Mechanistic Insights of a Non-specific Inosine-Uridine Nucleoside Hydrolase (IU-NH) found in *Arabidopsis thaliana*

By

Jason G. Griner

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy in Molecular Biosciences

Middle Tennessee State University
June 2024

Dissertation Committee:

Dr. Paul C. Kline, Chair

Dr. Justin M. Miller

Dr. Mary B. Farone

Dr. James B. Robertson

Dr. Rebecca L. Seipelt-Thiemann

Mechanistic Insights of a Non-specific Inosine-Uridine Nucleoside Hydrolase (IU-NH) found in *Arabidopsis thaliana*

Jason G. Griner 6/28/2024

Approved:

_____

Dr. Paul Kline, Chair

_____

Dr. Justin M. Miller

_____

Dr. Mary B. Farone

_____

Dr. James B. Robertson

_____

Dr. Rebecca L. Seipelt-Thiemann

_____

Dean of the College of Graduate Studies

I dedicate this dissertation to my wife and partner Felisia, my mother and father Donna and Benny, and my sister Kristin. Through their love and support over the years I was able to stay focused, plot a path, and execute this research to complete a life goal. I could not have done it without them.

ACKNOWLEDGMENTS

I would like to take the time to thank several people at MTSU who assisted with this research to make it possible. First and foremost, my advisor Dr. Paul Kline. Thank you for your mentorship and confidence to take me on as a student and work with me to complete this research. Dr. Brock Arivett who assisted with plasmid development and Dr. Dan Bryant who helped with plasmid validations. Dr. Anthony Farone and Dr. Mary Farone for use of their laboratory while at MTSU. They were essential in providing resources to help develop cultures for enzyme purifications. Dr. James Robertson who also provided resources for plasmid development with cultures. Dr. Justin Miller, Dr. Rebecca Seipelt-Thiemann, and Dr. Chad Brambley who provided crucial guidance for the computational work produced for this research. Finally, thanks to the MTSU MOBI program for travel and support to attend the 2022 ASBMB conference and the 2023 ASM TN/KY regional meeting to present posters for this research. Thank you all so very much for your help and contributions for this project.

ABSTRACT

Herbicides are crucial tools used to control unwanted vegetation to increase crop yields and landscape aesthetics. Current herbicides target metabolic pathways for protein production in plants using chemical inhibitors to inhibit metabolic pathways involved in protein synthesis, but an alternative strategy may involve targeting nucleoside salvage pathways in plants. Salvage pathways in plants recycle nucleosides to create the nucleotide monomer pools they need to synthesize DNA and RNA for protein production. The process by which nucleosides are produced varies from organism to organism, but nucleoside hydrolases are enzymes utilized by plants to catalyze purine and pyrimidine hydrolysis to create an energy efficient pathway allowing recycling of nucleobases in early seedling development. Salvage processes with nucleoside hydrolases have been found to occur in many organisms such as plants, yeast, bacteria, and protozoa, but are absent in humans. This research was conducted to understand the mechanism by which a non-specific inosine-uridine nucleoside hydrolase (IU-NH), designated URH1 from *Arabidopsis thaliana* catalyzes the hydrolysis of selected nucleosides. URH1 was cloned, overexpressed, and purified to greater than 95% homogeneity to determine the activity, optimum pH, oligomerization state, calcium content, and substrate specificity. *In silico* structural models and molecular dynamic simulations were performed with URH1 and several different nucleoside hydrolases from protozoa and plant models to determine if conserved residues were present in the active site and if any flexible loops were present. URH1 was discovered to be a non-specific nucleoside hydrolase with a broad pH range and a high activity for uridine at high concentrations. Several uninvestigated residues and flexible loops were discovered around the active site which may serve as key sites for mechanistic control.

TABLE OF CONTENTS

PAGE

LIST OF TABLES

TABLES PAGE

LIST OF FIGURES

FIGURE PAGE

LIST OF APPENDICES

CHAPTER I: INTRODUCTION

Nucleoside hydrolases (NHs) are crucial enzymes involved in the cleavage of the β-N-glycosidic bond of nucleosides creating the corresponding nucleobase and ribose to facilitate recycling and maintaining cellular nucleotide pools in organisms (Figure 1).[1] Nucleoside hydrolases have been classified based on their substrate specificities and tertiary structure/sequence.[2,3] Based on substrate specificities, there are four classifications of these enzymes: the non-specific inosine-uridine (IU-NH), purine-specific inosine-adenosine-guanosine (IAG-NH), pyrimidine- specific cytidine-uridine (CU-NH), and 6-oxopurine guanosine-inosine (IG-NH) nucleoside hydrolases.[4,5] They were further subcategorized into three sequence or homology-based groups to include: Group I (IU-NH and CU-NH), Group II (IAG- NH), and Group III (GI-NH and uncharacterized).[6–8]

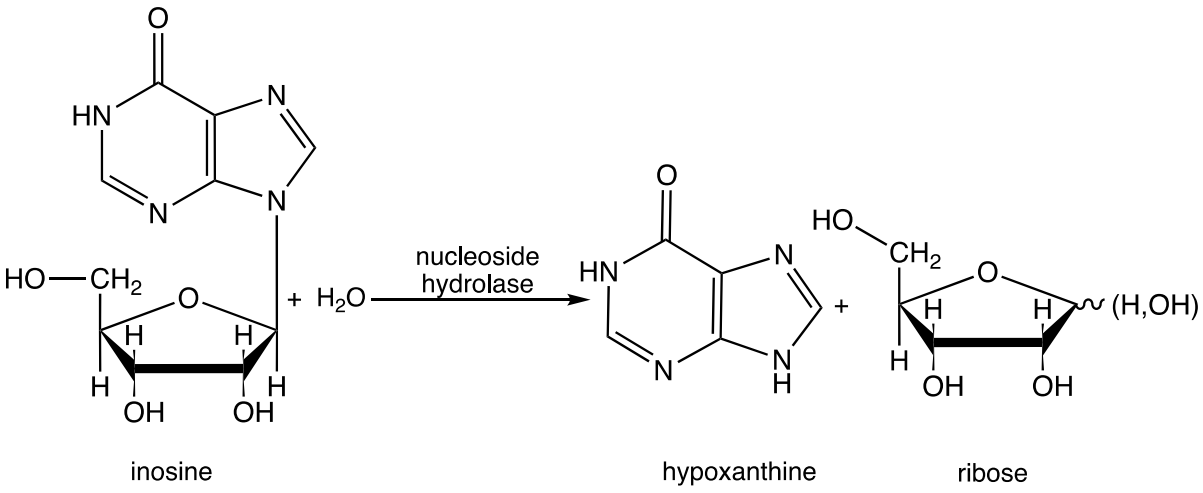


Figure 1. Nucleoside Hydrolysis. Hydrolysis reaction between the purine nucleoside inosine and water catalyzed by nucleoside hydrolase producing the nucleobase hypoxanthine and ribose. The initial stereochemistry of ribose is not known. Figure 1 was created using Chemdraw (Version 22.2.0).

Nucleoside hydrolases have been isolated from a wide variety of organisms including parasitic protozoans, such as *Trypanosoma*,[9–12] *Crithidia* and *Leishmania*,[1,13–16] as well as bacteria,[17–19] archaea,[20] baker's yeast,[21] insects,[22] and various species of plants.[23–28] A non-specific IU-NH found in *Arabidopsis thaliana* known as Uridine Ribohydrolase-1 (URH1) may be a promising target for herbicide design because of its specificity to plants and its critical role in plant metabolism during seedling development.[23,29]

1.1 The Impacts of Herbicides on Agriculture and Human Safety

The reliance on current herbicides to control weed populations in agriculture has raised significant economic and health concerns over the years. Current agricultural practices have created a large dependency on herbicides to eliminate unwanted plants to increase crop yields, but instead have increased demand because unwanted plants have become more tolerant.[30] Recent studies have linked herbicides to toxic effects on the nervous system and as endocrine disrupting chemicals raising concerns that they pose health risks to humans.[31,32] The most widely used herbicide glyphosate (Roundup) has even been linked to non-Hodgkin's lymphoma and involved in civil litigation resulting in a verdict worth billions of dollars.[30,33,34] We have continued to increase the amount of herbicides applied to crops without a true understanding of how they affect humans.

Since the 1980s, only a single herbicide has been introduced with a new mode of action creating a greater need for herbicides with a different mode of action to protect human health.[30,33] Herbicides are typically inhibitors that block important enzymatic functions in plants.[35] Glyphosate blocks plant growth by acting as an inhibitor to enolpyruvylshikimate-3-phosphate synthase (EPSPS) blocking production of the essential aromatic amino acids tryptophan, tyrosine, and phenylalanine.[36] Understanding the role of nucleoside hydrolases in plant growth may lead to new

herbicides that increase crop yields by more effectively controlling weeds and increase safety by targeting enzymes lacked by humans.

1.2 Biochemical Mechanism of Nucleoside Hydrolases

1.2.1 Types of Nucleosides

To understand how nucleoside hydrolases play a role in metabolic salvage pathways in organisms we must first understand the different types of nucleoside substrates they hydrolyze. Nucleosides have a cyclic pentose furanoside-type sugar (β-D-ribose) linked to a heterocyclic nucleobase on the C1' carbon and exist as ribonucleoside or deoxyribonucleoside depending on whether the C2' carbon of the ribose sugar is hydroxylated or not.[37–39] The N9 and N1 of purines and pyrimidines, respectively, are one of the five heterocyclic nucleobases attached via a glycosidic bond to C1' of the ribose sugar (Figure 2)..[37–39] The five most common nucleobases are adenine, guanine, cytosine, thymine, and uracil.[37–39]

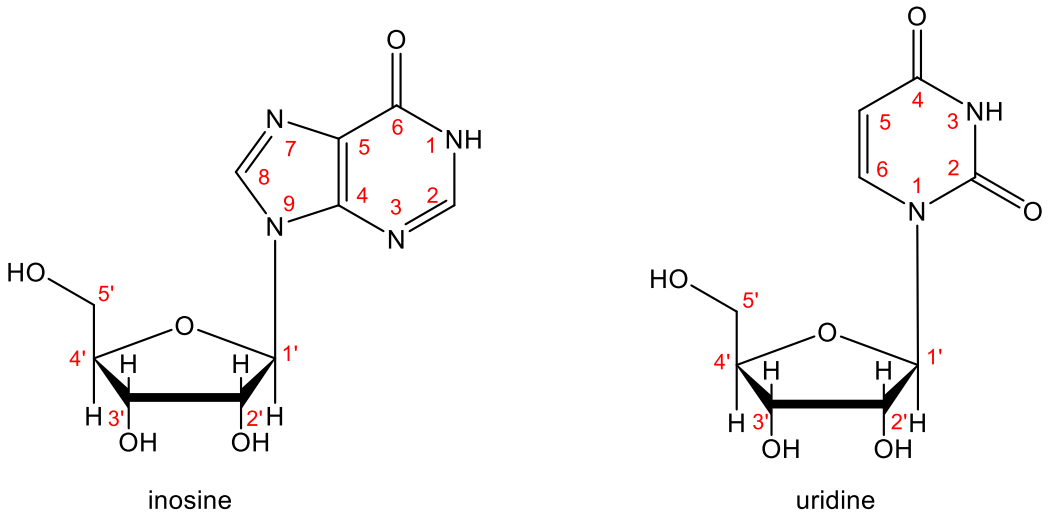


Figure 2. Ribonucleoside Numbering. Examples of purine (inosine) and pyrimidine (uridine) ribonucleosides with base and sugar reference numbers identified on each ring structure. Figure 2 was created using Chemdraw (Version 22.2.0).

Adenine and guanine, are purine nucleobases while cytosine, thymine, and uracil are pyrimidine nucleobases.[37–39] Purines consist of a six-membered heterocyclic ring fused to a second five-membered heterocyclic ring with variations at the C2 and/or C6 carbon of the six-membered heterocyclic ring resulting in adenosine or guanosine if adenine and guanine are linked with ribose (Appendix A). [37-39] Other purine nucleobases bases include hypoxanthine and xanthine which are found in the nucleosides, inosine and xanthosine respectively. These nucleosides/nucleobases play key roles as intermediates in the synthesis of adenine and guanine.[37-39] Pyrimidines consist of a single six-membered heterocyclic ring with variations at the C2 and C4 carbon resulting in uridine and cytidine if uracil or cytosine are combined with ribose (Appendix A).[37-39] Another pyrimidine nucleobase is thymine which is found in the 2'-deoxyribonucleoside thymidine.[37-39] The C5 carbon of the pyrimidine base contains a methyl group giving it unique chemical qualities compared to the other nucleobases.[37–39] Nucleosides take many derivative forms so understanding the structural differences of nucleosides as substrates allow us to see how enzymatic function can occur in the active site of nucleoside hydrolases to better understand how they function as salvage products in plants.

1.2.2 Nucleoside hydrolase catalysis

Enzymes are protein catalysts that drive biological function by speeding up reactions. To understand how enzymes work, experimental knowledge of the activity, specificity, optimal pH, and cofactor interactions are needed. The currently accepted mechanism for purine nucleoside hydrolysis is based on extensive studies from the non-specific, Group I IU-NH from the protozoa *Crithidia fasciculata* which uses a His/Asp amino acid residue pair for base protonation and hydrolysis.[7] Mutation experiments conducted by Gopaul et al., 1996 showed that

when His241 was replaced with a non-reactive Ala, a 2,100-fold loss of activity was observed.[40] Purine hydrolysis models contain an oxocarbenium ion based on kinetic isotope studies to explain how this reaction occurs.[16,41]

The enzyme uses three catalytic strategies which have developed from kinetic isotope studies to describe transition state properties in purine hydrolysis: steric and electrostatic stabilization of the oxocarbenium ion, activation of nucleophilic water, and activation of the nucleobase leaving group (Figure 3).[16,42–44] The ribose hydroxyl groups on the C2', C3', and C5' carbons all have a stabilizing effect on the glycosidic bond but much less is known about interactions which involve C5'.[1,9] The C2' and C3' hydroxyl groups act in coordination with calcium as binding sites to facilitate docking and keep the molecule stabilized to create the oxocarbenium transition state.[1,9] Based on solvent reactivity studies with *C. fasciculata*, the activated water molecule is a specific enzyme-bound water rather than a general solvent water which functions as a nucleophile to the C1' on ribose.[1] The nucleobase leaving group has undergone the most experimental analysis for specificity, mutagenesis, molecular dynamics, and X-ray crystallography.[10,16,40,44–46] Calcium binding residues help reduce the pK_a of water and align it in the proper orientation for attack.[9] Calcium has been determined to be a key cofactor that is required to keep the active site coordinated and stabilized.

1.2.3 Active site function with calcium

A calcium ion is observed to play a central role as a cofactor which tightly binds several Asp residues together in the active site of non-specific IU-NHs in protozoa and plants.[2,47] The archetype non-specific, Group I IU-NH found in the protozoa *C. fasciculata* was shown to have a sequence motif of DXDXXXDD, located at Asp8, Asp10, Asp14, and Asp15, used for binding

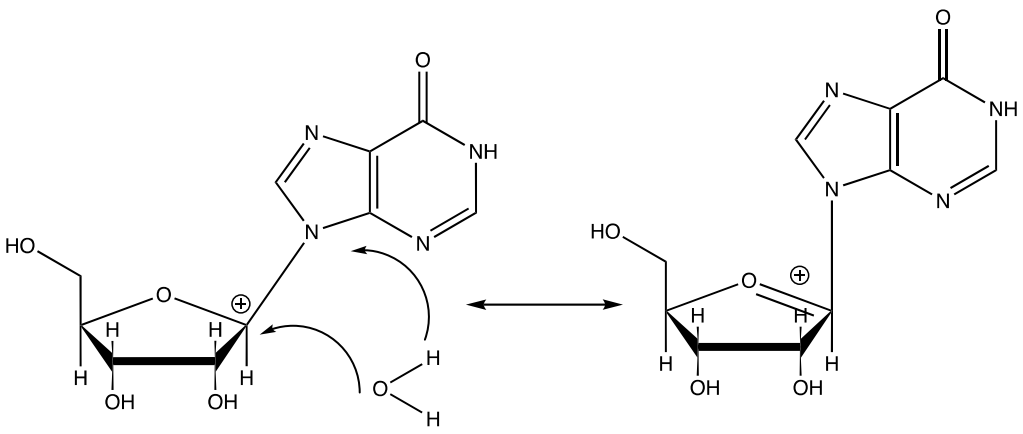


Figure 3. The oxocarbenium transition state. The left shows a resonance structure of inosine with a nucleophilic water attacking the C1' of a carbocation on the ribose. To facilitate the breaking of the N-glycoside bond and the leaving of the purine base, N7 is protonated by an essential histidine residue. The right shows an oxocarbenium ion resonance structure between C1' and O4' on the ribose. Figure 3 was created using Chemdraw (Version 22.2.0).

calcium as a cofactor in the active site**.**[2] Coordination sites on the calcium ion were found to interact with Asp10, Asp15, Thr126, and Asp242 amino acid residues when a p- aminophenyliminorbitol ligand was docked.[2] Both 2' and 3' hydroxyls on the ribose also coordinated with calcium assuming a C3'-*exo* conformation.[2] The O5' hydroxyl seemed to form hydrogen bonds between Glu166 while the nitrogen on p- aminophenyliminorbitol formed a hydrogen bond between Asn168.[2] A similar calcium binding motif can also be found in plants.

A homologous motif with the same sequence order of DXDXXXDD found in protozoa was also discovered in plant IU-NHs from *P. patens* and *Z. mays*.[47] When Asp residues at site locations were targeted for mutations, a 10^4-fold lower activity than the wild type was observed.[47] To date, there have been no attempts to try to understand and measure calcium content from

enzyme purifications for non-specific IU-NHs in plants. To develop a better understanding of how URH1 may function in *Arabidopsis* a comparison of URH1 from *Arabidopsis* can be made to other types of IU-NHs in other organisms.

1.3 Salvage Pathways in Protozoa

Parasitic protozoan have historically been the target of research since the diseases that result from parasitic protozoan infections are difficult to treat and have no antibiotic options. Since nucleoside salvage pathways are not present in humans, parasitic protozoan have no other means of nucleic acid production making them attractive targets for drug design.[2,7,9,10,13,14] Research has primarily targeted nucleoside hydrolases for drug design against leishmaniosis, trypanosomiasis, and plasmodium, the organisms responsible for causing leishmaniasis, African sleeping sickness, and malaria respectively.[5,13,48] Leishmaniasis and trypanosomiasis are estimated to cause anywhere from 50,000-500,000 deaths per year, while malaria causes over 600,000 deaths per year.[49–51] Protozoan have both Group I non-specific nucleoside hydrolases such as IU-NHs and Group II purine-specific nucleosidases, such as IAG-NH.

The most well characterized non-specific nucleoside hydrolases, IU-NH, are produced by *Crithidia fasciculata*[1,7,40,52], *Leishmania major*[13,14], and *Leishmania donovani*[15,53–56], while *Trypanosoma vivax*[9,10,12,44,57] and *Trypanosoma brucei*[11,57–59] more often produce specific IAG-NHs. The structural and mechanistic characteristics for an IU-NH in *C. fasciculata* and an IAG-NH in *T. vivax* were described in 2003 by Versées and Steyaert.[7] The structure of IAG-NHs in *T. vivax* were shown to be that of a homodimer, while the IU-NH in *C. fasciculata* was shown to be a homotetramer.[7] The homotetramer of IU-NH had two binding regions: one region with an affinity for ribose and the other with an affinity for the nitrogenous base.[7] The active sites for each of these

enzymes were located near the C-terminus and consisted of a single globular domain with two loops (loop I & II) positioned on each side of the active site.[7] A tightly-bound calcium ion was located on both active sites and coordinated with water to replace a hydroxyl group on the ribose, while a hydrophobic pocket assisted with nucleobase binding.[7]

Amino acid residue on purine specific IAG-NHs from *T. vivax* primarily involved with calcium binding occurred at Asp10, Asp15, and Thr137.[7] Similar amino acid residues from the non-specific IU-NH of *C. fasciculata* involved calcium binding with Asp10, sp15, and Thr126.[7] Two other important residue pairs implicated with reaction interactions for each species were Trp83/Trp260 in *T. vivax* and His241/Asp242 in *C. fasciculata*.[7,40] *T. vivax* used Trp83 and Trp260 on IAG-NHs used pi-pi arene stacking interactions to align ribonucleosides for hydrolysis, while *C. fasciculata* used His241/Asp242 as a proton donor with water for leaving group activation.[740] As previously mentioned earlier, Gopaul *et. al.*, 1996 mutated the His241 reside from the IU-NH enzyme of *C. fasciculata* to Ala and evaluated it as a proton donor for leaving group activation.[40] The resulting mutation of His241 caused a 2,100-fold loss in k_{cat} for inosine, and a 2.8-fold increase for k_{cat} with p-nitrophenyl-D-ribofuranoside which was used to mimic leaving group mechanics without protonation.[40] The residues described here in *T. vivax* and *C. fasciculata* were established as putative conserved amino acid residues for calcium binding and protonation during nucleoside hydrolysis establishing them as important residues to compare with other IU-NHs other organisms.[7,40] Bacteria are another group of organism observed to have nucleoside hydrolases homologous to protozoa and plant models which can be used for a comparative analysis.

1.4 Salvage Pathways in Bacteria

Nucleoside hydrolases have also been identified in bacteria and archaea isolated from

Bacillus cereus[60], *Bacillus suis*[60], *Bacillus anthracis*[61], *Bacillus thuringensis*[17], *Mycobacterium tuberculosis*[62,63], *Ochrobactrum anthropic*[6], *Sulfolobus solfataricus*[20], and *Escherichia coli*.[19,64–66] The purpose of nucleoside hydrolases in bacteria is not as well understood, but studies point to roles in nutrient scavenging, thermal stability, nucleoside regulation, and sporulation.[6,17,60–63,65,66] What is known is that similar homology and hydrolysis mechanics can be seen in the RihC enzyme of *E. coli* making it a favorable bacterial IU-NH to compare with *C. fasciculata* and *A. thaliana*.[19,64,66–69]

There are three ribonucleoside hydrolase genes isolated from the *E. coli* genome designated *ybk*, *yeik*, and *yaaf*. These genes are also known as *rihA*, *rihB*, and *rihC*, respectively.[66]The RihA and RihB hydrolases are pyrimidine-specific with an affinityforncytidine over that of uridine, while the gene encoding r*ihC* was identified as a non-specific IU-NH with an affinity for both purines and pyrimidines.[66] The RihC enzyme from *E. coli* exhibits transition-state characteristics similar to that of IU-NH in *C. fasciculata.*[19,64] RihC readily interacts with pyrimidines (cytidine and uridine), purines (adenosine and guanosine), and purine intermediates (inosine and xanthosine).[19,64]

RihC hydrolyzes purine and pyrimidine ribonucleosides in order of decreasing activity from uridine, xanthosine, inosine, adenosine, cytidine, and guanosine and exists as a dimer with a pH optimum of 7.5.[19,64] Substrate specificity experiments indicated that the 2’ and 3’ hydroxyl groups were needed for ribosyl binding to the active site whereas the 5’ hydroxyl group was not needed.[19,64] A motif found near the N-terminus consisted of a group of Asp residues with the sequence of DXDXXXDD and a His/Asp proton donor group similar to that found in *C. fasciculata* and *A. thaliana*.[19,23,64,66,67] Because of the similarity between all of these organisms, RihC also made a useful non-specific IU-NH to compare with protozoan and plant non-specific IU-NHs.

1.5 Salvage Pathways in Plants

Many plants have been extensively studied to understand metabolic pathways involving nucleosides and cytokinins in salvage processes with nucleoside hydrolases.[25] Nucleoside hydrolases have been isolated from many different species of plants including: a uridine-specific NH from mung bean (*Phaseolus radiatus*),[28] a guanosine-inosine specific NH from yellow lupin (*Lupinus luteus*)[70], and an adenosine-specific NH from coffee (*Coffea arabica*).[70–72] Non- specific nucleoside hydrolases have been isolated from Zea mays[47], *Physcomitrella patens,*[47] Alaska pea seeds,[73] and *Arabidopsis thaliana.*[23,26,29,74] *A. thaliana* has been well characterized in the literature and is an important model organism for plants with a non-specific IU-NH present for controlling salvage pathways.[75]

Non-specific IU-NHs have been implicated in controlling the ratio between nucleotide salvage and degradation in plants through many examples. In ureidic legumes, IU-NHs provided substrates needed for the synthesis of ureides such as allantoin and allantoic acid.[76] In starvation conditions, the purine bases released by nucleoside hydrolases served as a source of nitrogen.[24] Delgado-Garciá et. al. have shown that the activity of these nucleosidases increases in cotyledon and embryonic axes after radicle emergence.[77] Seeds with mutations in *URH1* activity exhibited delayed germination indicating this enzyme was essential during early phases of plant development.[23,74] Experiments performed by Jung et al. 2009 compared sequences of nucleosidases from *A. thaliana* and *Oryza sativa* (rice).[23,24] These two plants were used to obtain sequences for two key uridine ribohydrolases (URH1 and URH2) which were then expressed in recombinant yeast and bacteria.[8,21,23,24] URH1 was shown to play a very important role in regulating growth in plants and directly affected root and leaf growth, especially during the early growth phases of development.[23,24]

Nucleoside hydrolases have also been implicated in cytokinin metabolism, which are a group of plant hormones that are adenosine derivatives.[25] *Physcomitrella patens* and *Zea mays* demonstrated a metabolic function for purine, pyrimidine, and cytokinin activity with nucleoside N- ribohydrolases named PpNRH (*Physcomitrella patens*) and ZmNRH (*Zea mays*).[47] Delay in bud formation was observed in *Physcomitrella patens* with knockout plants lacking PpNRH resulting in elevated levels of purines, pyrimidines, and cytokinins.[47] Residues conserved in each of these plant species based on sequence alignments show similar reactive residues in *A. thaliana, Physcomitrella patens,* and *Zea mays*.[47] Each of these plant species all shared sequence similarity with a DXDXXXDD motif suspected of binding calcium and a residue pair which drive protonation of the nucleobase.[47] *Physcomitrella patens* and *Zea mays* were also optimal candidates for comparative analysis between other organism types with non-specific IU-NHs.

1.6 Research Goals

The goal of this research was to characterize ad describe the non-specific nucleoside hydrolase (URH1) found in *A. thaliana* to assist herbicide design for protecting economic and human health factors affecting crops. One of the primary questions we answered was how URH1 functions with regards to its state of oligomerization, equilibrium constant, pH profile, kinetic activity, substrate specificity, and calcium content. We then answered questions regarding critical amino acid residues around the active site that seemed to be impactful by comparing different organisms with nucleoside hydrolases found in putative plant and protozoa models using structural docking and molecular dynamic studies. Finally, we performed a mutagenesis study to determine if conserved residues around the active site play a key role with hydrolysis occurring with purine and pyrimidine nucleosides.

By answering these questions involving the functional and theoretical properties of URH1 we can begin to understand if these concepts apply across a broad spectrum of plant species and is a general type of salvage pathway in all plants or if it's just a specific pathway in certain species. This will allow us to understand if we can apply broad conceptualizations to herbicide design for targeting larger populations of plants to make more effective herbicides or if new ideas for herbicide design should be reimagined.

CHAPTER II: MATERIALS AND METHODS

Precast 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels were purchased from Lonza Bioscience (Rockland Inc., Rockland, ME). Supporting reagents including Laemmli sample buffer, 10X Tris- Glycine-SDS buffer, and Precision Plus Protein™ unstained protein standards were purchased from Bio-Rad Laboratories. Bio-Rad protein assay kit with bovine serum albumin as the standard was also purchased from Bio-Rad. BL21(DE3) Competent *E. cli* cells (2527H) were purchased from New England Biolabs. LongLife lysoyme was obtained from G BioSciences. Isopropyl-β-D- thiogalactopyranoside (IPTG), Zeba™ Desalting columns (MWCO 7 kD), and Gel Code Blue Safe Protein Stain were purchased from ThermoScientific. DNase I, HIS-Select™ affinity gel along with assorted nucleosides and bases were purchased from Sigma Aldrich. 1-β-D-arabinofuranosyluracil was obtained from Cayman Chemical. The Maxwell® 16 purification system and MagneHis™ polyhistidine protein purification kit were purchased from Promega**.** Gel filtration HPLC Protein Standard was purchased from CellMosaic. HPLC experiments were carried out on a Dionex Ultimate 3000 HPLC equipped with a four solvent delivery system, thermostatted autosampler, thermostatted column compartment, and variable wavelength UV/VIS detector. A Kinetex™ 5µm C18 100 Å (4.6 x 150 mm) HPLC column was obtained from Phenomenex and a Zorbax GF-250 4µm (9.4 x 250 mm) HPLC column was obtained from Agilent. VeriSpec Ca^{2+} standard solution (1,000 ppm) was purchased from Ricca Chemical Co. All other reagents were reagent grade.

2.1 Plasmid Construct

Plasmid construct work was conducted by Dr Brock Arivett from Middle Tennessee State University (MTSU). Two genes with close homology to plant non-specific IU-NHs used as inspiration for this work were from *C. fasciculata* (accession # U43371/Q27546) and *E. coli* (accession # EG11082/P22564.1).[7,19,64] BL21 (DE3) cell cultures containing a pET28b(+) expression vector with the *rihC* gene from previous studies produced by Dr Bock Arivett were used for experimental controls with URH1. The gene for URH1 found in *A. thaliana* (accession # At2g36310/Q9SJM7.2) was placed in a pET-28b(+) expression vector to produce high yield protein from BL21 (DE3) cell cultures. The company Genscript® was utilized to create a pET-28b(+) expression vector for the URH1 gene. Gene synthesis occurred via Genscript's patented phosphoramidite reaction cycle to build the sequence.[78] The URH1 expression vector included an N-terminal His_6 tag for purification with kanamycin resistance for culture selection. Restriction sites for gene insertion were *XhoI* and *NdeI* similar to *rihC* cut sites. The total gene length was 1020 base pairs (bp) and required transformation into BL21 cells once delivered to the laboratory. Information regarding the pET-28b(+) expression vector from Genscript® can be found in Appendix D.

2.2 Induction of Plasmid Construct

Dr Brock Arivett also assisted with protein expressions for quality assurance. The expression vector for URH1 from *A. thaliana* IU-NH was transformed in BL21 (DE3) competent cells for high-level protein production. A 50 µL tube of BL21 (DE3) competent cells was thawed on ice for 10 minutes. Once thawed, 5 µL of plasmid material was added to 50 µL of competent cells and mixed to yield a concentration of 100 ng/µL. The 55 µL tube of BL21 cells was kept on ice

for 30 min and then heat-shocked at 42 ^{o}C for 10 sec. After heat-shocking, the BL21 cells were placed on ice for 5 min. The contents of the starter tube were transferred to 950 µL of SOC media to make 1 mL SOC cultures and the combined solution was shaken at 250 rpm at 37 ^{o}C for 60 min. SOC media was composed of 0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose.

The 1 mL starter culture was placed on a warmed selection plate, containing 50 µg/mL kanamycin and held at 37 ^{o}C overnight. The next day colonies were selected and placed in 1 mL SOC media and grown overnight at 37 ^{o}C, while shaking at 220 rpm. Stock cultures were created by mixing overnight cultures with 1 mL sterile glycerol and placed in -80 ^{o}C freezer for long-term storage. A single stock culture was used to create small 5 mL starter cultures under antibiotic pressure with 50 µg/mL of kanamycin. Larger 1-L cultures were produced by creating four 20 mL starter cultures under antibiotic pressure with 50 µg/mL of kanamycin. Starter cultures were grown at 37 ^{o}C with shaking at 220 rpm overnight and then added to 50 mL cultures, without kanamycin the next day or four 250 mL cultures in the case of larger 1-L productions. The small batch 50 mL and large batch 250 mL cultures were grown to an OD_{600} of 0.6. Fifty (50) mL cultures were treated with 0.5 mL of 100 mM IPTG and 250 mL cultures were treated with 2.5 mL of 100 mM IPTG to reach a final concentration of 1 mM IPTG. IPTG-induced cultures were placed back on the shaker at 37 ^{o}C and shaken at 220 rpm for 3 hours. Once induction was complete, cultures were centrifuged at 9,000*xg* for 30 min at 4^{o}C creating cell pellets. Cell were either stored at -80 ^{o}C for long-term storage to be purified on a later date or were immediately purified.

2.3 Purification of His-tagged URH1

Cell pellets produced from small batch culture inductions were purified using a Promega Maxwell® 16 purification system in conjunction with a MagneHis™ polyhistidne protein purification kit. Cell pellets were resuspended in 2 mL of 100 mM HEPES buffer pH 7.5. Resuspensions were sonicated on ice 5X. Each sonication lasted 30 s with a 1 min rest period between sonication. For sonicated lysates with an OD_{600} between 4 and 20, 20 µL of 5 mg/mL DNase I were added to the 1 mL samples and inverted several times before being placed in the Maxwell purification system. Aliquots (1 mL) of the sonicated lysate were loaded into the first well of the Maxwell™ 16 protein purification cartridges. Each purification run used multiple cartridges with a maximum of 16 cartridges. A total volume of 300 µL of purified protein in 500 mM imidazole was produced by each cartridge. Buffer exchange was accomplished using Zeba™ Spin Desalting Columns (MWCO 7 kD) to remove imidazole and replace with 10 mM Tris buffer; pH 7.2. Zeba™ columns were first primed with 300 µL of 10 mM Tris buffer; pH 7.2 by centrifugation at 500 rpm for 30 s for a total of 3 times. After the columns were primed, purified enzyme preparations (300 µL) were loaded onto the column and centrifuged at 500 rpm for 30 s. The effluent from the column was collected and desalted an additional two times as described above. The eluent was collected, combined and stored at 4 ^{o}C.

Cell pellets produced from large batch culture inductions were purified by column chromatography containing His-Select™ Ni affinity gel (10 x 150 mm).[79] Cell pellets were resuspended in 10 mL of 50 mM sodium phosphate pH 8.0, 300 mM NaCl and 10 mM imidazole (Native Binding Buffer). LongLife lysozyme (10 mg) was added to cell pellets and incubated for 30 min. Resuspensions were sonicated 5X for 30 s with a 1 min rest period in the cold room at 4 ^{o}C. Lysates were centrifuged at 12000*xg* for 30 min at 4 ^{o}C to remove unwanted cell debris. The

Ni-affinity column was prepared by washing with at least 200 mL of Native Binding Buffer. Centrifuged lysate (10 mL) was loaded onto the column at 4 ^{o}C, and flow stopped for 30 min to allow binding of the protein to the resin. The column was eluted with 200 mL of Native Binding Buffer, collecting 15 mL fractions. The column was then washed with at least 200 mL of 50 mM sodium phosphate pH 8.0, 300 mM NaCl, and 20 mM imidazole (Native Wash Buffer) with a pH of 8.0. Fractions (10 mL) were collected. The column was eluted a third and final time with at least 200 mL with 50 mM sodium phosphate pH 8.0, 300 mM NaCl and 150 mM imidazole again collecting 10 mL fractions (Native Elution Buffer). Fractions produced from the binding, washing, and elution phases were analyzed for protein content by measuring the absorbance at 280 nm. Fractions were also tested for nucleoside hydrolase activity by reducing sugar assay described below. Related fractions with the highest protein concentration and activity were pooled. Pooled fractions were concentrated to 12 mL on an Amicon Ultrafiltration unit containing a regenerated cellulose filter (MWCO 5 kD).

2.4. Protein Concentration and Activity Determinations

Protein concentrations were determined in one of two ways; (1) Bradford assays using Bio-Rad Protein Assay Kit with bovine serum albumin as the standard[80] or (2) measurement of absorbance at 280 nm using the equivalency 1 AU = 1 mg/mL. Bradford assays were prepared by creating a standard assay calibration plot to compare purified protein fractions created from small or large batch productions. Each assay contained known amounts of protein standard, water, and Bio-Rad Dye concentrate. Protein standards ranged from 0-10 µg of protein in 2 µg increments with sufficient water to yield 800 µL. Bio-Rad dye concentrate (200 µL) was added to each protein assay. Water (790 µL) was added to 10 µL from pooled column fractions. Bio-Rad

dye concentrate (200 µL) was added to each protein assay. Assay mixtures were allowed to incubate for 15 min before being absorbance being measured. Each assay was measured in triplicate at 595 nm.

The activity of pooled fractions was determined using a reducing sugar assay to monitor activity colorimetrically.[1] Aliquots (100 µL) were added to a reaction mixture (1,000 µL) containing 1 mM inosine or uridine in 50 mM Tris pH 7.2. The reaction mixture was incubated for 2 hr at room temperature. The reaction was stopped by adding 250 µL copper reagent (4% Na_2CO_3, 1.6 % glycine and 0.045% $CuSO_4 \cdot 5H_20$). Neocuproine reagent (0.12% 2,4-dimethyl- 1,10-phenanthroline HCl pH 3.0; 250 µL) was added to the reaction mixture and the color developed by heating in a water bath at 95°C for 7 min. After allowing the reaction mixture to cool, its absorbance was measured at 450 nm. The amount of ribose produced was determined using a calibration curve relating absorbance at 450 nm to solutions containing known amounts of ribose.

2.5 Kinetic Measurements

Kinetic measurements were carried out in one of two ways: (1) spectrophotometrically using a temperature-controlled Shimadzu 1280 UV-Vis spectrophotometer or (2) chromatographically by HPLC on a Dionex Ultimate 3000 HPLC. Reaction mixtures assayed spectrophotometrically consisted of nucleoside (inosine or uridine) at various concentrations in 10 mM Tris pH 7.2 at 32 °C. (Total volume 990 µL). Initial velocity measurements were determined by adding 10 µL of enzyme (RihC or URH1; ~3 µg) for a total volume of 1 mL and the change in absorbance measured as a function of time. Standard plots relating absorbance to concentration of nucleoside/base (inosine/hypoxanthine and uridine/uracil) were created ranging from 10 µM to 1,000 µM of each nucleoside. Reaction velocities were determined from a linear regression of the

change in absorbance (nucleoside concentration) vs reaction time.

Each nucleoside and base concentration were measured at two wavelengths. Inosine concentrations above 50 µM were measured at 285 nm and concentrations at 50 µM and below were measured at 280 nm. Uridine concentrations above 50 µM were measured at 287 nm and concentrations at 50 µM and below were measured at 280 nm. Velocities were determined from a linear regression of the change in absorbance (nucleoside concentration) vs. reaction time. The amounts of nucleoside and nucleobase produced were determined based upon the linear regression of a standard plot relating nucleoside/base concentrations to absorbance. Samples were analyzed in triplicate and analyzed using a Michaelis-Menten plot to estimate the K_M and k_{cat}.

Activity was also determined on a Dionex Ultimate 3000 HPLC equipped with a quaternary pump, autosampler, thermostatted column compartment and variable wavelength UV-Vis detector. The reaction mixture was 950 µL of 1.05 mM nucleoside (except guanosine which was 525 µM) in 10 mM PIPES pH 6.6 to which 50 µL of enzyme (~15 µg) was added. An aliquot (10 µL) was analyzed on a Kinetex™ 5 µm C18 100 Å (150 x 4.6 mm) HPLC column from Phenomenex at various reaction times. The mobile phase was 98% 10 mM ammonium phosphate buffer pH 5.4: 2% methanol with a flow rate of 0.6 mL/min, with the exception of the analysis of adenosine and adenosine derivatives when a mobile phase of 90% 10 mM ammonium phosphate pH 5.4: 10% methanol was used. Nucleosides and bases were detected at 254 nm. The nucleoside and corresponding base were identified by their respective retention times. The amounts of unreacted nucleoside and base produced were determined based upon a standard curve relating nucleoside/base amount to peak area. Samples were analyzed in triplicate.

2.6. Subunit and Native Molecular Weight Determinations

Validation of protein purity and determination of subunit size were performed using pre-cast SDS-PAGE minigels (12%) with Precision Plus Protein™ unstained protein standards. The native molecular weight of URH1 was determined by gel filtration HPLC on the Dionex Ultimate 3000 HPLC described above. The sample was analyzed on an Agilent Zorbax GF-250 4 µm (9.4 x 250 mm) gel filtration column with a fractionation range of 4 kD to 400 kD. The mobile phase was 150 mM sodium chloride in 50 mM phosphate buffer pH 7.0 with a flow rate of 1.0 mL/min. Ten (10) µL of URH1 were injected onto the column and the effluent monitored at 218 nm. The native molecular weight of URH1 was determined using a calibration curve constructed by plotting the elution volume of proteins of known molecular weights vs the log of the protein molecular weight.

2.7. Protein Calcium Concentration Determination

Dr Paul Kline from MTSU performed calcium concentration determinations. The presence of calcium in URH1 or RihC was determined by a modification of the method of Nitta and Watanabe.[81] To minimize the background, trace metal grade water was used to prepare all reagents. URH1 or RihC (1 mL; 15.0 µM) was extensively dialyzed in 100 mL trace metal grade water 3X to remove calcium loosely associated with the protein. After dialysis was complete, the protein was transferred to 8M urea to denature it. Quin-2 (100 µM) in 8 M urea was added to the dialyzed protein solution and incubated for 30 min at room temperature. The solution was chromatographed on a BioBasic SEC-300 HPLC column (150 x 7.8 mm) eluted with water. The protein:Quin-2:calcium complex was detected using a fluorescence detector with an excitation wavelength of 325 nm and an emission wavelength of 492 nm.

2.8. Determination of Equilibrium Constant

Triplicate reaction mixtures consisted of 2 mM hypoxanthine and various concentrations of ribose from 0 to 3 M in 995 µL of 10 mM Tris pH 7.2 buffer with 0.5 mM dithiothreitol. The addition of 50 µL of URH1 (~ 15 µg) initiated the reaction. The reaction mixture was maintained at 37 ^{o}C. The concentrations of hypoxanthine and inosine were determined at various time intervals as described above by HPLC and the reaction was continued until the concentrations of inosine and hypoxanthine were constant for 24 hr.

2.9. Determination of Kinetic Mechanism

The activity of URH1 was determined by HPLC as described above. Reaction mixtures consisted of varying concentrations of inosine from 10 to 1000 µM in 10 mM PIPES pH 6.6. The reaction mixture additionally contained varying concentrations of either hypoxanthine (0-1 mM) or ribose (0-100 mM). The reaction was initiated by addition of enzyme (10 µL; ~3 µg). At appropriate times, 20 µL aliquots were withdrawn, and the reaction quenched by the addition of 20 µL 1 M HCl. The velocity of the reaction was calculated based on the decrease of the peak area representing inosine as a function of time. A double reciprocal plot of 1/v vs 1/S for each hypoxanthine or ribose concentration was prepared and the inhibition type determined by visual inspection of the plots.

2.10. pH Optimum

A series of sodium acetate, MES, CHES, PIPES, HEPES, and BICINE (300 mM) buffers covering the pH range from 3-9 were prepared. A combined buffer solution was prepared by adding 1 mL of the individual buffers to produce a common solution and adjusting the pH to between 3-9 in increments of 0.5 pH units as required by the addition NaOH or HCl. Reaction

mixtures consisting of the above buffer containing 1,000 µM inosine were prepared, and the reaction initiated by the addition of 50 µL URH1 (~15 µg) and incubated at 32 ^{o}C. After 20 min, the reaction was stopped by the addition of 100 µL 6 M HCl. The relative amounts of inosine and hypoxanthine were determined by HPLC as described above and the reaction velocity determined at each pH.

2.11. Dynamic Light Scattering to Determine Oligomerization State

Dr Paul Kline from MTSU performed dynamic light scattering experiments to determine the oligomerization state of URH1. The hydrodynamic radii of URH1 was determined at 10, 1, and 0.1 mg/mL of URH1. Protein was dialyzed against 10 mM Tris pH 7.2 buffer and diluted to the desired concentration based on the absorbance at 280 nm. The solutions were centrifuged at 12,000*xg* for 5 min at room temperature to remove any particulates. Samples (1,500 µL) were analyzed on a Zetasizer Nano ZS (Malvern Panalytical, Westborough, MA). The instrument was set at a size detection range of 0.3 nm to 10 µm.

2.12 Molecular Docking and Multiple Sequence Alignment

Protein structural files for the nucleoside hydrolases from two protozoa and two plant species were retrieved from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB).[82,83] Sequences for the non-specific CfIU-NH from *Crithidia fasciculata* (PDB ID: 1MAS) (accession # U43371/Q27546) and the purine specific TvIAG-NH from *Trypanosoma vivax* (PDB ID: 2FF1) (accession # AF311701.2/Q9GPQ4) were retrieved as representatives of protozoan nucleoside hydrolases. Similarly, sequences for the nucleoside N-ribohydrolases from *Physcomitrium patens* (PpNRH) (PDB ID: 4KPN) (accession # JQ649322.1 / A9TXA6) and *Zea*

mays (ZmNRH) (PDB ID: 4KPO) (accession # HQ825162.1 / B6T563) were used as representatives of plant nucleoside hydrolases. No structural models existed for URH1 in *Arabidopsis thaliana* and RihC in *E. coli* on the PDB, but FASTA protein sequence files were available from the National Center for Biotechnology Information (NCBI) website using accession #P22564.1 for *E. coli* and accession #Q9SJM7.2 for *A. thaliana*.[23]

The sequence file obtained from NCBI for URH1 was used for prediction of a three-dimensional structure using the SWISS-MODEL website server.[68,84–92] SWISS-MODEL utilizes homology modeling to predict the three-dimensional structure of an input protein based on its sequence and the structure(s) of other homologous proteins with known crystal structures.[68,85,88,91] In addition, the URH1 structural file was also built using the AlphaFold Protein Structural Database (AlphaFoldDB) version 2, and compared to the SWISS- MODEL structure.[93–95] A pairwise alignment of each structure yielded a 100% match using the EMBL-EBI (European Molecular Biology Laboratory-European Bioinformatics Institute) EMBOSS (European Molecular Biology Open Software Suite) Needle pairwise sequence alignment (PSA) tool.[96,97] A pairwise sequence analysis was also performed between the SWISS-MODEL URH1 sequence and each of the other crystal structures retrieved from the PDB website for a similarity match. A multiple sequence alignment (MSA) was performed using Clustal Omega

from EMBOSS for all six protein sequences retrieved from the PDB and NCBI websites to produce a comparative analysis of each sequence.[96]

Structural files for inosine and uridine were retrieved from the PubChem database as structured data files (SDF) and used for ligand docking.[98] SDF files were converted to PDB files using the UCSF (University of California San Francisco) Chimera software (version 1.15).[99,100] Molecular docking was separately performed with inosine and uridine using AutoDock Vina

embedded in the Chimera software package (version 1.15).[99–101] Inosine and uridine were independently docked to nucleoside hydrolases from *C. fasciculata* (PDB ID: 1MAS), *T. vivax* (PDB ID: 2FF1), *P. patens* (PDB ID: 4KPN), *Z. mays* (PDB ID: 4KPO), and the SWISS-MODEL of URH1. Inosine docks were used to produce active site images for each structure, while both inosine and uridine docks were used for molecular dynamics simulations. A 5 Å zone of all residues around inosine were mapped with each of the four crystal structural files and then compared to the homology file for URH1 to develop a visual sense of the active site residues that potentially play a key role in purine hydrolysis.

2.13 Molecular Dynamics Simulations

Dr Chad Brambley from MTSU provided recommendations and assistance developing molecular dynamic simulations. GROMACS (version 2021.7) software package was used to perform molecular dynamics (MD) simulations using nucleoside hydrolase models with inosine or uridine docked alongside the apo and holo states of URH1.[102–106] All simulation parameters and settings were performed according to The GROMACS protein-ligand complex tutorial from the website http://www.mdtutorials.com/gmx/complex/index.html .[107]

GROMACS converted the URH1 PDB files to GMX files using the pdb2gmx command. Coordinate files for the docked ligand files were created using the SwissParam small molecule force field generation tool.[92] Simulations used the Charmm-36 all atom force field and TIP3P water model for the topology parameters.[108–111] Each homology model and ligand was solvated with water in a rhombic dodecahedral box with neutralizing ions present. Solvated systems were subjected to energy minimization using a steepest descent protocol, which allows for system relaxation necessary to avoid inappropriate geometry and steric clashes. Each minimized structure

was thermally equilibrated for 100 ps as position restrained NVT and NPT ensembles using a Berendsen thermostat. Preliminary unrestrained MD simulations were then performed for 10 ns using Parinello-Rahman temperature and pressure coupling, respectively, to test the system and evaluate ligand stability for each homology model. After initial evaluation of each MD simulation, 200 ns production MD runs were performed for each model. Root-mean-square fluctuation (RMSF) plots were created to evaluate each simulation.

XMGrace was used to view data and exported to excel.[112] Visual Molecular Dynamics (VMD) (version 1.9.3) was utilized to visualize ligand-residue interactions occurring during each production MD simulation.[113] Trajectory and PDB files produced from GROMACS and VMD were used to perform principal component analysis (PCA) with each 200 ns simulation for the apo and holo states of URH1 as well as with inosine and uridine docked states of URH1. Bio3D version 2.3 data software package was used to generate Scree plots for each PCA model.[114–117] Gibbs-free energy (GFE) landscapes were produced using sham input from GROMACS (version 2021.7) MD simulations.[102–106]

2.14 Site-Directed Mutagenesis

Wild Type Plasmid. The gene for URH1 found in *A. thaliana* (accession # At2g36310 / Q9SJM7) was incorporated into an optimized pET-28b(+) plasmid by the company Genscript®.[118,119] The expression vector of the URH1 gene was then placed in BL21 (DE3) cells for transformation. The restriction sites for NdeI and XhoI were used to insert the *URH1* gene with an N-terminal His6 tag for purification. Total gene length was 1020 base pairs.

Mutant Construct. A site-directed mutagenesis was performed on the URH1 plasmid. The mutagenesis kit chosen to make mutations for these residues was the QuikChange II Site-Directed

Mutagenesis Kit purchased from Agilent Technologies (Santa Clara, CA). Primers used for mutations with the QuikChange II Kit were purchased from Invitrogen™. Primer design was performed using the Agilent's web-based QuikChange Primer Design Program (www.agilent.com/genomics/qcpd). A list of the primers used for the mutagenesis study can be found in Table 1.

Mutant Synthesis Protocol. The mutagenesis kit used three steps to produce mutations in the wild-type plasmid. The first step was a mutant strand synthesis using PCR techniques to denature, anneal, and extend mutant primers with *PfuUltra* DNA polymerase. The next step used DpnI to digest methylated, parental template and destroy any older plasmids without the mutation. The final step transformed mutated plasmids into BL21 cells. All primers purchased from Invitrogen ranged from 20-32 base pairs in length.

Mutant strand synthesis reaction was started by preparing reaction aliquots for PCR. Each aliquot contained 5 µL of 10x reaction buffer, 1.5 µL (25 ng) of dsDNA template (plasmid), 1 µL of oligonucleotide primer #1, 1 µL of oligonucleotide primer #2, 1 µL of dNTP mix, and ddH_2O to a final volume of 50 µL. One (1) µL of *PfuUltra* HF DNA polymerase (2.5 U/mL) was added to begin reactions. PCR reaction mixes were placed on a thermocycler to replicate plasmids. The reaction mixtures were run for 16 cycles and peaked at 95 ^oC followed by 55 ^oC for 1 minute and then followed by a 68 ^oC hold with the time dependent on the kb length of the plasmid. The pET-28b plasmid with the URH1 gene was approximately 5.4 kb long and was set to cycle at 68 ^oC for 6 mins. Reaction mixtures were then placed on a Bio-Rad Thermocycler to replicate reaction mixtures. The reaction mixture was then placed on ice and allowed to cool for 2 minutes at $\leq$37 ^oC.

After the creation of mutant strand templates using PCR, digestion of parental strands was performed. DpnI was used to digest the methylated parental templates leaving mutant template

behind. Each aliquot received 1 µL of DpnI (10 U/mL) and was mixed by pipetting solutions up and down several times. The DpnI restriction enzyme was then added to each amplification reaction mixture to digest the methylated parental DNA plasmid. Amplification reaction mixtures were gently mixed with a pipette and then quickly centrifuged for 10 secs. Amplification reaction mixtures were then allowed to digest at 37 ^{o}C for 1 hour. Following digestion all parental strand (non-mutated DNA) was destroyed leaving behind mutant plasmids which were used to transform new BL21 cells.

Transformations were used to validate plasmid mutations. A quick screen of each plasmid was performed via a transformation of XL1-Blue supercompetent cells. XL-1 Blue cells are extremely sensitive to temperature variations and must always be kept on ice. XL1-Blue supercompetent cells were gently thawed from frozen at -80 ^{o}C to room temperature while on ice. Twenty µL of supercompetent cell were placed in a 14 mL BD Falcon polypropylene round-bottom tube and mixed with 1 mL of the DpnI digested mutant plasmids. Reaction mixtures were gently mixed and incubated on ice for 30 minutes. Reaction mixtures were then heat pulsed at 42 ^{o}C for 45 seconds and placed on ice for 2 minutes. Reaction mixtures were added to 500 mL of SOC outgrowth media, preheated to 42 ^{o}C and incubated at 37 ^{o}C for 1 hour with shaking at 220 rpm. The heat pulsed SOC reaction mixtures were placed on a shaker for 1 hour and held at 37 ^{o}C with shaking at 225 rpm. Transformation mixture (100 mL) was then added to agar plates. Each plate contained 80 µL/mL of 5-bromo-4-chloro-3-indole-b-D- galactopyranoside (X-gal), 20 mM IPTG, and 50 µg/mL of kanamycin. Transformation mixtures were spread on plates and incubated at 37 ^{o}C for > 16 hours. Plate results ranged from 50-800 colonies with > 80% of the colonies appearing blue. Mutations met these requirements and validated successful mutations for each residue. A stock of each mutant passing the plate screen was grown in LB media with 50 µg/mL

kanamycin. Approximately 2 mL of each stock was placed in 10% sterile glycerol and stored long term in the -80 ^{o}C storage freezer.

2.15 Mutant Sequence Validations

Mutant sequence validations were conducted by Dr. Dan Braynt from MTSU. The wild type and mutant DNA plasmid sequences were validated using a SeqStudio Genetic Analyzer System with SmartStart from ThermoFisher Scientific. Each plasmid was sequenced using the BigDye Terminator V3.1 kit according to the manufacturer's instructions. Sanger sequencing data was analyzed using the Geneous Prime^{TM} software. Two primer locations were targeted on the plasmid. The T7 promoter and terminator sites were the outermost sites targeted with primers for a left and right sequence validation. Primer sequences can be found in Table 1. Validation results for each plasmid are explained in the Chapter V results section.

Table 1. Oligonucleotides used to generate mutant alleles.

Primer	Length (nt)	Melt Temperature (ºC)	Direction	Sequence
T7 - Terminator	20	61	Forward	5'-AAA CGG TCT GGG TGA TGT GA-3'
T7 - Promoter	20	61	Reverse	5'-CTT GCA GTT ACC CAG CTC CA-3'
Asp-27-Ala	32	80	Forward	5'-ACG AGA AAC TGA TTA TTG CCA CCG ACC CGG GC-3'
			Reverse	5'-GCC CGG GTC GGT GGC AAT AAT CAG TTT CTC GT-3'
Asp-29-Ala	29	81	Forward	5'-CGT CGA TGC CCG GGG CGG TGT CAA TAA TC-3'
			Reverse	5'-GAT TAT TGA CAC CGC CCC GGG CAT CGA CG-3'
Asp-33-Ala	27	83	Forward	5'-CGC CAT GCT ATC GGC GAT GCC CGG CTC-3'
			Reverse	5'-GAC CCG GGC ATC GCC GAT AGC ATG GCG-3'
Asp-34-Ala	29	83	Forward	5'-GAA TCG CCA TGC TAG CGT CGA TGC CCG GG-3'
			Reverse	5'-CCC GGG CAT CGA CGC TAG CAT GGC GAT TC-3'
His-260-Ala	34	81	Forward	5'-GAA GCT CAC CGG ATC GGC CAG ATA AAC GCC GTA C-3'
			Reverse	5'-GTA CGG CGT TTA TCT GGC CGA TCC GGT GAG CTT C-3'
Asp-261-Ala	31	79	Forward	5'-ACG AAG CTC ACC GGA GCG TGC AGA TAA ACG C-3'
			Reverse	5'-GCG TTT ATC TGC ACG CTC CGG TGA GCT TCG T-3'

CHAPTER III: CHARACTERIZATION RESULTS AND DISCUSSION

3.1 Purification and Characterization of RihC and URH1

To better understand the biochemical properties of non-specific nucleoside hydrolases URH1 was expressed and extracted from cultures and characterized to understand its function. Since RihC was available for culturing, we used it as a purification process control while producing URH1. Both *URH1* found in *A. thaliana* (accession # At2g36310) and *rihC* from *E. coli* (accession # EG11082) were expressed with similar pET28b plasmids and placed in cultures for induction and purification.[19,64,66] The *URH1* gene constructed by Genscript™ was 1020 bp which was ~336 amino acids in length and contained an additional N-terminal His_6-tag for purification. A Promega Maxwell® 16 Extraction system was successfully employed to purify both RihC from *E. coli* and URH1 from *A. thaliana* for small batches of enzyme. Each cartridge from the Maxwell™ 16 produced ~300 µL of total purified enzyme for URH1 and RihC which were validated using SDS-PAGE (Figure 4).

Molecular weight determinations help verify the proper protein was produced in cultures. The theoretical molecular weights for RihC and URH1 were calculated using the Expasy website based on the respective amino acid sequences and determined to be ~32 kD for RihC and ~36 kD for URH1.[120] Based on the distance RihC and URH1 traveled on the SDS-PAGE gel, the molecular weights of the enzymes RihC and URH1were determined to be ~33 kD and ~37 kD respectively. The extra weight was attributed to the His_6-tag at the end of the sequence. Inspection of the SDS-PAGE gel revealed the purity of both enzymes to be greater than 95%. Concentration data was also produced to assist with other parameter analyses. Purified enzyme concentrations were determined to be ~0.35 µg/µL (~ 100 µg) of URH1 and ~0.38 µg/µL (~ 110 µg) of RihC.35

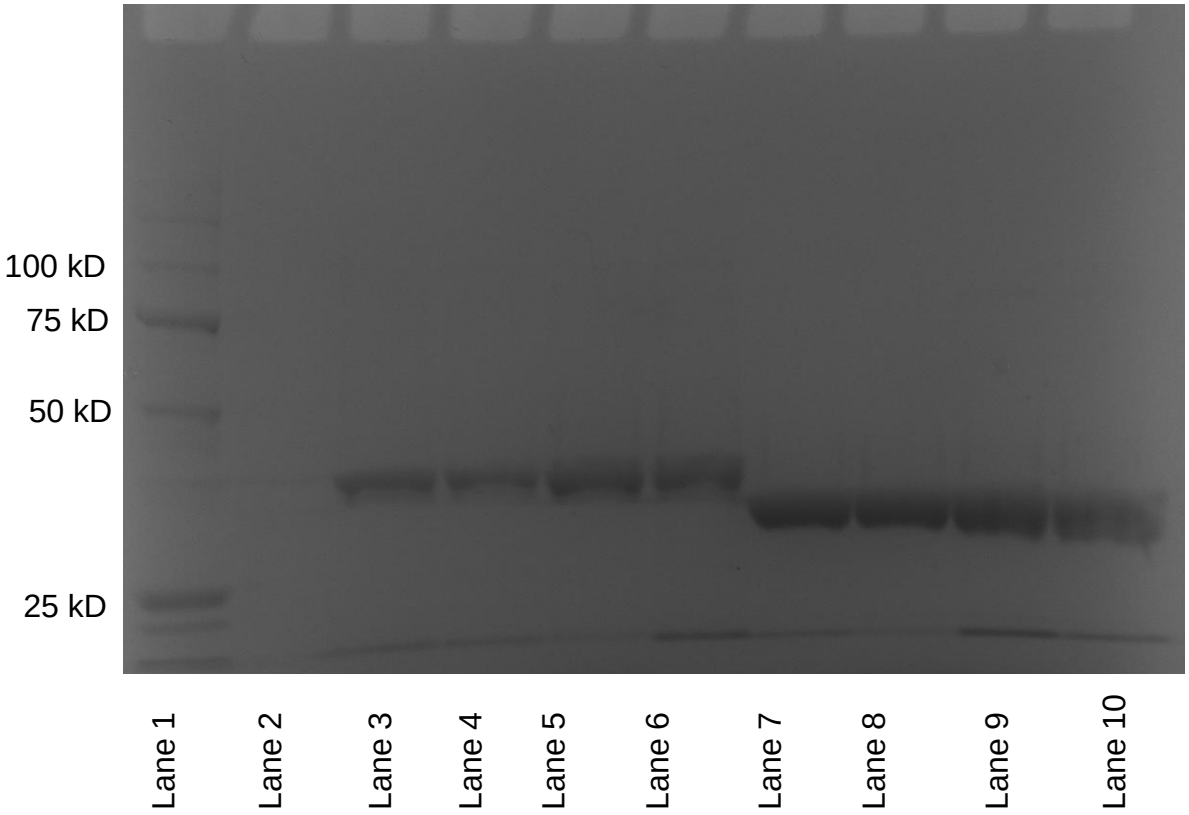


Figure 4. SDS-PAGE analysis of RihC and URH1. Final protein preparation from the Maxwell® 16 purification system for RihC and URH1. From left to right, lane 1 Bio-Rad Precision Plus Protein™ unstained protein standard; lanes 3-6 URH1 purifications; lanes 7-10 RihC. Visual inspections of the lanes containing RihC and URH1 revealed a purity greater than 95%. Based on the calibration curve using the Bio-Rad Precision Pus Protein standards the subunit molecular weight of URH1 was 37 kD and RihC was 33 kD.

The native molecular weight of a protein is a function of the subunit molecular weight and the number of subunits that aggregate. The native molecular weight of URH1 was determined using size exclusion chromatography. The native molecular weight assists researchers with developing better technique applications for separations and interactions for binding. Based on size exclusion chromatography data performed by Dr Paul Kline, the apparent native molecular weight was determined to be ~78.3 kD**.** It was determined from these experiments that the oligomerization state of the enzyme was a homodimer consisting of two identical subunits. Nucleoside hydrolases from baker's yeast and *Mycobacterium tuberculosis*[63,121] have been observed to be monomers, while protozoan nucleoside hydrolases have exhibited several different oligomerization states including dimeric, trimeric, and tetrameric.[1,4,14,14,52,122] *C. fasciculata* and *Leishmania* produced inosine-uridine nucleoside hydrolases which form tetramers, while the purine-specific inosine-guanosine nucleoside hydrolase from *C. fasciculata* form a trimer. In contrast, all plant nucleoside hydrolases with a determined crystal structure have been observed to be dimers**.**[3]

Oligomerization states for nucleoside hydrolases in different species may be dependent on several factors. Recombinant nucleoside hydrolases from *Mycobacterium tuberculosis* H37Rv were found to have different oligomerization states based on the concentration of hydrolysis products ribose, hypoxanthine, adenine, guanine, xanthine, and uracil.[63] The enzyme concentration may also affect the oligomerization state of nucleoside hydrolases in addition to the effects of hydrolysis products. Four subunits were described from preliminary crystal structures of RihC from *E. coli* which revealed two well-defined subunits and two subunits not as well-defined**.**[19,66]

Dr Paul Kline from MTSU performed measurements of the hydrodynamic radius to determine the oligomerization state of URH1 to determine if protein concentrations were important

to the oligomerization state of URH1. The hydrodynamic radii were measured using dynamic light scattering (DLS) as a function of protein concentration. Dynamic light scattering can measure protein structure by analyzing the diffusion behavior of proteins. Protein solutions with concentrations at 10 mg/mL, 1mg/mL, and 0.1 mg/mL were measured in a Malvern Zetasizer Pro dynamic light scattering system. The hydrodynamic radius of a protein depends on the size and shape of the protein. For globular proteins, there is a strong correlation between the hydrodynamic radius and molecular weight,[123,124] therefore, determination of the hydrodynamic radius allows study of the oligomerization state of a protein.

The state of oligomerization/aggregation displayed by URH1 was complex with multiple aggregates present at different concentrations (Figure 5). Four different enzyme subunits were observed in 10 mg/mL (red) of URH1. The hydrodynamic radius ranged from ~0.1-10 μm as major peaks for particle diameters with less than 20% abundance for each peak. As concentrations decreased by a factor of 10, less subunit aggregates were seen with smaller particle diameters, but with greater abundance. The hydrodynamic radius in a 1:10 dilution (blue) at 1 mg/mL of URH1 resulted in only two major subunits with particle diameters ~0.1-1 nm with 20-40% abundance while a 1:100 dilution (green) at 0.1 mg/mL resulted only one subunit with a major peak at ~0.1 nm, but with 65% abundance. When a single subunit for URH1 was measured computationally using the software Chimera (version 1.15) it was ~6.7 nm in diameter. Compared with the oligomerization data collected, URH1 potentially has a monomer/dimer type subunit. Understanding the behavior displayed by URH1 may have implications in future purification of this protein, since size exclusion chromatography has been used as a final polishing step in other protein purifications.[125]

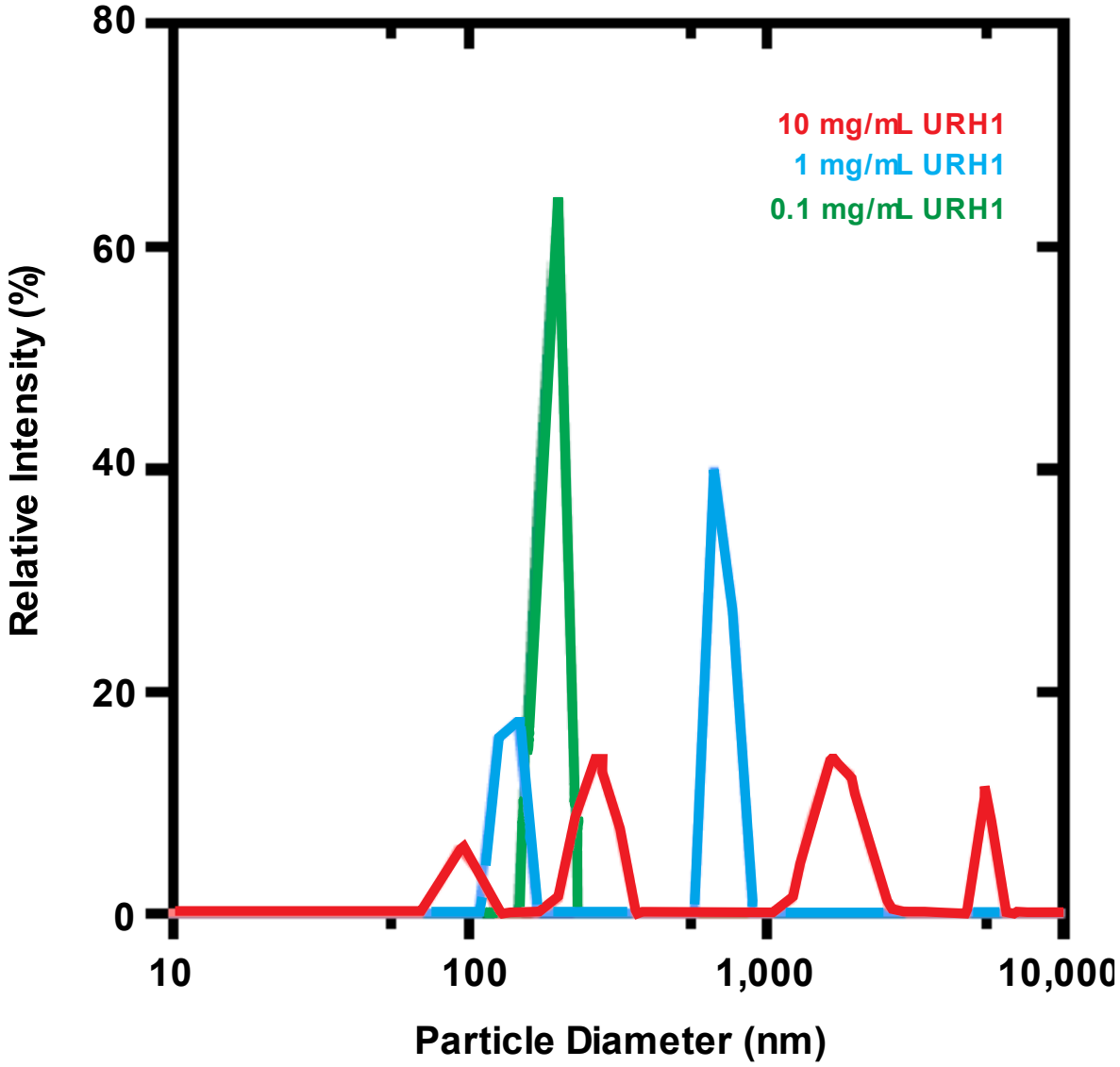


Figure 5. Hydrodynamic Size Distribution. Hydrodynamic radii were determined using dynamic light scattering with protein concentrations at 10 mg/mL, 1 mg/mL, and 0.1 mg/mL. Four different states of oligomerization (red) were observed at 10 mg/mL URH1. As concentrations were lowered by factors of 10 (blue) and 100 (green) only two states of oligomerization and one state of oligomerization were observed respectively.

3.2 Calcium Content of URH1

Dr Paul Kline from MTSU performed calcium measurements in URH1 to determine if calcium was present. Determining the calcium content in URH1 is essential to understanding how it regulates activity, substrate binding, and active site stability. The calcium ion in *C. fasciculata* interacts with Asp10, Asp15, and Thr126 creating a stable active site allowing a water molecule to interact with His241/Asp242 and cleave the glycosidic bond of a nucleoside.[1] The presence of calcium decreases the pK_a of the attacking water molecule allowing Asp-10 to accept a proton from the water molecule, thereby increasing the nucleophilicity of the activated water molecule.[7] This activated water molecule acting as a nucleophile attacks C1' of the ribose moiety, generating an oxocarbenium transition state.[7] Calcium is very predominant in plant nucleoside hydrolases such as a nucleoside N-ribohydrolases (NRH) isolated from maize, known as *Zea mays* and *Physcomitrella patens*, known as spreading earthmoss.[47] In *Physcomitrella patens* and *Zea mays*, calcium is coordinated by several aspartic acid residues in a manner similar to the DXDXXXDD motif reported for the parasitic protozoan enzymes.

To determine if calcium was present in URH1, the purified protein was treated with Quin-2, a fluorescent indicator that binds with a high affinity for calcium compared to other divalent cations such as magnesium, 2.9 ± 0.2 nM and 89 ± 5 µM respectively.[126,127] As can be seen based on computer modeling, the calcium ion is buried deeply in the active site.[47] To ensure that the Quin-2 reagent had physical access to the calcium ion, the protein was denatured by treatment with 8M urea. Size exclusion chromatography was used to separate the protein:Ca^{2+}:Quin-2 complex from other components present in the reaction mixture. The fluorescence chromatogram of URH1 shows A) the intensity of the fluorescence without Quin-2 and B) the intensity of the fluorescence of the protein after treatment with Quin-2 and 8M urea (Figure 6). The fluorescent intensity of

URH1 in the presence of Quin-2 is approximately 50 times the intensity of the protein alone which is consistent with the presence of a protein-bound calcium.

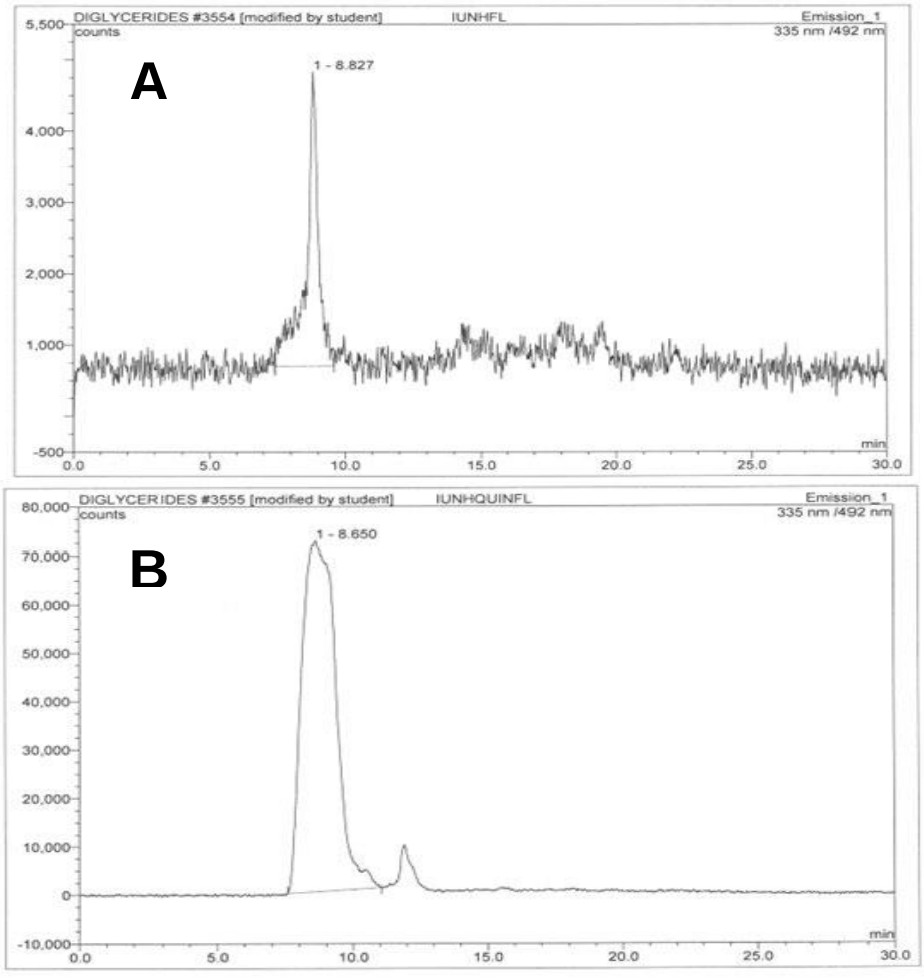


Figure 6. Calcium Determination in URH1. A) Fluorescence chromatograph of URH1 without the addition of Quin-2 fluorescence reagent or urea. B) Fluorescence chromatograph of URH1 in 8 M urea and 100 µM of Quin-2. The excitation wavelength was 335 nm and the emission wavelength was 492 nm.

3.3 Equilibrium Constant Determination with HPLC

The equilibrium constant for enzymes is important to understand to predict the relationship between reactants and products. The equilibrium constant for nucleoside formation in the presence of high concentrations of base and ribose were determined for URH1. Because of the high concentration of water in the reaction mixture compared to the other substrate concentrations, the equilibrium constant was measured in the direction of nucleoside formation, rather than in the direction of nucleoside hydrolysis. The equilibrium constants for nucleoside formation in the presence of nucleoside hydrolases previously determined were relatively constant, ranging from 45 ± 2 M for RihC and 106 ± 16 M for CfIU-NH, and to 263 ± 19 M for a nucleoside hydrolase from yellow lupin.[1,27,41] The reaction was initiated by the addition of ~ 15 µg URH1 to reaction mixtures containing 2 mM hypoxanthine and varying concentrations of ribose from 0 to 3 M and incubating at 37 ^{o}C until the amounts of inosine and hypoxanthine remained constant for a minimum of 24 hr as determined by HPLC. No inosine was detected after 120 hr indicating the reaction was essentially irreversible under these conditions and therefore at equilibrium.

3.4 Kinetic Mechanism

The kinetic mechanism for URH1 was essential to understand substrate interactions. The kinetic mechanism for URH1 was determined by examination of the product inhibition patterns for ribose and hypoxanthine. Both ribose and hypoxanthine exhibited competitive inhibition patterns that were consistent with a rapid equilibrium random mechanism (Figure 7). A similar mechanism was reported for AMP nucleosidase from *Azotobacter vinelandii*.[128] These experiments resulted in a simple ordered release of adenine followed by ribose-5-phosphate in

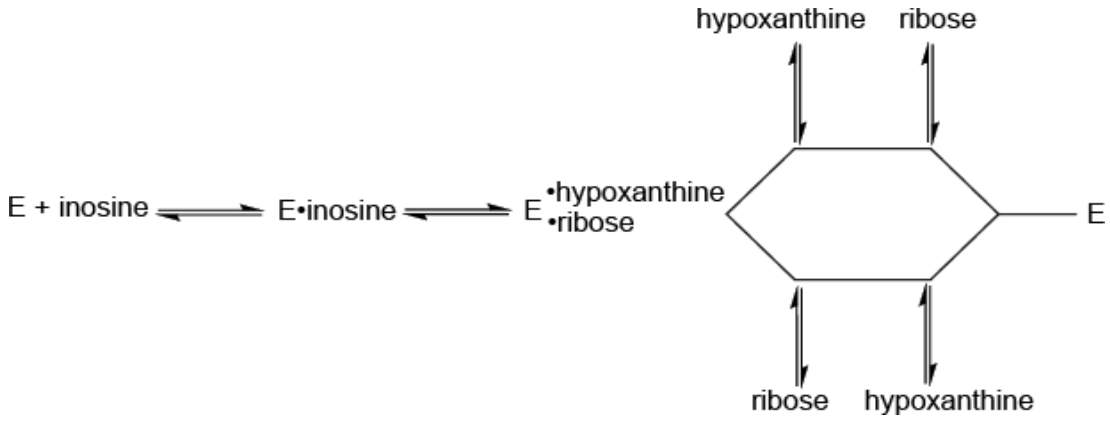


Figure 7. Kinetic mechanism for URH1. Based on competitive inhibition shown by ribose and hypoxanthine, the mechanism is consistent with a rapid equilibrium random kinetic mechanism. The mechanism is consistent with the low substrate specificity of the nucleobase and the high specificity for the ribose moiety.

rapid equilibrium. In contrast, however, AMP nucleosidase resulted in a dead-end complex of the enzyme, AMP, and adenine while our experiments resulted in rapid equilibrium between enzyme, inosine, ribose, and hypoxanthine.

The major barrier in the enzyme-catalyzed reaction is the breaking of the N-glycosidic bond in the enzyme-substrate complex rather than the binding of the substrates or release of the products. This is further supported by kinetic isotope effects which showed that the formation of the ribooxycarbenium transition state is the largest energetic barrier.[41] The low inhibition constant for ribose compared to that of hypoxanthine is consistent with the broad substrate specificity of the enzyme in which nucleosides with either purine or pyrimidine are substrates, while at the same time there is a strict requirement for ribose as the pentose moiety.

3.5. pH Optimum

Understanding pH interactions for enzymes are critical for comprehending residue interactions and active site function. The optimal pH provides insight into the necessary conditions for creating charged interactions and hydrogen bonding that are necessary to allow a protein to fold and maintain its structural stability for optimal function. The activity of URH1 was determined as a function of pH from 3- 9 using inosine as the substrate. The enzyme functioned over a wide pH range retaining 80% of its activity at pH 5 and 60% of its activity at pH 9 (Figure 8). The activity of the enzyme remained relatively constant from a pH 6 to 8. Below a pH of 3 and above a pH of 9, the enzyme did not exhibit significant activity. Overall, this behavior is similar to other nucleoside hydrolases which also functioned over a wide pH range. For example, a non-specific nucleoside hydrolase from Alaska pea seeds resulted a bell-shaped activity profile with a pH optimum at 6.4.[129]

3.6 Substrate Specificity

Nucleosides have many derivatives with different biological functions varying from their natural analogues. Nucleoside hydrolases have also been classified in different groups based on substrate analogues they break down. Group I nucleoside hydrolases are nonspecific nucleoside hydrolases which preferentially hydrolyze inosine and uridine (IU-NH), Group II nucleoside hydrolases are purine-specific (IAG-NH), Group III are 6-oxopurine nucleoside hydrolases (GI-NH), and Group IV are pyrimidine-specific nucleoside hydrolases (CU-NH).[7] Structure-activity relationships were determined for a number of different nucleoside analogues to evaluate the importance of different moieties of the nucleoside in controlling the ability of URH1 to hydrolyze the N-glycosidic bond. The ribonucleosides inosine, adenosine, guanosine,

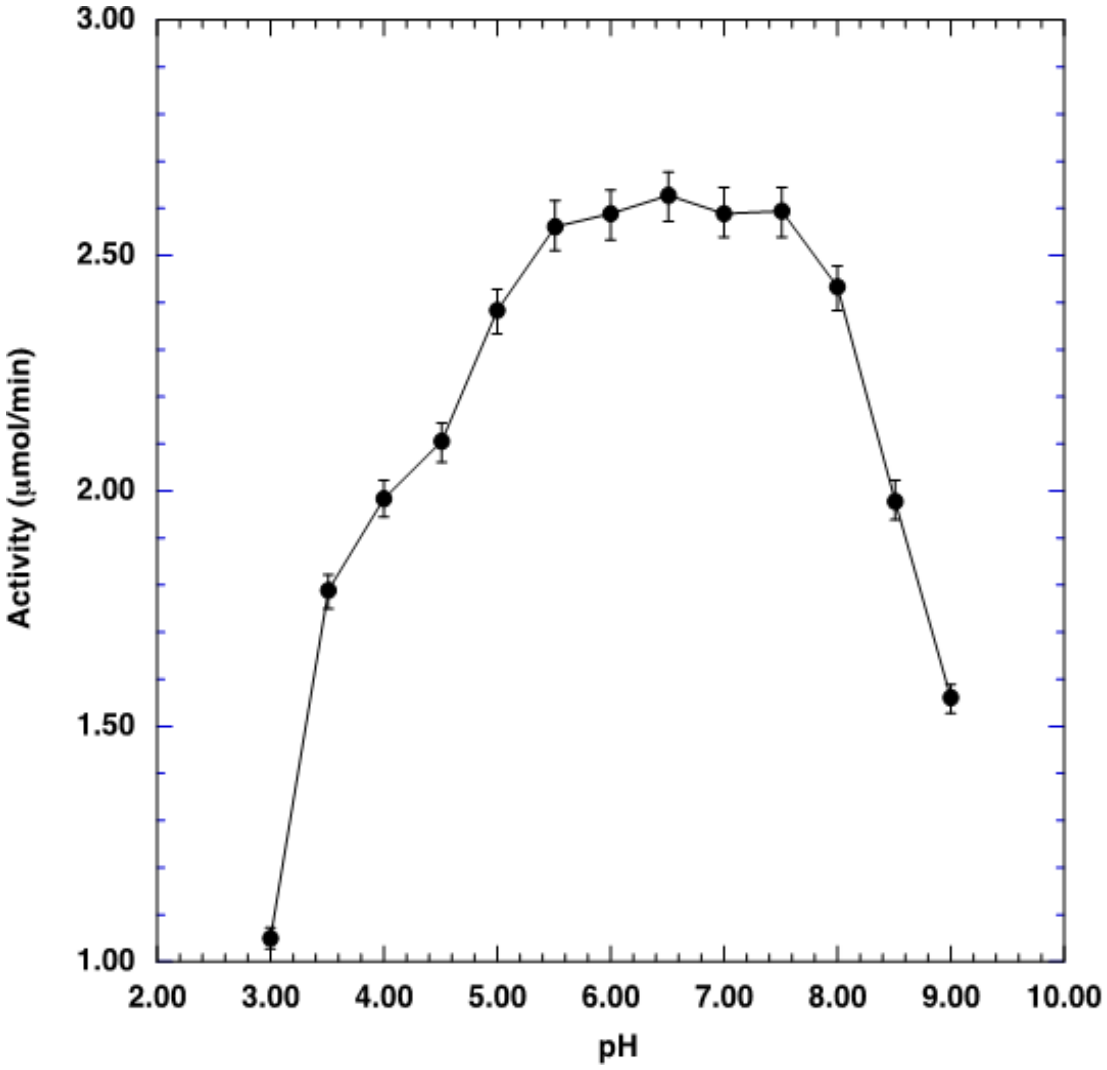


Figure 8. Effect of pH on activity of URH1. Using inosine (1 mM) as the substrate. The pH was maintained with a buffer consisting of sodium acetate, MES, CHES, PIPES, HEPES, and BICINE (300 mM) adjusted to the desired pH with HCl or NaOH. The enzyme maintained its activity over a wide pH range. Activity was essentially constant from a pH of 6 to 8 with a sharp decrease in activity above a pH of 8.

nebularine (purine riboside), xanthosine, uridine, 5-methylurdine, and cytidine were substrates for URH1 in the order of lower to higher activity: 5-methyluridine (ribothymidine) > xanthosine > uridine > adenosine > cytidine > inosine > nebularine > guanosine (Table 2 & 3). Based on the observed activities, URH1 is a non-specific nucleoside hydrolase for which both purine and pyrimidine ribonucleosides were substrates.

Based on this classification, URH1 belongs to the Group I non-specific nucleoside hydrolases exemplified by IU-NH isolated from *C. fasciculata*, although the relatively high activity with adenosine is unusual for this class of enzyme. The relatively high activity associated with the hydrolysis of adenosine may indicate this enzyme has a role in the synthesis and degradation of cytokinins, a group of phytohormones that are derivatives of adenosine. URH1 may directly control the amount of cytokinins present by controlling the rate at which these phytohormones are degraded or it may control the amount of adenosine available for the synthesis of cytokinins.

The 2'-deoxyribonucleosides, 2'-deoxyuridine, 2'-deoxyadenosine, 2'- deoxyinosine, 2'-deoxycytidine and thymidine exhibited no activity indicating the presence of the 2'-OH group is essential for catalytic activity. This is consistent with other nucleoside hydrolases which have resulted in a stringent specificity for the ribose moiety. The importance of -OH groups at other positions on the pentose moiety was studied by determining the activity of URH1 with cordycepin (3'-deoxyadenosine), 5'-deoxyadenosine, 2'-amino-2'-deoxyinosine, 9-β-D-arabinofuranosyladenine (Vidarabine), 1-β-D-arabinofuranosyluracil, and 1-β-D- erythrouridine. No activity was observed for any of the nucleosides in which one of the hydroxyl groups on the ribose moiety has been removed including 3'-deoxyadenosine, 1-β-D-erythrouridine and

Table 2. URH1 Purine Substrate Specificity.

Purine Substrates (1mM)	Velocity (μmol/min)	% Relative Velocity based on Inosine
Guanosine (0.5 mM)	1.280 ± 0.026	129
Nebularine	1.010 ± 0.015	102
Inosine	0.991 ± 0.017	100
Adenosine	0.952 ± 0.009	96
6-γ-γ-Dimethylallylamine purine riboside (Cytokinin)	0.550 ± 0.013	56
Xanthosine	0.448 ± 0.014	45
Kinetin riboside (Cytokinin)	*NM	-
Zeatin riboside (Cytokinin)	*NM	-
6-Benzylaminopurine riboside	*NM	-
5'-Deoxy-5'-methylthioadenosine	*NM	-
5'-Deoxyadenosine	*NM	-
2'-Deoxyadenosine	*NM	-
2'-Deoxycytidine	*NM	-
2'-Deoxyinosine	*NM	-
2'-Amino-2'-deoxyinosine	*NM	-
6-Chloropurine riboside	*NM	-
7-Methylguanosine	*NM	-
9-β-D-Arabinofuranysladenine (Vidarabine)	*NM	-
3'-Deoxyadenosine (Cordycepin)	*NM	-
7-Deazadenosine (Tubercidin)	*NM	-

*NM – Not metabolized

See Appendices A, B, & C for structural details of substrate

Table 3. URH1 Pyrimidine Substrate Specificity.

Pyrimidine Substrates (1mM)	Velocity (μmol/min)	% Relative Velocity based on Inosine
Uridine	0.451 ± 0.011	46
Cytidine	0.979 ± 0.014	99
5-Methyluridine	0.236 ± 0.005	32
2'-Deoxyuridine	*NM	-
Thymidine	*NM	-
3-Deazauridine	*NM	-
1-β-Erythrouridine	*NM	-
1-β-Arabinofuranosyluracil	*NM	-

*NM – Not metabolized

See Appendices A, B, & C for structural details of substrate

5’-deoxyadenosine, or has been altered including, 2’-amino-2’-deoxyinosine, 9-β-arabinofuranosyladenine and 1-β-D-arabinofuranosyluracil. In 9-β-arabinofuranosyladenine and 1-β-D-arabinofuranosyluracil, the 2’-OH group is present, but its stereochemistry has been inverted preventing the normal interaction between the calcium ion and the 2’-OH group. In 1-β-D-erythrouridine, the 5’-hydroxymethyl group (-CH_2OH) has been replaced with a hydrogen at C4’ and in 5’-deoxyadenosine, the 5’-OH group has been removed and replaced with a hydrogen yielding a methyl group at C5’. The 5’-OH group had previously been found to contribute greatly to the hydrolysis of the glycosidic bond.[44] The 5’-OH group participates in locking the C4’-C5’ bond in a conformation that allows the 5’- hydroxyl group to form a hydrogen bond with C8 of the purine group. *Ab initio* quantum chemical calculations indicate this hydrogen bond facilitates the protonation of N7, an essential step in the hydrolysis of purine nucleosides. For

pyrimidine nucleosides, the hydrolysis of 1-β-D-erythrouridine with URH1 was not observed over a 24 hr time period and in RihC was 500 times lower than the rate of uridine hydrolysis. As with other nucleoside hydrolases, URH1 has a stringent requirement for ribose. When bound to a nucleoside hydrolase, the ribose moiety is surrounded by a group of conserved residues that tightly bind the ribose moiety through the hydroxyl groups to correctly position the nucleoside. In addition to the importance of the structural aspects of the pentose ring, the effect of manipulating the nitrogenous base was also investigated. 7- Deazaadenosine (Tubercidin) in which the N7 of purine ring has been replaced with a carbon, was not a substrate. This is consistent with the results obtained with other nucleoside hydrolases from *C. fasciculata*, *Leishmania*, yellow lupin and Alaska pea among others in which N7 of the purine base is protonated early in the reaction to make the purine base a better leaving group. This is a common feature of purine hydrolysis among all nucleoside hydrolases characterized from many species. The activity of URH1 with 3-deazauridine, in which N3 of the pyrimidine ring has been replaced with a carbon in a manner similar to N7 of 7-deazaadenoine was determined. As with 7-deazaadenosine, 3-deazauridine was not a substrate indicating that protonation of N3 in a manner similar to that of protonation of N7 for purines is a required step during hydrolysis of pyrimidines. The chemical hydrolysis of purine nucleosides is acid-catalyzed and also involves protonation of N7. Protonation of purine bases destabilizes these bases by making them electron deficient.[42] The electron deficiency is made up by cleavage of the N-glycosidic bond.

7-Methylguanosine, an alkylated nucleoside was tested for substrate activity. This alkylated nucleoside mimics the N7-protonated nucleoside creating a permanent positive charge on N7, similar to its protonation. Unexpectedly 7-methylguanosine was not metabolized, the most probable reason being a steric effect between the exocyclic methyl group and the enzyme.

The chemical hydrolysis of pyrimidine nucleosides is more complex. The pK_a of uridine's N3 is much higher than the purine N7, on the order of 9, than the pK_a of adenosine N7, which is 2.4, indicating N3 is already protonated over the pH range at which the enzyme is active. The acid-catalyzed hydrolysis of uridine involves the hydration of the 5,6-double bond of uridine to yield 5,6-dihydrouridine which can undergo hydrolysis of the N-glycosidic bond.[130] Uracil can leave as a stable anion formed by resonance between O2 and O4. Additional study will need to be done to clarify the mechanism of enzyme-catalyzed hydrolysis of pyrimidines.

The structural requirements of the nitrogenous base are fairly non-specific with both purines and pyrimidines being substrates. The presence or absence of a specific group at the C6 position of the purine ring has a limited effect on substrate activity. Nebularine (purine riboside) which lacks an exocyclic group at C6 was a good substrate in which 80% of nucleoside was converted to purine in 2 hrs, a finding similar to that for RihC. URH1 tolerates both hydrogen bond donors and hydrogen bond acceptors in the C6 position of the purine ring. Adenosine containing the hydrogen bond donor NH_2 group and inosine containing the hydrogen bond acceptor =O group are both substrates with comparable activities. Xanthosine, which contains a keto group at C6 was hydrolyzed at a rate at approximately 57% of adenosine. The broad specificity of the nitrogenous base seen in purines is also present in pyrimidine nucleosides. Both uridine, containing a keto group at C4, and cytidine, containing an amino group at C4, are good substrates for URH1. In addition, 5-ribothymidine (5-methyluridine) is a substrate with an activity comparable to that of cytidine.

Plant hormones known as cytokinins are adenosine derivatives that help plants regulate cells and promote development. Cytokinins are N6- substituted adenine derivatives that are involved in a number of processes in plant growth and development including cell division, shoot

initiation, and leaf senescence among others. Two cytokinins, trans-zeatin riboside ((N^6- trans-4-hydroxy-3-methyl-2-buten-1-yl)adenosine) and kinetin riboside ((N^6-furfuryl)adenosine), and one cytokinin precursor, 6-(γ,γ- dimethylallylamino)purine riboside were tested as substrates for URH1 (Figure 9).

The cytokinins, trans-zeatin riboside and kinetin riboside, were not substrates, while unexpectedly 6-(γ,γ-dimethylallylamino)purine riboside was a reasonably good substrate. This is in contrast to RihC from *E. coli* in which compounds, such as 6-(γ,γ-dimethylallylamino)purine riboside and N^6-benzylaminopurine riboside, containing a bulky group at N6 were not substrates. The main limitation at the N6 exocyclic ring moiety appears to involve mainly steric effects. Purines that contain bulky substituents such as the benzyl group or furfuryl group which are found in cytokines were not substrates, although 6-(γ,γ-dimethylallylamino)purine riboside is an anomaly. This is in contrast to all purines that completely lack an exocyclic group or contained the relatively small - NH_2 or =O groups that are both substrates with comparable activities. Xanthosine, which contains a keto group at C6 was hydrolyzed at a rate at approximately 57% of adenosine. The broad specificity of the nitrogenous base seen in purines is also present in pyrimidine nucleosides. Both uridine, containing a keto group at C4, and cytidine, containing an amino group at C4, are good substrates for URH1. In addition, 5-ribothymidine (5-methyluridine) is a substrate with an activity comparable to that of cytidine.

The kinetic constants, Michaelis constant (K_m), and turnover number (k_{cat}) were determined for URH1 using inosine and uridine as substrates in triplicate (Table 4). Velocities for uridine and inosine were determined spectrophotometrically at concentrations of 10 µM, 50 µM, 100 µM, 250 µM, 500 µM, 750 µM, and 1,000 µM. The resulting velocity versus concentration data were fit to a Michaelis-Menten equation using a non-linear regression. Both substrates for each enzyme tested

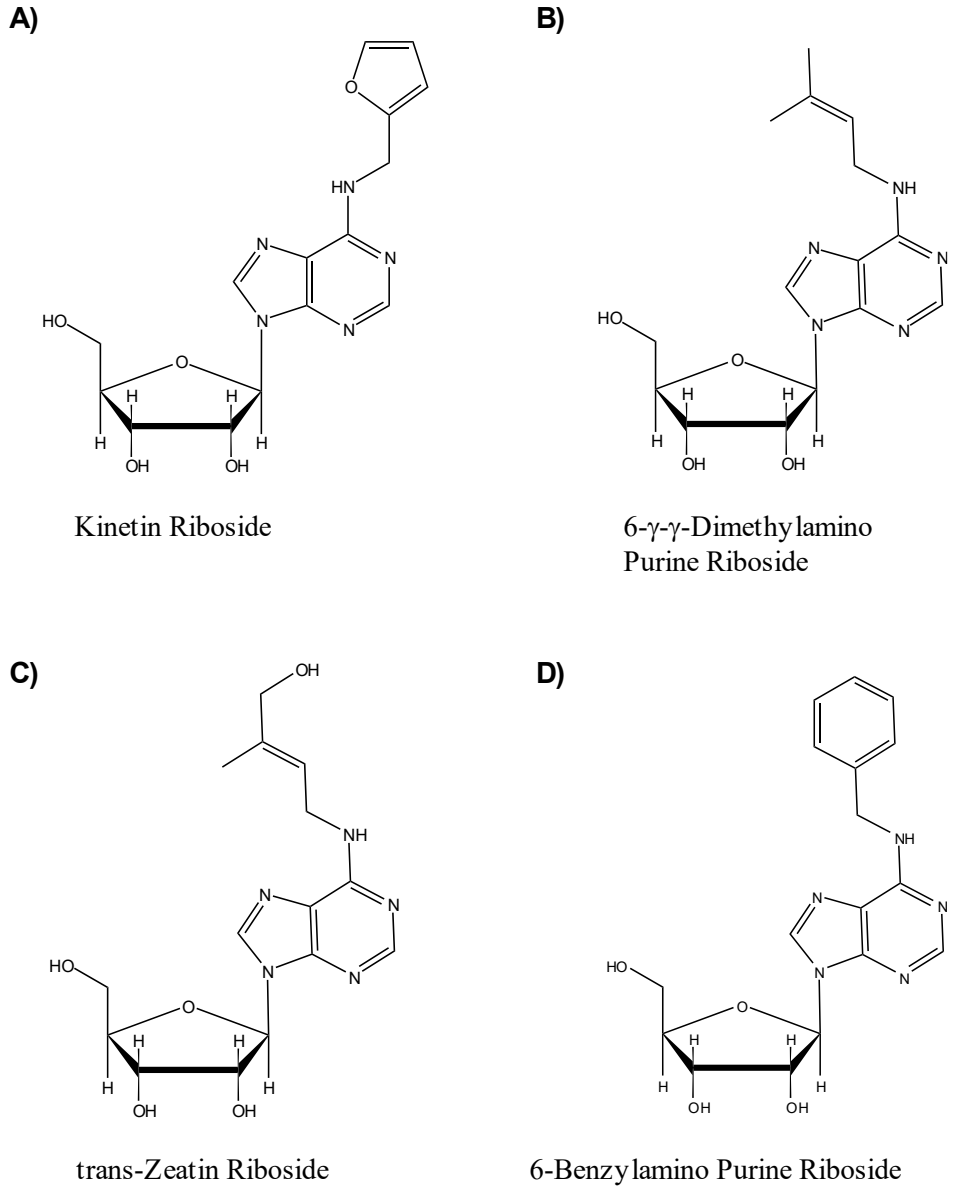


Figure 9. Cytokinin Chemical Structures. The structures above represent three cytokinin purine ribosides and a fourth aromatic riboside all tested as substrate for substrates. Further characterization needs to be carried out to ultimately determine the structure of the active site. URH1. The three cytokinin structures are A) Kinetin riboside B) 6-γ-γ-Dimethylaminopurine riboside and C) trans-Zeatin riboside. The fourth aromatic riboside is D) 6-Benzylamino purine riboside. 6-γ-γ-Dimethylamino purine riboside is a precursor in the biosynthesis of cytokinins.

Table 4. Kinetic parameters for RihC and URH1. Both values for RihC from *E. coli* and URH1 from *A. thaliana* with inosine and uridine. Values for RihC are taken from Arivett *et al.* 2014.

Enzyme	Substrate	K_M (μM)	k_{cat} per subunit (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
RihC	Inosine	422 ± 225	4.31 ± 0.22	1.02 x 10^4
	Uridine	408 ± 184	10.8 ± 0.23	2.6 x 10^4
URH1	Inosine	509 ± 177	61.87 ± 1.49	1.22 x 10^5
	Uridine	2,099 ± 389	629.9 ± 3.61	2.99 x 10^5

exhibited normal hyperbolic curvature consistent with Michaelis-Menten kinetics (Figure 10). There was no evidence of allosteric behavior with either inosine or uridine.

The Michaelis constants for URH1 with inosine and uridine were 509 ± 177 and 2,099 ± 389 μM respectively, while the turnover numbers were 61.87 ± 1.49 s^{-1} for inosine and 629.9 ± 3.61 s^{-1} for uridine. Jung et al have reported the Michaelis constant for uridine was 800 μM and for inosine was 1400 μM for URH1 which are consistent with the values reported in this study. The Michaelis constants for RihC with inosine and uridine were previously reported to be 422 ± 225 μM and 408 ± 184 μM respectively. The turnover numbers for inosine and uridine were 4.31 ± 0.22 s^{-1} and 10.8 ± 0.23 s^{-1}.

The turnover numbers for both inosine and uridine for URH1 were significantly higher compared to RihC. In addition, URH1 exhibited a greater difference between the turnover numbers of uridine compared to inosine, 10-fold, compared to the 2-fold difference of k_{cat} for

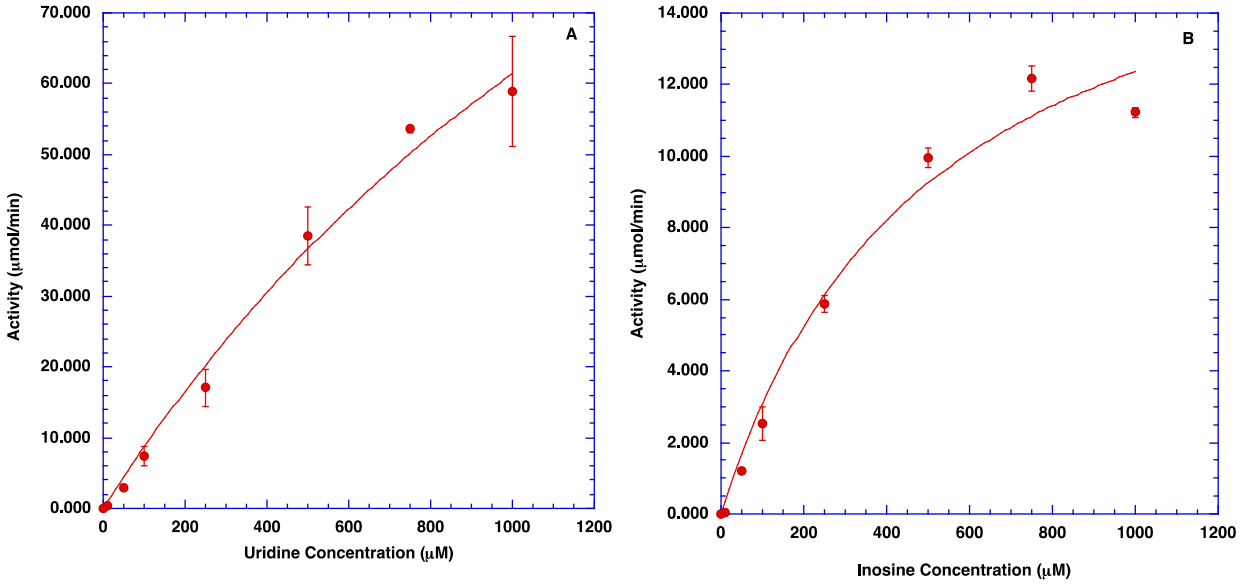


Figure 10. Michaelis-Menten plots for URH1. The average of triplicates for velocity measurements of uridine (A) and inosine (B) were plotted against nucleoside concentrations. A nonlinear regression analysis was used to fit the Michaelis-Menten equation to the observed rate determine K_M and V_{max}.

uridine and inosine in RihC. These variations may reflect the various metabolic roles of nucleoside hydrolases in different organisms. In *E. coli*, the role of nucleoside hydrolases is not well-understood as their metabolic role is duplicated by nucleoside phosphorylases[49]. In plants, a deficiency of enzymes of the salvage pathway exhibit reduced fertility and germination indicating they are important, but not necessarily essential for growth.[76,131]

The Michaelis constants for uridine and inosine from RihC and for inosine from URH1 are relatively constant. Uridine from URH1 has a significantly higher Michaelis constant, 4-fold, compared to the other reported values. URH1 from *A. thaliana* has a large Michaelis constant and

a high k_{cat}, indicating that while the enzyme efficiently hydrolyzes uridine its activity is significant only in the presence of high uridine concentrations. Despite the similarities in nucleoside hydrolases from *C. fasciculata*, *L. major*, *E. coli*, and *A. thaliana*, there are differences in the catalytic parameters and efficiency. The Michaelis constants for inosine and uridine in *E. coli* and *L. major* were similar, while in *A. thaliana* and *C. fasciculata,* the Michaelis constants for uridine were significantly higher than for inosine. It was observed that URH1 in *A. thaliana* was much more energetically favorable to pyrimidine hydrolysis compared to purine hydrolysis while the opposite was true for RihC in bacteria.

CHAPTER IV: MOLECULAR MECHANICS RESULTS AND DISCUSSION

4.1 Multiple Sequence Analysis

A multiple sequence analysis allows researchers to compare and contrast sequences from different organisms to identify conserved regions, predict functional domains, and infer evolutionary history. The crystal structural files were obtained for the non-specific CfIU-NH from *Crithidia fasciculata* (accession# Q27546) (PDB ID: 1MAS), the purine specific TvIAG-NH from *Trypanosoma vivax* (accession# Q9GPQ4) (PDB ID: 2FF1), the nucleoside N-ribohydrolases from *Physcomitrium patens* (PpNRH) (accession# A9TXA6) (PDB ID: 4KPN), and the nucleoside N-ribohydrolases from *Zea mays* (ZmNRH) (accession# B6T563) (PDB ID: 4KPO) from the RCSB PDB. In addition, the FASTA protein sequence files for RihC found from *E. coli* (accession # P22564.1) and URH1 found from *A. thaliana* (accession # Q9SJM7.2) were also retrieved to produce a multiple sequence analysis between each different nucleoside hydrolase. The % similarity and evolutionary relationship between each nucleoside hydrolase were ascertained by using the EMBOSS Needle PSA tool and Clustal Omega MSA tool.[96,97]

A PSA between the protozoan enzymatic sequences and URH1 resulted in the least similarity overall, as expected. CfIU-NH is a non- specific IU-NH which shared the closest homology and reaction mechanism with URH1, while TvIAG-NH is a purine specific inosine-adenosine- guanosine nucleoside hydrolase with less homology to URH1, but a very different reaction mechanism.[9,40] URH1 and CfIU-NH shared 49.9% similarity between each other, but when compared to TvIAG-NH, URH1 only shared 31.1% similarity. These results were expected since TvIAG-NH is a purine specific nucleoside hydrolase rather than a non-specific nucleoside hydrolase. These results were also consistent when CfIU-NH was compared to TvIAG-

NH by only sharing a 37.6% similarity. The bacterial enzymatic sequence RihC from *E. coli* was also compared to URH1 but resulted in a higher % similarity then protozoa. URH1 and RihC shared 50.3% similarity between each other. These results were also expected since RihC is a non-specific IU-NH with similar reaction characteristics to URH1.

As expected, plant protein sequences for PpNRH and ZmNRH resulted in the highest % similarity compared to URH1 using a PSA tool. PpNRH shared 62.5% similarity to URH1, while ZmNRH shared 75.5% similarity. PpNRH and ZmNRH each shared 63.8% similarity between each other, but when compared to URH1 the ZmNRH had the highest % similarity with regards to residue similarity. A multiple sequence alignment of each nucleoside hydrolase shows where conserved residues can be located between different species (Figure 11). Conserved active site residues have been highlighted to present the calcium binding motif early in each sequence followed by the protonation pair which comes later in the sequence. The evolutionary relationship between each species can also be viewed in a phylogenetic tree (Figure 12). The phylogenetic tree shows how each different nucleoside hydrolase may function the same mechanistically based on the similarity of amino acid residues which are present. Next we evaluated the structural characteristics of each enzyme by producing molecular docks of each protozoa and plant structural models retrieved from the PDB and compared them using an inosine ligand.

4.2 NUCLEOSIDE CONFORMATIONAL PARAMETERS

Characterization of hydrolysis reactions catalyzed by nucleoside hydrolases also requires an understanding of the bound conformation of the nucleoside ligand before analyzing the docked state of each enzyme. Nucleosides have complicated and extensive geometries between the ribose sugar and nucleobase. Conformational parameters were evaluated after docking inosine to the

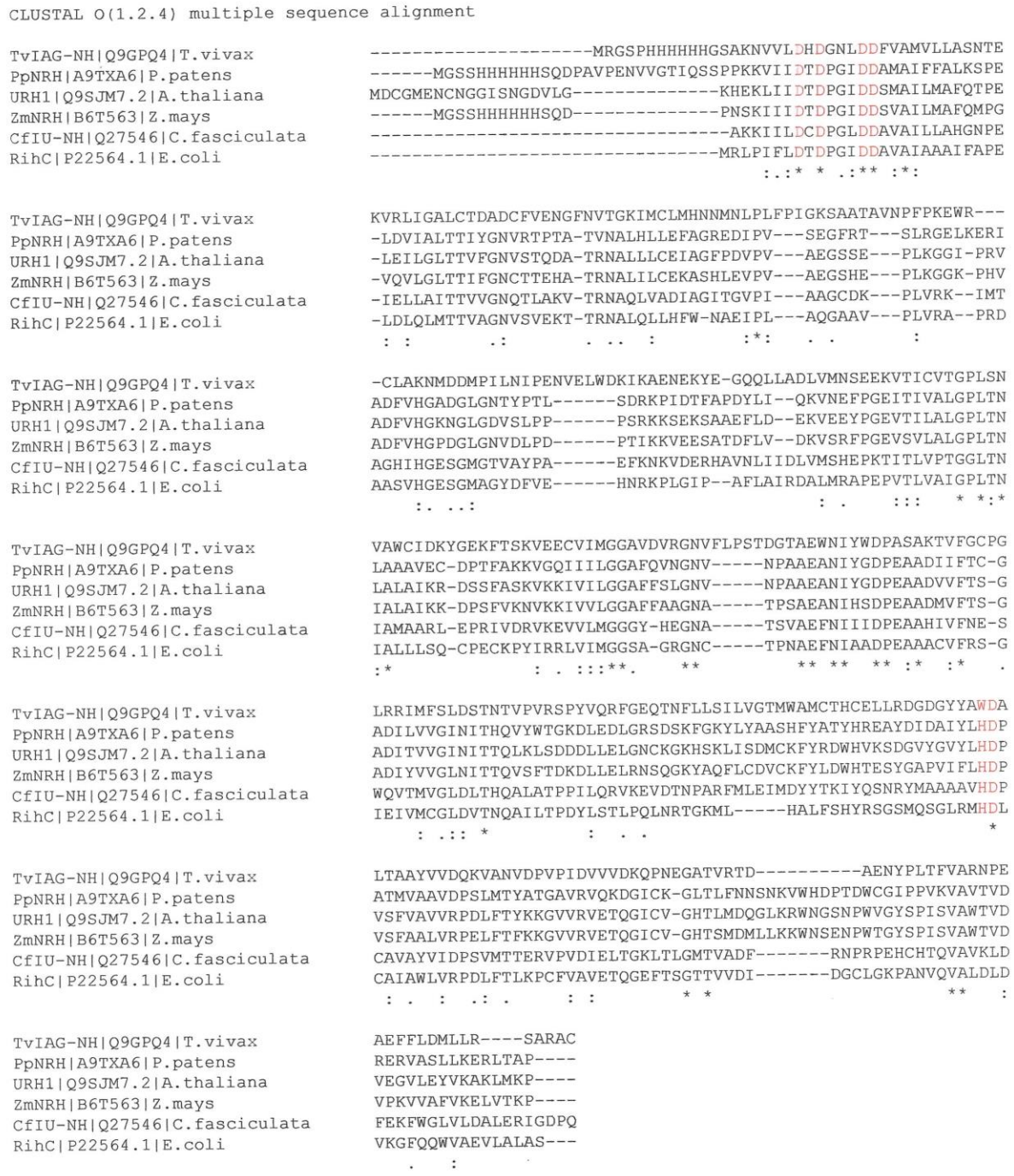


Figure 11. Multiple Sequence Analysis. The sequences for TvIAG-NH, CfIU-NH, PpNRH, ZmNRH, and URH1 were all compared using an MSA. Conserved residues for the DXDXXXDD motif for calcium binding and HD/WD protonation pair are identified with red letters.

Figure 12. Phylogenetic tree. Evolutionary relationships between each different nucleoside hydrolases for each species can be seen above.

active site of each PDB model and the homology model of URH1. Inosine was chosen because it has been well described experimentally in the literature and conformations are much easier to observe when viewing molecular docks. Three important conformational parameters for inosine to gauge favorable docks involved pseudorotation between the C2' and C3' on ribose, the exocyclic orientation of C4'-C5' on ribose, and the glycosyl conformation of the nucleoside base between ribose.[132–136] Very little is known about what residues affect each conformational pose so these parameters were evaluated with each dock to compare between each type of nucleoside hydrolase. Comparing these four related nucleoside hydrolase docks with crystal structures from the PDB allowed us to see if we can assume functionality between residues on the URH1 homology model.

The pseudorotation cycle can take two forms: an envelope (E) with four atoms in the same plane forming a boat structure or a twist (T) with two adjacent atoms displacing the other three atoms in a chair structure.[132–136] When the C2' or C3' atom is puckered in an upward direction in the same plane as the C5' hydroxyl it takes an -*endo* conformation.[132–136] When the C2' or C3' atom is puckered in a downward direction away from the C5' hydroxyl plane it takes an -*exo*

conformation..[132–136] The most energetically favorable pose for pyrimidine nucleosides is the C3'- *endo* conformation, while the most energetically favorable for purine nucleosides is the C2'- *endo* conformation (Figure 13).[132–136]

The second parameter involves the orientation of the exocyclic C4'- C5' bond.[132–136] The C4'-C5' bond can freely rotate the O5' hydroxyl in three conformations around the furanose ring as: gauche-gauche (g+g-), gauche- trans (g+t), and trans-gauche (t+g).[132–136] Each one of these conformational poses affect flexibility through sugar/base interactions, hydrogen bonding, and ultimately how the structure will come to function as a nucleic acid. [132–136] The gauche-gauche conformation is typically the most favorable pose and faced inside the pentofuranose ring at 60^{o} from the C2' while gauche-trans and trans-gauche face away from the pentofuranose ring at 180^{o} and 300^{o}(-60^{o}) from the C2' respectively (Figure 13).[132–136]

The third and final parameter involved two conformations according to where the nitrogenous base is oriented relative to the pentofuranose ring.[132–136] The *syn*-conformation resulted in the bulk of the nitrogenous base located over the pentofuranose ring while the *anti*-conformation resulted in the bulk of the nitrogenous base rotated away from the pentofuranose ring.[132–136] In solution, purine and pyrimidine nucleosides exist in the *syn* and *anti*-conformations with equal frequency at equilibrium, but the *anti*-conformation is energetically more favorable out of the two conformations due to less steric hindrance and the ease of hydrogen bonding with the exposed base pointing away from the pentofuranose ring (Figure 13).[132–136] All of these parameters were evaluated with inosine docked as a ligand in the active site of protozoan and plant nucleoside hydrolases from the PDB along with the URH1 homology model relate functionality. Next, we evaluated the active site residues for each protozoan model with inosine docked.

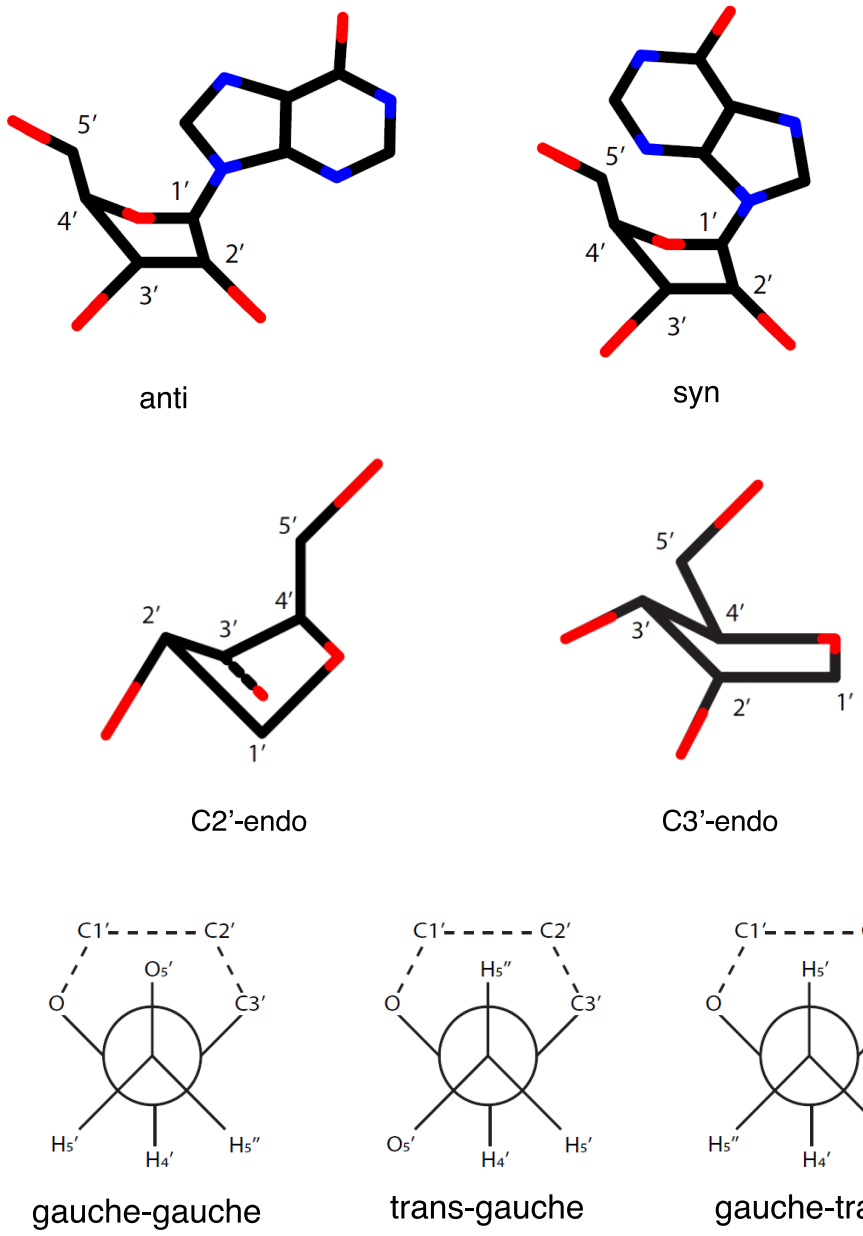


Figure 13. Nucleoside Conformational Parameters. Each nucleoside conformational parameter can be viewed above. Red areas note oxygen atoms in the ring while blue notes nitrogen atoms. The pseudorotation of the ribose A), exocyclic C4'-C5' Newman projection of O5' B), and glycosyl conformation C) are also presented. The figure was reproduced from An Vandemeulebroucke Dissertation, 2008.[136]

4.3 Molecular Docks Between Protozoan Models

Molecular docking of protein and ligand structures allowed the viewing of important spatial interactions that occur between amino acid residues when in close proximity to nearby ligands for a deeper understanding of active site interactions. The crystal structural files of the putative enzymes for CfIU-NH from *Crithidia fasciculata* (PDB ID: 1MAS) and *Trypanosoma vivax* (PDB ID: 2FF1) were retrieved from the RCSB PDB to perform molecular docks with inosine. CfIU-NH is a non-specific IU-NH which shares close homology and a reaction mechanism with URH1, while TvIAG-NH is a purine specific inosine-adenosine-guanosine nucleoside hydrolase with slightly less homology to URH1, but a very different reaction mechanism.[9,40] URH1 and CfIU-NH shared 49.9% similarity between each other, but when compared to TvIAG-NH, URH1 only shared 31.1% similarity. These results were expected since TvIAG-NH is a purine specific nucleoside hydrolase with less homology then CfIU-NH which is a non-specific nucleoside hydrolase. These results were also consistent when CfIU-NH was compared to TvIAG-NH by only sharing a 37.6% similarity. The crystalline structure for CfIU-NH and TvIAG-NH retrieved form the PDB can each be viewed in along with a close- up view of the active site docked with an inosine molecule and a close-up view of the residues around the N7 atom on inosine within ~5 Å (Figure 14 and 15).

Conserved residues were evaluated around the active site. Six residues have been associated with nucleoside hydrolysis experimentally in the past around the active site of protozoa.[7] Four aspartic acid residues from a motif of the protein sequence in the order DXDXXXDD are believed to be involved in calcium binding along with a single Thr residue. CfIU-NH and TvIAG-NH both shared these residues in sequence order with each other. A fifth, single Asp paired with a His residue later in the protein sequence believed to contribute to protonation during nucleoside hydrolysis was observed on CfIU-NH.[7] TvIAG-NH did not share these same residues. TvIIAG-NH had Trp260

located near the N7 atom on inosine where the histidine residue was located on CfIU-NH. The Trp residue is believed to function in the activation of the purine base through an arene-arene interaction with a second Trp residue to stack the purine base in position for hydrolysis.[44]

The distances between the six primary residues believed to be involved between calcium binding and base protonation were also measured. Potassium was used as a place holder ion to represent calcium for the CfIU-NH. Residues observed as potentially binding calcium (represented as potassium visually) on the CfIU-NH were Asp10, Asp15, Asp242, and Thr126. Each of these residues ranged at distances from 2.4-2.7 Å. The most feasible residues closest to N7 on inosine and involved with base protonation are His241 and Asp14. His241 and Asp14 were each 4.8 and 4.2 Å away from N7 respectively. Residues on TvIAG-NH closest to calcium based on the crystalline structure form the PDB are Asp10, Asp15, Asp261, and Thr137. The most feasible residues closest to N7 available for arene stacking on TvIAG-NH was Trp260. Trp260 was ~7.8 Å away while an Asp14 was measured 7.0 Å away for reference (Figure 14 & 15). Based on these solved structures, we used these residues to identify which residues of URH1 are likely to be involved in the reaction.

Several other residues, previously uninvestigated, were close enough for involvement with protonation for hydrolysis. Amino acid residues and local atoms around the N7 atom on inosine were selected and viewed in a 5 Å zone using Chimera (version 1.15).[99,100] Distances were measured for each species to find residue similarities to URH1 (Figures 14 and 15). The CfIU-NH active site structure was mapped and had four residues available for ligand interaction. The only significant residue located on CfIU-NH was His241 which acts as its primary electron donor during protonation allowing the formation of an oxocarbenium ion on the ribosyl moiety to use water in a nucleophilic reaction.[3,16,40] His 241 was measured to only be 4.8 Å from N7. Other residues within 5 Å with reactive side groups were Asp14, Asn160, and Phe167, but were of no interest as

reactive side groups to merit further interest.

The TvIAG-NH protonation residues did not share homology with any of the species examined in this study. Distances around the N7 atom on inosine were also mapped around TvIAG-NH as well with four residues found to be within 5 Å area nearby. TvIAG-NH was identified having Trp83 and Trp260 which both act on nucleoside ligands by stacking pi elections from the ring structures in parallel to align the nitrogenous base for protonation.[9,137] Both Trp residues were measured to be only 4.5 and 7.3 Å away from N7. Other residues within 5 Å with reactive side groups were Asn12 and Asp 40 which may both have the potential to play a role in hydrolysis.

The conformational poses of inosine were also examined on CfIU- NH and TvIU-NH. Inosine docked in the CfIU-NH and TvIU-NH active sites both showed the nucleobase facing the *anti*-conformation away from ribose. The pseudorotation of ribose found in the active site of both protozoa models were in a twist (T) pose in the C2'-*endo* conformation. The exocyclic C4'-C5' bond with O5' hydroxyl found on each protozoa model were different. The CfIU-NH favored a trans-gauche orientation while TvIU-NH favored the gauche-trans orientation. Protozoa models appear to both favor an *anti*-C2'-*endo* conformation, but only changed conformational poses with regards to the exocyclic C4'- C5' bond (See Figure 13).

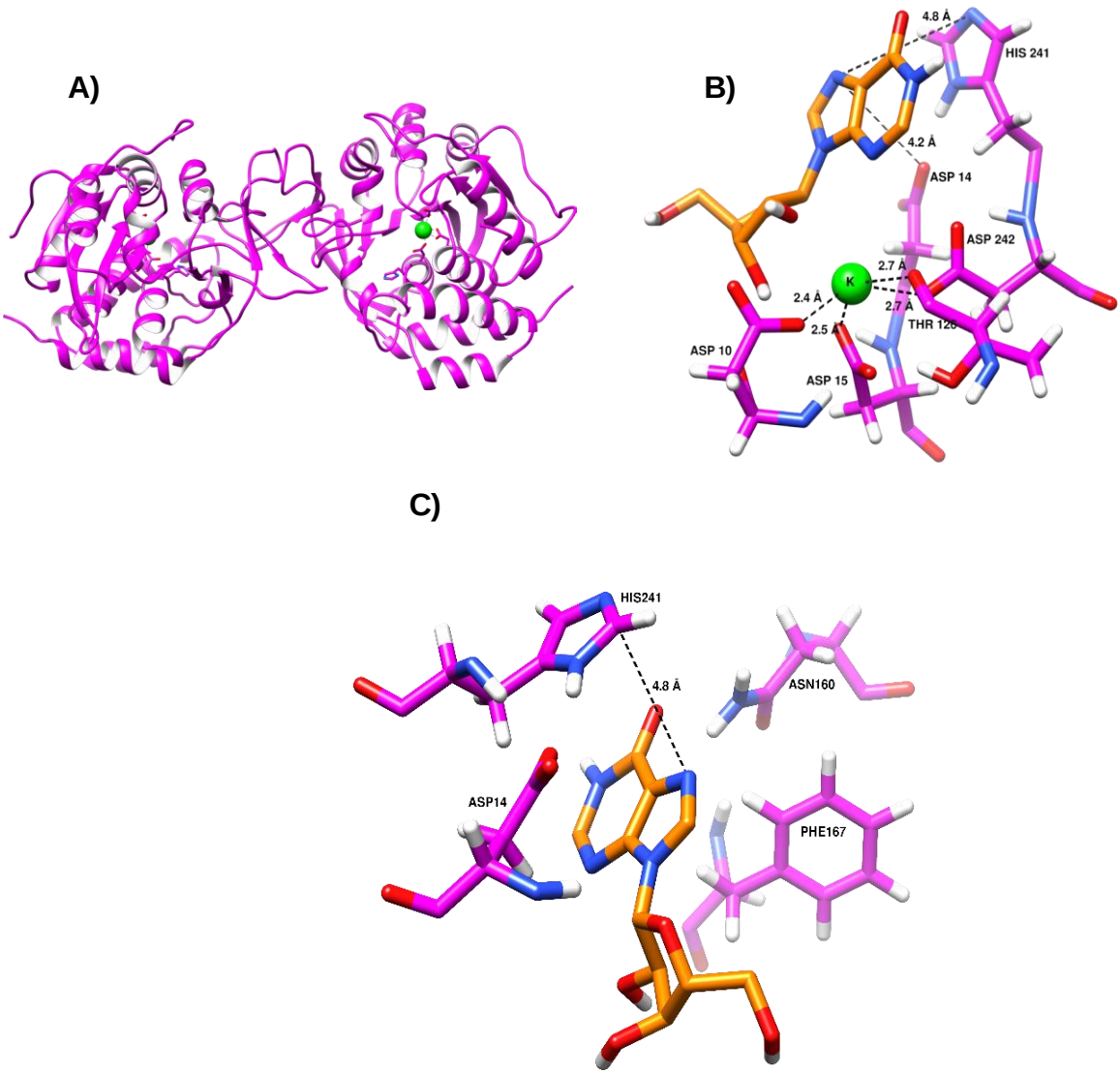


Figure 14. CfIU-NH Active Site (1MAS). The overall crystalline structure for CfIU- NH from the PDB can be seen in A. The active site of CfIU-NH with inosine docked and the distances for Asp10, Asp14, Asp15, Asp 242, Thr126, and His241 to relevant residues can also be viewed in B. A 5 Å zone was created around N7 of the inosine nitrogenous base to determine favorable residues available to interact for protonation as a potential electron donor C.

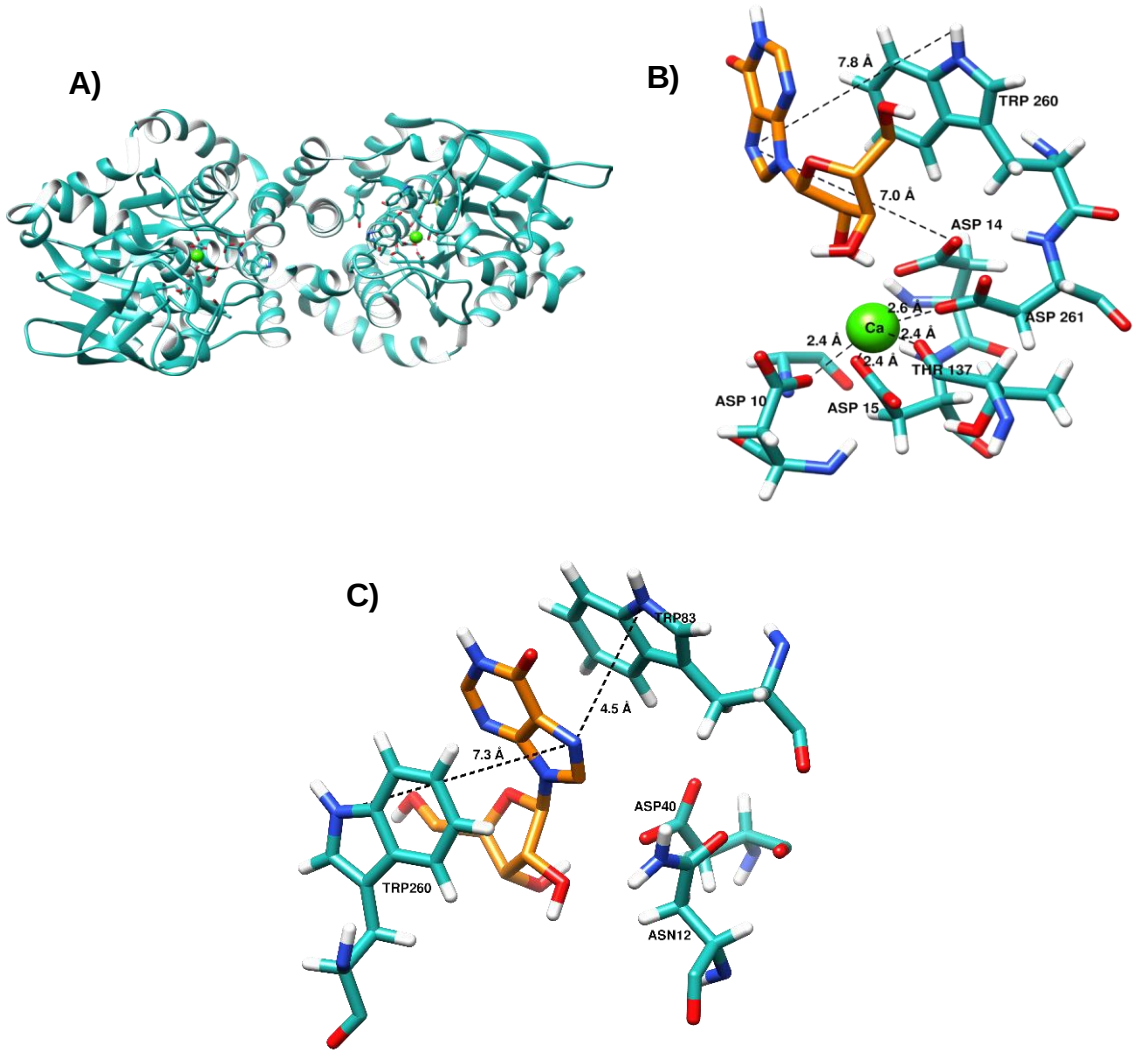


Figure 15. TvIAG-NH Active Site (2FF1). The overall structure of the crystalline structure for TvIAG-NH from the PDB can be viewed in A. The active site of TvIAG-NH with inosine docked and the distances for Asp10, Asp14, Asp15, Asp 261, Thr137, and Trp260 to relevant residues can also be viewed in B. A 5 Å zone was created around N7 of the inosine nitrogenous base to determine if favorable residues were available to base interactions C.

4.4 Molecular Docks Between Plant Models

In similar fashion, molecular docks using other plant models which have been experimentally observed to act as non-specific IU-NHs were also evaluated to relate the functionality of the URH1 homology model. Structural files for PpNRH from *Physcomitrium patens* (PDB ID: 4KPN) and ZmNRH from *Zea mays* (PDB ID: 4KPO) were retrieved from the RCSB PDB. Molecular docks of inosine were created with the plant PDB models. Plant models were shown to share very close homology with URH1 based on pairwise and multiple sequence analysis tools (Figures 11 and 12). The percent similarity between PpNRH and ZmNRH was only 63.8%, but when compared to URH1, PpNRH shared 62.5% similarity, while ZmNRH shared 75.5% similarity. The structural characteristics for each plant model was also viewed in Chimera (version 1.15).[47] The crystalline structure for PpNRH and ZmNRH retrieved from the PDB can each be viewed with a close-up view of the active site docked with an inosine molecule and a close-up view of the residues around the N7 atom on inosine within ~5 Å (Figures 16 and 17).

Conserved residues were also evaluated around the active site of each plant model. The same six homologous residues related to nucleoside hydrolysis present in protozoan models were also present in PpNRH and ZmNRH. Each plant model had four Asp residues from the DXDXXXDD motif, believed to be involved with calcium binding, along with a fourth Asn residue. It appeared Asn may have changed from the Thr residue found in protozoan models. A fifth, single Asp paired with a His residue later in the protein sequence were also present related to protonation for hydrolysis. The residues around the active site were also mapped as they were for the protozoan model (Figures 16 and 17).

Homologous residues attributed to calcium binding in PpNRH with tight coordination include: Asp25, Asp29, Asn54, and Asp258. Each of these residues measured in distance from 2.3

and 4.5 Å from calcium. The N7 atom from inosine was also measured for Asp29 and His257. Each of these residues were closest to coordinate with the N7 atom to provide a reasonable means for protonation for hydrolysis. Asp29 and His257 were measured to be 5.1 and 5.6 Å respectively. The residues associated with calcium binding on ZmNRH included: Asp8, Asp13, Asn37, and Asp240. Each of these residues were measured at distance ranging from 2.5 to 4.6 Å. Protonation residues near the N7 atom included Asp12 and His239 and were distance of 4.5 to 5.6 Å respectively.

Because of the close homology between plant IU-NHs from PpNRH and ZmNRH we expect to see very similar structural and functional characteristics from URH1. Other residues close to the N7 atom of the inosine purine ring were also selected on the plant models for PpNRH and ZmNRH as well (Figure 16 and 17). PpNRH had a His99 and His257 very close to the N7 atom also potentially close enough for protonation. His99 was only 5.6 Å from N7 while His257 was only 5.1 Å. Other residues of interest on PpNRH included Asn176, Tyr241, and Tyr244. ZmNRH also had a His81 and His239 very close to the N7 atom also potentially close enough for protonation. His81 was only 5.6 Å from N7 while His239 was only 4.5 Å. Other residues of interest on ZmNRH included Asn58, Tyr223, and Trp226.

The poses of inosine were also examined on PpNRH and ZmNRH. Inosine docked in the PpNRH and ZmNRH active sites both showed the nucleobase facing the *anti*-conformation away from ribose as well. The pseudorotation of ribose found in the active site of each plant model were also in a twist (T) pose and the C2'-*endo* conformation. The exocyclic C4'-C5' bond with O5' hydroxyl found on each plant model were both in the gauche-trans orientation. The plant models appear to both favor an *anti*- conformation with a (T) C2'-*endo* pseudorotation, and an exocyclic C4'-C5' bond gauche-trans orientation (See Figure 13).

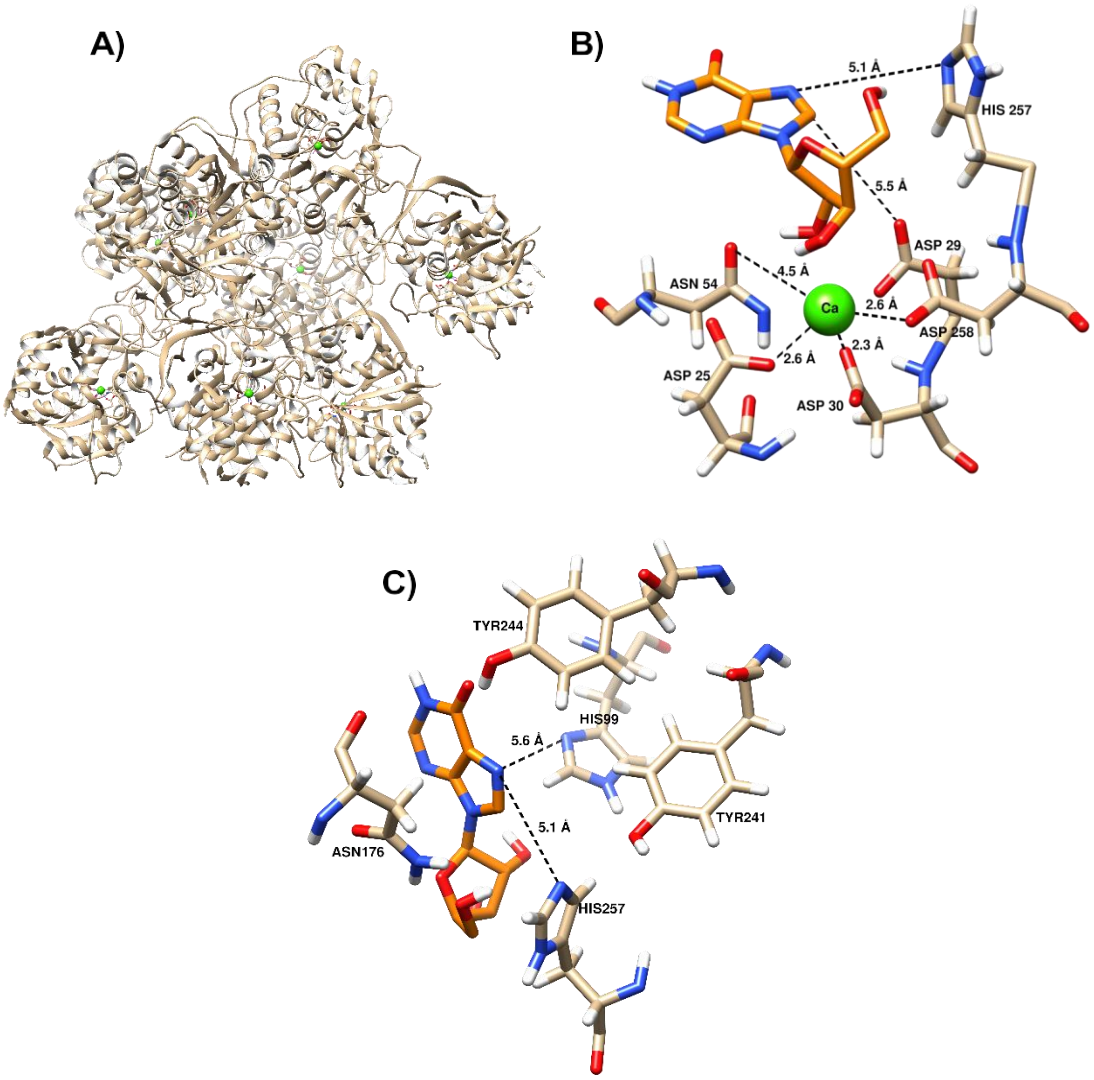


Figure 16. PpNRH Active Site (4KPN). The crystalline structure for PpNRH from the PDB can be viewed in A). The active site of PpNRH with inosine docked and the distances for Asp25, Asp29, Asp30, Asn54, His257, and Asp 258 to relevant residues can also be viewed in B). A 5 Å zone was created. A 5 Å zone was created around N7 of the inosine nitrogenous base to determine favorable residues available to interact for protonation as an electron donor in C).

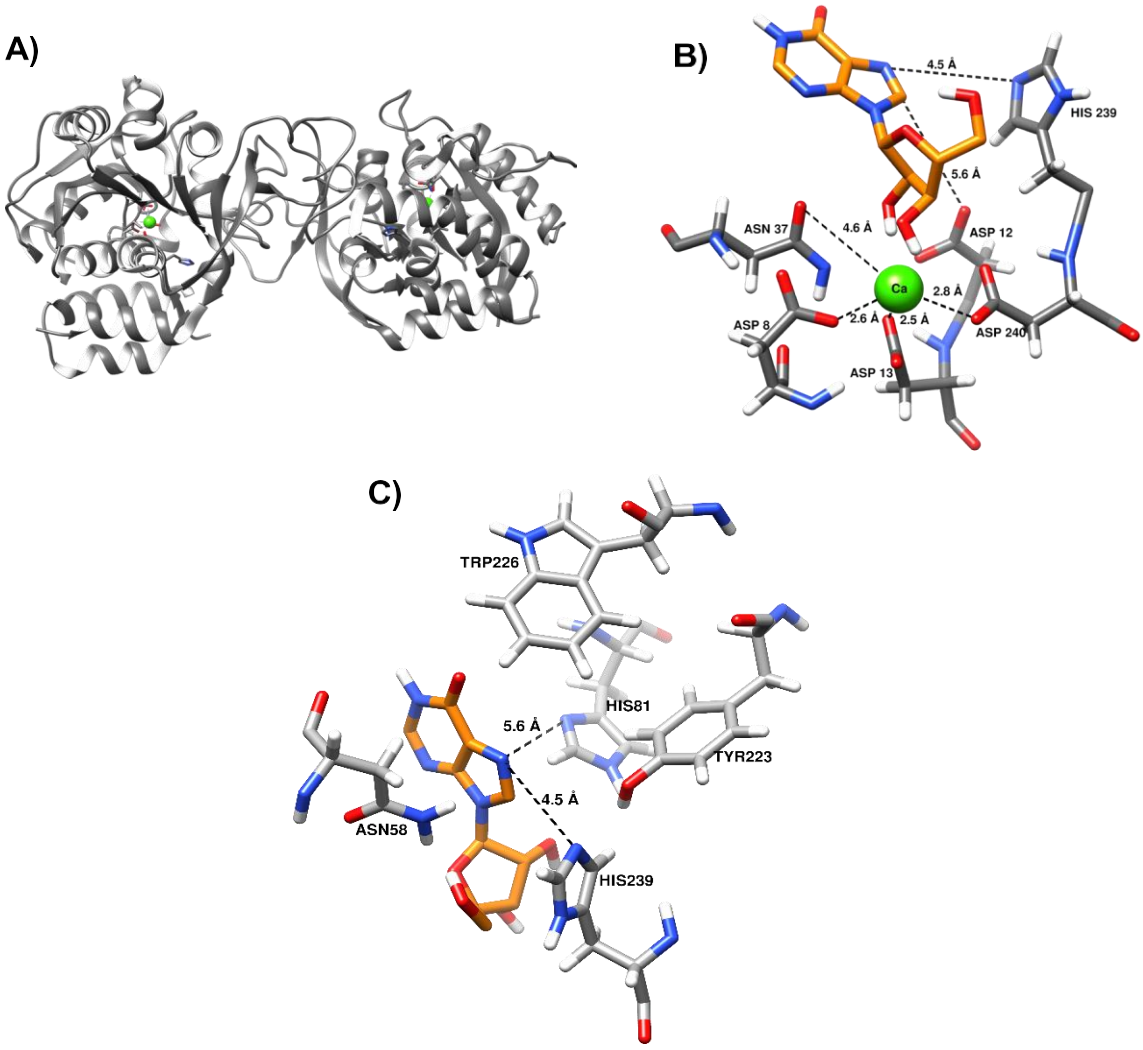


Figure 17. ZmNRH Active Site (4KPO). The crystalline structure for ZmNRH from the PDB can be viewed in A). The active site of ZmNRH with inosine docked and the distances for Asp8, Asp12, Asp13, Asn37, His239, and Asp 240 to relevant residues can also be viewed in B). A 5 Å zone was created around N7 of the inosine nitrogenous base to determine favorable residues available to interact for protonation as an electron donor in C).

4.5 Structural Analysis of URH1

The results produced from sequence alignments and structural analysis can all be used to relate functional mechanics to the URH1 homology model produced by the SWISS-MODEL website.[68,85,88,91,92] Visual models were created with the overall structure of URH1 along with a close view of the active site docked with an inosine and a close view of URH1 residues around the N7 atom on inosine (Figure 18). Conserved residues on six sites identified from two motifs for calcium binding and base protonation from previous analyses were also present on URH1 as well. Each of the residues and distance models were also applied to the URH1 model to draw comparisons with other models.

Residue distances were very consistent with crystal models retrieved from the PDB. Distances between residues from a motif shown to be involved with calcium binding on URH1 included Asp29, Asp34, Asn58, and Asp261. Each of these residues were tightly bound around calcium at distances of 2.7 to 4.7 Å. Two other homologous residues on a second motif close enough to the N7 atom on inosine that could potentially be attributed to protonation on URH1 were Asp33 and His260. Asp33 and His260 were the closest residues near N7 at a distance of 5.3 and 4.2 Å from the N7 atom on inosine, respectively. Several other residues were also observed around the N7 atom on inosine not previously attributed to hydrolysis but were close enough and had the charge (Figure 18). His102 was only 5.3 Å away from N7 which is as comparatively close as His260 which has been the conserved residue attributed to base protonation driving hydrolysis in reaction mechanics

The poses of inosine were examined for each docked ligand as well. Inosine docked in the URH1 active site resulted the nucleobase facing the *anti*- conformation away from ribose. The pseudorotation of ribose found in the URH1 active site was in a twist pose (T) in the C2'-*endo*

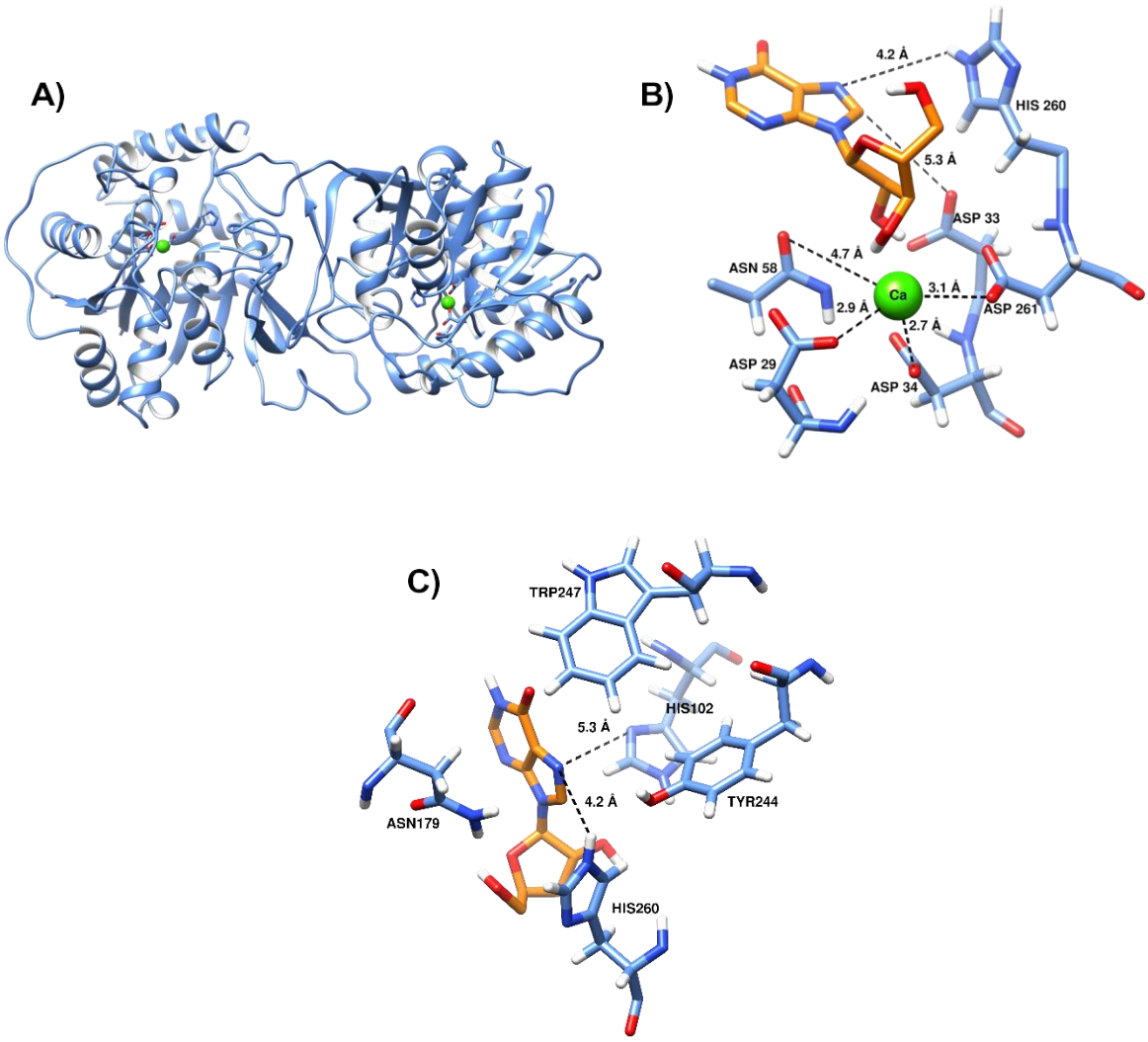


Figure 18. UR1 Active Site. Sequence data retrieved for URH1 from NCBI was used to create a SWISS-MODEL structure seen in A). The active site of the SWISS-MODEL URH1 with inosine docked and the distances for Asp29, Asp33, Asp34, Asp 261, Asn58, and His260 to relevant residues can also be viewed in B). characteristics to interact with protonation. These residues included: His102, Asn179, Tyr244, and Trp247. A 5 Å zone was created around N7 of the inosine nitrogenous base to determine favorable residues available to interact for protonation as an electron donor in C).

4.6 Molecular Dynamics Simulations with URH1

Molecular dynamics (MD) computational modeling for enzymatic studies have become a popular method to predict reaction mechanisms by using ligand docking, molecular mechanical (MM), and quantum mechanical (QM) analyses tools to predict regions that are flexible and static to target for mutations.[55,58,61,138–148] Computational simulations allow users to explicitly describe the Newtonian movements of all protein, ligand, and solvent atoms, thereby providing information on protein:ligand interactions in solution. Models of protein residues important to ligand binding can be developed based on relative distances of residues with respect to each other. We used MD simulations of URH1 to determine residue flexibility regions around the active site to assist targeting residues which mostly likely drive nucleoside hydrolysis function. To accomplish this, we produced four MD simulations with the software GROMACS using undocked apo and holo states along with docked inosine and uridine states of URH1.[102–106]

MD simulations were produced using the apo state without calcium or a ligand, the holo state with a calcium cofactor present, and two docked models using inosine and uridine docked with a calcium cofactor. We hoped to produce conformational states which revealed the areas of greatest flexibility and change while ligand and cofactors were bound and unbound in and around the active site. Each of these states were equilibrated and tested for favorable temperature and pressure conditions. A short 10 ns MD simulation was performed to determine if the system was relaxed and stable before producing longer 200 ns MD simulations. A plot of the RMSD and time was produced to present the overall system equilibration/relaxation (Appendix E).

After verifying the system was relaxed, we produced three different MD simulations to capture molecular mechanical movements of URH1 while in solution. We then used a principal component analysis (PCA), Gibbs-free energy (GFE) landscape, and a root-mean square

fluctuation (RMSF) of URH1 to analyze static and dynamic areas located around the regions on URH1. The information gathered allowed us to build a picture of the important residues around URH1 which may be driving the mechanism of hydrolysis with nucleosides.

To assess the conformational dynamics of URH1, a principal component analysis (PCA) was performed. The PCA evaluated multidimensional data from MD trajectories created in GROMACS and produced a simplified representation of regions with major motion moving in unison for each URH1 state. Scree plots present the proportion of variance plotted against the eigenvalue rank of the individual conformational motions to highlight dominant motions displayed by the simulated protein states. The principal component for the undocked apo and holo protein along with the docked inosine and uridine URH1 states can be viewed in four scree plots (Figure 19). Each scree plot illustrates the principal components of residue motion for each URH1 state. The principal components (PC1 and PC2) of the undocked apo state in the scree plot resulted in a 38.8% of the system variance as 27.3% and 11.5%. The undocked holo scree plots resulted in slightly similar principal component (PC1 and PC2) values of 38.6% of the system variance as 27.7% and 10.9%. The docked inosine and uridine scree plots resulted in different principal components when compared to undocked states. The docked inosine scree plot resulted in 32.5% of the system variance for PC1 and PC2 as 17.5% and 15%, respectively, while the docked uridine scree plot resulted in 27.7% of the system variance for PC1 and PC2 as 15.5% and 12.2%, respectively. The bound states seemed to produce less movement among its principal components. This most likely is due to the stability provided by the docked state holding the enzyme tightly coordinated. Out of each docked state inosine seemed to have a higher system variance equating to slightly more motion than the uridine docked state. The regions of greatest variance are found in the top left corner of each plot. Additional analysis was also performed to determine the energy

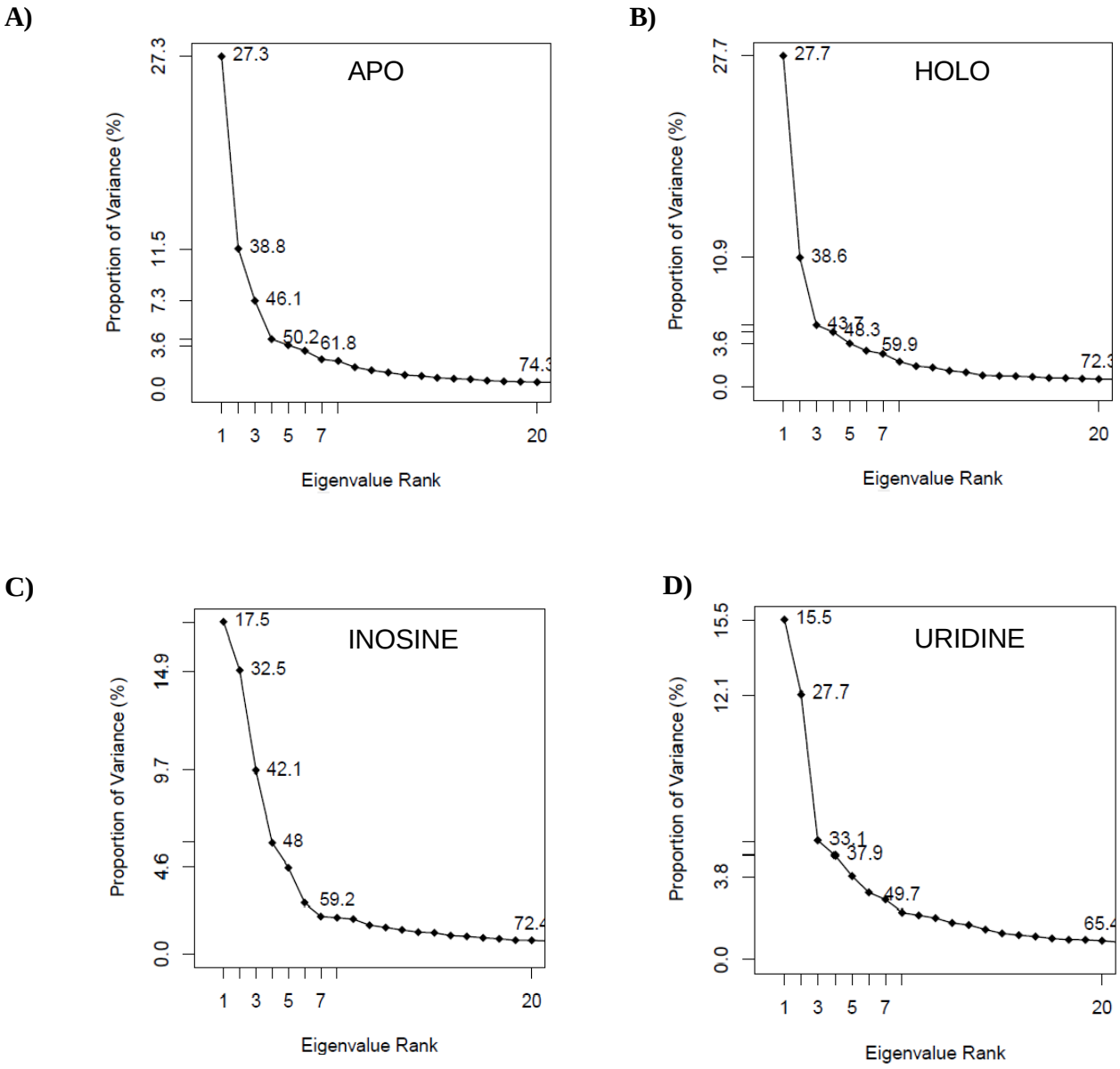


Figure 19. PCA scree plots for URH1. Four MD scree plots illustrate the undocked A) Apo and B) Holo states along with the C) Inosine D) Uridine docked states. Each MD simulation captured a 200 ns time frame of residue movement in a 3-dimensional water solution box. Scree plots show the proportion of variance of as the principal components of correlated residue motion to the system variance.

states from the principal component analysis.

The Gibbs free energy (FEL) landscape can capture the energy barriers of each enzyme state as a function of its conformational space in a three-dimensional plot. Each FEL plot illustrated a two-dimensional plot of residue movement from MD simulations on the xy-axis with the root-mean- square deviation (RMSD) and the radius of gyration (ROG), along with the energy state for each region of motion on the z-axis (Figure 20). High energy barriers indicate regions of flexibility, while low energy barriers suggest rigidity. This information can describe how an enzyme can adapt to substrate and reaction intermediates. Each undocked and docked URH1 state investigated from MD simulations were also analyzed with a FEL plot.

The undocked apo and inosine dock each showed consistent range of motion along the xy-axis and had similar energy barriers along the z-axis. The holo and uridine dock showed similar ranges of motion, but each went in an opposite direction, while also maintaining equivalent energy levels. Since the holo and uridine seem to highlight the same types of movement and energy it seems to imply that calcium may play a greater role in pyrimidine hydrolysis based on enzymatic folding. In all four cases the regions with the highest energy barriers and greatest flexibility seemed to occur on the outer regions of the enzyme while the inner regions associated with the active site were much more stable and rigid.

To better understand what residues are specifically involved with regions of flexibility and stability the root mean square fluctuation (RMSF) was performed for each residue from the MD simulation of each undocked and docked state of URH1. The RMSF helped develop more specificity according to the residues involved with flexible loops located on the enzyme based on the PCA data. The top three principal components of each state seem to align closely with each

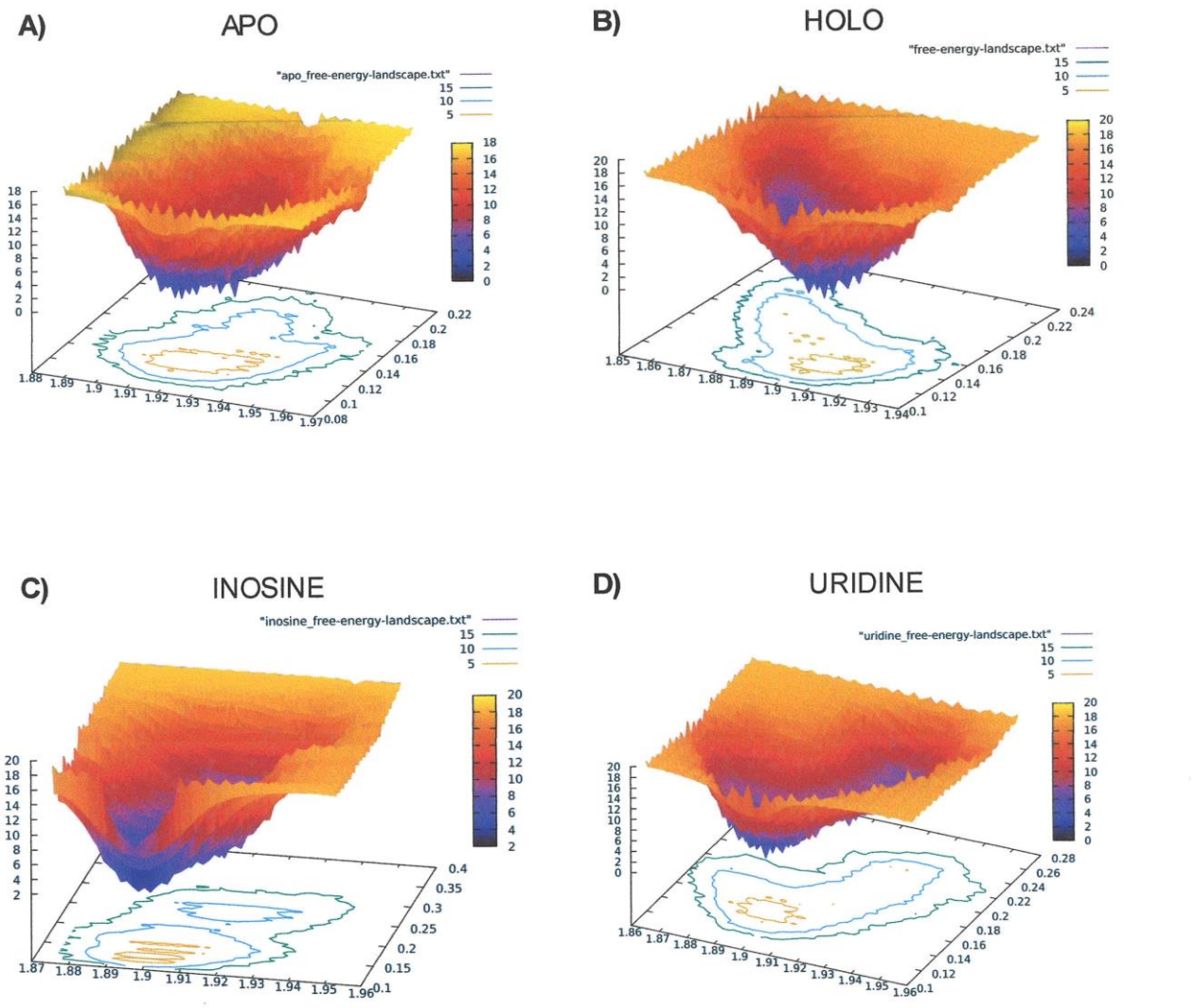


Figure 20. Gibbs free energy landscape for URH1. The three-dimensional plots for the free energy landscape of each URH1 state can be viewed above. The plots for the A) apo protein, B) holo protein, C) inosine bound, and D) uridine bound states of URH1 were produced for each 200 ns MD simulation.

other based on the RMSF plots from each undocked and docked state of URH1 (Figure 21). Flexible loops on each enzyme state have been identified by a colorized region on the enzyme in each plot. The top three principal components can be viewed according to its colored residue region. The red region represents residue positions 60-80, the blue region represents residue positions 220-240, and the green position represents residue positions 260-290. These three regions were consistently the most flexible regions in all analyzed cases.

Undocked states shared the most regions with high flexibility compared to the docked states. The region of greatest flexibility for URH1 occurred at the green residue positions 260-290 with the exception of the holo state. The residue region with the greatest flexibility for the undocked holo state occurred at the blue region for positions 220-240. Both of these regions were an α-helix which included the conserved His260-Asp261 residue positions on one end involved with driving hydrolysis through protonation of the nucleoside base. Other residues of interest that could potentially play a role in hydrolysis near the blue region included Tyr244 and Trp247 which were observed near the N7 of the purine base. The RMSF plots also show that conserved residues that occur early in the enzyme sequence attributed to binding calcium are very rigid and stable.

Other studies performed with MD and QM/MM (quantum mechanics/molecular mechanics) simulations have analyzed other nucleoside hydrolases to find flexible loops and NH motion. A QM/MM simulation performed for a pyrimidine-specific cytidine-uridine nucleoside hydrolase (CU-NH) from *E. coli* investigated two flexible loops involved with hydrolysis found along residue regions 90-110 and 190-220.[139] Several other QM/MM simulations have also been performed with purine-specific inosine-adenosine-guanosine nucleoside hydrolases (IAG-NH) found in *T. vivax*. These studies targeted calcium binding and protonation residues for IAG-NHs in *T. vivax* to study transitional states of purines which also

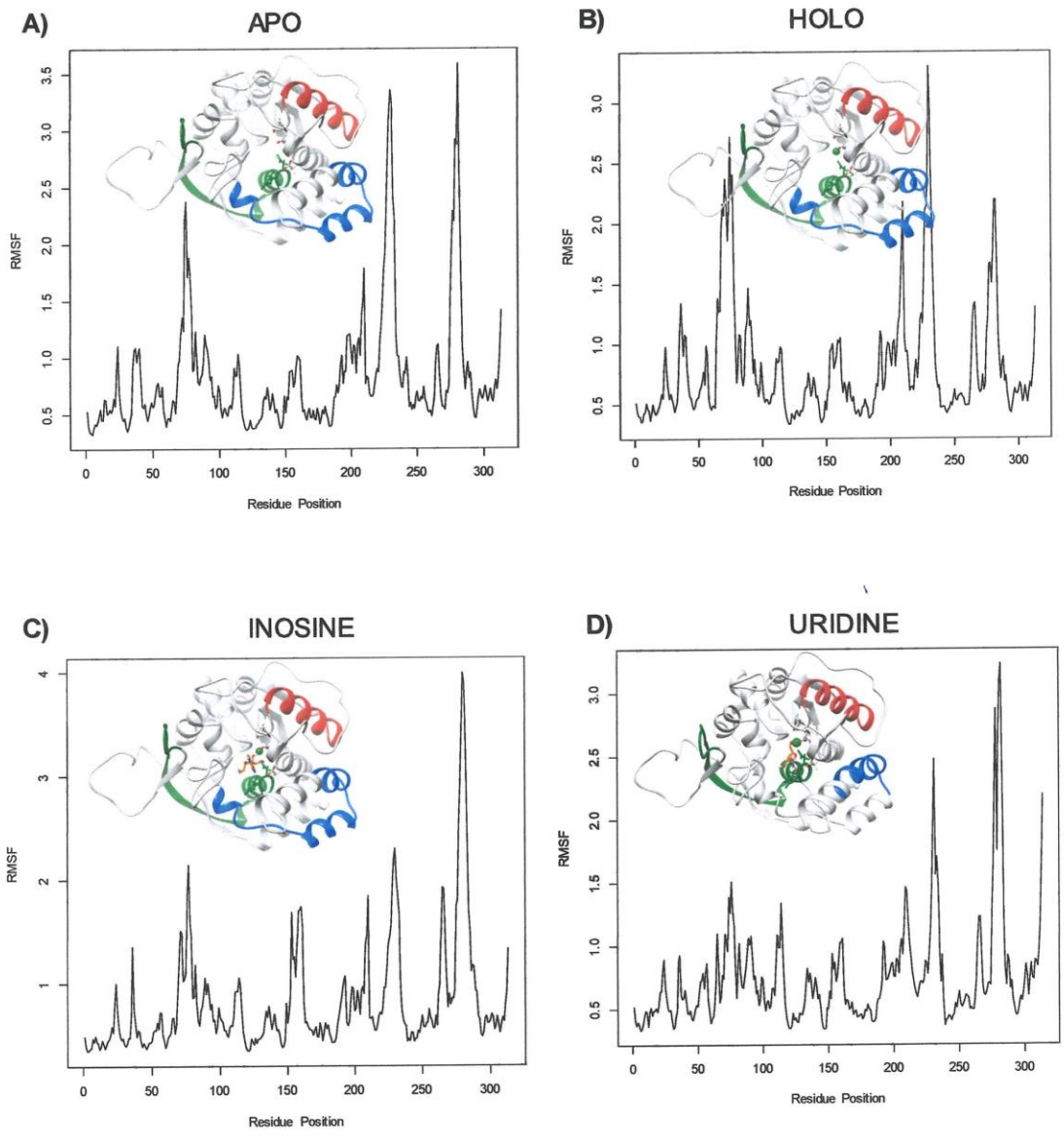


Figure 21. RMSF for URH1. The residues with the greatest flexibility for the apo A) and holo B) undocked states and the docked inosine C) and uridine D) states for URH1 based on MD simulations are presented above. Flexible regions are highlighted by color. Red areas highlight residue positions 60-80, blue regions represent positions from 220-240, and the green regions represent 260-290.

identified two loops at residue positions 75-85 and 244-258.[137,140,149] These studies seem to support MD mechanical studies we have conducted identifying several flexible loops on nucleoside hydrolases near active site residues. Several residues have been identified from this research that have previously been unacknowledged but may warrant further investigation.

If we analyze residue positions that have been identified from *C. fasciculata* through experimental analysis, we can begin to build a picture of how the reaction may occur in URH1 if inosine is bound to the active site.[7] A schematic for a proposed hydrolysis pathway occurring in URH1 was produced using conserved and newly viewed residues from this study potentially involved with nucleoside hydrolysis (Figure 22).

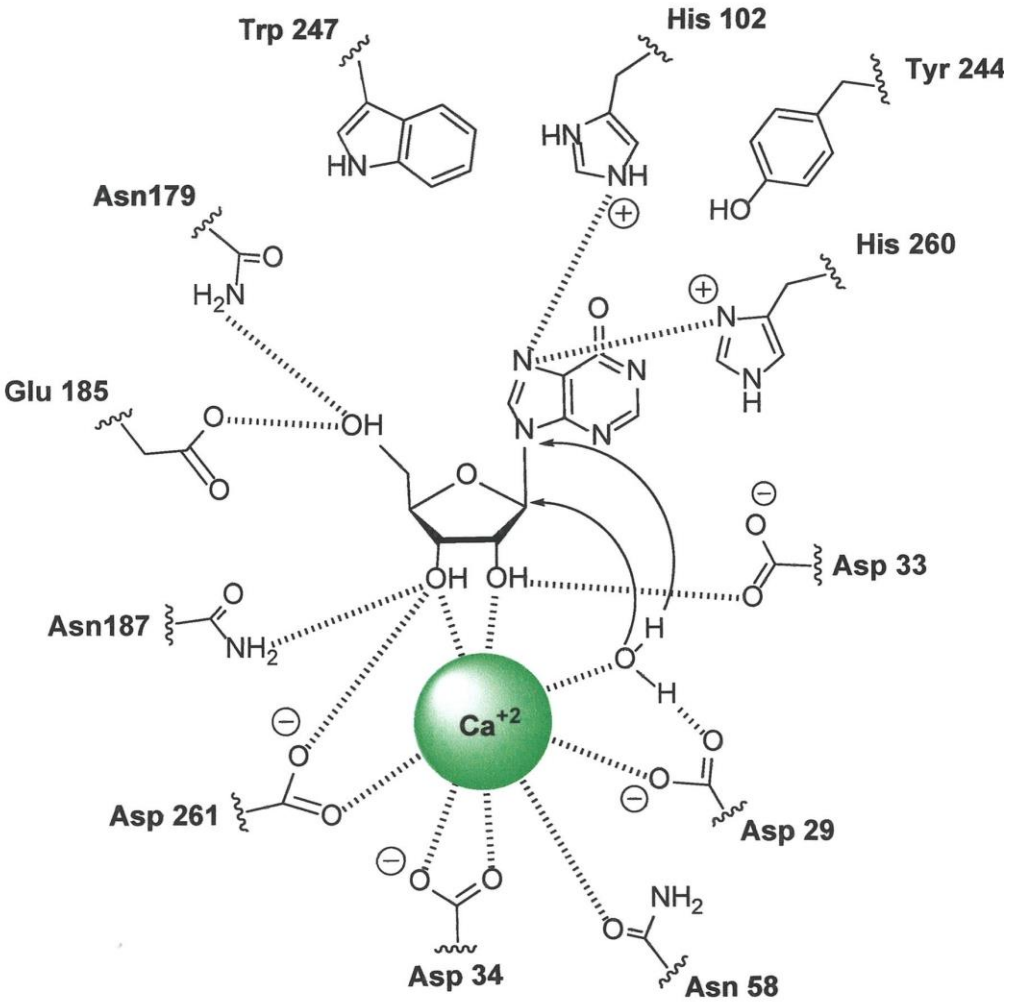


Figure 22. Proposed Hydrolysis Pathway for URH1. Several new and conserved residues were observed around the URH1 active site which may be linked to nucleoside hydrolysis. A 5 Å zone was created around a docked inosine ligand to find nearby residues that potentially play a role in hydrolysis consistent with the hydrolysis model presented for CfIU-NH.[7] These new residues included Asn58, His102, Asn179, Glu185, Asn187, Trp247, and Tyr244. Figure 22 was created using Chemdraw Version 22.2.0.

CHAPTER V: SITE-DIRECTED RESULTS AND DISCUSSION

5.1 Site-Directed Mutagenesis

Site-directed mutagenesis studies allow researchers to replace amino acid residues on proteins to discover how they function biologically. Molecular biologists can relate function by finding conserved amino acid residues that have shared the same function over generations of evolutionary history. Sequence alignments are useful tools used by molecular biologists to identify similar residues on proteins that may relate function based on evolutionary relationships. An MSA was produced for this study to relate nucleoside hydrolase activity to several different species to discover similarity in catalytic function. The Clustal Omega tool (version 1.2.4) from EMBOSS was used to perform an MSA for *Arabidopsis thaliana* (accession # At2g36310 / Q9SJM7.2), *Physcomitrium patens* (accession # JQ649322.1 / A9TXA6), *Zea mays* (accession # HQ825162.1 / B6T563) *Trypanosoma vivax* (accession # AF311701.2 / Q9GPQ4), *Crithidia fasciculata* (accession # U43371 / Q27546), and *E. coli* (accession # EG11082 / P22564.1) (Figure 11). The % similarity for URH1 between each species was examined and can be found in section 4.1.

The MSA in Figure 11 between each species of enzyme model used in this study provided a view of the key conserved residue sites identified with red letters. The *C. fasciculata* model for non-specific IU-NHs has been used as the archetype organism used to describe the putative catalytic mechanism which describes non-specific nucleoside hydrolysis.[7] Calcium has been proposed to be an important cofactor coordinating homologous residues to fold and stabilize the active site providing the function needed for the hydrolysis reaction to occur.[67] Homologous residues on each species are located on two motifs with a sequence of DXDXXXDD found early in the protein sequence attributed with calcium binding and a second motif with a His/Asp pair attributed with

base protonation all interacting to cleave the glycosidic bond of nucleosides.[1,19,23] These residues were conserved on each species and were targets for this mutagenesis study to find the impact they have for hydrolysis with URH1 in *A. thaliana*.

Conserved residues described by this hydrolysis reaction and chosen for mutation around the active site of URH1 were identified from previous experimental data from CfIU-NH and RihC. Several mutants were also available from RihC in *E. coli* performed by Arivett et al. 2014 for velocity comparisons as a control. Six conserved reactive residue sites from two motifs were mutated on URH1 to non-reactive Ala residues to compare velocities to see how mutations affect activity. The first four mutations from *URH1* were Asp27, Asp29, Asp33, and Asp34 occurring on the DXDXXXDD motif early in the sequence attributed to calcium binding and the other two mutations were carried out on His260 and Asp261 occurring later in the sequence attributed to base protonation. Four mutant residues were chosen from *rihC* and occurred at Asp 14, Asp15, His233, and Asp234 and were also mutated to non-reactive Ala residues (Table 4).

Each of these residues were changed to Ala residues to affect reactivity. Since these residues were conserved, as shown from an alignment study, they were targeted to show if function was consistent with URH1. If hydrolysis velocity values from mutants were consistent with velocities observed from wild type experiments, residues were not important to hydrolysis activity for function. If hydrolysis velocities were observed to be significantly lower than velocities seen from wild type experimental values then these residues were identified as being critical to hydrolysis activity for function. A sequence validation was also conducted to determine if mutations were successful.

5.2 Mutant Sequence Validation Results

Mutant sequence validation work was conducted by Dr Dan Bryant from MTSU. The site-directed mutagenesis performed on the wild type sequences of URH1 were validated with gene sequencing similar to previous studies with *rihC*.[64] Mutation locations for the six mutant residues on URH1 were targeted based on evolutionary relationships between other non-specific nucleoside hydrolases (Figures 11 and 12). Wild type reading frames were trimmed for quality and pairwise aligned to create a full sequence for the wild type and each mutant site. The missense mutations were validated by comparing the wild type amino acid sequence to the mutant amino acid sequence from the Geneous Prime software. Each of the sequence alignments performed for the mutant plasmids showed that missense mutations were successful, however they appeared to be a slightly off- set by one residue potentially due to an extra His residue mistakenly taken from the beginning of the His_6-tag on the N-terminus.

5.3 Mutant Velocity Comparison

Reaction data for wild type and mutants of URH1 and RihC were tabulated to determine if hydrolysis velocities (μmol/min) were different when compared. Experimental rates between inosine and uridine for the wild type and mutant enzymes were produced and tabulated (Tabe 4). Results from these mutation experiments suggested conserved residues from both RihC and URH1 were important for maintaining stability with a calcium cofactor and for active site hydrolysis for function because the rates from mutant experiments were substantially lower than wild type rates. While these results showed that the six conserved residues chosen for experiments played a key role in hydrolysis, certain residues did not drop as low as other residues were expected.

It was expected to observe greater losses of activity with mutants from His233 on 233 and

His260 on URH1. When His241 was mutated to alanine in experiments by Gopaul et al., 1996 a 2,100 fold loss in k_{cat} for inosine was observed indicating the importance of His in the active site of CfIU-NH.[40] Other URH1 mutations did not exhibit the same drop in velocities as mutations for His233 and His260 which are considered vital for hydrolysis. This seemed to indicate that there may be other residues driving hydrolysis reactions for non-specific IU-NHs potentially showing a more complicated mechanism than previously suspected. Based on the data from these experiments there may be other residues involved and possibly multiple mechanisms occurring between the different types of nucleosides.

Table 5. Mutagenesis data.

Organism	Homologous Residue	Suspected Function	Mutation	Velocity (µmol/min)	
				Inosine	Uridine
E. coli **(RihC)**	***WT**	-	-	**0.665**	**1.31**
	Asp-14	Ca Binding Site	Asp-14-Ala	0.103	0.150
	Asp-15	Ca Binding Site	Asp-15-Ala	0.053	0.001
	His-233	Base Protonation	His-233-Ala	0.100	0.521
	Asp-234	Base Protonation	Asp-234-Ala	0.072	0.094
A. thaliana **(URH1)**	***WT**	-	-	**11.23**	**58.84**
	Asp-27	Ca Binding Site	Asp-27-Ala	0.163	0.038
	Asp-29	Ca Binding Site	Asp-29-Ala	0.074	0.099
	Asp-33	Ca Binding Site	Asp-33-Ala	0.047	0.092
	Asp-34	Ca Binding Site	Asp-34-Ala	0.107	0.119
	His-260	Base Protonation	His-260-Ala	0.097	0.385
	Asp-261	Base Protonation	Asp-261-Ala	0.117	0.165

*WT = Wild Type

CHAPTER VI: CONCLUSIONS

The aim of this dissertation was to analyze a non-specific IU-NH (URH1) from the plant species *A. thaliana* to better understand the mechanism controlling nucleoside hydrolysis. One of the questions we hoped to answer from this study was how URH1 functions by characterizing its state of oligomerization, equilibrium constant, pH profile, kinetic activity, substrate specificity, and calcium content. We also hoped to answer questions with regards to what the critical amino acid residues around the active site are by comparing different species of non-specific IU-NHs found in putative plant, bacterial, and protozoa models through sequence alignment, ligand docking, and molecular dynamics simulations. Finally, we also hoped to show if putative conserved residues controlled hydrolysis through a mutagenesis study.

The characterization of this enzyme started with designing a plasmid for the gene of URH1 to be placed in a BL21 cultures for production and purification. The protein was determined to exist in a mixed oligomerization state around 37 kD. URH1 activity functioned over a broad pH range from 6-8 and the enzyme-catalyzed reaction was essentially irreversible, with no observation of nucleoside formation in the presence of high concentrations of hypoxanthine and inosine. The kinetic mechanism was in rapid random equilibrium which has been seen in a number of other nucleoside hydrolases. URH1 was also classified as a Group 1 non-specific nucleoside hydrolase based on its broad substrate specificity. The enzyme was nonspecific for both purine and pyrimidine bases and highly specific for ribose as the pentose moiety. The primary ribonucleosides shown to display activity included: guanosine, nebularine (purine riboside), inosine, adenosine, xanthosine, uridine, cytidine, 5-methylurdine. All ribose alterations and deoxyribosides tested resulted in no detectable activity. Zeatin and kinetin were not substrates, while the cytokinin

precursor 6-(γ,γ-dimethylallylamino) purine riboside was observed to produce slight activity on a lower level. Kinetic constants indicated URH1 was very active against uridine and was shown to be very efficient at high concentrations. URH1 had a lower activity against inosine, but remained active at lower concentrations and were consistent with rates seen form RihC in *E. coli*.

Amino acid residues for URH1 were examined using computational tools to evaluate the active site. Sequence alignments for the protozoan, bacterial, and plant models of *Crithidia fasciculata* (accession # U43371 / Q27546) (PDB ID: 1MAS), *Trypanosoma vivax* (accession # AF311701.2 / Q9GPQ4) (PDB ID: 2FF1), *Physcomitrium patens* (accession # JQ649322.1 / A9TXA6) (PDB ID: 4KPN), *Zea mays* (accession # HQ825162.1 / B6T563) (PDB ID: 4KPO), and *E. coli* (accession # EG11082 / P22564.1) were all obtained from the NCBI and PDB websites to compare to a SWISS-MODEL of URH1 from *A. thaliana* (accession # At2g36310 / Q9SJM7.2). A pairwise and multiple sequence analysis of each protein sequence showed CfIU-NH and TvIAG-NH had the least % similarity to URH1 with 49.9% and 31.1% respectively, while PpNRH and ZmNRH had the most % similarity to URH1 with 62.5% and 75.5% respectively. RihC was slightly consistent with CfIU- NH and shared 50.3% similarity with URH1. These results were expected since all of these nucleoside hydrolases were all non-specific IU-NHs with the exception of TvIAG-NH which was the least similar of all. Each of these model organisms were docked with inosine and the active site evaluated to find putative residues close enough for interaction to impact hydrolysis. Distances between six putative conserved residues on URH1 (Asp27, Asp29, Asp33, Asp34, His260, and Asp261) were compared to conserved residues on other CfIU-NH, TvIAG-NH, PpNRH, and ZmNRH to show stability between binding residues and function between hydrolytic residues. Several new residues were identified around the N7 atom of a docked inosine on the URH1 active site that were previously uninvestigated and may potentially impact hydrolysis.

Molecular Dynamic simulations revealed regions of stability and flexibility on URH1 through a principal component analysis, Gibbs free energy landscape, and RMSF data. MD simulations showed conserved residues binding calcium played a role in coordinating stability with a calcium cofactor in the active site, while other conserved residues were near regions of greater flexibility and motion assisting hydrolysis. Gibbs free energy data and RMSF data identified that regions and residues attributed to calcium binding were among some of the most stable regions on URH1, while other residues around the active site attributed to hydrolysis were near flexible loops on the enzyme. A hydrolysis pathway was proposed to describe the most prominent residues on URH1 that most likely influence mechanism.

Mutagenesis data showed conserved residues from alignment and docking experiments on URH1 do play an important role in nucleoside hydrolysis, but other residues may also be impacting the overall mechanism. Conserved residues mutated to Ala on URH1 (Asp27, Asp29, Asp33, Asp34, His260, and Asp261) and RihC (Asp14, Asp15, His233, and Asp234) associated with calcium binding and protonation were all shown to impact hydrolysis rates when inosine and uridine substrate were measured spectrophotometrically. Putative Asp residues (Asp27, Asp 29, Asp33, Asp34, Asp261 and Asp14, Asp15, and Asp234) all showed a large drop in the reaction velocities, while the protonation residues (His233 and His260), thought to be more important in hydrolysis, were slightly higher. These findings seem to indicate a more complex mechanism than previously thought.

Additional work is still needed to explore the mechanism occurring around the active site of URH1. More mutagenesis studies are needed to explore other residues discovered around the active site that seem to also play a role in hydrolysis. Calcium was shown to be very important to hydrolysis in coordinating residues around the active site. More quantitative studies should be

explored to see how impactful calcium concentrations can affect hydrolysis. Binding experiments involving additional substrates should also be explored to determine why some substrate had a greater affinity as opposed to others. Additional research should also focus on other inhibitors that would make better candidates to control plant growth as herbicides that do not pose human health concerns or become more tolerant to plant exposure while also exploring whether these candidates affect other organisms in the environment as well.

REFERENCES

(1) Parkin, D. W.; Horenstein, B. A.; Abdulah, D. R.; Estupiñán, B.; Schramm, V. L. Nucleoside Hydrolase from Crithidia Fasciculata. Metabolic Role, Purification, Specificity, and Kinetic Mechanism. *J. Biol. Chem.* **1991**, *266* (31), 20658–20665. https://doi.org/10.1016/S0021-9258(18)54759-1.

(2) Degano, M.; Almo, S. C.; Sacchettini, J. C.; Schramm, V. L. Trypanosomal Nucleoside Hydrolase. A Novel Mechanism from the Structure with a Transition-State Inhibitor. *Biochemistry* **1998**, *37* (18), 6277–6285. https://doi.org/10.1021/bi973012e.

(3) Degano, M.; Gopaul, D. N.; Scapin, G.; Schramm, V. L.; Sacchettini, J. C. Three-Dimensional Structure of the Inosine−Uridine Nucleoside *N* - Ribohydrolase from *Crithidia Fasciculata*. *Biochemistry* **1996**, *35* (19), 5971–5981. https://doi.org/10.1021/bi952999m.

(4) Estupiñán, B.; Schramm, V. L. Guanosine-Inosine-Preferring Nucleoside N- Glycohydrolase from *Crithidia Fasciculata*. *J. Biol. Chem.* **1994**, *269* (37), 23068–23073. https://doi.org/10.1016/S0021-9258(17)31620-4.

(5) Miller, R. L.; Sabourin, C. L.; Krenitsky, T. A.; Berens, R. L.; Marr, J. J. Nucleoside Hydrolases from *Trypanosoma Cruzi*. *J. Biol. Chem.* **1984**, *259* (8), 5073–5077.

(6) Ogawa, J.; Takeda, S.; Xie, S. X.; Hatanaka, H.; Ashikari, T.; Amachi, T.; Shimizu, S. Purification, Characterization, and Gene Cloning of Purine Nucleosidase from *Ochrobactrum Anthropi*. *Appl. Environ. Microbiol.* **2001**, *67* (4), 1783–1787. https://doi.org/10.1128/aem.67.4.1783-1787.2001.

(7) Versées, W.; Steyaert, J. Catalysis by Nucleoside Hydrolases. *Curr. Opin. Struct. Biol.* **2003**, *13* (6), 731–738.https://doi.org/10.1016/j.sbi.2003.10.002.

(8) Mitterbauer, R.; Karl, T.; Adam, G. Saccharomyces Cerevisiae URH1 (Encoding Uridine-Cytidine N-Ribohydrolase): Functional Complementation by a Nucleoside Hydrolase from a Protozoan Parasite and by a Mammalian Uridine Phosphorylase. *Appl. Environ. Microbiol.* **2002**, *68* (3), 1336–1343. https://doi.org/10.1128/AEM.68.3.1336-1343.2002.

(9) Versées, W.; Decanniere, K.; Van Holsbeke, E.; Devroede, N.; Steyaert, J. Enzyme-Substrate Interactions in the Purine-Specific Nucleoside Hydrolase from *Trypanosoma Vivax*. *J. Biol. Chem.* **2002**, *277* (18), 15938–15946. https://doi.org/10.1074/jbc.M111735200.

(10) Versées, W.; Barlow, J.; Steyaert, J. Transition-State Complex of the Purine- Specific Nucleoside Hydrolase of *T.Vivax*: Enzyme Conformational Changes and Implications for Catalysis. *J. Mol. Biol.* **2006**, *359* (2), 331–346. https://doi.org/10.1016/j.jmb.2006.03.026.

(11) Pellé, R.; Schramm, V. L.; Parkin, D. W. Molecular Cloning and Expression of a Purine-Specific N-Ribohydrolase from *Trypanosoma Brucei Brucei*. *J. Biol. Chem.* **1998**, *273* (4), 2118–2126. https://doi.org/10.1074/jbc.273.4.2118.

(12) Vandemeulebroucke, A.; De Vos, S.; Van Holsbeke, E.; Steyaert, J.; Versées, W. A Flexible Loop as a Functional Element in the Catalytic Mechanism of Nucleoside Hydrolase from *Trypanosoma Vivax*. *J. Biol. Chem.* **2008**, *283* (32), 22272–22282. https://doi.org/10.1074/jbc.M803705200.

(13) Figueroa-Villar, J. D.; Sales, E. M. The Importance of Nucleoside Hydrolase Enzyme (NH) in Studies to Treatment of *Leishmania* : A Review. *Chem. Biol. Interact.* **2017**, *263*, 18–27. https://doi.org/10.1016/j.cbi.2016.12.004.

(14) Shi, W.; Schramm, V. L.; Almo, S. C. Nucleoside Hydrolase from *Leishmania Major*: CLONING, EXPRESSION, CATALYTIC PROPERTIES, TRANSITION STATE INHIBITORS, AND THE 2.5-Å CRYSTAL STRUCTURE *. *J. Biol. Chem.* **1999**, *274* (30),

21114–21120. https://doi.org/10.1074/jbc.274.30.21114.

(15) Koszalka, G. W.; Krenitsky, T. A. Nucleosidases from Leishmania Donovani. Pyrimidine Ribonucleosidase, Purine Ribonucleosidase, and a Novel Purine 2'-Deoxyribonucleosidase. *J. Biol. Chem.* **1979**, *254* (17), 8185–8193. https://doi.org/10.1016/S0021-9258(19)86874-6.

(16) Horenstein, B. A.; Parkin, D. W.; Estupinan, B.; Schramm, V. L. Transition- State Analysis of Nucleoside Hydrolase from *Crithidia Fasciculata*. *Biochemistry* **1991**, *30* (44), 10788–10795. https://doi.org/10.1021/bi00108a026.

(17) Liang, L.; He, X.; Liu, G.; Tan, H. The Role of a Purine-Specific Nucleoside Hydrolase in Spore Germination of *Bacillus Thuringiensis*. *Microbiology* **2008**, *154* (5), 1333–1340. https://doi.org/10.1099/mic.0.2007/014399-0.

(18) Hansen, M. R.; Dandanell, G. Purification and Characterization of RihC, a Xanthosine–Inosine–Uridine–Adenosine-Preferring Hydrolase from *Salmonella Enterica Serovar Typhimurium*. *Biochim. Biophys. Acta BBA - Gen. Subj.* **2005**, *1723* (1–3), 55–62. https://doi.org/10.1016/j.bbagen.2005.01.012.

(19) Arivett, B.; Farone, M.; Masiragani, R.; Burden, A.; Judge, S.; Osinloye, A.; Minici, C.; Degano, M.; Robinson, M.; Kline, P. Characterization of Inosine- Uridine Nucleoside Hydrolase (RihC) from *Escherichia coli*. *Biochim. Biophys. Acta* **2014**, *1844* (3), 656–662. https://doi.org/10.1016/j.bbapap.2014.01.010.

(20) Porcelli, M.; Peluso, I.; Marabotti, A.; Facchiano, A.; Cacciapuoti, G. Biochemical Characterization and Homology Modeling of a Purine-Specific Ribonucleoside Hydrolase from the Archaeon *Sulfolobus Solfataricus*: Insights into Mechanisms of Protein Stabilization. *Arch. Biochem. Biophys.* **2009**, *483* (1), 55–65. https://doi.org/10.1016/j.abb.2008.12.005.

(21) Kurtz, J.-E.; Exinger, F.; Erbs, P.; Jund, R. The URH1 Uridine Ribohydrolase of *Saccharomyces*

Cerevisiae. *Curr. Genet.* **2002**, *41*, 132–141. https://doi.org/10.1007/s00294-002-0296-9.

(22) Ribeiro, J. M. C.; Valenzuela, J. G. The Salivary Purine Nucleosidase of the Mosquito, Aedes Aegypti. *Insect Biochem. Mol. Biol.* **2003**, *33* (1), 13–22. https://doi.org/10.1016/s0965-1748(02)00078-4.

(23) Jung, B.; Flörchinger, M.; Kunz, H.-H.; Traub, M.; Wartenberg, R.; Jeblick, W.; Neuhaus, H. E.; Möhlmann, T. Uridine-Ribohydrolase Is a Key Regulator in the Uridine Degradation Pathway of *Arabidopsis*. *Plant Cell* **2009**, *21* (3), 876–891. https://doi.org/10.1105/tpc.108.062612.

(24) Jung, B.; Hoffmann, C.; Möhlmann, T. Arabidopsis Nucleoside Hydrolases Involved in Intracellular and Extracellular Degradation of Purines. *Plant J. Cell Mol. Biol.* **2011**, *65* (5), 703–711. https://doi.org/10.1111/j.1365- 313X.2010.04455.x.

(25) Ashihara, H.; Stasolla, C.; Fujimura, T.; Crozier, A. Purine Salvage in Plants. *Phytochemistry* **2018**, *147*, 89–124. https://doi.org/10.1016/j.phytochem.2017.12.008.

(26) Mach, J. Uridine Ribohydrolase and the Balance between Nucleotide Degradation and Salvage. *Plant Cell* **2009**, *21* (3), 699. https://doi.org/10.1105/tpc.109.210312.

(27) Abusamhadneh, E.; McDonald, N. E.; Kline, P. C. Isolation and Characterization of Adenosine Nucleosidase from Yellow Lupin (*Lupinus Luteus*). *Plant Sci.* **2000**, *153* (1), 25–32. https://doi.org/10.1016/S0168- 9452(99)00240-X.

(28) Achar, B. S.; Vaidyanathan, C. S. Purification and Properties of Uridine Hydrolase from Mung-Bean (Phaseolus Radiatus) Seedlings. *Arch. Biochem. Biophys.* **1967**, *119*, 356–362. https://doi.org/10.1016/0003-9861(67)90465-1.

(29) Jung, B.; Hoffmann, C.; Möhlmann, T. *Arabidopsis* Nucleoside Hydrolases Involved in Intracellular and Extracellular Degradation of Purines. *Plant J. Cell Mol. Biol.* **2010**, *65*, 703–711. https://doi.org/10.1111/j.1365- 313X.2010.04455.x.

(30) The Herbicide Well Runs Dry. *CEN Glob. Enterp.* **2022**, *100* (22), 22–28. https://doi.org/10.1021/cen-10022-cover.

(31) de Araújo-Ramos, A. T.; Passoni, M. T.; Romano, M. A.; Romano, R. M.; Martino-Andrade, A. J. Controversies on Endocrine and Reproductive Effects of Glyphosate and Glyphosate-Based Herbicides: A Mini-Review. *Front. Endocrinol.* **2021**, *12*, 627210. https://doi.org/10.3389/fendo.2021.627210.

(32) Costas-Ferreira, C.; Durán, R.; Faro, L. R. F. Toxic Effects of Glyphosate on the Nervous System: A Systematic Review. *Int. J. Mol. Sci.* **2022**, *23* (9), 4605. https://doi.org/10.3390/ijms23094605.

(33) Stokstad, E. Jury Verdicts Cloud Future of Popular Herbicide. *Science* **2019**, *364* (6442), 717–718. https://doi.org/10.1126/science.364.6442.717.

(34) Wolfenbarger, L. L.; Phifer, P. R. The Ecological Risks and Benefits of Genetically Engineered Plants. *Science* **2000**, *290* (5499), 2088–2093. https://doi.org/10.1126/science.290.5499.2088.

(35) US EPA, O. *Herbicides*. https://www.epa.gov/caddis/herbicides (accessed 2024-07-02).

(36) Baer, K. N.; Marcel, B. J. Glyphosate. In *Encyclopedia of Toxicology (Third Edition)*; Wexler, P., Ed.; Academic Press: Oxford, 2014; pp 767–769. https://doi.org/10.1016/B978-0-12-386454-3.00148-2.

(37) Haschemeyer, A. E. V.; Rich, A. Nucleoside Conformations: An Analysis of Steric Barriers to Rotation about the Glycosidic Bond. *J. Mol. Biol.* **1967**, *27* (2), 369–384. https://doi.org/10.1016/0022-2836(67)90026-5.

(38) Levitt, M.; Warshel, A. Extreme Conformational Flexibility of the Furanose Ring in DNA and RNA. *J. Am. Chem. Soc.* **1978**, *100* (9), 2607–2613. https://doi.org/10.1021/ja00477a004.

(39) Olson, W. K.; Sussman, J. L. How Flexible Is the Furanose Ring? 1. A Comparison of

Experimental and Theoretical Studies. *J. Am. Chem. Soc.* **1982**, *104* (1), 270–278. https://doi.org/10.1021/ja00365a049.

(40) Gopaul, D. N.; Meyer, S. L.; Degano, M.; Sacchettini, J. C.; Schramm, V. L. Inosine−Uridine Nucleoside Hydrolase from *Crithidia Fasciculata*. Genetic Characterization, Crystallization, and Identification of Histidine 241 as a Catalytic Site Residue. *Biochemistry* **1996**, *35* (19), 5963–5970. https://doi.org/10.1021/bi952998u.

(41) Bates, C.; Kendrick, Z.; McDonald, N.; Kline, P. C. Transition State Analysis of Adenosine Nucleosidase from Yellow Lupin (*Lupinus Luteus*). *Phytochemistry* **2006**, *67* (1), 5–12. https://doi.org/10.1016/j.phytochem.2005.10.006.

(42) Berti, P. J.; McCann, J. A. B. Toward a Detailed Understanding of Base Excision Repair Enzymes: Transition State and Mechanistic Analyses of N- Glycoside Hydrolysis and N-Glycoside Transfer. *Chem. Rev.* **2006**, *106* (2), 506–555. https://doi.org/10.1021/cr040461t.

(43) Berti, P. J.; McCann, J. A. B. Toward a Detailed Understanding of Base Excision Repair Enzymes: Transition State and Mechanistic Analyses of N- Glycoside Hydrolysis and N-Glycoside Transfer. *Chem. Rev.* **2006**, *106* (2), 506–555. https://doi.org/10.1021/cr040461t.

(44) Loverix, S.; Geerlings, P.; McNaughton, M.; Augustyns, K.; Vandemeulebroucke, A.; Steyaert, J.; Versées, W. Substrate-Assisted Leaving Group Activation in Enzyme-Catalyzed N-Glycosidic Bond Cleavage. *J. Biol. Chem.* **2005**, *280* (15), 14799–14802. https://doi.org/10.1074/jbc.M413231200.

(45) Mazzella, L. J.; Parkin, D. W.; Tyler, P. C.; Furneaux, R. H.; Schramm, V. L. Mechanistic Diagnoses of N-Ribohydrolases and Purine Nucleoside Phosphorylase. *J. Am. Chem. Soc.* **1996**, *118* (8), 2111–2112. https://doi.org/10.1021/ja953537z.

(46) Mazumder, D.; Kahn, K.; Bruice, T. C. Computer Simulations of Trypanosomal Nucleoside Hydrolase: Determination of the Protonation State of the Bound Transition-State Analogue. *J.*

Am. Chem. Soc. **2002**, *124* (30), 8825–8833. https://doi.org/10.1021/ja020312x.

(47) Kopečná, M.; Blaschke, H.; Kopečný, D.; Vigouroux, A.; Končitíková, R.; Novák, O.; Kotland, O.; Strnad, M.; Moréra, S.; von Schwartzenberg, K. Structure and Function of Nucleoside Hydrolases from *Physcomitrella Patens* and Maize Catalyzing the Hydrolysis of Purine, Pyrimidine, and Cytokinin Ribosides. *Plant Physiol.* **2013**, *163* (4), 1568–1583. https://doi.org/10.1104/pp.113.228775.

(48) Degano, M.; Almo, S.C., Sacchettini, J.C.; Schramm, V.L. Trypanosomal Nucleoside Hydrolase. A Novel Mechanism from the Structure with a Transition-State Inhibitor. *Biochemistry* **1998**, 37, 6277-6285. https://pubs-acs- org.ezproxy.mtsu.edu/doi/pdf/10.1021/bi973012e (accessed 2022-10-28).

(49) *World malaria report 2022*. https://www.who.int/publications-detail-redirect/9789240064898 (accessed 2024-03-20).

(50) *Trypanosomiasis, human African (sleeping sickness)*. https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human- african-(sleeping-sickness) (accessed 2024-03-20).

(51) *Control of the leishmaniases WHO TRS n° 949*. https://www.who.int/publications-detail-redirect/WHO-TRS-949 (accessed 2024-03-20).

(52) Degano, M.; Gopaul, D. N.; Scapin, G.; Schramm, V. L.; Sacchettini, J. C. Three-Dimensional Structure of the Inosine-Uridine Nucleoside N- Ribohydrolase from *Crithidia Fasciculata*. *Biochemistry* **1996**, *35* (19), 5971–5981. https://doi.org/10.1021/bi952999m.

(53) Alves, M. A.; Nirma, C.; Moreira, M. M.; Soares, R. O.; Pascutti, P. G.; Noël, F.; Costa, P. R. R.; Sant'Anna, C. M. R.; Barreiro, E. J.; Lima, L. M.; Tinoco, L. W. Non-Competitive Inhibitor of Nucleoside Hydrolase from Leishmania Donovani Identified by Fragment-Based Drug Discovery. *RSC Adv.* **2016**, *6* (90), 87738–87744. https://doi.org/10.1039/C6RA15143D.

(54) McAtee, C. P.; Seid, C. A.; Hammond, M.; Hudspeth, E.; Keegan, B. P.; Liu, Z.; Wei, J.; Zhan, B.; Arjona-Sabido, R.; Cruz-Chan, V.; Dumonteil, E.; Hotez, P. J.; Bottazzi,

M. E. Expression, Purification, Immunogenicity and Protective Efficacy of a Recombinant Nucleoside Hydrolase from *Leishmania Donovani*, a Vaccine Candidate for Preventing Cutaneous Leishmaniasis. *Protein Expr. Purif.* **2017**, *130*, 129–136. https://doi.org/10.1016/j.pep.2016.10.008.

(55) Rennó, M. N.; França, T. C. C.; Nico, D.; Palatnik-de-Sousa, C. B.; Tinoco, L. W.; Figueroa-Villar, J. D. Kinetics and Docking Studies of Two Potential New Inhibitors of the Nucleoside Hydrolase *from Leishmania Donovani*. *Eur. J. Med. Chem.* **2012**, *56*, 301–307. https://doi.org/10.1016/j.ejmech.2012.07.052.

(56) Palatnik-de-Sousa, C. B. Nucleoside Hydrolase NH 36: A Vital Enzyme for the Leishmania Genus in the Development of T-Cell Epitope Cross- Protective Vaccines. *Front. Immunol.* **2019**, *10*, 813. https://doi.org/10.3389/fimmu.2019.00813.

(57) Ogbunude, P. O.; Ikediobi, C. O.; Ukoha, A. I. Adenosine Cycle in African Trypanosomes. *Ann. Trop. Med. Parasitol.* **1985**, *79* (1), 7–11. https://doi.org/10.1080/00034983.1985.11811883.

(58) Ha, C. H. H.; Fatima, A.; Gaurav, A. *In Silico* Investigation of Flavonoids as Potential Trypanosomal Nucleoside Hydrolase Inhibitors. *Adv. Bioinforma.* **2015**, *2015*, 1–10. https://doi.org/10.1155/2015/826047.

(59) Parkin, D. W. Purine-Specific Nucleoside N-Ribohydrolase from Trypanosoma Brucei Brucei: PURIFICATION, SPECIFICITY, AND KINETIC MECHANISM. *J. Biol. Chem.* **1996**, *271* (36), 21713–21719. https://doi.org/10.1074/jbc.271.36.21713.

(60) Setlow, P. Summer Meeting 2013 - When the Sleepers Wake: The Germination of Spores of *Bacillus* Species. *J. Appl. Microbiol.* **2013**, *115* (6), 1251–1268. https://doi.org/10.1111/jam.12343.

(61) Karami, M.; Jalali, C.; Mirzaie, S. Combined Virtual Screening, MMPBSA, Molecular Docking

and Dynamics Studies against Deadly Anthrax: An *in Silico* Effort to Inhibit *Bacillus Anthracis* Nucleoside Hydrolase. *J. Theor. Biol.* **2017**, *420*, 180–189. https://doi.org/10.1016/j.jtbi.2017.03.010.

(62) Ducati, R. G.; Basso, L. A.; Santos, D. S.; de Azevedo, W. F. Crystallographic and Docking Studies of Purine Nucleoside Phosphorylase from Mycobacterium Tuberculosis. *Bioorg. Med. Chem.* **2010**, *18* (13), 4769–4774. https://doi.org/10.1016/j.bmc.2010.05.009.

(63) Wink, P. L.; Sanchez Quitian, Z. A.; Rosado, L. A.; Rodrigues, V. da S.; Petersen, G. O.; Lorenzini, D. M.; Lipinski-Paes, T.; Saraiva Macedo Timmers, L. F.; de Souza, O. N.; Basso, L. A.; Santos, D. S. Biochemical Characterization of Recombinant Nucleoside Hydrolase from *Mycobacterium Tuberculosis* H37Rv. *Arch. Biochem. Biophys.* **2013**, *538* (2), 80–94. https://doi.org/10.1016/j.abb.2013.08.011.

(64) Arivett, B. A. Site-Directed Mutagenesis and HPLC Analysis of Inosine- Uridine Nucleoside Hydrolase RihC of *Escherichia coli*. *undefined* **2014**.

(65) Muzzolini, L.; Versées, W.; Tornaghi, P.; Van Holsbeke, E.; Steyaert, J.; Degano, M. New Insights into the Mechanism of Nucleoside Hydrolases from the Crystal Structure of the *Escherichia Coli* YbeK Protein Bound to the Reaction Product ᵎ. *Biochemistry* **2006**, *45* (3), 773–782. https://doi.org/10.1021/bi0511991.

(66) Petersen, C.; Møller, L. B. The RihA, RihB, and RihC Ribonucleoside Hydrolases of *Escherichia coli*: SUBSTRATE SPECIFICITY, GENE EXPRESSION, AND REGULATION*. *J. Biol. Chem.* **2001**, *276* (2), 884–894. https://doi.org/10.1074/jbc.M008300200.

(67) Iovane, E.; Giabbai, B.; Muzzolini, L.; Matafora, V.; Fornili, A.; Minici, C.; Giannese, F.; Degano, M. Structural Basis for Substrate Specificity in Group I Nucleoside Hydrolases. *Biochemistry* **2008**, *47* (15), 4418–4426. https://doi.org/10.1021/bi702448s.

(68) Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. The SWISS-MODEL Workspace: A Web-Based Environment for Protein Structure Homology Modelling. *Bioinforma. Oxf. Engl.* **2006**, *22* (2), 195–201. https://doi.org/10.1093/bioinformatics/bti770.

(69) Lenz, S. A. P.; Wetmore, S. D. Structural Explanation for the Tunable Substrate Specificity of an *E. Coli* Nucleoside Hydrolase: Insights from Molecular Dynamics Simulations. *J. Comput. Aided Mol. Des.* **2018**, *32* (12), 1375–1388. https://doi.org/10.1007/s10822-018-0178-y.

(70) Szuwart, M.; Starzyńska, E.; Pietrowska-Borek, M.; Guranowski, A. Calcium- Stimulated Guanosine–Inosine Nucleosidase from Yellow Lupin (*Lupinus Luteus*). *Phytochemistry* **2006**, *67* (14), 1476–1485. https://doi.org/10.1016/j.phytochem.2006.05.021.

(71) Guranowski, A. Plant 5-Methylthioribose Kinase: Properties of the Partially Purified Enzyme from Yellow Lupin (*Lupinus Luteus* L.) Seeds. *Plant Physiol.* **1983**, *71* (4), 932–935. https://doi.org/10.1104/pp.71.4.932.

(72) Campos, A.; Rijo-Johansen, M. J.; Carneiro, M. F.; Fevereiro, P. Purification and Characterisation of Adenosine Nucleosidase from *Coffea Arabica* Young Leaves. *Phytochemistry* **2005**, *66* (2), 147–151. https://doi.org/10.1016/j.phytochem.2004.11.018.

(73) Thicklin, L.; Shamsuddin, A.; Alahmry, F.; Gezley, C.; Brown, E.; Stone, J.; Burns-Carver, E.; Kline, P. C. Purification of a Non-Specific Nucleoside Hydrolase from Alaska Pea Seeds. *Protein Expr. Purif.* **2019**, *154*, 140–146. https://doi.org/10.1016/j.pep.2018.10.009.

(74) Riegler, H.; Geserick, C.; Zrenner, R. *Arabidopsis Thaliana* Nucleosidase Mutants Provide New Insights into Nucleoside Degradation. *New Phytol.* **2011**, *191* (2), 349–359. https://doi.org/10.1111/j.1469-8137.2011.03711.x.

(75) Wienkoop, S.; Baginsky, S.; Weckwerth, W. Arabidopsis Thaliana as a Model Organism for Plant Proteome Research. *J. Proteomics* **2010**, *73* (11), 2239–2248.

https://doi.org/10.1016/j.jprot.2010.07.012.

(76) Moffatt, B. A.; Ashihara, H. Purine and Pyrimidine Nucleotide Synthesis and Metabolism. *Arab. Book* **2002**, *1*, e0018. https://doi.org/10.1199/tab.0018.

(77) Delgado-García, E.; Piedras, P.; Gómez-Baena, G.; García-Magdaleno, I. M.; Pineda, M.; Gálvez-Valdivieso, G. Nucleoside Metabolism Is Induced in Common Bean During Early Seedling Development. *Front. Plant Sci.* **2021**, *12*, 651015. https://doi.org/10.3389/fpls.2021.651015.

(78) *Gene Synthesis Handbook*. https://www.genscript.com/gsh.html (accessed 2024-04-08).

(79) Bornhorst, J. A.; Falke, J. J. Purification of Proteins Using Polyhistidine Affinity Tags. *Methods Enzymol.* **2000**, *326*, 245–254. https://doi.org/10.1016/s0076-6879(00)26058-8.

(80) Bradford, M. M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **1976**, *72*, 248–254. https://doi.org/10.1006/abio.1976.9999.

(81) Nitta, K.; Watanabe, A. Determination of Calcium Ions Tightly Bound to Proteins. *J. Chromatogr.* **1991**, *585* (1), 173–176. https://doi.org/10.1016/0021-9673(91)85073-o.

(82) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–242. https://doi.org/10.1093/nar/28.1.235.

(83) *PDB-101: Educational resources supporting molecular explorations through biology and medicine - Zardecki - 2022 - Protein Science - Wiley Online Library*. https://onlinelibrary.wiley.com/doi/10.1002/pro.4200 (accessed 2024-07-04).

(84) Benkert, P.; Biasini, M.; Schwede, T. Toward the Estimation of the Absolute Quality of Individual Protein Structure Models. *Bioinforma. Oxf. Engl.* **2011**, *27* (3), 343–350. https://doi.org/10.1093/bioinformatics/btq662.

(85) Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Gallo Cassarino, T.; Bertoni, M.; Bordoli, L.; Schwede, T. SWISS- MODEL: Modelling Protein Tertiary and Quaternary Structure Using Evolutionary Information. *Nucleic Acids Res.* **2014**, *42* (Web Server issue), W252-258. https://doi.org/10.1093/nar/gku340.

(86) Bienert, S.; Waterhouse, A.; de Beer, T. A. P.; Tauriello, G.; Studer, G.; Bordoli, L.; Schwede, T. The SWISS-MODEL Repository-New Features and Functionality. *Nucleic Acids Res.* **2017**, *45* (D1), D313–D319. https://doi.org/10.1093/nar/gkw1132.

(87) Guex, N.; Peitsch, M. C.; Schwede, T. Automated Comparative Protein Structure Modeling with SWISS-MODEL and Swiss-PdbViewer: A Historical Perspective. *Electrophoresis* **2009**, *30 Suppl 1*, S162-173. https://doi.org/10.1002/elps.200900140.

(88) Kiefer, F.; Arnold, K.; Künzli, M.; Bordoli, L.; Schwede, T. The SWISS- MODEL Repository and Associated Resources. *Nucleic Acids Res.* **2009**, *37* (suppl_1), D387–D392. https://doi.org/10.1093/nar/gkn750.

(89) Mariani, V.; Biasini, M.; Barbato, A.; Schwede, T. lDDT: A Local Superposition-Free Score for Comparing Protein Structures and Models Using Distance Difference Tests. *Bioinforma. Oxf. Engl.* **2013**, *29* (21), 2722–2728. https://doi.org/10.1093/bioinformatics/btt473.

(90) Studer, G.; Rempfer, C.; Waterhouse, A. M.; Gumienny, R.; Haas, J.; Schwede, T. QMEANDisCo-Distance Constraints Applied on Model Quality Estimation. *Bioinforma. Oxf. Engl.* **2020**, *36* (6), 1765–1771. https://doi.org/10.1093/bioinformatics/btz828.

(91) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res.* **2018**, *46* (W1), W296–W303. https://doi.org/10.1093/nar/gky427.

(92) Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. SwissParam: A Fast Force Field Generation Tool for Small Organic Molecules. *J. Comput. Chem.* **2011**, *32* (11), 2359–2368. https://doi.org/10.1002/jcc.21816.

(93) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

(94) Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: Making Protein Folding Accessible to All. *Nat. Methods* **2022**, *19* (6), 679–682. https://doi.org/10.1038/s41592-022-01488-1.

(95) Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Žídek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein- Sequence Space with High-Accuracy Models. *Nucleic Acids Res.* **2022**, *50* (D1), D439–D444. https://doi.org/10.1093/nar/gkab1061.

(96) Yuan, D.; Ahamed, A.; Burgin, J.; Cummins, C.; Devraj, R.; Gueye, K.; Gupta, D.; Gupta, V.; Haseeb, M.; Ihsan, M.; Ivanov, E.; Jayathilaka, S.; Kadhirvelu, V. B.; Kumar, M.; Lathi, A.; Leinonen, R.; McKinnon, J.; Meszaros, L.; O'Cathail, C.; Ouma, D.; Paupério, J.; Pesant, S.; Rahman, N.; Rinck, G.; Selvakumar, S.; Suman, S.; Sunthornyotin, Y.; Ventouratou, M.; Vijayaraja, S.; Waheed, Z.; Woollard, P.; Zyoud, A.; Burdett, T.; Cochrane, G. The European

Nucleotide Archive in 2023. *Nucleic Acids Res.* **2024**, *52* (D1), D92–D97. https://doi.org/10.1093/nar/gkad1067.

(97) Madeira, F.; Madhusoodanan, N.; Lee, J.; Eusebi, A.; Niewielska, A.; Tivey, A. R. N.; Lopez, R.; Butcher, S. The EMBL-EBI Job Dispatcher Sequence Analysis Tools Framework in 2024. *Nucleic Acids Res.* **2024**, *52* (W1), W521–W525. https://doi.org/10.1093/nar/gkae241.

(98) Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E. E. PubChem 2023 Update. *Nucleic Acids Res.* **2023**, *51* (D1), D1373– D1380. https://doi.org/10.1093/nar/gkac956.

(99) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera--a Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612. https://doi.org/10.1002/jcc.20084.

(100) Meng, E. C.; Pettersen, E. F.; Couch, G. S.; Huang, C. C.; Ferrin, T. E. Tools for Integrated Sequence-Structure Analysis with UCSF Chimera. *BMC Bioinformatics* **2006**, *7*, 339. https://doi.org/10.1186/1471-2105-7-339.

(101) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455–461. https://doi.org/10.1002/jcc.21334.

(102) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91* (1), 43–56. https://doi.org/10.1016/0010- 4655(95)00042-E.

(103) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26* (16), 1701–1718. https://doi.org/10.1002/jcc.20291.

(104) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4* (3), 435–447. https://doi.org/10.1021/ct700301q.

(105) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: A High-Throughput and Highly Parallel Open-Source Molecular Simulation Toolkit. *Bioinforma. Oxf. Engl.* **2013**, *29* (7), 845–854. https://doi.org/10.1093/bioinformatics/btt055.

(106) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001.

(107) Lemkul, J. A. From Proteins to Perturbed Hamiltonians: A Suite of Tutorials for the GROMACS-2018 Molecular Simulation Package [Article v1.0]. *Living J. Comput. Mol. Sci.* **2019**, *1* (1), 5068–5068. https://doi.org/10.33011/livecoms.1.1.5068.

(108) Mackerell, A. D. Empirical Force Fields for Biological Macromolecules: Overview and Issues. *J. Comput. Chem.* **2004**, *25* (13), 1584–1604. https://doi.org/10.1002/jcc.20082.

(109) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; Mackerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone φ, ψ and Side- Chain χ(1) and χ(2) Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8* (9), 3257–3273. https://doi.org/10.1021/ct300400x.

(110) Huang, J.; MacKerell, A. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34*. https://doi.org/10.1002/jcc.23354.

(111) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14* (1), 71–73. https://doi.org/10.1038/nmeth.4067.

(112) *Grace Home*. https://plasma-gate.weizmann.ac.il/Grace/ (accessed 2024-07-04).

(113) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.

(114) Grant, B. J.; Rodrigues, A. P. C.; ElSawy, K. M.; McCammon, J. A.; Caves, L. S. D. Bio3d: An R Package for the Comparative Analysis of Protein Structures. *Bioinforma. Oxf. Engl.* **2006**, *22* (21), 2695–2696. https://doi.org/10.1093/bioinformatics/btl461.

(115) Grant, B. J.; Skjaerven, L.; Yao, X.-Q. The Bio3D Packages for Structural Bioinformatics. *Protein Sci. Publ. Protein Soc.* **2021**, *30* (1), 20–30. https://doi.org/10.1002/pro.3923.

(116) Skjærven, L.; Jariwala, S.; Yao, X.-Q.; Grant, B. J. Online Interactive Analysis of Protein Structure Ensembles with Bio3D-Web. *Bioinformatics* **2016**, *32* (22), 3510–3512. https://doi.org/10.1093/bioinformatics/btw482.

(117) Skjærven, L.; Yao, X.-Q.; Scarabelli, G.; Grant, B. J. Integrating Protein Structural Dynamics and Evolutionary Analysis with Bio3D. *BMC Bioinformatics* **2014**, *15* (1), 399. https://doi.org/10.1186/s12859-014-0399-6.

(118) Gene_synthesis_handbook.Pdf. https://www.genscript.com/gsfiles/gene_synthesis_handbook.pdf (accessed 2022-10-28).

(119) Caruthers, M. H.; Beaucage, S. L.; Becker, C.; Efcavitch, J. W.; Fisher, E. F.; Galluppi, G.; Goldman, R.; deHaseth, P.; Matteucci, M.; McBride, L. Deoxyoligonucleotide Synthesis via the Phosphoramidite Method. *Gene Amplif. Anal.* **1983**, *3*, 1–26.

(120) Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A. ExPASy: The Proteomics Server for in-Depth Protein Knowledge and Analysis. *Nucleic Acids Res.* **2003**, *31* (13), 3784–3788. https://doi.org/10.1093/nar/gkg563.

(121) Degano, M. Structure, Oligomerization and Activity Modulation in N- Ribohydrolases. *Int. J. Mol. Sci.* **2022**, *23* (5), 2576. https://doi.org/10.3390/ijms23052576.

(122) Asai, T.; O'Sullivan, W. J.; Tatibana, M. A Potent Nucleoside Triphosphate Hydrolase from the Parasitic Protozoan Toxoplasma Gondii. Purification, Some Properties, and Activation by Thiol Compounds. *J. Biol. Chem.* **1983**, *258* (11), 6816–6822. https://doi.org/10.1016/S0021-9258(18)32295-6.

(123) Gast, K.; Zirwer, D.; MÜller-Frohne, M.; Damaschun, G. Compactness of the Kinetic Molten Globule of Bovineα-Lactalbumin: A Dynamic Light Scattering Study. *Protein Sci.* **1998**, *7* (9), 2004–2011. https://doi.org/10.1002/pro.5560070917.

(124) Pizones Ruiz-Henestrosa, V. M.; Martinez, M. J.; Patino, J. M. R.; Pilosof, A. M. R. A Dynamic Light Scattering Study on the Complex Assembly of Glycinin Soy Globulin in Aqueous Solutions. *J. Am. Oil Chem. Soc.* **2012**, *89* (7), 1183–1191. https://doi.org/10.1007/s11746-012-2029-7.

(125) Hong, P.; Koza, S.; Bouvier, E. S. P. A Review Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and Their Aggregates. *J. Liq. Chromatogr. Relat. Technol.* **2012**, *35* (20), 2923–2950. https://doi.org/10.1080/10826076.2012.743724.

(126) Bryant, D. T. W. Quin 2: The Dissociation Constants of Its Ca2+ and Mg2+ Complexes and Its Use in a Fluorimetric Method for Determining the Dissociation of Ca2+-Protein Complexes. *Biochem. J.* **1985**, *226* (2), 613–616. https://doi.org/10.1042/bj2260613.

(127) Grynkiewicz, G.; Poenie, M.; Tsien, R. Y. A New Generation of Ca2+ Indicators with Greatly Improved Fluorescence Properties. *J. Biol. Chem.* **1985**, *260* (6), 3440–3450. https://doi.org/10.1016/S0021-9258(19)83641-4.

(128) DeWolf, W. E. Jr.; Emig, F. A.; Schramm, V. L. AMP Nucleosidase: Kinetic Mechanism and Thermodynamics. *Biochemistry* **1986**, *25* (14), 4132–4140. https://doi.org/10.1021/bi00362a022.

(129) Thicklin, L.; Shamsuddin, A.; Alahmry, F.; Gezley, C.; Brown, E.; Stone, J.; Burns-Carver, E.; Kline, P. C. Purification of a Non-Specific Nucleoside Hydrolase from Alaska Pea Seeds. *Protein Expr. Purif.* **2019**, *154*, 140–146. https://doi.org/10.1016/j.pep.2018.10.009.

(130) Prior, J. J.; Santi, D. V. On the Mechanism of the Acid-Catalyzed Hydrolysis of Uridine to Uracil. Evidence for 6-Hydroxy-5,6-Dihydrouridine Intermediates. *J. Biol. Chem.* **1984**, *259* (4), 2429–2434. https://doi.org/10.1016/S0021-9258(17)43370-9.

(131) Moffatt, B.; Somerville, C. Positive Selection for Male-Sterile Mutants of *Arabidopsis* Lacking Adenine Phosphoribosyl Transferase Activity. *Plant Physiol.* **1988**, *86* (4), 1150–1154. https://doi.org/10.1104/pp.86.4.1150.

(132) Saenger, W. *Principles of Nucleic Acid Structure*; Cantor, C. R., Series Ed.; Springer Advanced Texts in Chemistry; Springer: New York, NY, 1984. https://doi.org/10.1007/978-1-4612-5190-3.

(133) Sinnott, M. *Carbohydrate Chemistry and Biochemistry*; 2013.

(134) Lescrinier, E.; Froeyen, M.; Herdewijn, P. SURVEY AND SUMMARY: Difference in Conformational Diversity between Nucleic Acids with a Six- Membered ‘Sugar’ Unit and Natural ‘Furanose’ Nucleic Acids. *Nucleic Acids Res.* **2003**, *31* (12), 2975–2989.

(135) Structure and conformation of nucleosides and nucleotides and their analogs as

determined by x-ray diffraction - *PubMed*. https://pubmed.ncbi.nlm.nih.gov/1103684/ (accessed 2024-02-28).

(136) An Vandemeulebroucke. Kinetic Analysis of the Purine Nucleoside Hydrolase Mechanism Dissertation (2008).

(137) Chen, N.; Ge, H.; Xu, J.; Cao, Z.; Wu, R. Loop Motion and Base Release in Purine-Specific Nucleoside Hydrolase: A Molecular Dynamics Study. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2013**, *1834* (6), 1117–1124. https://doi.org/10.1016/j.bbapap.2013.02.005.

(138) Núñez, S.; Antoniou, D.; Schramm, V. L.; Schwartz, S. D. Promoting Vibrations in Human Purine Nucleoside Phosphorylase. A Molecular Dynamics and Hybrid Quantum Mechanical/Molecular Mechanical Study. *J. Am. Chem. Soc.* **2004**, *126* (48), 15720–15729. https://doi.org/10.1021/ja0457563.

(139) Fan, F.; Chen, N.; Wang, Y.; Wu, R.; Cao, Z. QM/MM and MM MD Simulations on the Pyrimidine-Specific Nucleoside Hydrolase: A Comprehensive Understanding of Enzymatic Hydrolysis of Uridine. *J. Phys. Chem. B* **2018**, *122* (3), 1121–1131. https://doi.org/10.1021/acs.jpcb.7b10524.

(140) Wu, R.; Gong, W.; Liu, T.; Zhang, Y.; Cao, Z. QM/MM Molecular Dynamics Study of Purine-Specific Nucleoside Hydrolase. *J. Phys. Chem. B* **2012**, *116* (6), 1984–1991. https://doi.org/10.1021/jp211403j.

(141) Ducati, R. G.; Basso, L. A.; Santos, D. S.; de Azevedo, W. F. Crystallographic and Docking Studies of Purine Nucleoside Phosphorylase from Mycobacterium Tuberculosis. *Bioorg. Med. Chem.* **2010**, *18* (13), 4769–4774. https://doi.org/10.1016/j.bmc.2010.05.009.

(142) Mazumder, D.; Bruice, T. C. Exploring Nucleoside Hydrolase Catalysis *in Silico* :

Molecular Dynamics Study of Enzyme-Bound Substrate and Transition State. *J. Am. Chem. Soc.* **2002**, *124* (49), 14591–14600. https://doi.org/10.1021/ja021088e.

(143) Mancini, D. T.; Matos, K. S.; da Cunha, E. F. F.; Assis, T. M.; Guimarães, A.P.; França, T. C. C.; Ramalho, T. C. Molecular Modeling Studies on Nucleoside Hydrolase from the Biological Warfare Agent *Brucella Suis*. *J. Biomol. Struct. Dyn.* **2012**, *30* (1), 125–136. https://doi.org/10.1080/07391102.2012.674293.

(144) Brambley, C. A.; Bolay, A. L.; Salvo, H.; Jansch, A. L.; Yared, T. J.; Miller, J. M.; Wallen, J. R.; Weiland, M. H. Structural Characterization of *Sphingomonas* Sp. KT-1 PahZ1-Catalyzed Biodegradation of Thermally Synthesized Poly(Aspartic Acid). *ACS Sustain. Chem. Eng.* **2020**, acssuschemeng.0c01158. https://doi.org/10.1021/acssuschemeng.0c01158.

(145) Brambley, C. A.; Yared, T. J.; Gonzalez, M.; Jansch, A. L.; Wallen, J. R.; Weiland, M. H.; Miller, J. M. *Sphingomonas* Sp. KT-1 PahZ2 Structure Reveals a Role for Conformational Dynamics in Peptide Bond Hydrolysis. *J. Phys. Chem. B* **2021**, *125* (22), 5722–5739. https://doi.org/10.1021/acs.jpcb.1c01216.

(146) Wang, S.; Hu, P.; Zhang, Y. Ab Initio Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulation of Enzyme Catalysis: The Case of Histone Lysine Methyltransferase SET7/9. *J. Phys. Chem. B* **2007**, *111* (14), 3758–3764. https://doi.org/10.1021/jp067147i.

(147) Hu, P.; Wang, S.; Zhang, Y. Highly Dissociative and Concerted Mechanism for the Nicotinamide Cleavage Reaction in Sir2Tm Enzyme Suggested by Ab Initio QM/MM Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2008**, *130* (49), 16721–16728. https://doi.org/10.1021/ja807269j.

(148) Ke, Z.; Guo, H.; Xie, D.; Wang, S.; Zhang, Y. Ab Initio QM/MM Free- Energy Studies of Arginine Deiminase Catalysis: The Protonation State of the Cys Nucleophile. *J. Phys. Chem. B*

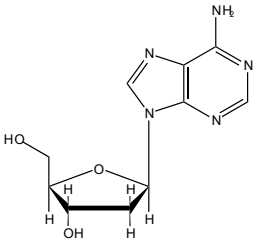
2011, *115* (13), 3725–3733. https://doi.org/10.1021/jp200843s.

(149) Chen, N.; Zhao, Y.; Lu, J.; Wu, R.; Cao, Z. Mechanistic Insights into the Rate-Limiting Step in Purine-Specific Nucleoside Hydrolase. *J. Chem. Theory Comput.* **2015**, *11* (7), 3180–3188. https://doi.org/10.1021/acs.jctc.5b00045.

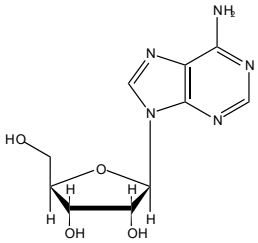
APPENDIX A

Purines

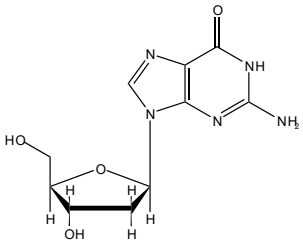
Deoxyadenosine



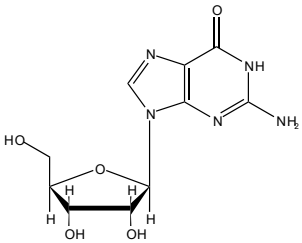
Adenosine



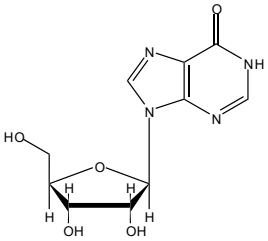
Deoxyguanosine



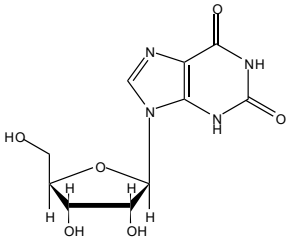
Guanosine



Inosine

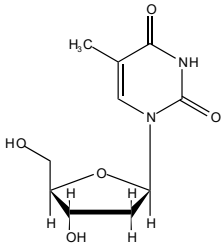


Xanthosine

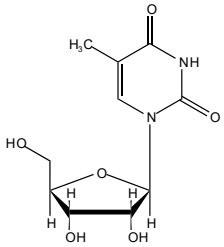


Pyrimidines

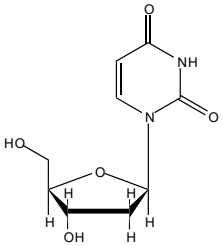
Thymidine



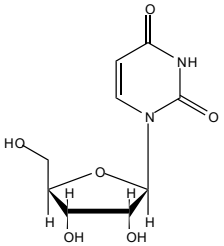
5-Methyluridine



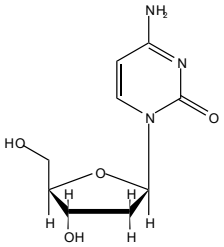
Deoxyuridine



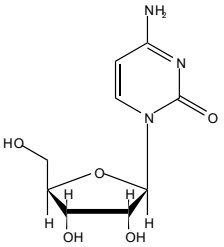
Uridine



Deoxycytidine

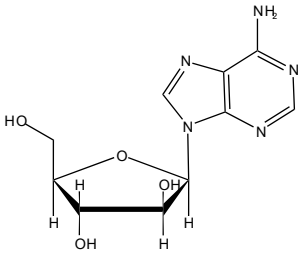


Cytidine

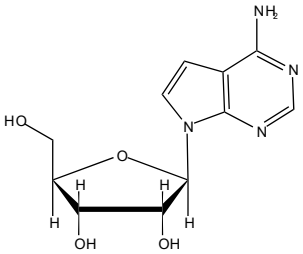


APPENDIX B

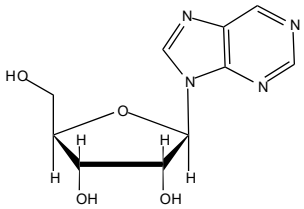
9-β-D-Arabinofuranosyladenine (Vidarabine)



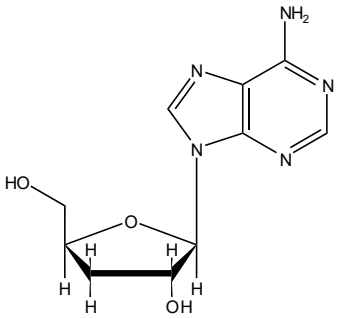
7-Deazaadenosine (Tubercidin)



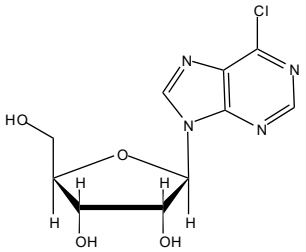
Purine Riboside (Nebularine)



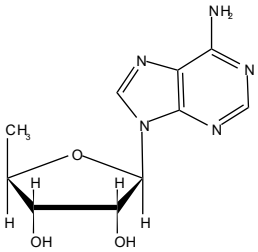
3'-Deoxyadenosine (Cordycepin)



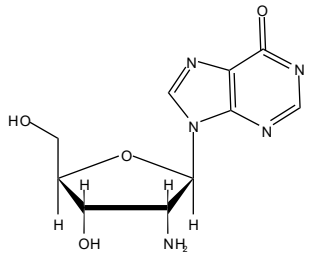
6-Chloropurine Riboside



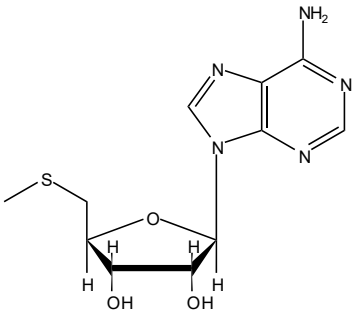
5'-Deoxyadenosine



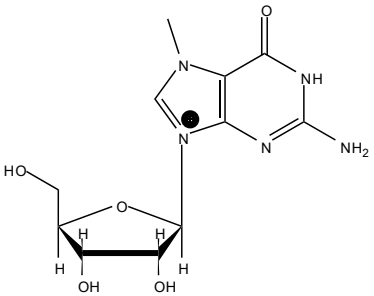
2'-Amino-2'-deoxyinosine



5'-Deoxy-5-'methylthioadenosine

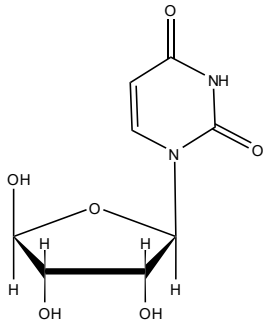


7-Methylguanosine

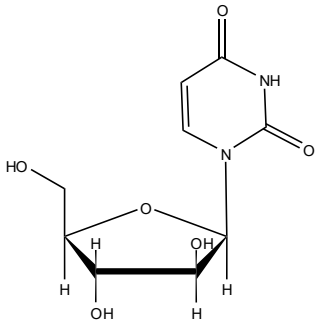


APPENDIX C

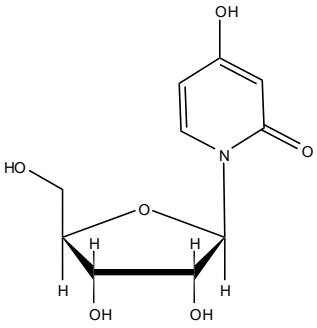
1β-Erythrouridine



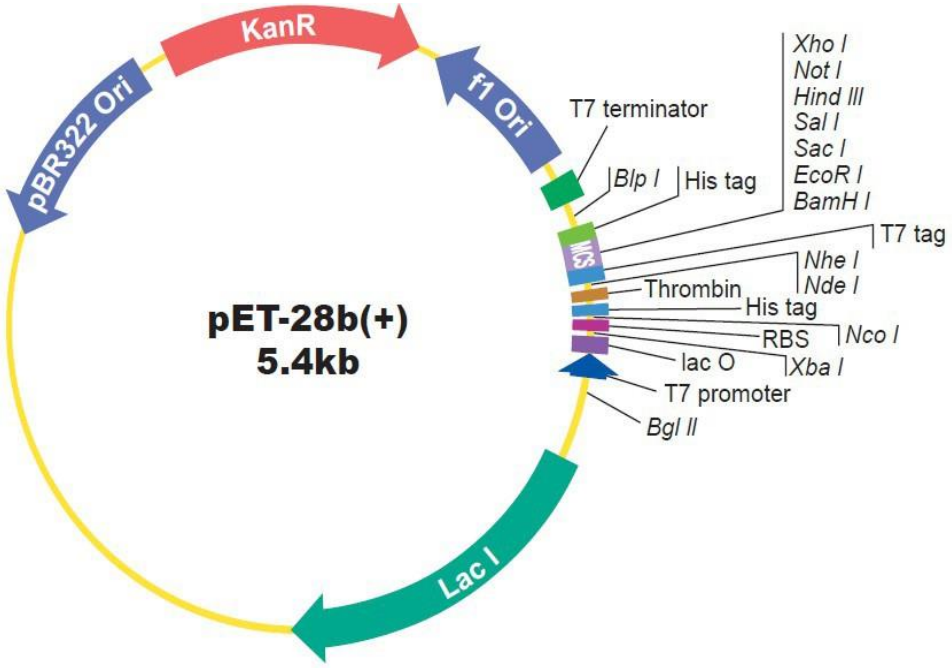
1β-D-Arabinofuranosyluracil



3-Deazauridine



APPENDIX D



Bgl II — T7 promoter → — lac operator — *Xba* I — rbs

--- AG ATC TCG ATC CCG CGA AAT TAA TAC GAC TCA CTA TAG GGG AAT TGT GAG CGG ATA ACA ATT CCC CTC TAG AAA TAA TTT TGT TTA ACT TTA AGA AGG AGA

Nco I — His tag — *Nde* I — *Nhe* I — T7 tag

TAT ACC ATG GGC AGC AGC CAT CAT CAT CAT CAT CAC AGC AGC GGC CTG GTG CCG CGC GGC AGC CAT ATG GCT AGC ATG ACT GGT GGA CAG CAA ATG GGT

M G S S H H H H H H S S G L V P R G S H M A S M T G G E E M G

thrombin site

BamH I — *EcoR* I — *Sac* I — *Sal* I — *Hind* III — *Not* I — *Xho* I — His tag

CGG GAT CCG AAT TCG AGC TCC GTC GAC AAG CTT GCG GCC GCA CTC GAG CAC CAC CAC CAC CAC CAC TGA GAT CCG GCT GCT AAC AAA GCC CGA AAG GAA

R D P N S S S V D K L A A A L E H H H H H H Stop

Blp I — T7 terminator

GCT GAG TTG GCT GCT GCC ACC GCT GAG CAA TAA CTA GCA TAA CCC CTT GGG GCC TCT AAA CGG GTC TTG AGG GGT TTT TTG ---

APPENDIX E

RMSD plots for URH1 simulations