Characterization of Metal Binding Sites Within the Protein Only RNase P

Ganga Nilmini Kumari Wijesinghe Mudiyanselage
Marquette University

CHARACTERIZATION OF METAL BINDING SITES WITHIN THE PROTEIN ONLY RNASE P

by

Ganga Nilmini Kumari Wijesinghe Mudiyanselage, B.Sc
University of Kelaniya, Sri Lanka

A Thesis submitted to the Faculty of the Graduate School,
Marquette University,
in Partial Fulfillment of the Requirements for
the Degree of Master of Science

Milwaukee, Wisconsin

August 2019

ABSTRACT

CHARACTERIZATION OF METAL BINDING SITES WITHIN THE PROTEIN ONLY RNASE P

Ganga Nilmini Kumari Wijesinghe Mudiyanselage, B.Sc

Marquette University, 2019

Transfer RNA (tRNA) maturation is an essential step that occurs after transcription and prior to the translation. In all domains of life, an RNA-based ribonuclease (RNase) P catalyzes the cleavage of the phosphodiester bond at the 5′ end of precursor tRNA. However, not all organelles utilize the RNA-based version of RNase P. A recent discovery altering the paradigm of RNase P has revealed that chloroplast organelles, parasitic trypanosomes, and the compact human mitochondrial (mt) genome rely solely on a protein-only 5′ tRNA processing enzyme. Structural and biochemical studies of the RNA-based and protein-based plant RNase P systems suggest that these distinct isoenzyme active centers are both comprised of conserved oxygen-rich, electronegative pockets that are stabilized by two catalytic metal ions. The similar mode of pre-tRNA substrate recognition and proposed similarities between the RNA-based and protein-only RNase P reaction mechanisms suggest an unprecedented example of convergent enzyme evolution. In the crystal structure of the **PR**oteinaceous **O**nly **R**ibonuclease **P** 1 enzyme (PRORP1) (PDB: 4G24), at least four aspartate residues ($D_{399}D_{474}D_{475}D_{493}$) are involved in coordinating two catalytic Mg^{2+} metal ions. However, there is a lack of information about PRORP metal binding and the organization of the catalytic metal-ligand environment. In this study, we characterize the metal binding properties of the PRORP enzyme family found in *Arabidopsis thaliana* using Electron Paramagnetic Resonance (EPR) and Isothermal Titration Calorimetry (ITC). Results from these studies will help us to better understand the metal-ligand environment of the metallonuclease domain of PRORP enzymes. ITC results suggest that the PRORP2 homolog binds to Mg^{2+} with a stoichiometric mole ratio of 1:1.88. Endothermic binding of the (ΔH = 218.3 cal/mol) Mg^{2+} suggests non-favorable metal binding and this data is consistent with a high μM metal binding dissociation constant (K_d). Similarly, PRORP2 binds approximately 1.53 equivalents of Mn^{2+}. Mn^{2+} binds 10-fold weaker than Mg^{2+} binding to PRORP2 and also undergoes non-favorable binding to PRORP2. Binding titrations of the PRORP3 homolog suggest that 1-2 metal ions are weakly associated with the active site. In addition to these metal binding experiments, analysis of EPR data suggest that PRORP2 bound Mn^{2+} exhibits monomeric binding, with Mn^{2+} oriented in a distorted octahedral geometry and a distinctive 90 G splitting pattern.

In addition, a newly discovered form of PRORP from archaea (*Aquifex aeolicus*) was successfully purified in the soluble form and characterized by Thermofluor analysis and NMR spectroscopy. Structural studies of *Aquifex aeolicus* PRORP has been initiated and preliminary crystals have also been obtained. *Aquifex aeolicus* PRORP is smaller in size and is thought to resemble a minimal metallonuclease domain of RNase P. Modelling studies of the *Aquifex aeolicus* PRORP suggest that protein domain will be a useful candidate for structural and metal binding studies and serves as a valuable system to understanding the structural diversity of the PRORP enzyme active site.

ACKNOWLEDGEMENTS

Ganga Nilmini Kumari Wijesinghe Mudiyanselage

During my research period, I have worked with a number of people who contributed in assorted ways to my research. First and foremost I owe my sincere appreciation and deepest gratitude to my supervisor Dr. Nicholas Reiter for providing me with academic knowledge and for his valuable guidance, assistance, and encouragement throughout my research.

I extend my gratitude to committee members Dr. Adam Fiedler, Dr. Chae S Yi, and Dr. Martin St. Maurice for being my committee members and their support given.

My heartfelt contribution also goes out to Dr. Brian Bennet, for helping me to get EPR data and for sharing his vast experience on EPR experiments, Dr. Martin St. Maurice for providing ITC instrument for my experiments and Dr. Francis Peterson for providing NMR facility in Medical College of Wisconsin.

I would like to thank my group members Dr. Danyun, Ainur, Dulmi, John, Kadhijah, Roshni and who helped me in various ways throughout the research work.

Further, I would like to convey my heartiest gratitude to my loving parents for helping me in every aspect and for their encouragement throughout my life. I am also thankful to my sisters and my loving husband Wasantha for assisting me in every way in all my activities. At last, I extremely thankful to my loving son Thusindu wholeheartedly for giving his milestones which release all my stress out.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1. General Overview

1.1.1. RNA and RNA processing

Cellular gene expression consists of three major events; replication, transcription, and protein translation, which is known as the central dogma (Figure 1) in biology.

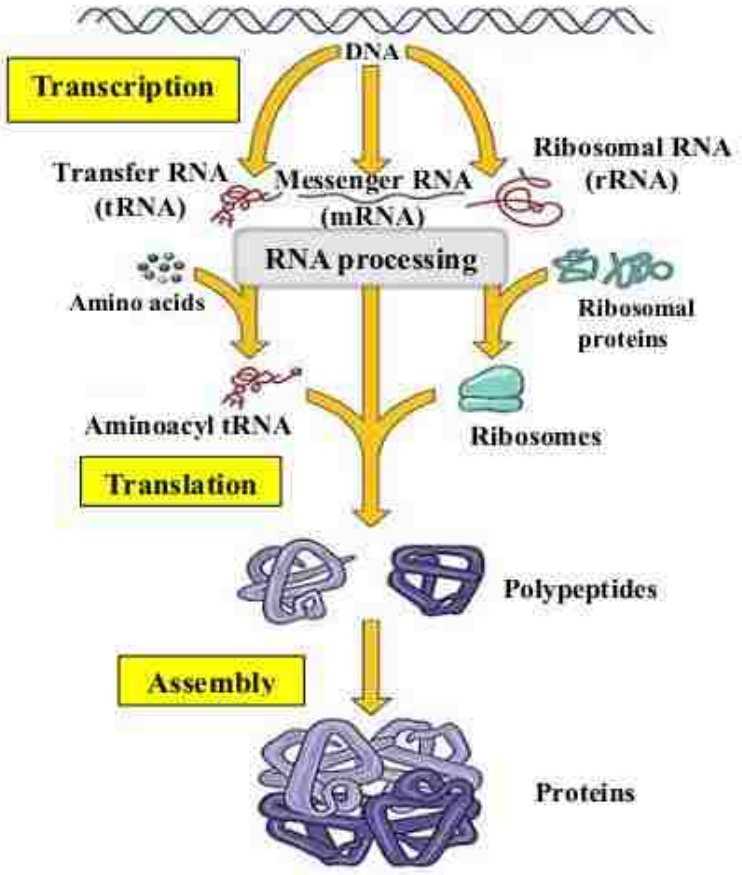

Figure 1. Central dogma; the flow of genetic information from DNA to RNA to protein involves transcription and translation.

DNA is transcribed into RNA, which mediates the flow of information stored in genes to a functional protein through the translation. There are three RNA types involved in translation, which include: messenger RNA (mRNA), ribosomal RNA (rRNA) and transfer RNA (tRNA). All three types of RNA molecules are essential mediators and perform specific functions in protein synthesis. In transcription, RNA is synthesized as a premature, and nonfunctional molecule. Therefore, RNA undergoes extensive RNA processing and post-transcriptional modifications (Figure 2) in order to become a fully functional form. These processing and modification steps serve as key factors in regulating the cellular gene expression [1].

Messenger RNA (mRNA) or coding RNA serves as the template for the gene translation and contains the protein-based genetic information. In eukaryotes, mRNA molecules transcribed from DNA is considered as precursor mRNA (pre-mRNA) and contains both coding and noncoding regions of the gene. In between transcription and translation, pre-mRNA undergoes RNA processing. RNA splicing is one such step which removes introns (noncoding) while linking coding regions (exons) [1]. Apart from RNA splicing, the mature mRNA is capped at the 5’ end with a modified guanine nucleotide and 3’ end with polyadenylation. These steps are important for gaining the ability to transport mRNA to the ribosome; to initiate translation, to increase the stability, to protect mRNA molecule against the cellular degradation, and for exporting mature mRNA from nucleus to the cytoplasm.

In eukaryotes, both tRNA and rRNA also undergo processing and post-transcriptional modifications prior to the translation. Ribosomal RNA (rRNA) in eukaryotes is transcribed into two long RNA molecules and undergoes RNA processing

to obtain mature rRNA molecule which is needed for synthesizing the functional ribosome. One such modification is the methylation of particular nucleotides that stabilize the rRNA. The ribosome is a complex macromolecular machine made up of mature rRNA and protein, and serve as the place where aminoacylated tRNA directly interact with the mRNA and peptides are synthesized.

Transfer RNAs (tRNAs) are essential adaptor molecules that serve to decode an mRNA and promote the formation of peptide bonds in ribosomal-based amino acid synthesis. tRNA is synthesized as immature precursors and is non-functional as a precursor molecule prior to translation. Therefore, precursor tRNA (pre-tRNA) must undergo numerous processing and post-translational modification steps in order to become a fully functional molecule. RNA processing in tRNA has been divided into multiple steps; removal of 5’ leader sequence, exonuclease trimming of the 3’ end, CAA addition at the 3’ end, nucleotide modification of the nitrogenous purine or pyrimidine nucleobases, and occasional tRNA splicing that is required to remove the intronic tRNA regions present in eukaryotes and archaea. The formation of functional and stable tRNA molecules is an essential step to accomplish specific functions during protein synthesis.

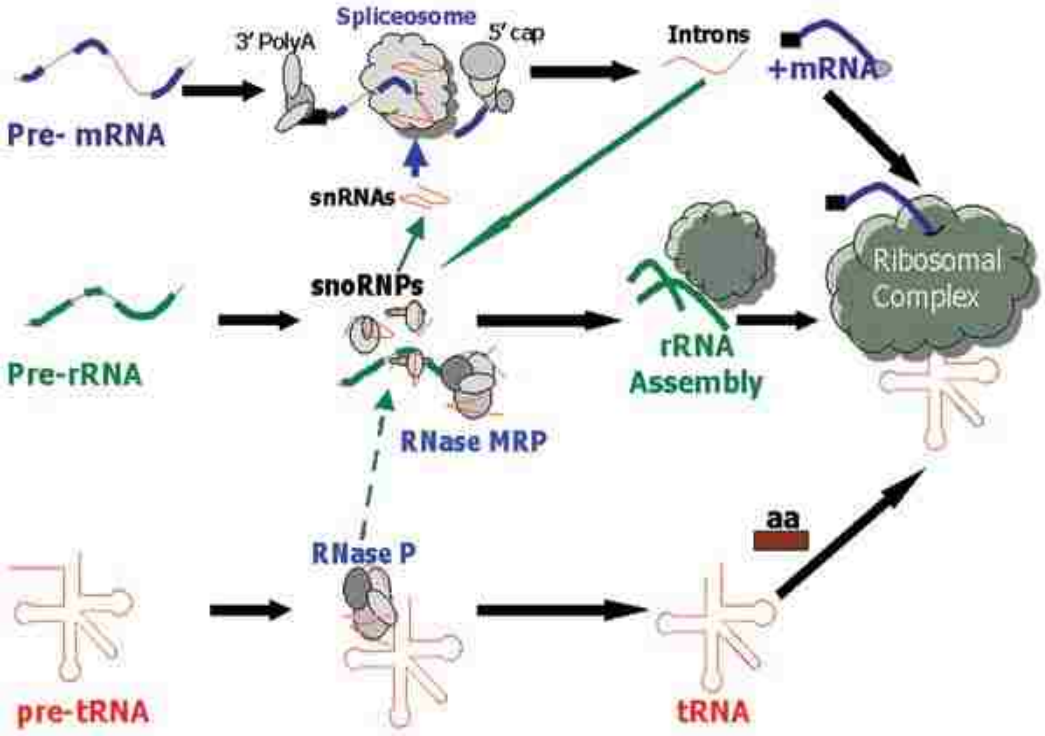

Figure 2. RNA processing in eukaryotes. The fate of messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA) after transcription and before translation [2].

1.1.2. RNase P

During tRNA maturation, specific enzymes remove the extra nucleotides at the 5' and 3' ends. In 1989 Sidney Altman received the Nobel Prize for the discovery of a unique catalytic RNA ("ribozyme") called Ribonuclease P (RNase P). RNase P is an important metalloenzyme involved in the tRNA processing, and catalyzes the cleavage of the 5'end sequence of pre-tRNA. From bacteria to archaea to eukaryotic organisms, RNase P is an RNA-based multi-turnover enzyme that interacts with the substrate in *trans*, relying on the recognition of RNA tertiary structural elements and RNA-metal ion enzyme active site [3, 4]. In eukaryotes, RNase P is the only enzyme that acts upon the 5'

end of pre-tRNA, whereas extra nucleotides at the 3’ end are primarily removed by the RNase Z exonuclease [5]. Hence, RNase P acts on pre-tRNA and tRNA-like structure early in maturation and is considered an ancient ribozyme that is found in all domains of life.

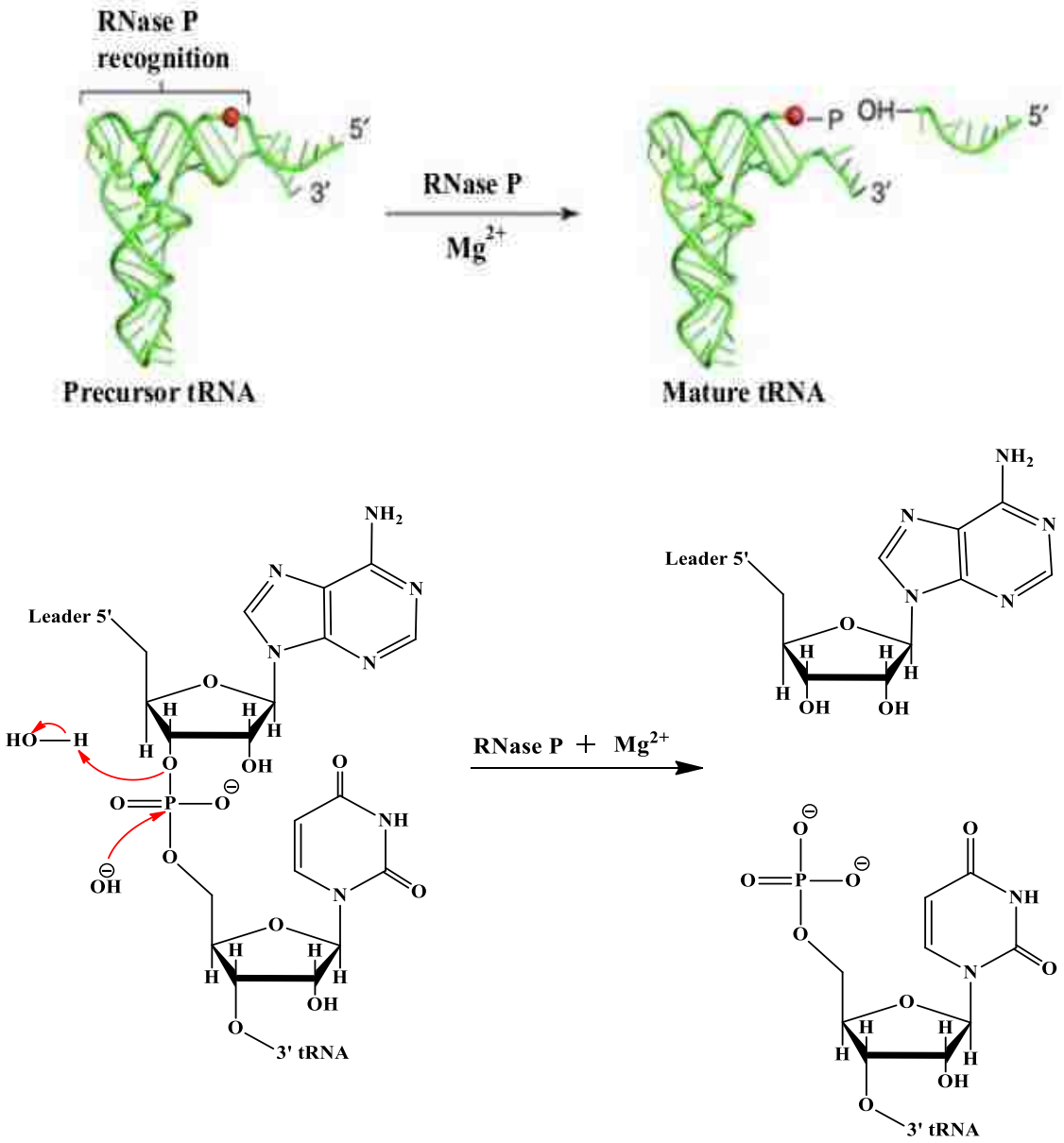

Figure 3. The reaction of RNase P, (A) Schematic diagram of the reaction of RNase P with precursor tRNA, (B) Chemical reaction showing the hydrolysis of the backbone phosphate group at the 5’ end of precursor tRNA [6].

1.1.3. Two types of RNase P

RNase P was first identified as a ribonucleoprotein (RNP) complex in bacteria. This ribonucleoprotein is comprised of a conserved single RNA subunit that shows ribozyme-like catalytic activity, and a variable number of protein subunits. As this ribonucleoprotein is found in all three domains of life (bacteria, archaea, and eukarya), it gives credence to believe that the RNA-based RNase P is a universal enzyme [7, 8]. A number of protein subunits associate with the catalytic RNA subunit in higher organisms with the primary goal of improving stabilization of the holoenzyme, enhancing the gene regulatory of RNase P, and accommodating diverse pre-tRNA and tRNA-like substrates. Although RNase P is considered a *bona-fide* RNA-metal catalyst, the protein component is essential for *in vivo* enzymatic activity.

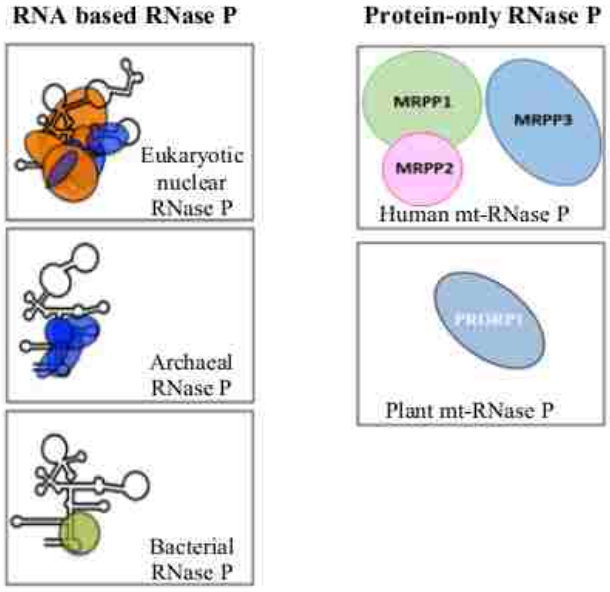

Figure 4. Types of RNase P; Two types of RNase P found all domains of life. Left: RNA based RNase P contains conserved catalytic RNA and variable protein subunits that associate with the RNA. Right: Protein only RNase P (PRORP) found in eukaryotes. A plant PRORP with single protein subunit and human mitochondrial RNase P with three protein complex termed MRPP1, 2, and 3 [9].

However, not all organelles utilize the RNA-based version of RNase P. A recent discovery has revealed that chloroplast organelles, parasitic trypanosomes, and the compact human mitochondrial (mt) genome rely solely on a protein-only 5′ tRNA processing enzyme (PRORP) [10-12]. This discovery demonstrated that RNase P activity does not require catalytic RNA, however, it does appear that both the RNA-based and the protein-only forms of RNase P (Figure 4) possess similar metalloendonuclease enzyme active sites [13, 14].

1.1.3.1. RNA based RNase P

Ribonucleases (RNases) catalyze the degradation of RNA. However, the RNase P is a unique enzyme compared to the other RNases and acts as a ribozyme to cleave pre-tRNA at 5’ end in the presence of Mg^{2+} ions [15]. RNA-based RNase P in bacteria is composed of one large catalytic RNA subunit (P RNA) and a single protein subunit (Figure 5A) [15]. With increasing organismal complexity, additional proteins have been added to the RNA-based RNase P system. Archaeal RNase P, consists of four or five different protein subunits [16], whereas RNase P in eukarya can contain up to ten or more protein subunits (eg; yeast and human RNase P) (Figure 5C) [17]. Thus, the number of subunits that associate with the RNase P enzyme increases with the complexity of life. However, RNA component in RNA based RNase P is conserved through bacteria, archaea and eukaryotic nuclear RNase P and, are structurally related. This suggests that the ribozyme found in bacterial RNase P could serve as a reasonable model system to understand the mechanism of pre-tRNA cleavage.

1.1.3.2. Structure of the RNA based RNase P

Structure of the RNA based RNase P is mainly composed of two components, a catalytic RNA component and protein subunit/s. Bacterial RNase P consist of one large catalytic RNA subunit (P RNA) with a size of 340-420 nucleotides, and a protein subunit with the size of 14 kDa (Figure 5A) [15]. The protein facilitates substrate binding and stabilizes the catalytic core of the RNase P RNA. In addition, the protein subunits associated with RNA can vary dramatically between bacteria, archaea, and eukarya. Although there is a single protein subunit in bacterial RNase P, archaeal RNase P integrates the RNA component with at least 3-4 protein subunits. In human nuclear RNase P, there are minimum ten protein subunits, while yeast contains nine protein subunits [18-21]. Interestingly, despite these differences in composition, X-ray and recent Cryo-EM structures of the RNase P holoenzyme strongly suggest that the mode of substrate recognition and the catalytic active sites appear to be structurally conserved between bacteria, archaea, and eukaryotes [4, 21-23].

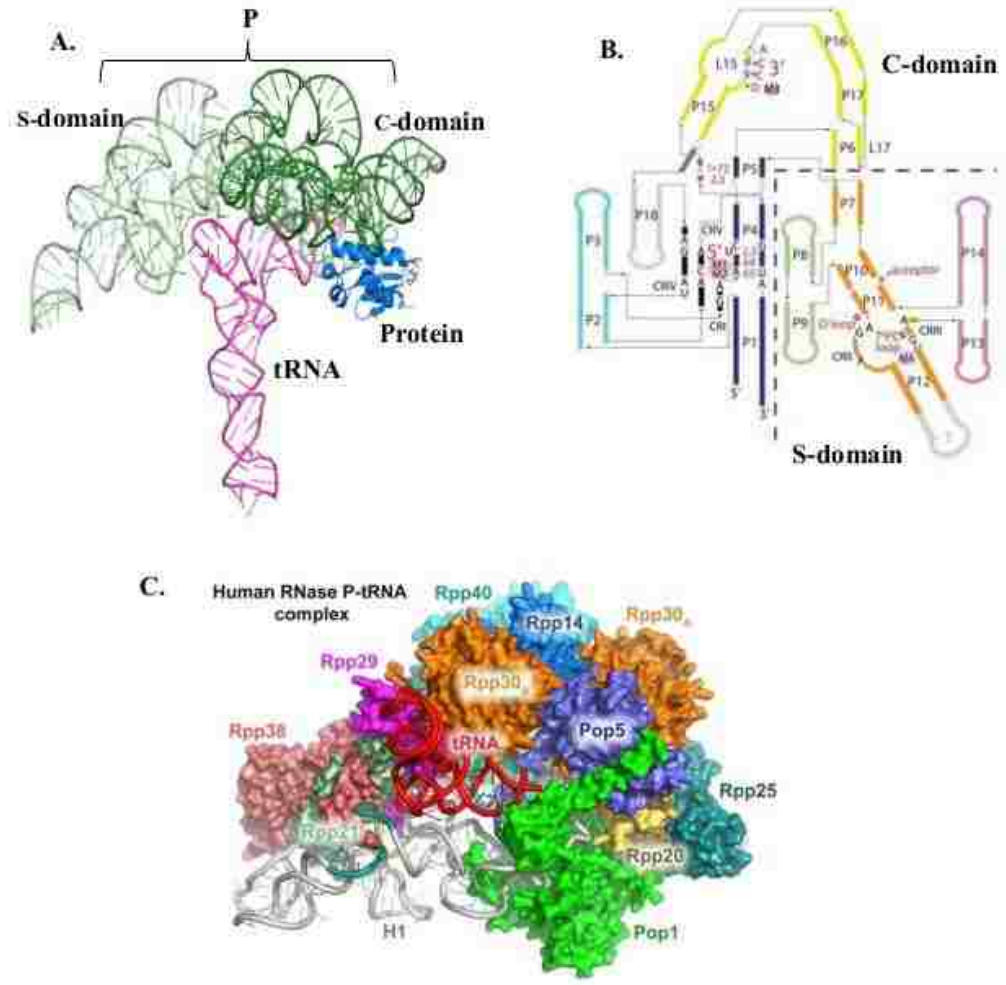

Figure 5. (A) Crystal structure of Bacterial RNase P with tRNA; PDB: 3Q1Q generated using PyMOL software. Bacterial RNase P has a catalytic RNA (P RNA) component with a single protein subunit (blue). RNA component is divided as C(catalytic) domain (dark green) and S(specificity) domain (lime green) [4]. (B) Secondary structure of RNA component. (C) CryoEM structure of Human nuclear RNase P with tRNA(PDB:6AHU); Human nuclear RNase P is RNA based RNase P and has identified with ten different protein subunits [21].

The RNA component of RNA based RNase P has been classified to subgroups based on the secondary structure of RNA along with the phylogeny [24]. So, depending on the secondary structure or RNA component researchers have grouped RNase P into different classes. There are three types of RNase P in bacteria, class A (Ancestral), B (*Bacillus*), and C (Chloroflexi), classified based upon phylogeny and secondary structural

similarities. There are two types in Archaeal RNase P (named type M (Methanococci) and P (Pyrobaculum) and highly variable and no distinct classifications defined in eukaryotic RNase P RNA.

Nonetheless, all RNase P RNA molecules in all organisms contain five universally conserved regions (CR). In bacterial RNase P RNA, three regions reside in the catalytic (C-domain) and include CR-I, CR-IV, and CR-V and two universally conserved regions exist in the specificity domain (S-domain) (CR-II, CR-III) (Figure 5B). The C domain represents 60% of P RNA in bacteria and catalyze cleavage of precursor tRNA in the presence or absence of RnpA protein. The S-domain is important in substrate shape recognition, and consists of remaining 40% of P RNA in bacteria. Thus, the P RNA has five domains which conserved through all RNA based RNase P. Based upon the X-ray and recent Cryo-EM RNase P holoenzymes, these conserved regions are structurally similar and help to properly recognize the pre-tRNA substrate and form the catalytic active site scaffold [4, 21-23].

1.1.3.3. Protein Only RNase P (Proteinaceous RNase P)

Due to the conservation of RNA based RNase P throughout the domains of life, there appears to be no need to rely on a catalytic protein for 5' pre-tRNA processing. However, unforeseen 5'-end pre-tRNA processing protein was originally identified in spinach chloroplast and has been subsequently identified in several organelles in nature [25]. This Protein Only RNase P (Proteinaceous RNase P) – **PRORP,** represents a different type of RNase P enzyme family that exists in Eukarya (animals (Metazoa), plant and algae (Archaeplastida), trypanosomes (Excavata), and heterokonts (SAR subgroup Stramenopiles) [26].

In *Arabidopsis thaliana*, three types of PRORP enzymes have been identified; PRORP1, PRORP2, and PRORP3. The PRORP1 is found in both mitochondria and chloroplast while PRORP2 and PRORP3 are found in the nucleus and found to be involved in tRNA maturation [8]. Crystal structures of PRORP proteins have shown three distinct domains; N-terminal pentatricopeptide repeats (PPR) RNA binding domain, central Zn^{2+} binding domain, and metalloendonuclease domain with two distinct metal ions [27].

In contrast to plant-protein only RNase P, human mitochondrial RNase P enzyme exhibits a complex structure and includes three distinct mitochondrial RNA processing proteins (MRPP); MRPP1, MRPP2, and MRPP3. These three proteins are involved in methyltransferase, dehydrogenase, and 5' metalloendonuclease activity respectively [9, 28, 29]. The MRPP3 subunit is homologues with PRORP enzymes and undergoes complex formation with MRPP1 and MRPP2 to become fully functional [30].

1.1.3.4. Structure of PRORP1

According to the crystal structure of PRORP1 *A. thaliana* (PDB: 4G24), the PRORP1 resembles a V shape and contains three major domains; pentatricopeptide domain (PPR), central domain, and a metallonuclease domain (Figure 6). However, the elements in the multi-domain PRORP1 protein do not show similarity with the protein component found in RNA based RNase P [27].

Pentatricopeptide Domain (PPR). The PPR domain (residues 95-292) includes repeats of the tandem helix-turn-helix fold of ~35 amino acids. This domain is proposed to be important for substrate binding affinity and activity. PPR domain in PRORP1 consists of 11 α helices and forms repeating helix- turn-helix hairpins [27].

Central Domain. The central domain (residue 328-357 and 534-570) is sandwiched between PPR domain and metallonuclease domain, and is composed of four antiparallel β sheets which connect both PPR and metallonuclease domains through two extended loops. There is a Zn^{2+} metal ion, coordinated with four residues (C344, C350, H548, and C565) that are supposed to be important in stabilizing the protein structure and orienting the three domains. Additionally, two positively charged patches in the central domain have been observed and could interact with the tRNA backbone [27]. Role of the Zn^{2+} site is similar to the common zinc finger motif found in proteins.

Metallonuclease Domain. The metallonuclease domain (residue 358-533) serves as the catalytic domain in PRORP enzymes and resembles the Nedd4-BP1, YacP nuclease (NYN) domain, that is a member of PilT N-Terminal (PIN) domain-like fold superfamily [31]. In PRORP1, the metallocenter is coordinated with four aspartate residues (D399, D474, D475, and D493) and mutagenesis of four aspartates to alanine exhibits a significant decrease in the cleavage activity, confirming the importance of these aspartate residues in enzyme function [27]. It has been proposed that the PRORP1 binds to two distinct metal ions and this result is largely based upon metal crystal soaking experiments with either Ca(II), Sr(II), or Mn(II). These results suggest that binding of Ca(II) or Sr(II) occurs with a mole ratio of approximately 1: 1, whereas binding of Mn(II) results in a mole ratio of 1: 2. (PDB:4G24) [27]. In the crystal structure of PRORP1 soaked with Mn^{2+} PDB: 4G24, it appears that a second Mn^{2+} ion exists with lower occupancy. This suggests that there may be a weak affinity for the second metal ion in the enzyme active site [27].

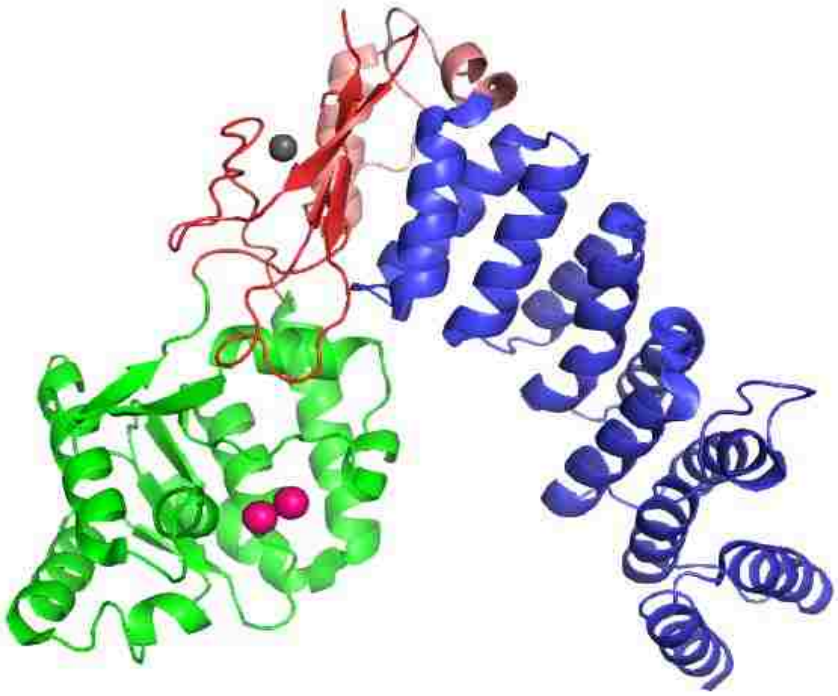

Figure 6. Crystal structure of PRORP1 *Arabidopsis thaliana* (PDB: 4G24) generated from PyMOL software. Contains three different domains metallonuclease domain (green) with two Mn^{2+} (magenta spheres), central domain (red) with Zn^{2+} (black sphere), and pentatricopeptide domain (PPR) (blue).

UV crosslinking and footprinting analysis of PRORP1 with tRNA has shown that the U16, G18, and G19 of mitochondrial precursor $tRNA^{Cys}$ D-loop, and C56 in TΨC loop (Figure 7B) were protected by recombinant PRORP1 from nuclease digestion [32]. These residues are conserved in tRNA binding with both the RNA-based and protein only RNase P enzymes.

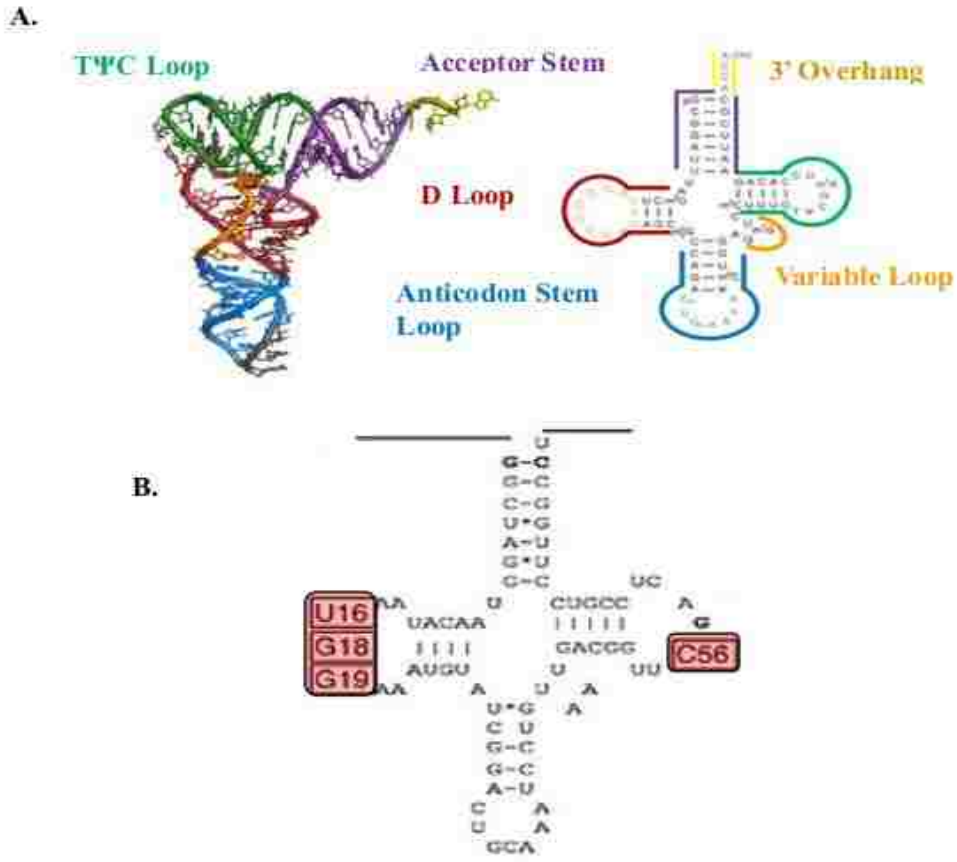

Figure 7. (A) Left; Tertiary structure of tRNA and Right; the secondary structure of tRNA (B) Conserved nucleotides in $tRNA^{cys}$ in red which interact with PRORP1 [32].

1.1.4. Comparison of RNA based RNase P and Protein Only RNase P (PRORP)

Biochemical and structural work has shown that both the RNA-based and protein-only isoenzymes adopt convergent evolution strategies to perform enzymatic catalysis. Though the function of both types of RNase P is similar, they are significantly different in structure. RNA based RNase P consist of ribonucleoprotein complex, while the PRORP consist of one (plant PRORP) or multiple (human PRORP) protein subunits.

Structural and biochemical evidence show that the metallocenter of PRORP is coordinated by carboxylated side chains of aspartate residues while the oxygens of non-bridging phosphate groups or nucleotide carbonyl groups coordinated the metallocenter

of the RNA based RNase P. It is interesting why nature has altered the form of RNase P into protein in some eukaryotic organisms and endosymbiotic organelles.

1.1.4.1. Ribonuclease mechanism of RNA based RNase P and protein only RNase P

The active site of the RNA based RNase P has been suggested to consist of two distinct metal ions; M1 and M2, required for catalysis. According to the proposed mechanism, water derived hydroxyl group that is coordinated by the M1 ion performs a nucleophilic attack on the phosphorus of the RNA backbone at precisely the 5' end. The M1 is responsible for positioning the phosphorous group and activates the nucleophilic water for the S_N2 type reaction. M2 appears to hold a water molecule which mediates the proton transfer on 3' oxygen atom in product releasing step (Figure 8).

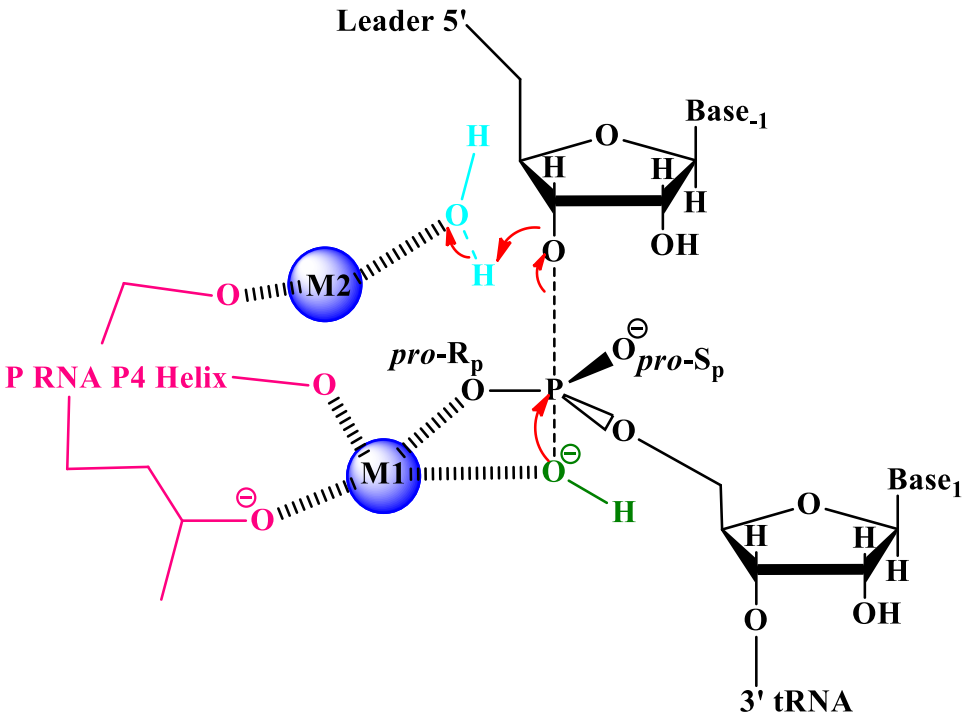

Figure 8. Mechanism of RNA based RNase P [4].

Compared to the RNA based RNase P, PRORP has also been suggested to have two distinct metal ions; M1 and M2 in the active site. A water molecule held by the M1 ion is suggested to undergo deprotonation by an aspartate group which acts as a general base to generate the nucleophile to undergo nucleophilic attack. Both M1 and M2 are involved in positioning the phosphate group, facilitate the nucleophilic attack, and stabilize the transition intermediate. Thus, a general acid in the active site of protein mediates the proton transfer on 3' oxygen in product releasing step (Figure 9).

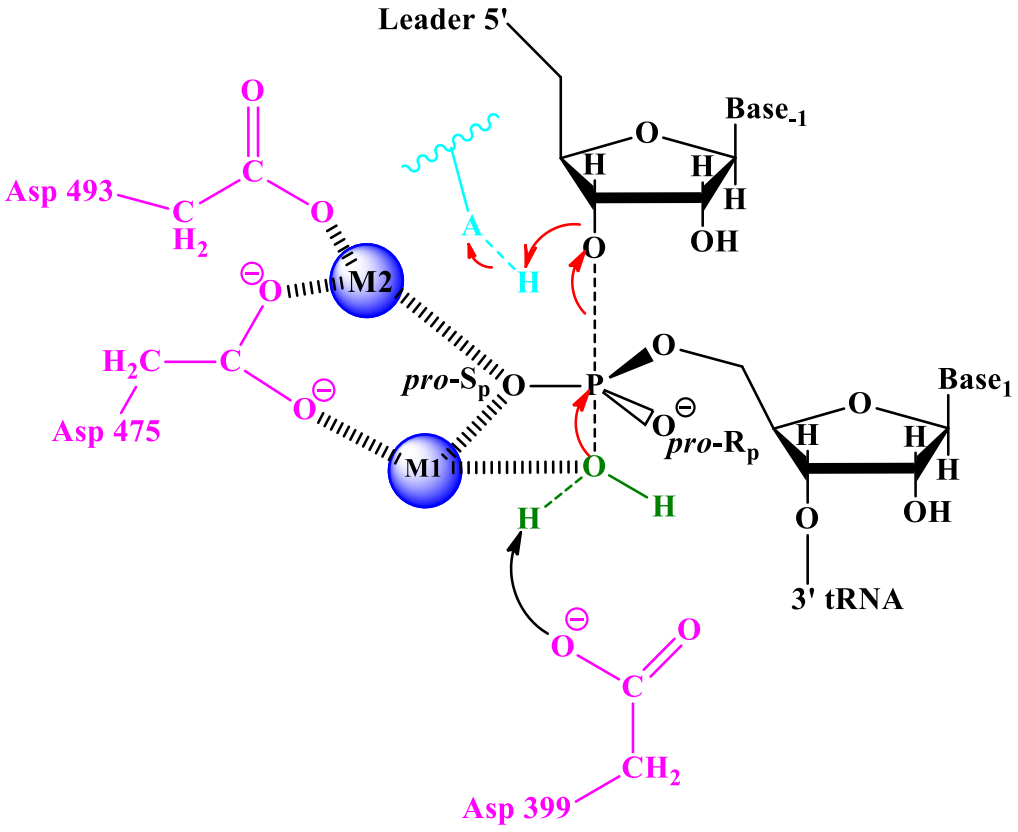

Figure 9. Mechanism of Protein Only RNase P [27].

1.1.5. Scope of the study

The discovery and biochemistry of the Protein only RNase P enzymes and their enzyme active site similarities to the RNA-based RNase P complex, suggest that these distinct isoenzymes adopt convergent evolution strategies to perform enzymatic catalysis. However, the structural diversity of the active site and the mode of substrate recognition by the PRORP enzyme family are not well understood. Understanding the properties of the protein-based RNase P complex as well as the molecular recognition of the different types of PRORP protein subunits will provide insight into the process of tRNA maturation.

The main goal of this study is to understand the metal ion environment of the protein-only RNase P from *A. thaliana*. Under this study, the metal binding properties of protein only RNase P enzymes (PRORP2 and PRORP3) were studied by ITC and EPR spectroscopy. In addition, structural analysis of RNase P form *Aquifex aeolicus* was also performed by Thermofluor screening, NMR spectroscopy, and X-ray crystallographic methods. This *Aquifex aeolicus* RNase P is thought to represent a minimal PRORP metalloendonuclease scaffold yet no structural information exists regarding its tertiary fold, catalytic active site environment, or its mode of substrate recognition.

CHAPTER2

COMPARISON OF PRORP2 AND PRORP3 METALLOENDONUCLEASE ENZYME ACTIVE SITES

2.1. Introduction

2.1.1. Structure of PRORP2

There are three proteins only RNase P homologs present in *A. thaliana*; PRORP1, PRORP2, and PRORP3 which are encoded by the nuclear DNA. PRORP1 is functioning in mitochondria and chloroplast, while PRORP2 and PRORP3 co-localize and functioning in the nucleus. Both PRORP1 and PRORP2 show similar structures that contain three distinct domains; metalloendonuclease domain, Zn binding central domain, and substrate binding pentatricopeptide (PPR) domain (Figure 10). The overall structure of both proteins are V-shaped and require the cofactor Mg^{2+} to properly cleavage of precursor tRNA substrates. Interestingly, PRORP2 (PDB-5DIZ/5FT9) crystallizes as a dimer whereas the PRORP1 crystallizes as a monomer.

Compared to the PPR domain in PRORP1, PRORP2 has a short helix-turn-helix structure that consists of five and a half PPR motif and 11 consecutive alpha helices which are similar to PRORP1. Differences in PRORP2 are mostly found in the loop regions. However, the central domains and the metalloendonuclease of PRORP1 and PRORP2 show high similarity (The central domain consists of conserved C281, C284, H494, and C511 residues). Like in PRORP1, Zn^{2+} in PRORP2 does not involve in catalysis. The active site of PRORP2 consist of four aspartate residues; D343, D421, D422, and D440 similar to the conserved aspartate residues in PRORP1; D399, D474,

D475, and D493 [33,27]. Mutations of active site aspartates to alanines in both PRORP1 and PRORP2 have caused significant decrements of the activities. Recent studies have shown that both PRORP1 and PRORP2 are functioning in the presence of Mg^{2+} and Mn^{2+}, and inactive with Ca^{2+} and Zn^{2+}, suggesting the endonucleolytic cleavage to be metal ion specific. In addition, a biochemical comparison of PRORP1 and PRORP2 suggest that PRORP2 preferentially interacts with substrates with shorter 3' trailer and 5' leader sequences [33]. An additional difference between the PRORP family is that PRORP2 functions only in the nucleus and does not cleave non-canonical mitochondrial substrates showing its site-specific functioning [33].

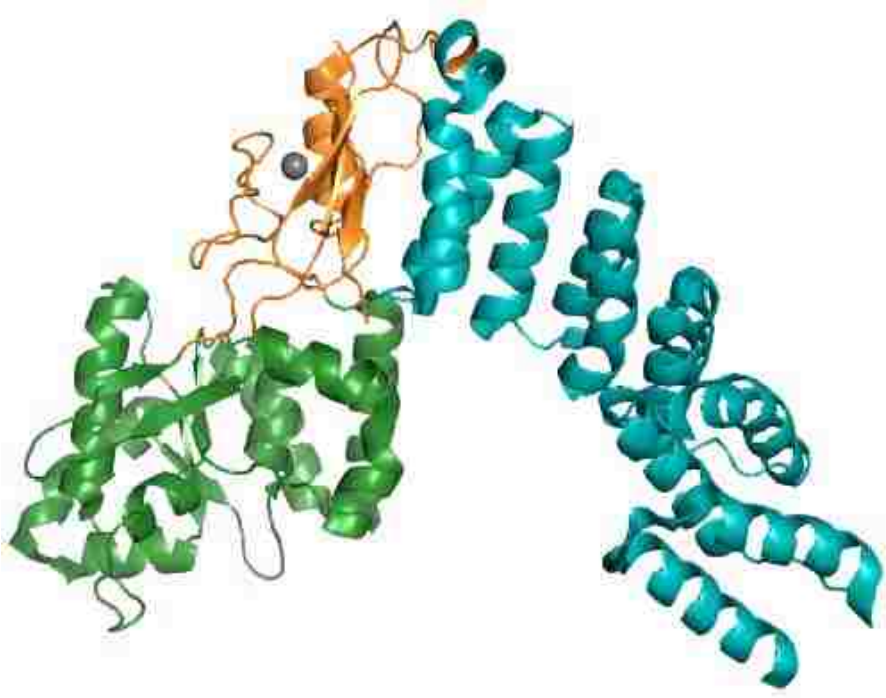

Figure 10. Crystal structure of PRORP2 *Arabidopsis thaliana* (PDB: 5FT9) generated from PyMOL software. Contains three different domains metallonuclease domain (green), central domain (orange) with Zn^{2+} (black sphere), and pentatricopeptide domain (cyan).

2.1.2. Homology of PRORP1, PRORP2, and PRORP3

Figure 11. Sequence alignment of *A. thaliana* PRORP1, PRORP2, and PRORP3 using clustalW multiple alignments- BioEdit software. Identical amino acids highlighted in magenta and similar nucleotides highlighted in cyan.

According to the sequence alignment by Cluster omega software, PRORP1 shows 47.61% similarity to PRORP2 and 45.83 % similarity to PRORP3. However, PRORP2 shows 79.88 % similarity to PRORP3. Four or five aspartate groups within the active site are conserved in all three types of PRORP found in *A. thaliana* (Figure 11).

2.2. Material and Methods

2.2.1. Materials

Agar, kanamycin, ampicillin, imidazole, glycerol, Tween 20, Isopropyl-β-D-1-thiogalactopyranoside (IPTG), Monosodium phosphate (NaH_2PO_4), tris(2-carboxyethyl)phosphine (TCEP), Beta-mercaptoethanol (BME), Ethylenediaminetetraacetic acid (EDTA), PMSF (phenylmethylsulfonyl fluoride) were purchased from Sigma-Aldrich, VWR or Fisher Scientific. All other reagents were purchased commercially and were the highest purity available.

Pre-packed immobilized-metal affinity chromatography nickel-nitrilotriacetic acid (IMAC Ni-NTA) columns (5 mL) and unpacked Ni-NTA resin were purchased from Bio-Rad. Plasmid preparation kits were purchased from Promega. All the solutions including buffer solutions were prepared using Nanopure water (~ 18.2 MΩ). Cells and proteins were handled at 4 ºC in or a temperature controlled incubator.

2.2.2. Instruments/Apparatus

The New Brunswick Scientific Innova 4330 Refrigerated Incubator Shaker was used for growing and inducing of cell cultures. Cell lysis was performed using a Branson Sonifier. Pre-packed and manually packed IMAC Ni-NTA columns were connected to the BIORAD NCG chromatography system and used for protein purification. EPR

spectra were obtained on an updated Bruker EMX-AA-TDU/L spectrometer equipped with an ER4112-SHQ resonator (9.5 GHz) and an HP 5350B microwave counter for precise frequency measurement. The temperature was maintained with a ColdEdge/Bruker Stinger S5-L recirculating helium refrigerator, and an Oxford ESR900 cryostat and MercuryITC temperature controller. EPR simulations were carried out using EasySpin. All metal binding titrations were performed on a MicroCal iTC200 ultrasensitive titration calorimeter. The metal content of all protein samples were determined using inductively-coupled plasma mass spectrometry (ICP-MS).

2.2.3. Preparation of LB-agar plates and LB growing media

For the preparation of LB agar plates, a 5 g portion of LB agar powder was dissolved in 250 mL of nanopure water. Autoclaved growing media were cooled and kanamycin was added up to a final concentration of 50 μg/mL before start the solidification. The mixture was transferred into sterile Petri dishes under the sterile conditions. The solidified dishes were stored at 4 ºC until needed.

For the preparation of LB growing media, a 25 g portion of LB agar powder was dissolved in 1 L of nanopure water followed by the autoclaving and left to cool under air.

2.2.4. Preparation of PRORP2, PRORP3, and PRORP3 D422N expression systems

The pET-28a$^+$ plasmids containing the PRORP2, PRORP3, and PRORP3 D422N genes from *A. thaliana* were separately transformed into BL21 (gold) cells using the heat-shocked at 42 ºC for 45 seconds for the preparation of individual expression systems for the PRORP2, PRORP3, and PRORP3 D422N proteins. Transformed cells were separately mixed with 500 μL of autoclaved LB media which was at 37 ºC and incubated

for 1 hour at 37 ºC while shaking at 225 r.p.m. LB-agar plates containing kanamycin were inoculated with the prepared expression systems and incubated overnight at 37 ºC.

2.2.5. Growth, expression, and harvesting of PRORP2, PRORP3, and PRORP3 D422N proteins

Starter cultures were prepared by adding a single colony from each freshly prepared expression system into 50 mL of autoclaved LB media with kanamycin (50 µg/mL). The cultures were incubated for 16-18 hours at 37 ºC while shaking at 220 r.p.m. These cultures were used to inoculate separate 6 L of LB media containing kanamycin (50 µg/mL). The cultures were grown at 37 ºC and constant shaking at 225 r.p.m. until the optical density of ~0.8 at 600 nm was reached. The cell cultures were chilled using a temperature controlled incubator and induced with 500 µM of IPTG. Induced cultures were incubated for additional 16-18 hours at 18 ºC with constant shaking at 225 r.p.m. The cells were harvested by the centrifugation at 7000 r.p.m. for 10 minutes under 4 ºC and cell pellets were stored at -80 ºC until needed.

2.2.6. Purification of the PRORP2, PRORP3, and PRORP3 D422N proteins

The purification of PRORP2, PRORP3, and PRORP3 D422N proteins were separately performed according to the optimized protocol. Cell pellets were re-suspended in the lysis buffer (50 mM HEPES, 250 mM NaCl, 5% glycerol, 10 mM BME, 0.2 % Tween 20, 1 mg/mL Lysozyme, 0.1 mg/mL PMSF, and protease inhibitor) and lysed by ultra-sonication at output power of 6 for 12 minutes on ice. The mixture was centrifuged at 16,000 rpm for 45 minutes under 4 ºC and separated supernatant was loaded on to the Ni-NTA column (5 mL) that was previously equilibrated with 100 mL of 50 mM HEPES buffer, pH 7.5 (250 mM NaCl, 5% glycerol, and 10 mM BME). The column was washed

using ~20 column volumes (CVs) of the wash buffer (50 mM HEPES buffer at pH 7.5, 250 mM NaCl, 5% glycerol, and 10 mM BME, 12.5 mM Imidazole) to remove nonspecifically bound proteins. The protein was eluted using ~20 CVs of the elution buffer (50 mM HEPES buffer at pH 7.5, 250 mM NaCl, 5% glycerol, and 10 mM BME, 100 mM Imidazole). All the fractions were analyzed by SDS-PAGE (10-12%) and the fractions containing targeted protein were pooled, and concentrated using 50 K Amicon ultra centrifugal units. Protein was exchanged into 50 mM HEPES buffer, pH 7.5 (250 mM NaCl, 5% glycerol and 1 mM TCEP). Protein concentrations were measured by Nano Drop (ThermoScientific) as well as a Bradford Assay.

2.2.7. The metal content of the PRORP2 and PRORP3 using ICP-MS

Each protein (PRORP2 and PRORP3) sample was passed through a desalting column and 1 mL fractions were collected. In each preparation, all the protein containing fractions were pooled together and concentrated using Amicon concentrator (50 K). The protein sample was digested using a mixture of 2 % HNO_3 and 0.5 % HCl, in metal-free polypropylene centrifuge tubes (15 mL). After 24 hours, the digested protein samples were filtered using 0.2 μm syringe filters. The metal content of the protein samples were determined by inductively-coupled plasma mass spectrometry (ICP-MS).

2.2.8. Isothermal titration calorimetric (ITC) analysis of the PRORP2, PRORP3, and PRORP3 D422N protein

Protein samples that passed through the desalting column were used for binding titrations. All the protein samples and metal samples were prepared in 50 mM HEPES buffer, pH 7.5 containing 1 mM TCEP. The protein samples (0.45 mM PRORP2, 0.25 mM PRORP3, and 0.25 mM PRORP3 D422N) were titrated with 20 mM Mg^{2+} and 20

mM Mn^{2+} metal ion solutions. The cell temperature was kept at 22 ºC throughout the titration process with a stirring speed of 750 r.p.m. and an injection volume of 2 μL. The data were analyzed by the Origin software provided by MicroCal and used a one binding site fitting model to calculate the stoichiometry and Mg^{2+} and Mn^{2+} dissociation constants. Blank titrations were done with identical conditions except that the buffer solution replaced the protein sample volume. A blank titration was subtracted from experimental titration to remove the heat generated from dilution. The relationship of K_d $=1/K_a$ was used to calculate dissociation constant.

2.2.9. EPR spectroscopy of PRORP2, PRORP3, and PRORP3 D422N proteins

Separate 1 mM protein samples and mixtures of 1 mM proteins/2 mM Mn^{2+} were transferred into EPR tubes. Each sample was gradually frozen in liquid nitrogen for EPR analysis. EPR Spectra were recorded under two sets of conditions, identified in the figure legends, each with 12 G magnetic field modulation at 100 kHz. Conventional EPR spectra were recorded at 35 K and 3.3 mW. In other cases, spectra were recorded at 5 K using a dual-mode technique, in which two channels simultaneously recorded the spectra with phase-sensitive detection (PSD) at 100 and 200 kHz, respectively, and with the PSD set out-of-phase with the field modulation. The microwave power was balanced (6.6 mW) to obtain the conventional EPR spectrum in the out-of-phase 100 kHz channel, devoid of the rapid-passage effects that are seen with the PSD and modulation in phase at 5 K, and those spectra are presented (Figure 17). The out-of-phase 200 kHz channel contained the derivative components of fast-relaxing species and the absorption components of slow-relaxing species and is therefore useful in deconvoluting the spectrum on the basis of relaxation; these latter spectra will not be considered herein.

Each of the spectra presented here, therefore, represents the usual first derivative of absorption display.

2.3. Results and Discussion

The tRNA maturation is an essential step in post-transcription and occurs prior to translation. In nature, there are two forms of RNase P enzymes; RNA based RNase P and protein-based RNase P. However, there is lack of information on the metal-protein and substrate-protein interactions of the protein only RNase P enzyme family. Therefore, studying the biochemistry of the protein-based RNase P will be important to better understand the mechanism of 5' precursor tRNA maturation. In this study, our main goal is to determine the metal ion dependence of protein-based RNase P forms found in *A. thaliana*.

Purification of A. thaliana recombinant PRORP2 protein using affinity column chromatography. For protein overexpression of PRORP2 from *A. thaliana*, the pET28a^{+} Kan-resistant plasmid was transformed into BL21 GOLD competent cells using the heat shock method. Transformed cells were grown in Luria-broth media in the presence of the 50 μg/mL kanamycin antibiotic and protein overexpression was induced using 1 mM IPTG, as previously defined [33]. In the protein purification, all the recombinant PRORP2 proteins were His_6 tagged and purified using immobilized metal affinity chromatography (IMAC) and Ni-NTA (Nickel-nitrilotriacetic acid) resin. The purity of proteins was assessed by SDS-PAGE and results are shown in Figure 12. The band at 63 kDa represents the purified PRORP2 protein.

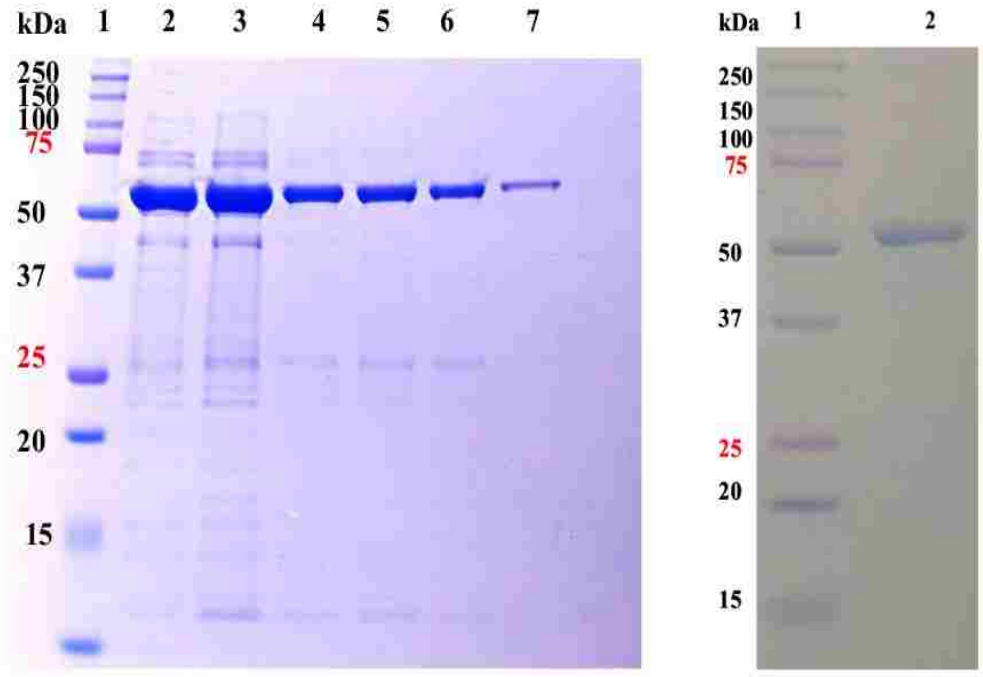

Figure 12. (A) Results of SDS-PAGE (12 %) of eluted fractions of *A. thaliana* PRORP2 protein after the Ni-column purification (Lane 1-Protein molecular weight marker, Lane 2 to 7-fraction 1 to 6) (B) Purified *A. thaliana* PRORP2 protein (Lane 1-Protein molecular weight marker, Lane 2- purified PRORP2 protein).

Determination of metal ion content in PRORP2 protein using Inductive coupled plasma mass spectrometry. Inductive coupled plasma mass spectrometry (ICP MS) analysis was performed to investigate the metal ions in the PRORP2 protein. The metal contents of the recombinant PRORP2 was shown in Table 1.

Table1: Inductively coupled plasma mass spectrometry (ICP-MS) data of PRORP2 protein.

Sample	Zn Mole ratio	Mn Mole ratio
Buffer	1.0 : 0.0	1.0 : 0.0
PRORP2 (10 μM)	1.0 : 1.1	1.0 : 0.0

The results show that the PRORP2 binds to the Zn with a mole ratio of 1:1. Interestingly this ratio was the same as the previously known data which has been observed by the crystal structure of *A. thaliana* PRORP2 (PDB: 5FT9) further confirming the 1: 1 Zn^{2+} binding in central domain. However, no information about the bound Mn^{2+} was observed by ICP-MS suggesting that the recombinant PRORP2 protein does not tightly associate with Mn^{2+} under standard purification conditions. This result enabled us to pursue alternative Mg^{2+} and Mn^{2+} ion reconstitution protocols to better understand metal-PRORP binding interactions.

Investigation of two metal ion binding in the active site of PRORP2 and PRORP3 protein using Isothermal titration calorimetry (ITC) titration. The crystal structure of the PRORP1 (PDB: 4G24) shows that two bound Mn^{2+} exist in the metallonuclease domain, suggesting that the active site of PRORP enzymes contains two metal ions (Figure 15). In previous kinetic studies, PRORP2 has shown catalytic activity in the presence of Mg^{2+} and Mn^{2+}. So, we selected Mg^{2+} and Mn^{2+} as the interested divalent metals that should be used in the metal binding studies. Therefore, we performed ITC experiments with PRORP2 and PRORP3 enzymes to determine the number of metal binding sites.

ITC data reveal two metal binding sites in PRORP2 with Mg^{2+} and Mn^{2+}.

Results from ITC experiments for PRORP2 reveal that approximately 1.88 ± 0.22 metal binding sites (Figure 13) for Mg^{2+}. This result supports the proposed mechanism with approximately two metal ions in the active site. ITC experiments are used to measure the heat absorbed or released when metal/ligand incorporates with protein and enables us to estimate the binding affinity and enthalpy according to the analysis of the binding curve fit. In the ITC experiment of PRORP2 with Mg^{2+}, analysis of the binding curve reveals an association constant (K_a) 6.78 x 10^3 M^{-1} and a dissociation constant (K_d) of 0.147 mM. The resulted binding curve gives a positive enthalpy of 218.3 cal/mol. When comparing the data obtained from the binding curve, it shows 2 binding sites but the positive enthalpy value suggests an endothermic binding. Comparatively, previous study has shown a K_d value of $K_{d,app}^{Ca(II)} = 40 \pm 10$ μM for the binding of Ca(II) with RNA based RNase P-tRNA complex [34]. In another study, a $K_{1/2}$ value of 6.7 ± 2.0 mM has been observed at the binding of Cd^{2+} with native RNA based RNase P, suggesting a strong metal binding affinity of RNA based RNase P [35].

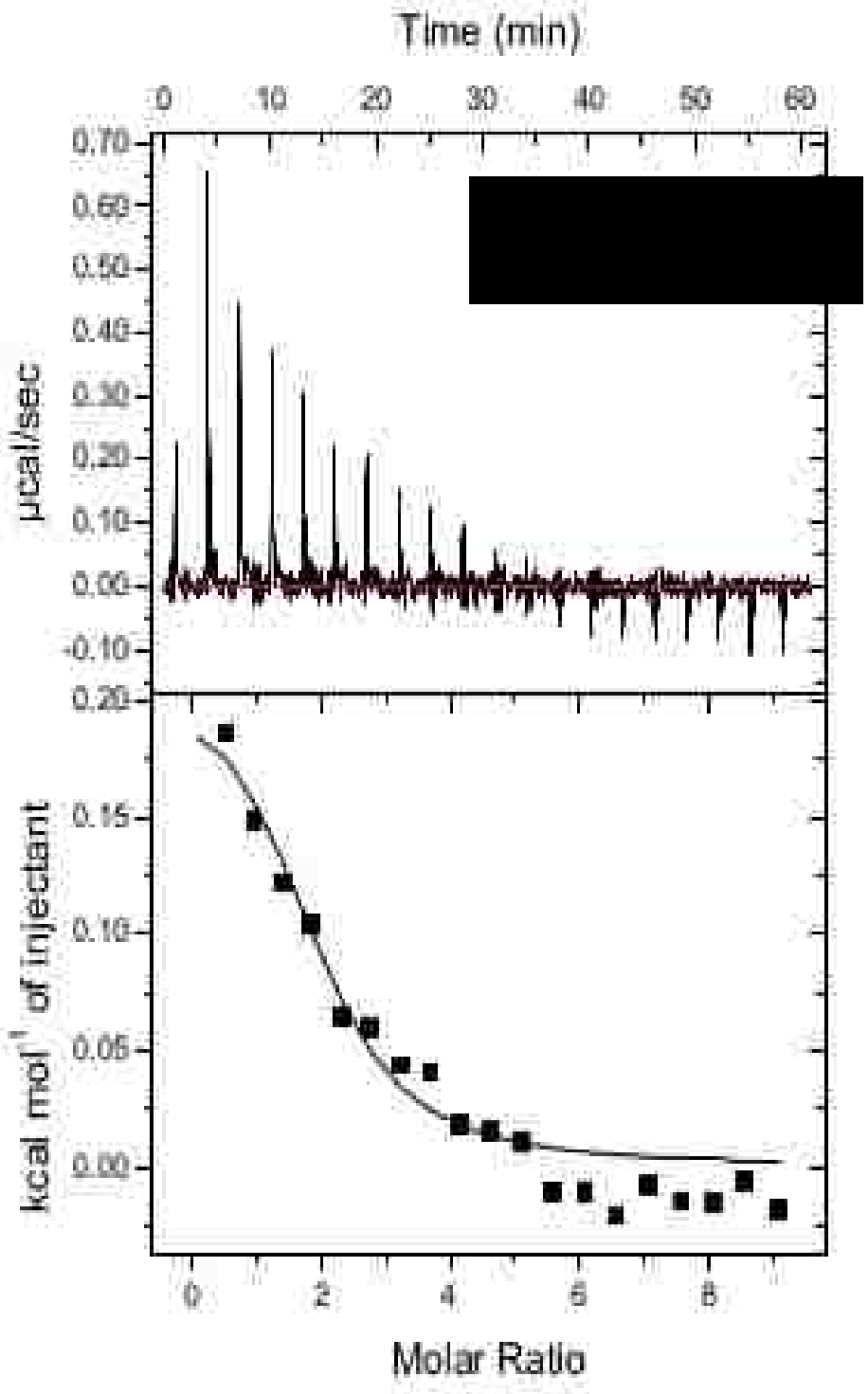

Figure 13. Isothermal titration calorimetry (ITC) titration with PRORP2 (0.45 mM) and Mg^{2+} (20 mM) at 22 ºC (Without fixing the *n* value).

Similarly, PRORP2 shows 1.53 ± 0.84 metal binding sites with Mn^{2+} (Figure 14). The binding curve of PRORP2 with Mn^{2+} shows comparatively higher dissociation constant of 1.196 mM. This ITC result suggests that that Mn^{2+} binds approximately 10x weaker affinity to PRORP2 compared to Mg^{2+}. The positive enthalpy value suggests that, like Mg^{2+}-PRORP, this metal binding interaction is endothermic.

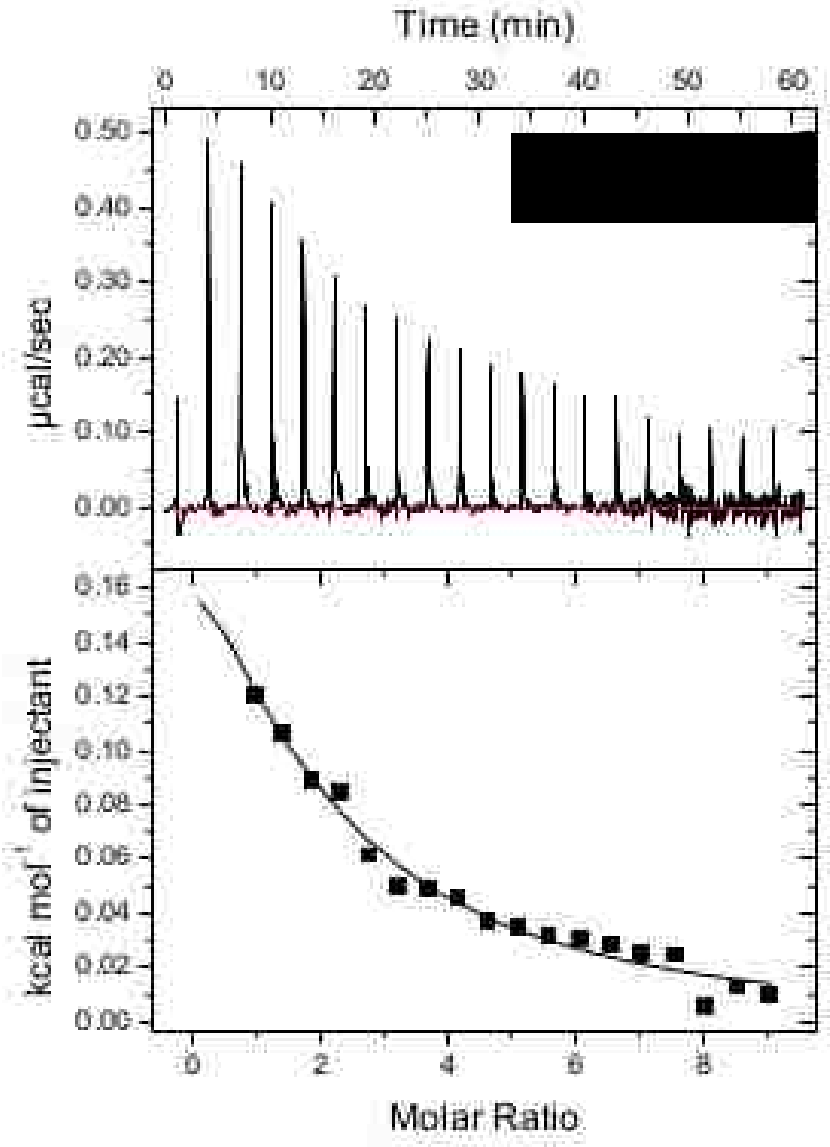

Figure 14. Isothermal titration calorimetry (ITC) titration with PRORP2 (0.45 mM) and Mn^{2+} (20 mM) at 22 ºC (Without fixing the *n* value).

Though both titrations of PRORP2 with Mg^{2+} and Mn^{2+} show two binding sites, the reactions have shown positive enthalpy suggesting that both $Mg^{2+}$and Mn^{2+} metal binding to PRORP2 are endothermic in the given conditions. These results suggest that the protein may require additional factors (substrate binding) or conformational changes that would enable the PRORP2 to incorporate the metals into the enzyme active site.

Investigation of two metal ion binding in the active site of PRORP protein isoforms using EPR. According to the previous crystal structure of PRORP1 metalloendonuclease (PDB: 4G24), two Mn^{2+} ions are positioned 3.5 Å from each other within the enzyme active site (Figure 15). This metal-ion enzyme active site is located in a surface accessible region of the protein, where inner-sphere metal ion coordination involves 4-5 aspartate amino acid residues as well as 3-6 water ligands.

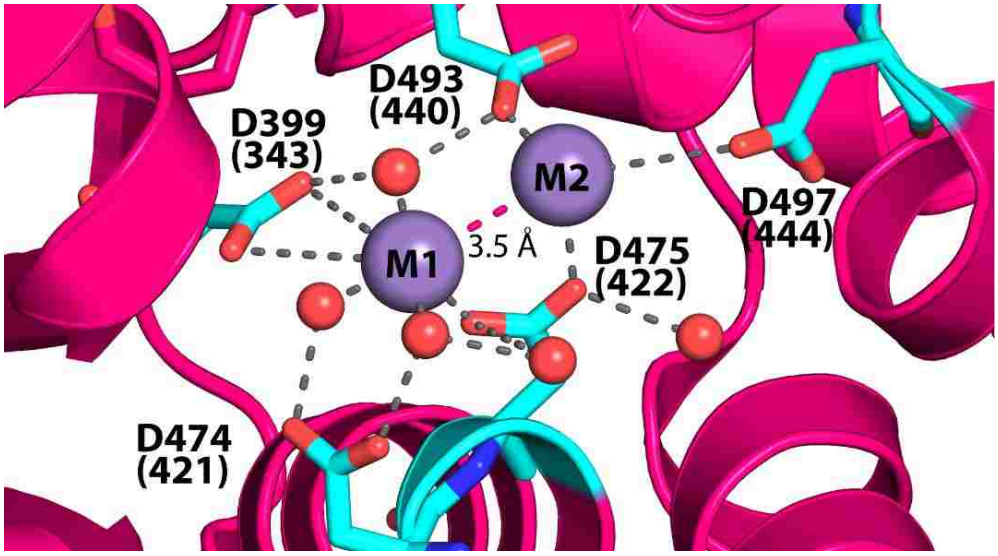

Figure 15. The active site of *A. thaliana* PRORP1 with two Mn^{2+} ions (purple spheres), labeled aspartate residues (cyan), water molecules (red spheres) generated using PyMOL software. Aspartate residues which are similar to *A. thaliana* PRORP3 are shown in brackets.

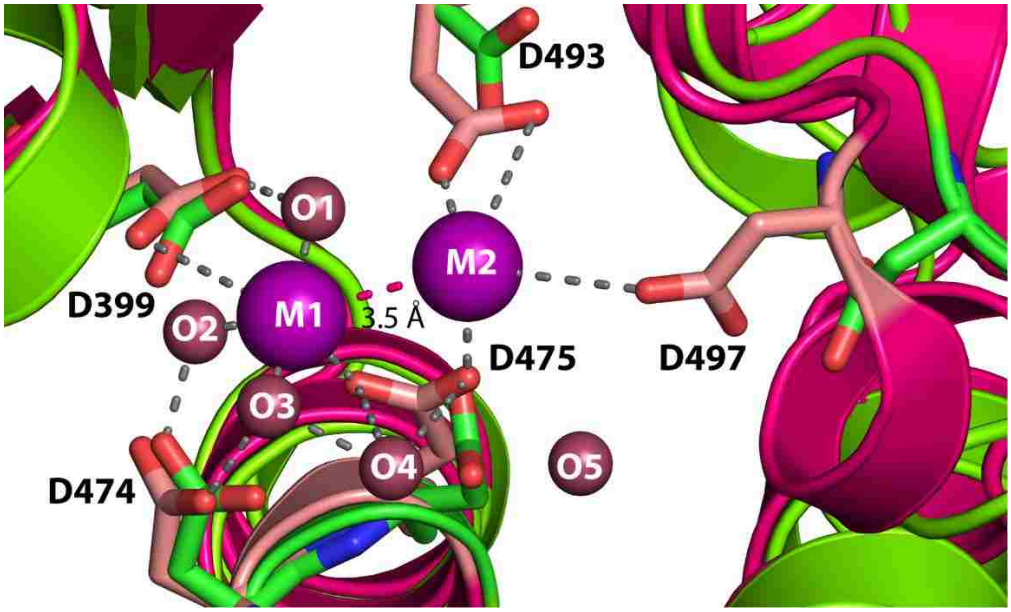

Figure 16. Overlay of the active sites of *A. thaliana* PRORP1 and PRORP2. The active sites of *A. thaliana* PRORP1 (magenta) with two Mn^{2+} ions (purple spheres), labeled aspartate residues, water molecules (red spheres) and aspartate residues in the PRORP2 active site (green) aligned using PyMOL software.

Based on this X-ray structure, a spectroscopic study aimed at defining the Mn^{2+} affinity and PRORP-metal ion environment were pursued. Characterization of mononuclear and a dinuclear spin-coupled (Mn^{2+}) of enzymes by electron paramagnetic resonance has proven to be useful in understanding many metalloenzyme systems [36-40] ; (1) to define the biochemical properties of the catalytic metal species and (2) to better understand the mechanistic roles of metal ligands to the dinuclear metal center. Here, we designed an EPR study of PRORP2 and PRORP3 (structural and functionally related homologs to PRORP1) in the presence of Mn^{2+} to examine the metal ion environment and the putative dinuclear interactions of the two catalytic metal ions. The reason for selecting Mn^{2+} instead of Mg^{2+}, which is the endogenous catalytic metal for all PRORP family members, is due to the paramagnetism observed in Mn^{2+} metal center as well as demonstrated PRORP1 catalytic activity shown in the presence of Mn^{2+} [27, 41]. For a

pure dimanganese species with antiferromagnetic coupling, there is no EPR spectrum from the $S' = 0$ ground state at very low temperatures. As the temperature is increased, excited spin states $S' = 1, 2,...$ 5 are populated according to the Boltzmann distribution, depending on the exchange coupling, J, where the relationship between temperature and J is 1.4 K ~ 1 cm^{-1}. The situation is complicated by the presence of aqueous or adventitiously-bound mononuclear Mn(II), as the $S = 5/2$ ground state EPR signal persists to very low temperature; however, at e.g. 5 K the mononuclear signal usually exhibits marked rapid-passage distortion unless special precautions are taken. Nevertheless, recording the EPR at various temperatures is important when trying to identify a coupled center. Despite the spectral complexity observed in Mn^{2+}-based EPR, the diagnostic features described above allow us to only distinguish and qualitatively characterize the metal and metal-ligand environment of the PRORP metalloendonuclease family.

EPR of Mn(II) is complex. Mn(II) is a high-spin d5 system with $S = 5/2$. The zero-field splittings (ZFS) are typically small, smaller than the Zeeman interaction, and five sets of transitions are possible: +5/2→+3/2, +3/2→+1/2, +1/2→-1/2, -1/2→-3/2, and -3/2→-5/2. The ZFS tensor, **D**, is typically rhombic, resulting in distinct sets of transitions in *x*, *y*, and *z* in the first-derivative powder spectrum of Mn(II) in frozen solution, i.e. 15 sets of transitions. The ^{55}Mn nucleus has a spin $I = 5/2$, with nuclear spin states M_I = +5/2, +3/2, +1/2, -1/2, -3/2 and -5/2 and, therefore, each set of transitions itself contains six individual transitions due to electron-nuclear coupling. Therefore, a *minimum* of 90 lines (i.e. $5 \times 3 \times 6$) in the EPR spectrum are expected. However, additional transitions are also observed. The small ZFS values allow the microwave quantum to access inter-doublet energy level spacings, providing inter-doublet transitions

between M_S levels. In addition, the Zeeman, ZFS, and hyperfine interaction energies are of the same order of magnitude and the quantum numbers are not well defined, the outcome of which is that hyperfine transitions that are predicted to be forbidden by perturbation theory are, in reality, partially allowed and, indeed, are observed. The typical Mn(II) signal, therefore, contains some hundreds of lines (simulated in the top trace of Figure 17; these so-called exact simulations were carried out using full matrix-diagonalization and not perturbation theory).

In the case of biological systems or, indeed, any frozen solution system, these are not resolved due to the heterogeneity in local fields due to (i) structural heterogeneity of the Mn(II) sites themselves and (ii) heterogeneity of the solvent matrix, that provides a heterogeneous electric field due to the electric polarization of e.g. water, and a heterogeneous magnetic environment due to disordered solvent protons. The result of these heterogeneities is a distribution, or "strain", in each of the spin Hamiltonian parameters that lead to a broadening of the lines and a much-simplified spectrum, from which the ZFS parameters, in particular, can be estimated from observation or more quantitatively determined by simulation.

For octahedral Mn(II)(aq) in frozen solution, typical values of the axial ZFS term, D, lie around 500 MHz, and the spectrum, at first glance, looks like a simple six-line pattern, due to the $+1/2 \rightarrow -1/2$ transitions, with some "wiggles" in-between the main lines (forbidden hyperfine transitions), and "wings" due to the strain-broadened absorption in the outer M_S doublets (simulated in middle trace of Figure 17). Some fine structure at the low field is sometimes observed due to inter-doublet transitions. Where the Mn(II) environment is geometrically constrained to be of lower symmetry, e.g. as in Mn(II)-

EDTA or in some proteins, values of D are higher by factors of around three, and individual sets of transitions from the individual M_S doublets can be observed at fields much higher and lower than the region around the $+1/2 \rightarrow -1/2$ transitions (simulated in bottom trace of Figure 17). These may contain resolved patterns of lines or may be observed merely as "lumps", depending on the degree of strain in the ZFS parameters. It should be noted that while a spectrum of this latter type is diagnostic for defined coordination of Mn(II) by constraining ligands, e.g. as would be found in a protein site, the observation of a spectrum of the former, simpler-looking type cannot be taken as evidence against binding of Mn(II), particularly if the coordination sphere is only partially occupied by proteinaceous ligands and the remaining sites are occupied by water or other non-constraining ligands.

Finally, in the case of a di-Mn(II) site, the isotropic natures of the $\boldsymbol{g}$ and $\boldsymbol{A}$ tensors lead to a remarkably simple outcome in which, instead of a six-line 90 G hyperfine pattern, an eleven-line hyperfine pattern is observed with transitions separated by half the isolated ion value, i.e. ~ 45 G. Multiple patterns from the di-Mn(II) spin states $S' = 1, 2, \ldots 5$ may lead to many more than eleven lines. The intensities of the various sets of these 45 G-split transitions exhibit a temperature dependence that is dependent on the net spin state in which the transitions arise and the inter-spin coupling, J. Observation of one or more temperature-dependent 45 G-split patterns is therefore diagnostic for a dinuclear Mn(II) center. Because of strains and the potential for overlapping patterns, care should be taken when interpreting the lack of a 45 G pattern as evidence of absence of a second Mn(II) ion (parallel-mode $\mathbf{B}_0 || \mathbf{B}_1$ EPR can be used to confirm suspected cases).

In all cases, Mn^{2+}-EPR is temperature-dependent, where each spin level is weighted according to the Boltzmann population distribution. Mn^{2+} species can be partially differentiated based on their temperature-dependence and electron-nuclear hyperfine patterns, comprising Mn^{2+} bound protein spectral features and a variable amount of free $Mn(H_2O)_6^{2+}$ that may be present in the sample.

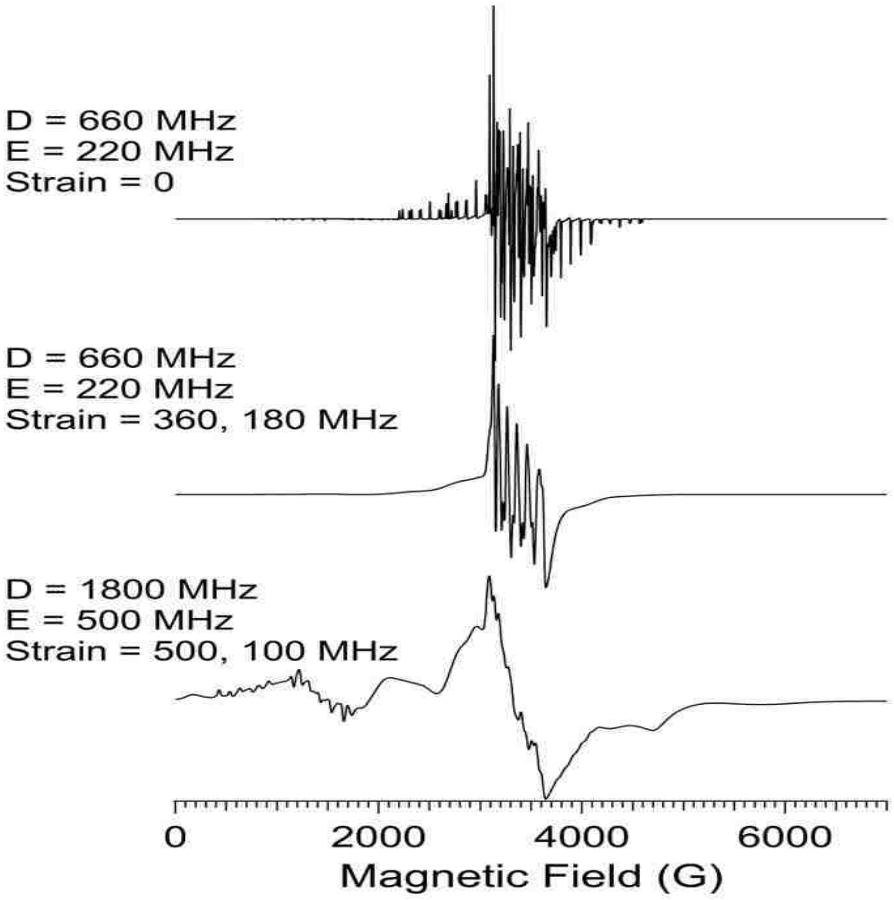

Figure 17. Exact matrix-diagonalization simulations of typical Mn(II) environments. Top: Modest zero-field splitting (ZFS) with no ZFS strain [corresponding to octahedral Mn(II) ions in a microcrystalline powder]; Middle: Modest ZFS with significant ZFS strains (σD, σE) [corresponding to octahedral Mn(II) ions in frozen aqueous solution]; Bottom: Large ZFS with significant ZFS strains [corresponding to hexacoordinate Mn(II) ions coordinated by geometry-constraining multidentate ligands]. ZFS parameters and strains are noted in the figure for each simulation. Other parameters were: g_{iso} = 2.0, A_{iso} = 270 MHz, σA = 10 MHz, ν = 9.5 GHz.

Evidence for a mononuclear Mn^{2+} species in PRORP2. Based upon the crystal structure of PRORP1, all metal-ligands derive from oxygen atoms that are either from aspartate amino acid residues or water molecules (Figure 15). The EPR spectrum obtained for PRORP2 (1 mM) sample that was incubated with 2 mM Mn^{2+} shows primarily a monomeric species at 5.5 K with 90 G splitting pattern (Figure 18A). In this sample, characteristic EPR features of a mononuclear Mn^{2+} species that are representative of metal bound PRORP include: (1) an intense, six-line ^{55}Mn-hyperfine pattern centered near 3300-3800 G ($g \approx 2$, A ~ 90 Gauss splitting); (2) a hyperfine-split resonance from 1400-2100 G that can be assigned to a "forbidden" $\Delta M_S = \pm 2$ field transition ($g = 4.05$, A = ~90 gauss splitting) (3) two broad featureless peaks from 2000-3000-3100 G; and (4) a very low field transition from 0-500 G ($g \geq 17$) that also exhibits ^{55}Mn hyperfine splitting. In addition, a control sample of 2mM Mn^{2+} in buffer (50 mM HEPES pH 7.5, 200 mM NaCl, 1mM TCEP, and 5% glycerol) reveals a characteristic six-line pattern for high spin Mn^{2+}center in the EPR spectrum, attributed to six coordinated $(Mn(H_2O)_6]^{2+}$ (Figure 18B). In addition, preparation of a reconstituted PRORP2 with 8 mM Mn^{2+} and subsequent passage over a G-25 desalting column resulted in spectral features that are indicative of an "adventitious" protein-bound Mn^{2+} species but it appears that the major species observed can be ascribed to six coordinated $[Mn(H_2O)_6]^{2+}$ (Figure 18C). Lastly, the PRORP2 protein by itself without any added metal ions and purified in the absence of EDTA does not show any bound Mn^{2+} in its EPR spectrum (Figure 18D). Instead, a weak axial signal that is diagnostic of aqueous Fe(III) (~1800 G) and two very weak signals that are likely attributed to Fe(III) bound to very minor sample impurities are observed in

addition to a signal (~3300 G, g ~ 2) that correlates to the Cu(II) cavity line of the instrument.

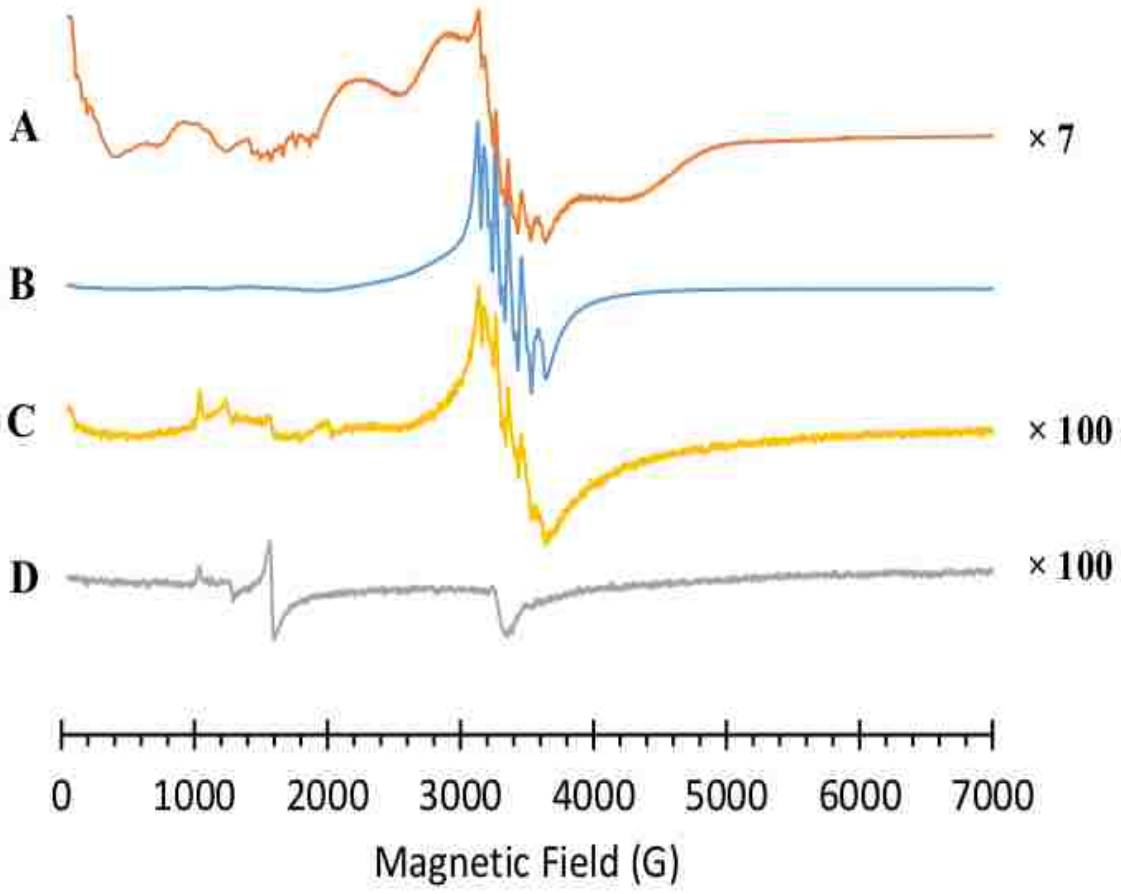

Figure 18. Low-temperature Electron paramagnetic resonance spectra of PRORP2 samples at 5.5 Kelvin. (A) 1 mM PRORP2 with 2 mM Mn^{2+} (B) Buffer with 2 mM Mn^{2+} (50 mM HEPES pH 7.5, 200 mM NaCl, 1 mM TCEP, 5 % glycerol) (C) Reconstituted PRORP2 with Mn^{2+} and (D) 1 mM PRORP2.

No evidence for an exchange coupled $(Mn^{2+})_2$ species in PRORP2. High-temperature EPR spectra of PRORP2 at 77 K reveal similar spectra features to the protein bound mononuclear Mn^{2+} at 5.5 K spectra, indicating that no spin-coupled $(Mn^{2+})_2$ cluster is present in the sample. No 40-45 G hyperfine splitting pattern that is diagnostic of a dinuclear species is present under these biochemical conditions and the reconstituted

sample (Figure 19C) is similar to spectrum collected at 5.5 K. This result suggests that the coordination environment of the Mn^{2+}-bound protein does not change.

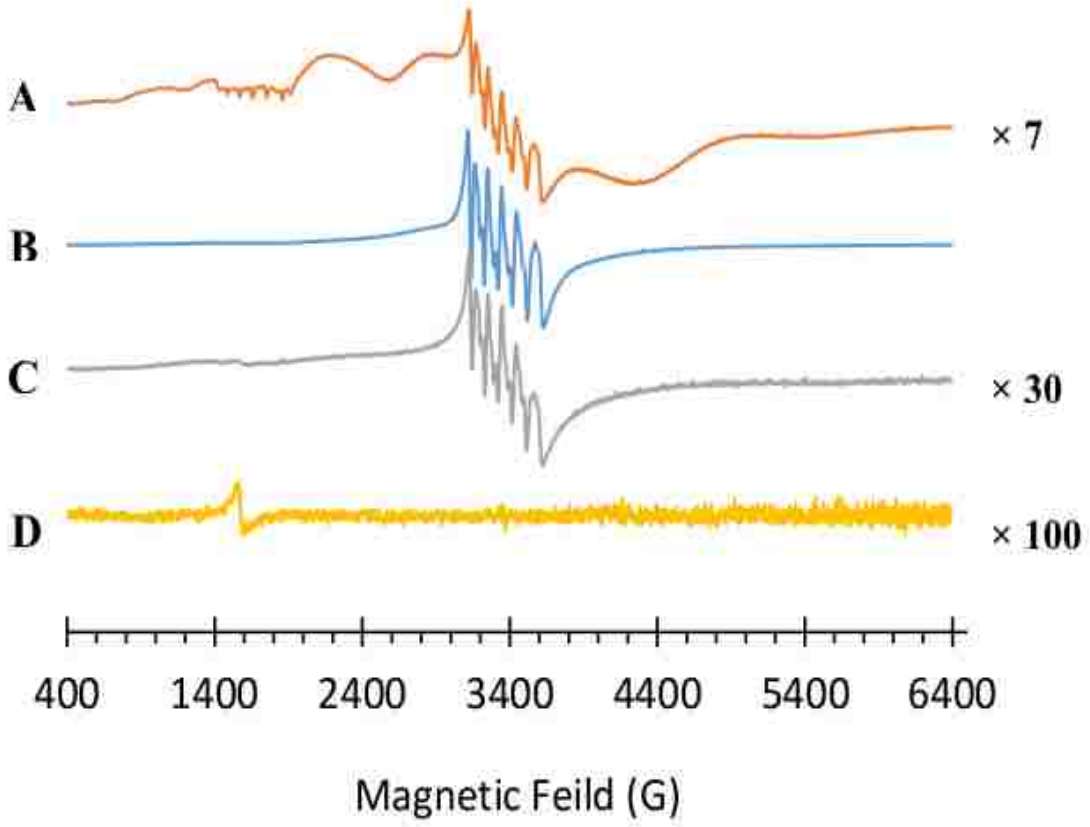

Figure 19. High-temperature EPR spectra of PRORP2 samples at 77 Kelvin. (A) PRORP2 with 2mM Mn^{2+} (same as Figure 18A), (B) Buffer with 2mM Mn^{2+} (same as Fig 18B), (C) reconstituted PRORP2 (same as Fig 18C), and (D) PRORP2-apo.

Purification of A. thaliana recombinant PRORP3 protein using affinity column. To date, there is no crystal structure of *A. thaliana* PRORP3 in Protein Data Bank. Although PRORP2 and PRORP3 in *A. thaliana* are isoenzymes and both function in the nucleus, the structures could be different from each other. We successfully purified His_6-tagged recombinant PRORP3 by one-step nickel affinity column chromatography (Figure 20).

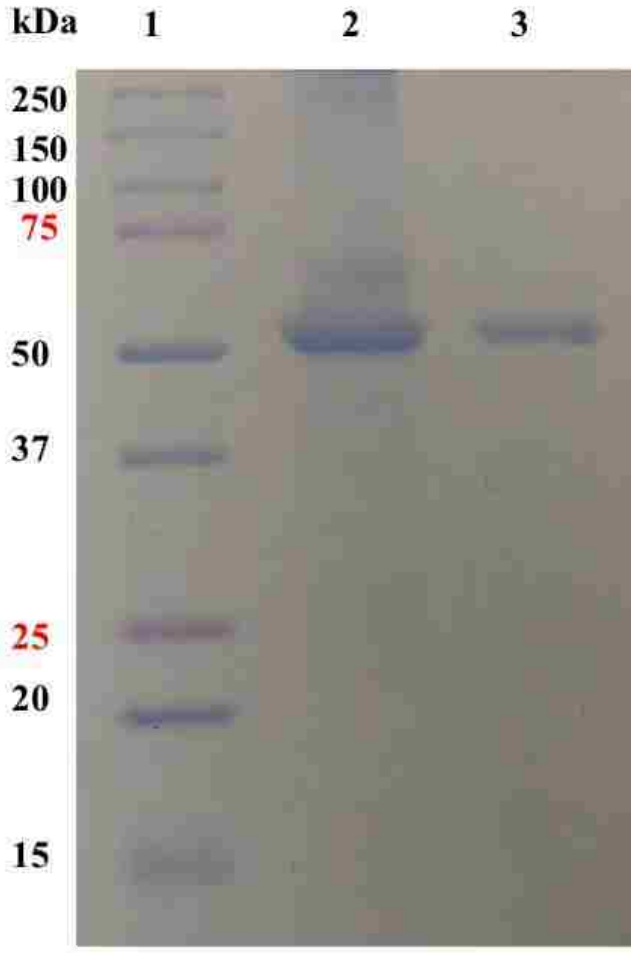

Figure 20. SDS PAGE gel (12 %) of purified PRORP3 protein (Lane 1-Protein molecular weight marker, Lane 2 purified protein of PRORP3, Lane 2 purified protein of PRORP3-D422N).

Determination of metal ion content in PRORP3 protein using Inductive coupled plasma mass spectrometry. Similarly, with PRORP2, ICP MS analysis was performed to investigate the metal ion content in the PRORP3 protein. The metal contents of the recombinant PRORP3 was shown in Table 2. The results show that there is a 1:2 stoichiometric mole ratio between PRORP3 and Zn^{2+}, respectively demonstrating that there are two Zn^{2+} ions in the monomeric form of PRORP3. In contrast, ICP MS data does not provide any Mn^{2+} metal binding sites in the PRORP3 under the tested

conditions. In both Zn^{2+} and Mn^{2+} ICP-MS studies PRORP3 protein was passed through a G-25 desalting column.

Table 2. Inductively coupled plasma mass spectrometry (ICP-MS) data of PRORP3.

Protein	Zn Mole ratio	Mn Mole ratio
Buffer	1.0 : 0.0	1.0 : 0.0
PRORP3 (10 μM)	1.0: 2.3	1.0 : 0.0

PRORP3 can fit to two metal binding sites with Mg^{2+} and Mn^{2+}. Similar to PRORP2, the energetics of Mg^{2+} and Mn^{2+} binding to PRORP3 were analyzed using ITC experiments at 22 ºC at pH 7.5 as mentioned in the experimental procedure. As shown in Figure 21 and 22, both titrations of Mg^{2+} and Mn^{2+} with PRORP3 show an endothermic reaction. Similar to PRORP2, the binding isotherm of Mg^{2+} and Mn^{2+} to PRORP3 best fitted to the two binding sites. However, it should be noted that PRORP3 required a manual fit of the data using Origin software whereas all PRORP2 results derived from automated data fitting Origin software. Three parameters (the hill coefficient (n), dissociation constant (K_d), change in enthalpy (ΔH)) were fitted and are shown in Table 3. Taken together, ITC experiment result suggests that both Mg^{2+} and Mn^{2+} bind to two sites on PRORP3 and that enthalpy values suggest that metal binding may not favorable with PRORP3. This positive enthalpic value suggests an endothermic binding and other factors or PRORP3 structural rearrangements may occur prior to metal binding. This

trend is consistent with PRORP2 and suggests that PRORP2 and PRORP3 may have similar structures and binding affinities.

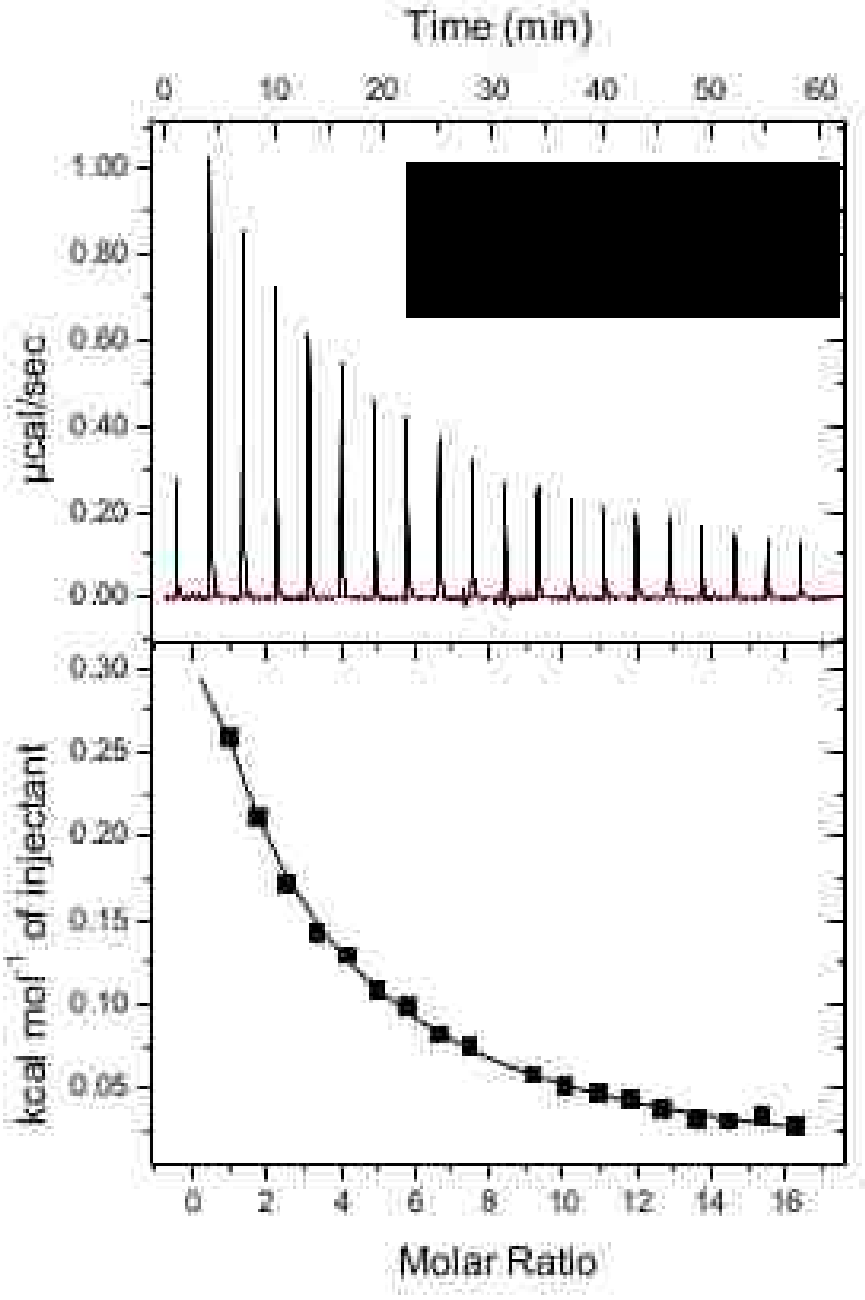

Figure 21. Isothermal titration calorimetry (ITC) titration with PRORP3 (0.25 mM) and Mg^{2+} (20 mM) at 22 ºC (With fixed *n* =2 value).

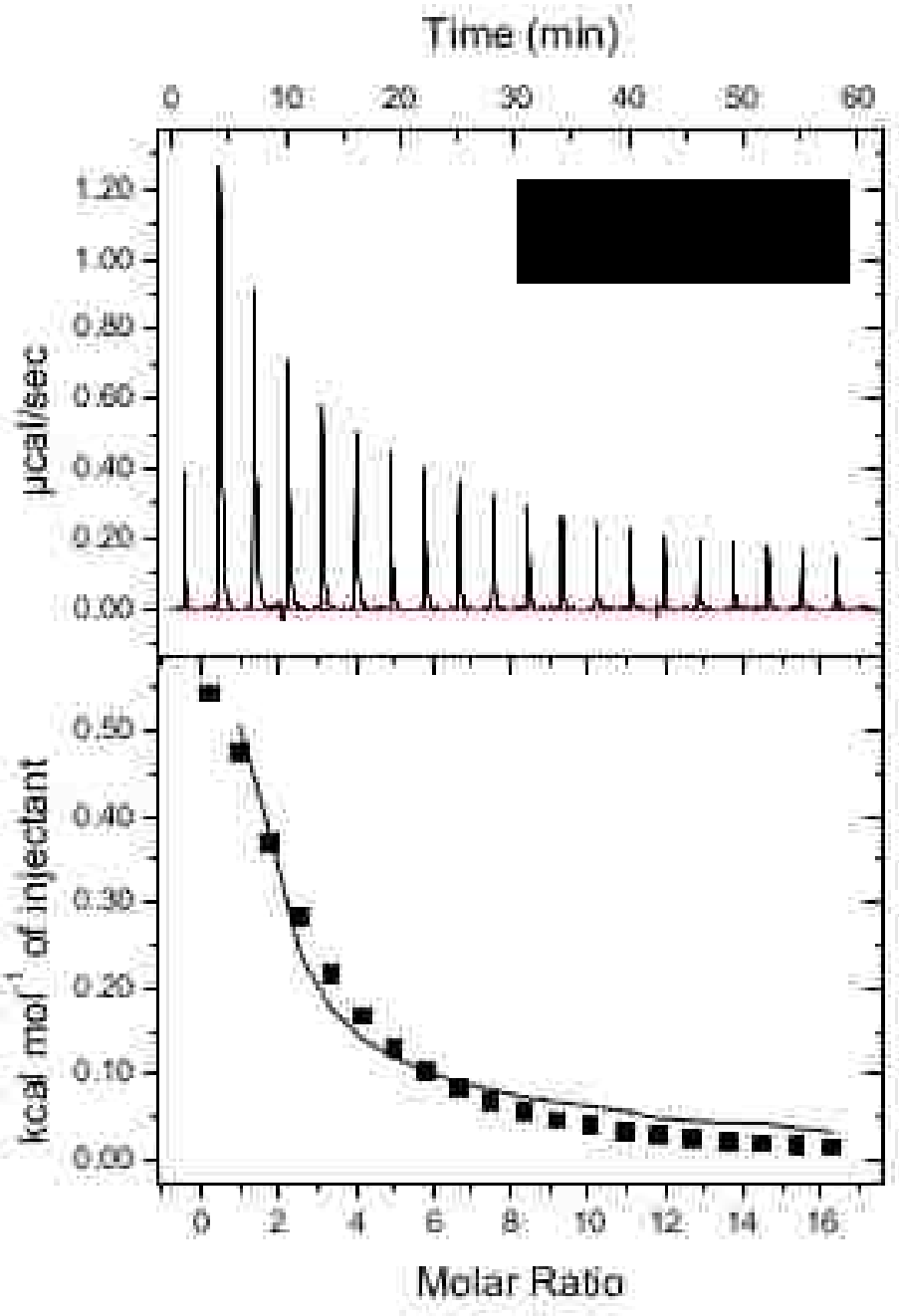

Figure 22. Isothermal titration calorimetry (ITC) titration with PRORP3 (0.25 mM) and Mn^{2+} (20 mM) at 22 ºC (With fixed *n* =2 value).

PRORP3-D422N does not fit two metal binding sites with Mg^{2+} and Mn^{2+}. A former student in our group has performed site-directed mutagenesis to create a putative active site mutant of the PRORP3-D422N protein. The Aspartate to Asparagine mutation at position 422 likely creates a different environment in the active site compared to PRORP3. After overexpressing and purifying this D422N PRORP3 mutant, we carried out similar ITC experiments with the PRORP3-D422N mutant protein. The protein was over-expressed and purified one step nickel affinity column (Figure 20-Lane 2). ITC experiments were performed at 22 ºC and pH 7.5. The titrations show endothermic binding curve to both Mg^{2+} and Mn^{2+}. However, the binding isotherms did not fit into the normal binding curve given by the Origin software provided by MicroCal ITC instruments (Figure 35 and 36 - appendix). In fact, it is difficult to investigate or fit any metal binding data with the PRORP3-D422N mutant protein using ITC experiments.

Table 3. Isothermal calorimetry titration results obtained for PRORP3 using Origin software provided by MicroCal.

Protein (0.25 mM)	**Metal (20 mM)**	**Number of binding sites**	**Dissociation constant (mM)**	**Enthalpy (cal/mol)**
PRORP3	Mg^{2+}	2	1.33	1.04×10^3
PRORP3	Mn^{2+}	2	0.49	1.10×10^3

Similar to PRORP2, EPR spectroscopy was applied to both PRORP3 and the PRORP3-D422N proteins to observe the potential metal-ligand environment of these proteins. Unfortunately, neither PRORP3 nor PRORP3-D422N show strong metal binding character in low (5.5 K) or high temperatures (30 K) EPR spectra (Figure 23). In fact, both wild type PRORP3 and PRORP3-D422N protein samples appear to behave the same, with Mn(II) showing no or little binding to protein and a free $Mn(H_2O)^{2+}$ species that likely predominates in solution.

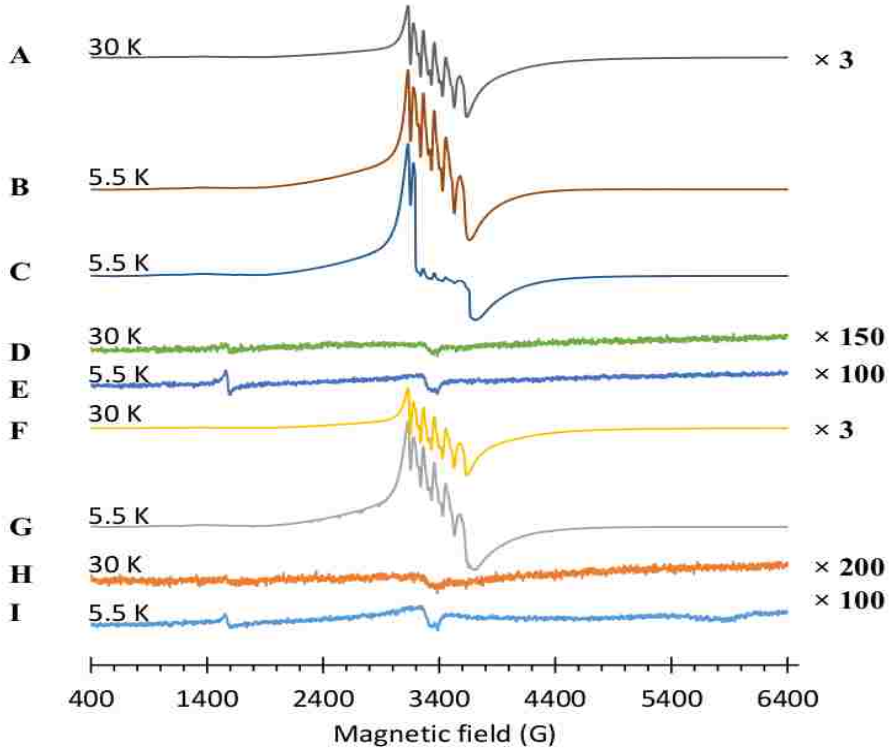

Figure 23. Electron paramagnetic resonance spectra of PRORP3 samples (A,B,C,D,E-PRORP3-D422N in the presence and absence of Mn^{2+}) in HEPES buffer pH 7.5 (50 mM HEPES, 200 mM NaCl, 1 mM TCEP, 5 % glycerol) (A) PRORP3-D422N (1 mM) with 2 mM Mn^{2+} at 30 K, (B) PRORP3-D422N (1 mM) with 2 mM Mn^{2+} at 5.5 K, (C) PRORP3-D422N (1 mM) with 2 mM Mn^{2+} at 5.5 K lower power (D) PRORP3-D422N (1 mM) at 30 K (E) PRORP3-D422N (1 mM) at 5.5 K (F,G,H,I) EPR spectra for PRORP3 in the presence and absence of Mn^{2+}) in HEPES buffer pH 7.5 (50 mM HEPES, 200 mM NaCl, 1 mM TCEP, 5 % glycerol) (F) PRORP3 (1 mM) with Mn^{2+}(2 mM) at 30 K (G) PRORP3 (1 mM) with Mn^{2+}(2 mM) at 5.5 K (H) PRORP3 (1 mM) at 30 K (I) PRORP3 (1 mM) at 5.5 K.

2.4. Conclusion

The primary goal of this study was to compare the metal binding properties of the RNase P proteins found in *A. thaliana* (PRORP1, PRORP2, and PRORP3). Though all 3 proteins are encoded by the nucleus, they are functioning in different locations in the cell. Previous investigations have proposed a dinuclear metal binding environment of the protein only RNase P and still, there are no direct evidence to prove the number of metal binding sites in the active site. For the first time, our ITC data has suggested the 1: 2 molar binding ratio of PRORP2 and PRORP3 with divalent metal ions. However, positive binding enthalpies suggest an endothermic binding and higher dissociation constants suggest that binding is less favorable under the experimental conditions we used. In contrast to ITC data, EPR data show strong metal binding character of one Mn^{2+} with the PRORP2 enzyme, suggesting that only one pre-arranged metal binding site exists in the active site of PRORP2. It is important to note that EPR data do not correlate with the proposed dinuclear metal binding in either PRORP2 or PRORP3. Instead, both ITC data together with EPR data support the existence of one metal binding site and perhaps a very weak second metal binding site. There may be a second metal binding site (M2) that is significantly weaker than site one (M1) and may require special conditions (conformational change or substrate binding) in order to form and for the PRORP family of enzymes to become fully functional.

CHAPTER 3

RNase P form in *Aquifex aeolicus*

3.1. Introduction

A new type of PRORP (minimal and RNA free RNase) was identified in *Aquifex aeolicus* which has shown a maximum activity when combined with another protein fraction of polynucleotide phosphorylase (PNPase) [42]. This protein is ~23 kDa in size (pI~5.1) and shows homology to the metallonuclease domain of PRORP. Though both *Aquifex aeolicus* RNase P and typical PRORP belongs to PIN domain-like superfamily, the relationship between them is still unknown [43]. This protein was proposed to form through a horizontal gene transfer event in *Aquifex aeolicus* and bioinformatics studies suggest that this protein gene exists in several archaea and some bacterium that also contain the ancient RNA-based RNase P enzyme. Thus, the *Aquifex aeolicus* RNase P may represent an evolutionary transitory enzyme that links the ribonucleoprotein-based RNase P form and the protein-only PRORP enzyme family. A primary goal of this project is to obtain structural information of this minimal protein-only RNase P.

3.1.1. Homology of RNase P of *Aquifex aeolicus* with PRORP1, PRORP2, and PRORP3

According to the sequence alignment data from Cluster omega software, *Aquifex aeolicus* PRORP shows similarities of 20.42 %, 22.80 %, and 21.24 % to PRORP1, PRORP2, and PRORP3 respectively. Interestingly, conserved active site residues; D399, D475, and D493 of PRORP1 were also found in *Aquifex aeolicus* RNase P suggesting their similar structural and functional properties (Figure 24).

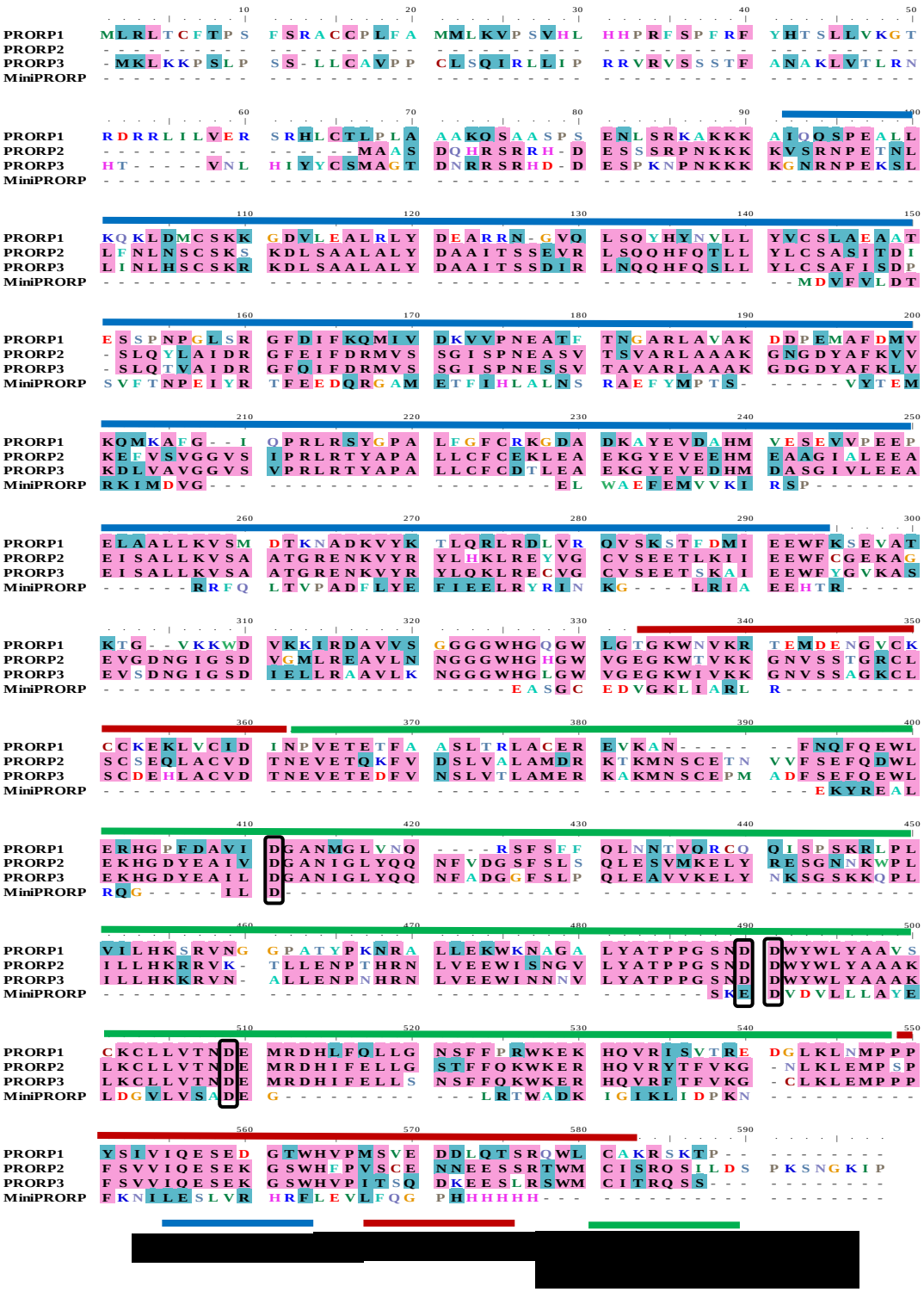

Figure 24. Sequence alignment of *A. thaliana* PRORP1, PRORP2, and PRORP3 with RNase P form found in *Aquifex aeolicus* using ClusterW multiple alignments- BioEdit. Identical amino acids highlighted in magenta and similar nucleotides highlighted in cyan.

3.2. Materials and Methods

3.2.1. Materials

Agar, kanamycin, ampicillin, imidazole, glycerol, Tween 20, Isopropyl-β-D-1-thiogalactopyranoside (IPTG), Monosodium phosphate (NaH_2PO_4), tris(2-carboxyethyl)phosphine (TCEP), Beta-mercaptoethanol (BME), PMSF (phenylmethylsulfonyl fluoride) were purchased from Sigma-Aldrich, VWR or Fisher Scientific. All other reagents were purchased commercially and were of the highest purity available.

Pre-packed immobilized-metal affinity chromatography nickel-nitrilotriacetic acid (IMAC Ni-NTA) columns (5 mL) and unpacked Ni-NTA resin were purchased from Bio-Rad. Plasmid preparation kits were purchased from Promega. All the solutions including buffer solutions were prepared using Nanopure water (~ 18.2 MΩ). Cells and proteins were handled on ice or at 4 ºC in a temperature controlled incubator.

3.2.2. Instruments/Apparatus

The New Brunswick Scientific Innova 4330 Refrigerated Incubator Shaker was used for growing and inducing of cell cultures. Cell lysis was performed using a Branson Sonifier sonicator. Pre-packed and manually packed IMAC Ni-NTA columns were connected with the BIORAD NCG chromatography system and used for the protein purification. Nuclear paramagnetic resonance (NMR) spectra were recorded on Agilent and Bruker 600 MHz NMR instruments.

3.2.3. Plasmid construction

A gene containing *Aquifex aeolicus* [42] was designed into a pET28a+ vector with cleavable C terminal His tag using C3 protease. The designed plasmid was purchased from Genescript and the highlighted yellow region indicates the C3 protease cleavage site.

MDVFVLDTSVFTNPEIYRTFEEDQRGAMETFIHLALNSRAEFYMPTSVYTEMRKI
MDVGELWAEFEMVVKIRSPRRFQLTVPADFLYEFIEELRYRINKGLRIAEEHTRE
ASGCEDVGKLIARLREKYREALRQGILDSKEDVDVLLLAYELDGVLVSADEGLR
TWADKIGIKLIDPKNFKNILESLVRHRFLEVLFQGP**HHHHHH**

3.2.4. Preparation of LB-agar plates and LB growing media

For the preparation of LB agar plates, a 5 g portion of LB agar powder was dissolved in 250 mL of nanopure water. Autoclaved growing media were cooled down and kanamycin was added up to a final concentration of 50 μg/mL before start the solidification. The mixture was transferred into sterile Petri dishes under the sterile conditions. The solidified dishes were stored at 4 ºC until needed.

For the preparation of LB growing media, a 25 g portion of LB agar powder was dissolved in 1 L of nanopure water followed by the autoclave treatment.

3.2.5. Preparation of the expression system of cells *Aquifex aeolicus* RNase P

The pET-28a^{+} plasmids containing the *Aquifex aeolicus* RNase P gene was transformed into BL21 (gold) cells using the heat-shocked at 42 ºC for 45 seconds. Transformed cells were mixed with 500 μL of autoclaved LB media which was at 37 ºC and incubated for 1 hour at 37 ºC while shaking at 225 r.p.m. kanamycin resistant LB-

agar plates were inoculated with the prepared expression system and incubated overnight at 37 ºC.

3.2.6. Test induction

A single colony from the freshly prepared expression system was transferred into two separate culture tubes containing sterile LB media (5 mL). The tube was incubated for 8 hours at 37 ºC for the cell growing. Gene expression was induced by adding IPTG up to a final concentration of 500 μM followed by the overnight incubation at 37 ºC. A 500 μL from each sample was centrifuged and cell pellets were separated. Cells were re-suspended in 200 μL of the lysis buffer and incubated at 95 ºC for 10 minutes. The resulting sample was analyzed by SDS-PAGE (10 %).

3.2.7. Growth, expression, and harvesting of *Aquifex aeolicus* RNase P

Starter culture was prepared by adding a single colony from the freshly prepared expression system into 50 mL of autoclaved LB media with kanamycin (50 μg/mL). The culture was incubated for 16-18 hours at 37 ºC while shaking at 220 r.p.m. This culture was used to inoculate 6 L of LB media containing kanamycin (50 μg/mL). The cultures were grown at 37 ºC and constant shaking at 225 r.p.m. until the optical density of ~0.8 at 600 nm was reached. The cell cultures were chilled using a temperature controlled incubator and induced with 500 μM of IPTG. Induced cultures were incubated for additional 16-18 hours at 18 ºC with constant shaking at 225 r.p.m. The cells were harvested by the centrifugation at 7000 r.p.m. for 10 minutes under 4 ºC. Cell pellets were stored at -80 ºC until needed.

3.2.8. Preparation of the minimal media (M9 media)

$FeCl_2$ solution. $FeCl_2.4H_2O$ (5 g), $CaCl_2.2H_2O$ (184 mg), H_3BO_3 (64 mg), $MnCl_2.4H_2O$ (40 mg), $CuCl_2.2H_2O$ (4 mg), $CoCl_2.6H_2O$ (18 mg), $ZnCl_2$ (340 mg), $NaMoO_4.2H_2O$ (605 mg), and concentrated HCl (8 mL) was added. The mixture was dissolved in water and make the volume up to 100 mL with water. The solution was stored in room temperature. ***SBM solution.*** KH_2PO_4 (16.5 g), K_2HPO_4 (87.5 g), and NaCl (18.25 g) were dissolved in 500 mL water. The solution was autoclaved and stored in room temperature. ***"O" solution***. $MgCl_2.6H_2O$ (28.8 g) and $FeCl_2$ stock solution (10 mL) were added into the water. The solution volume was made up to 500 mL using water. The solution was filter sterilized using 0.2 μm filters and stored in room temperature. ***"S" solution***. K_2SO_4 (4.8 g) was added into water (100 mL) and autoclaved. The solution was stored in room temperature. ***Thiamine solution.*** Thiamine hydrochloride (10 mg) was dissolved in water (10 mL). The solution was filter sterilized using a 0.2 μm filter.

Preparation of 1 L of minimal media. Autoclaved water (940 mL), SBM solution (40 mL) "S" solution (1 mL) were mixed and autoclaved. "O" solution (2 mL), Thiamine solution (1 mL) were added into the autoclaved growth medium of SBM and S solutions. $^{15}NH_4Cl$ (1 g) which was dissolved in water (5 mL) and glucose (4 g) which was dissolved in water (10 mL) was added into the growth medium.

3.2.9. Growing in minimal media (M9)

Starter culture was prepared by adding a single colony from the freshly prepared expression system into 5 mL of minimal media (M9) with kanamycin (50 μg/mL). The culture was incubated for 6 hours at 37 ºC while shaking at 220 r.p.m for cell growing. After 6 hours, 5 mL starter culture was added into fresh 200 mL of M9 media (scaling up) and culture was incubated for 16-18 hours at 37 ºC while shaking at 220 r.p.m for cell growing. Secondary starter culture was used to inoculate 6 L of M9 media containing kanamycin (50 μg/mL). The cultures were grown at 37 ºC and constant shaking at 225 r.p.m. until the optical density of ~0.8 at 600 nm was reached. The cell cultures were chilled using a temperature controlled incubator and induced with 500 μM of IPTG. Induced cultures were incubated for additional 16-18 hours at 18 ºC with constant shaking at 225 r.p.m. The cells were harvested by the centrifugation at 6000 r.p.m. for 10 minutes under 4 ºC. Cell pellets were stored at -80 ºC until needed.

3.2.10. Purification of the *Aquifex aeolicus* RNase P protein

Samples grown in both normal and minimal media were purified using the same optimized protocol. Cell pellets were re-suspended in the lysis buffer (50 mM HEPES, 250 mM NaCl, 5% glycerol, 10 mM BME, 0.2 % Tween 20, 1 mg/mL Lysozyme, 0.1 mg/mL PMSF, and protease inhibitor) and lysed by ultra-sonication at output power of 6 for 12 minutes on ice. The mixture was centrifuged at 16,000 rpm for 45 minutes under 4 ºC and separated supernatant was loaded on to the Ni-NTA column (5 mL) that was previously equilibrated with 100 mL of 50 mM HEPES buffer, pH 7.5 (250 mM NaCl, 5% glycerol, and 10 mM BME). The column was washed using ~20 column volumes

(CVs) of the wash buffer (50 mM HEPES buffer at pH 7.5, 250 mM NaCl, 5% glycerol, and 10 mM BME, 12.5 mM Imidazole) to remove nonspecifically bound proteins. The protein was eluted using ~20 CVs of the elution buffer (50 mM HEPES buffer at pH 7.5, 250 mM NaCl, 5% glycerol, and 10 mM BME, 100 mM Imidazole). The fractions containing *Aquifex aeolicus* RNase P were pooled and exchanged into 50 mM Tris buffer (50 mM Tris Buffer, 200 mM NaCl, 5 % glycerol, 2 mM TCEP) at pH 8.0. Protein sample was concentrated using an Amicon (10 K) ultracentrifugal concentrator and loaded onto a size exclusion column that was pre-equilibrated 50 mM Tris Buffer at pH 8.0 (200 mM NaCl, 5 % glycerol, and 2 mM TCEP). All the eluted fractions were analyzed by SDS-PAGE (10-12%) and the fractions containing targeted protein were pooled, and concentrated using 50 K Amicon ultra centrifugal units. Protein was exchanged into 50 mM HEPES buffer, pH 7.5 (250 mM NaCl, 5% glycerol and 1 mM TCEP). Protein concentrations were measured by the Nano Drop as well as the Bradford Assay.

3.2.11. Sample preparation for NMR

The protein sample was prepared in 10 % D_2O. A ^{1}H-^{15}N HSQC NMR experiment was performed in the Medical College of Wisconsin. The 2D ^{1}H-^{15}N HSQC spectrum was collected on a 600 MHz Bruker cryoprobe instrument.

3.2.12. Crystal screening of *Aquifex aeolicus* RNase P protein

Conditions for the crystal growth of Aquifex aeolicus RNase P protein were found using MD-LMB/HT 96 crystallization screen reagents. According to the optimized conditions, 2 μL of the 6 mg/mL protein solution was mixed with the crystallization

buffers in a volume ratio of 1:1 and allowed the crystal growing at 18 ºC by vapor diffusion hanging drop method. All the crystal preparations were performed in 24 well Hampton crystal trays.

3.3. Results and Discussion

Purification of RNase P form found in Aquifex aeolicus. Archaea domain is known to have the RNA based RNase P enzyme. However, recent data suggests that a protein-based RNase P also exists in the archaeal *Aquifex aeolicus* [42].

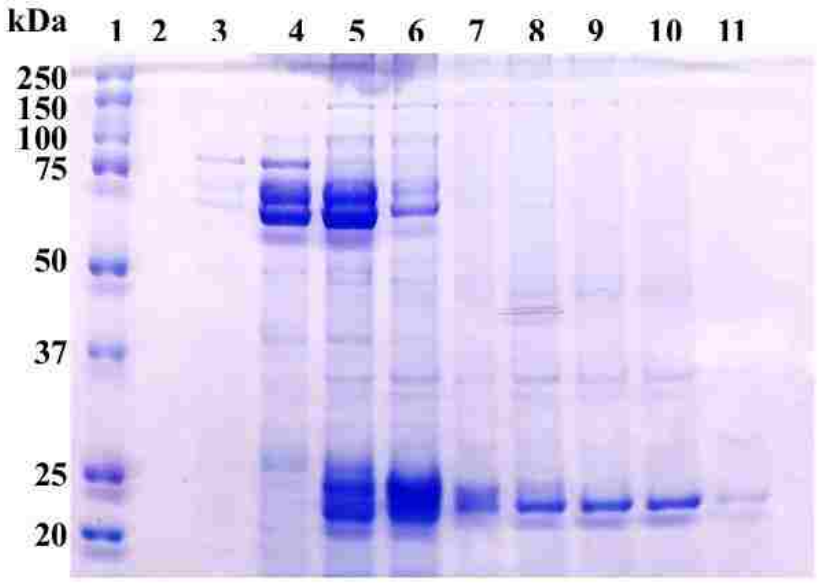

Figure 25. SDS PAGE gel (12 %) of eluted fractions of *Aquifex aeolicus* RNase P protein after Ni-column (Lane 1-Protein molecular weight marker, Lane 2 to 11-fraction 1 to 9).

This relatively small (23 kDa) putative protein appears to have some similarity to the metalloendonuclease domain of PRORP family yet it has poor sequence identity to the PRORP isoenzyme family. In this study, we prepared a His_6-tagged recombinant RNase P form from *Aquifex aeolicus* and the protein was overexpressed and purified

using a Ni-affinity chromatography according to the previous procedure [42]. SDS PAGE for the purification is shown in Figure 25. Though the molecular weight of this protein is 23 kDa, the size exclusion column elution profile (Appendix Figure 39) suggests the dimerization, trimerization, and extensive multimerization of the protein. Therefore, the purification is critical.

Protein fractions were passed through a 50 K Amicon filter unit and the filtrate was passed through an SEC to separate protein according to the molecular weight. Two protein bands appeared when passing through the size exclusion column and the SDS-PAGE analysis of eluted fractions are shown in Figure 26.

The first band appears (Appendix Figure 39) may contain multimerize species of the protein. However, both bands are likely forms of the *Aquifex aeolicus* RNase P.

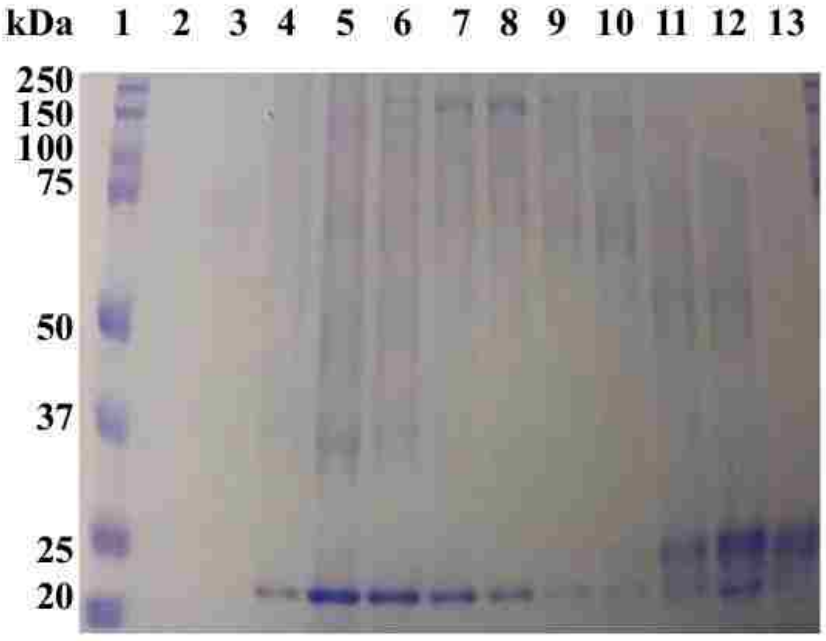

Figure 26. SDS PAGE gel (12 %) of eluted fractions of *Aquifex aeolicus* RNase P protein after size exclusion column (Lane 1, Protein molecular weight markers, Lane 2 to 13-fraction 1 to 12).

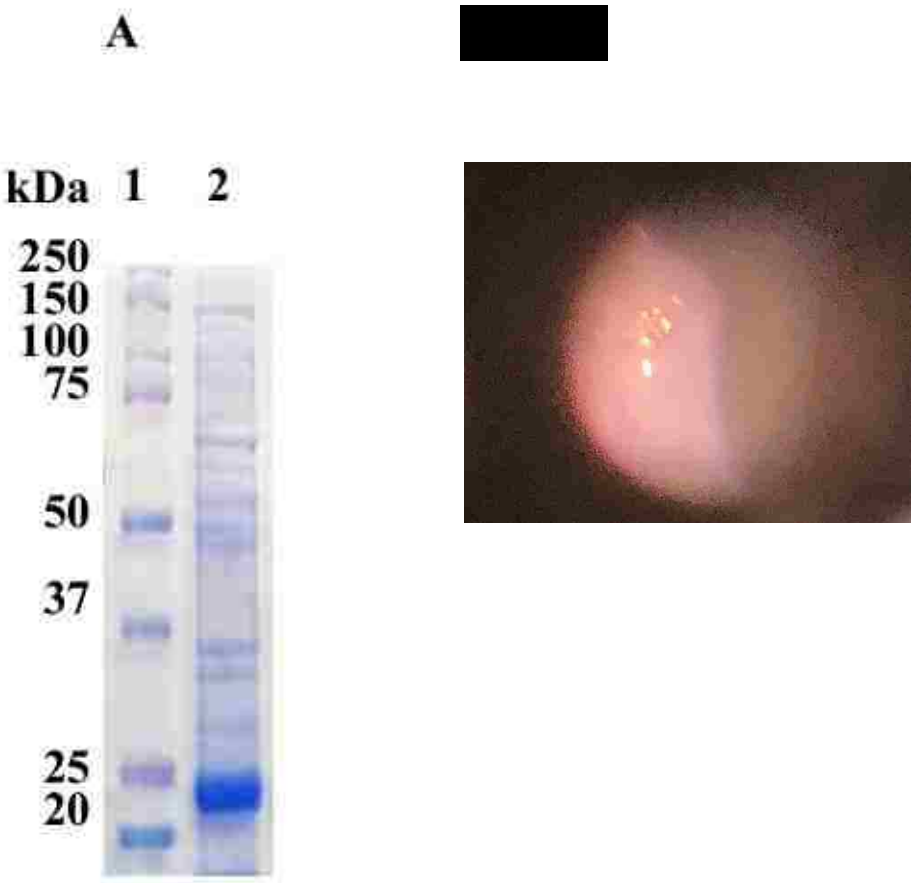

Figure 27. (A) SDS PAGE gel (12 %) of eluted fractions of *Aquifex aeolicus* RNase P protein. (Lane 1, Protein markers, Lane 2 Concentrated fractions of *Aquifex aeolicus* RNase P) (B) Image of crystals of the RNase P *Aquifex aeolicus* obtained from LMB screening.

Crystallization of RNase P form found in Aquifex aeolicus. As protein crystallization conditions for this protein are unknown, the purified protein was screened through LMB crystallization solutions using the hanging drop method. After one-month tiny crystals (Figure 27B) were observed in the conditions 28 % v/v PEG 300, 0.1 M Imidazole, 0.07 M Calcium acetate hydrate pH 7.0, crystals were mounted in 0.3 μm loop and was subjected to x-ray beam in Argonne national lab in Chicago. However, the size (10-20 μm in length and 5-10 μm in thickness) of the crystals was not large enough to diffract the x-ray.

Stability of RNase P form found in Aquifex aeolicus. To obtain suitable conditions for structural studies, a thermofluor assay was performed to find the stability of *Aquifex aeolicus* RNase P in different buffer conditions. Buffers that enhance the stability of the protein are shown in Table 4.

Table 4. Buffer conditions which show stable RNase P protein.

Buffer	pH	T_m (ºC)
KH_2PO_4	6.5	49.56
KH_2PO_4	7.2	50.00
MES	5.6	49.61
Sodium Citrate	5.7	49.97
Sodium Citrate	6.5	50.58

As the size of the protein is small (23 kDa), it is possible to perform protein NMR to determine the secondary structure of protein. Thus, 2D ^{1}H-^{15}N HSQC NMR was performed as the first step to get preliminary structural details of the protein (Figure 40 appendix). ^{15}N labeled protein was purified in M9 minimal media to characterize the ^{1}H-^{15}N bonds in the backbone of the protein. However, the preliminary NMR spectrum is of poor quality and needs to be reproduced on a new sample. It is possible that the *Aquifex aeolicus* RNase P protein undergoes dimerization or trimerization.

Homology model of the RNase P from Aquifex aeolicus. To date, no crystal structure is available for this minimal RNase P protein. Previous sequence analysis data has shown that the *Aquifex aeolicus* RNase P contains a "$D_{138}D_{142}E_{143}D_{160}$" residues which mimics the active site of PRORP protein ($D_{399}D_{474}D_{475}D_{493}$) [42]. Therefore, the structural information of the *Aquifex aeolicus* RNase P will be important to understand the function of this minimal protein-only RNase P.

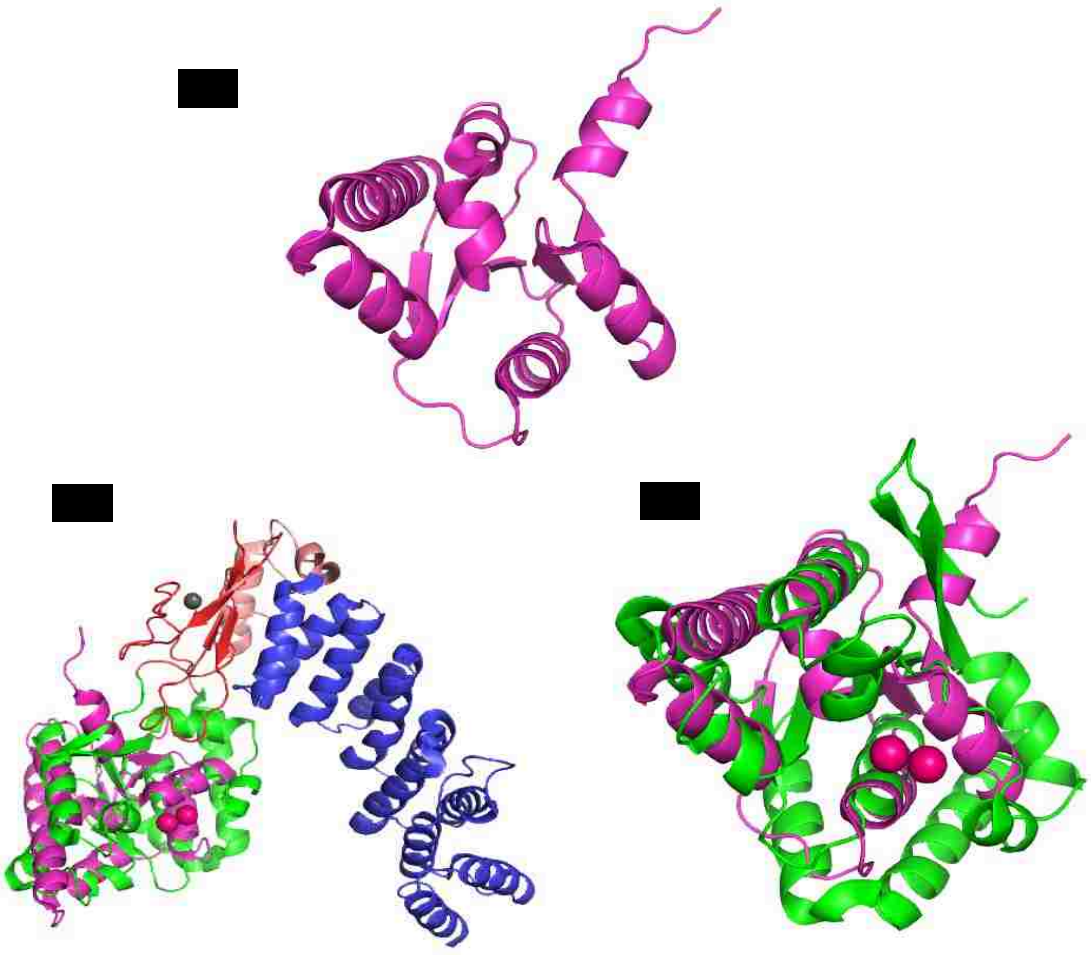

Figure 28. (A) Homology model of the RNase P from *Aquifex aeolicus* (HM-*Aa*RNase P), from Phyre2 web server, (B) Overlapped alignment of HM-*Aa*RNase P (magenta) with PRORP1 (PDB: 4G24) by PyMOL; PPR domain (blue) central domain (red), metallonuclease domain (green) Zn^{2+} (grey sphere) Mn^{2+} (red spheres), (C) enlarge view of the metallonuclease domain of PRORP1 (green) and HM-*Aa*RNase P (magenta).

A homology model of the *Aquifex aeolicus* RNase P was developed based on the structure of *Thermus thermophiles* HB8 (99 % confidence and 27 % similarity) using Phyre 2 protein folding recognition web server (Figure 28A). Interestingly, the homology model of the *Aquifex aeolicus* RNase P, partially overlapped with the metallonuclease domain of the crystal structure of PRORP1 (PDB: 4G24) suggesting that the RNase P found in *Aquifex aeolicus* could be a protein which resembles the metallonuclease domain of the PRORP enzymes. Future structural studies will be required to elucidate the tertiary fold and active site environment of *Aquifex aeolicus* RNase P. A key goal of this project will be to see if this minimal protein-inly RNase P isoform resembles the metalloendonuclease domain of the *A. thaliana* PRORP family.

3.4. Conclusion

It was known that the purification of *Aquifex aeolicus* RNase P is complex, due to the tendency of dimerization or trimerization. Interestingly, we have purified the *Aquifex aeolicus* RNase P in the soluble form and conditions were optimized for crystal preparation. However, crystals were small (10-20 μm in length and 5-10 μm in thickness) and not large enough for the X-ray diffraction. The homology model of the *Aquifex aeolicus* RNase P suggests some structural similarities to the metallonuclease domain of PRORP enzymes of *A. thaliana* and it is possible that the distinct isoforms utilize similar reaction mechanisms. Further screening and crystal optimizations to prepare larger crystals of *Aquifex aeolicus* RNase P will be important to reveal the tertiary fold and active site of the protein.

FUTURE PLANS

ITC experiments on PRORP enzymes with metal ions in the presence of substrate. Although ITC and EPR data have not provided definitive evidence about the metallocenter of PRORP proteins, it is important to understand the number of metal ions bound to the active center and the dynamic nature of the enzyme active site. It will be important to perform similar binding titrations in the presence of substrate or substrate analogues that may trap a specific E-S conformation and metal-ligand active site environment. Preparation of an E-S complex for structural studies and metal binding ITC experiments could be obtained by adding the cleavage site modified tRNA substrates (2' OH → 2' H, 2'F, or 2'OCH_3) to the purified PRORP proteins [44]. To obtain a non-hydrolyzable pre-tRNA substrate containing these site-specific modifications, two RNA molecules (5' leader RNA containing chemical modification and a near tRNA-like substrate) could be joined together using a DNA splint and the T4 DNA ligase enzyme. Here, the RNAs are bridged by a splint DNA in an RNA:DNA hybrid, allowing for insertion of a specific modification within an RNA and the assembly of a smaller synthetic RNA into the longer tRNA-like molecule [45]. Prior to the ITC titrations, each PRORP protein, including the *Aquifex aeolicus* RNase P, could also be detected using Circular Dichroism (CD) analysis to investigate the protein secondary structural changes that occur upon the mixing with substrate. Screening under CD will help to find the appropriate substrate concentration that could be used in ITC titrations. In addition, performing several ITC titrations with a series of divalent metal ions in the absence and presence of the substrate will be helpful to understand the metal binding properties of PRORP proteins. A set of optimal conditions in these binding studies will

be helpful for future EPR analysis of the enzyme-substrate in the absence and presence of various divalent metal ion concentrations.

Crystalizing of PRORP3 and the minimal RNase P found in Aquifex aeolicus. The PRORP3 and minimal RNase P (mini-PRORP) are two important protein-only RNase P enzymes with no structural information. Although we have prepared PRORP3 and the *Aquifex aeolicus* RNase P to high purity and have also optimized the conditions for *Aquifex aeolicus* RNase P crystals, we will need to obtain reproducible X-ray diffraction data and determine these crystal structures. Current conditions for minimal RNase P may not represent ideal crystal growth conditions and further optimization screens will be required. Future experiments where we use sparse matrix strategies will help us to identify new conditions, optimize the crystal size and diffraction properties around our current conditions.

Activity assay of PRORP2, PRORP3 and Aquifex aeolicus RNase P with metal ions. Activity of PRORP enzymes depend on the metal ion concentration (Eg: Mg^{2+} or Mn^{2+}. Therefore, activity assays of purified PRORP2, PRORP3 and *Aquifex aeolicus* RNase P proteins will provide an estimation of the degree of metalation.

REFERENCES

1. Tollervey, D.; Caceres, J. F., RNA processing marches on. *Cell* **2000,** *103*, 703-709.

2. Woodhams, M. D.; Stadler, P. F.; Penny, D.; Collins, L. J., RNase MRP and the RNA processing cascade in the eukaryotic ancestor. *BMC Evol. Bio.* **2007,** *7 Suppl 1*, S13-S13.

3. Kazantsev, A. V.; Pace, N. R., Bacterial RNase p: A new view of an ancient enzyme. *Nat. Rev. Microbiol.* **2006,** *4*, 729-740.

4. Reiter, N. J.; Osterman, A.; Torres-Larios, A.; Swinger, K. K.; Pan, T.; Mondragón, A., Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature* **2010,** *468*, 784-789.

5. Phizicky, E. M.; Hopper, A. K., tRNA biology charges to the front. *Genes Dev.* **2010,** *24*, 1832-1860.

6. Evans, D.; Marquez, S. M.; Pace, N. R., RNase P: Interface of the RNA and protein worlds. *Trends Biochem. Sci.* **2006,** *31*, 333-341.

7. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S., The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 1983, *35*, 849-857.

8. Pinker, F.; Giegé, P.; Sauter, C., Crystallization and crystallographic analysis of an *Arabidopsis* nuclear proteinaceous RNase P. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2015,** *71*, 1372-1377.

9. Walker, S. C.; Engelke, D. R., A protein-only RNase P in human mitochondria. *Cell* **2008,** *135*, 412-414.

10. Holzmann, J.; Frank, P.; Loffler, E.; Bennett, K. L.; Gerner, C.; Rossmanith, W., RNase P without RNA: Identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell* **2008,** *135*, 462-474.

11. Gobert, A.; Gutmann, B.; Taschner, A.; Gossringer, M.; Holzmann, J.; Hartmann, R. K.; Rossmanith, W.; Giege, P., A single *Arabidopsis* organellar protein has RNase P activity. *Nat. Struct. Mol. Biol.* **2010,** *17*, 740-744.

12. Taschner, A.; Weber, C.; Buzet, A.; Hartmann, R. K.; Hartig, A.; Rossmanith, W., Nuclear RNase P of *Trypanosoma brucei*: A single protein in place of the multicomponent RNA-protein complex. *Cell Rep.* **2012,** *2*, 19-25.

13. Rossmanith, W.; Holzmann, J., Processing mitochondrial (t)RNAs: New enzyme, old job. *Cell Cycle* **2009,** *8*, 1650-1653.

14. Holzmann, J.; Rossmanith, W., tRNA recognition, processing, and disease: Hypotheses around an unorthodox type of RNase P in human mitochondria. *Mitochondrion* **2009,** *9*, 284-288.

15. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S., The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **1983,** *35*, 849-857.

16. Tsai, H.-Y.; Pulukkunat, D. K.; Woznick, W. K.; Gopalan, V., Functional reconstitution and characterization of *Pyrococcus furiosus* RNase P. *Proc. Natl. Acad. Sci.* **2006,** *103*, 16147-16152.

17. Gopalan, V., Uniformity amid diversity in RNase P. *Proc. Natl. Acad. Sci.* **2007,** *104*, 2031-2032.

18. Jarrous, N.; Altman, S., Nicholson, A. W., Ed. Human ribonuclease P. *Methods in enzymol.* 2001; Vol. 342, pp 93-100.

19. Walker, S. C.; Engelke, D. R., Ribonuclease P: The evolution of an ancient RNA enzyme. *Crit. Rev. Biochem. Mol. Biol.* **2006,** *41*, 77-102.

20. Jarrous, N.; Reiner, R., Human RNase P: A tRNA-processing enzyme and transcription factor. *Nucleic acids res.* **2007,** *35*, 3519-3524.

21. Wu, J.; Niu, S.; Tan, M.; Huang, C.; Li, M.; Song, Y.; Wang, Q.; Chen, J.; Shi, S.; Lan, P.; Lei, M., Cryo-EM structure of the human ribonuclease P holoenzyme. *Cell* **2018,** *175*, 1393-1404.

22. Lan, P.; Tan, M.; Zhang, Y.; Niu, S.; Chen, J.; Shi, S.; Qiu, S.; Wang, X.; Peng, X.; Cai, G.; Cheng, H.; Wu, J.; Li, G.; Lei, M., Structural insight into precursor tRNA processing by yeast ribonuclease P. *Science* **2018,** *362*, 6678.

23. Wan, F.; Wang, Q.; Tan, J.; Tan, M.; Chen, J.; Shi, S.; Lan, P.; Wu, J.; Lei, M., Cryo-electron microscopy structure of an archaeal ribonuclease P holoenzyme. *Nat. Commun.* **2019,** *10*, 2617.

24. Mondragón, A., Structural studies of RNase P. *Annu. Rev. Biophys.* **2013,** *42*, 537-557.

25. Wang, M. J.; Davis, N. W.; Gegenheimer, P., Novel mechanisms for maturation of chloroplast transfer RNA precursors. *EMBO J.* **1988,** *7*, 1567-1574.

26. Lechner, M.; Rossmanith, W.; Hartmann, R. K.; Thölken, C.; Gutmann, B.; Giegé, P.; Gobert, A., Distribution of ribonucleoprotein and protein-only RNase P in eukarya. *Mol. Biol. Evol.* **2015,** *32*, 3186-3193.

27. Howard, M. J.; Lim, W. H.; Fierke, C. A.; Koutmos, M., Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5' processing. *Proc. Natl. Acad. Sci. U. S. A.* **2012,** *109*, 16149-16154.

28. Oerum, S.; Roovers, M.; Rambo, R. P.; Kopec, J.; Bailey, H. J.; Fitzpatrick, F.; Newman, J. A.; Newman, W. G.; Amberger, A.; Zschocke, J.; Droogmans, L.; Oppermann, U.; Yue, W. W., Structural insight into the human mitochondrial tRNA purine N1-methyltransferase and ribonuclease P complexes. *J. Biol. Chem.* **2018,** *293*, 12862-12876.

29. Holzmann, J.; Frank, P.; Löffler, E.; Bennett, K. L.; Gerner, C.; Rossmanith, W., RNase P without RNA: Identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell* **2008,** *135*, 462-474.

30. Klemm, B. P.; Karasik, A.; Kaitany, K. J.; Shanmuganathan, A.; Henley, M. J.; Thelen, A. Z.; Dewar, A. J. L.; Jackson, N. D.; Koutmos, M.; Fierke, C. A., Molecular recognition of pre-tRNA by *Arabidopsi*s protein-only ribonuclease P. *RNA* **2017,** *23*, 1860-1873.

31. Anantharaman, V.; Aravind, L., The NYN domains: Novel predicted RNases with a PIN domain-like fold. *RNA Biol.* **2006,** *3*, 18-27.

32. Gobert, A.; Pinker, F.; Fuchsbauer, O.; Gutmann, B.; Boutin, R.; Roblin, P.; Sauter, C.; Giegé, P., Structural insights into protein-only RNase P complexed with tRNA. *Nat. Commun.* **2013,** *4*, 1353.

33. Karasik, A.; Shanmuganathan, A.; Howard, M. J.; Fierke, C. A.; Koutmos, M., Nuclear protein-only ribonuclease P2 structure and biochemical characterization provide insight into the conserved properties of tRNA 5' end processing enzymes. *J. Mol. Biol.* **2016,** *428*, 26-40.

34. Hsieh, J.; Koutmou, K. S.; Rueda, D.; Koutmos, M.; Walter, N. G.; Fierke, C. A. *J. Mol. Biol.* **2010**, *400*, 38.

35. Christian, E. L.; Smith, K. M. J.; Perera, N.; Harris, M. E. *RNA (New York, N.Y.)* **2006**, *12*, 1463.

36. Reed, G. H.; Ray, W. J., Electron paramagnetic resonance studies of manganese(II) coordination in the phosphoglucomutase system. *Biochemistry* **1971,** *10*, 3190-3197.

37. Copik, A.; Nocek, B.; I Swierczek, S.; Ruebush, S.; Jang, S.; Meng, L.; D'Souza, V.; W Peters, J.; Bennett, B.; Holz, R., EPR and X-ray crystallographic characterization of the product-bound form of the Mn(II) -loaded methionyl aminopeptidase from *Pyrococcus furiosus*. *Inorg. Chem*. 2005; Vol. 44, p 121-129.

38. D'souza, V. M.; Brown, R. S.; Bennett, B.; Holz, R. C., Characterization of the active site and insight into the binding mode of the anti-angiogenesis agent fumagillin to the manganese(II)-loaded methionyl aminopeptidase from *Escherichia coli*. *J. Biol. Inorg. Chem.* **2005,** *10*, 41-50.

39. Rusnak, F.; Yu, L.; Todorovic, S.; Mertz, P., Interaction of bacteriophage λ protein phosphatase with Mn(II): Evidence for the formation of a [Mn(II)]2 cluster. *Biochemistry* **1999,** *38*, 6943-6952.

40. Sharma, A.; Gaidamakova, E. K.; Matrosova, V. Y.; Bennett, B.; Daly, M. J.; Hoffman, B. M., Responses of Mn(II) speciation in *Deinococcus radiodurans* and *Escherichia coli* to γ-radiation by advanced paramagnetic resonance methods. *Proc. Natl. Acad. Sci. U.S.A.* **2013,** *110*, 5945-5950.

41. Howard, M. J.; Karasik, A.; Klemm, B. P.; Mei, C.; Shanmuganathan, A.; Fierke, C. A.; Koutmos, M., Differential substrate recognition by isozymes of plant protein-only ribonuclease P. *RNA* **2016,** *22*, 782-792.

42. Nickel, A. I.; Wäber, N. B.; Gößringer, M.; Lechner, M.; Linne, U.; Toth, U.; Rossmanith, W.; Hartmann, R. K., Minimal and RNA-free RNase P in A*quifex aeolicus*. *Proc. Natl. Acad. Sci.* **2017**, 201707862.

43. Daniels, C. J.; Lai, L. B.; Chen, T.-H.; Gopalan, V., Both kinds of RNase P in all domains of life: Surprises galore. *RNA* **2019,** *25*, 286-291.

44. Cuzic-Feltens, S.; Weber, M. H. W.; Hartmann, R. K., Investigation of catalysis by bacterial RNase P via LNA and other modifications at the scissile phosphodiester. *Nucleic Acids Res.* **2009,** *37*, 7638-7653.

45. Kershaw, C. J.; O'Keefe, R. T. *Methods Mol. Biol.* **2012**, *941*, 257.

APPENDIX

Chapter 2 Appendix

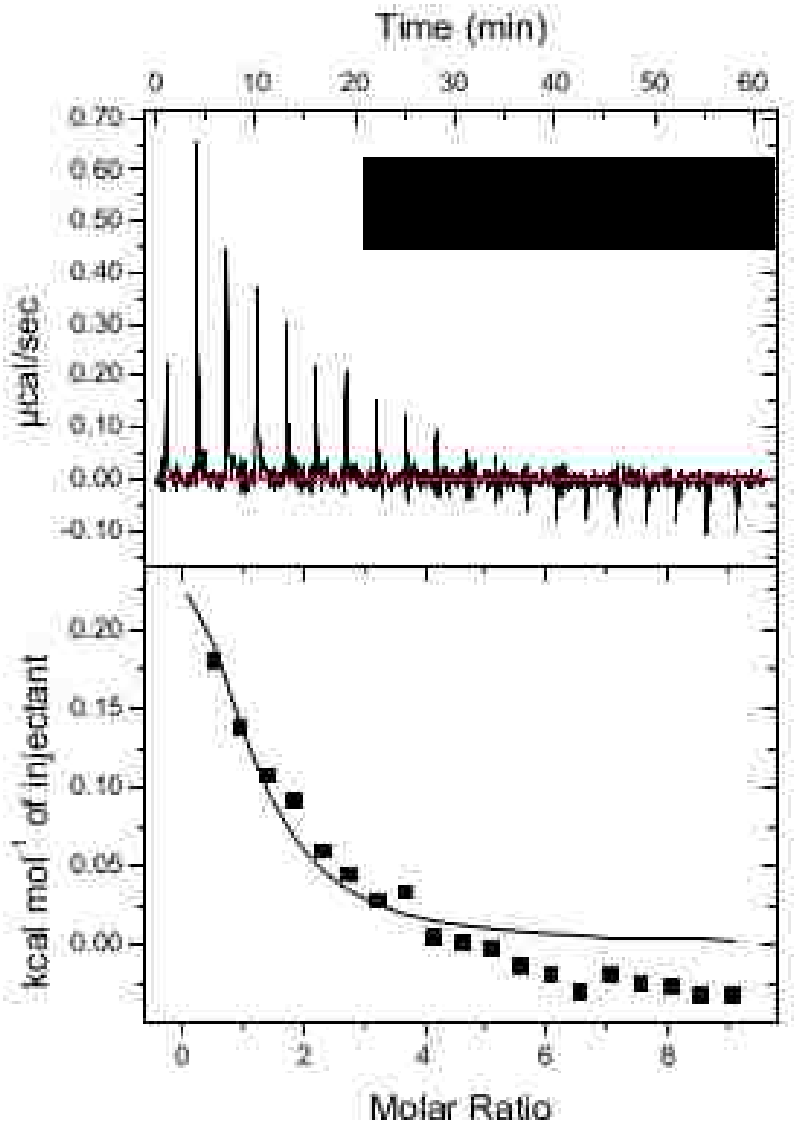

Figure 29. Isothermal titration calorimetry (ITC) titration with PRORP2 (0.45 mM) and Mg^{2+} (20 mM) at 22 ºC (With fixed *n*=1 value).

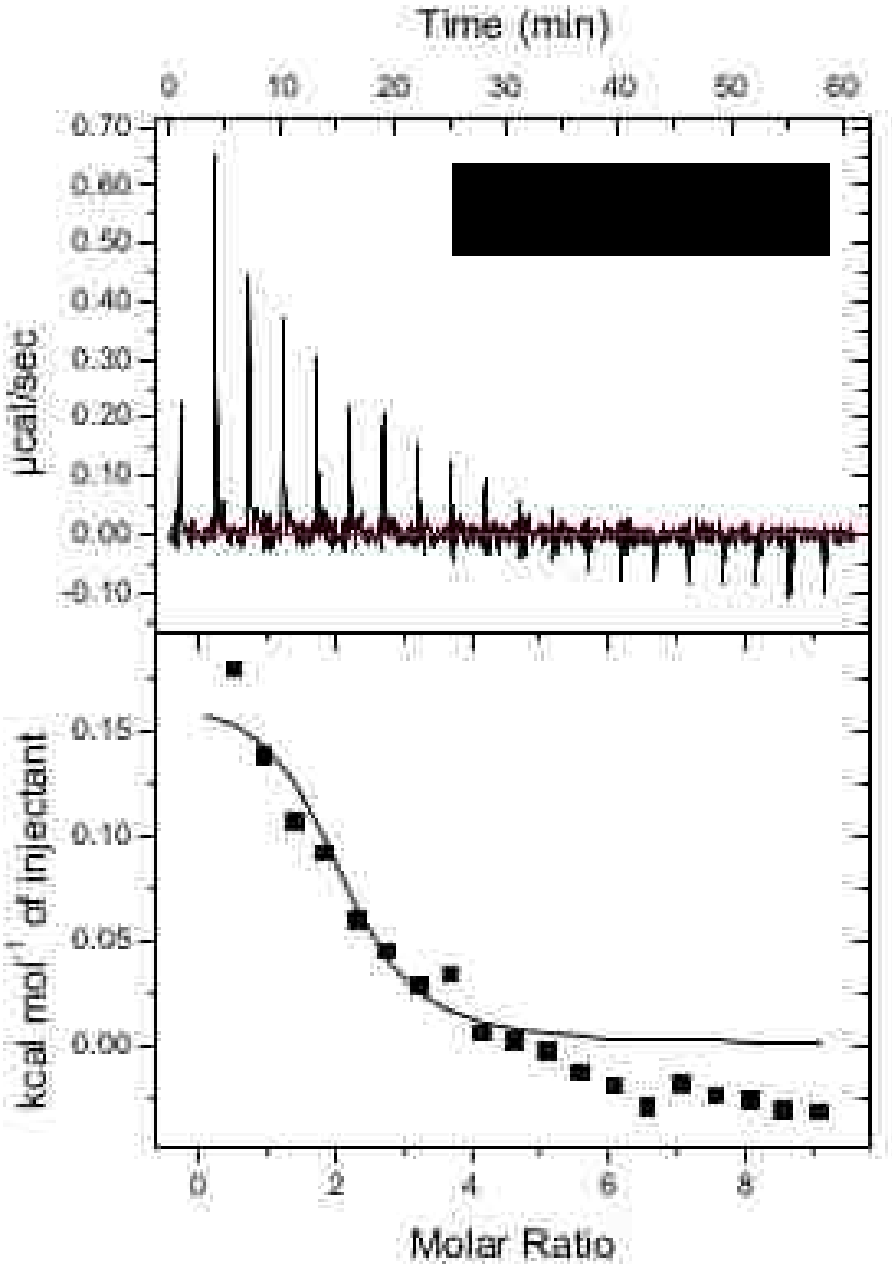

Figure 30. Isothermal titration calorimetry (ITC) titration with PRORP2 (0.45 mM) and Mg^{2+} (20 mM) at 22 ºC (With fixed *n*=2 value).

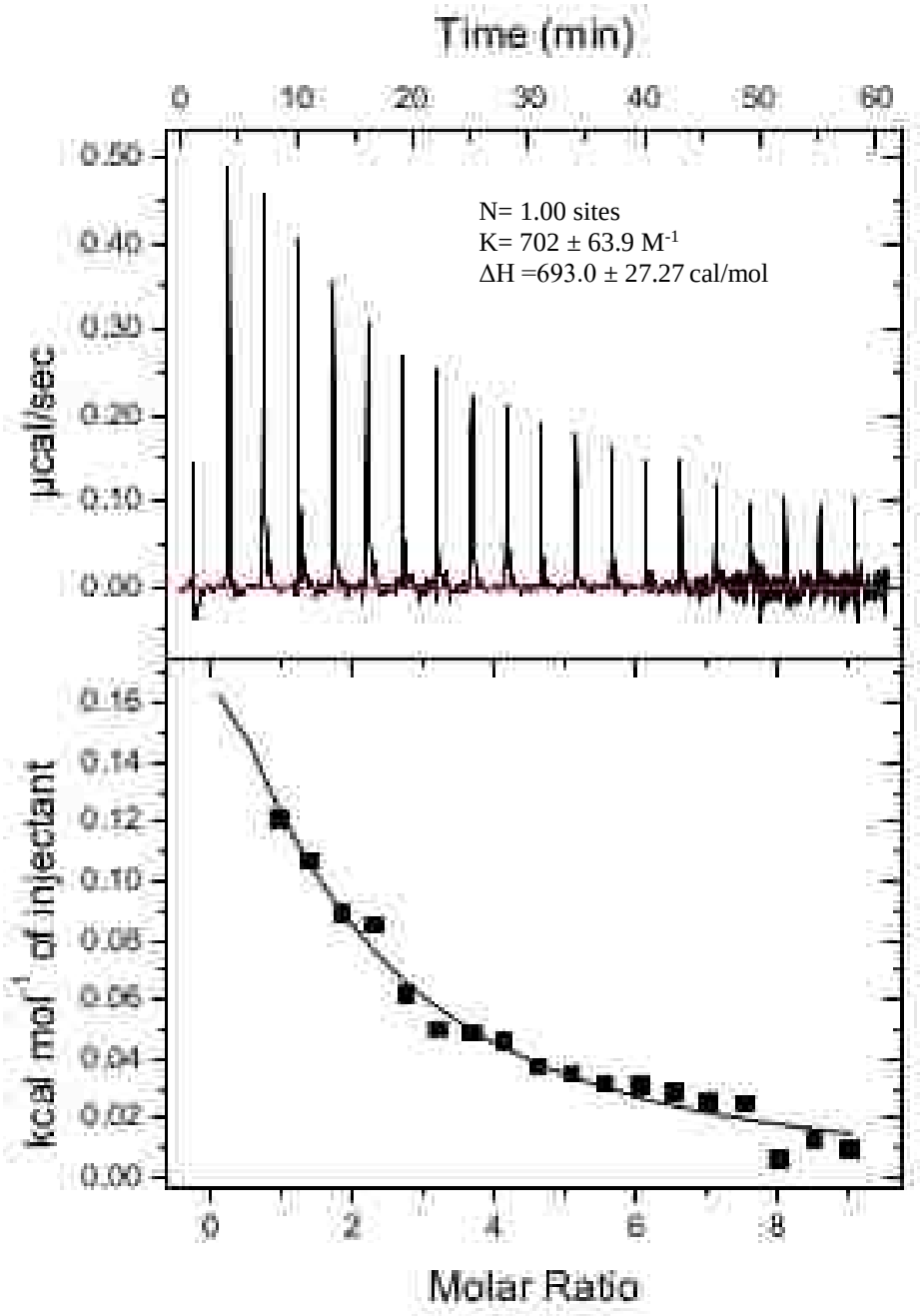

Figure 31. Isothermal titration calorimetry (ITC) titration with PRORP2 (0.45 mM) and Mn^{2+} (20 mM) at 22 ºC (With fixed *n*=1 value).

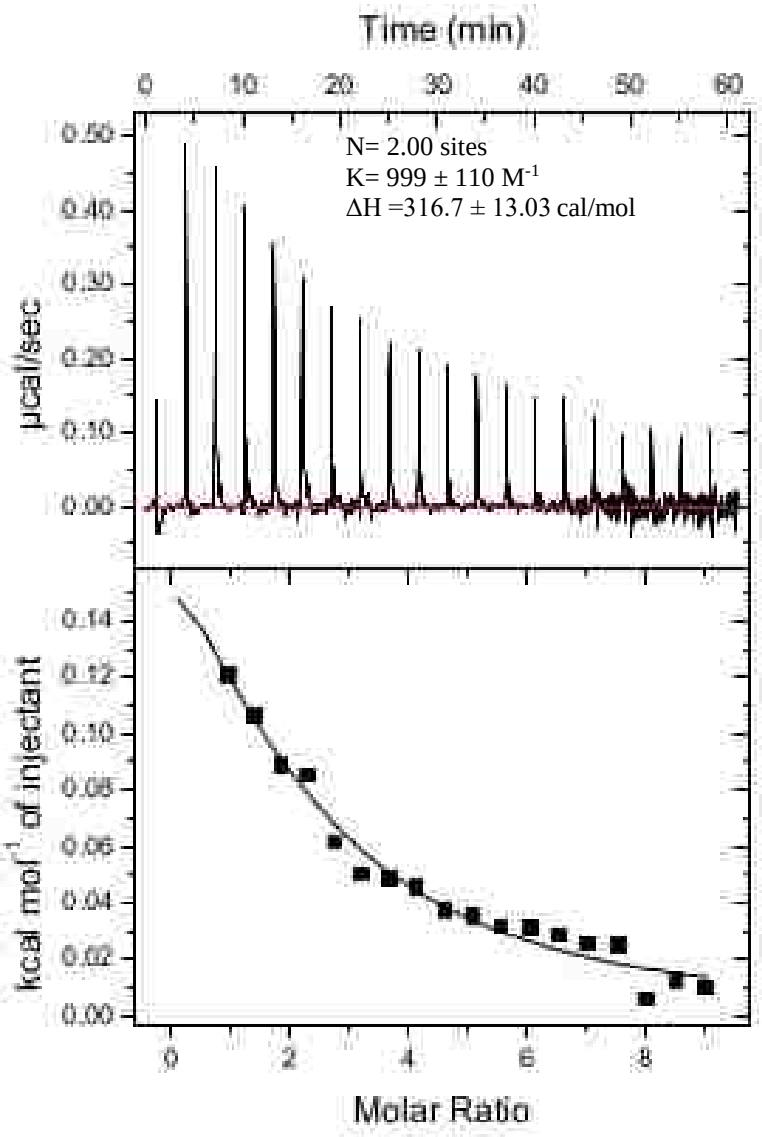

Figure 32. Isothermal titration calorimetry (ITC) titration with PRORP2 (0.45 mM) and Mn^{2+} (20 mM) at 22 ºC (With fixed *n*=2 value).

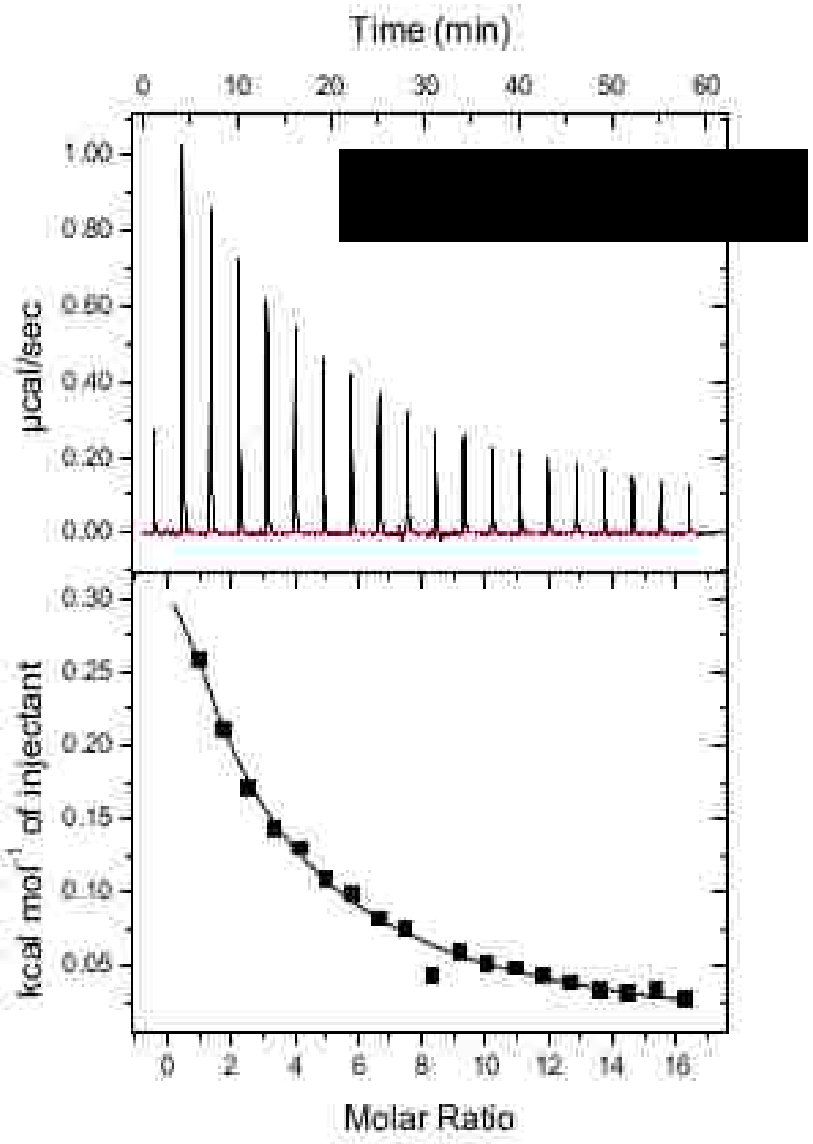

Figure 33. Isothermal titration calorimetry (ITC) titration with PRORP3 (0.25 mM) and Mg^{2+} (20 mM) at 22 ºC (Without fixing the *n* value).

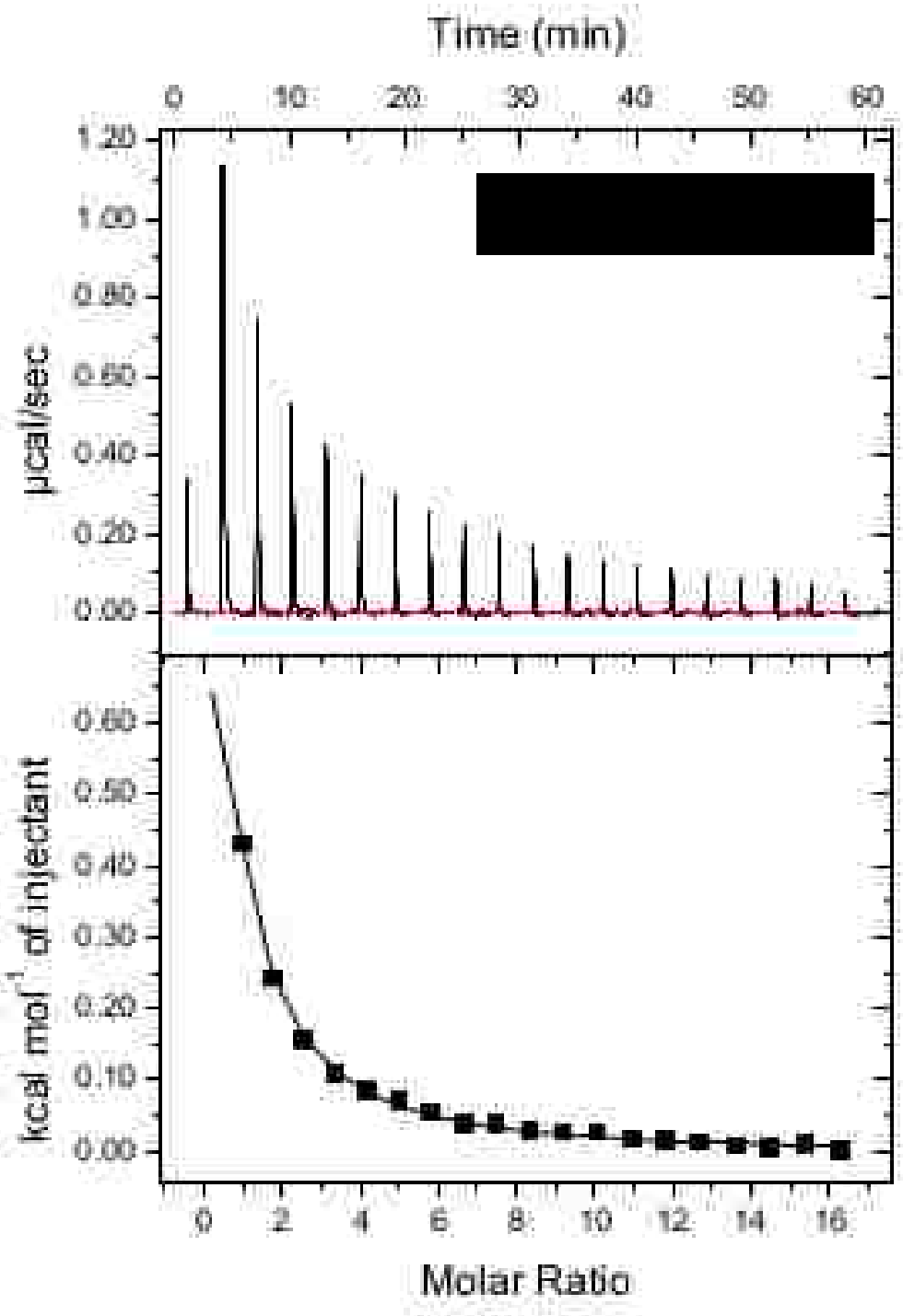

Figure 34. Isothermal titration calorimetry (ITC) titration with PRORP3 (0.25 mM) and Mn^{2+} (20 mM) at 22 ºC (Without fixing the *n* value)

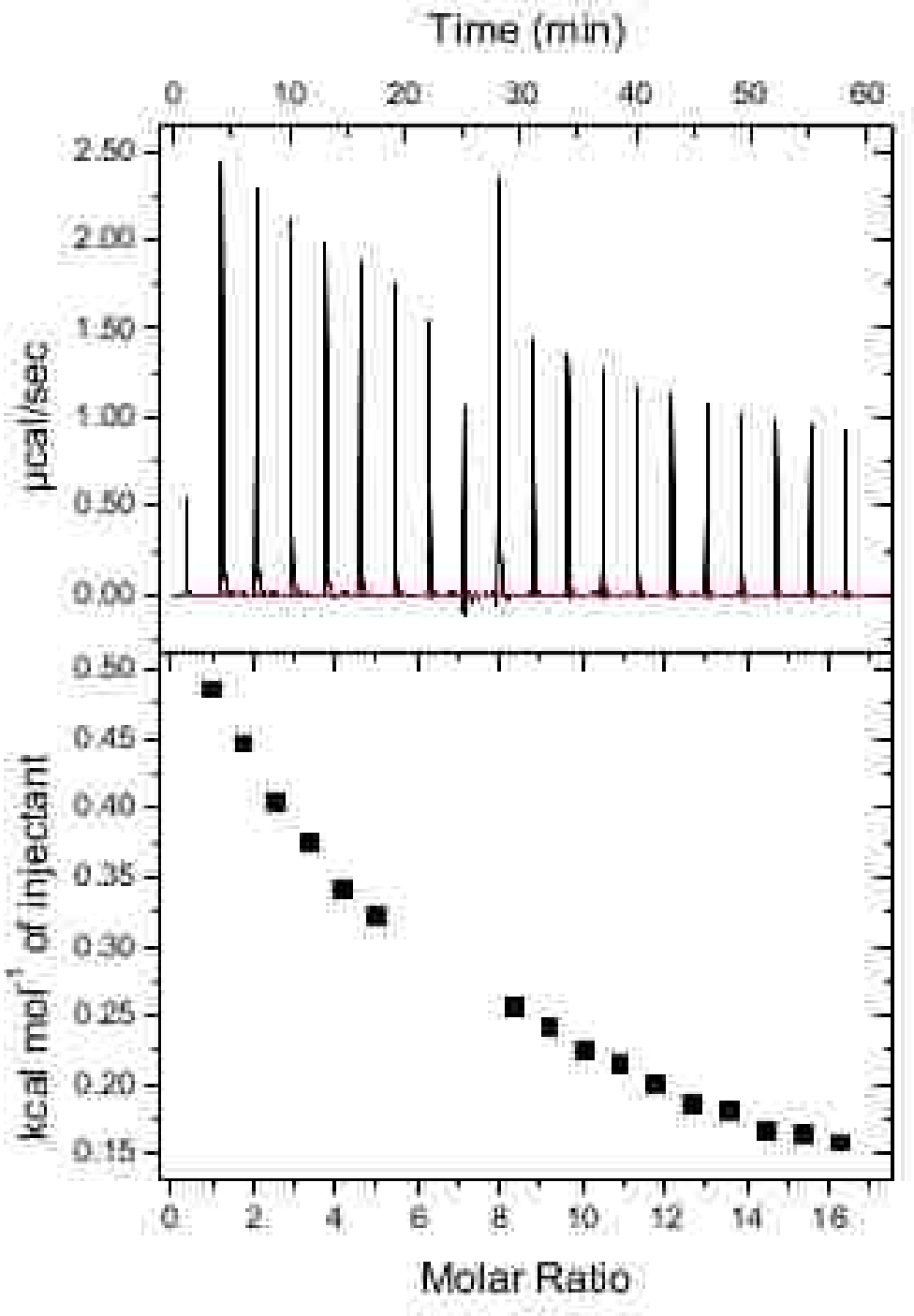

Figure 35. Isothermal titration calorimetry (ITC) titration with PRORP3-D422N (0.25 mM) and Mg^{2+} (20 mM) at 22 ºC (Without fixing the *n* value).

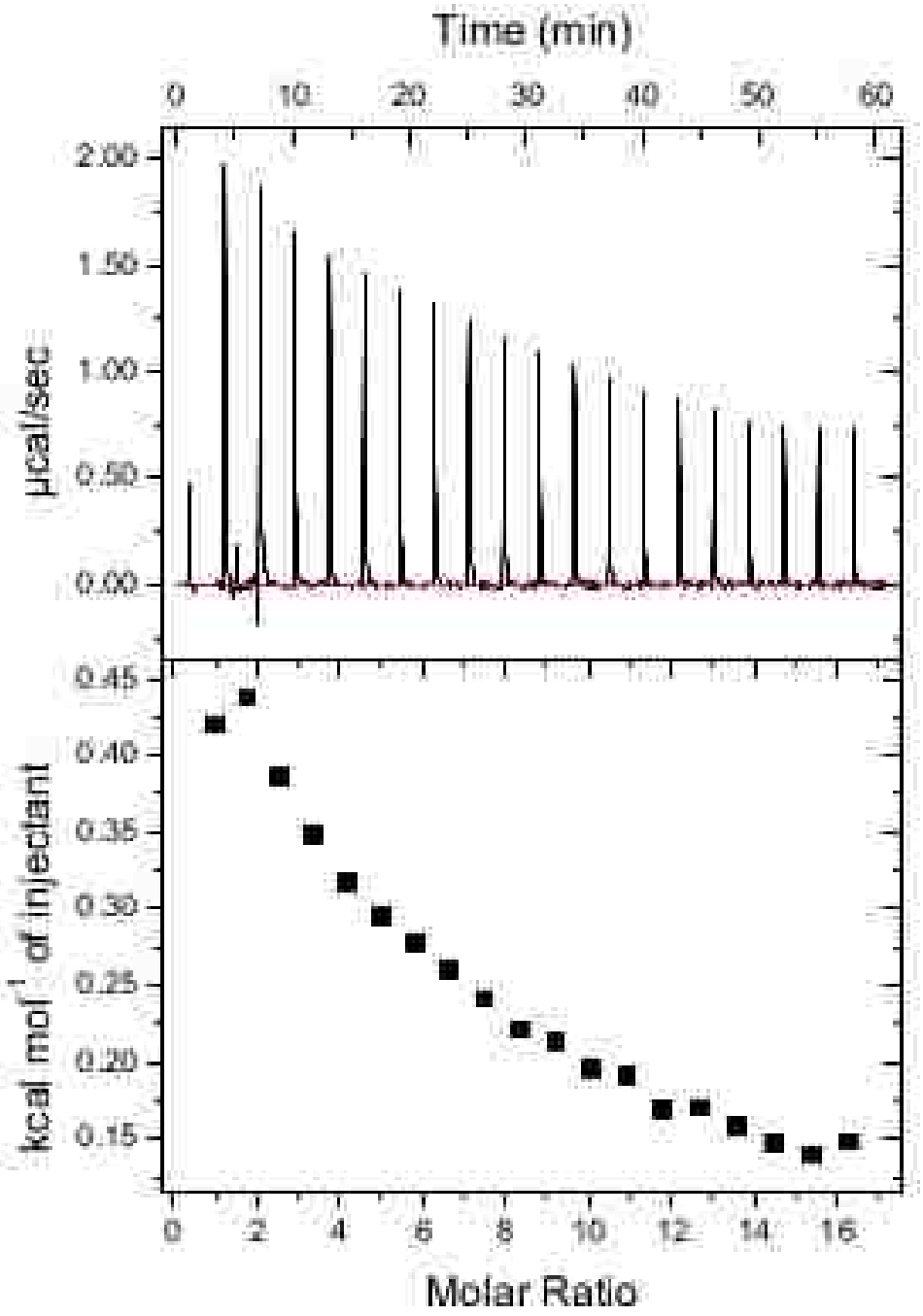

Figure 36. Isothermal titration calorimetry (ITC) titration with PRORP3-D422N (0.25 mM) and Mn^{2+} (20 mM) at 22 ºC (Without fixing the *n* value)

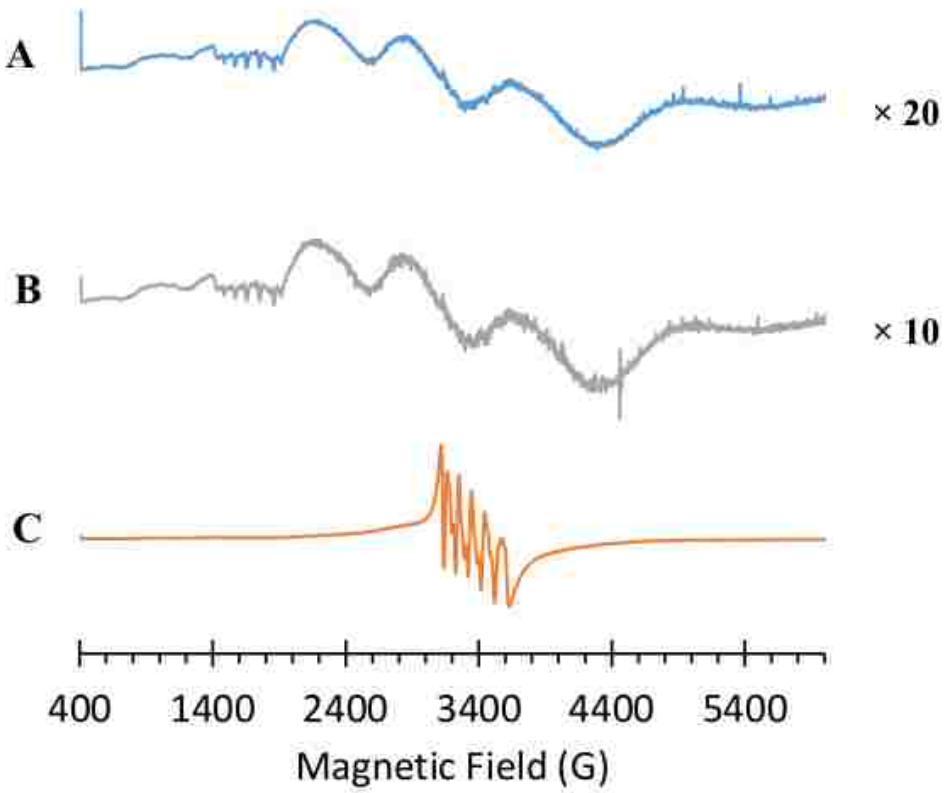

Figure 37. Electron paramagnetic resonance spectra at 77 K. All the samples are in HEPES buffer pH 7.5 (50 mM HEPES, 200 mM NaCl, 1 mM TCEP, 5 % glycerol) (A) 5mM EDTA with 1 mM Mn^{2+} (B) 10 mM EDTA with 2 mM Mn^{2+} (C) 0.2 mM EDTA with 2 mM Mn^{2+}.

Chapter 3 Appendix

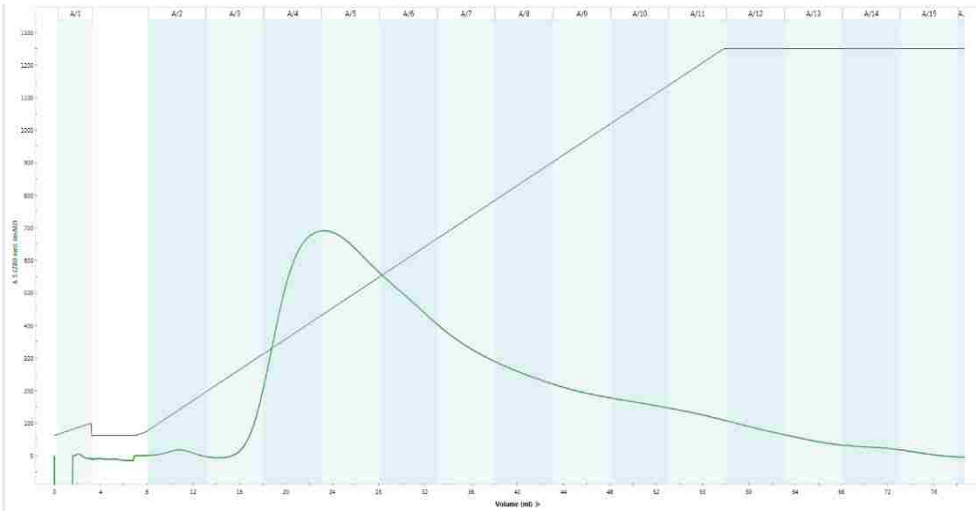

Figure 38. Elution profile of *Aquifex aeolicus* RNase P protein by gradient elution from Nickel affinity column (50 mM HEPES, 250 mM NaCl, 10% BME, 250 mM Imidazole, 5% glycerol).

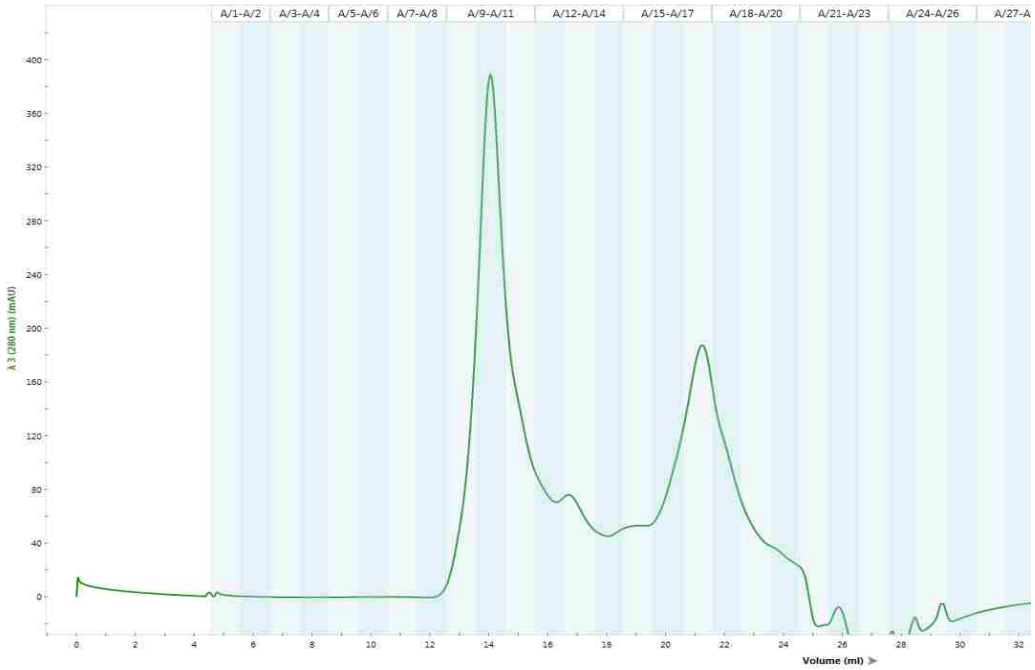

Figure 39. Elution profile of *Aquifex aeolicus* RNase P protein using buffer (50 mM Tris 8.0, 200 mM NaCl, 2 mM TCEP, 5 % Glycerol) by size exclusion column (flow rate 0.25 mL/min, 0.7 mL fractions).

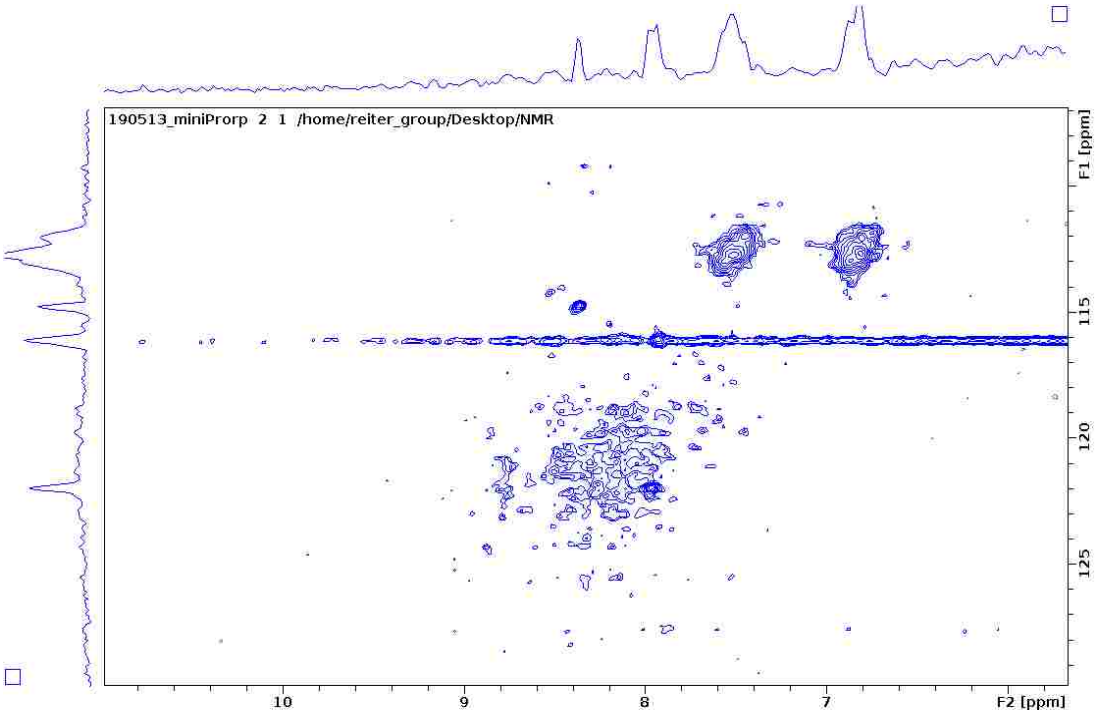

Figure 40. 2D ^{1}H-^{15}N HSQC NMR spectrum collected at 600 MHz with a cryogenic probe (Bruker Avance III) of purified ^{15}N-labled *Aquifex aeolicus* RNase P protein.