms time-scale, hRNase5 is a more conformationally restrained as compared to other homologs.\textsuperscript{146} While the conservation of global, collective flexibility manifested by RNases has been appreciated for this enzyme fold, the differences in their dynamical signatures at functionally relevant time-scale have not yet been analyzed in detail. Correlating the subtle difference in the dynamics of these human homologs with their catalytic efficiency would provide a better understanding of the role of conformational motions in enzyme functions.

1.5 Methods for studying protein dynamics

Internal dynamics and conformational sub-states associated with the function of an enzyme can be identified and characterized using experimental and computer simulations discussed below.

1.5.A Molecular dynamics simulations

Experimental techniques like NMR,\textsuperscript{147, 148} single molecule Fröster Resonance Energy Transfer (smFRET),\textsuperscript{149} small angle scattering (X-ray and neutron),\textsuperscript{150, 151} electron paramagnetic resonance (EPR) spectroscopy,\textsuperscript{152} hydrogen/deuterium exchange mass spectroscopy (HD/X-MS)\textsuperscript{153} have been useful in identify global motions and conformational fluctuations in proteins. However, these techniques cannot capture the behavior of proteins at the full atomic detail and over time. Molecular dynamics (MD) is a computer simulation method for studying the physical movements of individual atoms in a biological macromolecule that cannot be observed directly in laboratory experiments. MD simulations have played an important role in increasing our understanding of the dynamical aspects of protein structure and function.\textsuperscript{154-156}

MD simulations provides a detailed picture of atomic-level motions or particles and molecules as a function of time guided by Newton's laws of motions.\textsuperscript{157} Based on the forces calculated, position
and velocity of each atom are updated, and this process is iteratively repeated through time. The inter-atomic forces between the particles are calculated using interatomic potentials collectively referred to as molecular mechanics force fields. A typical force field incorporates electrostatic (coulombic) interactions between atoms, spring-like terms that model the preferred length of each covalent bonds, angles, dihedrals and other types of non-bonded interactions. The timestep between each step in a MD simulation is in femtoseconds. Therefore, interactions between atoms are computed for millions and billions of steps to understand the structurally important motions at functionally relevant time-scales. This makes MD simulations computationally demanding.

Conformational motions in proteins are observed over a wide range of time scales (fs-s). Historically, the time-scales accessible by MD simulation have been shorter than the time-scales at which biomolecular events of interest take place. However, recent advances in computer hardware, software, and simulation algorithms have enabled sampling of biomolecules at the functionally relevant μs - ms time-scales in a regular manner. Improvements in force fields and advances in sampling techniques have also made significant contributions towards overcoming barriers and sampling the vast conformational landscape of biomolecules. Careful analysis of the conformational snapshots sampled during MD simulation can provide information about underlying low-frequency motions which corresponds to significant conformational changes relevant to the function of biomolecule under investigation. MD simulations can also be used to understand how a biomolecule will respond to perturbation like removal or change of ligand, mutation of residues, post-translational modifications, change in protonation state of important amino acids, temperature, pressure, pH, voltage etc. MD simulations can further capture elements required for understanding the molecular basis of important functional processes like ligand binding, protein folding and allostery.

Simulations have proven valuable in deciphering functional mechanisms of biomolecules, in uncovering the structural basis function, however, a number of choices have to be made when using MD simulations. Starting from the hardware to use (central processing unit (CPU) vs. graphical processing unit (GPU)), the software package and its version, the forcefield (AMBER, CHARMM, GROMOS, OPLS) to simulation conditions (like temperature or pressure) and choice of starting structure can have a significant effect of the computed trajectory. While each of these components has its own strengths and weaknesses, they could significantly affect the results. Therefore, it is essential to perform multiple simulations and ensure reproducibility in the behavior of biomolecule under investigation. It is also crucial to analyze MD simulations...
meticulously, considering various sources of error that might affect the results and results should be compared with experimental data if available. Using MD simulation, we are interested in understanding the biophysical and biochemical basis of substrate binding and product release in human homologs of RNase family members.

1.5.B Hybrid QM/MM

Limitations of classical MD simulations include that no covalent bonds can be formed or broken during the simulations. Therefore, MD simulations cannot be employed for modeling reactions that involve changes to covalent bonds or chemical reaction. Quantum-mechanical (QM) methods are used to model chemical reactions and other electronic processes in biomolecular systems. However, QM methods are computationally expensive and can practically be applied to systems which typically consists of up to a few hundred atoms at most. The modeling of enzyme reactions (in explicit solvent) requires methods that can treat thousands of atoms. Thus, to model the reaction mechanism of large biomolecules like enzyme systems the preferred approach is to use the hybrid QM/MM approach. This approach combines the use of QM method for active site region and an MM treatment for the surrounding protein and solvent. With a reasonable computational effort hybrid QM/MM method provide the necessary accuracy to model reactive biomolecular systems.

Like MD simulations proper design and execution of QM/MM simulations are very important. A wide variety of QM methods is available, from fast semi-empirical methods (AM1, PM3, SCC-DFTB), to highly accurate but computationally expensive electronic structure methods (DFT and ab-initio). It is equally important to carefully choose the MM force field that can accurately treat the atoms outside the QM region and their interactions with the QM atoms. The method to treat covalent bonds that cross the QM/MM boundary is an important consideration in QM/MM simulations. Methods like introducing additional link atoms, hybrid orbitals on MM atoms or a pseudobond to replace the QM-MM covalent bond has provided reasonable results.

Further depending on the scientific question, (relative) potential energies, or alternatively free energies, for the reaction can be obtained. A potential energy profile or the minimum energy pathway of a reaction can be determined by minimizing the energy of the system at several points along the reaction coordinate. Deriving free energy profiles using QM/MM methods are generally expensive. One strategy to reduce the computational cost of QM/MM simulations is to treat the
QM region using cheap approximate methods that reduces the computational effort considerably. Unfortunately, such methods have been subjected to a critical and often controversial discussion in the literature. There is a need to further improve the accuracy and sampling achieved by QM/MM methods, but the existing methods are already good enough to provide realistic modeling of biomolecular reactions.

QM/MM studies have provided detailed mechanistic insights into the enzymatic reactions of cytochrome P450,\textsuperscript{183} peroxidases,\textsuperscript{184} nitric oxide synthase,\textsuperscript{185} horseradish peroxidase,\textsuperscript{186} chorismite mutase,\textsuperscript{187, 188} DHFR \textsuperscript{146} and many other enzymes. The information obtained from QM/MM studies is increasingly allowing the designing of better drug leads, predicting drug metabolism, understanding drug resistance, and understanding the effect of mutations (in active-site and distal) on reactivity.\textsuperscript{175} In RNases, determination of enzyme reaction pathway and TS using accurate QM/MM methods will significantly assist in the design of enzyme for specific application, structure-based drug design, and improving structure-activity relationship methods. Additionally, obtaining conformational fluctuations associated with the reaction mechanism from QM/MM methods will aid in allosteric regulation of enzymatic processes and designing allosteric drugs.

1.5.C Quasi Anharmonic Analysis (QAA)

Experimental techniques, including neutron scattering, have indicated that protein motions are anharmonic with multiple wells in the conformational landscape.\textsuperscript{189} These anharmonic structural changes activate functionally relevant motions in proteins. Second order clustering methods are unable to identify conformational sub-states of anharmonic protein motions, the use of higher-order methods is necessary for identifying them. Quasi anharmonic analysis (QAA) uses fourth-order statistics to describe the atomic fluctuations and summarizes the internal motions relevant to protein function.\textsuperscript{129}

In QAA, a coupling matrix which encodes an anharmonic mode of motion describing the higher-order correlations between different regions of the protein is computed. Anharmonic sources are fully decorrelated, and higher-order dependencies in positional fluctuations are minimized. Additionally, a linear model is used which ignores any non-linear coupling that may exist in the fluctuations between different parts of a protein. The motions described along QAA basis vectors
are more relevant to the intrinsic motions of atom-pairs in proteins since the directionality of the motions lead to energetically homogeneous sub-states.

In their work Chennubhotla, Agarwal and co-workers, have shown that using QAA it is possible to identify and characterize conformational sub-states that have low populations in the conformational landscape.\textsuperscript{122, 129} They have demonstrated that the QAA has unique ability to distinguish energetically homogenous sub-states. Further QAA allows clean separation between the conformational sub-states by projecting the conformations sampled during the MD simulations in a lower dimensional space represented by QAA vectors. These conformational sub-states can be characterized for any relevant biophysical properties (such as internal energy, distance order parameter, or reaction coordinate).

Using series umbrella sampling simulations Chennubhotla, Agarwal and co-workers modeled the cis/trans isomerization catalyzed by enzyme cyclophilin A (CypA).\textsuperscript{129} They obtained the conformational sub-states along the reaction pathway of CypA by applying QAA to all conformations generated from simulations (Figure 1.10). Characterization of conformational sub-states in CypA indicates the enzyme's intrinsic ability is to explore conformations correspond to various sections of the reaction pathway, corresponding to low energy reactant and product states as well as the high energy transition state. Analysis of movements that allow interconversion between sub-states revealed the structural and dynamical features required to attain the transition state.

1.5.D Carr-Purcell-Meiboom-Gill (CPMG) - NMR relaxation dispersion

NMR spectroscopy represents a powerful technique for studying dynamical features of proteins. NMR utilizes the magnetic properties of certain \textsuperscript{1}H, \textsuperscript{13}C, \textsuperscript{15}N and \textsuperscript{31}P nuclei commonly found in proteins.\textsuperscript{190} These isotopically nuclei are inserted in the protein structure in a non-invasive manner serving as site-specific probes of local structure and dynamics.\textsuperscript{191} Depending on the environment of atoms within the protein, the nuclei of individual atoms will absorb different frequencies of radio signals. The absorption signals of different nuclei may be perturbed by change in their chemical environment over time or in presence of a ligand. The NMR probes are exposed to at least two distinct chemical environment in a time dependent manner, a process called as chemical exchange. NMR spectra affected by chemical exchange is then used to determine the change in the behavior of individual nuclei, and the dynamics of the whole protein can be analyzed with
collective behavior of individual nucleotide.\textsuperscript{192} Therefore NMR spectroscopy is uniquely suited to study protein dynamics under equilibrium conditions, with no need for external perturbations such as changes of temperature or pressure.

Many biochemical events (including the ribonucleolytic cleavage catalyzed by RNases) occur on the \( \mu s \) - ms time-scale, and it is of considerable interest to characterize the conformational transitions that are involved in such processes.\textsuperscript{138, 193-196} Motion in this time-scale window generally includes side chain reorientation, loop motion, secondary structure changes and hinged domain movements. CPMG relaxation dispersion is a powerful approach to obtaining kinetic, thermodynamic and structural information for exchange processes in the 0.3 - 10 ms time window. The principle of NMR-CPMG experiment is to refocus exchange broadening (i.e., reduce or “disperse” \( R_{ex} \)) by applying pulse elements to transverse magnetization during a special relaxation delay.\textsuperscript{197} Relaxation describes how signals change with time. CPMG relaxation dispersion experiments most often make use of backbone amide \(^1\text{H}\) or \(^{15}\text{N}\) spin probes, focusing on proton or nitrogen magnetization (single-quantum transitions), respectively.

Dynamics on the \( \mu s \) - ms time scale can be characterized in terms of an exchange contribution, \( R_{ex} \), to the observed transverse relaxation rate, \( R_{2}^{obs} \), using the following expression:

\[
R_{2}^{obs} = R_{ex} + R_{2}^{0}
\]

Where \( R_{2}^{0} \) is the intrinsic relaxation rate of the nuclei under investigation. For a two-site exchange \( R_{ex} \) depends on the population of the states \( (\rho_a, \rho_b) \), the difference in chemical shifts \( (\Delta \omega) \), the exchange rate constant \( (k_{ex} = k_f + k_r, \text{ where } k_f \text{ and } k_r \text{ are the forward and reverse rate constants, respectively}) \) and the pulse repetition rate. Primarily recorded as a series of \(^{15}\text{N}\)- Heteronuclear Single Quantum Coherence (HSQC) experiments, the technique has an extremely good atomic-scale resolution, allowing the investigation of specific protons and other relevant atoms on selected residues in any protein or enzyme amenable to NMR investigation. The chemical shift of the investigated atoms can then be determined by combining \( \Delta \omega \) with information obtained from an Heteronuclear Multiple-Quantum Correlation (HMQC) spectrum which in turn can be used to determine the solution structure of the excited state.\textsuperscript{198} The experimental framework outlined above has been used extensively to detect motions on the \( \mu s \) - ms time-scale within proteins and understand their function.\textsuperscript{139, 199, 200}
Figure 1.10: QAA describes conformational sub-states leading to transition state during catalysis in cyclophilin A. Two levels of sub-states are shown (A) level 1 and (B) level 2. Conformational sub-states with homogenous populations are identified for the reactant (R), transition state (T) and product (P). Each colored dot in the top panels corresponds to a single conformation and is colored according to the reaction coordinate value in panel D. Arrows indicate interconversion between other sub-states and T and the corresponding CypA motions are shown below the plots. (C) The identified motions extend from the flexible surface region to the active site connected by hydrogen bond networks. (D) Free-energy profile of the CypA reaction. The color bar corresponds to reaction coordinate used for coloring conformations in panel A. (This figure is published in 129)
1.6 Proposed study

The overview of the biophysical, biochemical and physiological properties of RNases offer promising therapeutic applications in several diseases including amyotrophic lateral sclerosis and Parkinson’s Disease. However, an accurate understanding of the enzymatic properties of most hRNases is still unavailable. Limited information is available about the mechanistic basis of substrate binding and chemical step in various hRNases, while the atomistic details of product release step has not yet been investigated. Moreover, the contributions of distal regions in the steps of the enzymatic cycle of hRNases remain uncharacterized. It also unexplored which of the two proposed mechanism is more relevant to catalysis by hRNases.

The similarity in range of enzyme efficiency and dynamics in these enzymes raises the question of the role of dynamics in catalysis and provides a unique opportunity to investigate human ribonucleases as prototypical systems for understanding relationship between dynamics and enzyme catalysis. Using computational (Classical MD simulations and hybrid QM/MM approach) and experimental (NMR, steady-state kinetics) approaches this study quantitatively characterizes functionally important dynamics in the catalytic cycle of RNases. As the rate-determining step is currently unclear, and also may be different for all hRNases, conformational dynamics is characterized in each step of RNase catalytic cycle.

The four steps of the RNase catalytic cycle studied as a part of this thesis are organized in the chapters 2-5 (Figure 1.11). Chapter 2 investigates µs - ms dynamics in 23 RNases in their apo form using $^{15}$N-CPMG NMR experiments and computer simulations. Chapter 3 focuses on characterizing the nucleotide substrate binding properties in hRNases using computer simulations at the µs time-scales. Chapter 4 concentrates on identifying conformational dynamics and sub-states associated with product release in RNase family members. Umbrella sampling simulations along with quasi-anharmonic analysis (QAA) are employed for this study. In Chapter 5 additional aspects of RNase enzyme mechanism are explored, the dynamic properties of RNases in presence of ligands (product analogs) is determined using NMR titrations and chemical shift perturbation analysis (CHESPA). Conformational fluctuations in the chemical turnover step are identified for bRNaseA using hybrid QM/MM and Nudged Elastic Band (NEB) simulations. In addition, kinetic parameters for hRNases are obtained using steady-state kinetic experiments. Finally, the major findings of this thesis are summarized in chapter 6 with suggestion for future studies.
Figure 1.11: Schematic representation of the RNase catalytic cycle and proposed work. (A) Simple scheme indicating steps involved in an enzyme’s catalytic cycle. MD = modeled using classical simulations. QM/MM = modeled with hybrid quantum-classical simulations. (B) Summary of proposed work.
Chapter 2  Conservation of Dynamics Associated with Biological Function in an Enzyme Superfamily
This chapter is a slightly revised version of a manuscript by the same title published in the journal, Structure published in 2018

Reprinted (adapted) with permission from Chitra Narayanan,1,# David N. Bernard,1,# Khushboo Bafna,2 Donald Gagné,1 Chakra S. Chennubhotla,3 Nicolas Doucet,1,4,5,* Pratul K. Agarwa16,7,*

1INRS – Institut Armand-Frappier, Université du Québec, 531 Boulevard des Prairies, Laval, QC, H7V 1B7, Canada.
2Genome Science and Technology, University of Tennessee, Knoxville, TN 37996.
3Department of Computational & Systems Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260.
4PROTEO, the Québec Network for Research on Protein Function, Engineering, and Applications, 1045 Avenue de la Médecine, Université Laval, Québec, QC, G1V 0A6, Canada.
5GRASP, Groupe de Recherche Axé sur la Structure des Protéines, 3649 Promenade Sir William Osler, McGill University, Montréal, QC, H3G 0B1, Canada.
6Computational Biology Institute and Computer Science and Engineering Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN 37830.
7Department of Biochemistry, Cellular and Molecular Biology, University of Tennessee, Knoxville, TN 37996

#These authors contributed equally
*Corresponding authors: Nicolas.Doucet@iaf.inrs.ca and pratul@agarwal-lab.org

DOI:10.1016/j.str.2018.01.015

This author contributed the following to the manuscript: (1) homology modeling, (2) molecular dynamics simulations, and (3) data analysis and organization.
2.1 Abstract

Enzyme superfamily members that share common chemical and/or biological functions also share common features. Conservation of structural features such as critical active-site residues and the entire protein fold, even in cases of low sequence similarity, is well known. However, the link between function and dynamics is not well understood. We present a systematic characterization of the intrinsic atomic-scale dynamics of over twenty members of the pancreatic-type ribonuclease (RNase) superfamily, which share a common structural fold. This study is motivated by the fact that the range of chemical activity as well as molecular motions exhibited by RNase homologs span >10^5 fold. Dynamics on the microsecond-millisecond time-scale was characterized using a combination of nuclear magnetic resonance (NMR) relaxation dispersion experiments and computer simulations. Phylogenetic clustering of the RNases led to the grouping of sequences into sub-families, with members within sub-families sharing similar biological functions. Detailed characterization of the dynamical properties of the diverse RNases showed conserved dynamical traits for enzymes grouped into functionally distinct sub-families. These results suggest that selective pressure for the conservation of dynamical behavior, among other factors, may be linked to the distinct chemical and biological functions in an enzyme superfamily.

2.2 Introduction

The linkage between dynamics and enzyme catalysis has been intensely debated. Enzymes, like other molecules, experience internal motions over a wide range of time-scales. These motions are driven by temperature and surrounding environment, including the solvent. Evidence continues to emerge from an increasing number of systems emphasizing the involvement of dynamics in enzyme function. However, the presence of motions even near/in the active site, or their occurrence at time-scales that coincide with rate-limiting events in enzyme mechanisms, does not automatically imply that these motions play a role in enzyme function. One of the challenges associated with investigating enzyme dynamics is that structure and dynamics are inter-related, thus making it difficult to decouple their individual contributions to enzyme catalysis.

The enzyme fold has been proposed to serve as a scaffold which enables optimal positioning of the catalytic residues to provide a unique environment, which is very different from solvent, and provide structural and electrostatic features necessary for transition state stabilization.
Catalytically important residues are conserved across enzymes that share the same fold and catalyze the same (or similar) chemical reaction. It has been suggested that if dynamics plays a functional role in enzyme catalysis, then it must also be subjected to evolutionary pressure and be conserved as a part of the enzyme fold topology. Therefore, characterization of dynamics of enzymes within a superfamily could provide vital information on the evolutionary conservation of dynamics (or lack thereof) among structural and functional homologs that catalyze the same chemical reaction. This approach could provide much needed information for teasing apart the role of structure and dynamics in enzyme catalysis.

Pancreatic-type ribonucleases (RNases) are a superfamily of enzymes found in bacteria and vertebrates that catalyze the hydrolysis and transphosphorylation of ribonucleic acid (RNA) substrates. Nearly 650 sequences corresponding to members of this superfamily are found in the Pfam database. Eight catalytically active canonical RNases (and five supplemental pseudo-genes that lack the conserved active-site residues essential for ribonucleolytic activity) were identified in the initial sequencing of the human genome. All structurally characterized RNases display a conserved structural fold and share the common chemical function of cleaving the phosphodiester bond of ribonucleic acid (RNA) substrates, albeit with varying degrees of catalytic efficiencies and specificities. The canonical RNases have further been reported to perform a variety of biological functions such as angiogenesis, anti-pathogenicity and immuno-suppressivity, in addition to the common ribonucleolytic activity. While the biological functions are broadly known and have been characterized for some members identified in the human genome, a significant number of members remain largely uncharacterized to date. The unique similarity in structure (fold), whilst displaying significant diversity in function, makes the RNase superfamily an excellent candidate for characterizing the effect of conserved dynamics on function.

A variety of biophysical, NMR and computational approaches have indicated the role of conformational dynamics in regulating catalysis in bovine RNase A (BtRA), the archetypal member of this family. Millisecond motions of a loop located over 20 Å away from the active site, suggested to be propagated by a conserved hydrogen-bonding network, were shown to be essential for efficient catalysis. Mutations disrupting this hydrogen bonding network were shown to result in the uncoupling of dynamics on the catalytic time-scale and a corresponding decrease in the catalytic turnover. Further, reduced catalytic efficiency relative to bovine RNase A was observed in two naturally occurring RNase A homologs lacking this
hydrogen bonding network \cite{137, 138, 216}. Interestingly, the global conformational exchange rates of the diverse RNase members vary over five orders of magnitude (microseconds to 100s of milliseconds) coinciding with the range of substrate turnover (and chemical efficiency), which also varies up to five orders of magnitude between the most efficient and least efficient members of the human canonical forms \cite{215}. For example, human RNase 5 (angiogenin), which lacks the dynamical loop 4 region of RNase A shown to impact catalysis, displays 10^5 fold reduction in ribonucleolytic activity compared to RNase A \cite{217}.

As a step towards characterizing the relationship between dynamics, chemical and biological functions of various members of this broad vertebrate family, we performed a systematic characterization of the intrinsic dynamics of over 20 RNase homologs. The selected RNases correspond to a representative set of homologous vertebrate sequences whose structures have been determined previously by X-ray crystallography or NMR and are available in the protein data bank. Using a combination of NMR and computational approaches, we characterized the dynamics of these proteins over the microsecond to millisecond (\( \mu s-ms \)) time-scale to probe the relationship between sequence diversity and dynamics, and its correlation with the designated biological functions. The present study focuses on characterizing and comparing the intrinsic dynamics of apo proteins, as these enzymes bind to and catalyze a wide variety of RNA substrates. Our results show that while the diverse RNase homologs used for the analysis share a common structural fold, the dynamical properties of these proteins vary significantly. Clustering superfamily members into evolutionarily distinct sub-families showed conservation of dynamical properties within sub-family members and notable variations between different sub-families. Categorizing homologous RNase superfamily sequences into sub-families, which also exhibit similar dynamics, provides a unique opportunity for further investigation to tease apart the role of structure and dynamics in enzyme function.

2.3 Materials and methods

**Genes, Expression and Purification:** *Escherichia coli* codon-optimized sequences of *Pongo pygmaeus* ribonuclease 3 (PpR3) and *Macaca fascicularis* ribonuclease 3 (MfR3) (UniProt entries P47781 and P47779) were synthesized in IPTG-inducible pJExpress414 vectors containing the ampicillin resistance gene (DNA2.0, Newark, CA). These vectors were transformed into E. coli BL21 (DE3) for all protein expression experiments. All human RNases were cloned, expressed, and purified according to previous protocols \cite{138, 142, 146, 218}. \(^{15}\)N- and \(^{15}\)N/\(^{13}\)C-labeled protein
expression and purification was conducted as described previously \cite{138,142}, with the following modifications. The temperature was lowered to 30°C following addition of IPTG, the volume of culture media was of 1 L, and bacteria were grown overnight before being harvested by centrifugation. Protein concentrations were determined using extinction coefficients of 11,960 and 14,940 M⁻¹cm⁻¹, respectively, as estimated by ExPASy ProtParam.

**Solution NMR Experiments:** 2D ¹⁵N-HSQC, 3D-HNCA CB and 3D-CBCA(CO)NH assignment experiments were performed on a Varian INOVA 500 MHz spectrometer (298 K), while ¹⁵N-CPMG relaxation dispersion experiments were performed on the same spectrometer as well as a Varian INOVA 800 MHz spectrometer (298 K), as described earlier \cite{138,142,219}. NMR processing and analysis was performed using NMRPipe \cite{220}, CcpNmr Analysis \cite{221} and Sparky \cite{222}.

**Phylogenetic analysis:** Multiple sequence alignment of the 23 RNase sequences was performed using ClustalΩ \cite{223}. A complete list of sequences and sequence identifiers is provided in Table 2.1. The phylogenetic tree of the aligned sequences was determined based on maximum likelihood approach using RaxML v8.0.26 with the WAG amino acid substitution model \cite{224}. Reliability of the branching was assessed using the bootstrapping method over 100 iterations. The resulting phylogenetic tree in Newick format was visualized using Figtree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

**System preparation for computer simulations:** Coordinates of 21 RNases were obtained from the corresponding PDB files for the apo model preparations (see Table 2.1). The structures for two Hominidae RNase 3 are not available; therefore, homology models were generated as described below. System preparation for molecular dynamics (MD) simulations were performed using the leap module in AMBER \cite{225} v14 and AMBER’s ff14SB force-field. Equilibrated systems were used for each of the 23 proteins to run a total of 1μs of production MD under constant energy conditions (NVE ensemble). A PME cut-off of 8 Å was used as suggested by AMBER input files for GPU-enabled simulations \cite{226}. To address an important concern regarding the possible bias of results from a single MD trajectory and ensure reproducibility, for each of the 23 systems two independent MD trajectories each 0.5 μs were performed.

**Homology modeling of Hominidae RNase 3 sequences:** Structures of two Hominidae RNase 3 sequences, *Pongo pygmaeus* (PpR3) and *Macaca fascicularis* (MfR3), were generated through homology modeling using MODELLER v9.16 \cite{227}. The two Hominidae sequences share 84.3%
and 85.8% sequence identity with human RNase 3 (PDB ID: 1QMT), which was used as the template for modeling. The automodel option in MODELLER was used to generate models, of which two models with the lowest DOPE scores were chosen for each of the two sequences.

**Molecular dynamics simulations:** The protein systems were neutralized by the addition of counter ions and the resulting systems were then immersed in a rectangular box of SPC/E water, with a 10 Å minimum distance between the protein and the edge of the periodic box. The prepared systems were equilibrated using a protocol described previously 129. The equilibrated system was used for each of the 23 proteins to run a total of 1 μs of production MD (two independent 0.5 μs MD trajectories, see details below) under constant energy conditions (NVE ensemble). The Particle-mesh Ewald (PME) method was used for long-range electrostatics and a time-step of 2 femtoseconds was used during production runs. Coordinates of the system were saved every 5 picoseconds, for a total of 200,000 conformational snapshots for each protein (as a sum of 100,000 snapshots from each independent trajectory). All simulations were performed using graphical processing units (GPU) enabled version of AMBER’s MD simulation engine pmemd (v14). A PME cut-off of 8 Å was used as suggested by AMBER input files for GPU-enabled simulations 226. To address an important concern regarding the possible bias of results from a single MD trajectory and ensure reproducibility, for each of the 23 systems two independent MD trajectories each 0.5 μs were performed. The first trajectory after equilibration was used for production at 300 K, while for the second trajectory the structure was slowly heated to 310 K using MD with 50,000 steps and minimized to generate conformational and structural diversity. Production run for the second trajectory was also performed at 300 K.

**RMSF and RMSF10 calculations:** All-atoms root-mean-square fluctuations (RMSF) were computed based on the conformational snapshots collected during the MD simulations. To identify global motions on slower time-scales from MD, for each of the 23 systems we separately computed and aggregated the fluctuations associated with the first (slowest) 10 quasi-harmonic modes (RMSF10). It is well known that slowest 10 modes contribute to the majority of fluctuations in proteins (>80%) and the use of RMSF10, instead of all modes (RMSF), removes the faster stochastic motions of the protein, allowing focus on intrinsic dynamics of proteins 228. Both these calculations were performed using AMBER’s ptraj analysis program. All trajectory conformations were first aligned to a common structure, to remove any translation and overall molecular rotation during the simulations.
Table 2.1: PDB IDs of RNase homologs used for dynamical characterization

<table>
<thead>
<tr>
<th>PDB code</th>
<th>Abbreviation</th>
<th>Species</th>
<th>RNase type</th>
</tr>
</thead>
<tbody>
<tr>
<td>7RSA</td>
<td>BtRA</td>
<td><em>Bos taurus</em></td>
<td>RNase A</td>
</tr>
<tr>
<td>2K11</td>
<td>HsR1</td>
<td><em>Homo sapiens</em></td>
<td>RNase 1</td>
</tr>
<tr>
<td>3T5R</td>
<td>MmR1</td>
<td><em>Musculus</em></td>
<td>RNase 1</td>
</tr>
<tr>
<td>1RRA</td>
<td>RnRA</td>
<td><em>Rattus norvegicus</em></td>
<td>RNase A</td>
</tr>
<tr>
<td>1GQV</td>
<td>HsR2</td>
<td><em>Homo sapiens</em></td>
<td>RNase 2</td>
</tr>
<tr>
<td>1QMT</td>
<td>HsR3</td>
<td><em>Homo sapiens</em></td>
<td>RNase 3</td>
</tr>
<tr>
<td>PpECP*</td>
<td>PpR3</td>
<td><em>Pongo pygmaeus</em></td>
<td>RNase 3</td>
</tr>
<tr>
<td>MIECP*</td>
<td>Mfr3</td>
<td><em>Macaca fascicularis</em></td>
<td>RNase 3</td>
</tr>
<tr>
<td>2HKY</td>
<td>HsR7</td>
<td><em>Homo sapiens</em></td>
<td>RNase 7</td>
</tr>
<tr>
<td>RNase6#1</td>
<td>HsR6</td>
<td><em>Homo sapiens</em></td>
<td>RNase 6</td>
</tr>
<tr>
<td>1ANG</td>
<td>HsR5</td>
<td><em>Homo sapiens</em></td>
<td>Angiogenin</td>
</tr>
<tr>
<td>1AGI</td>
<td>BtR5</td>
<td><em>Bos taurus</em></td>
<td>Angiogenin</td>
</tr>
<tr>
<td>2BWK</td>
<td>MmR5</td>
<td><em>Musculus</em></td>
<td>Angiogenin</td>
</tr>
<tr>
<td>4QFI</td>
<td>RnR5</td>
<td><em>Rattus norvegicus</em></td>
<td>Angiogenin</td>
</tr>
<tr>
<td>3ZBV</td>
<td>MmR52</td>
<td><em>Musculus</em></td>
<td>Angiogenin</td>
</tr>
<tr>
<td>3ZBW</td>
<td>MmR53</td>
<td><em>Musculus</em></td>
<td>Angiogenin</td>
</tr>
<tr>
<td>1RNF</td>
<td>HsR4</td>
<td><em>Homo sapiens</em></td>
<td>RNase 4</td>
</tr>
<tr>
<td>1KM9</td>
<td>RcRx</td>
<td><em>Rana catesbeiana</em></td>
<td>Bullfrog RNase</td>
</tr>
<tr>
<td>3SNF</td>
<td>RpRx</td>
<td><em>Rana pipiens</em></td>
<td>Onconase</td>
</tr>
<tr>
<td>4PER</td>
<td>GgR5</td>
<td><em>Gallus gallus</em></td>
<td>Angiogenin</td>
</tr>
<tr>
<td>2VQ9</td>
<td>DrR1</td>
<td><em>Danio rerio</em></td>
<td>RNase ZF-3E</td>
</tr>
<tr>
<td>3LJE</td>
<td>DrR5</td>
<td><em>Danio rerio</em></td>
<td>Angiogenin</td>
</tr>
<tr>
<td>2ZPO</td>
<td>CmRx</td>
<td><em>Chelonia mydas</em></td>
<td>Turtle RNase</td>
</tr>
</tbody>
</table>

* = Homology model, # = Structure not deposited

1 The crystal structure of human RNase 6 was obtained in our collaborators research group. While this structure has not yet been deposited in the PDB, it shows a significant structural identity with the structure available in the PDB (4X09), with the root mean square deviation of 0.20 Å.
Quantitative comparison of dynamical properties: Pearson’s correlations were used for the quantitative characterization of dynamical similarities (and differences) between protein pairs. Correlation between RMSFs for each residue \( a_i \) and \( b_i \) of RNases A and B, respectively, were calculated using the equation:

\[
\frac{\sum_{i=1}^{n} (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^{n} (a_i - \bar{a})^2} \sqrt{\sum_{i=1}^{n} (b_i - \bar{b})^2}}
\]

(2.1)

where \( n \) corresponds to the number of residues including the gaps in the alignment. \( \bar{a} \) and \( \bar{b} \) correspond to the average RMSF values for RNase A and B. Pairwise-correlations were calculated for each of the 23 RNases using RMSF10 excluding residues corresponding to gaps in the pairwise alignment. Average correlations within and between sub-families were calculated by averaging correlation coefficients of sequences being compared.

2.4 Results

2.4.A Phylogenetic analysis of ribonuclease superfamily

As a representative set of the pancreatic ribonuclease superfamily, 21 diverse RNases, whose three-dimensional structures are available in the protein data bank, were selected for detailed dynamical characterization. Human RNases 1-7 were included in this analysis. Additionally, two Hominidae RNase 3 sequences that are closely related to human RNase 3 were also analyzed using structures generated based on homology modeling. A list of all RNase sequences and sequence identifiers is provided in Table 2.1. Comparison of the 23 RNase dataset shows that the sequences share an average sequence identity of \( \sim 37.4\% \), and an average RMSD of 2.47 Å for the alignment of all backbone atoms of the 21 members with solved structures. Sequence and structural alignments indicated that most variable regions are, unsurprisingly, associated with loop regions of the protein fold (Figure 2.1 A, B). Catalytic triad residues (H12, K41 and H119, bovine RNase A residue sequence numbering), which are essential for the ribonucleolytic activity, are conserved in all 23 proteins analyzed.

Phylogenetic analysis of the 23 RNase sequences led to the identification of four distinct sequence clusters, henceforth referred to as sub-families (Figure 2.1 C). The RNase A-like sub-family (4 members) includes members that share similarity with bovine RNase A (BtRA), whose
primary function is associated with the cleavage of RNA substrates. Human and bovine ribonucleases of this sub-family are the most efficient ribonucleolytic catalysts characterized to date within the RNase superfamily. Members of the Eosinophil associated ribonuclease-like (EAR-like) sub-family (6 members) are characterized by insertion of a loop (L7), and have been suggested to play an important role in host defense and inflammatory response \(^{229,230}\), in addition to preserving the common ribonucleolytic function. Members of this sub-family show catalytic turnover rates 5- to >400-fold lower relative to BtRA, for those with known catalytic turnover rates (data not shown).

The Angiogenin-like sub-family (6 members) is characterized by a truncated loop 4 (L4), a loop region that displays conformational dynamics on the millisecond time-scale in bovine RNase A and was shown to play an important role in its ribonucleolytic function. Catalytic efficiency of human angiogenin (HsR5) is 105-fold lower than bovine RNase A. However, its ribonucleolytic activity was shown to be essential for biological function (angiogenesis) \(^{231}\). Previous studies also showed that replacing the truncated loop 4 of HsR5 with that of bovine RNase A increases catalytic efficiency of this chimeric protein by over 200-fold \(^{232}\), while swapping loop 4 from RNase A with that of HsR5 reduced the activity of this chimera by 60-fold \(^{233}\). The other seven RNases not forming a part of the three groups described above were classified as Others, and include sequences that are not necessarily closely related to each other. The functional characteristics of this cluster are not well understood and could therefore display functional and dynamical properties similar to the three sub-families described above.

### 2.4.B Conformational dynamics on the microsecond-millisecond (μs - ms) time-scale

The time-scale of catalysis varies widely in the RNase superfamily. BtRA is known to catalyze cleavage of RNA with a \(k_{\text{cat}} \approx 1700\) s\(^{-1}\), with product release being the rate-limiting step of the reaction \(^{137,216}\). Other members are slower catalysts, showing up to \(10^5\)-fold weaker catalytic efficiency \((k_{\text{cat}}/K_M)\). We used a combination of NMR relaxation dispersion CPMG experiments and MD simulations to probe the μs - ms dynamical properties of these enzymes. Figure 2.2 shows the conformational exchange of various RNases from each of the four sub-families probed using \(^{15}\)N-CPMG relaxation dispersion experiments. Residues exhibiting conformational exchange on the μs-ms time-scale are represented as spheres. Comparison of millisecond motions of these selected RNases showed significant differences in regions displaying conformational exchange.
Figure 2.1: Structural and phylogenetic analysis of RNase homologs. (A) Consensus sequence alignment of the 23 RNases, with respective sequence identifiers displayed to the left. Detailed information about the sequences is provided in Table 2.1. The secondary structural elements $\alpha$-helices (orange rectangles), $\beta$-strands (cyan arrows) and loop regions (L), corresponding to the bovine RNase A structure, are displayed above the alignment. Conserved active-site residues H12, K41 and H119 (BtRA numbering) are identified with an asterisk above the alignment. BtRA (PDB ID 7RSA) sequence numbering (BtRA#) is shown at the top and consensus sequence numbering (Cons#) below the multiple sequence alignment. A list of all consensus sequence positions is presented in Figure 2.10 (See Appendix 2.7). Gaps in the alignment are represented using dotted lines. Alignment was prepared using EsPript. (B) Structural alignment of RNase homologs. Structures include 7RSA (green); 1QMT (blue), 2HKY (cyan); 1ANG (salmon); 3SNF (yellow) and 1RNF (gold). The V1 and V2 arms as identified using boxes with dotted lines. Active-site residues are displayed as sticks and labeled with BtRA numbering. (C) Phylogenetic clustering of the 23 RNases generated based on maximum likelihood analysis. Sequences are classified into four clusters, identified using different colors.
for members from different sub-families. In contrast, residues exhibiting conformational exchange are clustered in the same regions for four members within a sub-family (Figure 2.2 A). Members of the EAR-like RNase sub-family display motions primarily clustered in the V1 arm and loop 4. In contrast, BtRA from the RNase A-like sub-family showed conformational exchange in both V1 and V2 arms, including loops L1 and L4 (Figure 2.2 B). Human RNase 5 (HsR5) of the angiogenin-like sub-family exhibits no conformational exchange on the CPMG time-scale (Figure 2.2 C), confirming results from a previous report. To investigate whether HsR5 experiences motions on a faster µs - ms time-scale window than the one probed by the CPMG experiment, we also analyzed elevated transverse relaxation rates at short spin-echo pulse delays in the CPMG relaxation dispersion profiles (i.e. \( R_2, \text{eff} \) at \( \nu_{\text{CPMG}} = 1600 \text{ s}^{-1} \), or \( \tau_{\text{cp}} = 0.625 \text{ ms} \)). Analysis for apo and ligand-bound HsR5 showed that only one residue per state exhibited elevated \( R_2, \text{eff} \) values (Asn68 in the apo form, and Asn109 in the ligand-bound form), which could indicate that these two residues experience conformational exchange on a faster µs - ms time-scale (Figure 2.3). Overall, these observations further confirm that HsR5 is a fairly rigid protein on the µs - ms time-scale and that no significant conformational exchange is observed on this regime. Human RNase 4 (HsR4), from the Others cluster, displayed extensive conformational exchange throughout the protein (Figure 2.2 D). NMR data from other members of the RNase sub-family is presently unavailable on the µs - ms time-scale.

Interestingly, insertion of loop 1 (L1) from an EAR-like RNase (HsR3) into BtRA (BtRA-HsR3, Figure 2.2 B) switched the dynamical profile of BtRA into that of HsR3 (Figure 2.2 A) in the chimeric protein. The motivation for the design of this chimera was to develop and characterize a control system for investigating the dependence of dynamics on a region based on type of residues rather than secondary structure. Previous studies showed a detailed comparison of the rates of ligand binding and release, in addition to conformational exchange profiles of BtRA, BtRA-HsR3, and HsR3. To gain insights into changes in the relaxation dispersion of these enzymes in the L1 region, here we compared the fitted relaxation dispersions for residues of loop 1 from HsR3 (EAR-like sub-family) with BtRA (RNase A-like sub-family) and the chimera BtRA-HsR3 (Figure 2.4 A). Our results show that loop 1 residues of BtRA exhibit a two-site conformational exchange at these positions, as observed from the presence of relaxation dispersion profiles (solid lines in Figure 2.4, left panel). In contrast, no motions on this time-scale were observed for loop 1 positions in HsR3 and the chimera BtRA-HsR3, as seen from the flat dispersion profiles (shown as dashed lines in Figure 2.4 A, middle and right panels). Residues in structurally equivalent positions in the three enzymes are represented using the same colors. Further, a chemical shift comparison between
Figure 2.2: Millisecond conformational exchange in Hominidae and bovine RNases. Cartoon representation of (A) EAR-like sub-family members - H. sapiens EDN (HsR2), H. sapiens ECP (HsR3), P. pygmaeus ECP (PpR3), M. fascicularis ECP (MfR3); (B) RNase A-like sub-family member B. taurus RNase A (BtRA) and the BtRA\textsubscript{HsR3} chimera; (C) Angiogenin-like sub-family member H. sapiens angiogenin (HsR5); and (D) H. sapiens RNase 4 (HsR4) from the Others cluster. Structures are colored based on the phylogenetic group shown in Figure 2.1 C. Residues identified by NMR as undergoing conformational exchange on the millisecond time-scale are depicted as spheres and numbered according to their position in the sequence alignment (see Figure 2.1) to facilitate comparison. Conformational exchange was probed by $^{15}$N-CPMG NMR relaxation dispersion experiments at 500 and 800 MHz (298 K). Residues were considered for further analysis only if the difference in measured $R_2 (1/\tau_{cp})$ values at fast ($\tau_{cp} = 0.625$ ms) and slow ($\tau_{cp} = 10$ ms) refocusing pulse delays was greater than 2 s$^{-1}$, similar to previous reports. BtRA\textsubscript{HsR3}, which displays sub-sector conformational exchange similar to HsR3, is an artificial chimeric hybrid of BtRA in which the 12-residue long loop 1 of BtRA (14-DSSTSAASSSNY-25, BtRA numbering) was replaced by the 6-residue long loop 1 of HsR3 (17-SLNPPR-22, HsR3 numbering). Structures color-coded based on conformational exchange rates are shown in Figure 2.9 (See Appendix 2.7).
Figure 2.3: Conformational exchange in HsR5 residues. Transverse relaxation rates ($R_2$) at $\nu_{\text{CPMG}} = 1600 \text{ s}^{-1}$ ($\tau_{\text{CP}} = 0.625 \text{ ms}$) in free and ligand-bound HsR5. $R_{2,\text{eff}}$ values at $\nu_{\text{CPMG}} = 1600 \text{ s}^{-1}$ are shown for all assigned residues of apo (left) and 3'-UMP-bound (right) HsR5. Dashed lines indicate the average $R_{2,\text{eff}}$ value and dotted lines indicate standard deviation. Residues exhibiting elevated $R_{2,\text{eff}}$ are identified on each graph. Upon closer examination of the full relaxation dispersion curves for these residues, most were disregarded since their $R_{2,\text{eff}}$ values were either inconsistently elevated and/or simply resulted from bad raw data. Only two residues displayed consistently elevated values: Asn68 in the apo protein, and Asn109 in the 3'-UMP bound protein. This could indicate the presence of dynamics at positions Asn68 (free form) and Asn109 (3'-UMP-bound form) on a microsecond time-scale faster than that measurable by CPMG experiments. These observations further confirm the lack of conformational exchange on the CPMG time-scale, and the overall likelihood that no significant functional dynamics is observed on the faster microsecond time-scale in HsR5.
Figure 2.4: Dynamical effects of loop 1 swapping on BtRA. (A) Relaxation dispersion profiles of various loop 1 residues obtained from $^{15}$N-CPMG experiments at 800 MHz and 25°C for bovine RNase A (BtRA), human RNase 3 (HsR3) and a previously described $^{138}$ chimera of RNase A (BtRA$\text{HsR3}$) in which loop 1 has been replaced by the corresponding residues of HsR3. $^{15}$N-CPMG data was also acquired at 500 and/or 600 MHz for these proteins (not shown for clarity). Solid curves are shown for residues displaying conformational exchange, and dashed curves belong to non-exchanging residues. Curves of the same color represent structurally equivalent residues in the sequence alignment. (B) Compounded $^1$H-$^{15}$N chemical shift variations induced on BtRA by loop 1 swapping in BtRA$\text{HsR3}$. Residues in gray are either unassigned or belong to the swapped loop. Chemical shift variations for each residue are indicated by both the color and the width of the cartoon putty. Figure was made with PyMOL using the bovine RNase A structure (PDB ID 7RSA).
BtRA and the chimera showed that chemical shift variations arising from loop swapping are primarily localized in the V2 domain (Figure 2.4 B). However, residues that acquire conformational exchange in the V1 domain (Figure 2.2 B) do not exhibit significant chemical shift variations. These observations suggest that alterations in regions displaying functionally relevant motions results in the modulation or loss of conformational exchange for residues in these regions. Overall, these results suggest that conformational exchange profiles are conserved between members within the EAR-like sub-family, perhaps suggesting a role in the shared biological or chemical function, while distinctly different dynamical fluctuations are observed for functionally distinct homologs from other phylogenetic clusters.

2.4.C Conformational dynamics on the microsecond (µs) time-scale

Atomic level dynamical properties of all 23 RNase homologs were probed using MD simulations to determine the detailed dynamical characteristics on the µs time-scale (Figure 2.5). Root mean square fluctuations (RMSF) associated with the slowest ten quasi-harmonic modes (RMSF₁₀) were used to characterize the intrinsic dynamics of the proteins. Previously, such an approach was used for other family members and was shown to match the dynamics from NMR studies and other techniques.¹¹⁰ One of the issues with comparing dynamical patterns between proteins sequences of different lengths is the presence of gaps associated with insertions and deletions. To address this, we represented dynamical profiles of consensus sequences that include gaps in the multiple sequence alignment (Figure 2.1 A).

A comparison of the conformational dynamics on the µs time-scale showed similarity in the RMSF profiles of RNases within the same sub-family members, while differences were observed between sub-families (Figure 2.5), consistent with the qualitative comparison of the conformational exchange profiles from experiments (Figure 2.2). The RNase A-like proteins, which share significant sequence similarity with BtRA, showed greater similarity in dynamical patterns between each other. Differences in the dynamical patterns were restricted to the loop L1 region (residues 20-30, consensus sequence numbering, Figure 2.10 (see Appendix 2.7)), which showed the largest fluctuations for most members in this group except for the Mus musculus (mouse) RNase A (MmR1), which showed diminished dynamics in this region relative to other members. Characterization of hydrogen bonding lifetimes in MmR1 showed the interaction between the side chains of Ser19 of loop1 and Thr82 of the β4 strand, occurring in ~81% of the
simulation ensemble. This strong hydrogen bonding interaction may be dampening the dynamics of loop1 observed for MmR1.

**EAR-like** sub-family members exhibit similar dynamical patterns in all regions. RNases of this group are characterized by a truncated loop L1, which incidentally showed the largest fluctuations in the **RNase A-like** sub-family. Swapping loop 1 from HsR3 into BtRA resulted in a chimera (BtRA_{HsR3}), which displayed a loss of dynamics in this loop, in addition to an overall reduction in the dynamics throughout the protein relative to BtRA (Figure 2.6 A). Human RNase 7 (HsR7) showed larger fluctuations throughout the protein while preserving the dynamical profile exhibited by other members of the sub-family. **Angiogenin-like** members are characterized by a truncated loop L4, and showed differences in dynamics of loop L2. Members not classified in the groups above (clustered into **Others**), as expected, showed large variations within the group compared to members within the other three sub-families. Large dynamical fluctuations were observed in loops L2, L4 and L6 for some members of the group, while other members showed deletions in these regions.

### 2.4.D Quantitative characterization of dynamical conservation

To determine if the similarities (and differences) observed within (between) sub-families are statistically significant, we performed a quantitative comparison of the dynamical profiles of sequences by performing pairwise correlations for all pairs of sequences (Figure 2.7). The pairwise Pearson’s correlation coefficient for all 23 sequences (Figure 2.7 A) showed larger correlations between members within, rather than between the three functionally distinct phylogenetic sub-families, except the **Others** cluster which showed small intra-cluster correlations. To better elucidate these comparisons within and between sub-families, we calculated the average value for these correlations, with the standard deviations presented in parentheses (Figures 2.7 B-D). Comparison of these averages showed that correlations within sub-families are significantly larger than those between sub-families for the three functionally distinct RNase sub-families (Figure 2.7 B). Overall, these observations highlight the conservation of dynamical patterns in homologous sequences sharing distinct functional properties, further supporting NMR experiments within and between sub-family members.

Sequence and structural alignments showed insertions/deletions of residue positions in different regions along the primary sequence of the RNases (Figure 2.1). To determine if these insertions
**Figure 2.5: Microsecond conformational dynamics of RNase homologs.** Root mean square fluctuations corresponding to the C$_\alpha$ displacements in the top 10 quasi-harmonic modes (RMSF$_{10}$), for the 23 proteins grouped into the four phylogenetic clusters described in Figure 2.1. A consensus sequence with gaps is used (see Figure 2.1 A), where gaps are represented as dotted regions for each protein. The panel on the right shows representative RNase homologs from each phylogenetic group using a tube representation where the thickness of the tube corresponds to the flexibility of residues in each protein, with thicker tubes corresponding to flexible regions and thinner tubes representing less flexible regions. The dynamical range represented using the color spectrum is consistent across all sequences with the blue and red ends of the spectrum corresponding to low and high dynamic regions, respectively. The representative structures correspond to *B. taurus* RNase A (BtRA) and *H. sapiens* RNase 1 (HsR1); *H. sapiens* RNase 3 (HsR3) and *P. pygmaeus* RNase 3 (PpR3); *H. sapiens* angiogenin (HsR5) and *Mus musculus* RNase 5 (MmR5); and *H. sapiens* RNase 4 (HsR4) and *Rana pipiens* RNase (RpRx), respectively. N-terminal residues, which showed larger RMSFs (see left panel), are hidden for clarity. Tube representations for all RNases are shown in Figure 2.11 (see Appendix 2.7).
Figure 2.6: Dynamical properties of BtRA-HsR3 chimera. (A) Comparison of the root mean squared fluctuations of the ten slowest modes (RMSF$_{10}$) between BtRA (black) and the chimera (red). On the right, RMSFs are shown using a tube representation for BtRA and the chimera, where the thickness of the tube corresponds to the flexibility of residues in each protein, with thicker tubes corresponding to flexible regions and thinner tubes representing less flexible regions. B) Dynamical cross-correlation maps (DCCMs) for BtRA (left), with regions displaying significant (anti-)correlations displayed identified using boxes, and the BtRA$_{HsR3}$ chimera (right). Loops 1, 2, 4 and 6 are highlighted in yellow.
and deletions influence the functional identity of these homologous sequences, we calculated the average correlations for consensus and non-consensus residue positions separately. Consensus positions correspond to residue positions that are preserved in all sequences (86 amino acid positions) while non-consensus positions represent positions with insertions or deletions in at least one of the sequences in the dataset (66 amino acid positions). A comparison of the correlations within sub-families showed similar trends for both consensus and non-consensus positions (Figure 2.7 C, D). Interestingly, a comparison of correlations between sub-families for non-consensus residues showed markedly smaller correlations (grey in Figure 2.7) between the three functionally distinct sub-families, except Angiogenin-like sub-family which showed similar average correlations within the sub-family and with RNase A-like sub-family on this time-scale (Figure 2.7 D).

Overall, these observations suggest that non-consensus positions influence the distinct dynamical patterns observed among the functionally divergent phylogenetic sub-families. It could potentially be counter-argued that dynamical patterns will always be most influenced at non-consensus residue positions when comparing phylogenetically distinct proteins. However, it is important to remember that the physicochemical identity of residues determines the extent and nature of dynamical motions and rotameric conformer populations. Experimental characterization of the BtRA_{HsR3} chimera (with L1 loop region from HsR3 inserted into BtRA) revealed significant changes in the dynamics of this region and the V2 domain relative to BtRA (Figure 2.4). It is tempting to speculate that as loop regions are flexible, any sequence in the L1 loop region of BtRA could be expected to show large dynamical motions. On the contrary, the results from chimera characterization indicate that L1 loop residues in BtRA_{HsR3} show dynamical behavior similar to L1 in HsR3 rather than BtRA.

Analysis of MD simulations allow the characterization of atomic-scale details, including residue level dynamical interactions. A comparison of dynamic residue-residue cross-correlations for the different RNase homologs shows distinct variations among members. Detailed characterization of dynamical cross-correlations was performed for members with experimental dynamical information (correlation maps with black outlines in Figure 2.12 see Appendix 2.7). BtRA of the RNase A sub-family showed strong anti-correlations (value < -0.4) for residues 65-72 of loop L4 (positions 75-84, consensus sequence numbering, Figure 2.10 in Appendix 2.7) with residues near the active site and the catalytic K41 (position 51) and correlations (value > 0.4) with residues near H119 (position 142). Anti-correlations were also observed between residues near K41 and
Figure 2.7: Quantitative characterization of dynamical similarities. Pairwise Pearson's correlations for the 23 RNase homologs classified into four phylogenetic clusters. (A) Correlation coefficients for each pair of sequences, color coded based on strong correlations (red) to weaker or no correlations (grey). (B-D) Average correlations, calculated by averaging the correlation coefficients for sequences compared for: all residue positions (B), consensus positions (C), and non-consensus positions, corresponding to insertions or deletions in sequences (D). The diagonal elements in B-D correspond to correlations within phylogenetic sub-families while off-diagonal elements correspond to correlations between the sub-families. Standard deviations for average correlation calculations (B-D) are shown in parentheses. Abbreviations R, E, A and O used in B-D correspond to phylogenetic clusters RNase A-like, EAR-like and Angiogenin-like, and Others, respectively.
residues 105-110 (positions 119-124) and those near catalytic H119. Further, strong correlations were observed for residues near, and including H12 (position 17) and residues 13-23 of loop 1 (positions 18-29) with residues around H48 and T82 (positions 58 and 94, respectively), which form a hydrogen bonding network shown previously to be involved in the allosteric modulation of catalysis and product release in this enzyme.\textsuperscript{137,216}

Interestingly, loop swapping in the chimera (BtRA\textsubscript{HsR3}) resulted in changes throughout the protein, in regions far from loop 1, including significant changes near active-site residues. Specifically, we observed a loss of correlated motions between loop 4 and the C-terminal region in the vicinity of the active-site residue His119, and loss of anti-correlated motions of loop 4 and the vicinity of Lys41 (another active-site residue) and between residues near His119 and Lys41 (Figure 2.6 B). HsR3 of the \textit{EAR-like} sub-family showed strong anti-correlations between residues near the catalytic residue H15 and residues surrounding residues 86-97 of loop 6 (positions 99-110, Figure 2.12) and residues 27-38 (positions 36-51), while correlations were observed with residues 42-46 of \beta-strand 1 (positions 55-59). Residues in the vicinity of and including the catalytic K38 (position 51) showed anti-correlations near residues 60-65 of loop 4 (positions 73-78) and residues near catalytic residue H128 (position 142). Further, H128 also showed correlations with residues near H15 (position 17) and loop 4, and anti-correlations with residues near K38. Residues 28-36 (positions 37-49) showed anti-correlations with residues 98-102 (positions 111-115).

Contrary to members from the other phylogenetic sub-families, HsR5 of the \textit{Angiogenin-like} sub-family showed significantly weaker (anti)correlations between residues. The lack of conformational dynamics of HsR5 and the resulting absence of correlated motions is consistent with our observations from relaxation dispersion experiments, which showed the absence of motions on the millisecond time-scale in this enzyme (Figure 2.2 C). HsR4 of the \textit{Others} sub-family showed numerous dynamic (anti)correlations between different regions of the protein, suggesting enhanced correlated dynamics throughout the protein, consistent with NMR relaxation dispersion experiments, which showed extensive conformational exchange in this protein (Figure 2.2 D). Dynamical correlations of all 23 RNase homologs is shown in Figure 2.12 (Appendix 2.7).
2.4.E Faster time-scale conformational dynamics in functionally distinct RNase sub-families

To determine if MD simulations capture the motions observed experimentally on the faster ps-ns time-scale, we compared the square of the generalized order parameter (S2). Using ShiftX2 (Han et al., 2011), we calculated the backbone chemical shifts for two representative members each from the three functionally distinct sub-families defined in Figure 2.1 C. NMR chemical shifts for the selected RNases were obtained from the BMRB (Ulrich et al., 2008). Order parameters were calculated from the chemical shifts using TALOS+ (Shen et al., 2009). Comparison of the order parameters showed qualitative agreement between the experimental (black) and MD (red) data (Figure 2.8), suggesting that MD simulations capture the motions observed experimentally.

2.5 Discussion

Increasing evidence from characterization of discrete enzyme systems suggests the important role of conformational motions in the catalytic function in these systems. Despite tremendous progress in our understanding of the role of conformational motions in function, evolutionary conservation of dynamical properties on functionally relevant time-scales among structural and functional homologs within enzyme families remains largely uncharacterized. Numerous studies using normal mode analysis have shown the conservation of backbone flexibility in homologous proteins sharing a common structural framework. Experimental and computational approaches further showed the conservation of dynamics in selected proteins within enzyme families. However, systematic characterization of superfamily-wide conservation of dynamical properties on the functionally relevant µs - ms time-scales is limited.

In this study, we combined $^{15}$N-CPMG NMR relaxation dispersion experiments and detailed MD simulations to systematically characterize dynamical properties of diverse pancreatic-type RNases on the µs - ms time-scale. Using a set of >20 representative RNases, chosen based on the availability of three-dimensional structures in the protein data bank, we probed the conservation of dynamical patterns in the homologous RNases. Phylogenetic analysis led to the grouping of sequences into four clusters – 16 of these sequences forming part of three functionally distinct sub-families (Figure 2.1 C), where individual members within sub-families also share common biological function distinct from those of other sub-families. Little is known about the biological function of the remaining seven members that were grouped into the Others cluster.
Figure 2.8: Faster time-scale motions in representative RNases. Comparison of the calculated order parameters for the experimental (black) and MD (red) data for two representative RNases each from the three functionally distinct RNase sub-families (RNase A-like, EAR-like and Angiogenin-like) defined in Figure 2.1 C.
Slowest dynamical modes were previously shown to display the largest evolutionary conservation
with these modes being the most robust to amino acid substitutions. Through quantitative
characterization of the dynamical properties of the slowest modes from MD simulations, we show
distinct variations in the dynamical patterns between different members of the RNase superfamily.
Conservation of dynamical patterns was observed for sequences clustered into phylogenetic sub-
families (Figure 2.5) that incidentally also shared similar biological functions. This conservation in
the dynamical profiles within sub-families was also observed on the ms time-scale probed by
NMR for select RNases (Figure 2.2). These observations are consistent with previous studies
which showed that enzymes sharing similar functions also share similar reaction-promoting
dynamics.

Rapid changes through evolution in dynamical properties were previously shown to be linked to
functional divergence in proteins. Sequence and structural alignments of RNases used in this
study showed variations, corresponding to insertions/deletions, in different regions along the
primary sequence of RNases (Figure 2.1). A comparison of dynamical patterns of just these non-
consensus amino acid positions showed notable differences between sub-families (Figure 2.7 D).
These observations suggest that dynamical differences observed between sequences sharing
distinct functional properties could be linked to the type of residues in non-consensus amino acid
positions, which may influence the functional identity of these sequences.

The use of chimeric constructs such as BtRA<sub>HsR3</sub> investigated in this study serve as an important
control system. It is widely discussed that the overall enzyme dynamics (particularly at slow time-
scales) is determined by the structural fold. Comparison of the conformational exchange profiles
for the chimeric and wild-type RNases (Figure 2.2, 2.4) indicate that for enzymes sharing the
same fold, significant differences could arise from changes in non-conserved or variable regions.
However, it is not sufficient to say that loop regions are the most dynamical regions. The extent
and range of motions depends on the nature and type of residues in the loop regions. Indeed,
characterization of dynamical correlations from molecular dynamics simulations of the BtRA<sub>HsR3</sub>
chimera shows the long-range effect of loop 1 swapping on loop 4 and residues in the vicinity of
active-site residues His119 and Lys41 (Figure 2.6 B).

Each enzyme could provide a unique set of local dynamical motions depending on the sequence
(such as seen in the case of loop 1 in the two parents and the chimera, Figure 2.4). Characterization of these differences and its relation to changes in catalytic efficiency requires
further investigation. Further, it should be noted that even in the conserved regions of fold, different family members could exhibit distinct dynamics. For example, see Figure 2.2 and related text for differences in the dynamics of the V2 domain in different sub-families. These observations open new avenues for probing the effect of conformational exchange associated with loop swapping between members from the different sub-families. Of notable interest is the effect on the conformational dynamics and catalytic turnover of swapping loop 4 from BtRA on HsR5, which is characterized by a truncated loop 4 and a $10^5$-fold lower catalytic efficiency relative to BtRA. While swapping the loop 4 residues from HsR5 on to BtRA and *vice versa* have been performed before, the effect of this swapping on the conformational exchange and the resulting impact on catalysis have not been characterized \(^{232, 233, 245}\).

HsR5 exhibits at least $10^5$-fold weaker ribonucleolytic activity than bovine RNase A against standard RNA substrates \(^{232}\), although this ribonucleolytic activity remains essential for angiogenesis. This low enzyme activity has been structurally and mutationally rationalized by the obstructive positioning of a C-terminal residue segment of HsR5 (residues 116-123, primarily Glu117), which is occluding a pyrimidine binding sub-site that is normally solvent exposed in homologous RNases \(^{246}\). This ‘trademark’ blockage of the active site in HsR5 demonstrates that a conformational change of this C-terminal tail is required to alleviate the occluded pocket for proper RNA substrate recognition and ribonucleolytic activity to occur. Since movement of the C-terminal tail is essential for the ribonucleolytic activity of HsR5, this was not totally unexpected considering that the catalytic rate of this enzyme occurs on a much slower time-scale than its RNase homologs. For instance, HsR5 exhibits a $k_{cat}/K_m = 1.1 \text{ M}^{-1}\text{s}^{-1}$ against dinucleotide RNA substrate UpA, relative to $k_{cat}/K_m = 4,000,000 \text{ M}^{-1}\text{s}^{-1}$ in RNase A \(^{232}\). Additionally, no less than 19 unique protein partners have been experimentally reported to interact with HsR5 for a wide variety of biological processes involving cell migration, adhesion, invasion, proliferation, and apoptosis \(^{52}\). This could suggest that the ribonucleolytic activity of HsR5 is carefully programmed in the cellular context, perhaps only triggered in the presence of the proper protein binding partner and/or after translocation to the right cellular compartment. Conformational exchange of the C-terminal tail could thus be closely modulated or allosterically controlled by an uncharacterized protein partner. Although we did not observe $^1\text{H}-^{15}\text{N}$ HSQC resonances that could indicate residues experiencing exchange on a slower time-scale, future chemical exchange saturation transfer (CEST), ZZ-exchange, and/or H/D exchange experiments could provide additional clues on slow time-scale motions experienced by HsR5.
In summary, characterization of dynamical properties of diverse RNases showed their conservation among sequences that also share similar biological functions. Interestingly, mutations of sub-family sequences in regions showing large dynamical motions have been shown to alter their biological function (Table 2.2). The similarity in dynamics for members with common biological function requires further investigation, especially for sub-families where the connection between the biological and chemical function is unclear. The present results provide vital clues that conservation of dynamics as a part of a protein fold may have a broader significance beyond enzyme catalysis.
Table 2.2: Amino acid substitutions and their effect on the catalytic and biological activities of select RNase homologs.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Mutation</th>
<th>Localization on structure</th>
<th>Effect on $k_{cat}/K_M$</th>
<th>Other effects</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>HsR3</td>
<td>R121A</td>
<td>Loop 7 (V1)</td>
<td>No effect</td>
<td>Reduced bactericidal activity against Gram-positive <em>S. aureus</em></td>
<td>42</td>
</tr>
<tr>
<td>HsR3</td>
<td>W10K</td>
<td>Helix 1</td>
<td>2-fold decrease</td>
<td>Reduced bactericidal activity against Gram-positive <em>S. aureus</em> and reduced membrane-permeating capacity</td>
<td>42</td>
</tr>
<tr>
<td>HsR3</td>
<td>R75A/F76A</td>
<td>Loop 5 (V1)</td>
<td>2-fold decrease</td>
<td>Reduced bactericidal activity against Gram-negative <em>E. coli</em></td>
<td>42</td>
</tr>
<tr>
<td>HsR3</td>
<td>R121A/Y122A</td>
<td>Loop 7 (V1)</td>
<td>1.25-fold increase</td>
<td>Reduced bactericidal activity against Gram-positive <em>S. aureus</em></td>
<td>42</td>
</tr>
<tr>
<td>HsR3</td>
<td>W35A/R36A</td>
<td>Loop 2 (V2)</td>
<td>1.6-fold increase</td>
<td>Reduced bactericidal activity against both <em>E. coli</em> and <em>S. aureus</em>, and reduced membrane-permeating capacity</td>
<td>42</td>
</tr>
<tr>
<td>HsR3</td>
<td>R101A/R104A</td>
<td>β-strand 4 (V2)</td>
<td>2-fold increase</td>
<td>Reduced bactericidal activity against Gram-negative <em>E. coli</em> and reduced membrane-permeating capacity</td>
<td>42</td>
</tr>
<tr>
<td>HsR3</td>
<td>(115-122)EDN</td>
<td>Loop 7 (V1)</td>
<td>No effect</td>
<td></td>
<td>42</td>
</tr>
<tr>
<td>HsR3</td>
<td>Δ(115-122)</td>
<td>Loop 7 (V1)</td>
<td>1.5-fold decrease</td>
<td>Reduced bactericidal activity against Gram-positive <em>S. aureus</em></td>
<td>42</td>
</tr>
<tr>
<td>Mouse EAR 11</td>
<td>K35R</td>
<td>Loop 2 (V2)</td>
<td>Catalytically inactive (catalytic lysine)</td>
<td>No effect on leukocyte chemotaxis</td>
<td>247</td>
</tr>
<tr>
<td>HsR2</td>
<td>W7A</td>
<td>Helix 1</td>
<td>Abolishment of EDN C-mannosylation (W7 was identified as the C-mannosylation site)</td>
<td>248</td>
<td></td>
</tr>
<tr>
<td>HsR2</td>
<td>W10A</td>
<td>Helix 1</td>
<td>Abolishment of W7 C-mannosylation</td>
<td>248</td>
<td></td>
</tr>
<tr>
<td>HsR2</td>
<td>W10F</td>
<td>Helix 1</td>
<td>Partial rescue of W7 C-mannosylation abolishment</td>
<td>248</td>
<td></td>
</tr>
<tr>
<td>HsR2</td>
<td>R132T</td>
<td>β-strand 6 (V1)</td>
<td>12-fold decrease</td>
<td></td>
<td>249</td>
</tr>
<tr>
<td>HsR2</td>
<td>S64R</td>
<td>Loop 4 (V1)</td>
<td>74-fold decrease</td>
<td></td>
<td>249</td>
</tr>
<tr>
<td>HsR2</td>
<td>S64R/R132T</td>
<td>V1</td>
<td>12-fold decrease</td>
<td>Reduced antiviral activity against RSV</td>
<td>249</td>
</tr>
<tr>
<td><em>Aotus trivirgatus EDN</em></td>
<td>T132R</td>
<td>β-strand 6 (V1)</td>
<td>1.7-fold decrease</td>
<td></td>
<td>249</td>
</tr>
<tr>
<td><em>Aotus trivirgatus EDN</em></td>
<td>R64S</td>
<td>Loop 4 (V1)</td>
<td>1.5-fold increase</td>
<td></td>
<td>249</td>
</tr>
<tr>
<td><em>Aotus trivirgatus EDN</em></td>
<td>R64S/T132R</td>
<td>V1</td>
<td>15-fold increase</td>
<td>No increase in antiviral activity against RSV (opposite mutation than in HsR2 S64R/R132T)</td>
<td>249</td>
</tr>
<tr>
<td>HsR2</td>
<td>K38R</td>
<td>Loop 2 (V2)</td>
<td>Catalytically inactive (catalytic lysine)</td>
<td>No antiviral activity against RSV</td>
<td>36</td>
</tr>
<tr>
<td>HsR2</td>
<td>K38R</td>
<td>Loop 2 (V2)</td>
<td>Catalytically inactive (catalytic lysine)</td>
<td>No antiviral activity against RSV</td>
<td>36</td>
</tr>
</tbody>
</table>
Table 2.2 continued

<table>
<thead>
<tr>
<th>Protein</th>
<th>Mutation</th>
<th>Localization on structure</th>
<th>Effect on $k_{cat}/K_M$</th>
<th>Other effects</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>HsR3</td>
<td>R97T (natural variant)</td>
<td>Loop 6 (V2)</td>
<td>Reduced cytotoxicity and neurotoxicity</td>
<td></td>
<td>45</td>
</tr>
<tr>
<td>HsR3</td>
<td>R22A</td>
<td>Loop 1 (V2)</td>
<td>1.6-fold decrease</td>
<td>Large reduction in cytotoxicity; small reduction in bactericidal activity against Gram-negative E. coli; large reduction in bactericidal activity against Gram-positive B. subtilis; large reduction of antiparasitic activity</td>
<td>111</td>
</tr>
<tr>
<td>HsR3</td>
<td>R34A</td>
<td>Loop 2 (V2)</td>
<td>1.25-fold decrease</td>
<td>Large reduction in cytotoxicity; small reduction in bactericidal activity against Gram-negative E. coli; large reduction in bactericidal activity against Gram-positive B. subtilis; large reduction of antiparasitic activity</td>
<td>111</td>
</tr>
<tr>
<td>HsR3</td>
<td>R61A</td>
<td>Loop 4 (V1)</td>
<td>1.4-fold decrease</td>
<td>Reduction in cytotoxicity; small reduction in bactericidal activity against Gram-negative E. coli; large reduction in bactericidal activity against Gram-positive B. subtilis; reduction of antiparasitic activity</td>
<td>111</td>
</tr>
<tr>
<td>HsR3</td>
<td>H64A</td>
<td>Loop 4 (V1)</td>
<td>1.2-fold increase</td>
<td>Reduction in cytotoxicity; small reduction in bactericidal activity against Gram-negative E. coli; large reduction in bactericidal activity against Gram-positive B. subtilis; reduction of antiparasitic activity</td>
<td>111</td>
</tr>
<tr>
<td>HsR3</td>
<td>R77A</td>
<td>Loop 5 (V1)</td>
<td>1.9-fold increase</td>
<td>Reduction in cytotoxicity; small reduction in bactericidal activity against Gram-negative E. coli; large reduction in bactericidal activity against Gram-positive B. subtilis; reduction of antiparasitic activity</td>
<td>111</td>
</tr>
<tr>
<td>HsR3</td>
<td>R22A/R34A</td>
<td>V2</td>
<td>2-fold increase</td>
<td>Large reduction in cytotoxicity; large reduction in bactericidal activity against Gram-negative E. coli; large reduction in bactericidal activity against Gram-positive B. subtilis; large reduction of antiparasitic activity</td>
<td>111</td>
</tr>
<tr>
<td>HsR3</td>
<td>R22A/R61A</td>
<td>2-fold increase</td>
<td>Reduction in cytotoxicity; large reduction in bactericidal activity against Gram-negative E. coli; large reduction in bactericidal activity against Gram-positive B. subtilis; reduction of antiparasitic activity</td>
<td></td>
<td>111</td>
</tr>
<tr>
<td>HsR3</td>
<td>R22A/R77A</td>
<td>4-fold increase</td>
<td>Reduction in cytotoxicity; abolition of bactericidal activity against Gram-negative E. coli; large reduction of bactericidal activity against Gram-positive B. subtilis; reduction of antiparasitic activity</td>
<td></td>
<td>111</td>
</tr>
<tr>
<td>Protein</td>
<td>Mutation</td>
<td>Localization on structure</td>
<td>Effect on $k_{cat}/K_M$</td>
<td>Other effects</td>
<td>Ref</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>---------------------------</td>
<td>--------------------------</td>
<td>---------------</td>
<td>----</td>
</tr>
<tr>
<td>HsR3</td>
<td>R22A/H64A</td>
<td>5-fold increase</td>
<td>No reduction in cytotoxicity; small reduction in bactericidal activity against Gram-negative E. coli; no reduction in bactericidal activity against Gram-positive B. subtilis; reduction of antiparasitic activity</td>
<td></td>
<td>111</td>
</tr>
<tr>
<td>HsR3</td>
<td>R34A/H64A</td>
<td>1.25-fold decrease</td>
<td>Large reduction in cytotoxicity; small reduction in bactericidal activity against Gram-negative E. coli; large reduction in bactericidal activity against Gram-positive B. subtilis; large reduction of antiparasitic activity</td>
<td></td>
<td>111</td>
</tr>
<tr>
<td>HsR3</td>
<td>R61A/H64A</td>
<td>V1</td>
<td>Large reduction in cytotoxicity; large reduction in bactericidal activity against Gram-negative E. coli; large reduction in bactericidal activity against Gram-positive B. subtilis; large reduction of antiparasitic activity</td>
<td></td>
<td>111</td>
</tr>
<tr>
<td>HsR3</td>
<td>H64A/R77A</td>
<td>V1</td>
<td>Reduction in cytotoxicity; large reduction in bactericidal activity against Gram-negative E. coli; large reduction in bactericidal activity against Gram-positive B. subtilis; large reduction in antiparasitic activity</td>
<td></td>
<td>111</td>
</tr>
</tbody>
</table>
2.7 Appendix

**Figure 2.9: Conformational exchange rates of selected RNases.** Cartoon representation of *Homo sapiens* EDN (HsR2), *Homo sapiens* ECP (HsR3), *Pongo pygmaeus* ECP (PpR3), *Macaca fascicularis* ECP (MfR3), *Bos taurus* RNase A (BtRA), *Homo sapiens* angiogenin (HsR5) and *Homo sapiens* RNase 4 (HsR4). Structures are colored using a blue-white-red spectrum corresponding to residues exhibiting exchange rates ($k_{ex}$) ranging from 0 to 1000 s$^{-1}$. Residues experiencing enhanced conformational exchange appear in the red end of the spectrum while residues displaying no conformational exchange are in the blue end of the spectrum. Exchange rates ($k_{ex}$) were calculated for each individual residue by dual fitting 500 and 800 MHz CPMG data to the full single quantum relaxation-compensated CPMG equation (Manley and Loria, 2012).
Figure 2.10: Consensus sequence corresponding to the multiple sequence alignment.
Sequences for 23 RNases shown in Figure 2.1 A are included. The first column and row labels correspond to sequence identifiers and amino acid positions, respectively.
Figure 2.11: μs dynamics of all 23 RNase homologs. RMSFs for each protein from each of the four phylogenetic groups are shown using a tube representation where the thickness of the tube corresponds to the flexibility of residues in each protein, with thicker tubes corresponding to flexible regions and thinner tubes representing less flexible regions. The dynamical ranges represented using the color spectrum is consistent across all sequences with the red and blue ends of the spectrum corresponding to more and less dynamic regions, respectively. Note that N-terminal residues, which showed larger RMSFs (see left panel), are hidden for clarity.
Figure 2.12: Dynamic cross-correlation maps (DCCMs) for the 23 RNase homologs. Values >0.4 (blue end of the spectrum) represent correlated motions in the protein, while values < -0.4 (red end of the spectrum) represent anti-correlated motions. DCCMs identified using black outlines are discussed in the section “Microsecond dynamics” of the manuscript. Regions displaying (anti-)correlations are highlighted in the boxed DCCMs in A and B.
Chapter 3  Nucleotide Substrate Binding Characterization in Human Pancreatic-type Ribonucleases
This chapter is a slightly revised version of a manuscript by the same title published in the journal, *PLoS ONE* published in 2019

Reprinted (adapted) with permission from Khushboo Bafna,¹ Chitra Narayanan,² S. Chakra Chennubhotla,³ Nicolas Doucet,²,⁴ Pratul K. Agarwal⁵.*

¹Genome Science and Technology, University of Tennessee, Knoxville, TN 37996, USA.
²INRS-Institut Armand-Frappier, Université du Québec, 531 Boul. des Prairies, Laval, Québec H7V 1B7, Canada.
³Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, Pennsylvania, 15260, USA.
⁴PROTEO, the Quebec Network for Research on Protein Function, Structure, and Engineering, 1045 Avenue de la Médecine, Université Laval, Québec, QC, G1V 0A6, Canada.
⁵Department of Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee, 37996, USA.

*Corresponding author: pratul@agarwal-lab.org

This author contributed the following to the manuscript: (1) molecular modeling, (2) molecular dynamics simulations, (3) data analysis and organization and (4) manuscript writing.
3.1 Abstract

Human genome contains a group of more than a dozen similar genes with diverse biological functions including antiviral, antibacterial and angiogenesis activities. The characterized gene products of this group show significant sequence similarity and a common structural fold associated with binding and cleavage of ribonucleic acid (RNA) substrates. Therefore, these proteins have been categorized as members of human pancreatic-type ribonucleases (hRNases). hRNases differ in cell/tissue localization and display distinct substrate binding preferences and a wide range of ribonucleolytic catalytic efficiencies. Limited information is available about structural and dynamical properties that influence this diversity among these homologous RNases. Here, we use computer simulations to characterize substrate interactions, electrostatics and dynamical properties of hRNases 1-7 associated with binding to two nucleotide substrates (ACAC and AUAU). Results indicate that even with complete conservation of active-site catalytic triad associated with ribonucleolytic activity, these enzymes show significant differences in substrate interactions. Detailed characterization suggests that in addition to binding site electrostatic and van der Waals interactions, dynamics of distal regions may also play a role in binding. Another key insight is that a small difference in temperature of 300 K (used in experimental studies) and 310 K (physiological temperature) shows significant changes in enzyme-substrate interactions.

3.2 Introduction

A set of related human genes encoding for enzymes that cleave ribonucleic acid (RNA) substrates also exhibit diverse biological functions such as angiogenesis, antiviral, antibacterial, and/or cytotoxic activities. Eight of the encoded enzymes (*canonical members*) show conservation of two active-site histidine residues and a lysine associated with the catalytic mechanism of ribonucleolytic cleavage. An additional five catalytically inactive (*non-canonical members*) pseudogenes were also identified in the human genome. Collectively, these 13 proteins exhibit 28% sequence identity (40% among the 8 canonical members). All structurally characterized members share a common structural fold while displaying a diverse range of catalytic efficiencies and biological activities. These proteins are referred to as human ribonucleases (hRNases) due to their conserved fold and ribonucleolytic function, which was first discovered and characterized for bovine ribonuclease A (bRNaseA). As enzymes, hRNases catalyze the transphosphorylation and hydrolysis of the phosphodiester bond in single-stranded RNA substrates, although cleavage of double-stranded RNA has also been reported.
broadly, the hRNases and bRNaseA belong to pancreatic-type ribonucleases (RNases) superfamily of proteins with about 1500 members, mostly from mammals.

RNase superfamily members bind and catalyze the phosphodiester bond cleavage in a wide variety of nucleotide substrates of different lengths and sequences. Substrate binding and kinetic assays for some hRNases have been performed using a number of substrates with varying nucleotide sequences and lengths, resulting in a wide variety of binding affinities and catalytic efficiencies being reported. Structural studies have provided detailed information for binding sites for the phosphate groups (P₀, P₁, P₂), pyrimidine (B₁) and purine (B₂) nucleotide bases in the active-site of bovine RNase A, the archetypal member of this superfamily (Figure 3.1). The past studies have suggested substrate preferences at the B₁ pyrimidine binding site for some of the human RNases: hRNase1 showed a preference for cytosine (C) over uridine (U), while hRNases 2, 3, 4 and 6 showed a preference for U over C. In contrast, the B₂ purine binding site was shown to prefer adenosine (A) in these hRNases. Substrate preferences for the other hRNases remains unclear. While RNase homologs exhibit specific nucleotide sub-site preferences in the context of structurally distinct active-site pockets, reports have illustrated that binding affinities differ over several orders of magnitude between all hRNase members (Kₘ ranging from about 2 µM to 250 mM and kₗₑ₅ ranging from 0.003 to 2400 s⁻¹), see Table 1.2 in chapter 1 for details. As hRNases exhibit distinct substrate specificities and bind a wide range of non-specific RNA sequences with micromolar affinities, the diverging ribonucleolytic preferences and catalytic activities has created challenges in understanding their chemical and designated biological functions at the molecular level. Therefore, considerable interest exists in the characterization of the structure, dynamics and catalytic mechanisms of hRNases, in order to gain a better understanding of their biological functions in the cell.

One of the most intriguing aspects of hRNases is the wide range of catalytic efficiencies associated with ribonucleolytic activities observed for these homologs, which differ by as much as 10⁶-fold. Interestingly, the time-scales associated with their intrinsic dynamics (corresponding to large-scale global conformational exchange) also varies over 5-6 orders of magnitude. The similarity in range of catalytic efficiencies as well as the observed intrinsic dynamics raises an interesting question regarding the relevance of dynamics in the catalytic mechanism. Recently, structural and dynamical characterization of a group of 23 members of the RNase superfamily based on nuclear magnetic resonance (NMR) and computational simulations (of apo proteins) provided unique insights into the conservation of dynamical
Figure 3.1: RNase-substrate binding sub-sites. Schematic representation based on the bRNaseA active-site. P0, P1 and P2 represent binding sub-sites for the phosphate groups of the RNA substrate backbone, while B1 and B2 correspond to nucleotide base sub-sites. The cleaved phosphodiester bond is indicated by an arrow. The nucleotide bases in pockets B1 (cytosine drawn, but uracil can also bind) and B2 (adenosine) are rendered based on published information about nucleotide preference in bRNaseA; for some hRNases, the information about nucleotide preference is unclear. Figure redrawn based on 7. The substrates investigated in this study are 5'-ACAC-3' (shown) and 5'-AUAU-3'. In the present work, simplified nucleotides numbering scheme A(-2)C/U(-1)A(1)C/U(2) is used with negative numbers corresponding to nucleotides away from cleaved phosphodiester bond decreasing in the 5' direction and positive numbers for nucleotides away from cleaved phosphodiester bond increasing in the 3' direction.
properties within this enzyme family. Members within sub-families, identified based on phylogenetic classification, showed similarity in structure, and more interestingly, similarity in biological function and intrinsic dynamics. Furthermore, the different sub-families corresponding to branches of evolution associated with gene duplication followed by functional adaptation show largely different dynamics. For a direct comparison of the catalytic efficiencies and dynamics of these hRNases, the use of identical substrates in similar conditions is essential to eliminate effects caused by inhomogeneous nucleotide sequences and/or experimental variations.

In this study, we use computer modeling to investigate the binding and interactions of 8 canonical members of the RNase superfamily (hRNases 1-7 and bovine RNase A) in the presence of two model single-stranded RNA substrates: the tetraribonucleotide sequences ACAC and AUAU. The selection of these model substrates is based on previous experimental data on nucleotide preferences in the B₁/B₂ sub-sites as well as structural and dynamical characterization of hRNases from our group. Using microsecond time-scale molecular dynamics (MD) simulations, we characterized the binding stability, hydrogen bonding, van der Waals and electrostatic interactions between the enzymes in complex with the two model substrates. The detailed comparative analysis provides new insights into the structural and dynamical properties that contribute to enzyme-nucleotide substrate interactions in these RNases. The results indicate that while the 8 enzymes display conservation of the catalytic triad and a number of active-site residues involved in substrate interactions and stability, their substrate preference and binding affinities differ significantly. These results further indicate that, in addition to the conserved catalytic triad, other active-site residues play important roles in enzyme-substrate interaction and stability, and therefore, could have important consequences for differences in the biological function between functional and structural enzyme homologs. Somewhat surprisingly, a temperature difference from 300 K to 310 K shows significant differences in substrate stability behavior in these enzymes. This is intriguing because the vast majority of prior experimental enzyme characterization were performed at room temperature (~300 K), while the relevant physiological temperature for these enzymes is closer to the average human (~310 K) and bovine (~312 K) body temperatures. These observations highlight the effect of subtle changes in temperature on the structural and dynamical properties of enzymes and further illustrate the importance of the selection of temperature in experimental and computational studies that are representative of physiological conditions.
3.3 Materials and methods

Atomic coordinates: Seven human ribonuclease substrate complexes were prepared based on X-ray/NMR structures with the following PDB codes: 2K11 (hRNase1), 1GQV (hRNase2), 1QMT (hRNase3), 1RNF (hRNase4), 1ANG (hRNase5), 2HKY (hRNase7) and 7RSA (bRNaseA), while hRNase6 was crystallized in our group. The sequence alignment of these 8 proteins is provided in Figure 1.2 of Chapter 1. hRNase8 was not investigated in this study as structural information about this enzyme is not yet available. Each of these proteins were investigated complexed with two model substrates AUAA and ACAC, under explicit water conditions based on the procedure described below.

Computational systems preparation: Computational system preparation and all-atom MD simulations were performed using the AMBER simulation package. For preparation of the nucleotide substrate, the coordinates for the phosphate and ribose sugar backbone were adopted from the DNA ligand (ATAA) in PDB structure 1RCN. The nucleotide bases of the substrate template were modified to create the desired nucleotide (ACAC or AUAA), using AMBER's tleap modeling preparation module. AMBER's ff14SB force-field were used for modeling both protein residues and RNA substrates. The protein-ligand complex was neutralized with the addition of chloride (Cl⁻) counterions and solvated by immersing in a rectangular box (SPC/E water model) with periodic boundary conditions. The prepared system was equilibrated using the protocols developed in our group and described previously. These protocols were effective in fully resolving the steric clashes caused by the introduction of modeled substrate in the active site. AMBER's pmemd simulation engine was used for equilibration and production runs, with PME-method for long range electrostatics. All hydrogen bonds were constrained using SHAKE.

Histidine protonation states: There are two conserved histidine residues in the active-site of all RNases that participate in the phosphodiester bond cleavage mechanism. The protonation states for these residues were determined based on the past pH profiles for bRNaseA as well as structural interaction with modeled substrate and other residues, the two histidine residues His12 and His119 (bRNaseA numbering), were modeled in single (proton on N₆₁) and double protonated states respectively. All the other histidine residues were modeled in single protonated state for all 8 RNases.
**MD simulations:** Equilibration runs were performed under constant volume and temperature (NVT) conditions with a thermostat to bring the system to target temperature of 300 K or 310 K for production run (see below). Production runs of 0.5 µs (microsecond) were performed under constant volume and energy (NVE) simulation conditions using a 2 femtoseconds (fs) time-step. Production runs under NVE ensemble were performed as this ensemble offers better computational stability and performance for longer MD simulations. All simulations were performed using NVIDIA graphics cards with CUDA-enabled version of AMBER’s *pmemd* MD simulation engine. The default precision in the *pmemd.cuda_SPFP* binary was used from AMBER v14 package.

Two independent sets of simulations were performed at 300 K and 310 K (no thermostat was applied during production runs). Furthermore, to ensure the reproducibility of the results from this study, two independent trajectories were computed at each temperature (labeled as 300 K 1, 300 K 2, 310 K 1, and 310 K 2 in the remaining text and Table 3.1). For the set 2, after equilibration the system was minimized (100,000 steps) followed by MD (1000 ps) to slowly readjust the temperature before the production runs were performed. At each temperature, for each of the unique RNase-substrate complex, 2,000 structures (1,000 from each trajectory, 500 picoseconds apart) were stored and analyzed for structural, electrostatic and dynamical analysis, as described below. A subset of stored structures was used for hRNase5 (1,400 structures) and hRNase7 (1,360 structures) analysis as the substrates were ejected mid-way during the simulations; therefore, only the structures where the substrate stays bound in the sub-sites were used for analysis. Specifically, only the first 400 structures from hRNase5-AUAU 300 K 2 trajectory and first 360 structures from hRNase7-AUAU 300 K 1 trajectory were used for structural (H-bonding) and interaction energy analyses as described below. For the two systems where the substrate was ejected (hRNase5-AUAU at 300 K and hRNase7-AUAU at 300 K), additional MD simulations were performed to ensure the reproducibility of the behavior. To generate the starting coordinate for this third set of trajectories for these two systems, the equilibrated structures were heated for 1000 ps to 305 K (for simulation set for 300 K) and 315 K (for simulation set for 310 K) to generate conformational diversity. The ending structures were minimized for 100,000 steps and heated to final temperature before production runs were performed similar to the other two sets.

**Structural analysis using H-bonding patterns:** H-bonding interactions between the substrate and RNases were analyzed using *ptraj*, with the following criteria: less than 3.3 Å for donor to acceptor heavy atom distance at an angle between 135° and 180° for the donor-hydrogen-
acceptor. The percentage occupancies of H-bonds, defined as fraction (percentage) of total number of structures (as described in last paragraph) where the interaction meets the H-bonding criterion mentioned above, were also calculated.

**Enzyme-substrate interaction:** The energy for the enzyme-substrate interactions \( E_{enz-subs} \) were calculated as a sum of electrostatic \( E_{el} \) and van der Waals energy \( E_{vdw} \) between atom pairs, based on an approach developed in our group.\textsuperscript{278, 284}

\[
E_{enz-subs} = \sum (E_{el} + E_{vdw}) \tag{3.1}
\]

\( E_{el} \) is the electrostatic contribution, \( E_{vdw} \) is the van der Waals term and the summation runs over all atom pairs for the enzyme and nucleotide substrate. The \( E_{el} \) and \( E_{vdw} \) terms were computed as follows

\[
E_{el} = \frac{q_i q_j}{\varepsilon(r) r_{ij}} \quad \text{and} \quad E_{vdw} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \tag{3.2}
\]

where \( q_i, q_j \) are partial charges, and \( A_{ij}, B_{ij} \) are Lennard-Jones parameters. These parameters were obtained from AMBER \textit{ff14SB} force field. A distance-dependent dielectric function was used:

\[
\varepsilon(r_{ij}) = A + \frac{B}{1 + k \exp(-\lambda Br_{ij})} \tag{3.3}
\]

\( B = \varepsilon_0 A; \varepsilon_0 = 78.4 \) for water; \( A = -8.5525; \lambda = 0.003627 \) and \( k = 7.7839. \)

All enzyme and substrate (ACAC/AUAU) atom pairs were included in the calculations and resulting interaction energies were summed for each residue-nucleotide pair. For each system, the reported interaction energies are an average over total number of structures (as described above). This method has successfully been used to predict the experimentally obtained binding preference/affinities of substrates for a number of proteins including human cyclophilin A, xylanases, and maltose binding proteins similar to the substrate preference of hRNases described in this study.\textsuperscript{278, 284, 285}

**Electrostatic surface calculations:** The enzyme electrostatic surface was computed using the Adaptive Poisson-Boltzmann Solver (APBS) software.\textsuperscript{286} APBS requires a protein conformation.
The ideal case of using averaged enzyme structure for the entire MD simulation was not feasible as the averaged structure shows artificial bonds and angles, particularly to the sampling of different conformations over the 0.5 µs simulations. Therefore, a conformation from the MD ensemble closest to the averaged structure (with least RMSD to the average) was used to generate the electrostatic surface.

**RMSF$^{10}$ calculations:** For dynamical analysis, backbone (C$_\alpha$) and all-atom flexibility of simulation trajectories was determined from the root mean square fluctuation (RMSF), computed by aggregating the magnitude of displacement eigenmodes computed using the quasi-harmonic analysis (QHA) in the *ptraj* analysis module in AMBER. As described previously,228, 287 only the top 10 QHA modes (RMSF$^{10}$) were used in the analysis to focus on the principal dynamics or long time-scale functionally relevant fluctuations in the proteins. For comparison, the results for apo systems were obtained from our previous study.267

### 3.4 Results

#### 3.4.A Stability of the substrate-bound RNase complexes

Table 3.1 summarizes the interactions of hRNases 1-7 in complex with model tetranucleotide substrates ACAC and AUAU. Starting from the enzyme conformation with substrate-bound in the active-site, 0.5 µs MD simulations were used to characterize the stability of interactions between the enzyme and nucleotide substrates. Note, the starting coordinates for the substrate phosphate and ribose sugar backbone were adopted from a bRNaseA-substrate complex published previously (PDB code 1RCN), and the nucleotide bases were added computationally. Two independent MD simulation trajectories for each complex were performed, labeled 1 and 2 in Table 3.1, to ensure reproducibility of the results reported. The characterization of enzyme-substrate complex stability is based on interaction analysis between the two central nucleotides and enzyme residues. Only the two central nucleotides were considered for stability analysis as these nucleotides are adjacent to the phosphodiester bond cleaved by RNases and have well defined interactions with active-site residues forming the P$_1$/B$_1$ and P$_2$/B$_2$ binding sites (see Figure 3.1). The other/terminal two nucleotides in the model substrates do not have well defined interactions and show significantly lower stability over the course of the MD simulations. In Table 3.1, simulation trajectories where the two central nucleotides show stable binding in their respective sub-sites (for entire 0.5 µs) are represented with a green box, while trajectories where one of the two central nucleotides is ejected from the sub-site are shown using a yellow box. The
stable nucleotide in the binding sub-site is identified in the yellow boxes of Table 3.1. The cases where both the central nucleotides lose all contacts with the sub-site residues and are considered to be ejected out of the binding pocket are marked in red boxes. Ejection out of the binding pocket is defined as no direct hydrogen bond or van der Waals contact between the enzyme residues and the two central nucleotides (any water or ion-mediated indirect interactions were excluded from this analysis).

The interaction analysis shows that the majority of RNases show stable nucleotide binding throughout the simulations for the two alternate temperatures at 300 K and 310 K, reproduced in both the alternative MD trajectories. Notable exceptions include hRNase5 and hRNase7, where the substrate AUUA is completely ejected out of the active-site. In the case of hRNase3, the enzyme displays interactions with only one of the two central nucleotides, while the second central nucleotide is ejected out of the binding pocket (Table 3.1). There are marked differences associated with the two temperatures as well as two independent trajectories at each temperature, but overall the results are qualitatively similar. Note that these are human proteins and 310 K is close to the physiological temperature, while a number of experimental enzyme kinetics parameters are reported for these systems at temperature close to 300 K. 54, 110, 112, 114, 116, 117, 254, 255, 264, 265 For hRNase5 and hRNase7, additional data was collected to confirm the observations of the ejection of substrate from the active-site. The third MD set (data only collected for these two systems) indicated that for hRNase7-AUUA at 300 K, the behavior was reproduced, as the substrate was completely ejected from the active-site at 370 ns. For hRNase5-AUUA the behavior was partially reproduced, as 3 out of the 4 nucleotides broke all contacts, with only central U maintaining contact with the protein.

Characterization of the relative substrate positions and interactions with enzymes over the course of MD simulations revealed significant variations among the different RNases. The results are depicted in Figure 3.2 for simulations at 300 K and Figure 3.3 for simulations at 310 K (the trajectory labeled 1 in Table 3.1 is shown). These figures depict three substrate conformations (relative positions); at the start (0 µs, green sticks for carbon), half-way (0.25 µs, yellow sticks for carbon) and endpoint (0.5 µs, magenta sticks for carbon) of the simulations. In cases where the substrate is completely ejected from the active-site, the mid-way and ending frames have been adjusted according to the time of ejection (listed in Table 3.1).
In the cases of bRNaseA, hRNase1, and hRNase6, both model substrates (ACAC and AUAU) stay bound in the active-site for the entire 0.5 µs simulations (300 K as well as 310 K), close to the original orientation, suggesting equal stability of these substrates in these enzymes. Note that the central nucleotides remain close to their initial orientation in the active-site (in P1/B1 and P2/B2 sites), while the two peripheral nucleotides show considerable variation from their initial orientation. In the case of hRNase2, both substrates stay in their initial orientation at 300 K, with the central adenosine base showing minor movement out of the B2 pocket while phosphate stays bound in the P2 pocket. For simulations at 310 K, substrate ACAC is partially displaced from its initial orientation, with the central cytosine in trajectory 1 (adenine in trajectory 2) moving completely out of the P1/B1 (P2/B2) pockets (Figure 3.3). In the case of hRNase4, substrate ACAC at 300 K stays localized in the active-site while at 310 K adenine leaves the binding pocket; for substrate AUAU at both temperature adenosine moves out of its binding pocket. In hRNase5, the substrates show considerable movements in simulations at 300 K, particularly in the case of AUAU, where the substrate completely leaves the active-site in one of the MD trajectories (labeled as trajectory 2 in Table 3.1). For hRNase3 and hRNase7, at least one of the nucleotide binding to B1 or B2 site (for both ACAC and AUAU) show significant movement away from the binding pocket over the course of MD simulations at both temperatures investigated.

3.4.B Structural interactions

To quantitatively characterize the substrate interactions with each of the 8 proteins in this study, hydrogen bonding interactions were calculated for the ensemble of conformations sampled during the MD simulations. The average distance between heavy atoms of the enzyme and substrate nucleotide, and percentage occupancy of the hydrogen bonding interactions over the course of aggregate 1 µs MD sampling (by combining two simulations labeled as 1 and 2 in Table 3.1) were calculated; the results are summarized in Table 3.2 for 300 K and Table 3.3 for 310 K, for the two substrates. The results indicate that the conserved catalytic triad (His12, Lys41 and His119, bRNaseA numbering) does not form stable hydrogen bonds, and any definite patterns in the H-bonding interactions between these residues (across all 8 enzymes investigated) and substrates were not observed. Among the non-catalytic active-site residues, the equivalent of bRNaseA residues Thr45 and Asn71 showed most stable H-bond interactions with the substrates in several human RNases. These residues were previously shown to be part of the B1 and B2 binding sites, respectively. Unsurprisingly, Thr45 and Asn71 are fully conserved in all 8 RNases investigated in this study, possibly illustrating their role in active-site substrate stabilization.
Table 3.1: Stability of nucleotide substrates ACAC and AUAU in the active-site of RNases. Stable substrates in the active-site for entire 0.5 µs (500 ns) are colored as green boxes; cases where one nucleotide is ejected are colored yellow and the stable nucleotide base is labeled (A – adenine, U – uracil, C – cytosine); and cases where both nucleotides are ejected are indicated as red boxes (time in nanoseconds when the substrate is fully ejected is specified). Note that only the two central nucleotides were considered for this analysis (in red, underlined). a = additional (3rd) trajectory indicates that only central U remained in active-site, b = additional (3rd) trajectory indicates that the substrate was ejected from active-site at 370 ns.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>PDB ID</th>
<th>ACAC</th>
<th></th>
<th></th>
<th>AUAU</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>300 K</td>
<td>310 K</td>
<td>300 K</td>
<td>310 K</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>7RSA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hRNase1</td>
<td>2K11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hRNase2</td>
<td>1GQV</td>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hRNase3</td>
<td>1QMT</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>hRNase4</td>
<td>1RNF</td>
<td>C</td>
<td>U</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hRNase5</td>
<td>1ANG</td>
<td></td>
<td>C</td>
<td>A</td>
<td></td>
<td>200 ns</td>
<td>U</td>
</tr>
<tr>
<td>hRNase6</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hRNase7</td>
<td>2HKY</td>
<td>C</td>
<td>C</td>
<td></td>
<td>180 ns</td>
<td>U</td>
<td>U</td>
</tr>
</tbody>
</table>
Figure 3.2: Substrate behavior over the course of 0.5 µs MD simulations. The enzymes are shown as gray cartoons and the 3 orientations of the substrate at start (0 µs, green), mid-way (0.25 µs, yellow) and end (0.5 µs, magenta) of the simulations are shown as sticks. The results for the two substrates ACAC and AUAU are shown separately. The central two nucleotides of the substrate that interact with two binding sub-sites (B1/P1 or B2/P2) are shown in dark colors, while the two terminal nucleotides (one on each side) are depicted with faded color. The results depicted are from the 300 K simulation trajectories and represent the MD simulations set 1. The results from trajectory 2 at 300 K are qualitatively similar (Table 3.1). * indicates simulations where the substrate is ejected and leaves the active-site; the end and mid-way frames for these cases are adjusted accordingly. a indicates that the substrate was stable in depicted trajectory (300 K 1); however, it was ejected out from the pocket in second trajectory (300 K 2).
Figure 3.3: Substrate behavior over the course of 0.5 μs MD simulations at 310 K. The enzymes are shown as gray cartoons and the 3 relative positions of the substrate at start (0 μs, green), mid-way (0.25 μs, yellow) and end (0.5 μs, magenta) of the simulations are shown as sticks. The results for the two substrates ACAC and AUAU are shown separately. The central two nucleotides of the substrate that interact with two binding sub-sites (B1/P1 or B2/P2) are shown in dark colors, while the two terminal nucleotides (one on each side) are depicted with faded color. The results depicted are from the 310 K simulation trajectories and represent the MD simulations set 1. The results from trajectory 2 at 310 K are qualitatively similar (Table 3.1).
For enzymes bRNaseA, hRNase1, hRNase4 and hRNase6, strong H-bonding interactions are observed with both substrates at 300 K (Figure 3.4 and Table 3.2). In this study, strong H-bonds are defined as H-bonds with bond length less than 3 Å between the heavy atoms; and H-bond occupancies greater than 80% is considered significant. hRNase2 forms at least one H-bond interaction with relatively higher occupancy (79.4%) with substrate ACAC while it forms H-bonds with AUAU that have lower occupancies. On the other hand, hRNase3 and hRNase7 form H-bonds with both substrates that have lower occupancies (≤ 70%). It was also observed that hRNase5 shows no H-bonding interactions with the substrate AUAU. This might be due to the complete ejection of substrate from the binding site of hRNase5 after ~200 ns (0.2 µs) of simulation (Table 3.1).

A number of differences were observed when comparing the enzyme-substrate H-bonding interactions at experimental (300 K) and physiological (310 K, see Figure 3.5 and Table 3.3) temperatures with the most obvious difference in H-bonding pattern appears in hRNase3, hRNase5 and hRNase7. Interestingly, hRNases 2 and 3 show lower occupancies with both ligands relative to bRNaseA at both temperatures. For hRNase2, strong H-bonding interactions with substrate ACAC that were observed at 300 K simulations exhibit lower occupancies at 310 K. Simulations indicated that hRNase3 does not form any H-bonds with ACAC at 310 K. A comparison of the interactions at 310 K suggest significantly more interactions for hRNases 3 and 4 with the AUAU ligand. Note when compared to lower temperature, hRNase3 forms H-bonding interactions with AUAU that have lower occupancies. On the other hand, it is observed that hRNase7 does not show any H-bonding interactions with the substrate AUAU at 310 K, although stronger H-bonds are observed with ACAC at 310 K, suggesting a potential preference for cytidine (C) in the B1 sub-site. hRNase5, which did not show any H-bonding interactions with substrate AUAU at 300 K, exhibits H-bonds with substrate AUAU at 310 K through the conserved Asn43 and Thr44 residues. The time evolution behavior of H-bonds for various complexes is shown in Figures 3.12 – 3.19 (Appendix 3.7).

Overall, results suggest that even though all the 8 enzymes have conserved catalytic triad and several other conserved residues in the active-site, the model substrates show significant differences in their interactions with the enzymes, with some of them showing weak interactions and/or lower occupancies. An interesting observation is that even for substrates which stay bound to the enzymes, the conserved catalytic triad residues do not exhibit strong H-bonding interactions, while other residues show stronger interactions, including residues reported to be
Figure 3.4: Hydrogen-bonding interactions between RNases and model substrates at 300 K. H-bonds (with >45% occupancy) are indicated with black dotted line and averaged bond length (Å) and percentage occupancy are shown in red colored text. Enzyme and nucleotide atoms participating in the H-bonds are labeled, with enzyme residues shown in blue and substrate nucleotides shown in green for carbon atoms, blue for nitrogen, red for oxygen, and orange for phosphorus. Enzyme residue numbers correspond to enzyme sequence, while substrate nucleotides are numbered as A_{i-2}C/U_{i-1}A_{i}C/U_{i+2}, as described in the legend of Figure 3.1. hRNase5-AUAU did not show any H-bonds meeting the criteria and is only shown for comparison. Note that the substrates orientation (marked by 3’ and 5’ for bRNaseA) appear opposite to the depiction in Figure 3.1, as enzymes are shown from view used commonly in literature. The percentage occupancies have been rounded off to the nearest whole number.
Figure 3.5: Hydrogen-bonding interactions between RNases and model substrates at 310 K. H-bonds (with >45% occupancy) are indicated with black dotted line and averaged bond length (Å) and percentage occupancy are shown in red colored text. Enzyme and nucleotide atoms participating in the H-bonds are labeled, with enzyme residues shown in blue and substrate nucleotides shown in green for carbon atoms, blue for nitrogen, red for oxygen, and orange for phosphorus. Enzyme residue numbers correspond to enzyme sequence, while substrate nucleotides are numbered as $A_{i-2}C/U_{i-1}A_{i}C/U_{i+2}$, as described in the legend of Figure 3.1. Note that the substrates orientation (marked by 3’ and 5’ for bRNaseA) appear opposite to the depiction in Figure 3.1, as enzymes are shown from view used commonly in literature. The percentage occupancies are rounded off to the nearest whole number.
Table 3.2: H-bonding properties of bovine and human pancreatic-type RNases at 300 K.
The H-bonds that meet the criteria of 3.3 Å, angle between 135 - 180° and more than 45% occupancy are listed. Results are based on analysis of both the alternate MD trajectories for each system (see methods for details).

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Residue</th>
<th>bond length (% occupancy)</th>
<th>Substrate</th>
<th>Residue</th>
<th>bond length (% occupancy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hRNaseA</td>
<td>C(1) (N3) Thr45 (Oγ1)</td>
<td>2.87 (95.7%)</td>
<td>U(1) (O2) Thr45 (Oγ1)</td>
<td>2.88 (95.4%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C(1) (O2) Thr45 (N)</td>
<td>2.87 (88.5%)</td>
<td>U(1) (N3) Thr45 (Oγ1)</td>
<td>2.97 (85.8%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (N1) Asn71 (Nε2)</td>
<td>3.01 (65.0%)</td>
<td>A(1) (O2) Gln11 (Nε2)</td>
<td>2.84 (73.2%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C(1) (O2) Phe120 (O)</td>
<td>2.74 (47.2%)</td>
<td>A(1) (N6) Asn71 (Oε2)</td>
<td>2.95 (69.2%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (Ne6) Asn71 (Oε1)</td>
<td>3.00 (46.4%)</td>
<td>A(1) (N1) Asn71 (Nε2)</td>
<td>3.02 (68.8%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C(1) (N3) Thr45 (Oγ1)</td>
<td>2.89 (93.6%)</td>
<td>C(1) (O2) Thr45 (N)</td>
<td>2.89 (72.2%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C(1) (O2) Phe120 (O)</td>
<td>2.75 (49.2%)</td>
<td>C(1) (O2) Phe120 (O)</td>
<td>2.73 (87.0%)</td>
<td></td>
</tr>
<tr>
<td>hRNase1</td>
<td>C(1) (N3) Thr43 (Oγ1)</td>
<td>2.88 (79.4%)</td>
<td>U(1) (N3) Thr45 (Oγ1)</td>
<td>2.91 (98.4%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (O2) Gln15 (Nε2)</td>
<td>2.87 (65.6%)</td>
<td>A(1) (N1) Asn71 (Nε2)</td>
<td>2.98 (94.2%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C(1) (O2) Phe120 (O)</td>
<td>2.75 (49.2%)</td>
<td>A(1) (O1) Asn71 (Oε1)</td>
<td>2.98 (85.6%)</td>
<td></td>
</tr>
<tr>
<td>hRNase2</td>
<td>C(1) (N3) Thr43 (Oγ1)</td>
<td>2.88 (79.4%)</td>
<td>U(1) (N3) Thr43 (Oγ1)</td>
<td>2.95 (66.1%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (O2) Gln15 (Nε2)</td>
<td>2.87 (65.6%)</td>
<td>A(1) (O2) Gln15 (Nε2)</td>
<td>2.88 (55.8%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C(1) (O2) Phe120 (O)</td>
<td>2.75 (49.2%)</td>
<td>U(1) (O4) Ile134 (N)</td>
<td>2.98 (47.7%)</td>
<td></td>
</tr>
<tr>
<td>hRNase3</td>
<td>A(1) (N6) Asn71 (Oε1)</td>
<td>2.95 (51.6%)</td>
<td>A(1) (N6) Asn71 (Oε1)</td>
<td>2.96 (67.4%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (N1) Asn71 (Nε2)</td>
<td>3.01 (50.8%)</td>
<td>A(1) (N6) Asn71 (Nε2)</td>
<td>3.01 (67.0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (O2) Phe118 (N)</td>
<td>2.91 (92.9%)</td>
<td>A(1) (O2) Phe118 (N)</td>
<td>2.89 (60.0%)</td>
<td></td>
</tr>
<tr>
<td>hRNase4</td>
<td>A(1) (O1) Phe118 (N)</td>
<td>2.91 (92.9%)</td>
<td>U(1) (N3) Thr45 (Oγ1)</td>
<td>2.95 (93.8%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (O1) Phe118 (N)</td>
<td>2.80 (66.3%)</td>
<td>U(1) (O2) Thr45 (N)</td>
<td>2.93 (86.2%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C(1) (N4) Thr45 (Oγ1)</td>
<td>2.91 (63.1%)</td>
<td>A(1) (O1) Phe118 (N)</td>
<td>2.89 (59.0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (N1) Asn71 (Nε2)</td>
<td>3.00 (61.0%)</td>
<td>A(1) (O1) Phe118 (N)</td>
<td>2.89 (59.0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (N6) Asn71 (Oε1)</td>
<td>2.94 (59.9%)</td>
<td>A(1) (O1) Phe118 (N)</td>
<td>2.89 (59.0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (O1) Phe118 (N)</td>
<td>2.96 (54.2%)</td>
<td>A(1) (O1) Phe118 (N)</td>
<td>2.89 (59.0%)</td>
<td></td>
</tr>
<tr>
<td>hRNase5</td>
<td>C(1) (N4) Thr44 (Oγ1)</td>
<td>2.95 (63.0%)</td>
<td>U(1) (O2) Thr45 (N)</td>
<td>2.93 (86.2%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (O1) His114 (Nε1)</td>
<td>2.77 (58.6%)</td>
<td>A(1) (N6) Asn69 (Oε2)</td>
<td>2.93 (66.2%)</td>
<td></td>
</tr>
<tr>
<td>hRNase6</td>
<td>A(1) (O1) His114 (Nε1)</td>
<td>2.77 (58.6%)</td>
<td>A(1) (N6) Asn69 (Oε2)</td>
<td>2.93 (66.2%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (N1) Asn69 (Nε2)</td>
<td>3.02 (84.8%)</td>
<td>A(1) (N1) Asn69 (Nε2)</td>
<td>3.05 (88.0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (N6) Asn69 (Oε1)</td>
<td>2.98 (84.4%)</td>
<td>A(1) (N1) Asn69 (Nε2)</td>
<td>3.05 (88.0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (O1) Trp11 (Nε1)</td>
<td>2.91 (79.7%)</td>
<td>A(1) (N6) Asn69 (Oε1)</td>
<td>2.94 (70.7%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(1) (N6) Asn65 (Oε1)</td>
<td>2.98 (50.6%)</td>
<td>A(1) (O2) His123 (Nε2)</td>
<td>2.94 (52.0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C(1) (O2) Leu124 (N)</td>
<td>3.00 (46.9%)</td>
<td>C(1) (N3) Thr43 (Oγ1)</td>
<td>2.85 (53.2%)</td>
<td></td>
</tr>
<tr>
<td>hRNase7</td>
<td>C(1) (N3) Thr43 (Oγ1)</td>
<td>2.85 (53.2%)</td>
<td>U(1) (O2) Leu125 (O)</td>
<td>2.67 (70.51%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5' A(2) (N6) Thr43 (Oγ1)</td>
<td>2.95 (48.01%)</td>
<td>5' A(2) (N6) Thr43 (Oγ1)</td>
<td>2.95 (48.01%)</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.3: H-bonding properties of bovine and human pancreatic-type RNases at 310 K.
The H-bonds that meet the criteria of 3.3 Å, angle between 135 - 180° and more than 45% occupancy are listed. Note that the results are based on analysis of both MD alternate trajectories for each system (see methods for details).

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Residue</th>
<th>bond length (% occupancy)</th>
<th>Substrate</th>
<th>Residue</th>
<th>bond length (% occupancy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bRNaseA</td>
<td>C(-1) (N3)</td>
<td>Thr45 (Oγ1)</td>
<td>2.88 (91.4%)</td>
<td>A(1) (N1)</td>
<td>Asn71 (Nδ2)</td>
</tr>
<tr>
<td></td>
<td>C(-1) (O2)</td>
<td>Thr45 (N)</td>
<td>2.88 (86.4%)</td>
<td>U(-1) (O2)</td>
<td>Thr45 (N)</td>
</tr>
<tr>
<td></td>
<td>A(1) (N1)</td>
<td>Asn71 (Nδ2)</td>
<td>3.00 (58.7%)</td>
<td>A(1) (N6)</td>
<td>Asn71 (Oδ1)</td>
</tr>
<tr>
<td></td>
<td>C(-1) (O2)</td>
<td>Phe120 (O)</td>
<td>2.74 (48.7%)</td>
<td>U(-1) (N3)</td>
<td>Thr45 (Oγ1)</td>
</tr>
<tr>
<td></td>
<td>A(1) (N6)</td>
<td>Asn71 (Oδ1)</td>
<td>2.99 (46.3%)</td>
<td>A(1) (N1)</td>
<td>Thr45 (Oγ1)</td>
</tr>
<tr>
<td>hRNase1</td>
<td>C(-1) (N3)</td>
<td>Thr45 (Oγ1)</td>
<td>2.87 (93.6%)</td>
<td>U(-1) (N3)</td>
<td>Thr45 (Oγ1)</td>
</tr>
<tr>
<td></td>
<td>C(-1) (O2)</td>
<td>Thr45 (N)</td>
<td>2.88 (84.4%)</td>
<td>A(1) (N1)</td>
<td>Asn71 (Nδ2)</td>
</tr>
<tr>
<td></td>
<td>A(1) (N1)</td>
<td>Asn71 (Nδ2)</td>
<td>3.02 (70.1%)</td>
<td>U(-1) (O2)</td>
<td>Thr45 (N)</td>
</tr>
<tr>
<td></td>
<td>C(-1) (O2)</td>
<td>Phe120 (O)</td>
<td>2.74 (45.4%)</td>
<td>A(1) (N6)</td>
<td>Asn71 (Oδ1)</td>
</tr>
<tr>
<td>hRNase2</td>
<td>A(1) (Oδ2)</td>
<td>Gln15 (Nε2)</td>
<td>2.92 (56.4%)</td>
<td>U(-1) (O2)</td>
<td>Thr43 (Oγ1)</td>
</tr>
<tr>
<td></td>
<td>C(-1) (N3)</td>
<td>Thr43 (OY1)</td>
<td>2.84 (49.6%)</td>
<td>A(1) (Oδ2)</td>
<td>Gln15 (Nε2)</td>
</tr>
<tr>
<td>hRNase3</td>
<td></td>
<td></td>
<td></td>
<td>A(1) (N1)</td>
<td>Asn71 (Nδ2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A(1) (O)</td>
<td>His129 (Nε1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A(1) (Oδ2)</td>
<td>Gln15 (Nε2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A(1) (N6)</td>
<td>Asn71 (Oδ1)</td>
</tr>
<tr>
<td>hRNase4</td>
<td>C(-1) (N4)</td>
<td>Thr45 (Oγ1)</td>
<td>2.90 (86.0%)</td>
<td>U(-1) (N3)</td>
<td>Thr45 (Oγ1)</td>
</tr>
<tr>
<td></td>
<td>A(1) (O)</td>
<td>Phe118 (N)</td>
<td>2.92 (58.0%)</td>
<td>U(-1) (O2)</td>
<td>Thr45 (N)</td>
</tr>
<tr>
<td></td>
<td>A(1) (Oδ1)</td>
<td>Asn71 (Nδ2)</td>
<td>2.78 (49.4%)</td>
<td>A(1) (N6)</td>
<td>Asn71 (Oδ1)</td>
</tr>
<tr>
<td></td>
<td>A(1) (N1)</td>
<td>Asn71 (Oδ1)</td>
<td>2.95 (48.2%)</td>
<td>A(1) (O)</td>
<td>His117 (Nε1)</td>
</tr>
<tr>
<td>hRNase5</td>
<td>C(-1) (N3)</td>
<td>Thr44 (Oγ1)</td>
<td>2.85 (76.3%)</td>
<td>U(-1) (O4)</td>
<td>Thr44 (N)</td>
</tr>
<tr>
<td></td>
<td>C(-1) (O2)</td>
<td>Thr44 (N)</td>
<td>2.90 (70.7%)</td>
<td>U(-1) (N3)</td>
<td>Asn43 (Oδ1)</td>
</tr>
<tr>
<td>hRNase6</td>
<td>A(1) (N6)</td>
<td>Asn69 (Oδ1)</td>
<td>2.97 (93.6%)</td>
<td>A(1) (N6)</td>
<td>Asn69 (Oδ1)</td>
</tr>
<tr>
<td></td>
<td>A(1) (N1)</td>
<td>Asn69 (Nε2)</td>
<td>3.03 (90.4%)</td>
<td>A(1) (N6)</td>
<td>Asn69 (Nε2)</td>
</tr>
<tr>
<td></td>
<td>A(1) (O)</td>
<td>Trp11 (Nε2)</td>
<td>2.87 (87.2%)</td>
<td>A(1) (N6)</td>
<td>Asn65 (Oδ1)</td>
</tr>
<tr>
<td></td>
<td>A(1) (O)</td>
<td>Gln15 (Nε2)</td>
<td>2.97 (70.6%)</td>
<td>A(1) (O)</td>
<td>Gln15 (Oε1)</td>
</tr>
<tr>
<td></td>
<td>C(-1) (N4)</td>
<td>Thr43 (Oγ1)</td>
<td>2.95 (48.2%)</td>
<td>A(1) (O)</td>
<td>Gln15 (Oε1)</td>
</tr>
<tr>
<td></td>
<td>A(1) (N6)</td>
<td>Asn65 (Oδ1)</td>
<td>2.97 (47.2%)</td>
<td>A(1) (O)</td>
<td>Gln15 (Oε1)</td>
</tr>
<tr>
<td>hRNase7</td>
<td>C(-1) (N3)</td>
<td>Thr43 (Oγ1)</td>
<td>2.83 (81.3%)</td>
<td>C(-1) (O2)</td>
<td>Thr43 (N)</td>
</tr>
<tr>
<td></td>
<td>C(-1) (O2)</td>
<td>Leu125 (O)</td>
<td>2.71 (60.2%)</td>
<td>C(-1) (O2)</td>
<td>Leu125 (O)</td>
</tr>
<tr>
<td></td>
<td>C(-1) (N4)</td>
<td>Thr43 (Oγ1)</td>
<td>3.04 (50.6%)</td>
<td>C(-1) (N4)</td>
<td>Thr43 (Oγ1)</td>
</tr>
</tbody>
</table>
involved in substrate recognition and stabilization. This suggests that while other conserved residues may hold the substrate in place, the catalytic triad shows a large degree of flexibility which is possibly required for the catalytical cleavage step. This behavior is different than other enzyme systems where it is shown that the catalytic active-site residues are fairly rigid and make strong interactions with the substrates. Note that the structural analysis based on H-bonds does not provide information about complete set of interactions between substrate and hRNases. It helps in identification of the most important direct hydrophilic interactions; however, long range electrostatic and hydrophobic interactions are not captured by H-bonding analysis. In RNases particularly the strongly charged phosphate backbone and the hydrophobic interactions of the nucleotide base in the binding pocket provide important contributions. Therefore, a detailed and quantitative analysis based on full electrostatic and van der Waals interaction energy was also performed as described below.

3.4.C Electrostatic interactions

ssRNA substrates are highly negatively charged molecules, particularly due to the presence of a phosphate backbone. It has been documented that long-range electrostatic interactions play an important role in the binding of negatively-charged RNA substrates to the highly cationic surface of this RNase superfamily. For characterizing and comparing the charge distribution on the surface of these homologs, the averaged electrostatic surface for each enzyme over the course of MD trajectories was calculated (Figure 3.6 for results obtained from simulations at 300 K and Figure 3.7 for results from simulations at 310 K). hRNases have been reported to bind to similar substrates and share the same overall shape, but the electrostatic surface shows variations due to the difference in protein sequence. In addition, the charge distribution also showed differences based on the type of bound substrate, possibly indicating changes in orientation of enzyme residues on the surface due to the nature of the substrate. The most prominent feature is that all RNases in this study have highly positively charged surfaces (blue regions in Figures 3.6 & 3.7). This is not surprising as the RNA substrates are negatively charged, and the positively charged surface is expected to guide the substrate into the active-site pocket. An interesting observation is that the electrostatic surfaces for human enzymes are similar when bound to the two substrates ACAC and AUUA, but different in the case of bRNaseA. Note that bRNaseA is catalytically the most efficient enzyme (for ribonucleolytic cleavage activity) of 8 enzymes investigated in this study. Further studies could investigate the correlation between these enzymes to adapt
conformations for optimal electrostatic based on bound substrate and efficiency of ribonucleolytic cleavage.

3.4.D Quantitative characterization of interaction energies

Figure 3.8 depicts an overview of the interaction energy between the enzyme and the model substrates, based on the 300 K simulations. The results for 310 K are shown in Figure 3.9. The total interaction energy was computed as a sum of van der Waals and electrostatic interaction energies between each enzyme residue and each nucleotide of the substrate molecules. Tables 3.5 and 3.6 (Appendix 3.7) provide the quantitative values, using a cut-off value of < -3 kcal/mol per residue-nucleotide pair being considered as significant. Overall, these results from this analysis indicate that both model substrates show favorable interactions (blue to yellow regions in the figure panels) with the enzyme. Note that H-bonding based structural analysis presented above, describes only a sub-set of key interactions but interaction energy analysis provides a complete picture. More specifically, favorable interactions with the catalytic His119 (bRNaseA numbering) were observed for both the substrates with all RNases. Most RNases also showed favorable interactions with Thr45 (bRNaseA numbering) in the β1 strand, a residue shown to bind to the pyrimidine nucleotide. bRNaseA residue Lys7 further showed favorable interactions with both the substrates and Lys66 (from loop L4) with substrate ACAC. Similar interactions were observed for hRNase1, which showed additional favorable interactions by Asn71 and Lys7 with AUAU. In case of hRNase2 residues Arg133 and Trp8 show additional interactions in case of substrate AUAU; the latter residue is previously reported to interact with the ribose group. In contrast to hRNase2 where a significant number of interactions are present in the region 39-43, hRNase3 showed significantly fewer interactions with the two substrates. The interaction energy profile of hRNase4 was similar to that observed for bRNaseA for the two substrates. The C-terminal region of hRNase5 displayed favorable interactions with the nucleotide substrates in addition to the catalytic Lys40, however, significantly fewer interactions of hRNase5 were observed with AUAU. In contrast to other RNases, hRNase6 showed favorable interactions with several residues of loop L4, suggesting a potential effect on the dynamical properties of this loop upon substrate binding. hRNase7 showed very few favorable interactions localized to the catalytic His124 with substrate AUAU while substrate ACAC showed additional interactions with the Lys39 and residues of β1 strand. Similar results were observed for all RNases at 310 K except hRNase5 and hRNase7, which showed more favorable interactions with substrate AUAU. See text in supporting information for more details of substrate interactions with each RNase.
Figure 3.6: Electrostatic surface of RNases with the two model substrates at 300 K. Enzyme conformation from the MD ensemble closest to the averaged structure for the entire trajectory (conformation with smallest RMSD with averaged structure) was used to calculate the representative electrostatic potential (+5kT/e in blue and –5kT/e per electron in red).
Figure 3.7: Electrostatic surface of RNases with the two model substrates at 310 K. Enzyme conformation from the MD ensemble closest to the averaged structure for the entire trajectory (conformation with smallest RMSD with averaged structure) was used to calculate the representative electrostatic potential (+5kT/e in blue and –5kT/e per electron in red).
A comparison of the total van der Waals ($E_{\text{vdw}}$) and electrostatic ($E_{\text{el}}$) energies summed up for the enzyme-substrate complexes provide a number of interesting observations (Table 3.4). Even though the nucleotide substrates are highly charged molecules (particularly the backbone phosphate), the largest contribution to the favorable interactions comes from van der Waals interactions, indicating that the active-site residues have preferential interactions with the bases. These quantitative estimates provide additional insights, indicating that bRNaseA possibly has a slight preference for substrate AUAU relative to ACAC, as $E_{\text{total}}$ is lower by 3-5 kcal/mol. In the case of human enzymes, hRNase1 has a better interaction with ACAC at higher temperature ($E_{\text{total}}$ is lower by ~9 kcal/mol); hRNase4, RNase5, RNase6 and RNase7 show a preference for ACAC substrate at 300 K, with $E_{\text{total}}$ lower more than 10 kcal/mol; hRNase3 prefers AUAU at 310 K ($E_{\text{total}}$ is lower by almost 9 kcal/mol); while hRNase2 has similar interactions with both substrates. An interesting observation is that hRNase5 and hRNase7, particularly for substrate AUAU, show very weak interactions compared to all the other enzymes in this study. The reason appears to be weak (and even repulsive for hRNase7) electrostatic interactions seen for the AUAU model substrate at 300 K. These quantitative estimates are consistent with the observations that the substrate is ejected midway during the simulations (note the energies were calculated, and averaged, only for conformations where substrates were bound to the enzymes). In the cases of hRNase5 and hRNase7, this observation possibly indicates that active-site residues other than the conserved catalytic triad may not be optimal for the enzyme-substrate interactions (see Figure 3.8 and 3.9, and discussion above). However, as indicated by the H-bond analysis with bound substrate, the non-catalytic residues appear to make modest interactions in these two enzyme systems, a different behavior from other enzymes investigated in this study. Therefore, the binding preference and stable interactions in these two systems may be related to non-catalytic residues. This behavior is very different than other enzyme systems such as human cyclophilin A, *E. coli* dihydrofolate reductase and horse liver alcohol dehydrogenase where the catalytic residues make the strongest interactions with the substrate.291-293
Figure 3.8: Enzyme-substrate interaction energy at 300 K. Total interaction energy was computed as a summation of electrostatic and van der Waals interaction energy between atom pairs of enzyme residue and substrate nucleotide (ACAC and AUUA, respectively). The areas of largest negative energy (blue) represent favorable interactions, while positive energy (red) represents less favorable or unfavorable interactions. For ease of comparison, sequence of the model substrate has been adjusted to match the orientation depicted in Figure 3.2 and Figure 3.4 (CACA and UAUA). Enzyme residues showing total interaction energy ($E_{\text{total}}$) contributions < -3 kcal/mol were considered significant and are marked with ellipses. (See Table 3.5 in Appendix 3.7 for individual $E_{\text{total}}$ of selected residues)
Figure 3.9: Enzyme-substrate interaction energy at 310 K Total interaction energy was computed as a summation of electrostatic and van der Waals interaction energy between atom pairs of enzyme residue and substrate nucleotide (ACAC and AUAU, respectively). The areas of largest negative energy (blue) represent favorable interactions, while positive energy (red) represents less favorable or unfavorable interactions. For ease of comparison, sequence of the model substrate has been adjusted to match the orientation depicted in Figure 3.3 and Figure 3.5 (CACA and UAUA). Enzyme residues showing total interaction energy ($E_{\text{total}}$) contributions < -3 kcal/mol were considered significant and are marked with ellipses. (See Table 3.6 in Appendix 3.8 for individual $E_{\text{total}}$ of selected residues)
Table 3.4: Averaged total interaction energy between nucleotide substrates and RNases.

The total interaction energy ($E_{\text{total}}$) was computed as a sum of van der Waals ($E_{\text{vdw}}$) and electrostatics ($E_{\text{el}}$) energy, averaged over 2,000 snapshots (see Methods for details). Difference $>5$ kcal/mol in $E_{\text{total}}$ is considered significant for preference of substrate ACAC and AUAU binding to the same RNase; such cases are underlined. Note that the case of bRNaseA-ACAC at 300 K is close to this criterion (5.1 kcal/mol) but is not marked. The repulsive $E_{\text{vdw}}$ observed in case of hRNase7-AUAU at 310 K is marked in bold.

<table>
<thead>
<tr>
<th></th>
<th>ACAC</th>
<th></th>
<th>AUAU</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>300 K</td>
<td>310 K</td>
<td>300 K</td>
<td>310 K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$E_{\text{vdw}}$</td>
<td>$E_{\text{el}}$</td>
<td>$E_{\text{total}}$</td>
<td>$E_{\text{vdw}}$</td>
<td>$E_{\text{el}}$</td>
<td>$E_{\text{total}}$</td>
<td>$E_{\text{vdw}}$</td>
<td>$E_{\text{el}}$</td>
</tr>
<tr>
<td>bRNaseA</td>
<td>-62.7</td>
<td>-34.4</td>
<td>-97.1</td>
<td>-59.1</td>
<td>-30.0</td>
<td>-89.1</td>
<td>-66.6</td>
<td>-35.5</td>
</tr>
<tr>
<td>hRNase1</td>
<td>-64.8</td>
<td>-36.2</td>
<td>-101.0</td>
<td>-69.0</td>
<td>-39.5</td>
<td>-108.5</td>
<td>-65.5</td>
<td>-35.9</td>
</tr>
<tr>
<td>hRNase2</td>
<td>-64.1</td>
<td>-44.0</td>
<td>-108.1</td>
<td>-59.4</td>
<td>-45.6</td>
<td>-105.0</td>
<td>-69.0</td>
<td>-36.3</td>
</tr>
<tr>
<td>hRNase3</td>
<td>-60.5</td>
<td>-51.9</td>
<td>-112.3</td>
<td>-51.9</td>
<td>-50.3</td>
<td>-102.2</td>
<td>-68.9</td>
<td>-43.1</td>
</tr>
<tr>
<td>hRNase4</td>
<td>-70.4</td>
<td>-45.6</td>
<td>-116.0</td>
<td>-66.9</td>
<td>-41.1</td>
<td>-108.1</td>
<td>-68.5</td>
<td>-34.4</td>
</tr>
<tr>
<td>hRNase5</td>
<td>-63.7</td>
<td>-36.6</td>
<td>-100.3</td>
<td>-55.9</td>
<td>-32.8</td>
<td>-88.7</td>
<td>-39.9</td>
<td>-8.4</td>
</tr>
<tr>
<td>hRNase6</td>
<td>-77.0</td>
<td>-39.8</td>
<td>-116.8</td>
<td>-79.4</td>
<td>-41.5</td>
<td>-120.9</td>
<td>-66.2</td>
<td>-38.1</td>
</tr>
<tr>
<td>hRNase7</td>
<td>-59.3</td>
<td>-41.9</td>
<td>-101.2</td>
<td>-66.7</td>
<td>-47.8</td>
<td>-114.5</td>
<td>-31.2</td>
<td>11.3</td>
</tr>
</tbody>
</table>

Note that the case of bRNaseA-ACAC at 300 K is close to this criterion (5.1 kcal/mol) but is not marked. The repulsive $E_{\text{vdw}}$ observed in case of hRNase7-AUAU at 310 K is marked in bold.
3.4.E Dynamical characterization of the enzyme-substrate complexes

The conformational flexibility (or intrinsic dynamics) of the RNases in presence of the two model substrates was quantified using quasi-harmonic analysis (QHA) of the snapshots sampled during the MD simulations. The aggregated root means square fluctuation of top 10 slowest QHA modes (RMSF\textsubscript{10}) provides a good measure of a protein’s conformational flexibility. Such an analysis has been successfully used for relating dynamics to function for a number of other proteins.\textsuperscript{284, 294} Figure 3.10 depicts the RMSF\textsubscript{10} for the 8 enzymes when bound to the two model substrates, as well as enzymes without substrate (apo) at 300 K. The corresponding results from 310 K simulations are shown in Figure 3.11. Results indicate that the most significant differences occur in the loop regions, similar to apo enzyme systems, as recently reported.\textsuperscript{267} For RNases, and other enzymes, it is widely reported that motions of surface loops play important roles in function through various steps in the catalytic cycle.\textsuperscript{138, 210}

The dynamical characterization at 300 K indicates distinct differences in RNases flexibility upon ligand binding. The intrinsic dynamics of bRNaseA is very similar for the substrate-bound and the apo forms with the largest fluctuations observed for residues of loop L1 (residues 13-21). Note that bRNaseA has the highest catalytic activity among these 8 enzymes, and the long-range dynamics of loop L1 was previously shown to play a functional role in enzyme activity.\textsuperscript{138, 252} hRNase1 showed overall reduction in fluctuations upon binding of the two substrates with additional reduction in fluctuations of loop L1 (residues 13-21) observed in the AUAU-bound state. The truncated loop L1 in hRNase2, hRNase3, hRNase6 and hRNase7 displayed diminished dynamics in these enzymes relative to bRNaseA. The large fluctuations in L4 (residues 58-71 for both enzymes) observed in the apo forms of hRNase2 and hRNase3 were diminished upon binding of the two substrates. Further, hRNase3 showed reduced conformational motions of L2 (residues 35-40) and L4 upon binding of the two substrates while loop 6 (residues 85-94) displayed enhanced motions upon binding of AUAU. hRNase4 showed significant reduction in the dynamical profile throughout the protein upon binding of the two substrates, consistent with similar changes in the conformational exchange patterns observed experimentally (data not shown). hRNase5 showed modest changes near L6 (residue 85-94) upon substrate binding. In the case of hRNase6, large fluctuations observed in L4 (residues 58-68) in the apo state were diminished upon binding of the two substrates. Further, the dynamics observed in loop L7 (residues 109-119) in the apo and AUAU-bound states was diminished in the ACAC-bound state. hRNase7 showed larger fluctuations upon binding of the two substrates in L7 (residues 109-119) while the reduced dynamics upon substrates binding was observed in L4 (residues 58-68).
comparison to fluctuations observed at 300 K, the dynamical profiles at 310 K (Figure 3.11) displayed minor differences in the fluctuations of loops of hRNase2, hRNase3 and hRNase4 for the two substrates. However, in the case of hRNase7, about a 2-fold increase in fluctuations are observed in loop L4 when bound to substrate ACAC, while the fluctuations in loop L7 are decreased for substrate AUAU at the higher temperature.

In a joint NMR-computational study of the RNase superfamily, we recently reported that the intrinsic dynamics of apo proteins is similar for members within phylogenetic sub-families that share common biological activities and different between sub-families. Comparison of intrinsic dynamics when the two different substrates are bound also provides some unique insights. Comparison of members within phylogenetic sub-families showed different dynamical properties upon binding of the two substrates. For example, AUAU binding led to reduced flexibility of L1 in hRNase1 relative to the apo state, while this change was not observed in bRNase1 at both 300 K and 310 K. Similarly, a comparison of hRNase2 and hRNase3, which belong to the same phylogenetic sub-family, suggests that while regions displaying large fluctuations are similar, there are notable differences in the changes in the dynamical profiles upon substrate binding between the two enzymes. These observations are consistent with recent studies which showed distinctly different effects of ligand binding for members within phylogenetic sub-families. However, further investigations are required, as the current work cannot fully rule out whether the observed changes in loop motions help in differentiating binding of substrates or are simply a result of substrate binding.
Figure 3.10: Dynamical behavior of RNases bound to model substrates at 300 K. In the left panel, for each enzyme-substrate [ACAC (red), AUAA (blue), and apo (black)], the flexibility of C\textsubscript{\alpha} atoms of each residue calculated from the slowest 10 quasi-harmonic modes (RMSF\textsubscript{10}) is shown as a function of enzyme residue numbering. Secondary structures [alpha-helix (α), beta sheets (β) and loops (L)] are identified (top); the important loop regions are highlighted in gray. On the right side, RMSF\textsubscript{10} are projected on cartoon representations of each enzyme, with flexible loop regions identified. The thickness of the cartoon tube corresponds to the amplitude of RMSF\textsubscript{10}, and the color corresponds to the dynamic range observed: least dynamic (0 Å, blue) to most dynamic (5 Å, red). These fluctuations are on the same scale, allowing direct comparison between different enzymes and substrates. For clarity, fluctuations in the terminal regions are removed (bRNase 1-2; hRNase1 1-2 and 126-127; hRNase2 134-135; hRNase3 1-3; hRNase4 1-3; hRNase5 1-3 and 121-123; hRNase6 1-4; hRNase7 1-4) from the cartoon representation.
Figure 3.11: Dynamical behavior of RNases bound to model substrates at 310 K. In the left panel, for each enzyme-substrate [ACAC (red), and AUAU (blue), the flexibility of $C_\alpha$ atoms of each residue calculated from the slowest 10 quasi-harmonic modes (RMSF$_{10}$) is shown as a function of enzyme residue numbering. Secondary structures [alpha-helix ($\alpha$), beta sheets ($\beta$) and loops (L)] are identified (top); the important loop regions are highlighted in gray. On the right side, RMSF$_{10}$ are projected on cartoon representations of each enzyme, with flexible loop regions identified. The thickness of the cartoon tube corresponds to the amplitude of RMSF$_{10}$, and the color corresponds to the dynamic range observed: least dynamic (0 Å, blue) to most dynamic (5 Å, red). These fluctuations are on the same scale, allowing direct comparison between different enzymes and substrates. For clarity, fluctuations in the terminal regions are removed (bRNase 1-2; hRNase1 1-2 and 125-127; hRNase4 1-3 and 119-120; hRNase5 1-2 and 121-123; hRNase6 1-3; hRNase7 1-3) from the cartoon representation.
3.5 Discussion

Ribonucleases are a superfamily of enzymes that catalyze the endonucleolytic cleavage of RNA substrates. The eight canonical human pancreatic-type RNases, identified in the initial sequencing of the human genome, perform distinct biological functions such as anti-pathogenicity, angiogenesis and host defense, among others. In addition, they also share the common chemical function of ribonucleolytic activity of RNA substrates of non-specific nucleotide sequences. The human RNases investigated so far also display a diverse range of substrate specificities and catalytic efficiencies. While numerous studies have characterized the nucleotide substrate binding properties, rate kinetics and efficiencies of human RNases, these studies were performed by different research groups with diverse substrates under various experimental conditions. Other studies also attempted to characterize the nucleotide preference based on available crystal structures. Unfortunately, these comparisons also suffer the same limitations due to different substrates and varying conditions. These differences prevent a direct comparison of the nucleotide specificities across hRNases, which in turn limits our current understanding of the underlying mechanism for the observed nucleotide specificities as well as the detailed mechanism of biological and chemical function of these enzymes.

Characterizing the nucleotide binding properties of RNases with the same substrate(s) under identical conditions would provide a benchmark for comparison of members of this superfamily. This in turn would provide a framework for the systematic identification of structural and energetic contributions in the enzyme function. In this study, computer simulations and detailed analysis were used for characterization of human RNases 1-7 and bovine RNase A to probe the nucleotide binding properties in the presence of two model substrates – ACAC and AUAU. Differences in structural stability and hydrogen bonding interactions were characterized (Figure 3.2-3.5, and Tables 3.2, 3.3). Further, the interaction energies between enzyme residues and substrate nucleotide pairs were calculated (as sum of van der Waals and electrostatics energy, see Figure 3.8, 3.9 and Table 3.4). The results indicate that both AUAU and ACAC substrates stay either in or close to the binding pocket for most enzymes on the microsecond time-scale (Table 3.1). hRNase5 and hRNase7, however, show the weakest interactions with substrate AUAU. In one of the two alternate trajectories, the substrate was ejected from the binding pockets of these two enzymes, consistent with weaker binding affinities. Note that we consider the set of two alternative trajectories as a representation of the expected behavior. A large set of alternate/longer trajectories (for each enzyme-substrate complex) will provide better quantitative
estimates of how often the substrate is ejected out of the pocket; however, we expect the results presented here will be qualitatively similar to the results from the bigger set of trajectories.

Detailed description of model substrates ACAC and AUAA interactions with bRNaseA and hRNase1-7 are provided in supporting information text. Our results indicate that while all RNases used in this work share the conserved active-site catalytic triad (two histidine residues and one lysine), subtle changes in the binding site result in distinct binding interactions for the two substrates. Interestingly, the non-catalytic active-site residues show the strongest H-bonding interactions with the substrates. However, when the detailed interaction energies are calculated, some of these residues do not show the most favorable interactions; instead, the catalytic triad shows the most favorable interaction energy. One possible interpretation of these results is that rigid interactions are required between the enzymes and the substrates for stability. However, the interactions with the catalytic triad may need to be flexible for mechanistic aspects. In other words, strong H-bonds between the substrate nucleotides and non-catalytic active-site residues holds the substrate in place, while the catalytic residues with most favorable interactions need to be non-rigid. This is consistent with a number of previous reports, whereby dynamics of several conserved catalytic residues play an important role in the mechanism of bovine RNase A.138, 216, 252

The results reported here require further comparison with experiments to confirm these observations.

Substrate preferences of bRNaseA and several human RNases were reported previously using a variety of RNA substrates.9, 211, 254 A comparison of the H-bonding properties of the different RNases with the two model substrates AUAU and ACAC in this work suggest that bRNaseA and hRNase1 show similar H-bonding interactions with the two substrates. Interestingly, the H-bonding occupancy for hRNase2 and hRNase3 is significantly lower, consistent with the lower binding affinities and catalytic rates of these enzymes relative to bRNaseA at 300 K and 310 K. Interestingly, hRNase3 shows no interactions with the substrate ACAC at 310 K, suggesting a lower preference for this substrate. Similar results are observed for hRNase4 at 310 K, which shows a significant increase in H-bonding occupancy for AUAU while a reduction in the occupancy is observed for ACAC-bound state. These observations for hRNase3 and hRNase4 at 310 K are consistent with the substrate preference reported for these enzymes.211, 254 Our results further highlight the significant differences in the observed substrate binding properties of the different RNases between 300 K and 310 K. These observations suggest the importance of selecting temperatures for experimental and computational characterization of conformational
properties and enzyme-substrate interactions that are representative of physiological conditions. As summarized in Table 1.2 in Chapter 1, the majority of previous studies have been conducted at room temperature (298 K).

Overall, the results of this study show that all the 8 investigated RNase family members have conserved active-sites residues; however, there is not a uniform preference of a substrate across all of these enzymes. The difference in substrate preferences observed in this study are broadly consistent with previous observations using a variety of RNase substrates. While our results provide structural and energy-based insights that form the basis for this preference for the different pyrimidine bases, it is not clear how this relates to the diverse biological functions of these RNases. Further investigations are required to gain insights into sequence dependence on designated biological functions of host-defense (antibacterial and antiviral), angiogenesis and RNA cleavage.

3.6 Summary of observations for each RNases investigated in this study

**bRNaseA:** Both the substrates ACAC and AUAU show stable binding in the active-site on the microsecond time-scale. The substrates show strong H-bond interactions with residues Thr45, Asp71, and Phe120. In addition, Gln11 shows H-bond interaction with substrate AUAU. The dynamical motions of the enzyme are similar in the case of two substrates, exhibiting highest flexibility in loops L1 and L6. The electrostatic surface of the bRNaseA is mostly positively charged but shows slight differences in the case of ACAC and AUAU, which appear to be an anomaly for the set of proteins investigated in this study. The residue Lys7, Val43, Thr45, Lys66, His119, and Phe120 show large favorable van der Waals and electrostatic interactions with these substrates. While the dynamical behavior at the temperatures 300 K and 310 K are qualitatively similar for both substrates, the overall strength of the favorable enzyme-substrate interactions increase at higher temperature (corresponding to more negative energy in Table 3.4); however, the number of H-bond interactions with substrate AUAU are reduced at 310 K.

**hRNase1:** Similar to the case on bRNaseA, the two substrates ACAC and AUAU show stable binding in the active-site of hRNase1 on the microsecond time-scale. However, the substrate ACAC forms strong H-bond interactions only with Thr45, while AUAU is observed to form strong H-bonds with Thr45, Asn71 and Phe120. The electrostatic surface of hRNase1 is mostly positively...
charged and very similar in the case of both substrates. Residues Arg4, Val43, Thr45, Lys66, His119 and Phe120 show the most favorable van der Waals and electrostatic interactions with both substrates. The dynamical motions of the enzyme are very similar in the case of two substrates, with large flexibility in loop L1; however, this loop shows higher flexibility when bound to ACAC than AUAU. An increase in number of H-bonds and their occupancies is observed at higher temperature (310 K) for both substrates investigated. The electrostatic and van der Waals interactions are more favorable for substrate ACAC and less favorable for substrate AUAU at the higher temperature investigated. The dynamical behavior at the temperatures 300 K and 310 K are similar in presence of both substrates.

**hRNase2:** Substrate AUAU shows stable binding in the hRNase2 active site at both temperatures, while the substrate ACAC shows stable binding only at 300 K. However, strong hydrogen bonds are observed with residues Gln15 and Thr43 in both cases. hRNase2 has mostly positively electrostatic surface when bound to both substrates as well. The van der Waals and electrostatic interactions show that residues Lys39, Gln41, Thr43 and His130 make the largest favorable contributions to interaction energy with both substrates. The dynamical motions of the enzyme are observed to be very similar for the two substrates, with high flexibility in loops L4, L6 and L7. The dynamical behavior is mostly similar at the two temperatures, except for loop L7 where increased fluctuations are observed with substrate AUAU at 310 K. The H-bond occupancies with substrate ACAC is reduced at 310 K as compared to simulations at 300 K. The interactions for substrate ACAC are weakened at higher temperature while it remains the same for AUAU at both temperatures.

**hRNase3:** Both the substrates ACAC and AUAU show significantly less stability in the active-site of hRNase3 on the microsecond time-scale as compared to the other members of the family investigated in this study. Only residue Asn71 (and Thr133 for AUAU) shows hydrogen bonding interactions with the substrates. Much like other RNases, the electrostatic surface of the hRNase3 shows the presence of mostly positively charged residues and is also similar in the case of both substrates investigated. Residues Lys39 and His129 make the largest contribution to van der Waals and electrostatic interaction energy. The protein dynamics is very similar for the two substrates, with largest flexibility in loops L2, L4 and L6; however, loop L6 has higher fluctuation when bound to AUAU than ACAC. At 310 K, no strong H-bonds are observed with substrate ACAC and the interactions become less favorable; however, the H-bond occupancies for substrate AUAU are considerably reduced, and the overall interaction energy remains the same.
Dynamics of the enzyme remains the same for most regions at the two temperatures for both substrates.

**hRNase4:** On the microsecond time-scale, both substrates are mostly stable in the active-site of hRNase4. Substrate ACAC forms strong hydrogen bonds with residues Thr45, Asn71, His117 and Phe118; while substrate AUAU forms hydrogen bonds only with Thr45 and Phe118. The electrostatic surface of hRNase4 is also mostly positively charged and similar in the case of both substrates. His117 and Phe118 are the largest contributors to van der Waals and electrostatic interaction energy with the two substrates. Lys8, Lys41 and Thr45 also make large contributions towards electrostatic and van der Waals interaction energy for both substrates. The dynamical motions are observed to be very similar for the two substrates. While the dynamical behavior is similar at the two temperatures, considerable difference is observed in H-bond occupancies. At higher temperature, the number of H-bonds with substrate ACAC is reduced significantly, whereas an increase is observed in the H-bonds and their occupancies with substrate AUAU, consistent with decrease in the interaction energy (less negative value in Table 3.4) in case of ACAC but increase in the case of AUAU. Further, these observations are consistent with the preference of the uridine nucleotide in the B₁ sub-site reported previously for hRNase4⁹.

**hRNase5:** For hRNase5, substrate ACAC is more stable in the active-site than the substrate AUAU on the microsecond time-scale. In one of the trajectories investigated, the substrate AUAU is completely ejected out of the active-site. This behavior is different than other hRNases (except for hRNase7, as discussed below), where either both or one central nucleotide of the substrate stays bound in the active-site. This is also indicated by the observation that substrate ACAC forms strong hydrogen bonds with Thr44 and His114, while no hydrogen bonds are observed between substrate AUAU and the enzyme. Much like other RNases in this study, the electrostatic surface of hRNase5 is also mostly positively charged. Residues Lys40 and His140 make the largest contribution to van der Waals and electrostatic interaction energy. The dynamical motions are observed to be very similar for the two substrates with flexibility in loops L₁, L₂ and L₆ when bound to both substrates. The dynamical behavior at two temperatures 300 K and 310 K is similar. At the higher temperature, H-bonds are formed with substrate AUAU, contrary to the observations at 300 K. Additionally, an increase in the favorable interaction is observed, while in the case of substrate ACAC, diminished interactions are observed.
**hRNase6:** Both the substrates ACAC and AUAU are stable in the hRNase6 active-site on the microsecond time-scale. Residue Gln15, Asn69 and Trp11 form strong hydrogen bonds with substrate ACAC, while substrate AUAU shows strong hydrogen bonding with Asn65, Arg67 and Asn69. The electrostatic surface of hRNase6 is mostly positively charged and similar in the case of both substrates. The van der Waals and electrostatic interactions show that His123 and Leu124 make the largest favorable contributions to interaction energy in case of both substrates. Additionally, large contributions are also observed from residues Asn65, Arg67 and Asn69 in presence of both substrates. The dynamical motions are very similar for the two substrates, except for loop L7 that shows higher fluctuations when bound to AUAU than ACAC. The dynamical behavior and hydrogen bond occupancies are alike at the two temperatures, while the favorable interactions become more pronounced at higher temperature for both substrates.

**hRNase7:** The substrates ACAC and AUAU are less stable in interacting with the active-site of hRNase7, forming only one strong hydrogen bond with Thr43 and Leu125, respectively. Similar to the case of hRNase7, in one of the trajectories investigated, substrate AUAU is completely ejected out of the active-site. The electrostatic surface of hRNase7 appears more positively charged when substrate ACAC is bound when compared to AUAU. Residues Lys39 and His124 make the largest contribution to van der Waals and electrostatic interaction energy. The enzyme dynamics is observed to be different for the two substrates with hRNase7: loop L2 and L7 show higher fluctuations when bound to ACAC; while loop L6 shows higher fluctuations when bound to AUAU. The overall dynamics of hRNase7 at two temperatures 300 K and 310 K is observed to be similar, except for loop L4 that shows higher fluctuation when substrate ACAC is bound. However, considerable change is observed in H-bond patterns at the two temperatures investigated. At higher temperature, more H-bonds with higher occupancies are formed with substrate ACAC, while no strong H-bonds are formed with substrate AUAU. Also, considerable increase in the favorable interaction energy is observed for both substrates at 310 K.
3.7 Appendix

Table 3.5: Residues showing significant favorable interactions with the substrates at 300 K. An interaction is considered as significant if the enzyme-nucleotide pair with interaction energy < -3 kcal/mol of $E_{\text{total}}$. Results for central two nucleotides and terminal nucleotides are shown separately. The enzyme residues are listed in the order of decreasing energy of interaction with the nucleotide making the favorable interaction is listed after value.

<table>
<thead>
<tr>
<th></th>
<th>center nucleotides (ACAC)</th>
<th>terminal nucleotides (ACAC)</th>
<th>center nucleotides (AUAU)</th>
<th>terminal nucleotides (AUAU)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>bRNaseA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>His119 (-10.9, A; -4.0, C)</td>
<td>Pro42 (-4.3, A)</td>
<td>His119 (-14.5, A)</td>
<td>Pro42 (-3.9, A)</td>
</tr>
<tr>
<td></td>
<td>Thr45 (-4.6, C)</td>
<td>Val43 (-3.0, A)</td>
<td>Lys7 (-3.3, C)</td>
<td>Val43 (-3.3, A)</td>
</tr>
<tr>
<td></td>
<td>Phe120 (-4.5, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Val43 (-4.4, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lys7 (-4.1, A)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lys66 (-4.1, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asn44 (-3.4, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>hRNase1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>His119 (-11.3, A; -3.7, C)</td>
<td>Pro42 (-4.4, A)</td>
<td>His119 (-11.6, A; -3.5, U)</td>
<td>Pro42 (-4.5, A)</td>
</tr>
<tr>
<td></td>
<td>Arg4 (-4.6, A)</td>
<td>Val43 (-3.2, A)</td>
<td>Lys66 (-4.8, U)</td>
<td>Val43 (-3.2, A)</td>
</tr>
<tr>
<td></td>
<td>Lys66 (-4.5, C)</td>
<td></td>
<td>Phe120 (-4.5, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thr45 (-4.5, C)</td>
<td></td>
<td>Arg4 (-3.9, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Val43 (-4.1, C)</td>
<td></td>
<td>Val43 (-3.9, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phe120 (-4.1, C)</td>
<td></td>
<td>Lys7 (-3.7, A)</td>
<td></td>
</tr>
<tr>
<td><strong>hRNase2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>His130 (-11.5, A)</td>
<td>Arg133 (-4.3, A)</td>
<td>His130 (-10.0, A)</td>
<td>Arg133 (-3.8, A)</td>
</tr>
<tr>
<td></td>
<td>Lys39 (-7.5, C; -4.3, A)</td>
<td></td>
<td>Lys39 (-5.9, U; -4.7 A)</td>
<td>Arg37 (-3.2, U)</td>
</tr>
<tr>
<td></td>
<td>Gln41 (-5.3, C)</td>
<td></td>
<td>Gln41 (-4.9, U)</td>
<td>Gln41 (-3.0, A)</td>
</tr>
<tr>
<td></td>
<td>Thr43 (-4.3, C)</td>
<td></td>
<td>Leu131 (-3.8, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asn42 (-3.7, C)</td>
<td></td>
<td>Trp8 (-3.4, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arg69 (-3.0, A)</td>
<td></td>
<td>Arg133 (-3.1, U)</td>
<td></td>
</tr>
<tr>
<td><strong>hRNase3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>His129 (-11.0, A; -3.2, C)</td>
<td>Arg2 (-7.8, C)</td>
<td>His129 (-11.8, A; -3.8, U)</td>
<td>Met1 (-6.5, U)</td>
</tr>
<tr>
<td></td>
<td>Lys39 (-5.2, C; -4.1, A)</td>
<td>Met1 (-6.7, C)</td>
<td>Lys39 (-5.4, U)</td>
<td>Thr133 (-3.7, A)</td>
</tr>
<tr>
<td></td>
<td>Gln41 (-3.2, A)</td>
<td>Gln41 (-3.2, A)</td>
<td>Lys39 (-4.7, U)</td>
<td>Gln41 (-3.7, A)</td>
</tr>
<tr>
<td></td>
<td>Arg35 (-3.2, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>hRNase4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>His117 (-15.4, A; -7.0, C)</td>
<td>Phe43 (-5.6, A)</td>
<td>His117 (-12.3, A; -3.6, U)</td>
<td>Phe43 (-6.3, A)</td>
</tr>
<tr>
<td></td>
<td>Phe118 (-6.0, C)</td>
<td></td>
<td>Phe118 (-6.2, U)</td>
<td>Arg8 (-4.4, U)</td>
</tr>
<tr>
<td></td>
<td>Lys41 (-5.7, C)</td>
<td></td>
<td>Phe43 (-4.4, U)</td>
<td>Arg8 (-3.5, A)</td>
</tr>
<tr>
<td></td>
<td>Arg8 (-3.9, A)</td>
<td></td>
<td>Lys41 (-3.9, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lys66 (-3.4, C)</td>
<td></td>
<td>Arg8 (-3.8, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thr45 (-3.1, C)</td>
<td></td>
<td>Thr45 (-3.6, U)</td>
<td></td>
</tr>
</tbody>
</table>

110
### Table 3.5 continued

<table>
<thead>
<tr>
<th>hRNase5</th>
<th>center nucleotides (ACAC)</th>
<th>terminal nucleotides (ACAC)</th>
<th>center nucleotides (AUAA)</th>
<th>terminal nucleotides (AUAA)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>His114 (-14.4, A; -4.6, C)</td>
<td>Arg5 (-6.6, C)</td>
<td>His114 (-11.7, A; -3.2, U)</td>
<td>Arg5 (-7.2, U)</td>
</tr>
<tr>
<td></td>
<td>Lys40 (-5.8, C)</td>
<td>His8 (-3.4, C)</td>
<td>Lys40 (-5.3, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gln117 (-5.2, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leu115 (-4.0, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thr44 (-3.8, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arg121 (-3.2, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hRNase6</td>
<td>His123 (-9.8, A; -5.4, C)</td>
<td>Arg5 (-4.8, C)</td>
<td>His123 (-11.5, A; -5.2, U)</td>
<td>Arg67 (-3.9, U)</td>
</tr>
<tr>
<td></td>
<td>Lys8 (-6.2, A)</td>
<td>Lys8 (-4.6, C)</td>
<td>Lys64 (-5.0, U)</td>
<td>Lys8 (-3.5, U)</td>
</tr>
<tr>
<td></td>
<td>Lys39 (-5.5, C)</td>
<td>Gln41 (-4.1, A)</td>
<td>Lue124 (-4.6, U)</td>
<td>Gln41 (-3.5, A)</td>
</tr>
<tr>
<td></td>
<td>Leu124 (-5.3, C)</td>
<td></td>
<td>Asn69 (-3.7, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trp11 (-4.0, A)</td>
<td></td>
<td>Asn65 (-3.6, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asn69 (-3.4, A)</td>
<td></td>
<td>Arg67 (-3.3, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>His16 (-3.2, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arg67 (-3.2, A)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asn65 (-3.0, A)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gln15 (-3.0, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hRNase7</td>
<td>His124 (-7.9, A)</td>
<td>Leu41 (-5.7, A)</td>
<td>His124 (-4.2, A)</td>
<td>Met1 (-3.2, U)</td>
</tr>
<tr>
<td></td>
<td>Thr43 (-4.2, C)</td>
<td>Arg37 (-3.9, C)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lys39 (-4.0, C)</td>
<td>Met1 (-3.4, C)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asn42 (-3.8, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leu41 (-3.8, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.6: Residues showing significant favorable interactions with the substrates at 310 K. An interaction is considered as significant if the enzyme-nucleotide pair with interaction energy < -3 kcal/mol of $E_{total}$. Results for central two nucleotides and terminal nucleotides are shown separately. The enzyme residues are listed in the order of decreasing energy of interaction with the nucleotide making the favorable interaction is listed after value.

<table>
<thead>
<tr>
<th></th>
<th>center nucleotides (ACAC)</th>
<th>terminal nucleotides (ACAC)</th>
<th>center nucleotides (AUAU)</th>
<th>terminal nucleotides (AUAU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bRNaseA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>His119 (-10.2, A; -4.0, C)</td>
<td>Pro42 (-3.3, A)</td>
<td>His119 (-11.1, A; -4.0, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thr45 (-4.4, C)</td>
<td></td>
<td>Phe120 (-5.2, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phe120 (-4.2, C)</td>
<td>Val43 (-3.8, U)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lys66 (-4.1, C)</td>
<td>Thr45 (-3.3, U)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Val43 (-4.0, C)</td>
<td>Asn71 (-3.1, A)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lys7 (-3.9, A)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asn44 (-3.1, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hRNase1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>His119 (-11.5, A; -4.2, C)</td>
<td>Pro42 (-4.4, A)</td>
<td>His119 (-11.6, A; -3.5, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lys66 (-5.1, C)</td>
<td>Val43 (-3.3, A)</td>
<td>Lys66 (-5.3, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thr45 (-4.5, C)</td>
<td></td>
<td>Val43 (-4.0, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phe120 (-4.4, C)</td>
<td></td>
<td>Phe120 (-3.9, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Val43 (-4.4, C)</td>
<td></td>
<td>Lys41 (-3.4, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arg4 (-4.3, A)</td>
<td></td>
<td>Thr45 (-3.4, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Arg4 (-3.4, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Asn71 (-3.1, A)</td>
<td></td>
</tr>
<tr>
<td>hRNase2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>His130 (-8.3, A; -5.7, C)</td>
<td>Asn40 (-3.02, C)</td>
<td>His130 (-11.9, A; -4.4, U)</td>
<td>Arg69 (-4.1, U)</td>
</tr>
<tr>
<td></td>
<td>Lys39 (-6.1, C; -4.1, A)</td>
<td></td>
<td>Lys39 (-5.5, U; -4.4, A)</td>
<td>Arg133 (-4.4, A)</td>
</tr>
<tr>
<td></td>
<td>Met1 (-3.3, A)</td>
<td></td>
<td>Arg133 (-4.8, U)</td>
<td>Gln41 (-4.4, U)</td>
</tr>
<tr>
<td></td>
<td>Gln41 (-3.1, C)</td>
<td></td>
<td>Gln41 (-4.4, U)</td>
<td></td>
</tr>
<tr>
<td>hRNase3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>His129 (-9.8, A; -6.6, C)</td>
<td>Met1 (-4.78, C)</td>
<td>His129 (-9.3, A; -3.7, U)</td>
<td>Arg2 (-4.4, U)</td>
</tr>
<tr>
<td></td>
<td>Arg35 (-5.4, A)</td>
<td></td>
<td>Lys39 (-4.9, U; -3.6, A)</td>
<td>Arg35 (-3.5, U)</td>
</tr>
<tr>
<td></td>
<td>Lys39 (-4.8, C)</td>
<td>Met1 (-3.9, U; -3.3, A)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Met1 (-4.4, A; -4.1, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arg2 (-4.0, A)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hRNase4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>His117 (-10.0, A; -4.9, C)</td>
<td>Phe43 (-6.8, A)</td>
<td>His117 (-11.8, A; -4.4, U)</td>
<td>Phe43 (-5.9, A)</td>
</tr>
<tr>
<td></td>
<td>Lys41 (-6.2, C; -3.5, A)</td>
<td>Arg8 (-4.4, C)</td>
<td>Phe118 (-6.5, U)</td>
<td>Arg8 (-3.9, U)</td>
</tr>
<tr>
<td></td>
<td>Phe118 (-5.7, C)</td>
<td></td>
<td>Arg8 (-4.3, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arg8 (-5.6, A)</td>
<td></td>
<td>Lys41 (-4.2, U; -3.7, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thr45 (-3.2, C)</td>
<td></td>
<td>Phe43 (-4.2, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lys66 (-3.2, C)</td>
<td></td>
<td>Thr45 (-3.5, U)</td>
<td></td>
</tr>
<tr>
<td>hRNase5</td>
<td>center nucleotides</td>
<td>terminal nucleotides</td>
<td>center nucleotides</td>
<td>terminal nucleotides</td>
</tr>
<tr>
<td>---------</td>
<td>---------------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td></td>
<td>(ACAC)</td>
<td>(ACAC)</td>
<td>(AUAU)</td>
<td>(AUAU)</td>
</tr>
<tr>
<td>hRNase5</td>
<td>His114 (-7.0, A; -4.2, C)</td>
<td>Arg5 (-4.5, C)</td>
<td>Gln117 (-6.8, U)</td>
<td>Arg121 (-3.1, A)</td>
</tr>
<tr>
<td></td>
<td>Lys40 (-6.6, C)</td>
<td></td>
<td>His114 (-6.7, A; -6.5, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thr44 (-4.3, C)</td>
<td></td>
<td>Arg5 (-4.8, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ile42 (-4.0, C)</td>
<td></td>
<td>Leu115 (-4.2, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leu115 (-3.6, C)</td>
<td></td>
<td>Lys40 (-3.6, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phe120 (-3.8, C)</td>
<td></td>
<td>His13 (-3.5, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asn43 (-3.5, C)</td>
<td></td>
<td>Asn43 (-3.2, U)</td>
<td></td>
</tr>
<tr>
<td>hRNase6</td>
<td>His123 (-9.2, A; -4.5, C)</td>
<td>Arg5 (-6.4, C)</td>
<td>Gln117 (-10.9, A; -5.3, U)</td>
<td>Gln41 (-3.9, A)</td>
</tr>
<tr>
<td></td>
<td>Lys39 (-6.6, C)</td>
<td>Gln41 (-3.5, A)</td>
<td>Leu124 (-5.3, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lys8 (-6.6, A)</td>
<td>Lys8 (-3.4, C)</td>
<td>Lys8 (-4.1, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leu124 (-4.5, C)</td>
<td></td>
<td>Lys64 (-4.0, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trp11 (-4.1, A)</td>
<td></td>
<td>Asn69 (-3.7, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thr43 (-3.8, C)</td>
<td></td>
<td>Asn65 (-3.4, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asn69 (-3.6, A)</td>
<td></td>
<td>Arg67 (-3.4, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arg67 (-3.4, A)</td>
<td></td>
<td>Thr43 (-3.0, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gln15 (-3.3, C)</td>
<td></td>
<td>Gln41 (-3.9, A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asn65 (-3.1, A)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hRNase7</td>
<td>His124 (-9.2, A; -4.0, C)</td>
<td>Met1 (-4.6, C)</td>
<td>His124 (-7.3, A; -3.8, U)</td>
<td>Trp11 (-4.0, U)</td>
</tr>
<tr>
<td></td>
<td>Lys39 (-5.5, C)</td>
<td>Leu41 (-4.7, A)</td>
<td>Met1 (-4.7, A)</td>
<td>Leu41 (-3.5, A)</td>
</tr>
<tr>
<td></td>
<td>Thr43 (-4.4, C)</td>
<td>Arg37 (-3.6, C)</td>
<td>Lys64 (-3.9, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leu125 (-4.2, C)</td>
<td>Lys2 (-3.9, C)</td>
<td>Lys39 (-3.7, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leu41 (-3.5, C)</td>
<td></td>
<td>Leu125 (-3.2, U)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asn42 (-3.2, C)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.12: Important H-bond behavior over time for bRNaseA-substrate complexes. Results from ACAC are shown on top, and from AUAC complex on the bottom. The two alternate trajectories are marked as (1) and (2) for two temperatures 300 K and 310 K. Red lines mark the distance range used for defining H-bonds.
Figure 3.13: Important H-bond behavior over time for hRNase1-substrate complexes. Results from ACAC are shown on top, and from AUAC complex on the bottom. The two alternate trajectories are marked as (1) and (2) for two temperatures 300 K and 310 K. Red lines mark the distance range used for defining H-bonds.
Figure 3.14: Important H-bond behavior over time for hRNase2-substrate complexes. Results from ACAC are shown on top, and from AUAC complex on the bottom. The two alternate trajectories are marked as (1) and (2) for two temperatures 300 K and 310 K. Red lines mark the distance range used for defining H-bonds.
Figure 3.15: Important H-bond behavior over time for hRNase3-substrate complexes. Results from ACAC are shown on top, and from AUAC complex on the bottom. The two alternate trajectories are marked as (1) and (2) for two temperatures 300 K and 310 K. Red lines mark the distance range used for defining H-bonds.
Figure 3.16: Important H-bond behavior over time for hRNase4-substrate complexes. Results from ACAC are shown on top, and from AUAC complex on the bottom. The two alternate trajectories are marked as (1) and (2) for two temperatures 300 K and 310 K. Red lines mark the distance range used for defining H-bonds.
Figure 3.17: Important H-bond behavior over time for hRNase5-substrate complexes. Results from ACAC are shown on top, and from AUAC complex on the bottom. The two alternate trajectories are marked as (1) and (2) for two temperatures 300 K and 310 K. Red lines mark the distance range used for defining H-bonds.
Figure 3.18: Important H-bond behavior over time for hRNase6-substrate complexes. Results from ACAC are shown on top, and from AUAC complex on the bottom. The two alternate trajectories are marked as (1) and (2) for two temperatures 300 K and 310 K. Red lines mark the distance range used for defining H-bonds.
Figure 3.19: Important H-bond behavior over time for hRNase7-substrate complexes. Results from ACAC are shown on top, and from AUAC complex on the bottom. The two alternate trajectories are marked as (1) and (2) for two temperatures 300 K and 310 K. Red lines mark the distance range used for defining H-bonds.
Chapter 4  Conformational Sub-states Associated with Product Release in Ribonucleases
4.1 Abstract

Conformational fluctuations between different sub-states with unique configurations allow enzymes to efficiently catalyze reactions with high turnover. However, establishing a direct link between protein motions and enzymatic function remains challenging. The structural basis of RNA hydrolysis has been extensively studied in bovine ribonuclease A (bRNaseA), and product release is the suggested rate-limited step. However, conformational motions that affect the turnover rate have received relatively little attention for bovine, as well as human ribonucleases (hRNases). This work identifies conformational sub-states and conformational fluctuations that play a role in product release step in hRNases and bRNaseA. Conformations collected from molecular dynamics simulations with umbrella sampling were analyzed using higher order statistical clustering method; this method, quasi anharmonic analysis, allows identification of conformational sub-states. The results revealed the existence of functionally relevant conformational sub-sates in product release step of these closely related enzymes. Enzyme residues that interact with cleaved RNA products were also identified. Conformational fluctuations were observed in distal regions away from the reactive center of the enzymes. Residues in these regions have previously been reported to experience perturbations during $^{15}$N-CPMG relaxation dispersion experiments, which correspond to the time-scale of RNase function.

4.2 Introduction

Members of the ribonuclease (RNase) superfamily catalyze the hydrolysis of ribonucleic acid (RNA). Bovine ribonuclease A (bRNaseA), which is the prototype of the family, is one of the most widely studied enzyme. In recent years, interest in human members of the RNase superfamily has increased due to their reported clinical importance and increasing therapeutic applications. A number of studies have focused on the intracellular localization, tissue distribution and biological role of human RNases (hRNases). Studies have also focused on the structure-function relationship in these enzymes. However, the wide range of catalytic efficiencies found in hRNases cannot be explained in terms of structure alone. Enzymes are intrinsically dynamic, and conformational motions in enzymes are known to play an essential role in their function. The role of dynamics in the catalytic cycle of hRNases has not yet been explored in detail. Understanding functionally relevant dynamics in hRNases is will open new avenues for therapeutic interventions in these enzymes.
Enzymes undergo a wide range of motion in terms of space and time. The time-scales of motion in enzymes range from picoseconds to seconds (ps - s).\textsuperscript{300} The faster atomic fluctuations occur on the ps time-scales. Motions in enzyme backbone and amino acid side chains are observed in the picoseconds to nanoseconds (ps - ns) time-scales. Conformational rearrangements of loops and domains occur on the milliseconds to seconds (ms - s) time-scale.\textsuperscript{128, 301} Any of these motions can be functionally significant and directly related to catalysis.\textsuperscript{299, 300} These motions promote catalysis by allowing the enzymes to overcome the activation barrier and reach its active form. The time-scale of catalytically important motion is dependent on the enzyme. Correlation between the time-scale of conformational fluctuation and catalytic turnover has been established for a number of enzyme systems including bRNaseA.\textsuperscript{128, 136, 137, 140, 302, 303}

The rate of conformational exchange has been shown to correlate with the rate limiting step, product release, in bRNaseA.\textsuperscript{136, 140} The structure of bRNaseA is well-studied by both homonuclear solution NMR and X-ray crystallography.\textsuperscript{7} The mechanism of RNA cleavage by RNases has been extensively studied by both experimental and theoretical approaches,\textsuperscript{7, 85} which suggest that microsecond to millisecond (µs - ms) motions occur throughout the bRNaseA enzyme, including fluctuations in the active site residues.\textsuperscript{136, 267} The rate of conformational kinetics measured in NMR spin relaxation experiments (average $k_{ex} = 1600 \text{ s}^{-1}$) is similar to the enzyme turnover rate ($k_{cat} = 1900 \pm 130 \text{ s}^{-1}$).\textsuperscript{136} Note, that the NMR experiments were performed in the apo state. Additionally, NMR line shape analysis of titrations of product analogue (3′-CMP) into bRNaseA yields a dissociation constant ($K_{off}$) of $\sim 1200 \text{ s}^{-1}$ suggesting that product release ($K_{off}$) is the rate-limiting step in bRNaseA.\textsuperscript{138} It was further implicated that the motion of loop 1 enables efficient release of product, and disrupting these motions results in 10-fold decrease in the product release rate.\textsuperscript{138, 216} These evidences presents a clear connection between enzyme motions, catalytic turnover and product release in bRNaseA.

Dynamical fluctuations occurring over a wide range of time-scales allow enzymes to sample a wide variety of distinct conformations (also called as conformational sub-states).\textsuperscript{129, 304} These conformational sub states are sometimes rarely populated, high energy states that can contain functionally important features.\textsuperscript{122} Conformational transitions between low energy ground states to higher energy sub-states are vital for the events occurring in an enzyme’s reaction pathway; substrate recognition and binding, the chemical step and product release.\textsuperscript{204, 305} Conformational fluctuations that govern the catalytic mechanism in enzyme systems like DHFR,\textsuperscript{306} CypA,\textsuperscript{204} HIV-protease,\textsuperscript{302} and DNA polymerase\textsuperscript{307} have been identified. Interestingly in these systems,
Conformational fluctuations in distal regions (sometimes more than 10 Å away from the active site) modulate the rate-limiting step. Note that the rate-limiting step is not same for all enzymes, for example the rate limiting-step in CypA is the catalytic step, while in DHFR product release is rate-limiting.

Conformational fluctuations are essential for enzyme function, and therefore characterizing the conformational sub-states in the rate limiting step is of paramount importance to understanding of the role dynamics in enzyme function. A deeper understanding the conformational motions in each step of reaction process of hydrolysis of RNA by these enzymes is essential for the rational design of novel therapies toward diseases associated with RNase action. Chapter 3 demonstrates the occurrence of conformational motions in the substrate bound form of RNases. Nevertheless, detailed characterization of conformational change coupled to the product release in these enzymes remains unknown. Very few studies have focused on the release of nucleotide products in bRNaseA, and to the best our knowledge none for human RNases. Note, that the rate-limiting step in hRNases is still unknown. Therefore, this study is focused on understanding how the protein conformations and dynamical properties evolve during product release; what are the key residues involved; and whether these motions and residues involved in product release are conserved in members of this diverse superfamily.

Studies from chapter 2 report that members of the RNase superfamily can be classified into sub-families based on their primary sequence and intrinsic dynamics, and the biological function of the members within a sub-family is same\textsuperscript{308}. To characterize the role of dynamics in function of individual sub-family, it is required to determine the conformational sub-states in at least one representative member in each family. Comparative analysis of conformational dynamics associated with product release in each sub-family will aid in understanding the role of dynamics in the function of the entire RNases superfamily. The select RNases in this study include hRNase3 from the \textit{EAR-like} sub-family, hRNase5 from the \textit{Angiogenin} sub-family and bRNaseA from the \textit{RNase A-like} sub-family. Conformational sub-sates in these enzymes occurring during product release are characterized using quasi anharmonic analysis (QAA)\textsuperscript{129}, a higher order statistical method, with conformations obtained from umbrella sampling molecular dynamics simulations. Conformational transitions associated with product release provide insights into the existence of separate sub-states sampled by these closely related enzymes.
4.3 Materials and methods

Modeling enzyme-product(s) complexes: The cleaved nucleotide products (Figure 4.1) were modeled into the active site of two human ribonucleases, hRNase3 (PDB ID: 1QMT) and hRNase5 (PDB ID: 1ANG), as well as the active site of bRNaseA (PDB ID: 7RSA). The coordinates for the phosphate and ribose sugar backbone for the products were adopted from the DNA ligand in the crystal structure of bRNaseA (PDB id 1RCN). The nucleotide bases of the substrate were modified to match the desired nucleotide using AMBER’s AMBER’s xleap modeling and preparation module. The parameters for both protein residues and RNA products were obtained from AMBER’s ff14SB force-field.

Histidine protonation states: The two catalytic histidine residues, His12 and His119 (bRNaseA numbering), were modeled in single and double protonated states respectively in all three RNases. The protonation states for these residues were determined based on pH profiles for bRNaseA, as well as structural interaction with the modeled substrates and products. All the other histidine residues were modeled in the single protonated state for all three RNases.

Umbrella sampling simulations: All modeled complexes were neutralized with the addition of chloride (Cl-) counterions and subsequently solvated in a SPC/E water box with 15 Å distance between the edges on all sides of the protein and the water box. Periodic boundary conditions were applied, and the solvated complexes were equilibrated using a previously described protocol. To monitor the release of product from the active site of RNases, the distance between the geometric center of mass (COM) of the side chain of three catalytic residues (two histidine residues and one lysine residue) and COM of each of the products were increased from 4 Å up to 18 Å (Figure 4.2 & 4.3). Ten ns of distance restrained simulations were performed to increase the distance between the two COMs in 1 Å increments and generate starting points for 15 windows. This increase in the distance between the two COMs is referred to as the reaction coordinate of the product release pathway. Conformations in each window along the reaction coordinate were sampled for a total of 70 ns with an appropriate biasing harmonic potential to ensure a reasonable overlap for all windows. Distance restraints were applied between the two COMs and a 2 femtoseconds (fs) time-step was used. All hydrogen atoms involved in bonds were constrained with the SHAKE algorithm. All simulations were performed at 300 K using AMBER’s pmemd.sander simulation engine. Time-averaged distance restraints between the two COMs were saved after every 10 steps. In each window, for each of the unique RNase-
product(s) complex, 3500 structures (20 picoseconds apart) were saved and analyzed for structural effects and conformational dynamics as described below.

**Analysis:** To obtain an accurate description of the conformational dynamics associated with the product release process, continuous sampling along the reaction coordinate is required. To ensure a reasonable overlap for all windows along the reaction coordinate, the distribution probability of all the time-averaged distance restraints values between the two COMs were binned into histograms. Using *xmgrace* plotting too histograms were binned from 3 – 20 Å into 1000 bins. Additionally, to ensure that the enzyme samples reasonable conformations in all windows along the reaction coordinate, the root mean square deviation (RMSD), referenced to the first structure of first window (4 Å) in the reaction coordinate, were computed for all trajectories. RMSDs were computed from the simulation trajectories using AMBER's *cpptraj* program. Simulation trajectories where the RMSD was beyond 4 Å were repeated until the sampled conformations had RMSD below the cutoff. Hydrogen atoms were excluded in the RMSD calculations.

**Hydrogen bond interactions:** Hydrogen bond (H-bond) interactions between the RNases and the product nucleotides were computed using AMBER’s *cpptraj* program. H-bonds with bond length less than 3.3 Å for the donor to acceptor heavy atom distance at an angle between 135° and 180° between the donor-hydrogen-acceptor were considered for analysis. The percentage occupancies of H-bonds, defined as a fraction (percentage) of the total number of structures where the interaction meets the H-bond criterion mentioned above, were also calculated.

**Hydrophobic interactions:** Hydrophobic interactions between the substrate and backbone atoms of hydrophobic residues glycine, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan, and tyrosine were computed using AMBER’s *cpptraj* program. A distance cutoff of 4.5 Å was used. Hydrophobic interactions that were present in more than 50% of all simulation snapshots were considered for analysis.

**Potential of mean force along the reaction coordinate:** To remove the biasing due to harmonic restraints in the umbrella sampling, the potential of mean force (PMF) was computed using the weighted histogram analysis method (WHAM). The weighted histogram calculation was performed from 4 - 20 Å with 1000 bins. All calculation were performed at 300 K. *WHAM* version 2.0.9 was used for the analysis.
Figure 4.1: Schematic representation of ribonuclease reaction. Tetranucleotide RNA substrate (AUAU) is shown on right and cleaved products are shown on left. A- Adenine, U – Uracil, 5’ A - 5’ Adenine, 3’ U - 3’ Uracil, U-p - Uracil with phosphate group.
Figure 4.2: Representation of geometric center of mass (COM) of catalytic residues and products. Enzyme ribonuclease is shown as a gray ribbon, while the catalytic triad residues are shown as sticks; COM of catalytic residues are represented in cyan sphere. The two products obtained after cleavage of tetranucleotide substrate are shown as sticks, and their respective COMs are shown as cyan spheres. 5’ A - 5’ Adenine, 3’ U - 3’ Uracil, U-p - Uracil with phosphate group.
Figure 4.3: Schematics of product release umbrella sampling simulations. Enzyme ribonuclease (RNase) is represented as a pink sphere, the catalytic triad is represented as a green sphere, the cleaved products are represented as blue rectangles and the center of mass (COM) of catalytic triad and products are represented as cyan spheres. Dotted arrows represent the distance between the COM of product(s) and catalytic triad, and the corresponding distance is labelled.
Intrinsic dynamics along the reaction coordinate: The simulation trajectories from 15 windows along the reaction coordinate were grouped into three categories; the first includes simulation trajectories from windows 4 Å to 8 Å, i.e. when the products are in the near vicinity of the catalytic triad residues. The second includes simulation trajectories from windows 9 Å to 13 Å, when the products are on the surface of the protein, but not in direct contact with the catalytic triad residues. And third is when the products are in the bulk solvent (14 – 18 Å), making no direct contact with the enzyme. The fluctuations of Cα atoms for all residues were computed to analyze the intrinsic dynamics profile for each of the three categories. An aggregate of root mean square fluctuation (RMSF) of the top 10 (RMSF10) quasi-harmonic analysis (QHA) modes were used in the analysis to focus on the principal dynamics. The magnitude of displacement eigenmodes of QHA were computed using AMBER’s cpptraj program.

Conformational sub-states using quasi-anharmonic analysis (QAA): The 52,500 conformational snapshots obtained from the umbrella sampling simulations were used for QAA. The original 3N dimensional space for the Cα atoms was projected onto 100 - dimensional space using QHA modes and then analyzed using QAA. This method provides an exploration of a non-equilibrium process as compared to the equilibrium state that is explored in free MD. Using a mixture of gaussian modes three clusters representing various regions of the reaction coordinate were identified. The three QAA-independent component vectors were analyzed in detail to characterize the conformational sub-states associated with the release of products in each of the three RNases studies. Note that the protein Cα atom positions were used for this analysis.

4.4 Results

4.4.A Evaluating the conformational coverage of umbrella sampling simulations

Umbrella sampling simulations along the reaction coordinate pertaining to product release were performed for hRNase3, hRNase5 and bRNaseA. Conformations obtained from the simulation trajectories were binned into histograms to ensure continuous sampling along the reaction coordinate. Histogram binning was based on the distance between the COMs of the catalytic triad and the nucleotide bases of the product(s). In general, all binned histograms followed a gaussian distribution for both products of all three enzymes (Figure 4.4). For hRNase3 and bRNaseA, the histogram binning show overlap in adjacent windows, indicating continuous and ample sampling of the reaction coordinate. In hRNase5, for product 1 there is a gap between the histograms.
around 8 Å, indicating that additional sampling is required. The sampling for other areas of the hRNase5 reaction coordinate falls within the acceptable limits.

To ensure that the force constants and restraints applied in the umbrella sampling simulations do not render the enzyme structure unphysical, RMSD with reference to the first structure of the first window (4 Å) were computed for individual trajectories (Figure 4.5) of all the windows along the reaction coordinate. For all three RNases the maximum RMSD for the conformations sampled in the trajectory was less than 4 Å. A visual inspection of equally space conformations from the simulation trajectories ensured that each of the three enzymes adopts its near-native structure. The cysteine residues continued to form disulfide bonds during all simulations maintaining the signature kidney shape fold of these enzymes.

### 4.4.B Computational Analysis of enzyme-product interactions

To quantitatively characterize the non-bonded interactions of product nucleotides with the enzyme residues along the reaction coordinate, H-bond and hydrophobic interactions were calculated for all conformations of the umbrella sampling simulations. The average distance between heavy atoms of the enzyme and product nucleotide, and the percentage occupancy of the H-bonds along the reaction coordinate were calculated (Table 4.1). Weak interactions of the product nucleotides with hydrophobic amino acids are reported in Table 4.2. These H-bond and hydrophobic interactions have been reported to increase the binding between protein and ligands, be essential for full activity of enzymes and be an important aspect of rational design of potential drugs.\(^{318-322}\) Thus, product nucleotides must overcome these interactions to completely release themselves from the enzyme after catalysis.

The results indicate occurrence of more H-bonds in all three enzymes when the products are in the near vicinity of the catalytic triad (4 - 8 Å). Most of the residues listed in Table 4.1 in the 4 - 8 Å trajectories were reported to strongly interact with the substrate nucleotides (Chapter 3, Table 3.2). As seen in Chapter 3, the substrate nucleotides strongly interacted with the residues of their respective sub-sites (B₁ and B₂) for all three enzymes. These interactions remain conserved until the products are in the near vicinity of the catalytic triad. In other words, same interactions are present for the substrate and products, until the products leave the near vicinity of the catalytic triad. The B₁ sub-site residues Thr44 (for hRNase5) and Thr45 (for bRNaseA) and the B₂ sub-site residue Asn71 (for hRNase3 and bRNaseA) strongly interact with over 50% occupancy with the
Figure 4.4: Overlap of histograms between umbrella sampling windows. To ensure a reasonable overlap for all windows along the reaction coordinate (4 - 18 Å of COM spacing), the distribution probability of all the time-averaged distance restraints values between the two COMs were binned into histograms. Using *xmgrace* plotting tool histograms were binned from 3 – 20 Å into 1000 bins. The histograms follow a gaussian distribution.
products residues. The catalytic residue His129 (hRNase3), His114(hRNase5) and His119 (bRNaseA) form strong H-bonds when the product nucleotides are close to the active site. Note that this observation is contradictory to our observation of the substrate bound state where the catalytic triad residues do not form strong H-bonds with the substrate nucleotide. Therefore, upon hydrolysis of the phosphodiester bond, the catalytic histidine residue interacts with the newly formed product.

When the product nucleotides move along the reaction coordinate, 9 - 13 Å away from the COM of the catalytic triad, they interact with the surface of the enzyme. At this point along the reaction coordinate, H-bonds between the products and most residues of the active site and the binding sub-sites (B₁ and B₂) are lost for all three RNases. In hRNase3, product nucleotides form new interactions with residues Arg2, Arg8 (N terminal helix) and Arg35. In hRnase5, one of the nucleotide products interacts with C-terminal residues Phe120 and Arg121. These residues are away from the catalytic triad, on the surface of enzyme, suggesting possible interaction of the product nucleotide with other regions of the enzyme before they are completely dissociated and released in the bulk solvent. Such interactions are not observed for bRNaseA. Table 4.2 reports a number of hydrophobic interactions for hRNase3 and bRNaseA when the products are in near vicinity of the catalytic triad and on the surface of the enzyme further along the reaction coordinate.
coordination. Comparatively, the extent of hydrophobic interactions is lesser for hRNase5. Finally, when the product nucleotides are more than 13 Å away on the reaction coordinate they do not form any H-bonds or hydrophobic contacts with the enzyme residues suggesting that they are completely released from the enzyme into the bulk solvent.

4.4.C Conformational dynamics associated product release

The conformational dynamics of RNases during the product release were quantified by QHA of the conformations obtained from the umbrella sampling simulations along the reaction coordinate. RMSF\textsubscript{10}, the aggregate root means square fluctuation of the top 10 slowest QHA modes, provides a good measure of a protein’s conformational flexibility.\textsuperscript{284, 294} RMSF\textsubscript{10} was computed for all the three categories of the reaction coordinate; when the products are in near vicinity of the catalytic triad (4 - 8 Å), when the products are on/near the surface of the enzyme (9 - 13 Å), and when the products are in bulk solvent making no contact with the enzyme (14 - 18 Å). The RMSF\textsubscript{10} profiles of the product release simulations indicate that the most significant differences in dynamics occur in the loop regions of RNases (Figure 4.6), similar to the apo and substrate bound studies (Chapters 2 and 3).

The dynamical characterization along the reaction coordinate indicates distinct differences in flexibility of the three RNases during product release. Difference in flexibility along the reaction coordinate is also observed for hRNase3, hRNase5 and bRNaseA. Specifically, in hRNase3, loop L4 (residues 58-71) has higher fluctuations when the products are in the near vicinity (4 - 8 Å), of the catalytic triad. As the products move away (9Å and beyond) from the catalytic triad along the reaction coordinate, dynamics in L4 decrease considerably. Dynamical fluctuations in L4 is considerably lower in hRNase5 and bRNaseA for all three categories of the reaction coordinate. In hRNase5, loops L1, L4 and L6 show fluctuations of equal magnitude throughout the reaction coordinate. In bRNaseA, loop L1 (residues 13-21) exhibits considerably higher fluctuations when the products are in the near vicinity of the catalytic triad. Motions in L1 subside when the products move away from the catalytic triad. Loop L1 of bRNaseA has been shown to affect product dissociation, and thereby modulate the catalytic efficiency.\textsuperscript{138} Loop L1 displays considerably lower fluctuations in hRNase5 while it is truncated in hRNase3. For all the three enzymes, fluctuations in the other loop regions remain constant while the product is moving along the reaction coordinate.
Table 4.1: H-bonding interaction of pancreatic-type RNases with product nucleotides. The H-bonds that meet the criteria of 3.3 Å, angle between 135 - 180° and more than 45% occupancy are listed. P1 - product 1; P2 - product 2; A5' - 5' adenine nucleotide; U3' - 3' uracil nucleotide; U-p - Uracil nucleotide with the cleaved phosphate group (refer to Figure 4.2)

<table>
<thead>
<tr>
<th>RC (Å)</th>
<th>hRNase3 Residue</th>
<th>Product Nucleotide</th>
<th>Bond Length (Å)</th>
<th>hRNase5 Residue</th>
<th>Product Nucleotide</th>
<th>Bond Length (Å)</th>
<th>bRNaseA Residue</th>
<th>Product Nucleotide</th>
<th>Bond Length (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Glu15 (Nε2)</td>
<td>U3' (O2) (P2)</td>
<td>2.82 (59.83)</td>
<td>Gln117 (Nε2)</td>
<td>U-p (O4) (P1)</td>
<td>2.82 (72.06)</td>
<td>Asn71 (Oδ1)</td>
<td>A5' (Nε) (P2)</td>
<td>2.88 (52.17)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>His114 (Nδ1)</td>
<td>U3' (O2) (P2)</td>
<td>2.82 (63.86)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>U3' (O2) (P2)</td>
<td>2.76 (60.51)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Thr44 (N)</td>
<td>U-p (O2) (P1)</td>
<td>2.85 (63.71)</td>
<td>Gln117 (Nε2)</td>
<td>U-p (O4) (P1)</td>
<td>2.82 (72.06)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>U-p (O4) (P1)</td>
<td>2.82 (63.86)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>U3' (O2) (P2)</td>
<td>2.76 (60.51)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Val128 (O)</td>
<td>A5' (O5) (P2)</td>
<td>2.69 (72.14)</td>
<td>Val128 (O)</td>
<td>A5' (O5) (P2)</td>
<td>2.85 (63.71)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arg2 (N)</td>
<td>U3' (O2) (P2)</td>
<td>2.86 (55.51)</td>
<td>Arg2 (Nε)</td>
<td>U3' (O2) (P2)</td>
<td>2.85 (63.71)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Arg8 (Nε)</td>
<td>U3' (O2) (P2)</td>
<td>2.86 (63.71)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>His129 (Nδ1)</td>
<td>A5' (O5) (P2)</td>
<td>2.86 (55.77)</td>
<td>Thr44 (N)</td>
<td>U-p (O2) (P1)</td>
<td>2.86 (59.34)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asn71 (Oδ1)</td>
<td>U-p (O3) (P1)</td>
<td>2.86 (51.66)</td>
<td></td>
<td>U-p (O4) (P1)</td>
<td>2.82 (72.06)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leu130 (O)</td>
<td>U-p (O3) (P1)</td>
<td>2.86 (51.66)</td>
<td></td>
<td>U-p (O4) (P1)</td>
<td>2.82 (72.06)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>U-p (O4) (P1)</td>
<td>2.82 (72.06)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Asn71 (Oδ1)</td>
<td>A5' (O5) (P2)</td>
<td>2.86 (56.66)</td>
<td>Leu115 (N)</td>
<td>U-p (O4) (P1)</td>
<td>2.82 (86.40)</td>
<td>Asn71 (Nδ2)</td>
<td>A5' (Nε) (P2)</td>
<td>2.87 (54.43)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Thr44 (N)</td>
<td>U-p (O4) (P1)</td>
<td>2.82 (86.40)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>U-p (O4) (P1)</td>
<td>2.82 (86.40)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Arg35 (Nε)</td>
<td>U-p (O4) (P1)</td>
<td>2.81 (70.89)</td>
<td>His114 (Nδ2)</td>
<td>U3' (O2) (P2)</td>
<td>2.78 (56.51)</td>
<td>Asn71 (Nδ2)</td>
<td>A5' (Nε) (P2)</td>
<td>2.90 (57.86)</td>
</tr>
<tr>
<td></td>
<td>Arg35 (Nδ2)</td>
<td>U-p (O4) (P1)</td>
<td>2.80 (67.77)</td>
<td></td>
<td>U3' (O2) (P2)</td>
<td>2.78 (56.51)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>U-p (O4) (P1)</td>
<td>2.80 (67.77)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RC: Reaction Coordinate


Table 4.1 continued

<table>
<thead>
<tr>
<th>RC (Å)</th>
<th>Residue</th>
<th>Product Nucleotide</th>
<th>Bond Length (%) occupancy</th>
<th>Residue</th>
<th>Product Nucleotide</th>
<th>Bond Length (%) occupancy</th>
<th>Residue</th>
<th>Product Nucleotide</th>
<th>Bond Length (%) occupancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Arg2 (Nε)</td>
<td>U3' (O5') (P2)</td>
<td>2.79 (87.91)</td>
<td>hRNase5</td>
<td>Arg8 (Nε)</td>
<td>U3' (O5) (P2)</td>
<td>2.82 (81.60)</td>
<td>bRNaseA</td>
<td>Leu115 (N)</td>
</tr>
<tr>
<td></td>
<td>Arg2 (Nε)</td>
<td>U3' (O5') (P2)</td>
<td>2.81 (77.23)</td>
<td></td>
<td>Arg35 (Nε2)</td>
<td>U-p (O5) (P1)</td>
<td>2.75 (69.23)</td>
<td></td>
<td>Phe120 (O)</td>
</tr>
<tr>
<td></td>
<td>Arg2 (Nε)</td>
<td>U3' (O5') (P2)</td>
<td>2.79 (61.57)</td>
<td></td>
<td>Arg121 (Nε)</td>
<td>U-p (O5) (P1)</td>
<td>2.83 (57.94)</td>
<td></td>
<td>Asn71 (Nδ2)</td>
</tr>
<tr>
<td>12</td>
<td>Arg2 (Nε)</td>
<td>U3' (O5') (P2)</td>
<td>2.80 (78.43)</td>
<td></td>
<td>Phe120 (O)</td>
<td>U-p (N3) (P1)</td>
<td>2.84 (60.34)</td>
<td></td>
<td>Asn71 (Nδ2)</td>
</tr>
<tr>
<td></td>
<td>Arg2 (Nε)</td>
<td>U3' (O5') (P2)</td>
<td>2.80 (78.00)</td>
<td></td>
<td>Arg121 (Nε)</td>
<td>U-p (O5) (P1)</td>
<td>2.81 (52.97)</td>
<td></td>
<td>Asn71 (Nδ2)</td>
</tr>
<tr>
<td></td>
<td>Arg2 (Nε)</td>
<td>U3' (O5') (P2)</td>
<td>2.83 (57.94)</td>
<td></td>
<td>Arg121 (Nε)</td>
<td>U-p (O5) (P1)</td>
<td>2.81 (52.97)</td>
<td></td>
<td>Asn71 (Nδ2)</td>
</tr>
<tr>
<td>13</td>
<td>Arg35 (Nε2)</td>
<td>U-p (O5) (P1)</td>
<td>2.78 (85.46)</td>
<td></td>
<td>Arg2 (Nε)</td>
<td>U3' (O5) (P2)</td>
<td>2.79 (53.63)</td>
<td></td>
<td>Asn71 (Nδ2)</td>
</tr>
<tr>
<td></td>
<td>Arg35 (Nε2)</td>
<td>U-p (O5) (P1)</td>
<td>2.82 (77.26)</td>
<td></td>
<td>Arg2 (Nε)</td>
<td>U3' (O5) (P2)</td>
<td>2.79 (53.63)</td>
<td></td>
<td>Asn71 (Nδ2)</td>
</tr>
</tbody>
</table>

RC: Reaction Coordinate
Table 4.2: Interaction of product nucleotides with hydrophobic amino acids. Contacts less than 4.5 Å with hydrophobic residues are listed. P1 - product 1; P2 - product 2; A5' - 5' adenine nucleotide; U3' - 3' uracil nucleotide; U-p - Uracil nucleotide with the cleaved phosphate group (see Figure 4.2)

<table>
<thead>
<tr>
<th>Reaction Coordinate (Å)</th>
<th>hRNase3 Products</th>
<th>hRNase5 Products</th>
<th>bRNaseA Products</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Trp11</td>
<td>U3' (P2)</td>
<td>Ile42</td>
</tr>
<tr>
<td></td>
<td>Phe12</td>
<td>A5' (P2)</td>
<td>Leu115</td>
</tr>
<tr>
<td></td>
<td>Leu130</td>
<td>A5' (P2)</td>
<td>U-p (P1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A5' (P2)</td>
</tr>
<tr>
<td>5</td>
<td>Met1</td>
<td>U3' (P2)</td>
<td>Ile42</td>
</tr>
<tr>
<td></td>
<td>Trp11</td>
<td>U3' (P2)</td>
<td>U-p (P1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Met1</td>
<td>U3' (P2)</td>
<td>Ile42</td>
</tr>
<tr>
<td></td>
<td>Trp11</td>
<td>A5' (P2)/U3' (P2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phe12</td>
<td>A5' (P2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Val128</td>
<td>A5' (P2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Met1</td>
<td>U3' (P2)/U-p (P1)</td>
<td>Ile42</td>
</tr>
<tr>
<td></td>
<td>Val128</td>
<td>A5' (P2)</td>
<td>Leu115</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Met1</td>
<td>U-p (P1)/U3' (P2)</td>
<td>Ile42</td>
</tr>
<tr>
<td></td>
<td>Leu69</td>
<td>A5' (P2)</td>
<td>Leu115</td>
</tr>
<tr>
<td></td>
<td>Val128</td>
<td>A5' (P2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Met1</td>
<td>U-p (P1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leu69</td>
<td>A5' (P2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Met1</td>
<td>U3' (P2)</td>
<td>Ile42</td>
</tr>
<tr>
<td></td>
<td>Leu69</td>
<td>A5' (P2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Met1</td>
<td>U3' (P2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 - 18</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.6: Conformational dynamics along the reaction coordinate. The conformational flexibility of the Cα atoms of each residue is obtained from the aggregate root mean square fluctuations from the top 10 modes (RMSF₁₀). The 2D plots represent the RMSF₁₀ as a function of enzyme residues. The important loop regions are highlighted in pink. RMSF₁₀ are also projected on cartoon representations of each enzyme, with flexible loop regions identified. The thickness of the cartoon tube corresponds to the amplitude of RMSF₁₀, and the color corresponds to the dynamic range observed: least dynamic (blue) to most dynamic (red).
4.4.D Conformational sub-states associated with product release

QAA was employed to obtain insights into the conformational sub-states associated with the release of products in the members of RNase superfamily. The original 3N-dimensional space for the C\textsubscript{\alpha} atoms of 52,500 conformations obtained from the umbrella sampling simulations along the reaction coordinate were projected onto 100-dimensional space using QHA modes for QAA. Using a mixture of gaussian modes, three clusters representing the three categories of reaction coordinate (mentioned above) were identified from QAA (Figure 4.7). Interestingly, these clusters identified using QAA contain homogenous populations of conformational sub-states. Motion involved in product release along the reaction coordinate are illustrated in wireframe, showing only the C\textsubscript{\alpha} trace of the enzyme residues, and colored according to their respective sub-states.

QAA applied to umbrella sampling simulations along the reaction coordinate of product release for the three RNases reveal the existence of distinct sub-states for these closely related enzymes. Moreover, the conformational sub-states also differ among the reaction coordinate for the three RNases studied here. In the modes, when the products are in near vicinity of the catalytic triad (blue conformations in Fig 4.7), large dynamical fluctuations are observed in loop L4 of hRNase3 and in loop L1 of bRNaseA. However, motions in these loops subside in other modes along the reaction coordinate for both hRNase3 and bRNaseA. There were no particular region showing motions with large change in amplitude in hRNase5.

4.4.E Potential of Mean Force (PMF)

The free energy profile of product(s) release along the reaction coordinate was calculated using WHAM and is represented in Figure 4.8. For all three enzymes studied here the free energy profile is different for both products. Energy difference between the release of two products is 2 folds for hRNase3 and 3 folds for hRNase5 and bRNaseA. Also, there is a difference of \sim 4 kcal/mol in the release of product 1 (Figures 4.1 and 4.2) from bRNaseA and hRNase5. The free energy for release of product 1 increases along the reaction coordinate (7 Å - 18 Å) for hRNase5, suggesting that additional energy is required to drive release of product 1 and the process is less likely to be spontaneous. On the other hand, there is a decrease in the free energy for release product 1 in bRNaseA, suggesting that product 1 suggesting that release of this product is more spontaneous in bRNaseA as compared to hRNase5. The free energy for release of product 2 decreases for both hRNase3 and hRNase5. However, this observation also contrary for bRNaseA where the free energy for release of product 2 increases (7 Å - 18 Å). These observations suggest that the
Figure 4.7: Conformational sub-states associated with release of products in ribonucleases. Higher order statistical clustering of conformations sampled in the molecular dynamics trajectories is shown in the box plots above. Each dot corresponds to a single conformation and is colored according to the reaction coordinate (see color bar). The ellipses represent conformational sub-states identified along the reaction coordinate. The black arrows indicate transitions from heterogenous population to a well-defined homogenous sub-state. $\gamma$ represents anharmonic modes identified by QAA and the subscripts denotes the rank of the QAA vector. Motions along the anharmonic mode ($\gamma$) are illustrated in the panels below in a wire frame representation, showing only the C$_\alpha$ trace of the protein and color coded according to the reaction coordinate.
Figure 4.8: Free energy profile of product release. The potential for mean force along the reaction coordinate was constructed using WHAM. The free energy profile for product 1 is shown in solid lines and for product 2 is shown in dotted lines. Results are shown for hRNase3 (red), hRNase5 (blue) and bRNaseA (black).

energetics of product(s) release is different for hRNases when compared to bRNaseA. Note, when the product(s) are in the bulk solvent (> 14 Å) away from the catalytic residues, they are free to diffuse in the simulation box. However due to limited size of the simulation box the product nucleotides stay close to the enzyme surface. This can certainly affect the shape of the PMF.

It is reported that a decrease of 1.37 kcal/mol in activation energy results in 10 fold improvement in the rate constant. The free energy of product release for hRNase5 is 4.5 kcal/mol more than bRNaseA (Figure 4.8). Suggesting a ~10^4 folds difference in the activation energy between these two enzymes. This observation corresponds with the 10^5 folds difference in catalytic efficiency between bRNaseA and hRNase5, with bRNaseA being the faster enzyme. However, whether this difference in the catalytic efficiency is due to difference in activation energy required for product release step needs further experimental validation.
4.5 Discussion

Enzymes are not rigid structures but sample an ensemble of conformations driven by internal motions occurring over a wide range of length and time-scales. The role of these conformational fluctuations in enzyme catalysis has been recognized, and in some cases even considered vital for the designated function.\textsuperscript{324, 325} Experimental characterization of conformations sampled by enzymes provide information only on a narrow time-scale window, with low a probability of finding functionally relevant conformational sub-states.\textsuperscript{326} Computer simulations, enhanced sampling techniques and higher order statistical methods have enabled the identification of functionally important conformational fluctuations and sub-states over a wide range of time-scales.\textsuperscript{129}

In this study, we observed that there were strong non-bonded interactions between the products and the residues lining the active site. When the product nucleotides are in the near vicinity of the catalytic triad (4 - 8 Å), the enzyme residues of all three RNases (Tables 4.1 and 4.2) strongly interacts with the cleaved products. As the product nucleotides move 9 - 13 Å away from the catalytic triad, the number of residues interacting with the product nucleotides decrease substantially. Further down the reaction coordinate (beyond 13 Å), the product nucleotide overcomes these strong interactions to completely release themselves into the bulk solvent. Many of these residues have been previously identified as part of networks of amino acid residues that show conformational exchange upon binding of single nucleotides in these RNases.\textsuperscript{146, 327} However, mutagenesis studies and further experiments are necessary to gain insights into the individual contribution of these residues to the catalytic turnover of select RNases.

In hRNase3, the product nucleotides interact with the residues on the surface of enzyme, away from the catalytic triad, suggesting possible interaction of the product nucleotide with other regions of the enzyme before they are completely dissociated and released in the bulk solvent. However, such interactions in distal regions are not observed for bRNaseA, suggesting that distal residues on the surface of this faster enzymes do not interact with the cleaved products post catalysis.

This study characterizes the internal dynamics in the product release step of three RNases. The RMSF\textsubscript{10} profiles indicate that the most significant differences in dynamics occur in the loop regions for these RNases (Figure 4.6). This further implies that loop regions, which are energetically more favorable to movements within a protein structure, play a major role in catalysis.\textsuperscript{328} However, for each RNases studied here different loops show fluctuations with higher amplitude, indicating a
marked difference in the conformational dynamics of product release in RNases. Long range effects in RNases have also been identified to modulate ligand binding, and these effects were also distinct for individual RNases. As seen in chapter 2, the three RNases studied here belong to different phylogenetic sub-family with unique function, distinctness in their conformational motions suggests that the even closely related RNases are uniquely adapted to their respective functions. These observations are also consistent with previous studies which show that enzymes/proteins sharing common fold can also have diverse functions and dynamics.

This study further characterizes the conformational sub-states and motions associated with the product release step in RNase family members, using atomistic umbrella sampling molecular dynamics simulations and QAA. QAA has been used to study the conformational sub-states and conformational fluctuation for a number of protein systems and the motions identified by QAA are intrinsic property of each protein studied. The results from this study indicate the presence of separate conformational sub-states in the three RNases studies here (Figure 4.7). In these sub-states, loops L4 of hRNase3 and L1 of bRNaseA, show fluctuations with high amplitude, suggesting a potential role of these loops in product release. It is also interesting to note that the conformational dynamics in product release are significantly different for members of the individual phylogenetic sub-family. Further, conformational exchange probed by $^{15}$N-CPMG relaxation dispersion experiments on the micro-millisecond time-scale were compared to the conformational fluctuations observed in the product release step from QAA (Figure 4.9). The residues of the functionally important loops, L4 in hRNase3 and L1 in bRNaseA, that show conformational fluctuations in product release simulations also undergo conformational exchange in $^{15}$N-CPMG experiments of apo protein. This observation supports the suggestion that free protein can samples conformations that resemble the ligand bound conformations.

Overall this study identified structural and dynamical features important product release in RNase family members. Modulating the dynamical fluctuations identified here, can help fine-tuning the conformational landscape of these enzymes to favor efficient catalysis. Subsequent experimental effort will be required to bring home the points raised by this computational study and to prove that modulating internal dynamics in these enzymes can greatly enhance their catalytic efficiencies. Apart from modulating enzyme function, insights gained form this work will have implications in the design of novel inhibitors and modulators of RNase function in the context of neurotoxicity, angiogenesis and anti-pathogenicity. This study further demonstrated that the
umbrella sampling simulations combined with QAA technique can generate reliable results for the study of product release in proteins at the atomistic level and can therefore be applied to study other enzymatic systems.
Figure 4.9: Comparison of computational conformational sub-states and NMR conformational exchange. Conformations sampled along reaction coordinate allowed identification of conformational sub-states (top panel) and transitions into these states (wire frames) as products move from active sites into bulk solvent. Residues that experience ms conformational exchange in NMR $^{15}$N-CPMG experiments are shown as spheres. Results are shown for hRNase3, hRNase5 in comparison with bRNaseA.
Chapter 5  Specialized Techniques to Identify Conformational Fluctuations Associated with Enzyme Catalysis
Part of this chapter is a slightly revised version of a manuscript by title “Ligand-Induced Variations in Structural and Dynamical Properties Within an Enzyme Superfamily” published in the journal, Frontiers in Molecular Biosciences in June 2018

Reprinted (adapted) with permission from Chitra Narayanan,1 David N. Bernard,1 Khushboo Bafna,2 Donald Gagné,1 Pratul K. Agarwal,3,4 and Nicolas Doucet1,5,*

1INRS-Institut Armand-Frappier, Université du Québec, 531 Boul. des Prairies, Laval, Québec H7V 1B7, Canada.
2Genome Science and Technology, University of Tennessee, Knoxville, TN 37996, USA.
3Computational Biology Institute and Computer Science and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States
4Department of Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee, 37996, USA.
5PROTEO, the Quebec Network for Research on Protein Function, Structure, and Engineering, 1045 Avenue de la Médecine, Université Laval, Québec, QC, G1V 0A6, Canada.

DOI : 10.3389/fmolb.2018.00054

*Corresponding authors.

This author contributed the following to the manuscript: (1) Calculating compounded chemical shifts (2) Chemical shift projection analysis, calculating projection angle and fractional shift and (3) data analysis
5.1 Abstract

It is now widely accepted that enzymes are dynamic entities and conformational fluctuations in enzymes is essential for efficient catalysis. Whether these conformational fluctuations affect one/all step of the enzymes’ catalytic cycle is still a point of debate. By combining NMR chemical shift titration experiments with the chemical shift projection analysis (CHESPA) this study characterizes the dynamics in ribonucleases (RNases) during ligand binding. Results indicate that members of the hRNase family display discrete chemical shift perturbations upon ligand binding. The amino acid networks exhibiting long range dynamics upon ligand binding are unique to each of the RNases analyzed. In addition, steady-state kinetic parameters for RNases was obtained. The results reveal that catalytic efficiency in the human homologs differs by a more than $10^5$ folds. Further using hybrid quantum classical simulations (QM/MM) and QAA, conformational sub-states and dynamics associated with the chemical (RNA cleavage) step of RNase function are identified. Interestingly the dynamics in bRNaseA during the chemical step were considerably reduced and the conformational sub-states did not show high amplitude motions even for functionally important loops. These observations suggest that high amplitude motions in RNase function facilitate the formation of transition state species while the actual chemical step is a result of reaction-ready active site configuration.

5.2 Introduction

Enzymes are biological macromolecules that enhance the rate of reaction in a cell by $>10^{20}$ folds. Catalytic cycle of enzymes involves several steps, including substrate recognition and binding to the active-site, chemical step involving the conversion of substrate to product(s), and product(s) release. Any of these steps could be the rate-limiting step affecting the overall rate of catalytic turnover in enzyme function. An enzyme progresses through these steps in its catalytic cycle by the virtue of conformational fluctuation occurring over a wide range of time-scales. These conformational fluctuations affect structural rearrangement of the ligand binding site allowing the reactive partners to orient themselves for the catalytic process. Evidence from computational and experimental approaches have revealed the existence of correlation between rate of catalytic turnover and time-scale of conformational flexibility in a number of enzyme systems including bovine ribonuclease A (bRNaseA).
Members of the RNase enzymes family cleave 3’- 5’ phosphodiester bond in ribonucleic acid (RNA) substrates. In humans, eight functionally important RNases (hRNases 1-8) have been identified. All RNases have a conserved kidney shaped fold, catalytic residues (2 histidine and one lysine) and possess two-nucleotide base binding sub-sites that interact with the substrate molecules. A pyrimidine base binds in the B1 sub-site and a purine base binds in the B2 sub-site of RNases. However, they display a wide range of substrate specificities and the catalytic efficiency in these enzymes differ by a million fold. In addition, recent studies report the existence of diverse biological function in the human members. In humans, RNases have evolved to perform other biological functions such as host defense, immunosuppressivity, angiogenesis, and antipathogenic activity. The detailed understanding of factors that contribute to the divergence of function in these structurally related enzymes remains limited. The existence of distinct dynamic behavior within hRNases (Chapter 2, 3 and 4) suggest a significant role of dynamics in facilitating their diverse catalytic efficiencies and biological functions.

Efforts have been made to relate dynamics with ribonucleolytic function in bRNaseA. NMR experiments reveal the existence of ms time-scale conformational motions in RNases which also corresponds to the time-scale of catalytic turnover reported for bRNaseA. It was further uncovered that conformational motions in bRNaseA corresponds to the product release step. Conformational exchange during product release was observed in residues of distal loop. Mutating these distal residues not only reduced the global motions in bRNaseA but also decreased the rate of product release by 10 folds. These evidences suggest a possible role of conformational fluctuations in function of bRNaseA, particularly the product release step. Whether the motions in distal regions also affect the chemical step of bRNaseA needs to be examined. Moreover, the effect of dynamics in ligand binding and chemical step of hRNases has not been characterized in such details. Therefore, several areas where dynamics could potentially play a promoting role in the function of hRNases and bRNaseA remain unexplored. This study aims to compare and characterize the structural and conformational changes associated with ligand binding for select hRNases. Further, this study seeks to address if RNases have specific motions corresponding to the chemical step in their enzymatic cycle and identify whether these motions are localized around the active site or occur in the distal regions of the enzyme. And finally, establish whether motions in ligand binding or catalytic step affect catalytic turnover.
The rate of catalytic turnover in these structurally related RNases varies by a million-fold. However, these studies were performed by different groups under different experimental conditions using a wide variety of substrates ranging from dinucleotide RNA substrates to tRNA of unknown sequence and lengths. The diversity in the previously used method for characterizing kinetic rates for RNases rather makes it difficult to comparatively analyze the difference in the efficiency of these enzyme. For quantitative and relevant comparisons, it is essential to remove the errors and variations that could be imparted due to differences in experimental conditions. And therefore, this study systematically characterizes the kinetic rates in hRNases and bRNaseA under same experimental conditions and with same dinucleotide substrate (UpA).

In this study, three aspects of RNase function were investigated: Kinetic parameters for various hRNases were obtained using steady-state kinetics with same substrate (UpA) under identical conditions; Conformational dynamics induced by ligand binding in the members of the RNase family were probed by NMR titrations and chemical shift projection analysis (CHESPA); Conformational dynamics and sub-states associated with phosphor-diester bond cleavage (the chemical step) were identified using hybrid QM/MM modeling of the reaction pathway and quasi anharmonic analysis (QAA).

5.3 Materials and methods

**Michaelis-Menten steady-state kinetics:** The rate of enzymatic reaction is measures over a range of substrate concentration ([S]). The initial rate $V_0$ increases with the increase in [S]. However, as [S] gets higher, the enzyme becomes saturated with substrate and the initial rate reaches $V_{max}$, the enzyme’s maximum rate. At $V_{max}$, increase in substrate concentration does not cause any increase in reaction rate as there is no more enzyme (E) available for reacting with substrate (S). Using the steady-state method kinetic parameters are obtained from the Michaelis-Menten equation described below (Eq 5.1). However, the Michaelis-Menten equation is based on two assumptions; first, the concentration of the substrate-bound enzyme changes much more slowly than those of the product and substrate and thus the change over time of the complex can be set to zero. And second, the total enzyme concentration does not change over time.
Michaelis-Menten equation:
\[ v = \frac{V_{\text{max}} [S]}{K_M + [S]} \]  

Where
\[ V_{\text{max}} = k_{\text{cat}} [E] \]

\( K_M \) is the Michaelis constant, experimentally defined as the concentration at which the rate of the enzyme reaction is half \( V_{\text{max}} \). \( k_{\text{cat}} \) denotes the maximum number of enzymatic reactions catalyzed per second also called turnover number.

Steady-state kinetic parameters (\( k_{\text{cat}} \) and \( K_M \)) were obtained on a Perkin Elmer Lambda 35 UV-vis spectrophotometer. Activity of RNases was measured in 50 mM sodium acetate, 1 mM EDTA at pH 5.5 and 25°C. Dinucleotide substrate UpA was obtained from Dharmacon at GE Healthcare. The concentration of UpA was measured using the extinction coefficient 24600 M\(^{-1}\)cm\(^{-1}\) supplied from Dharmacon. Change in UpA absorbance was measured in 1 or 2 mm path length cuvettes at 286 nm for 10 mins. Enzyme and substrate concentrations used for the study are indicated in Table 5.1, and all measurements were performed in duplicate. Water was incubated with 1% DEPC for 1 hr and autoclaved. All equipments were cleaned with DEPC-treated water before starting the experiment to inhibit the environmental contaminant RNases from interfering with the kinetic assay. Buffer was made with DEPC treated water as well. The data were fitted to Michaelis-Menten equation using sigma plot.

Table 5.1: Enzyme concentrations and range of substrate (UpA) concentrations used in steady-state kinetics experiments.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Enzyme concentration (µM)</th>
<th>Range of UpA concentration (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bRNaseA</td>
<td>0.002</td>
<td>0.1 – 1.5</td>
</tr>
<tr>
<td>hRNase2</td>
<td>0.05</td>
<td>0.0 – 3.0</td>
</tr>
<tr>
<td>hRNase3</td>
<td>1.0</td>
<td>0.5 – 4.2</td>
</tr>
<tr>
<td>hRNase4</td>
<td>0.05</td>
<td>0.3 – 2.1</td>
</tr>
<tr>
<td>hRNase5</td>
<td>9.0</td>
<td>0.8 – 4.0</td>
</tr>
<tr>
<td>hRNase6</td>
<td>0.1</td>
<td>0.1 – 1.6</td>
</tr>
</tbody>
</table>
NMR chemical shift titrations: NMR chemical shifts of atoms are sensitive to changes in their chemical environment and are therefore used to detect structural and conformational changes associated with ligand binding. In this study, NMR chemical shift titration experiments were performed to compare and characterize the effect ligand binding (two ligands, 3′-UMP and 5′-AMP) on the conformational properties of the five RNases, hRNase2, hRNase3, hRNase4, hRNase5 and bRNaseA. All RNases were expressed and purified in our collaborator Dr. Nicolas Doucet’s group using the protocols described in 138, 142, 146, 327. 15N-carr-purcell-meiboom-gill (CPMG) NMR relaxation experiments and NMR titration experiments were also performed in Dr. Nicolas Doucet’s group as described in 327.

Chemical shift perturbations (Δδ_{obs}): were calculated as the difference in the weighted average chemical shift of the ligand-bound (5′-AMP and 3′-UMP) and apo states, using the equation below.

$$\Delta \delta_{obs} = \sqrt{\frac{(\Delta \delta_H)^2 + (0.2 \Delta \delta_N)^2}{2}}$$

Chemical shift projection analysis (CHESPA): CHESPA was performed based on the protocol described by Selvaratnam et al. 336 The chemical shift perturbations (Δδ_{obs}) of the ligand-bound state relative to the apo state corresponds to the shifts for the highest enzyme: ligand molar ratios for each enzyme. For all enzymes studied, residues with a chemical shift variation Δδ_{obs} > 0.05 ppm were selected for further analysis. The two CHESPA parameters, projection angle (cos(θ)) and fractional shift (X), were calculated according to the equations (5.4 and 5.5 respectively) below using the 1H and 15N peak coordinates in their free (apo) form and upon binding to the two mononucleotide ligands, 3′-UMP (vector A) and 5′-AMP (vector B) at saturation conditions (Figure 5.1).

$$\cos(\theta) = \frac{A \cdot B}{|A||B|}$$

$$X = \frac{A \cdot B}{|B|^2}$$

The projection angle (θ) is defined as the angle between the chemical shift peak displacements (for residues with Δδ_{obs} > 0.05 ppm at the highest ligand concentration) upon binding 3′-UMP (A
vector) and 5′-AMP (B vector) relative to the apo state. The fractional shift (X) corresponds to the magnitude of displacement of A relative to B.

Residues that show a gain (or loss) of millisecond exchanges in both the 3′-UMP- and 5′-AMP-bound states relative to the apo state are defined as displaying coordinated changes in motions (Figure 5.2 A). Residues displaying a gain (or loss) of dynamics in only one of the two (5′-AMP or 3′-UMP) ligand-bound states relative to the apo state are defined as displaying uncoordinated changes (Figure 5.2 B).

**Reaction pathway sampling using Hybrid quantum mechanics/molecular mechanics (QM/MM) and nudged elastic band (NEB) method:** RNA cleavage step for bRNaseA was modeled using hybrid QM/MM and NEB method in AMBER. In NEB method, the pathway for conformational change between two end points is calculated using a series of images (called as nodes) of the molecule/system. Each image between the end points is connected to the previous and next image by springs that force each node to remain at an average separation between its partner images along the current path preventing the nodes from sliding onto each other. Minimization of the nodes between the end points provides a minimum energy pathway. Further optimization using simulated annealing estimates the global minimum energy pathway. The result is a minimum potential energy path that represents the conformational change between the two endpoint structures, independent of time-scale. The NEB method requires the users to define the start and end points of the reaction pathways to be studied. Also, the pathway endpoints are fixed during the NEB simulation in AMBER. Therefore, the reactant state (RS) and product state (PS) were modeled and extensively simulated to obtain a desirable start and end points for the QM/MM-NEB simulations.

**Structure preparation of RS and PS:** The catalytic mechanism of phosphodiester cleavage has been well studied for bRNaseA. Based on these studies, the reactant and product states for bRNaseA were modeled (Figure 5.3). Since there are no published X-ray structures of bRNaseA with substrate, tetranucleotide substrate was modeled in the active site of bRNaseA as mention in chapter 3. The bRNaseA enzyme with cleaved product(s) was modeled as mention in chapter 4. The RS and PS complexes were solvated in a SPC/E water box with periodic boundary conditions and neutralized with the addition of chloride (Cl⁻) counterions. The systems were equilibrated using the previously described protocol. PME-method was implemented for long range electrostatics. All hydrogen bonds were constrained using SHAKE.
Figure 5.1: Illustration of the CHESPA approach. The 1H-15N peak is represented for apo (gray), and bound forms 5'-AMP (blue) and 3'-UMP (orange), respectively. Arrows indicate the movement of the 1H-15N chemical shift for each peak from its apo (in red) to its saturated position (blue and orange). The compounded chemical shift upon ligand binding is calculated as the magnitude of vectors A and B, respectively (see text for details). θ represents the angle between vector A and B. Figure adapted from 340

Figure 5.2: The NMR chemical shift projection analysis (CHESPA). Graphical representation of scenarios representing (A) coordinated and (B) uncoordinated displacements of the chemical shifts upon ligand binding. Figure adapted from 340
Production runs of 0.5 µs (microsecond) were performed under constant volume and energy (NVE) conditions using a 2 femtoseconds (fs) time-step. Simulations to obtain starting conformation of RS and PS systems were performed using NVIDIA graphics cards with CUDA-enabled version of AMBER’s pmemd MD simulation engine. The default precision in the pmemd.cuda_SPFP binary was used from AMBER v14 package. The final structures of the production trajectories were used as input for QM/MM optimization.

**Structure preparation of RS and PS for QM/MM optimization:** A QM region and a MM region were identified for RS and PS. The QM region includes 47 atoms selected from the side chains of three catalytic residues (two histidine and one lysine residues), ribose sugar rings of the two nucleotides next to the cleaved phosphate group, and three water molecules (Figure 5.4). The remaining parts of the tetra-nucleotide substrate, enzyme, counterions and the water box were modeled as classical region and treated by molecular mechanics (AMBER’s ff14SB force field). A total of 38162 atoms constitute the MM region. QM/MM optimization of RS and PS was performed for 2000 steps using SCC-DFTB level of theory for QM atoms. The SCC-DFTB level...
of calculation was chosen because of its manageable computational complexity. QM compatible PME approach was used for calculating the long-range QM-MM electrostatic and forces. The optimized QM/MM structures of the RS and PS were used as endpoints for reaction pathway sampling using QM/MM-NEB method.

**Reaction pathway sampling using QM/MM - NEB method:** NEB methods attempt to find the lowest pathway on the potential energy surface between the two endpoints provided. The NEB protocol created six additional replicates (making a total of 8 nodes including the two endpoints provided) that will stretch along our pathway and be joined by springs. First all eight nodes were heated from 0 – 300 K in 20,000 steps with a 0.1 fs timestep. During the heating process hydrogen atoms were not constrained by SHAKE. Relatively small spring constant of 10 kcal/mol/Å² was used. Next simulated annealing was performed with the temperature profile increasing from 300 K to 500 K and back to 300 K. A total of 150 ps MD was performed in this step with a time-step of 0.25 fs. In the next step the system was cooled down, so that that images freeze out along the low energy pathway. The nodes were cooled in 2 stages. At first stage the nodes were gradually cooled from 300 K to 0 K step over 6 ps. Finally, a long cooling was performed at 0 K for 10 ps with a time-step of 0.1 fs. Conformations were saved after every 10 steps per NEB node. Final QM/MM - NEB production run was performed for 200,000 steps at the time step of 1 fs for all 8 NEB nodes. Coordinates were saved after every 2 ps.

**Intrinsic dynamics:** The intrinsic dynamics in bRNaseA during the chemical step was deduced from the root mean square fluctuations of the Cα atoms of the top 10 QHA modes (RMSF_{10}). Fluctuations were computed using the ptraj analysis module in AMBER. Conformations from only the final production trajectory of all 8 NEB nodes were used for RMSF_{10} calculations. For comparison RMSF_{10} results for apo, substrate bound and product release simulations from chapters 2, 3 and 4 respectively were used.

**Conformational sub-states using quasi anharmonic analysis (QAA):** The 80000 conformations (10000 conformations per NEB node) obtained from the hybrid QM/MM and NEB simulations were used for QAA. The original 3N dimensional space for the Cα atoms was projected onto 60-dimensional space using QHA modes and then analyzed using QAA. Using a mixture of Gaussian modes three clusters representing distinct NEB nodes were identified. The three QAA independent component vectors were analyzed in detail to characterize the conformational sub-states associated with the chemical step in bRNaseA.
Figure 5.4: Reactive center for reactant and product states. The side chain of catalytic residues and the sugar-phosphate backbone that form the reactive center for the RS and PS are shown. Atoms included in the QM region (blue) and MM region (black) are identified. Note that water molecules that form the part of the QM region are not shown to increase the clarity of the Figure (See Figure 5.3 to identify the water molecules (red spheres)).
5.4 Results

5.4.A Elucidating the differences in kinetics parameters for human ribonucleases

Kinetic parameters (\(k_{\text{cat}}\) and \(K_M\)) for cleavage of dinucleotide substrate UpA by hRNases and bRNaseA were obtained using Michaelis-Menten steady-state kinetics at pH 5.5 and 25 °C. To facilitate comparison between the different enzymes the enzymatic activity of the hRNases was quantified using the same substrate UpA. The kinetic parameters of bRNaseA was quantified for reference. The results indicate that all RNases were able to cleave UpA, as expected for the canonical members of this superfamily. Most RNases exhibited \(K_M\)s above 1 mM, with bRNaseA, hRNase2, hRNase4, hRNase6 showing better affinities (Table 5.2). hRNase2 showed the highest \(k_{\text{cat}}\) (~88 s\(^{-1}\)) among hRNases and a catalytic turnover ~ 6 times lower than bRNaseA. The catalytic efficiency \((k_{\text{cat}}/K_M)\) of hRNase3 was three to four orders of magnitude lower than that of the bRNaseA. Approximately 10\(^3\)-fold difference was observed in the \(k_{\text{cat}}\) values, between the fastest (bRNaseA) and the slowest (hRNase5) enzyme. Together, these results indicate a nearly 10\(^5\)-fold difference in catalytic efficiency \((k_{\text{cat}}/K_M)\) between hRNases, further exemplifying the distinct roles of the ribonucleolytic activity in their unique and varied biological functions.

Table 5.2: Comparing the steady-state kinetic parameters for RNases. The parameters were obtained at pH 5.5, 25 °C substrate concentration was varied, and the data were fitted to Michaelis-Menten equation.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>(K_M) (µM)</th>
<th>(k_{\text{cat}}) (s(^{-1}))</th>
<th>(k_{\text{cat}}/K_M) (M(^{-1})s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>bRNaseA</td>
<td>120 ± 40</td>
<td>450 ± 30</td>
<td>3.8 ± 0.7 (\times) 10(^6)</td>
</tr>
<tr>
<td>hRNase2</td>
<td>520 ± 50</td>
<td>88 + 3</td>
<td>1.7 ± 0.2 (\times) 10(^5)</td>
</tr>
<tr>
<td>hRNase3*</td>
<td></td>
<td></td>
<td>6.7 ± 1.6 (\times) 10(^2)</td>
</tr>
<tr>
<td>hRNase4</td>
<td>440 ±160</td>
<td>16 ± 2</td>
<td>3.7 ± 0.6 (\times) 10(^4)</td>
</tr>
<tr>
<td>hRNase5</td>
<td>2500 ± 880</td>
<td>0.031 ± 0.006</td>
<td>12 ± 2</td>
</tr>
<tr>
<td>hRNase6</td>
<td>140 ± 10</td>
<td>5.7 ± 0.1</td>
<td>4.1 ± 1.3 (\times) 10(^4)</td>
</tr>
</tbody>
</table>

* saturation was not obtained
5.4.B Ligand-induced variations in structural and dynamical properties of RNases

The effect of binding of two mononucleotide ligands 3'-UMP and 5'-AMP (also product analogs), on conformational properties of RNases was probed using NMR shift titrations. The compounded chemical shift changes $\Delta \delta_{\text{obs}}$ (at the highest ligand concentration) upon binding of 3'-UMP and 5'-AMP for the five RNases studied (hRNase2, hRNase3, hRNase4, hRNase5, bRNaseA) are shown in Figure 5.5. Residues displaying $\Delta \delta_{\text{obs}} > 0.1$ ppm for all RNases are reported in Table 5.3. Notable differences were observed in the number of residues affected upon binding of 3'-UMP to the functionally distinct RNases. In bRNaseA, 21 residues displayed $\Delta \delta_{\text{obs}} > 0.1$ ppm and largest chemical shift perturbations were observed near the residues Thr45, His119 and Lys41 primarily the B₁ pyrimidine binding sub-site. Additional large perturbations ($\Delta \delta_{\text{obs}} > 0.1$ ppm) were observed for some residues far from the active site, suggesting long-range conformational changes or motions in these regions upon 3'-UMP binding. Similar to bRNaseA, in hRNase2 large perturbations were observed in residues near the B₁ sub-site, as well as residues of distal loops (Ser64 and Lys66 of L4 and Cys37 of L2). However, the total number of residues that displayed $\Delta \delta_{\text{obs}} > 0.1$ ppm were less when compared to bRNaseA. Further, in hRNase4 only few residues, primarily localized to the pyrimidine binding B₁ sub-site, displayed chemical shift perturbations upon ligand binding. This observation is significantly different from its behavior in apo state, where most residues undergo conformational exchange. On the contrary, in hRNase3 perturbations were dispersed throughout the protein with large chemical shift variations in loops L4 (positions 67–76) and L5 (positions 80–83). These loops however were minimally perturbed in the other RNases upon binding of 3'-UMP. In contrast to bRNaseA, the catalytic residues of hRNase3 (Lys39 and His128) did not display significant perturbations upon 3'-UMP binding. Similarly, in hRNase5, residues in the B₁ sub-site did show perturbations while residues far from the active site were affected.

Titration of ligand 5'-AMP resulted in perturbations of residues near the purine binding sub-site (B₂) in all RNases, except hRNase5. In, bRNaseA, hRNase2 and hRNase4 residues of loop L4 displayed significant perturbations. Loop L4 has previously been reported to interact with the purine base. Additional perturbations are observed in N-terminal alpha helix in hRNase2 and C-terminal regions of hRNase4. In contrast perturbations in L4 were diminished in hRNase3 upon binding of 5'-AMP and even fewer residues displayed perturbations relative to 3'-UMP binding. In hRNase5, no changes were observed in loop L4 and overall very few residues displayed large
Figure 5.5: Effect of ligand binding on functionally distinct RNases. Compounded chemical shift perturbations ($\Delta \delta_{\text{obs}}$) relative to the apo form upon binding of two mononucleotides 3'-UMP and 5'-AMP are plotted as function of consensus sequence. Chemical shift perturbations were calculated by comparing the shifts at the largest enzyme:ligand molar ratios relative to the apo state for each enzyme. $\Delta \delta_{\text{obs}}$ are depicted on the 3D structures using the putty representation to the right of the plots, the width of the tube corresponds to larger $\Delta \delta_{\text{obs}}$. The 3'-UMP and 5'-AMP ligands are shown in the structures of bRNaseA only, they bind at same site in other RNases. Catalytic residues (His12, Lys41, His119, bRNaseA numbering) are highlighted using gray lines. Secondary structure of bRNaseA is shown on top of the graphs. The consensus sequence (obtained including gaps in the multiple sequence alignment) is used for easier comparison of proteins.
Table 5.3: Chemical shift perturbations upon ligand binding. Residues displaying $\Delta \delta_{\text{obs}} \geq 0.1$ ppm upon 3'-UMP and 5'-AMP binding to RNases is reported.

<table>
<thead>
<tr>
<th>Residues displaying $\Delta \delta_{\text{obs}} \geq 0.1$ ppm upon 3'-UMP binding</th>
<th>bRNaseA</th>
<th>hRNase2</th>
<th>hRNase3</th>
<th>hRNase4</th>
<th>hRNase5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residue</td>
<td>$\Delta \delta_{\text{obs}}$</td>
<td>Residue</td>
<td>$\Delta \delta_{\text{obs}}$</td>
<td>Residue</td>
<td>$\Delta \delta_{\text{obs}}$</td>
</tr>
<tr>
<td>11</td>
<td>0.145</td>
<td>7</td>
<td>0.147</td>
<td>4</td>
<td>0.281</td>
</tr>
<tr>
<td>12</td>
<td>0.137</td>
<td>11</td>
<td>0.241</td>
<td>11</td>
<td>0.134</td>
</tr>
<tr>
<td>17</td>
<td>0.149</td>
<td>15</td>
<td>0.195</td>
<td>14</td>
<td>0.194</td>
</tr>
<tr>
<td>41</td>
<td>0.223</td>
<td>16</td>
<td>0.102</td>
<td>16</td>
<td>0.162</td>
</tr>
<tr>
<td>44</td>
<td>0.490</td>
<td>37</td>
<td>0.229</td>
<td>31</td>
<td>0.129</td>
</tr>
<tr>
<td>45</td>
<td>1.197</td>
<td>42</td>
<td>0.213</td>
<td>40</td>
<td>0.335</td>
</tr>
<tr>
<td>46</td>
<td>0.258</td>
<td>43</td>
<td>0.247</td>
<td>48</td>
<td>0.118</td>
</tr>
<tr>
<td>47</td>
<td>0.110</td>
<td>64</td>
<td>0.145</td>
<td>55</td>
<td>0.109</td>
</tr>
<tr>
<td>51</td>
<td>0.180</td>
<td>66</td>
<td>0.101</td>
<td>57</td>
<td>0.105</td>
</tr>
<tr>
<td>66</td>
<td>0.152</td>
<td>82</td>
<td>0.162</td>
<td>62</td>
<td>0.123</td>
</tr>
<tr>
<td>82</td>
<td>0.163</td>
<td>83</td>
<td>0.209</td>
<td>64</td>
<td>0.302</td>
</tr>
<tr>
<td>96</td>
<td>0.149</td>
<td>103</td>
<td>0.108</td>
<td>66</td>
<td>0.238</td>
</tr>
<tr>
<td>97</td>
<td>0.108</td>
<td>129</td>
<td>0.249</td>
<td>69</td>
<td>0.113</td>
</tr>
<tr>
<td>99</td>
<td>0.101</td>
<td>130</td>
<td>0.477</td>
<td>75</td>
<td>0.279</td>
</tr>
<tr>
<td>100</td>
<td>0.130</td>
<td>131</td>
<td>0.361</td>
<td>77</td>
<td>0.189</td>
</tr>
<tr>
<td>101</td>
<td>0.187</td>
<td>84</td>
<td>0.346</td>
<td></td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>0.107</td>
<td>129</td>
<td>0.568</td>
<td></td>
<td></td>
</tr>
<tr>
<td>119</td>
<td>0.225</td>
<td>130</td>
<td>0.136</td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>0.403</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>122</td>
<td>0.206</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residues displaying $\Delta \delta_{\text{obs}} \geq 0.1$ ppm upon 5'-AMP binding</th>
<th>bRNaseA</th>
<th>hRNase2</th>
<th>hRNase3</th>
<th>hRNase4</th>
<th>hRNase5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residue</td>
<td>$\Delta \delta_{\text{obs}}$</td>
<td>Residue</td>
<td>$\Delta \delta_{\text{obs}}$</td>
<td>Residue</td>
<td>$\Delta \delta_{\text{obs}}$</td>
</tr>
<tr>
<td>11</td>
<td>0.166</td>
<td>7</td>
<td>0.101</td>
<td>10</td>
<td>0.279</td>
</tr>
<tr>
<td>43</td>
<td>0.107</td>
<td>8</td>
<td>0.208</td>
<td>41</td>
<td>0.112</td>
</tr>
<tr>
<td>45</td>
<td>0.227</td>
<td>11</td>
<td>0.244</td>
<td>64</td>
<td>0.118</td>
</tr>
<tr>
<td>46</td>
<td>0.133</td>
<td>14</td>
<td>0.250</td>
<td>110</td>
<td>0.132</td>
</tr>
<tr>
<td>51</td>
<td>0.239</td>
<td>15</td>
<td>0.164</td>
<td>128</td>
<td>0.235</td>
</tr>
<tr>
<td>58</td>
<td>0.249</td>
<td>39</td>
<td>0.109</td>
<td>129</td>
<td>0.169</td>
</tr>
<tr>
<td>64</td>
<td>0.134</td>
<td>64</td>
<td>0.211</td>
<td>130</td>
<td>0.194</td>
</tr>
<tr>
<td>65</td>
<td>0.450</td>
<td>66</td>
<td>0.101</td>
<td>70</td>
<td>0.201</td>
</tr>
<tr>
<td>66</td>
<td>0.193</td>
<td>82</td>
<td>0.100</td>
<td>71</td>
<td>0.157</td>
</tr>
<tr>
<td>68</td>
<td>0.139</td>
<td>83</td>
<td>0.105</td>
<td>115</td>
<td>0.110</td>
</tr>
<tr>
<td>69</td>
<td>0.145</td>
<td>95</td>
<td>0.190</td>
<td>116</td>
<td>0.137</td>
</tr>
<tr>
<td>71</td>
<td>0.190</td>
<td>128</td>
<td>0.236</td>
<td>117</td>
<td>0.625</td>
</tr>
<tr>
<td>72</td>
<td>0.194</td>
<td>129</td>
<td>0.460</td>
<td>118</td>
<td>0.152</td>
</tr>
<tr>
<td>101</td>
<td>0.116</td>
<td>130</td>
<td>0.332</td>
<td>119</td>
<td>0.208</td>
</tr>
<tr>
<td>112</td>
<td>0.104</td>
<td>131</td>
<td>0.464</td>
<td></td>
<td></td>
</tr>
<tr>
<td>118</td>
<td>0.149</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>119</td>
<td>0.259</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>0.325</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>121</td>
<td>0.107</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
chemical shift perturbations, suggesting few interactions with 5′-AMP. Overall, these observations indicate that closely related RNases display distinctly different conformational rearrangements of local chemical environment upon binding of the two single-nucleotide RNA product mimics. These observations further confirm the difference in affinity of these RNase for a common ligand as seen in Chapter 2. Additional dynamics in distal regions is different for these enzymes even though they share the same fold, consistent with our observation of distinct conformational sub-states in product release (Chapter 4).

The structural rearrangements and local/global conformational motions in proteins that occur upon ligand binding can be identified by probing the changes in chemical shift perturbations. In this study, the chemical shift projection analysis (CHESPA) is employed to delineate the changes associated with 3′-UMP and 5′-AMP binding to RNases. The schematic representation of the CHESPA approach is shown in Figure 5.1. The projection angle and fractional shift determined for each of the five RNases is shown in Figure 5.6. The projection angle ($\theta$) is the angle between the chemical shift peak displacements (for residues with $\Delta\delta_{\text{obs}} > 0.05$ ppm at the highest ligand concentration) upon binding of 3′-UMP and 5′-AMP (represented by A vector B vector respectively in Figure 5.1) relative to the apo state. The fractional shift ($X$) corresponds to the magnitude of displacement of A vector relative to B vector. Chemical shift changes are defined as coordinated when $\cos(\theta) \geq 0.9$ and the chemical shift perturbations upon binding of the two ligands is similar, and uncoordinated when $\cos(\theta) < 0.9$. Residues that display coordinated chemical shift displacements also showed fractional shift ($X$) $\geq 1$, suggesting similar or larger changes in magnitude upon binding of 3′-UMP relative to 5′-AMP in all RNases.

Most residues of bRNaseA were involved in both 3′-UMP and 5′-AMP ligand binding. Lys66 ($P_0$), Thr45 ($B_1$) and the catalytic His12, showed coordinated displacements (Figure 5.6). While, the residues His119 and Lys41 (of active-site), Lys7 ($P_1$) binding site, Thr17 (loop 1) and residues of the C-terminal $\beta$ strand showed uncoordinated displacements of the chemical shifts, indicating the different effects upon binding of the two ligands (Figure 5.6). In hRNase2, residues in the ligand binding sites displayed both coordinated and uncoordinated displacements, with residues of loop L4 displaying uncoordinated displacements. Uncoordinated displacements were observed throughout hRNase3, with the exception of residues Gln4 and His15 that displayed coordinated behavior. The effect of ligand binding on hRNase3 was markedly different from hRNase2 and bRNaseA. The active site residues of hRNase4 showed coordinated chemical shift displacements, while uncoordinated displacements were observed in the C-terminal $\beta7$ strand.
Figure 5.6: Chemical shift projection analysis of ligand binding to RNases. The projection angle ($\cos(\theta)$, left) and the fractional shift ($X$, right), calculated for residues with $\Delta \delta_{\text{obs}} > 0.05$ ppm at the highest ligand concentration are shown. Coordinated ($\cos(\theta) \sim 0.9$, green) and uncoordinated ($\cos(\theta) < 0.9$, red) residues are identified on the three-dimensional structures of individual RNases using tube representations. Residues displaying positive and negative fractional shifts are displayed using black and white bars, respectively. Catalytic residues (His12, Lys41, His119, bRNaseA numbering) are highlighted using gray lines. Secondary structure of bRNaseA is shown on top of the graphs. The consensus sequence (obtained including gaps in the multiple sequence alignment) is used for easier comparison of proteins.
Effect of ligand binding in hRNase4 was quite similar to bRNaseA. Coordinated chemical shift displacements were predominant in the V1 domain of hRNase5, consistent with previous observations\textsuperscript{146} while uncoordinated displacements were observed for residues His8, Ser28 and His84 (Figure 5.6). Overall, these results suggest that subtle variations in the ligand binding pocket or the active-site environment in these closely related RNases can lead to significantly different long-range effects throughout the structures upon ligand binding.

5.4.C Conformational dynamics in the chemical step of bRNaseA

The conformational flexibility (or intrinsic dynamics) of bRNaseA during the chemical (RNA cleavage) step was quantified using the aggregated root means square fluctuation of top 10 slowest QHA modes (RMSF\textsubscript{10}) sampled in the hybrid QM/MM - NEB simulations. QM/MM simulations have been previously used to identify conformational fluctuations in enzymatic systems.\textsuperscript{343} The conformational fluctuations in the chemical step of bRNaseA is shown in Figure 5.7. The dynamics of bRNaseA in the chemical step is further compared to fluctuations observed in the apo state, substrate bound and product release simulations (Figure 5.7 B). Higher fluctuations are observed in the loop regions particularly loops L1 and L4, as compared to other secondary structure elements (Figure 5.7 A). Loop L1 is previously reported to play a role in bRNaseA function.\textsuperscript{138} Residues of Loop L4 (particularly Asn71) are reported to interact with purine residue of RNA substrate (Chapter 3). These loops are therefore functionally important in bRNaseA catalytic cycle and continue to be dynamic in the chemical step. However, fluctuations in these loops is considerably reduced for the chemical step when compared to other steps of the bRNaseA catalytic cycle (Figure 5.7 B). This variation in the fluctuations could be due to the difference in simulation methods used for each step. Dynamics for apo, and substrate bound systems were computed using long time-scale classical MD simulations, product release simulations were performed with umbrella sampling simulation, while the dynamics in chemical step is computed using hybrid QM/MM method which are relatively of shorter. Further, the amount of sampling varies in the computed trajectory of each step. Therefore, in order to firmly conclude that the reduced dynamics in chemical step is not by chance but a feature of bRNaseA function more sampling is required for the QM/MM - NEB simulations.
Figure 5.7: Dynamical behaviour of bRNaseA in the chemical step. Root mean square fluctuations corresponding to the $C_{\alpha}$ displacements in the top 10 quasi-harmonic modes (RMSF$_{10}$) for bRNaseA (A) RMSF$_{10}$ is projected on cartoon representations of the enzyme (B) Comparison of conformational dynamics in each step of the bRNaseA catalytic cycle; apo enzyme (red), substrate bound (orange), chemical step (turquoise) and product release (green). Flexible loop regions of bRNaseA are identified.
5.4. D Conformational sub-states in the chemical step of bRNaseA function

The RNA cleavage step in bRNaseA was modelled using hybrid QM/MM and NEB simulations in AMBER simulation package. Eight nodes were computed along the pathway from reactant to product states (Figures 5.3 and 5.4). QAA was used to determine the conformational sub-states in the chemical step of bRNaseA. Coordinates of 80800 conformations obtained along reaction pathway were projected on to 60-dimensional space using QHA for QAA. Using a mixture of gaussian modes, three clusters representing three nodes of the NEB pathway were identified (Figure 5.8). Conformational sub-states in bRNaseA during the cleavage of RNA molecule are illustrated in wireframe, showing only the Cα trace of the enzyme residues, and colored according to their respective sub-states. Detailed analysis of the conformational fluctuations reveals that no particular region in the enzyme show motions with large amplitude in the conformational sub-states identified along the chemical step pathway from reactant to product states. Whereas the conformational sub-states associated with the product release step revealed the existence of high amplitude fluctuation in loop L1 of bRNaseA (Figure 4.7). This variation in conformational motions associated with chemical step in bRNaseA could be characteristic of intrinsic dynamics in the catalytic pathway. Suggesting that functionally relevant conformational sub-states with high amplitude of fluctuations are present in product release rather than the chemical step. However, it is quite possible that reduced motions in the chemical step are a result of scarce sampling in the QM/MM - NEB method. Therefore to validate the preliminary findings of this study, profound sampling of the NEB nodes in the chemical step pathway is required.

5.5 Discussion

The rate of enzymatic reactions is traditionally measured during the steady-state period. During this time the reaction rate changes relatively slowly and therefore steady-state kinetics methods provide a good approximation of the catalytic efficiency. To facilitate comparison between the different hRNases the enzymatic activity of these enzymes was quantified using the same substrate and the kinetic parameters of bRNaseA was quantified for reference (Table 5.3). The results indicate that the catalytic efficiency in hRNases vary by ~10^5 fold with hRNase5 being the slowest enzyme. While hRNase2 is fastest among all hRNases compared in this study, it is however ~20 times slower than bRNaseA. And hRNase5 is ~ 10^6 folds slower than bRNaseA. These observations correlate to the previously identified kinetic parameters (Table 1.2) for these enzymes indicating that variation in substrate only slightly affects catalytic efficiency in these enzymes.
Figure 5.8: Conformational sub-states in the chemical step of bRNaseA function. Quasi anharmonic analysis was used for higher order statistical clustering of conformations sampled in the QM/MM-NEB simulations is shown in the box plots above. Each dot corresponds to a single conformation and is colored according to the NEB node it belongs to (see color bar). The ellipses represent identified conformational sub-states. The black arrows indicate transitions from heterogenous population to a well-defined homogenous sub-state. Motions along the anharmonic mode (γ) are illustrated in the panels below in a wire frame representation, showing only the Cα trace of the protein and color coded according to their respective NEB node.
Characterizing conformational motions that affect enzyme catalysis is essential to understand enzyme function. These motions are known to occur in each step of the enzymes’ catalytic cycle: apo, ligand binding, ground-state, product release and the chemical event itself. Ligand binding perturbs the energy landscape of the enzymes resulting in conformational changes that affect enzyme catalysis. The structural and dynamical changes induced as a result of ligand binding have been identified for a few enzymatic systems. However how motions in RNases facilitate ligand recognition and catalysis remains unknown. It will be interesting to identify conformational motion induced by ligand binding in the members of the RNase family, where different members have distinct ligand specificities. Comparison of motion in ligand binding would further provide information on the role of dynamics in the wide range of specificities and catalytic efficiencies in this enzyme family.

In this study, NMR chemical shift titration experiments were combined with the chemical shift projection analysis (CHESPA) to characterize conformational motions induced upon binding of two mononucleotides (3'-UMP and 5'-AMP) to gain insights into the mechanism and function of RNases with different biological functions. Long-range conformational rearrangements are observed in hRNase3 upon binding of 3'-UMP consistent with previous studies. While, in hRNase4 perturbations were confined only to the pyrimidine binding site consistent with the low binding affinity observed in this enzyme. The results indicate the occurrence of distinct long-range and local conformational rearrangement in each of the enzyme studied. Further using CHESPA, amino acid residues that exhibit long-range coordinated chemical shift perturbations upon ligand binding were identified. Interestingly not all RNases displayed the same coordinated amino acid perturbations previously shown in hRNase5. Difference in the amino acid displaying concerted motions was observed even for closely related enzymes, hRNase2 and hRNase3. These observations suggest that dynamical pattern in RNases is not conserved and might be adapted to the biological function of each enzyme. To confirm if these long-range motions affect the function of individual enzyme, further mutagenesis experiments and development of allosteric modulators that can directly affect the ribonucleolytic activity or biological function is required.

Evidence suggest millisecond motions affect catalysis by bRNaseA. Whether these functionally relevant motions occur in one/all steps of the enzymes’ catalytic cycle remains unknown. The conformational sub-states that modulate enzyme function are not yet identified for RNases. To obtain insights into the conformational dynamics associated with the chemical step the pathway for RNA hydrolysis by bRNaseA was modelled using hybrid QM/MM and NEB simulations. Higher
order QAA is further employed to identify the conformational sub-states associated with the conversion of substrate to products. This type of analysis has not been previously employed for other enzyme systems. The conformational sub-states/motions identified from QAA in the chemical step show reduced dynamics as compared to apo, substrate bound and product release steps (chapters 2, 3, and 4). These observations suggest that motions in RNases might not affect the chemical step directly, consistent with the proposition that efficient enzymes minimize dynamic coupling in their transition states.348 Previous studies suggest that conformational fluctuations in enzymes facilitate the formation of transition state species while the actual chemical step is a result of reaction-ready active site configuration.348-350 This argument could stand true for bRNaseA, where conformational motions in ground state or product releases effect catalysis, while the dynamics is considerably reduced in the chemical step. However, this observation would have been more obvious if the time-scales QM/MM simulations and the chemical events were same. Therefore, correlations between the identified conformational sub-states/motion and chemical step must be interpreted with caution. To identify the direct effect of dynamics on the chemical step, the kinetic isotope effect in the chemical step of a ‘dynamic knockout’ (mutations that result in loss of flexibility) of bRNaseA could be measured.350
Chapter 6  Conclusions and Future Directions
6.1 Conclusions

The motivation of the work described in this thesis was to characterize the role of internal motions in the function of closely related RNases family, including human RNases. Dynamical features that allow the progress of enzymes’ catalytic cycle through various stages were identified for several hRNases and bRNaseA, using a combination of computational and experimental methods. The major contributions of the work are summarized below.

In chapter 2, sequence base phylogenetic analysis classifies 23 RNases from the vertebrate family into four sub-families. Members of each sub-family incidentally also shared similar biological functions. The millisecond time-scale motions of representative members from each sub-family were characterized using $^{15}$N-CPMG relaxation dispersion experiments. The dynamical profile of all 23 RNases, on the µs time-scale, were probed using computer simulations. Results indicate, members within a sub-family not only perform similar biological functions but also exhibit remarkably similar dynamics. Hence, in this enzyme superfamily, dynamics is conserved in association to biological function. Characterization of dynamical properties of diverse RNases showed higher fluctuations in the loop regions as compared to other secondary structural elements in all RNases. However, distinct variations in the dynamical pattern of each sub-family arise from higher fluctuations of different loops in each sub-family. Loop L1 exhibits motions with highest amplitude in the members of the RNase A-like sub-family. On the other hand, L1 is truncated in members of EAR-like sub-family and residues of loop L4 have fluctuations higher than other regions. Further it is noted that the difference in dynamics between different sub-families arise from non-conserved residues, which may influence the functional identity of these sequences.

Large protein families found in humans and other organisms are an outcome of gene duplication events. It is unlikely that two genes with same function are stably maintained in the genome unless that extra amount of gene product is advantageous. Therefore, the paralogous genes either adopts a new/related functions (neofunctionalization) or specializes in performing one of the function of the parent genes (sub-functionalization). For instance in the RNase A superfamily, a novel antibacterial activity emerged in hRNase3 of the EAR like sub-family while hRNase1 of the RNaseA like sub-family represents the progenitor gene before duplication. Mutations arising from functional divergence appears to confer optimization of protein structure and thereby protein function. Even a small structural change due to a single mutation can lead to a large difference in
A recent experimental study has shown that mutations in the active site and distant region generating an ensemble of different conformations can evolve new enzyme function within the same fold. \(^{355}\) Results from these examples and chapter 2 highlight that conformational dynamics can provides a link between structural and functional diversity in protein/enzyme super-families.

Chapter 3 focuses on systematically characterizing the nucleotide binding properties of 7 hRNases and bRNaseA with two model substrates ACAC and AUAAU using all atom MD simulations at microsecond time-scale and theoretical analysis. The results reveal that the interaction of all 7 hRNases with the model substrates is very different, ranging from stable enzyme-substrate complex (at μs time-scale) to weakly interacting. Detailed characterization reveals that different non-catalytic residues play a significant role in enzyme-substrate stabilization and therefore, could have important consequences for differences in the biological function in these enzymes. The interactions with the catalytic triad may need to be flexible for mechanistic aspects in all RNases. Further, insights into the role of enzyme dynamics in interaction with substrates unveil dynamical pattern similar to apo enzyme with variations in the amplitude of fluctuations in some loops for only a few RNases.

Computer simulations performed at two temperatures, first is the room temperature (300 K) as most RNases experiments are performed at room temperature, and second is the human body temperature (310 K) as most RNases characterized in this study are human variants. It was noted that subtle change in temperature can have profound effects on structural and dynamical properties of RNases. Variation in temperature also resulted in a change in substrate affinity and specificity for few RNases. However, the effect of temperature on substrate binding to RNases needs to be determined experimentally. Often discrepancy is found in in-vitro and in-vivo experiments. In several cases successful drug candidates from in-vitro experiments have failed to show any results in-vivo. This is in part due to difference in temperature of the two systems. \(^{356}\) Therefore to attain conclusive insights all studies, in-vivo in-vitro and in-silico, should be performed at the body temperature of individual organism.
Conformational motions critical to enzyme function have been identified in a number of enzymes. For enzymes like DHFR and CypA, it is now established that conformational motions regulate enzyme turnover. In both these enzymes, networks that connect distal loops to active site residues are identified. Further the population of conformations that assist in the progress of catalytic cycle in known. However, these intricate details that connect dynamics to catalysis remain unknown for hRNases. Chapter 4 successfully used a computational approach to identify conformational sub-states and fluctuations associated with the RNase function. Previous evidence has indicated that conformational motions in the product release step is the rate-limiting for bRNaseA. A combination of umbrella sampling simulation and QAA identifies sub-states associated with product(s) release in hRNase3, hRNase5 and bRNaseA, one member from each sub-family for which the biological function is known (chapter 1). Interestingly QAA reveals the presence of separate conformational sub-states for the three RNases studied. Even motions associated with the sub-states were different in each enzyme. In addition, the structural interactions that allow product release were different for three RNases. These observations suggest that RNases have evolved to achieve their respective function by effectively utilizing the motions unique to each enzyme.

The combination of umbrella sampling with QAA could further be used to characterize conformational motion unique to each enzyme in other protein families, enabling a better understanding of the role of dynamics in the evolution of diverse functions in structurally related super-families. Further, identifying conformational sub-states that pre-exist in the conformational landscape of related enzymes will allow selection of appropriate starting conformation required for successful enzyme design for a target reaction. In addition, the combination of umbrella sampling with QAA along the reaction pathway could be applied to identify conformational sub-states and dynamical features critical to the catalytic mechanism other therapeutic enzymes. Identifying the motions that enable the formation of transition state or are critical to overcoming the rate limiting step, will open new avenues for developing next generation therapeutics. Small molecules can be designed to either enhance or inhibit enzyme function by modulate enzyme flexibility. In other words, use of this methodology can enable better understanding of the biophysical nature of enzymes and can generate applications in medicine, industry and enzyme design.

In enzymes, inherently dynamic process is involved in substrate binding product release, and the chemical event itself. Further, studies suggest that transition between conformations of various
catalytic steps is achieved as a result of coupling between slower long-range motions and faster active site dynamics. Long range motions in enzymes expand from the enzymes surface to the active site through a network of amino acids connected via non-bonding interactions. Conformational fluctuations that enable transitions between different catalytic cycle events limit the rate of catalysis in enzyme DHFR. To identify rate limiting motions in RNases function, it is therefore required to characterize intrinsic dynamics and sub-states in each step of the catalytic cycle of various members of the RNases family. Interesting insights have been obtained in the dynamic associated with the apo enzyme, ground state, and product release steps of RNases in Chapters 2 - 4. The work in chapter 5, focuses on characterizing dynamics in the ligand binding and the chemical step in select RNases. NMR chemical shift titrations and CHESPA reveal the existence of distinct long-range networks in hRNases and bRNaseA, suggesting that enzymes with diverse biological functions also have diverse dynamics. Further considerable reduced dynamics was observed in bRNaseA during the chemical step. Identified conformational sub-states of the chemical step do not high amplitude fluctuations in the distant loops of bRNaseA as seen in the conformational sub-states of the product release step. This observation indicates that the chemical step in bRNaseA occurs in a relatively static environment.

Overall the findings of this thesis reveal the occurrence of distinct long-range conformational fluctuations in the apo and ground-state enzymes, ligand binding and product release steps of hRNases with different biological functions and bRNaseA. Conformational sub-states with reduced fluctuations were identified in the chemical step of bRNaseA (Figure 6.1). In addition, pre-existence of product release conformations in the ensemble of apo enzyme is established. Taken together, observations for this thesis suggest that in RNases, conformational fluctuations facilitate substrate binding, formation of transition state species and product release, while the actual chemical step is a result of reaction-ready active site configuration. However, the correlation between rate-limiting steps of catalysis and rates of conformational change is yet to be established.
Figure 6.1: Conformational dynamics in the catalytic cycle of bRNaseA. (A) Root mean square fluctuations of Cα atoms in the top 10 QHA modes (RMSF_{10}) in each step of bRNaseA catalytic cycle reveal existence of high amplitude fluctuations in apo (red), substrate bound (orange) and product release steps (green), while reduced dynamics is observed for the chemical step of RNA cleavage (turquoise). (B) All atom RMSF_{10} plotted on the structure of bRNaseA (PDB ID: 7RSA) using tube representation for each step of catalytic cycle. The dynamical range represented using the color spectrum is consistent across all steps with the blue and red ends of the spectrum corresponding to low and high dynamic regions, respectively. The thickness of the tube corresponds to the flexibility of residues in each step with thicker tubes corresponding to flexible regions and thinner tubes representing less flexible regions. Important loops are identified.
6.2 Future directions

RNase structural homologs in humans are involved in a distinct biological function with promising therapeutic applications in several diseases including cancer, ALS and Parkinson’s disease. As these closely related enzymes share similar binding pockets and/or active sites, developing drugs targeting individual members is difficult. To target these enzymes individually, it is required to develop highly specific small-molecule binders and modulators that can alter enzymatic activity through long-range dynamical motions away from the active site. In this study, using a combinations NMR experiments and computer simulation functionally relevant motions unique to select RNases have been identified for events such as ligand binding and product release. Modulating the dynamical motions that can efficiently funnel the enzyme through its catalytic cycle, particularly in the rate limiting step will allow regulating these enzymes individually. However, it remains unknow if the kinetics and motions in rate limiting step of these closely related homologs is same. Outlined below is a summary of how knowledge obtained in this thesis could be combined with further computational and experimental studies to better understand distinct dynamics hRNases.

1. In chapters 2 and 3 all atom simulation trajectories for hRNases in the apo state and substrate bound forms were computed. The conformations obtained from these MD runs can be further analyzed with QAA to identify the conformational sub-states and populations in apo and substrate bound forms respectively. These studies will also provide insights into conformational fluctuations and motions critical for molecular recognition of substrates or other binding partners. Populations of identified conformational sub-states could further be validated using conformational probabilities obtained from NMR relaxation dispersion experiments of hRNases in apo state and in presence of substrate analogs.

2. In chapter 4 it was determined that members in individual sub-families of the RNase superfamily sample separate conformational sub-states in the product release step of the catalytic cycle. Whether the members within a sub-family access same conformational sub-states during product release remains unknown. Therefore, exploring conformational sub-states in product release step for more RNases from each sub-family would be an important step towards understanding the homogeneity in dynamics within each sub-family. It would further be helpful in distinguishing how dynamics has evolved across sub-families of RNase superfamily.
3. Using hybrid QM/MM modeling with nudged elastic band (NEB) method, chapter 5 in this thesis identifies conformational sub-states and fluctuations that assist in phosphodiester cleavage step for bRNaseA of RNase A-like sub-family. This approach could be implemented to identify conformational sub-states in chemical turnover (RNA cleavage) step in members of other three sub-families as well. The results obtained will allow comparative analysis of conformational or dynamical changes associated with the enzyme mechanism in all identified sub-families.

4. Enzymatic reactions proceed from substrate to product(s) via a transient high energy transition state (TS). The enzyme’s ability to make the reaction faster depends on the fact that it stabilizes the TS. Hybrid QM/MM approach have proven useful to determine the dynamics of TS formation in enzymes. Identification and characterization of motions that facilitate attainment of TS would provide insights into the mechanistic details pertaining to the wide range of catalytic efficiencies observed in these closely related RNases. Modulating motions that promote TS formation could further be used to modulate enzyme activity in RNase family members. Obtaining information about TS in RNases will also provide a wealth of information necessary to design transition state analogs, which are also useful enzymatic inhibitors, unique to individual RNases.

5. Studies in other enzymatic systems have indicated that rate of enzyme conformational changes coincides with the rate limiting step in the enzymatic catalytic cycle. The rate limiting step for bRNaseA is debated. While, the rate limiting state for hRNases remains unknown. In the light of wide range of catalytic efficiency in these enzymes, it is very much possible that rate limiting step in all human RNases might not be same. Therefore, to understand the role of dynamics in the function of hRNases it is first required to identify the rate limiting step in all hRNases. Thermodynamic and kinetic characterization of substrate binding, chemical step, and product release, using hybrid QM/MM simulations, will provide a detailed picture of the complete catalytic process in RNases. The heights of energetic barrier for each step would allow the determination of the rate-limiting step. Stopped-flow is one experimental technique that could also be used for this purpose.

6. Detailed analysis of the conformational sub-states will provide information about location and amplitude of the functionally relevant dynamical motions associated with the RNA cleavage by hRNases. This information could be used to identify mutational hot-spots that potentially control
the function of divergent RNases. The effect of these mutations on the enzymatic activity could be evaluated using steady-state Michaelis-Menten kinetics experiments. Further, characterizing the dynamics and conformational sub-states in mutant enzymes could further emphasize on the interrelation between primary sequence and enzyme dynamics.

7. As observed in this study, unique dynamical properties and changes in sequence are localized to the distal loop regions emphasizing on their role in diversification of RNases superfamily. Engineering the distal loops by either inserting or deleting amino acids can induce changes in dynamics of these enzymes, modulating their respective function. Increase the length of loops would allow harnessing more energy from the bulk solvent to increasing the dynamics in enzyme thereby increasing catalytic turnover, as seen in \(^{358}\). On the other hand, reducing the length of the loop might work in a vice versa fashion. Allosteric modulation of enzyme function could also be achieved by identification/design of small molecules that can bind to specific loop in individual RNases. Whether it is required to increase/decrease catalytic turnover and dynamics in induvial RNases depends on the goal of study. For example, antibacterial activity of hRNase3 does not depend on its ribonucleolytic activity, there improving catalytic turnover of hRNase3 is not useful for enhancing its biological function.

8. It is suggested that human members of the EAR-like sub-family possess the ability to disrupt bacterial membrane integrity. However, it still remains unclear how these RNases specifically target the microbial cell wall and how it permeabilizes the membrane. Computational modelling of how hRNases interact with liposaccharide membrane could provide preliminary information on the mechanism of antibacterial activity in the enzymes. Understanding the bactericidal mechanisms is crucial for its development of hRNases as a novel antimicrobial agent.

9. The \(^{15}\text{N-CPMG}\) investigation does not report conformational changes in hRNase5 on the millisecond time-scale. In combination with computer simulations, use of other experimental methods like hydrogen/deuterium exchange mass spectrometry could provide additional clues on slow time-scale motions experienced by hRNase5.
6.3 Assessing the quality of computer simulations

Computer simulations are valuable tools that have immensely contributed towards understanding biological systems. Like experimental techniques, computer simulations have their own limitations. Simulations approaches include many different types of theoretical foundation each with their own approximations, creating an uncertainty associated with the prediction. Unlike experiments, it is proven hard to assess how much of the simulation is representative of the actual biophysical properties. Quantification of the uncertainties in a simulation can increase the confidence and significance of the observables obtained from simulations.

The most important issues that arise while using computer simulations to describe the properties of an enzymatic system include: First, accurate description of the starting point for the simulations to ensure that the structure is at a reasonable minima; Second, the time-scale of computer simulations are still not fully comparable to the time-scale of enzymatic function creating an uncertainty in the amount of conformational sampling achieved during simulations; Third, reproducibility of the observation/quantity of interest, i.e. statistically significant conclusion(s). Some suggestions to assess the quality of computed simulation trajectories are outlined below:

1. It is required to dedicate ample amount of time in carefully modeling the system(s) of interest and setting up the simulations. Among many other things choice of starting structure, desired ensemble, temperature, pressure and the force field to accurately describe your system are important factors that can change the desired predictions considerably. If the time-scale of biological process cannot be reached by simplistic simulations, including enhanced sampling techniques to access the conformational landscape is a logical approach. Therefore, a simulation protocol and subsequent analysis of its results should take into account the intent of the study.

2. Multiple trajectories should be computed starting from different starting structures. A variety in starting structures can be created by either minimizing the system or heating it at higher temperatures. Another way to generate new starting structures is to run a relatively small MD simulation and extract conformations at different time points each serving as an independent starting structure. The trajectories computed from all the different starting structures should be critically analyzed for reproducible behavior of any observables which can also be
measured with biophysical technique. To avoid bias, include all simulation data in analysis except the trajectories that were either not equilibrated or incorrectly setup.\textsuperscript{362}

3. The initial part of the simulation where the system is equilibrated to reach a local energy minimum relaxed state can introduce a bias in the analysis.\textsuperscript{365} Therefore, this equilibration portion of the simulation should not be included while computing the observables from simulation trajectory.

4. Single simulation trajectories can be analyzed using \textit{time correlation analysis}.\textsuperscript{366} Correlation time separates statistically independent values of the observable, having many statistically independent measurements of the observable suggests good sampling for that particular observable.

5. \textit{Block averaging}\textsuperscript{362, 366, 367} is another good method to calculate the statistical uncertainty of the ensemble averages calculated from single simulation trajectory. In this method the simulation trajectory is divided into subsets. Mean of an observable is computed for each segment and then a standard error is calculated for all means providing a direct estimation of statistical uncertainty.

6. Inadequate sampling in simulation trajectories can be checked by identifying if the simulation has converged. This can be done by either monitoring RMSD or certain dihedral angles in the simulation trajectory with respect to reference structure.\textsuperscript{366} A simulation is expected to be converged when the fluctuations in these quantities reduce along time.

7. To analyze if each trajectory adequately samples the conformational landscape a \textit{combined clustering} analysis can be useful.\textsuperscript{368} In this approach the conformations sub-states are grouped into clusters based on RMSD or dihedral space. If the simulations are converged all clusters should contain equal population of conformation from individual trajectory.

8. Observables from simulations should be reported with level of confidence instead of error bars.\textsuperscript{362} Confidence level quantifies the frequency of interval in a sample data that the true value of unknown parameter lies in within the interval. This interval is defined as the point estimate plus/minus the standard deviation. Observables (that follow a gaussian distribution)
with 95 % level of confidence passes the qualitative test. For non-gaussian observables use of bootstrapping method is suggested.
List of References


82. Elsasser, B.; Fels, G.; Weare, J. H., Qm/mm simulation (b3lyp) of the mase cleavage-
transesterification reaction supports a triester a(n) + d-n associative mechanism with an o2'
Q.; Miyagi, M.; Anderson, V. E.; York, D. M.; Piccirilli, J. A.; Harris, M. E., Experimental
and computational analysis of the transition state for ribonuclease a-catalyzed rna 2'-o-
84. Kamerlin, S. C.; Sharma, P. K.; Prasad, R. B.; Warshel, A., Why nature really chose
85. Elsasser, B.; Fels, G.; Weare, J. H., Qm/mm simulation (b3lyp) of the rnase a cleavage-
transesterification reaction supports a triester a(n) + d(n) associative mechanism with an o2'
86. Breslow, R.; Chapman, W. H., Jr., On the mechanism of action of ribonuclease a: Relevance
of enzymatic studies with a p-nitrophenylphosphate ester and a thiophosphate ester. Proc.
87. Langen, R.; Schweins, T.; Warshel, A., On the mechanism of guanosine triphosphate
6 (5), 399-412.
89. Cleland, W. W.; Hengge, A. C., Enzymatic mechanisms of phosphate and sulfate transfer.
Chem. Rev. 2006, 106 (8), 3252-78.
90. Findlay, D.; Herries, D. G.; Mathias, A. P.; Rabin, B. R.; Ross, C. A., The active site and
mechanism of action of bovine pancreatic ribonuclease. 7. The catalytic mechanism.
91. Fersht, A., Structure and mechanism in protein science: A guide to enzyme catalysis and
92. Thompson, J. E.; Kutateladze, T. G.; Schuster, M. C.; Venegas, F. D.; Messmore, J. M.;
93. Breslow, R.; Chapman, W. H., On the mechanism of action of ribonuclease a: Relevance of
enzymatic studies with a p-nitrophenylphosphate ester and a thiophosphate ester. Proc.
94. Lopez, X.; York, D. M.; Dejaegere, A.; Karplus, M., Theoretical studies on the hydrolysis of
phosphate diesters in the gas phase, solution, and mase a. Int. J. Quantum Chem. 2002,
86 (1), 10-26.
95. Loverix, S.; Winqvist, A.; Strömberg, R.; Steyaert, J., Mechanism of mase t1: Concerted
2000, 7 (8), 651-658.
96. Cuchillo, C. M.; Pares, X.; Guasch, A.; Barman, T.; Travers, F.; Nogues, M. V., The role of
2',3'-cyclic phosphodiesters in the bovine pancreatic ribonuclease a catalysed cleavage of
97. Thompson, J. E.; Venegas, F. D.; Raines, R. T., Energetics of catalysis by ribonucleases:
Fate of the 2',3'-cyclic phosphodiester intermediate. Biochemistry 1994, 33 (23), 7408-14.
(21), 6267-74.
modulates active site dynamics in extremophilic alcohol dehydrogenases. J. Biol. Chem.
2013, 288 (20), 14087-97.
100. Reynolds, K. A.; Mclaughlin, R. N.; Ranganathan, R., Hot spots for allosteric regulation on


172. Van Der Kamp, M. W.; Chaudret, R.; Mulholland, A. J., Qm/mm modelling of ketosteroid isomerase reactivity indicates that active site closure is integral to catalysis. *FEBS J.* **2013**, *280* (13), 3120-31.


199. Larion, M.; Hansen, A. L.; Zhang, F.; Bruschweiler-Li, L.; Tugarinov, V.; Miller, B. G.; Bruschweiler, R., Kinetic cooperativity in human pancreatic glucokinase originates from


Amber 14, University of California, San Francisco, 2014.

311. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C.,  
Ff14sb: Improving the accuracy of protein side chain and backbone parameters from ff99sb.  

312. Agarwal, P. K., Cis/trans isomerization in hiv-1 capsid protein catalyzed by cyclophilin a:  

313. Ramanathan, A.; Agarwal, P. K.; Kurnikova, M.; Langmead, C. J., An online approach for  

Koossváry, A. Kovalenko, T.S. Lee, S. L., T. Luchko, R. Luo, B. Madej, K.M. Merz, F.  
University of California, San Francisco, 2014.

315. Turner, P. Xmgrace, version 5.1. 19, Center for Coastal and Land-Margin Research, Oregon  
Graduate Institute of Science and Technology, Beaverton, OR, 2005.

316. Roe, D. R.; Cheatham Iii, T. E., Ptraj and cpptraj: Software for processing and analysis of  

317. Grossfield, A. Wham: The weighted histogram analysis method, 2.0.9.

318. Lu, Y.; Wang, Y.; Xu, Z.; Yan, X.; Luo, X.; Jiang, H.; Zhu, W., C-x...H contacts in  
B 2009, 113 (37), 12615-21.

319. Desiraju, G. R., C-h...O and other weak hydrogen bonds. From crystal engineering to virtual  

320. Shan, S. O.; Herschlag, D., Hydrogen bonding in enzymatic catalysis: Analysis of energetic  

(2).

of hydrophobic residues in the catalytic cleft of streptococcus pneumoniae hyaluronate  


motions correlated to hydride transfer in dihydrofolate reductase. J. Am. Chem. Soc. 2014,  
136 (6), 2575-82.

Fraser, J. S., Rescue of conformational dynamics in enzyme catalysis by directed evolution.  


327. Narayanan, C.; Bernard, D. N.; Bafna, K.; Gagne, D.; Agarwal, P. K.; Doucet, N., Ligand-  
induced variations in structural and dynamical properties within an enzyme superfamily.  
Front Mol Biosci 2018, 5, 54.

328. Kurkcuoglu, Z.; Bakan, A.; Kocaman, D.; Bahar, I.; Doruker, P., Coupling between catalytic  
loop motions and enzyme global dynamics. PLoS Comp. Biol. 2012, 8 (9), e1002705.


**Vita**

Khushboo Bafna was born on October 18th, 1988 in Kolkata, India, to Raj Kumar and Rajshree Bafna. She attended junior and middle school at Mahadevi Birla Girls higher secondary school in Kolkata before moving to Jaipur for further schooling in 2003. After completing high school in 2007, she enrolled in a Bachelor of Science degree in Bioinformatics at Jaipur National University in Jaipur, India. Upon graduation, Khushboo decided to pursue a Master's degree in Bioinformatics and traveled to Puducherry in Southern India, where she joined the Pondicherry Central University. She completed her master's program in 2012 and started working as a Junior Research Fellow at the Indian Institute of Technology (IIT) Delhi, in New Delhi India for two years. In 2014, Khushboo joined the Graduate School of Genome Science and Technology at the University of Tennessee Knoxville in the USA. She earned her Ph.D. in Life Sciences in 2019. Upon completion of her doctoral degree, Khushboo plans to pursue research at the interface of experimental and computational biology.