Study on the mechanism of substrate recognition by proteinaceous ribonuclease Ps in Arabidopsis thaliana

今井, 崇喜

出版情報：九州大学, 2015, 博士（農学）, 課程博士
バージョン：

Study on the mechanism of substrate recognition by proteinaceous ribonuclease Ps in *Arabidopsis thaliana*

Takayoshi Imai

Laboratory of Biochemistry
Department of Bioscience and Biotechnology
Faculty of Agriculture, Graduate School
Kyushu University, Japan

2016

CONTENTS

Acknowledgements

I deeply express my sincere gratitude to my research supervisor, Professor **Makoto Kimura**, for his patience, encouragement, and constant guidance throughout this project.

I am very much obligated to Professor **Yoshimitsu Kakuta (**Institute of Biophysics, Kyushu University) for his helpful discussions and critical comments on the thesis specially X-ray crystallographic study.

My appreciated to Assistant Professor **Takashi Nakashima** for his remarkable help and suggestions, lessons for usage, and maintenance work.

I wish to express my gratitude to Associated Professor **Takahiro Nakamura** (Laboratory of Plant Biosciences, Kyushu University) for his valuable discussion and critical reading this thesis.

I offer my thanks to my co-workers Ms. **Kaoru Nakayama**, Mr. **Taku Maeda**, Mr. **Xuzhu Gao**, and Ms. **Hisako Tanigawa** for their technical helps.

Finally, I am extremely thankful to all members in the laboratory of Biochemistry for encouragement, help and support.

Abbreviations

A280: absorbance at 280 nm

AU: arbitrary unit

C-domain: catalytic domain

CD: circular dichroism

DEAE: diethylaminoethyl

DTT: dithiothreitol

EDTA: ethylenediaminetetraacetic acid

Exo1: exonuclease 1

HAT: half-a-tetratricopeptide repeat

IPTG: isopropyl-β-D-thiogalactopyranoside

LB: Luria-Bertani

MR: molecular replacement

MRPP: mitochondrial ribonuclease P protein

*M*r: relative molecular mass

NSRs: nucleotides specifying residues

NYN domain: Nedd4 binding protein1, Yac P nucleases domain

Nedd4 - BP1: Nedd4 binding protein1

OD: optical density

OPR: octotricopeptide repeat

PAGE: polyacrylamide gel electrophoresis

PCR: polymerase chain reaction

PDB: protein data bank

PEG: polyethylene glycol

PIN: PilT N terminus

PMSF: phenylmethylsulfonyl fluoride

PPR: pentatricopeptide repeat

PRORP: proteinaceous ribonuclease P

PUF domain: pumilio and FBF homology domain

PUM1: pumilio RNA binding family member 1

RISC: RNA induced silencing complex

RNP: ribonucleoprotein

RNase: ribonuclease

S-domain: specificity domain

SDS: sodium dodecyl sulfate

THA8: thylakoid assembly 8

UV: ultra violet

VS: varkud satellite

cDNA: complementary DNA

glmS: glucosamine-6-phospate synthase

kDa: kilo dalton

kbp: kilo base pairs

mRNA: messenger RNA

mTERF: mitochondrial transcription termination factors

ncRNA: non-coding RNA

pre-tRNA: precursor tRNA

rRNA: ribosomal RNA

tRNA: transfer RNA

Chapter Ⅰ

General Introduction

Transfer RNAs (tRNAs) are known to be the adaptors, which play an essential role in converting genetic information to amino acids information, at the time when translation occurs on ribosomes (Leder and Nirenberg, 1964). Although the number of assembly factors and the size of ribosomes are different between prokaryote and eukaryote, structural and functional features of tRNAs are highly conserved in all living cells (Ban *et al*., 2000). The length of tRNA genes are almost 80 nt, and they contain four stems (known and described as D stem, TΨC stem, acceptor stem, and anticodon stem), leading to form the cloverleaf-shaped secondary structure (Holley *et al*., 1965; Robertus *et al*., 1974; Shi and Moore, 2000) (Figure -1). Additionally, their three-dimensional structures form L-shaped structures, caused by a variety of hydrogen bondings between TΨC loop and D loop (Robertus *et al*., 1974; Shi and Moore, 2000) (Figure -1).

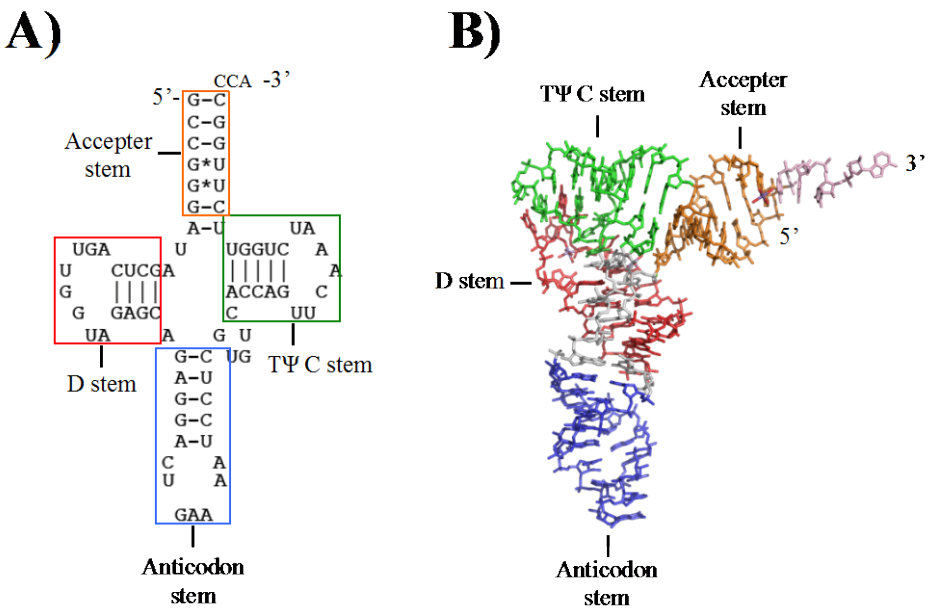

Figure -1: Structure of tRNA. (A) The secondary structure of chloroplast $tRNA^{Phe}$ in *A. thaliana* (Gobert *et al*., 2010). The accepter stem, D stem, anticodon stem, and TψC stem are colored by orange, red, blue, and green, respectively. Watson-Crick base pairs and Hoogsteen base pairs are marked by black lines and asterisks, respectively. (B) The crystal structure of yeast $tRNA^{Phe}$, reported by Shi and Moore, 2010 (Shi and Moore, 2010) (PDB ID: 1EHZ). The structure of accepter stem, TψC stem, D stem, and anticodon stem are colored by orange, green, red, and blue, respectively. The crystal structure was illustrated by Pymol (www. Pymol.org)

RNA molecules, including tRNA, ribosomal tRNA (rRNA), messenger RNA (mRNA), and non-coding RNA (ncRNA), are transcribed by RNA polymerases using their corresponding DNA molecules as templates. However, the produced primary transcripts do not exhibit their functions without post-transcriptional modification (*e.g.* splicing, RNA processing, adding other RNA sequences, and RNA editing) (Nishikura, 2010; Guerrier-Takada *et al.*, 1983; Fica *et al.*, 2013; Schafield *et al.*, 1977). Chemical modification of tRNAs, including pseudouridylation and dihydrouridylation, are involved in stabilization of tRNA structure and improving the affinities between codon and anticodon sequences located on tRNAs and its cognate mRNAs, respectively (Kersten, 1983). Recently studies also indicated that the modification improves the accuracy of amino acylation (Estella *et al.*, 2008).

In the case of RNA processing, a certain number of tRNA molecules in *E. coli* are transcribed as the single RNA sequence (Lander *et al.*, 2001). Furthermore, some of the RNA transcripts encode rRNA genes as well as tRNA genes (Morgan *et al.*, 1978). Thus, it is essential steps of cleaving the 5' and 3' leader sequences of each tRNA from the single RNA sequences. Previous studies showed the order of events in tRNA maturation (Kunzmann *et al.*, 1998). First, ribonuclease P (RNase P) hydrolyses tRNA precursors at the 5' terminus of tRNA precursors (pre-tRNAs), leading to produce the metabolic intermediates, which contain tRNA sequences and their 3' extra sequences (Guerrier-Takada *et al*, 1983). Second, ribonuclease Z (RNase Z) cleaves their 3' sequences (Möri and Marchfeider, 2001). Finally, CCA adding enzymes add the CCA sequences to the produced tRNA molecules in eukaryote; in prokaryote, the CCA sequences are already encoded in large number of tRNAs (Hou *et al.*, 2010). In addition, pre-tRNAs contain introns; in bacteria, the introns are cleaved by self-splicing, whereas in archaea and eukaryote, the extra sequences are removed by tRNA splicing enzymes (Trotta *et al.*, 1997; Westaway, 1988).

RNase P is a ribonucleoprotein complex that catalyzes the processing of 5' leader sequences from pre-tRNAs and other non-coding RNAs in all domain of life (Walker and Engelke, 2006; Torres-Larios *et al.*, 2006). Ever since Altman and co-workers discovered that the *E. coli* RNase P RNA (M1 RNA) is a ribozyme (Guerrier-Takada *et al.*, 1983), biochemical and structural studies have focused mainly on eubacterial RNase P RNAs and a large amount of information on structure-function relashionships has

become available for eubacterial RNase Ps (Kazantsev *et al*., 2006; Esakova *et al*., 2010). Recently, the X-ray structure of *Thermotoga maritima* RNase P in complex with tRNA was determined, and a structural basis for the catalytic activity of eubacterial RNase P has been proposed (Reiter *et al*., 2010) (Figure 1-2). In addition, information on the structure-function relationships of archaeal and eukaryotic RNase Ps has begun to emerge (Jarrous and Goparan, 2010; Kimura and Kakuta, 2012). Despite rigorous investigations, the presence of RNase P in plant cells and organelles in eukaryotic cells had remained largely unknown until the discovery of a proteinaceous enzyme in human mitochondria in 2008 (Holzmann *et al*., 2008).

Holzmann *et al*., using a combinational purification/proteomics approach, revealed that the protein complex comprising MRPP1, MRPP2, and MRPP3 can process the 5'-leader sequence of pre-tRNA in human mitochondria (Holzmann *et al*., 2008). This finding demonstrated that human mitochondrial RNase P is a proteinaceous enzyme that does not require a *trans*-acting RNA component for catalysis.

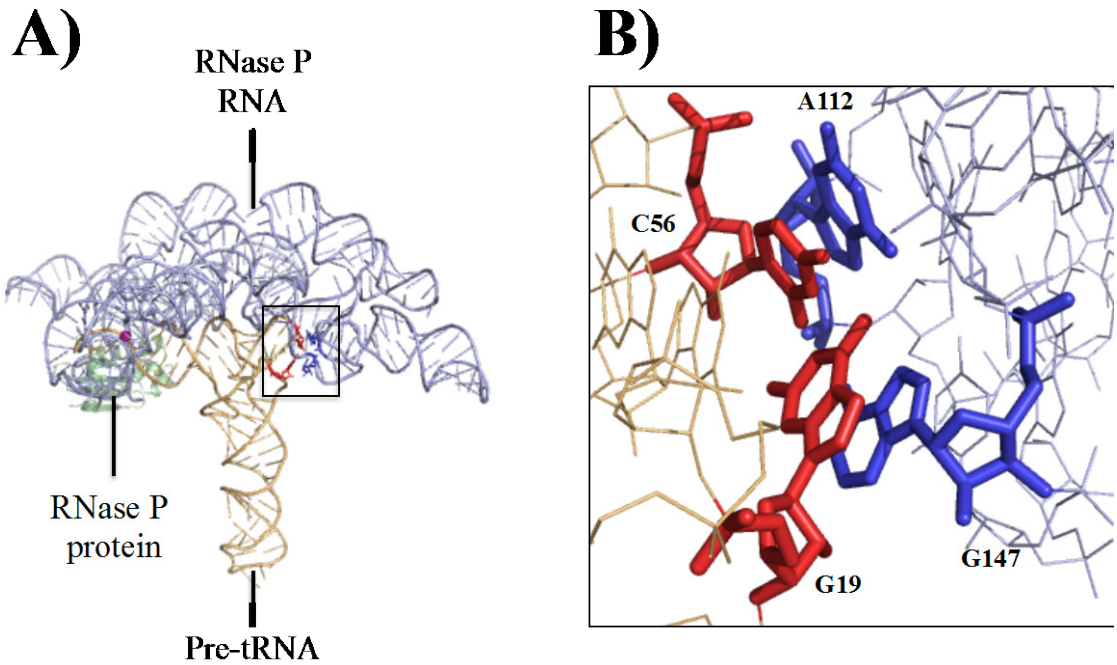

Figure 1-2: The crystal structure of *T. maritima* RNase P in complex with tRNA (PDB ID: 3QIR). (A) Overall structure of the RNase P/pre-tRNA complex from *T. maritima*. RNase P RNA, RNase P protein, pre-tRNA are shown in blue, green, and orange, respectively. The coordinated manganese ion are depicted as a magenta sphere. (B) The recognition site of *T. martitima* RNase P. A112 and G147 (*blue*) in the S-domain of RNase P interacts with G19 and C56 (*red*) in the D and TΨC loops, respectively.

Subsequently, Gobert *et al*. found that *Arabidopsis thaliana* has three homologues, PRORP1, PRORP2, and PRORP3, of MRPP3, and that PRORP1 is localized in mitochondria and chloroplasts, while PRORP2 and PRORP3 are in the nucleus (Gobert *et al*., 2010; Gutmann *et al*., 2012). It was further found that PRORP1 could perform the endonucleolytic maturation of pre-tRNA and could complement RNase P activity in *E. coli* cells (Gobert *et al*., 2010). Recently, the crystal structure of PRORP1 was determined at a resolution of 1.75 Å (Howard *et al*., 2013) (Figure I-3). PRORP1 comprises five tandem pentatricopeptide repeat (PPR) motifs, a central linker domain, and a metallonuclease domain belonging to the NYN family (Anantharaman *et al*., 2006). Furthermore, cross-liking and mutational analyses suggested that G19 and C56 in tRNA are involved in the interaction with PRORP1 (Gobert *et al*. 2013; Pinker *et al*. 2013). Even though the crystal structure of PRORP1 was elucidated, the mechanism of substrate recognition by PRORPs remains unknown.

Ever since the discovery of bacterial ribonuclease P by Altman and coworkers in 1983, it is generally accepted that current life on the earth emerged from RNA molecules (known and described as RNA world hypothesis) (Figure I-4). Hoogstraten and Sumita collected some biological catalysts and categorized them with structural components as described in Table I-5. Based on the RNA world hypothesis, they postulated that the biological catalysts have been evolved from pure RNA-based enzymes, ribonucleoprotein complexes with RNA active site, ribonucleoprotein complexes with protein active site, protein enzyme with nucleotides-bases cofactors, to pure polypeptide catalysts (Hoogstraten and Sumita, 2007) (Table I-1). However, the fascinating hypothesis has not yet been verified, because of the lack of suitable ribozymes and enzymes that catalyze an identical reaction. Thus, the detailed comparison of RNase P and PRORP1 has been expected to shed light on fundamental issues in the evolution of biological catalysis.

In order to elucidate the mechanisms of substrate recognition by PRORPs and also to gain insight into molecular evolution of biocatalysts, I prepared PRORP1, PRORP2, and PRORP3 from *A. thaliana* and characterized their enzymatic properties by utilizing pre-tRNA processing activity assay. Furthermore, I obtained the N-terminal successive deletion mutants of PRORP1 and characterized them in terms of pre-tRNA cleavage activity and its binding activity.

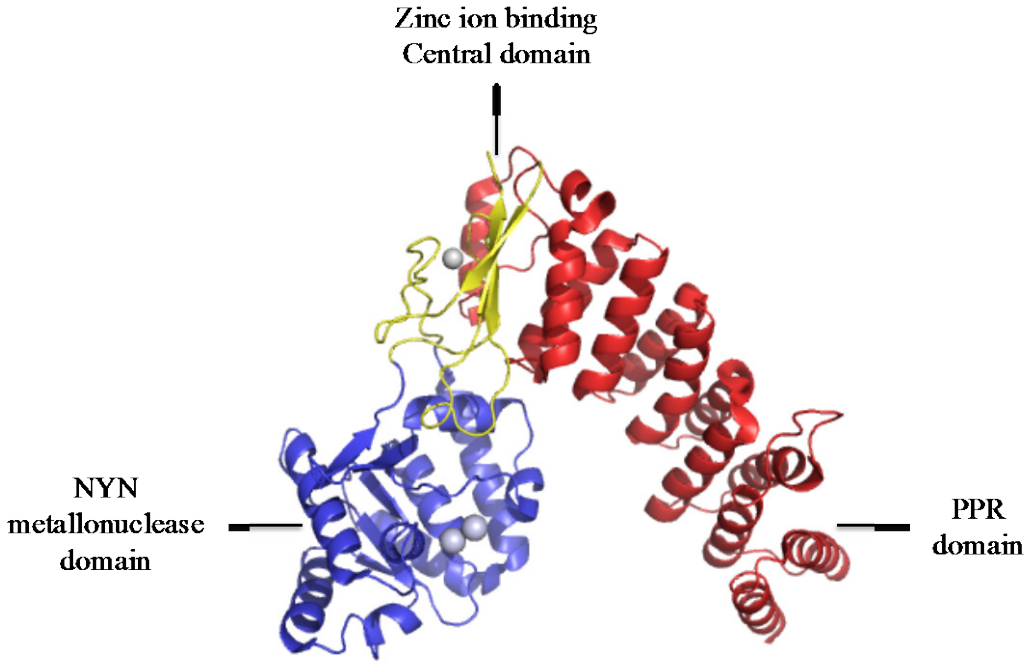

Figure I-3: The crystal structure of PRORP1 from *A. thaliana*. Ribbon diagram of PRORP1 des76 mutant. A PPR domain (pos. 95–327), a zinc-finger like central domain (pos. 328–354; pos. 534-570), and a NYN metallonuclease domain (pos. 355-533) are shown in red, yellow, and blue, respectively. The coordinated Zn^{2+} ion located on the central domain is depicted as a white sphere. Likewise, the two coordinated Mn^{2+} ions located on the NYN domain are depicted as blue spheres.

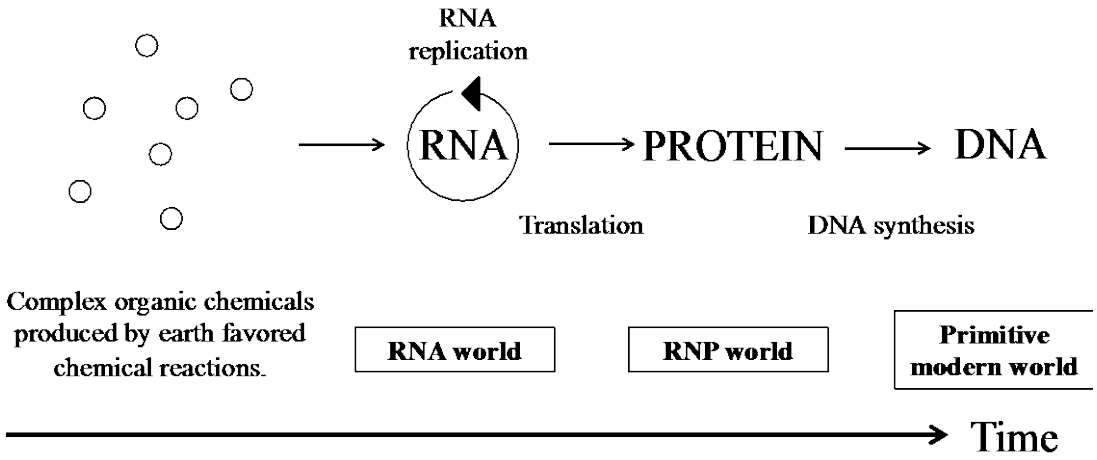

Figure I-4: An RNA world model for the consecutive appearance of RNA, proteins, and DNA during the evolution of life on earth. (This figure was referred to the figure 1 in Cech, 2011) Based on the experiments by Urey and Miller in 1953, complex organic materials were made from the in-organic chemicals by applying an earth favored chemical reactions (Urey and Miller, 1953). A part of produced material had the capacity of self-replication and becomes primordial RNA. Some of self-replicating RNAs were available to catalyze peptidyl transferase activity, leading to produce polypeptides. At the present, we can not identify the enzymes composed of DNA moiety. In 1970, Temin and Baltimore found that RNA-dependent DNA polymerase, which is composed of polypeptides, promotes the reverse transcription activity (Temin and Baltimore, 1972). Furthermore, RNA was utilized to synthesis DNA *in vivo* (Shiomi, 2013). Because of the reasons given, it is probably likely that DNA molecules appeared after the emergence of proteins.

Table -1: A Continuums of Ribonucleoprotein Machines

Components	Systems	Comments
RNA machines	Self-cleaving ribozymes [Hammerhead, hairpin, hepatitis delta, varkud satellite (VS), Glucosamine-6-phosphate synthase (glmS)]	RNA alone reacts in multiple turnovers.
RNA machines with protein cofactors	Bacterial RNase P Group □and □ introns	RNA alone reacts *in vitro,* but not *in vivo.* They catalyze in multiple turnover reactions.
RNP machines with RNA active sites	Archaeal and enkaryotic RNase P Ribosome Spliceosome	RNP machines do not react without either RNA or protein moiety.
RNP machines with protein active sites	Telomerase RNA induced silencing complex (RISC)	RNA moiety serves structural, recognition/ template functions.
Protein machines with nucleotides-based cofactors	Pyruvate dehydrogenase Alcohol dehydrogenase, *etc.*	The nucleotide-based cofactors are essential for the reaction.
Pure Protein machines	RNase A Lysozyme Trypsin, *etc.*	The polypeptide alone reacts in multiple turn over reaction.

Hoogstraten and Sumita, *Biopolymers* **87** (2007) 317-328.

In the chapter Ⅱ, the resutls o
pre-tRNAs by PRORP1, I predicted on the basis of the proposed recognition code that the PPR motifs PPR2, PPR3, PPR4, and PPR5 in PRORP1 recognize C, A/G, A, and U, respectively. This prediction and biochemical data are presented in the chapter Ⅲ. Moreover, the existing crystallographic structural information was combined to generate a fitting model of tRNA onto PRORP1. In the chapter Ⅳ, I attempted to structure in complex with pre-tRNA or pre-nad6 to gain more information about the mechanism for RNA recognition of PRORP1. Although the crystals, containing D399A alone or pre-tRNA alone, were produced in many conditions, that of D399A complexed with its RNA substrates was not appeared until now. In addition to the observation that the catalytic pocket of D399A was collapsed by mutation, it is required to change the mutation site, where this mutation does not lead to disrupt its catalytic structure.

Chapter Ⅱ

The Involvement of PPR Motifs on Catalytic Activity of PRORP1

INTRODUCTION

In *A. thaliana*, three PRORPs are encoded within the nuclear genome. PRORP2 and PRORP3 are co-localized to the nucleus, whereas PRORP1 functions within both mitochondria and chloroplasts (Gobert *et al*., 2010; Gutmann *et al*., 2012). In general, PRORPs are composed of three domains: a Nedd4-BP1, Yac P nuclease (NYN) metallonuclease domain, a central structural zinc-ion binding domain, and a pentaricopeptide repeat (PPR) domain involved in RNA-binding (Figure Ⅱ-1). The NYN domain is a novel metallonuclease domain sharing a structural homology to the PIN (PilT N terminus) and flap (FENs, also known as 5' nucleases in older references) nuclease families (Anantharaman and Aravind, 2006) (Figure Ⅱ-1). In general, PPR proteins contribute to RNA metabolism, including RNA editing, splicing of many genes, and mRNA stabilization and are widely distributed in plants (Forner *et al.*, 2007; Gobert *et al*., 2010). PPR proteins are eukaryotic specific and form a large family of RNA binding proteins with more than ~450 members in *A. thaliana* (Gobert *et al.*, 2010). Recently, the crystal structures of the PPR domain in human mitochondrial RNA polymerase and in PRORP1 from *A. thaliana* demonstrated that PPR motif adopts two antiparallel α-helices (Howard *et al.,* 2012; Ringle *et al.,* 2011) (Figure Ⅱ-2). An interesting question is how PPR proteins specifically recognize their target sequences. Several PPR proteins and their endogenous specific RNA recognition sequences have been characterized biochemically (Nakamura *et al*., 2012; Barkan *et al.,* 2012; Yagi *et al.,* 2013a, Yagi *et al.,* 2013b). Kobayashi *et al.* proposed the modular recognition mechanism of RNA by PPR proteins on the basis of results obtained from biochemical and *in silico* bioinformatic studies (Nakamura *et al.,* 2012). In this mechanism, one PPR motif recognizes one RNA base. Moreover, three amino acid residues at position 1 and 4 on one α helix in PPR motif and the position -ii on the other α helix in the motif play crucial roles in recognition of the RNA-base specifically (Manna, 2015; Barkan *et al.,* 2012; Yagi *et al.,* 2013a, Yagi *et al.,* 2013b) (Figure Ⅱ-2).

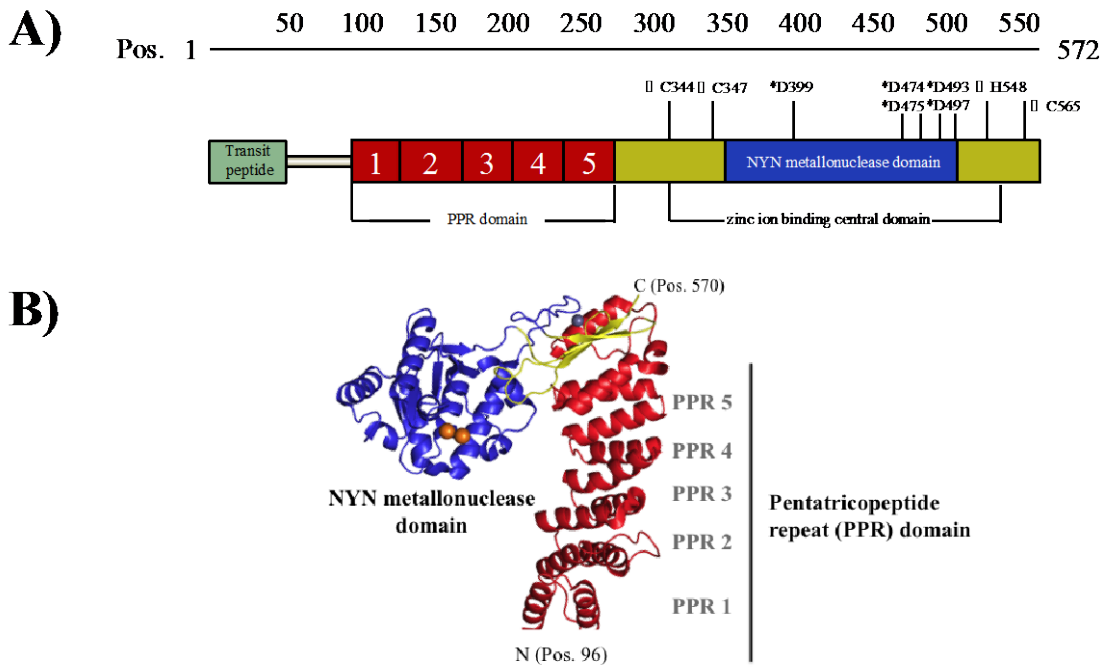

Figure 1: Domain organization of the PRORP1 from *A. thaliana*. (A) PRORP1 contains a transit peptide region (*green*), five PPR motifs (*red*), a central domain (*yellow*), and a NYN metallonuclease domain (*blue*). The linker region (pos. 38-94) is represented by a gray bar. The five catalytic aspartates (pos. 399, 474, 475, 493, and 497) are depicted as asterisks. The cysteines and a histidine at positions 344, 347, 548, and 565, indicated by black arrowheads, binds to one zinc ion to stabilize the catalytic domain of PRORP1. (B) The crystal structure of PRORP1 (pos. 96-570). Its PPR domain, central domain, and NYN metallonuclease domain are colored red, yellow, and blue, respectively, with the same color scheme as in (A). The coordinated Mn^{2+} ions are colored orange and the coordinated Zn^{2+} ion is represented by a gray sphere. (Gobert *et al.,* 2013)

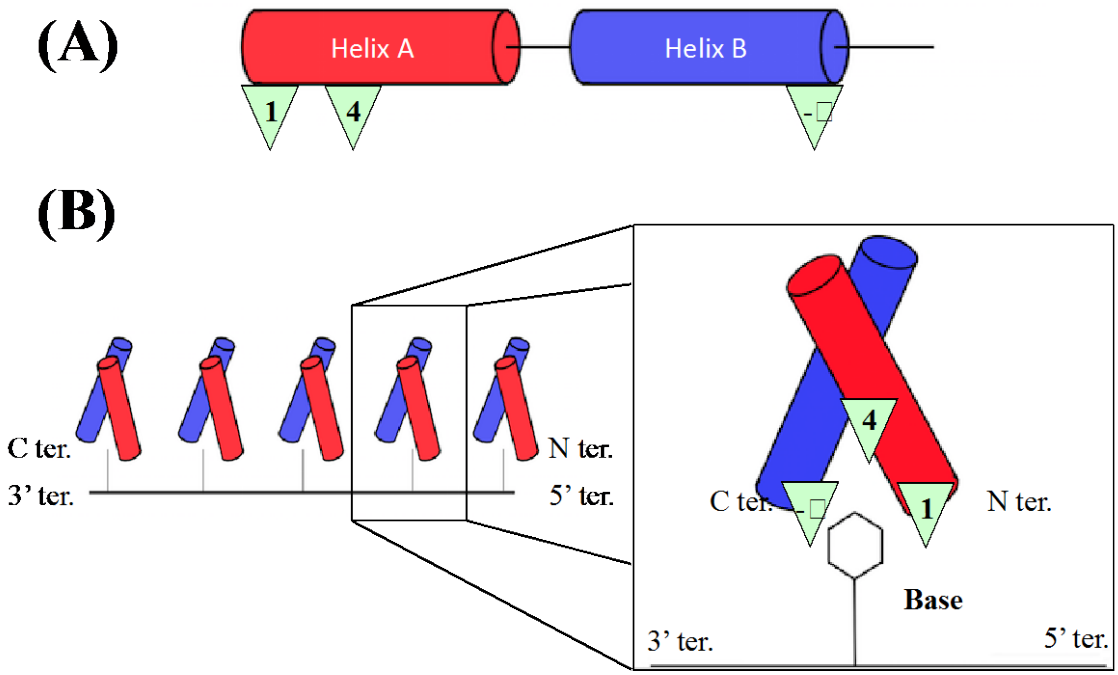

Figure Ⅱ-2: A pentatricopeptide repeat (PPR) motif. (A) The secondary structure arrangement of 35 amino acids in PPR motif. Helix A, helix B and the loop region are shown in red, blue and black, respectively. The nucleotides-specifying residues (NSRs) (known and described as 1, 4 and -ii) are depicted as green reverse-triangles. (B) Schematic representation of the RNA recognition of PPR motifs. Each PPR repeat binds to a single base in a modular fashion.

The crystal structures of PPR10 / *atpH* and THA8 / *ycf3* binary complex validate a proposed combinational amino acid code of RNA recognition (Yin *et al.,* 2013; Ke *et al.,* 2013) (Figure Ⅱ-3). Although structural information has become available for PRORP1, and then the RNA recognition mechanism of PPR proteins has been proposed, the molecular mechanism, by which the PPR motifs in PRORPs recognize substrates, has remained unclear (Nakamura *et al.*, 2012). To address this issue, I overexpressed PRORP1, PRORP2, PRORP3, and their mutants in *E. coli* cells and purified them with several steps of column chloromatographies. They were characterized in terms of the pre-tRNA^{Phe} cleavage activity and the pre-tRNA^{Phe} binding activity.

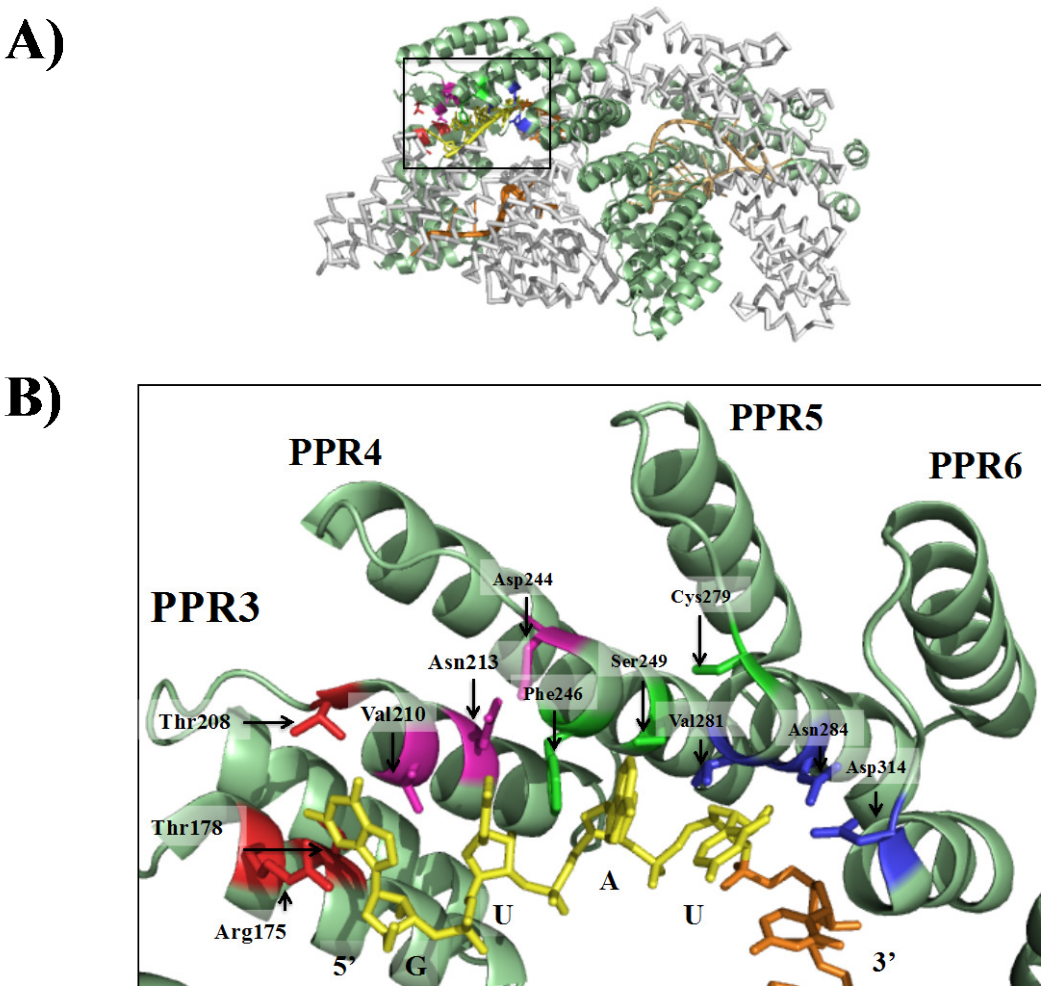

Figure Ⅱ-3: The crystal structure of the maize chloroplast protein PPR10 in complex with the

single-stranded RNA. (A) The overall structure of PPR10/atpH RNA complex from *Zea mays*. One PPR10 monomer is represented by cartoon representation (*green*). Likewise, the other PPR10 monomer is colored white and represented by ribbon representation. The two bound RNA molecules are shown in cartoon representation and colored by orange. (B) Close-up view of the mutual interaction sites between PPR10 and its RNA substrates. The bound RNA is colored yellow and depicted as stick representation. The three residues (pos. 1, 4, and -ii) on each PPR repeats play a crucial role in RNA base recognition and selection (known and described as nucleotide specifying residues (NSR)). The NSRs of PPR3 are colored red and denoted as stick representation. Likewise, those of PPR4, PPR5, and PPR6 are done magenta, green, and blue, respectively. Arg175 at 1, Thr178 at 4, and Thr208 at -ii on its PPR3 recognize guanine. Val210 at 1, Asn213 at 4, and Asp244 at -ii on its PPR4 recognize uracil. Phe246 at 1, Ser249 at 4, and Cys279 at -ii on its PPR5 recognize adenine. Val281 at 1, Asn284 at 4, and Asp314 at -ii on its PPR6 recognize uracil. (These figures were referred by Figure 3 in Yin *et al*., 2013).

MATERIALS AND METHODS

Materials

The full-length cDNAs encoding PRORP1, PRORP2, and PRORP3 (Accession ID, AT4G32230, AT2G16650, and AT4G21900, respectively) were purchased from the RIKEN Bio Resource Center (Tsukuba, Japan). DNA polymerase and DNA ligation kit were obtained from MBI Fermentas (Waltham, Massachusetts, United States of America) and Takara Bio (Prefecture of Shiga, Japan), respectively. Restriction enzymes were purchased from MBI Fermentas. All other common chemicals and reagents were purchased at the highest purity available.

Bacterial strains and plasmids

E. coli strain JM109 (Genotypes: *end*A1, *rec*A1, *gyr*A96, *thi*, *hsd*R17(r_k^-, m_k^+), *rel*A1, *sup*E44, Δ(*lac* - *pro*AB), 〔F' *tra*D36, *pro*AB, laql^qZΔM15〕) was used form routine transformations and plasmid preparation. *E. coli* strain BL21 (DE3) Codon Plus RIL purchased from Agilent Technologies (Tokyo, Japan) containing an IPTG-inducible gene for T7 RNA polymerase was used as a host cell producing the recombinant enzymes. Expression plasmid pET-15b obtained from Merck KGaK (Darmstadt, Germany).

Agarose gel electrophoresis

The desired amount of electrophoresis-grade agarose was added to a volume of TAE buffer (40 mM Tris, 40 mM acetic acid, and 1 μM EDTA, pH 8.0) sufficient for constructing the gel in an erlenmeyer flask. The agarose was solubilized in a microwave oven and swirled to ensure even mixing. The agarose solution was cooled to about 55 ℃

and then poured gently into the holding chamber designed for the gel electrophoresis apparatus. After the gel hardened, it was mounted in the gel electrophoresis tank. Samples of DNA were prepared by addition of the appropriate amount of 10 × loading buffer, containing 0.5% bromophenol blue, 0.5% xylene cyanol, and 50% ficol. The prepared DNA samples were slowly loaded into the wells and a voltage of 100 V was subjected to electrophoresis. After electrophoresis, the gel was stained by ethidium bromide for 10-15 min. The stained DNA bands were detected using a UV trans-illuminator.

Construction of expression plasmids

The amplified PCR products were separated on 1% agarose gel. The target DNAs were extracted from the gel, after which the purified genes were ligated onto pET-15b vector. The ligated plasmids were introduced into *E. coli* strain JM109. The transformants were selected by colony-directed PCR method. After conformation of the nucleotide sequencing, the plasmids were used to transform the *E. coli* strain BL21 (DE3) codon Plus RIL. The transformant harboring expression plasmid were grown to the optical density at 600 nm (OD600) of 0.6 in LB-medium containing 50 μg/ml of ampicillin and stored in 20% glycerol at -80 ℃.

Overexpression

Overexpression was done by the standard method described by Gobert *et al.*, (2010) (Gobert *et al.,* 2010). The *E. coli* cells BL21 (DE3) codon Plus RIL, bearing the expression vector, were grown in 40 ml of LB-medium in the presence of 50 μg/ml of ampicillin to an OD600 of 0.6-1.0 and stored at 4 ℃overnight. The harvested cells were used to inoculate in 2 liter of the same medium at 37 ℃. The cells were induced by the addition of IPTG to a final concentration of 1 mM when the OD600 of the culture reached 0.6-0.8. The culture was further incubated for 18 hours at 16 ℃. The overexpression of the proteins was confirmed by SDS-PAGE (Laemmli, 1970).

Purification of the recombinant proteins

All procedures were carried out at 4 ℃. Cultured cells were harvested, suspended to wash the cells in 40 ml of sonication buffer (30 mM Tris-HCl, pH 7.5, 100 mM NaCl,

1 mM DTT, 10% glycerol, and 5 mM $MgCl_2$), and then centrifigated at 6,000 rpm for 10 min. After washing, the pellet of *E. coli* cells was suspended again in 100 ml of sonication buffer and the cells were completely disrupted by sonication on ice at output 1 for 5 seconds followed by 20 seconds cooling for a total of 15 min. The total extract was centrifigated at 20,000 g for 30 min, and the supernatant was first put onto a SP-sepharose Fast Flow column (1.5 × 10 cm) obtained from GE Healthcare (Little Chalfont, United Kingdom) previously equilibrated with the same buffer. The recombinant protein was elutied with a linear gradient of NaCl up to 1 M in the same buffer, using gradient volume of 200 ml. The fractions being rich of PRORP1 des38, where the N-terminal region (pos. 1-38) was truncated, were evaluated by SDS-PAGE. Subsequently, the interested fractions were pooled and dialyzed against 30 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10% glycerol, 1 mM DTT, 5 mM $MgCl_2$, and 1.5 M ammonium sulfate. The protein was further purified with a RESOURCE ISO column (GE Healthcare) previously equilibrated with the same buffer. The absorbed protein was eluted with a linear gradient of ammonium sulfate from 1.5 M to 0 M. The purified protein was extensively dialyzed against sonication buffer. The homogeneity of PRORP1 des38 was confirmed by SDS-PAGE using 10% polyacrylamide gels. PRORP2 desN9 and PRORP3 desN10 were purified in the same manner as PRORP1 des38, except for using a COSMO GEL His Accept column purchased from Nakarai Tesuque (Kyoto, Japan) and a Resource Q column purchased from GE Healthcare.

Assay for pre-tRNA cleavage

Cleavage reactions of pre-$tRNA^{Phe}$ were performed with three replicates using 2 μg of transcripts by an increasing amount of the enzymes for 60 min at 37 ℃. The resulting RNA fragments were separated on 15% polyacrylamide denaturing gel in TBE buffer at 50 W for 1h. After electrophoresis, the reaction products were visualized by staining in a 0.1% toluidine blue solution, and then the resulting image was used to obtain values for the processing activity. The cleavage activity was calculated as follows: the quantity of (matured $tRNA^{Phe}$ + leader fragment) × 100 / the quantity of (pre-$tRNA^{Phe}$ + matured $tRNA^{Phe}$ + leader fragment), and the percentage was plotted against the incubation time.

Primers used for mutagenesis

The mutagenesis primers were purchased from Eurofins genomics (Tokyo, Japan). Oligonucleotide primers used in mutagenesis are shown in (Table -Ⅱ). In the sequences of the primers, the mutation sites are indicated by underlines.

Site-directed mutagenesis

The PRORP1 des38 cDNA fragment ligated to pET-15b vector was used as a template for the site-directed mutagenesis. All mutants used in this study were generated using the Prime star Master mix kit supplied by Takara Bio. This procedure employs two mutagenic oligonucleotide primers. The synthesized plasmid containing single-point mutation was introduced into the *E. coli* strain for cloning of plasmid, JM109. The plasmid isolated from the harvested JM109 was verified with DNA sequencing, followed by introducing into the *E. coli* strain, BL21 (DE3) Codon Plus RIL in the same manner as PRORP1 des38. The PROPR2 mutants were constructed in the same manner as those described for PRORP1.

Table □-1: Oligonucleotides used in this study.

Primers	Sequences of forward primers (5'-3')	Sequences of reverse primers (5'-3')
PRORP1 desS8	5'-AAAACATATGCCATTTCGTTTCTACCATACTTCC-3'	5'-AAAAGGATCCTCATCAAGGTGTTTTGGATCTTTTTGCGCATAACC-3'
PRORP2 full length	5'-GGGAATTCCATATGGCTGCTTCTGATCAACACCGG-3'	5'-AAAAGGATCCTCATCACTAAGGAATCTTCCCATTACTC-3'
PRORP2 desN9	5'-CCATATGCGCCGTCACGACGA-3'	5'-TGACGGCGCATATGGCTGCC-3'
PRORP3 desN10	5'-AAAACATATGTCTTCATTGCTCTGCGCCGTCCCGCC-3'	5'-AAAAGGATCCTCATCATGAACTCTGCCTTGTAATGCACATCC-3'
PRORP1 desS8 D399A	5'-GCAGTCATTGCTGGAGCCAACATGGGCTTA-3'	5'-GTTGGCTCCAGCAATGACTGCATGAAACGG-3'
PRORP1 desS8 D474A	5'-GGCTCAAACGCTGATTGGTATTGGCTTTAC-3'	5'-ATACCAATCAGCGTTTGAGCCAGGAGGGGT-3'
PRORP1 desS8 D475A	5'-TCAAACGATGCTTGGTATTGGCTTTACGCT-3'	5'-CCAATACCAAGCATCGTTTGAGCCAGGAGG-3'
PRORP1 desS8 D493A	5'-GTAACAAACGCTGAGATGAGAGACCATCTC-3'	5'-TCTCATCTCAGCGTTTGTTACTAACAAACA-3'
PRORP1 desS8 D497A	5'-GAGATGAGAGCCCATCTCTTCCAGTTACTT-3'	5'-GAAGAGATGGGCTCTCATCTCATCGTTTGT-3'
PRORP2 desN9 D343A	5'-GCTATGGTAGCTGGAGCAAATATTGGGCTC-3'	5'-TTTGCTCCAGCTACTATAGCTTCGTAATC-3'
PRORP2 desN9 D440A	5'-ACAAATGCTGAGATGAGAGACCACATTTTTGAAC-3'	5'-TCTCATCTCAGCATTTGTCACAAGCAAACACTTGA-3'
PRORP1 desS9	5'-AAAACATATGAAAGCTATTCAACAGTCTCC-3'	5'-AAAAGGATCCTCATCAAGGTGTTTTGGATCTTTTTGCGCATAACC-3'
PRORP1 desI29	5'-AAAACATATGCTTAGCCAGTACCATTATAATGTG-3'	5'-AAAAGGATCCTCATCAAGGTGTTTTGGATCTTTTTGCGCATAACC-3'
PRORP1 desI49	5'-AAAACATATGGAGTCGTCTCCAAATCCAGGG-3'	5'-AAAAGGATCCTCATCAAGGTGTTTTGGATCTTTTTGCGCATAACC-3'
PRORP1 desI72	5'-AAAACATATGGTTCCCAATGAAGCTAC-3'	5'-AAAAGGATCCTCATCAAGGTGTTTTGGATCTTTTTGCGCATAACC-3'
PRORP1 desS220	5'-AAAACATATGTTCTGTAGGAAAGGTGATGC-3'	5'-AAAAGGATCCTCATCAAGGTGTTTTGGATCTTTTTGCGCATAACC-3'
PRORP1 desS95	5'-AAAACATATGGCAGTCATTGATGGAGCC-3'	5'-AAAAGGATCCTCATCAAGGTGTTTTGGATCTTTTTGCGCATAACC-3'

Mutation sites were indicated by underlines.

Expression and purification of mutant enzymes

The expression plasmids containing D474A/D475A double point mutations, where the catalytic residues Asp474 and Asp475 were simultaneously replaced by Ala, in PRORP1 des38 gene, and D343A and D440A single point mutations, where Asp343 and Asp440 were individually replaced with Ala, in PROPR2-desN9 gene were introduced into *E. coli* strain BL21 (DE3) codon plus RIL. Likewise, the expression vectors in which they contain D399A, D474A, D475A, D493A, and D495A single point mutations, where Asp399, Asp474, Asp475, Asp493, and Asp497 were individually replaced with Ala, in PRORP1 des38 gene, were prepared. The transformants were cultivated as described above. After confirmation of the expression of the mutant proteins by SDS-PAGE, these proteins were purified using a series of column chromatography according to the purification method for the wild-type.

Preparation of N-terminal successive deletion mutants of PRORP1

PRORP1 des89, PRORP1 des129, PRORP1 des149, and PRORP1 des220, where the N-terminal region (pos. 38-89), PPR1 motif (pos. 95-129), PPR1-2 motifs (pos. 95-170), and PPR1-3 motifs (pos. 95-226) were deleted, respectively, were prepared. The cDNA fragments encoding the N-terminal deletion mutants were PCR amplified using two sets of primers, as described in Table -Ⅱ (Table -Ⅱ). The resulting PCR products were ligated into expression vector pET-15b, and their expression was induced in *E. coli* BL21(DE3) Codon Plus RIL in a conventional manner. The overproduced proteins were purified homogeneity using COSMOGEL His Accept column chromatography (Nakarai Tesque), followed by RESOURCE S column chromatography (GE Healthcare). Subsequently, they were further purified with HiLoad 16/60 Superdex 200 gel filtration column chromatography (GE Healthcare).

Gel mobility shift assay

The purified pre-tRNA was heated to 70 ℃ for 10 min, fo
room temperature before use. The binding reaction was performed with the increased amounts of PRORP1 des38, PRORP1 des89, or PRORP1 des220, and 2 μg of the pre-tRNA^{Phe} in the reaction buffer containing 30 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM DTT, 5 mM $MgCl_2$, and 10% glycerol. Reactions were incubated for 10 min

under 4 °C. Subsequently at 150 V at room temperature. The gel was stained by ethidium bromide solution and detected by UV absorbance detection.

Circular dichroic (CD) spectroscopy

During purification of PRORP1 des89 by a RESOURCE S column, the two peaks containing PRORP1 des89 were obtained (denoted PRORP1 des89 A and B, hereafter). To verify structural integrity of PRORP1 des89 A and PRORP1 des89 B, I measured their circular dichroism (CD) spectra using a JASCO J 720 spectropolarimeter at room temperature. PRORP1 des89 A (1 mg/ml) and B (1 mg/ml) were dissolved in 30 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM DTT, 5 mM $MgCl_2$, and 10% glycerol. The path lengths of the cells used for the wavelength region (190-250 nm and 260-320 nm) were 1 cm.

Limited proteolysis

Limited trypsin proteolysis was performed using 45 μg of purified PRORP1 des38. The protein solution was incubated with trypsin purchased from Sigma Aldrich (St. Louis, State of Missouri, United States of America) at a mass ratio of 1 : 450 and the partial proteolysis was conducted at 37 °C. Aliquots (10 were removed at 1, 3, 5, 10, and 15 min. The reactions were stopped by the addition of SDS-PAGE loading buffer and further analyzed by SDS-PAGE on 10% polyacrylamide gel, followed by staining with Coomasie Brilliant Blue G-250 (Nakarai Tesuque). The limited proteolysis using increasing amount of chymotrypsin supplied with Sigma Aldrich was conducted in the same manner as limited trypic digestion of PRORP1 des38 as described above except that the amount of PRORP1 des38 was 7.6 μg and then incubated with chymotrypsin at a mass ratio of 1: 152,000.

RESULTS

PCR amplification and construction of the expression vectors

The transit peptide containing 38 amino acids is truncated at the time when PRORP1 is transported into chloroplasts and mitochondria, giving rise to maturation of

PRORP1 (Gobert *et al.,* 2010). Accordingly, PRORP1 lacking the signal sequence (pos.1-38) (referred to as PRORP1 des38, hereafter) was used in this study (Figure -Ⅱ). To construct the expression plasmid, the amplification reaction by PCR was carried out using two sets of primers, as described in Table -Ⅱ. Analysis of the amplification products by agarose gel electrophoresis revealed that approximately 1,600 bp fragments were amplified by PCR (Figure -Ⅳ). The band corresponding to 1,600 bp was an expected size based on the length of the cDNA encoding the PRORP1 des38. Thus, the amplified fragment was extracted from the gel, and then digested with the restriction enzymes, *Nde* Ⅰ and *Bam*H Ⅰ, followed by ligated into pET-15b vector previously digested with the same restriction enzymes. The ligated plasmid was introduced the *E. coli* strain JM109 and the cloned plasmid was verified by DNA sequencing. After verification of its sequence, the plasmid DNA was introduced into the *E. coli* expression strain, BL21 (DE3) codon Plus RIL.

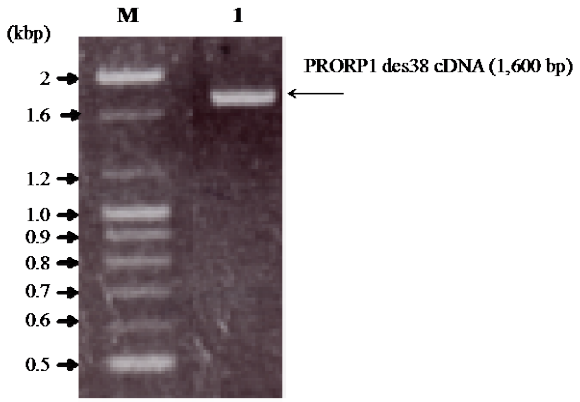

Figure -Ⅳ: PRORP1 des38 DNA fragments amplified by PCR. The DNA size marker is a commercial 1 kbp ladder. M and 1 represent molecular marker and the synthesized DNA fragment encoding PRORP1 des38 (1,600 bp), respectively.

Plasmid encoding PRORP2 and PRORP3, where the N-terminal 9 and 10 amino acids are truncated, respectively (referred to as PRORP2 desN9 and PRORP3 desN10), were constructed in the same manner as PRORP1 des38. I optimized over-expression systems of these proteins in *E. coli* cells BL21 (DE3) codon Plus RIL, and then checked by SDS-PAGE (Figure -Ⅴ). This result showed that PRORP1 des38, PRORP2 desN9, and PRORP3 desN10 were produced in a soluble form in *E. coli* cells by adding 1 mM IPTG under 16 ℃ for 18 hours.

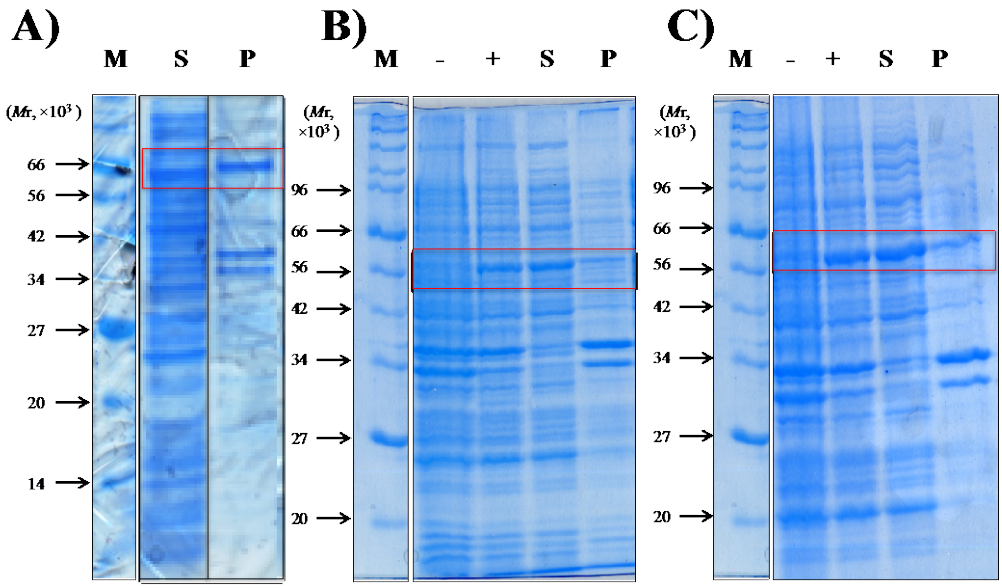

Figure II-5: Overexpression of the recombinant proteins by utilizing the expression system in *Escherichia coli*. The *E. coli* cells harboring the plasmid encoding PRORP1 des38 (A), PRORP2 desN9 (B), and PRORP3 desN10 (C) were harvested in 16 °C for 8, 18, and 40 h. M represent molecular marker, cell cultivation performed without adding IPTG, cell cultivation performed with adding 1 mM IPTG, the centrifuged disrupted cell dispersion (supernatant), and the centrifuged disrupted cell dispersion (pellet), respectively. The electrophoresis bands corresponding with PRORP1 des38, PRORP2 desN9, and PRORP3 desN10 are marked with red boxes.

Purification of PRORP1 des38, PRORP2 desN9, and PRORP3 desN10

The recombinant PRORP1 des38, fusing with hexa-histidine tag at the N-terminus, was purified by a series of column chromatographies. The harvested cells containing PRORP1 des38 were lysed by sonication, and then centrifuged. The supernatant fractions containing the soluble PRORP1 des38 were put onto a SP-sepharose Fast Flow column, previously equilibrated with the buffer A (30 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, and 1 mM DTT) (Figure II-6). The absorbed protein was eluted with a linear gradient from 0.1 M to 1 M NaCl and then interested fractions were pooled and dialyzed against the buffer A. During this purification, PRORP1 des38 was accidentally degraded by trypsin-like proteinases despite of existing a serine proteinase inhibitor, 1 mM PMSF (Figure II-7). In order to separate intact PRORP1 des38 from the polypeptide fragment produced by proteinase, I put onto a hydrophobic interaction column, RESOURCE ISO (GE Healthcare). PRORP1 des38 was absorbed onto the column, whereas the fragment produced by proteinase was flowed through the column (Figure II-7). The absorbed protein was

eluted by a linear gradient of ammonium sulfate from 1.5 M to 0 M, and then dialysis against the buffer A and stored at -80 ℃ before we use.

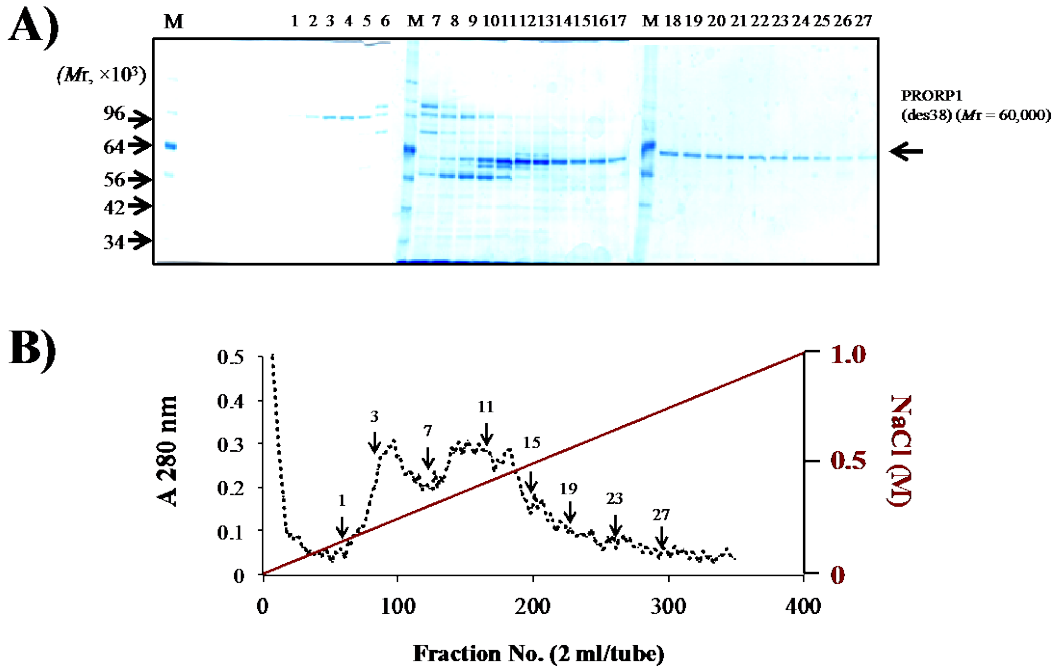

Figure 4: Preparation of PRORP1 des38 by a RESOURCE S column. (A) SDS-PAGE of the purified PRORP1 des38 by a SP-sepharose Fast Flow column (2 × 5 cm). M represents molecular marker. The electrophoresis band of PRORP1 des38 is depicted as black arrows. (B) The chromatogram of the purified PRORP1 des38 by the column. The electric conductivity is shown by brown lines. The analyzed fractions by SDS-PAGE represent by black arrows with their numbers.

I purified PRORP2 desN9 by a series of column chromatographies. The overexpression was done in the same manner as PRORP1 des38. The *E. coli* BL21 (DE3) codon Plus RIL, bearing the expression vector was used in this study. The harvested cells were lysed by sonication, and then purified with a COSMOGEL His Accept column (Figure 5). In order to prevent aggregation of PRORP2 desN9, the eluted fractions were pooled and then immediately diluted up to the absorbance at 280 nm (A280) of 0.2 by the same buffer. Subsequently, the eluted protein was dialyzed against the buffer C (30 mM Tris-HCl, pH 8.5, 100 mM NaCl, 10% glycerol, and 1 mM DTT), followed by purification with a RESOURCE Q anion exchange column (GE Healthcare) (Figure 6). The absorbed protein was eluted with a linear gradient from 0.1 M to 0.45 M NaCl. The interested protein was digested by thrombin to exclude the hexa-histidine tag from the protein at 4 ℃ for 12 hours. The digested protein was further purified with HiLoad 16/60 Superdex 200 gel filtration column chromatography,

followed by adding 20 mM DTT in order to prevent an intermolecular disulfide-bond interaction (Figure II-10). The yield was 0.3 mg per 1 liter cultivated LB-medium. Likewise, PRORP3 desN10 was overexpressed and purified in the same manner as PRORP2 desN9. The yield was about 0.3 mg per 1 liter induced medium (Figures II-8 ~ 10).

Characterization of PRORP1 des38, PRORP2 desN9, and PRORP3 desN10

To confirm whether it is enough to perform an assay for pre-tRNA cleavage by PRORP1 des38, PROPR2 desN9, and PRORP3 desN10, or not, their enzymatic activities were characterized using chloroplast pre-tRNA^{Phe} in *A. thaliana.* The study showed that PRORP2 desN9 did not exhibit its catalytic activity at 37 ℃, whereas PRORP1 des38 cleaved the 5' flanking sequences from chloroplast pre-tRNA^{Phe} in *A. thaliana* at 37 ℃. Our further study suggested that the optical temperature for PRORP1 and PRORP3 are around 37 ℃, whereas that of PRORP2 is around 28 ℃. This finding was also supported by Gutmann *et al.* (Gutmann *et al.,* 2012). Based on this information, the pre-tRNA processing activities for PRORP1 des38 and PROPR3 desN9 were assayed at 37 ℃, while PRORP2 desN9 was assayed at 28 ℃ (Figure II-11).

Preparation of mutant enzymes

Comparing the primary sequence of PRORP1 with those of PRORP2 and PRORP3, the five-aspartate amino acids (Asp399, Asp473, Asp474, Asp493, and Asp497 in PRORP1) are conserved among PRORPs in *A. thaliana* (Figure II-12). Additionally, these amino acids are also conserved between the PIN and Flap nuclease proteins (Anantharaman and Aravind, 2006). In 2010, Gobert and coworkers demonstrated the importance of Asp474 and Asp475 for the pre-tRNA processing activity of PRORP1 using the pre-tRNA processing assay *in vitro* (Gobert *et al.,* 2010). Based on this finding, I first constructed the expression vector, harboring the cDNAs of PRORP1 des38 mutants where Asp475 and Asp474 were individually replaced with Ala. The mutants were overexpressed in *E. coli* cells and purified with several column chromatographies in the same manner as PRORP1 des38. Their homogeneities were checked by SDS-PAGE. Their yields from 1 liter of culture broth were approximately 0.5 mg. Moreover, I also prepared the mutants, including PRORP1 des38 D399A,

D474A, D475A, D493A, and D497A, where Asp399, Asp474, Asp475, Asp493, and Asp497 were in turn replaced with Ala, respectively. Their expression and purification were conducted in the same manner as PRORP1 des38. Furthermore, to demonstrate the involvement of Asp343 and Asp440 in the catalysis of PRORP2 (corresponding to Asp399 and Asp493 in PRORP1, respectively), Asp343 and Asp440 were individually replaced with Ala by the site-directed mutagenesis kit purchased from Takara Bio (denoted as PRORP2 desN9 D343A and PRORP2 desN9 D440A, respectively). Mutants were over-expressed in the same manner as PRORP2 desN9. Although PRORP2 desN9 D343A mutant was not expressed in *E. coli* cells, PROPR2 desN9 D440A mutant was obtained as soluble proteins. Consequently, PRORP2 desN9 D440A was overproduced, and then purified in the same manner as PRORP2 desN9.

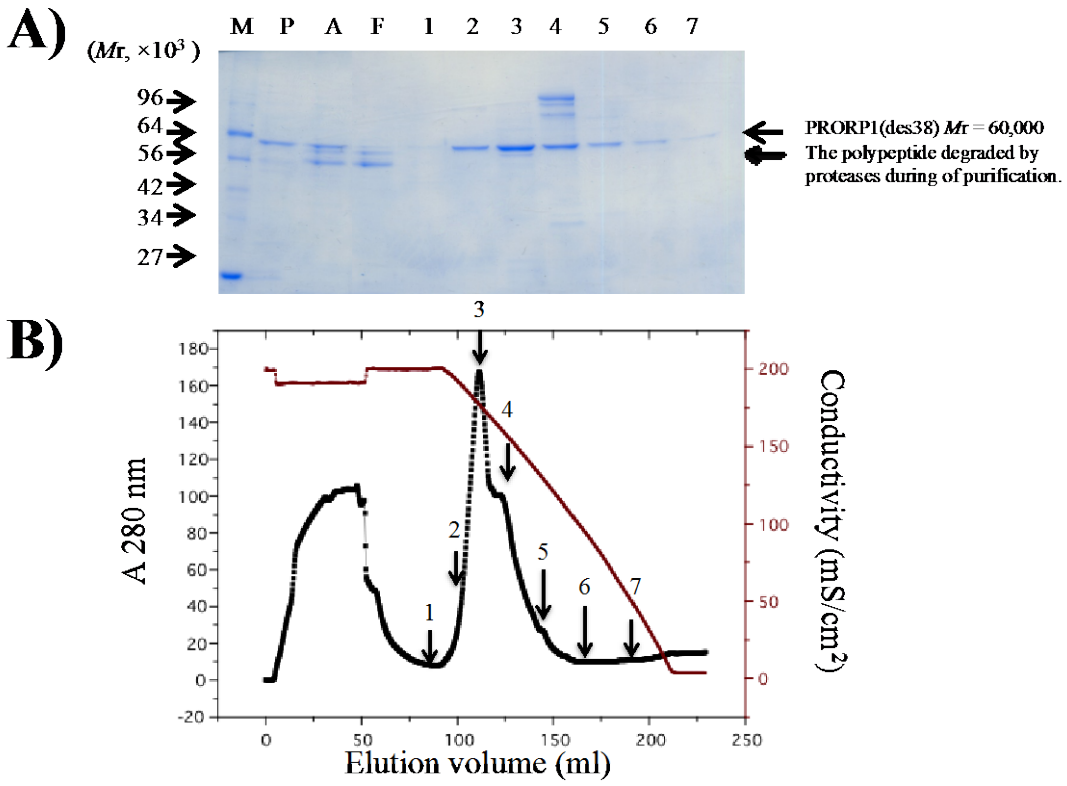

Figure Ⅱ: Preparation of PRORP1 des38 by a RESOURCE ISO column . (A) SDS-PAGE of the purified PRORP1 des38 by a RESOURCE ISO column. M, Molecular marker; P, Purified PRORP1 des38 by RESOURCE S column; A, Applied sample; F, Unbound fractions. The electrophoresis bands of PRORP1 des38 and the polypeptide fragment are depicted as black arrows. (B) The chromatogram of purified PRORP1 des38 by a RESOURCE ISO column. The electric conductivity is shown by brown lines. The analyzed fractions by SDS-PAGE are represented by black arrows with their numbers.

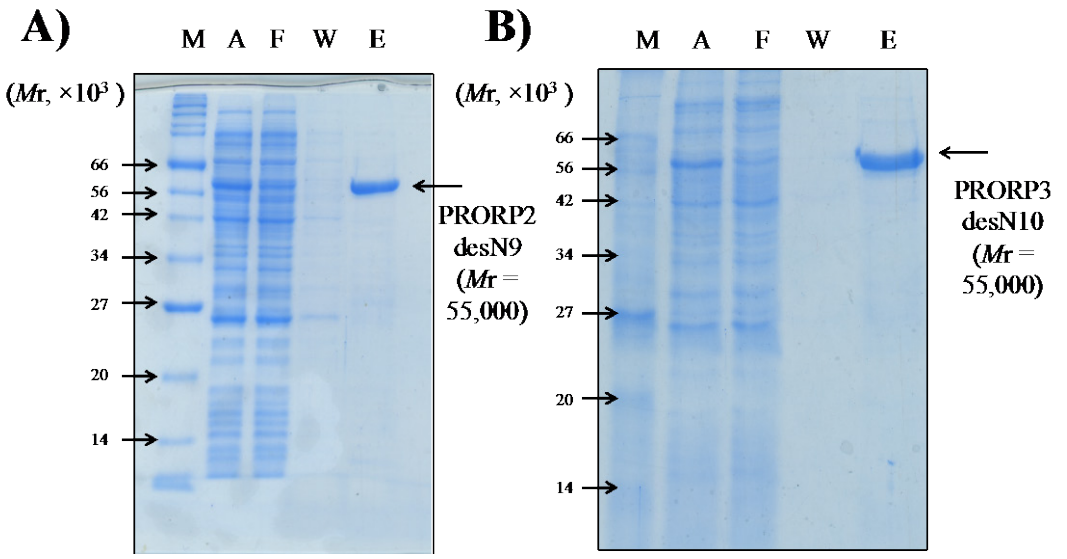

Figure 3: SDS-PAGE of the purified PRORP2 desN9 and PRORP3 desN10. (A) Nickel affinity column chromatography of PRORP2 desN9. M, Molecular marker; A, Applied sample; F, Unbound fractions; W, Washed fractions; E, Bound fractions. The electrophoresis band of PRORP2 desN9 is depicted as black arrows. (B) Nickel affinity column chromatography of PRORP3 desN10. M, Molecular marker; A, Applied sample; F, Unbound fractions; W, Washed fractions; E, Bound fractions. The electrophoresis band of PRORP3 desN10 is depicted as black arrows.

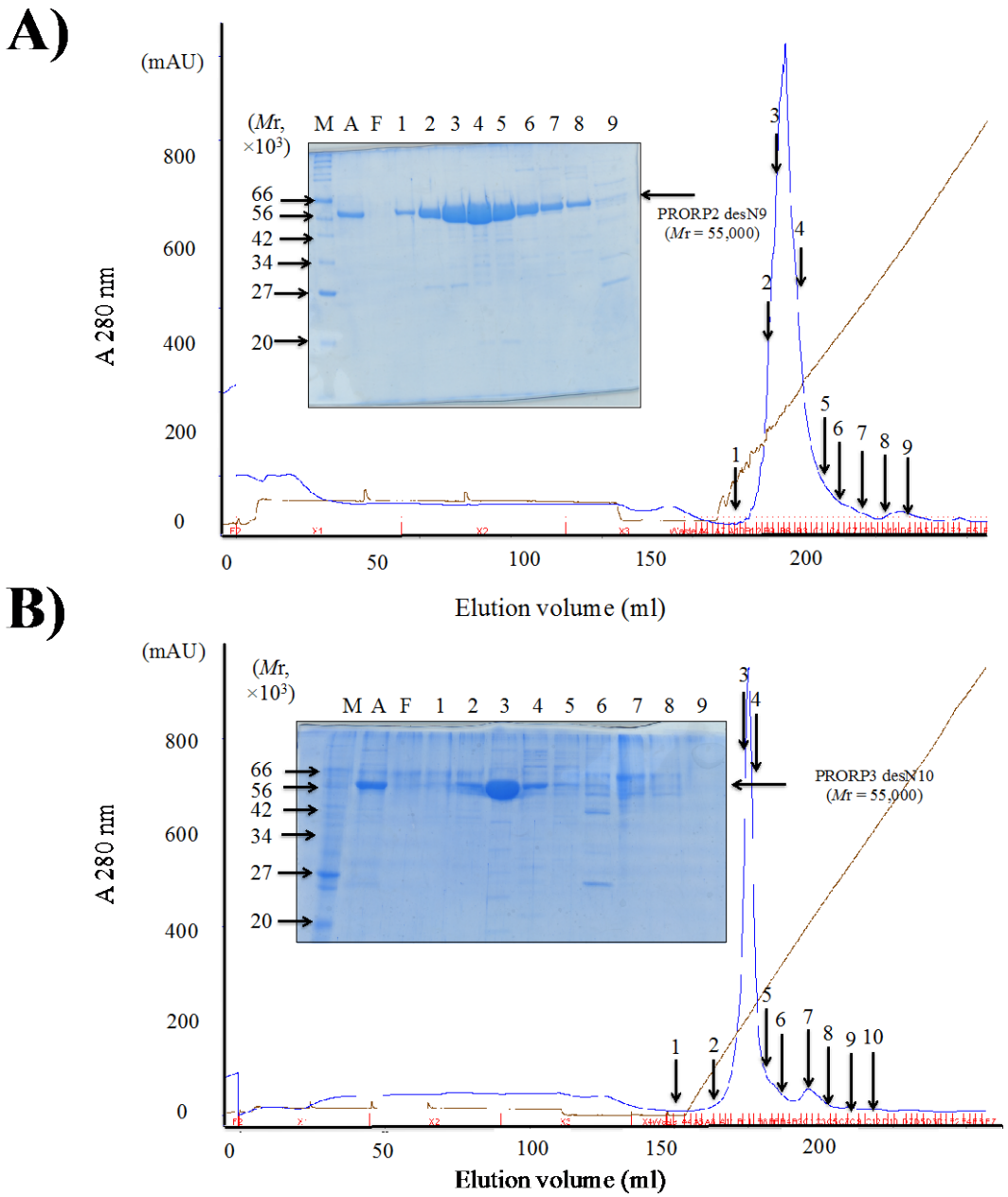

Figure 11: Ion exchange column chromatography of PRORP2 desN9 (*A*) and PRORP3 desN10 (B) by a RESOURCE Q column. M, Molecular marker; A, Applied sample; W, Washed fractions. The electronic conductivities are depicted as brown lines. The analyzed fractions by SDS-PAGE are represented by black arrows with their numbers in the chromatograms. The electrophoresis bands of PRORP2 desN9 and PROPR3 desN10 are represented by black arrows.

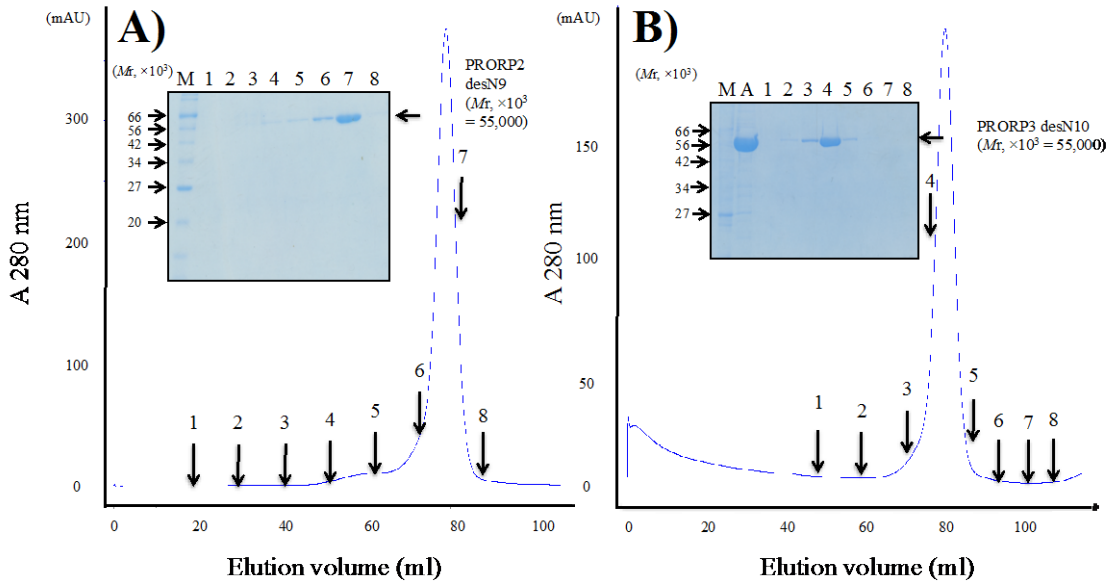

Figure 10: Preparation of PRORP2 desN9 (A) and PRORP3 desN10 (B) by a HiLoad 16/60 Superdex200 column. The homogeneity of these proteins was checked by SDS-PAGE. M represents molecular marker. A represents applied sample. The values of A 280 nm are depicted by blue lines on the chromatograms. The analyzed fractions by SDS-PAGE are marked with black arrows with their numbers on the chromatograms.

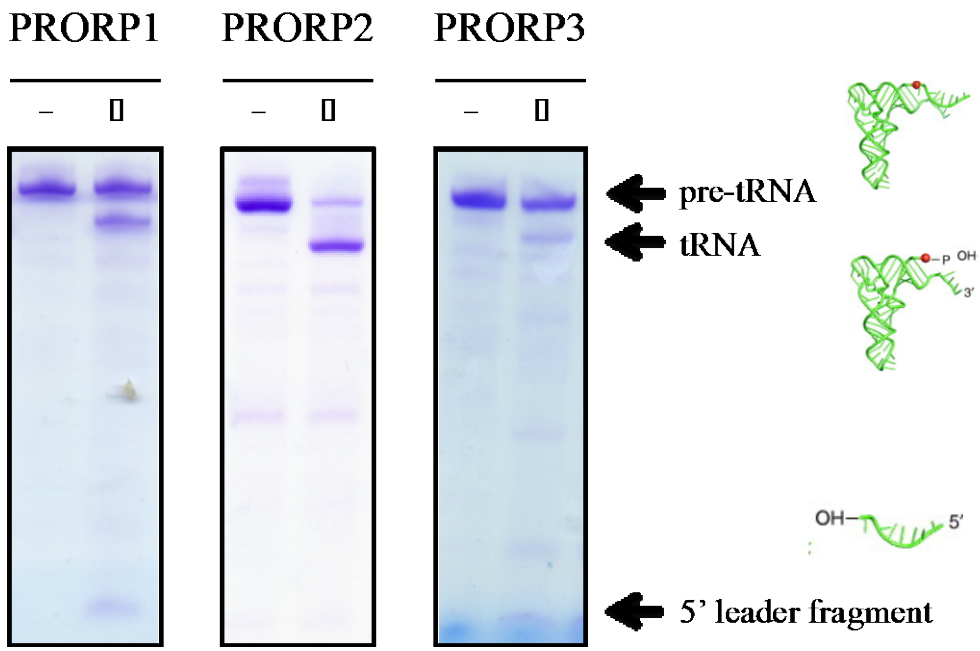

Figure 11: pre-tRNA processing activity of PRORP1 des38, PRORP2 desN9, and PROPR3 desN10. The assay for pre-tRNA cleavage by PRORP1 and PRORP3 were performed at 37 °C for 1 h, whereas that by PRORP2 was done at 28 °C for 1 h. –and + represent pre-tRNA cleavage assay performed without PRORP enzymes and the assay performed with PRORP enzymes, respectively. The electrophoresis gels were stained by 0.1% toluidine blue N solution. The electrophoresis bands of pre-tRNA, tRNA, and 5' leader fragment, are marked by black arrows.

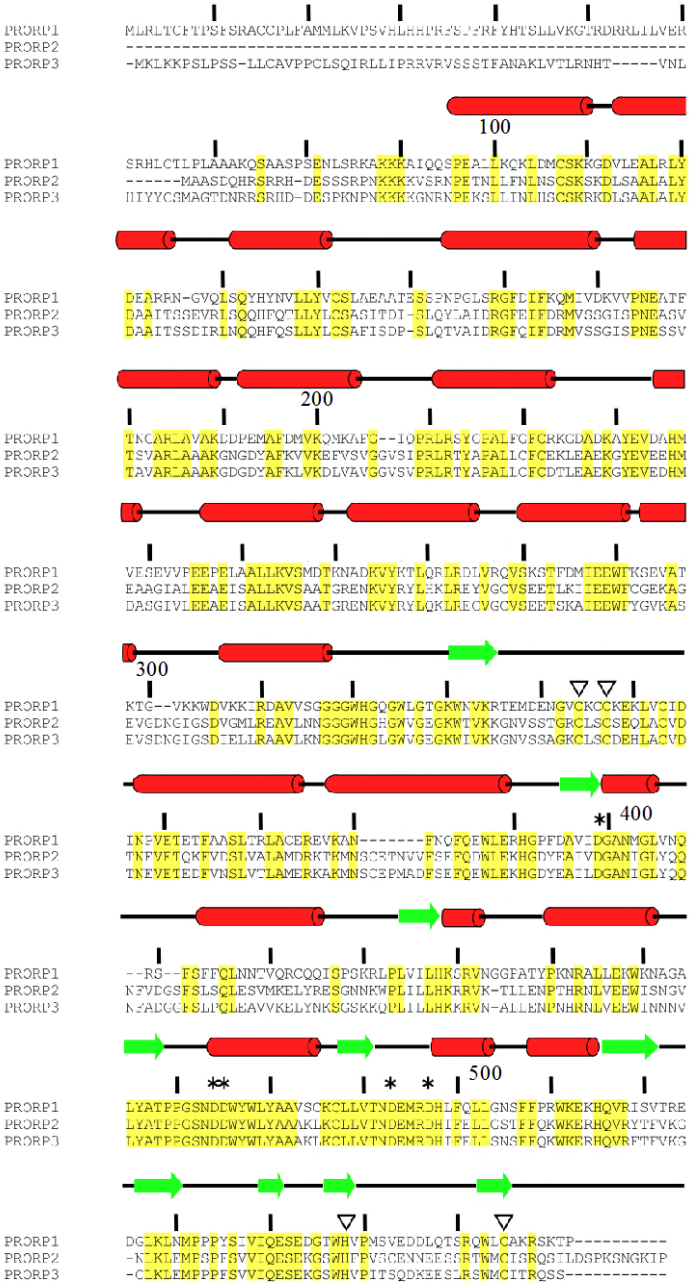

Figure II-2: Sequence alignment of PRORP1, PRORP2, and PRORP3 from *A. thaliana*. The secondary structure was predicted from the crystal structure of PRORP1 (PDB ID: 4G23). The α helices are colored in red, and the β sheet structures are done in green. The yellow highlighted box represents completely conserved regions. The CCCH Zn binding site is designated by black arrowheads. The asterisks indicate residues that were mutated to Ala in this study.

Pre-tRNA processing activity of mutant enzymes

I conducted the pre-tRNA processing assay by PRORP1 des38 D474A/D475A double mutant and compared it with that of PRORP1 des38 (Figure II-3). The result

showed that this mutant exhibited little activities, being in agreement with the previous data reported by Gobert *et al.* (Gobert *et al.*, 2010). To obtain more insight into the catalytic residues of PRORP1, the five mutants including D399A, D474A, D475A, D493A, and D497A were characterized in cleaving pre-tRNA^{Phe} from *A. thaliana* chloroplasts (Figure -Ⅱ3). Their mutants exhibited little activities, indicating the importance of these aspartate residues for PRORP1. Likewise, the PRORP2 D343A mutant exhibited little activity (Figure -Ⅱ3). It is thus likely that PRORP1, PRORP2, and PRORP3 catalyze pre-tRNA cleavage reaction in the same manner.

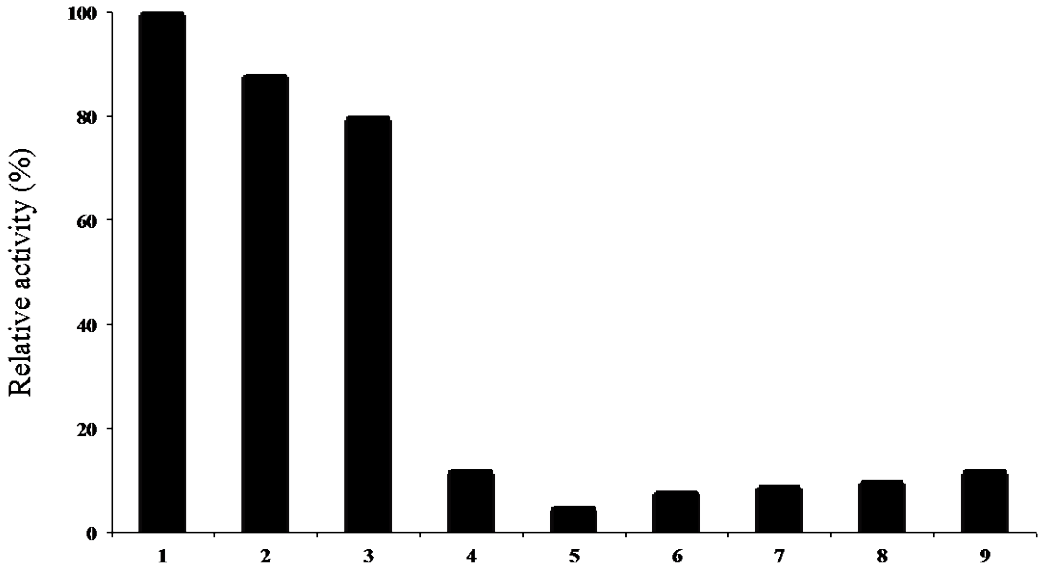

Figure -Ⅱ3: Pre-tRNA cleavage activity of PRORP1, PRORP2, PRORP3, and their mutants. Pre-tRNA cleavage activities of the enzymes were assayed at 37 °C using *A. thaliana* pre-tRNA^{Phe}, while that of PRORP2 was conducted at 28 °C. Their cleavage PRORP1 as 100%. Column 1, PRORP1 des38 wild-type; column 2, PRORP2 desN9 wild-type; column 3, PRORP3 desN10 wild-type; column 4, PRORP1 des 38 D399A; column 5, PRORP1 des38 D474A; column 6, PRORP1 des38 D475A; column 7, PRORP1 des38 D493A; column 8, PRORP1 des38 D497A; column 9, PRORP2 desN9 D343A.

Preparation of the N-terminal deletion mutants of PRORP1.

During purification of PRORP1 des38, I found that peptide bond locating at Lys-Gly (positions 89-90 amino acids) was highly sensitive to proteolysis. It is suggested that this region located on the molecular surface of PRORP1. To corroborate the proteolysis, limited trypic digestion of PRORP1 des38 was conducted (Figure -Ⅱ4 and 15). This result showed that the C-terminus of Arg220 as well as that of Lys89 was sensitive to the digestion. The limited chymotrypic digestion of PRORP1 indicated that the scissile bond at the C-terminus of Phe222 is highly sensitive to chymotrypsin. Thus, I named the obtained polypeptides as fragment 1: F1 (pos. 90-572) and fragment 2: F2

(pos. 220-572), as described in Figure Ⅱ-5. These observations are well consistent with the crystal structure of PRORP1 (pos. 95-570) reported recently, in which the N-terminal residues up to 90 residues are highly disordered. Based on the observation, the DNA fragments encoding pos. 90-572 corresponding to F1 and pos. 221-572 corresponding to F2, were PCR-amplified to characterize the functional roles in the PPR domain of PRORPs. The amplified DNA fragments were ligated to pET-15b expression vector, and then introduced into *E. coli* cells, BL21(DE3) CodonPlus RIL (Table Ⅱ-1).

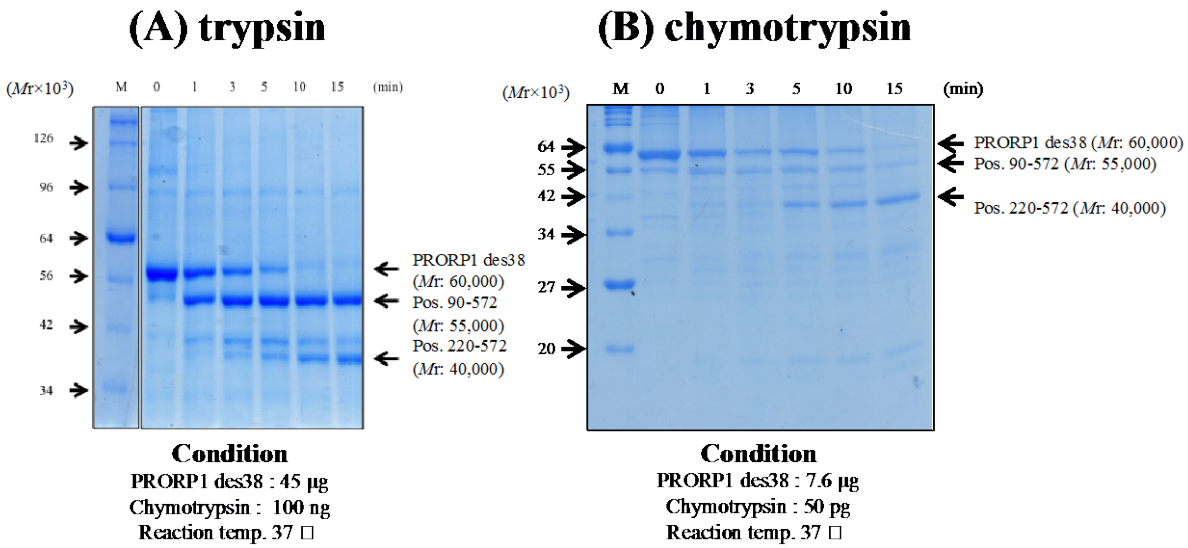

Figure Ⅱ-4: Limited proteolysis of PRORP1 des38. (A) Trypic digestion of PRORP1 des38 was carried out in the buf. A at 37 ℃. Aliquots were withdrawn at appropriate intervals (0, 1, 3, 5, 10, 15 min), and then subjected to SDS-PAGE. (B) PROPR1 des38 was limited digested with chymotrypsin. Aliquots were withdrawn at appropriate intervals (min) and subjected to SDS-PAGE. M represents molecular marker. The electrophoresis bands containing PRORP1 des38, the fragment (pos. 90-572), and the fragment (pos. 220-572) are depicted as by black arrows.

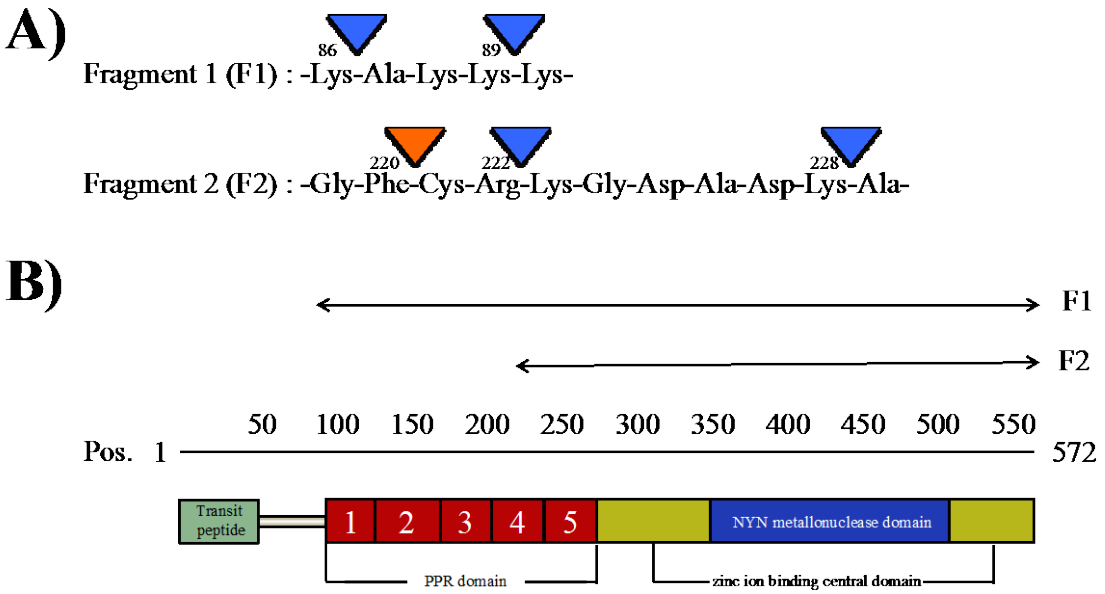

Figure Ⅱ-5: Cleaved sites of PRORP1 by trypsin and chymotrypsin. (A) The cleaved sites by trypsin and those of chymotrypsin located on fragments 1 and 2 (denoted F1 and F2) are depicted as blue arrowheads and orange arrowheads, respectively. (B) The cleaved sites F1 and F2 are mapped onto the domain structure of PRORP1.

These N-terminal deletion mutants (named PRORP1 des89 and PRORP1 des220, hereafter) were over-produced in *E. coli* cells and purified to homogeneity (Figures Ⅱ-16, 17, and 18). During purification of PRORP1 des89 by a Resource S column, two peaks containing PRORP1 des89 were obtained (Figure Ⅱ-17). The first peak was referred to as PRORP1 des89 A, whereas the back peak was called as PRORP1 des89 B.

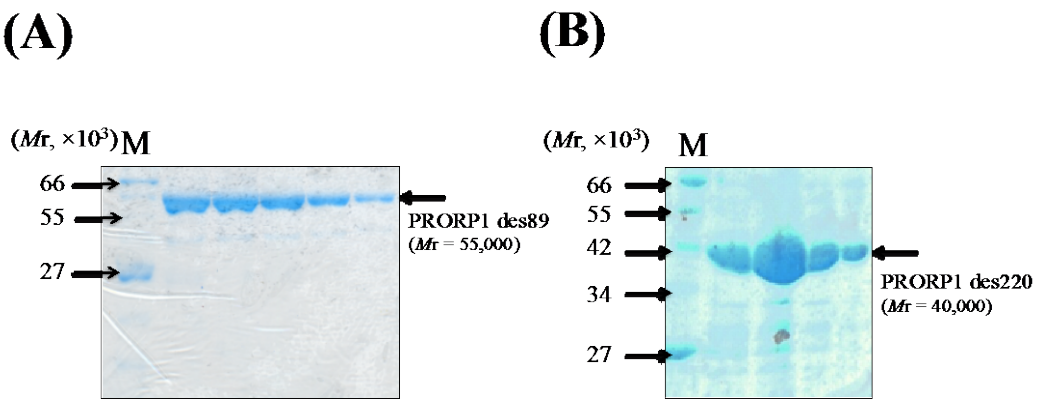

Figure Ⅱ-6: SDS-PAGE gels of PROPR1 des89 and PRORP1 des220 purified by Ni-affinity column chromatography. (A) SDS-PAGE of PRORP1 des89 purified by a COSMOGEL His Accept column. M represents molecular marker. The electrophoresis band of PRORP1 des89 is marked by black arrows. (B) SDS-PAGE of PRORP1 des220 purified by the same column. The electrophoresis band of PRORP1 des220 is depicted by black arrows.

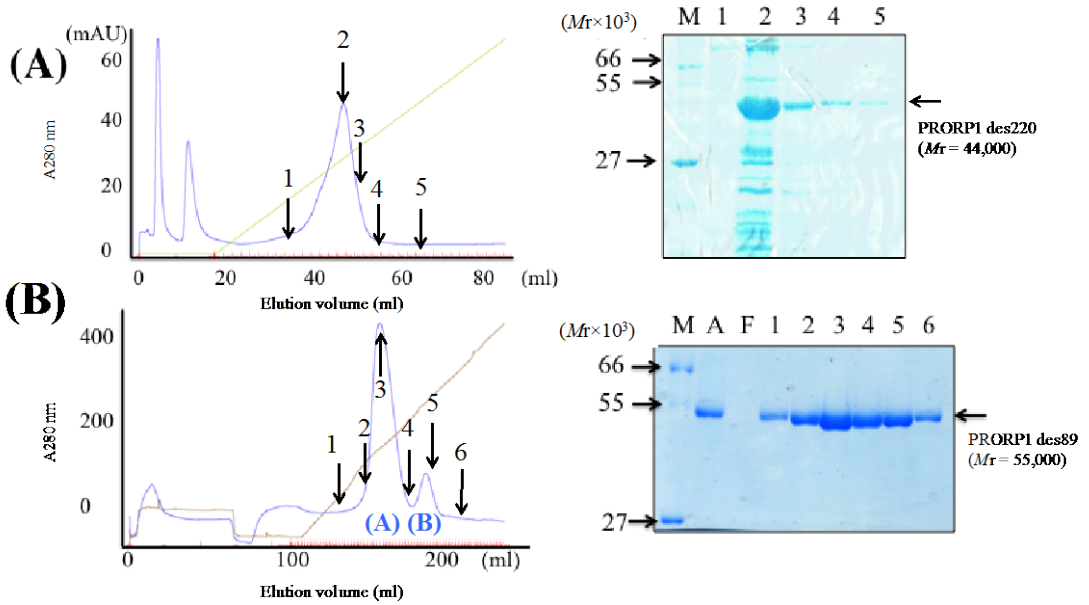

Figure -Ⅱ7: Ion exchange column chromatography of PRORP1 des89 and PRORP1 des220. (A) Purification of PRORP1 des220. Green line represents electric conductivity. The analyzed fractions by SDS-PAGE are marked with black arrows with their numbers on the chromatograms. (B) Purification of PRORP1 des89. Brown line represents electric conductivity. M, Molecular marker; A, Applied sample; F, Unbound fractions. The first peak was denoted as PRORP1 de89 A and the second peak referred to PRORP1 des89 B. The analyzed fractions by SDS-PAGE are marked with black arrows with their numbers on the chromatograms. The electric bands containing PRORP1 des89 and PRORP1 des220 are marked with black arrows.

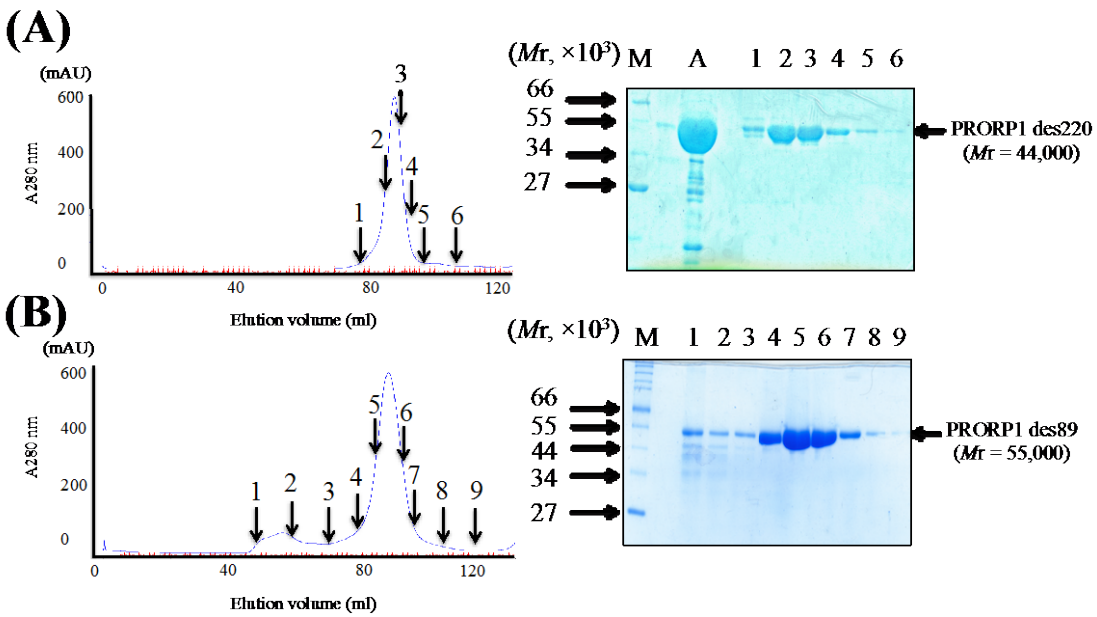

Figure -Ⅱ8: Gel filtration column chromatography of PRORP1 des89 and PRORP1 des220. (A) Purification of PRORP1 des220 by a HiLoad 16/60 Superdex200 column. (B) Purification of PRORP1 des89 by a HiLoad 16/60 Superdex200 column. M and A represent molecular marker and applied sample, respectively. The analyzed fractions by SDS-PAGE are marked with black arrows with their numbers on the chromatograms. The electric bands containing PRORP1 des220 and PRORP1 des89 are depicted as black arrows.

To obtain more information about the difference between PRORP1 des89 A and PRORP1 des89 B, I examined the pre-tRNA processing activity. These proteins exhibited almost the same reaction efficiency. Subsequently, CD spectroscopy was measured, indicating that although the CD spectrums in a range of a short wavelength 〔λ(nm): 190-250〕of these proteins were identical, the CD spectrums in a range of long wavelength〔λ(nm): 250-320〕of them were slightly different (Figure Ⅱ-9). This result suggested that the three dimensional structures of PRORP1 des89 A and PRORP1 des89 B were slightly different. As described below in detail, it is probably attributed to the N-terminal capping structure, which serves in structural stabilization of PPR motifs in PRORP1. Hence, the deletion of the N-terminal region resulted in producing the two types of PRORP1 des89. The question is needed to be elucidated.

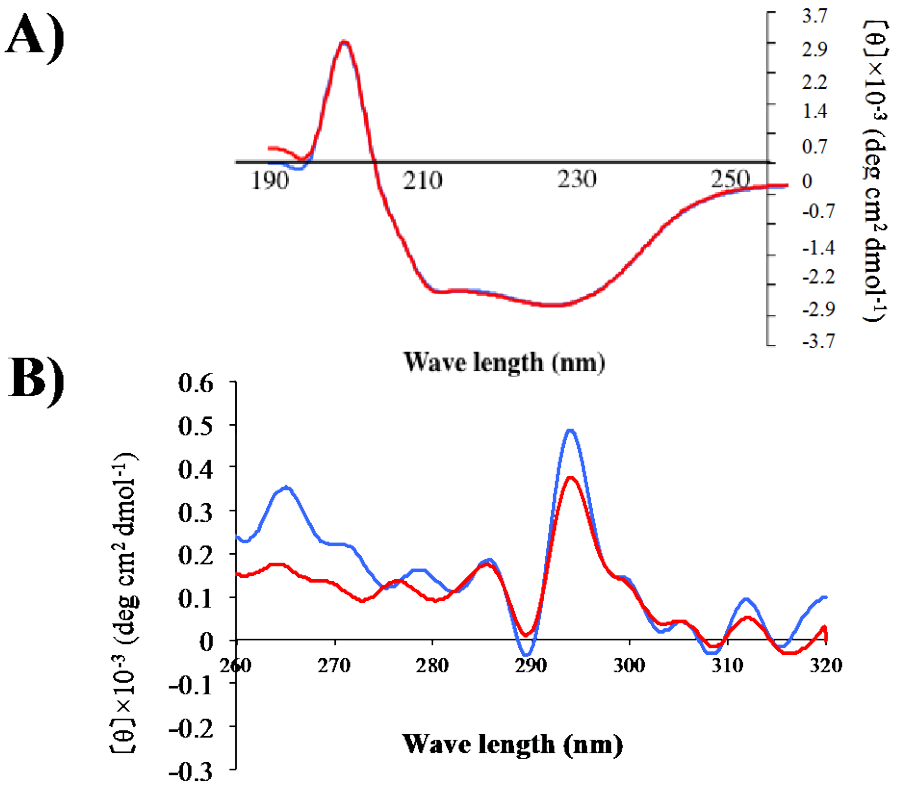

Figure Ⅱ-9: Circular dichroic (CD) spectra of PRORP1 des89 A and PRORP1 des89 B. CD spectrums of PRORP1 des89 A and B were measured in a wavelength of (A) 190–250 nm and (B) 260–320 nm, respectively. Blue: PRORP1 des89 A; red: PRORP1 des89 B.

Pre-tRNA^{Phe} cleavage activity of PRORP1 des38 and its N-terminal deletion mutants

Subsequently, the purified PRORP1 des89 and PRORP1 des220, in addition to PRORP1 des38, were characterized in terms of the pre-tRNA^{Phe} processing activity using chloroplast tRNA^{Phe} from *A. thaliana* (Figure 10). Although PRORP1 des89 retained the cleavage activity of pre-tRNA, it did slightly less efficiently than PRORP1 des38, indicating that the N-terminal residues (pos. 38-89) partly contribute to the enzymatic activity (40%). On the other hand, PRORP1 des220 exhibited little cleavage activity, suggesting that PPR domains play an essential role in pre-tRNA processing activity of PRORP1 (Figure 10).

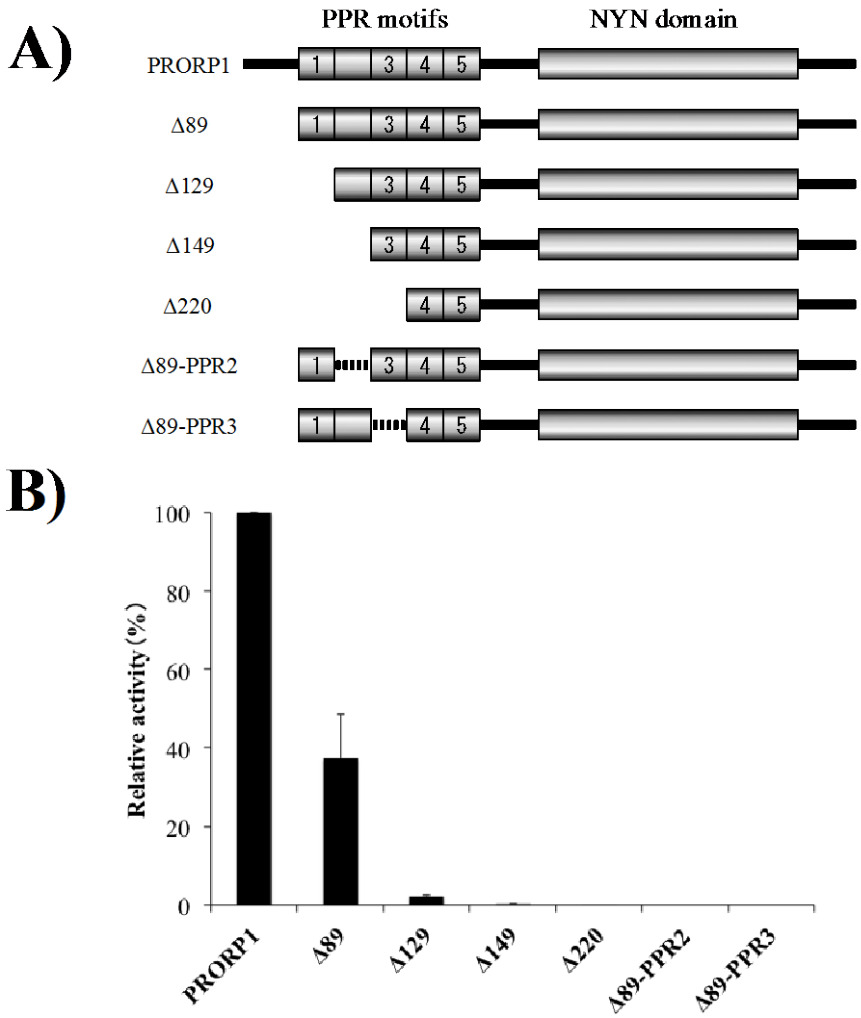

Figure 10: The pre-tRNA cleavage activity of PRORP1 and its successive N-terminal deletion mutants. (A) Schematic diagram of a modular architecture of PRORP1 with approximate domain boundaries indicated. The N-terminal deletion mutants used in this study are shown. (B) The pre-tRNA^{Phe} cleavage activity of PRORP1 and its N-terminal deletion mutants. Pre-tRNA^{Phe} cleavage activities of the

enzymes were assayed at 37 ℃ using *A. thaliana* pre-tRNA^{Phe} as described in MATERIALS AND METHODS. The cleavage products were resolved on 10% polyacrylamide gel in 8 M urea. Subsequently, the resulting gel was visualized by staining the gels in a 0.1% toluidine blue N solution. The individual bands were quantified using Multi Gauge V3.1. The cleavage efficiency was calculated as (matured-tRNA^{Phe}) / (pre-tRNA^{Phe}+ tRNA+5'- leader sequence) × 100. The experiments were conducted in triplicate, and then the mean values are presented.

To gain more insight into the function of each PPR motif in PRORP1, PRORP1 des38, PRORP1 des89, and PROPR1 des220 were characterized in terms of RNA-binding activity using gel mobility shift assay. This result showed that PRORP1 des38 and PRORP1 des89 have a slightly higher affinity for the pre-tRNA^{Phe} than PRORP1 des220. It is thus likely that PPR domains in PRORP1 play an important role in pre-tRNA binding (Figure 11).

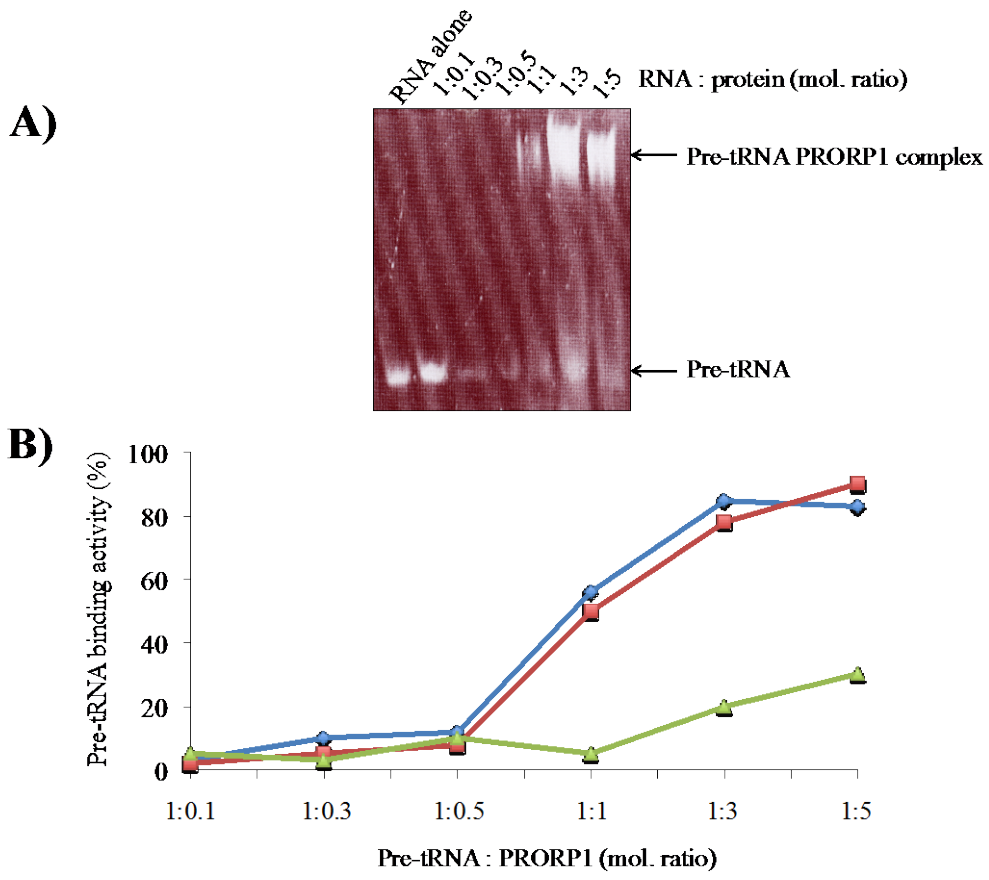

Figure 11: Gel mobility shift assay. (A) Gel mobility shift of pre-tRNA by adding an increasing amount of PRORP1 des38. The binding reaction contained the increased amount of PRORP1 des38 plus 2 μg of pre-tRNA (RNA : Protein (mol. ratio) = 1 : 0.1, 1 : 0.3, 1 : 0.5, 1 : 1, 1 : 3, and 1 : 5). After the electrophoresis, the gel was stained with ethidium bromide solution. The electric bands of pre-tRNA alone and pre-tRNA in complex with PRORP1 des38 were depicted as black arrows. (B) Each binding reaction contained the same number of molecules of PRORP1 des38, PRORP1 des89, or PRORP1 des220 plus 2 μg of pre-tRNA^{Phe}. The reaction buffer contains 30 mM Tris–HCl, pH 7.5, 100 mM NaCl,

10% glycerol, and 1 mM DTT. The reacted samples were run on a 4% native PAGE gel and run for 40 min at 150 V. The RNA molecules were detected by soaking with ethidium solution. The pre-tRNA binding activity was calculated as follows: the quantity of (the upper band, containing pre-tRNA in complex with PRORP1) × 100 / the quantity of (the upper band, containing pre-tRNA in complex with PRORP1 ＋ the lower band, containing pre-tRNA alone). PRORP1 des38: blue; PRORP1 des89: red; PRORP1 des220: green.

To obtain more insight into the function of PPR motifs in PRORP1, we prepared two mutants, PRORP1 des129 and PRORP1 des149, in which the 129 and 149 N-terminal residues comprising PPR1 (pos. 95-129), PPR1-2 (pos. 95-170), were deleted, respectively (Figure 2-10). Likewise, I attempted to produce PRORP1 des172 and PRORP1 des395, where PPR1 and 2 (pos. 95-170) and all of the PPR motifs and the central domain (pos. 95-394) were deleted, respectively. However, PRORP1 des172 and PRORP1 des395 could not express as soluble proteins in *E. coli* cells. Thus, the resulting mutant proteins, containing PRORP1 des129 and PRORP1 des149 were purified in the same manner as PRORP1 des89 and characterized with respect to pre-tRNA cleavage activity (Figure 2-10). The truncation of the first PPR motif (PPR1) and further successive deletions (PRORP1 des149 and PROPR1 des220) abolished the pre-tRNA cleavage activity (Figure 2-10). This result suggested that, at least, PPR1 plays a crucial role in substrate recognition and/or stabilization of PRORP1 for the pre-tRNA cleavage activity. To examine further the involvement of another PPR motif in PRORP1 in catalysis, we prepared two mutants, PRORP1 des89-PPR2 and PRORP1 des89-PPR3, in which residues 130-169 and 175-204 comprising PPR2 and PPR3, respectively, were deleted from PRORP1 des89 (Figure 2-10). The resulting mutants were characterized with respect to the pre-tRNA^{Phe} cleavage activity. Both mutants exhibited little pre-tRNA^{Phe} cleavage activity, indicating that PPR2 and PPR3 in addition to PPR1 play a crucial role in the recognition of pre-tRNAs in catalysis.

DISCUSSION

The catalytic mechanism of NYN domain adopts two metal ion mechanism

Howard and coworkers revealed the crystal structure of PRORP1 from *A. thaliana* (pos. 95-570) (Howard *et al.,* 2013). The structure revealed the catalytic pocket of PRORP1 where Asp399, Asp474, Asp475, and Asp493 are coordinated to the two manganese ions for catalysis *via* direct or *via* water-mediated hydrogen bonding.

Recently, they proposed the function of these aspartate amino acids in accord with the detailed biochemical assays based on the crystal structure of PRORP1 from *A. thaliana* (Howard *et al.,* 2015). They supposed that the carboxyl groups of Asp474 and Asp475 bind to two metal ions locating in the catalytic pocket of PRORP1. Furthermore, Asp399 binds and promotes one water molecule for nucleophilic-attack targeting to scissile phosphodiester bond in its RNA substrates. Asp493 binds the main chain of Met495 and plays an important role in formation of the catalytic pocket of PRORP1. According to their crystal structure, the carboxyl group of Asp497 flipped out without the coordinating two metal ions, whereas in the case of existing the metal ion (*e.g.* manganese ion) located at the catalytic pocket, the carboxyl group of Asp497 binds one of them. It is thus likely that Asp497 binds the metal ion to arrange and facilitate its catalytic efficiency (Figure 12) (Howard *et al.,* 2013). These results were in good agreement with our biochemical assay using PROPR2 desN9 D343A mutant conducted at 28 °C, indicating the c

thaliana. As described above, the NYN metallonulcease domain shares a structural homology to the PIN (PilT N terminus) and flap nuclease families. Recent studies revealed that PIN and Flap nucleases adopt two-metal ion mechanism as their catalytic mechanism (Orans *et al.,* 2011). Taken together, it could be thus assumed that the catalytic mechanism of NYN metallonuclease domain in PRORPs in *A. thaliana* is similar to that of PIN and Flap nuclease family.

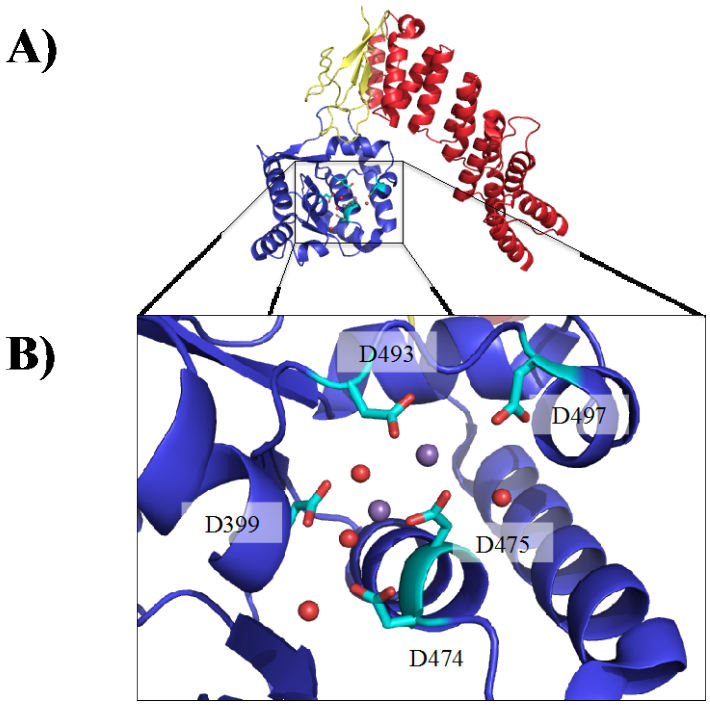

Figure 12: The proposed catalytic mechanism of PRORP1 from *A. thaliana*. (A) The overall structure of PRORP1. Its PPR domain, central domain, and NYN domain are colored as red, yellow, and blue, respectively. (B) The catalytic pocket of PRORP1, where Asp399, Asp474, Asp475, Asp493, and Asp497, are coordinated to the two manganese ions necessary for its catalysis. These aspartate residues are depicted as stick representation and colored in cyan. The coordinated water molecules are depicted as red spheres and the coordinated manganese ions are represented by silver spheres. These residues bind the two metal ions *via* directed or water-mediated hydrogen bonding network (Howard *et al.,* 2013).

The helical structure located on the region (pos. 38-89) plays an important role in stabilization of the PPR motif in PRORP1.

In general, the polypeptide immediately preceding or following an α-helix is known as Nt- and Ct-capping, respectively, and plays a crucial role in stabilization of α-helices due to lack of inter-helical hydrogen bonds in the first and last turn. Moreover, terminal helices are known to contribute to ligand specificity for repeat proteins such as tetratricopeptide repeat, which can be found in all domains of life, and then regulate various biological processes (*e.g.* organelle targeting, protein import, vesicle fusion, and biomineralization; they are referred to as a capping motif (Zeytuni and Zarivach, 2012; D' Andrea and Regan, 2003). It is thus likely that the polypeptide chain located in pos.38-89 may an important role in stabilization of the PPR motifs in PRORP1, serving as the capping motif in PRORP1. This assumption was supported by CAPS-DB (http://www.bioinsilico.org/CAPSDB) (Segura *et al.,* 2011), a database to search Nt- or

Ct-capping structure that predicted the polypeptide to form a capping helix (Figure Ⅱ-23). Moreover, the gel shift assay, carried out by Dr. Teramoto (NIH/NIEHS), revealed that the pre-tRNA^{Phe} binding affinity of PRORP1 des92 was slightly dicreased (*K*d value = 6-10 μM) rather than that of PRORP1 des76 (*K*d value = 1-3 μM). It is thus likely that the stabilization and proper arrangement of each PPR motif in PRORP1 by the N-terminal region (pos.76-89) were importance of binding chloroplast pre-tRNA^{Phe} from *A. thaliana.*

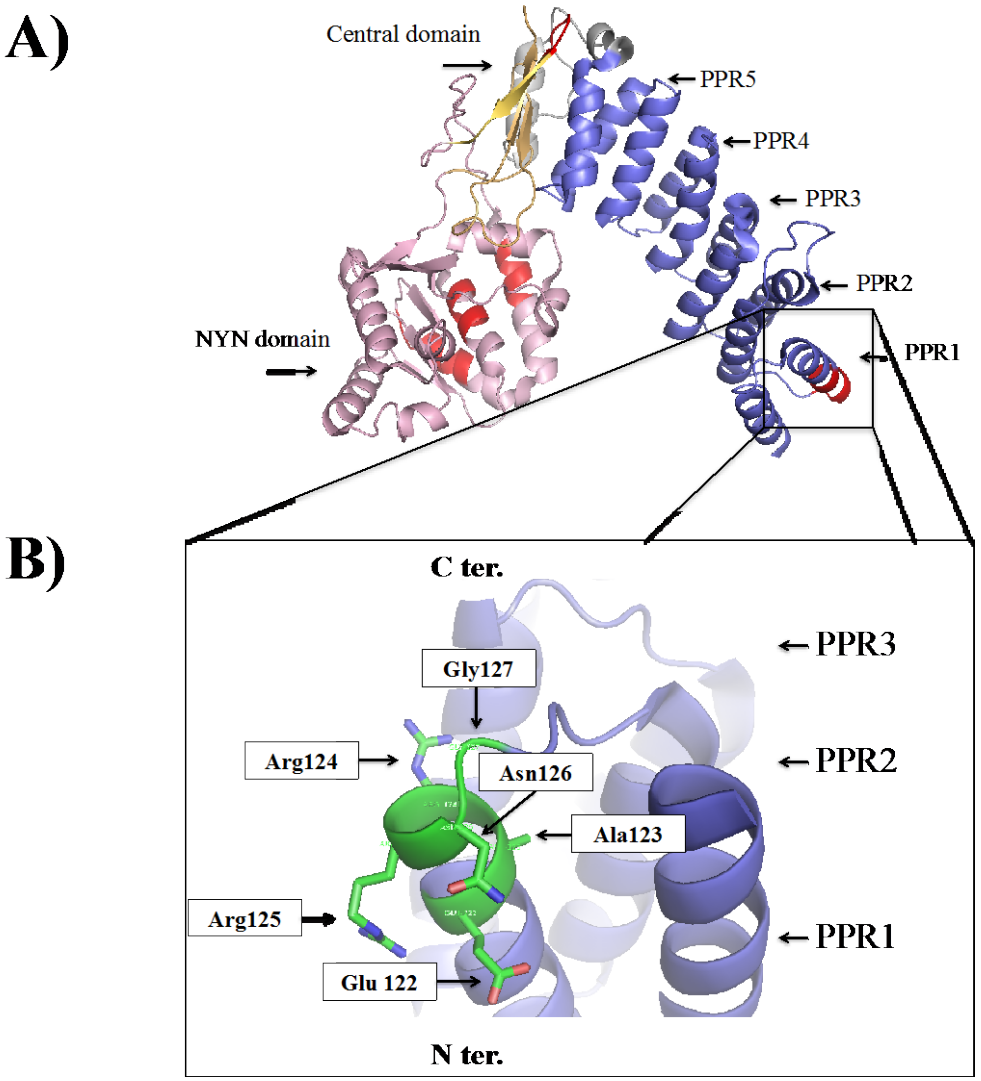

Figure Ⅱ-23: The PPR1 motif may serve as the capping motif in PRORP1. (A) The Nt- and Ct-capping motifs in PRORP1 was identified by CAPS-DB (http://www.bioinsilico.org/CAPS-DB) and illustrated on the crystal structure of PRORP1 from *A. thaliana*. The PPR domain, the central domain, and the NYN domain are colored blue, yellow, and pink, respectively. The proposed Nt- and Ct- capping structure in PRORP1 is colored red. (B) The predicted Ct-capping structure (pos. 122-127) in PPR1 motif is colored green. The amino acid residues located on the region are illustrated by stick representation.

SUMMARY

The recombinant PRORP1 des38, PRORP2 desN9, and PRORP3 desN10 were overproduced in *E. coli* BL21 (DE3) codon Plus RIL and purified to homogeneity by a series of column chromatography. Approximately 0.5 mg of PRORP1 des38 was obtained from 1 liter of culture broth. Likewise, 0.3-0.5 mg of PRORP2 desN9 and the same amount of PRORP3 desN10 were obtained from 1 liter of culture broth. The purified enzymes exhibited the pre-tRNA processing activity *in vitro*. It is suggested that the expression and purification system of PRORP1, PRORP2, and PRORP3 is sufficient to study the function and structure relationship of PRORPs in *A. thaliana.* To characterize the function of the catalytic residues in PRORPs, PRORP1 des38 D474A/D475A and PRORP2 D343A were overexpressed in *E. coli* cells and purified with column chromatography in the same manner as those wild-type proteins. Furthermore, the five mutants containing Asp399, Asp474, Asp475, Asp493, and Asp497 in PRORP1, where these amino acids were individually replaced with Ala, were prepared and characterized with the pre-tRNA processing activity. These mutants exhibited little enzymatic activity, indicating that the conserved aspartate amino acids in PIN and Flap nuclease family as well as NYN metallonuclease family are important to cleave pre-tRNAs at the 5' end. Taken together, these results showed that the catalytic mechanism of NYN metallonuclease domain in PRORP1 possibly utilizes two metal ion mechanism as PIN and Flap nucleases do.

During the purification of PRORP1 des38, the peptide bond locating at Lys-Gly (pos. 89-90) was highly sensitive to proteolysis, resulting in producing the polypeptide fragment where the N-terminal residues (pos. 38-89) were deleted. Limited trypsin and chymotrypsin proteolysis with PRORP1 des38 showed that Arg220 in addition to Lys89 are highly sensitive. Hence, I constructed the expression vector, harboring PRORP1 des89 and PRORP1 des220, where the N-terminal regions pos. 1-89 and pos. 1-220 were truncated, respectively, in order to characterize PPR motifs in PRORP1. I introduced these constructs into the host *E. coli* cells BL21 (DE3) Codon Plus RIL. Additionally, the successive deletion mutants, including PRORP1 des129, PRORP1 des149, where the PPR1 (pos. 95-129) and PPR1-PPR2 (pos. 95-170) were truncated, respectively, were prepared in the same manner as PRORP1 des89. These purified

proteins were characterized in terms of the pre-tRNA^{Phe} processing activity and binding activity. These results indicate that the N-terminal fragment containing pos. 76-89 plays an important role in stabilization and proper arrangement of PPR domain of PRORP1. Additionally, this region is likely to be partially contributed to pre-tRNA processing activity and binding affinity. Furthermore, at least, PPR1, PPR2, and PPR3 located in PRORP1 play a crucial role in binding its substrate, pre-tRNA^{Phe}.

Chapter Ⅲ

Mutation Analysis of PRORP1 and pre-tRNAs to Elucidate the Mechanism for Substrate Recognition

INTRODUCTION

The RNA binding motifs that are composed of 2-3 α helices (*e.g.* Pumilio and FBF homology domain, PUF domain; Pentatricopeptide repeat, PPR) are known to recognize the specific bases on the single stranded RNA in a modular fashion (Wang *et al*., 2002; Yin *et al*., 2013). As shown in Figure Ⅱ-3 and Figure Ⅲ-1, PPR motifs and PUF domain adopt the similar fashion in RNA base recognition (Figures Ⅱ-3 and Ⅲ-1). Two hydrophobic residues or arginine residues bind the specific RNA base *via* hydrophobic interactions in sandwich-like manner, and then the polar residues serve in improving RNA recognition specificity *via* Watson-Crick base pairing-like hydrogen bonding (Wang *et al.,* 2002; Yin *et al.,* 2013).

Recently, it has become increasingly evidence that other helical repeat proteins, including half-a-tetratricopeptide repeat (HAT), octotricopeptide repeat proteins (OPR), and mitochondrial transcription termination factors (mTERF), play a crucial role in RNA metabolisms (*e.g.* mRNA stabilization, RNA editing, DNA replication, and RNA processing) (Hammari *et al.,* 2014). Genomic and functional studies revealed that although the primary sequence of these helical repeats proteins are not conserved, the secondary structure and their arrangement are highly conserved. It could thus be assumed that despite these proteins are evolutionally not related to each other, these helical repeat motifs adopt the same RNA-recognition mechanism (Hammari *et al.,* 2014). Up to now, among the helical repeat proteins, the PPR family has clearly attracted the most attention. Especially, its mechanism for specific target selection has been identified with the PPR recognition code as described below in detail (Yagi *et al.,* 2013a; Yagi *et al.,* 2013b; Barkan *et al.,* 2012).

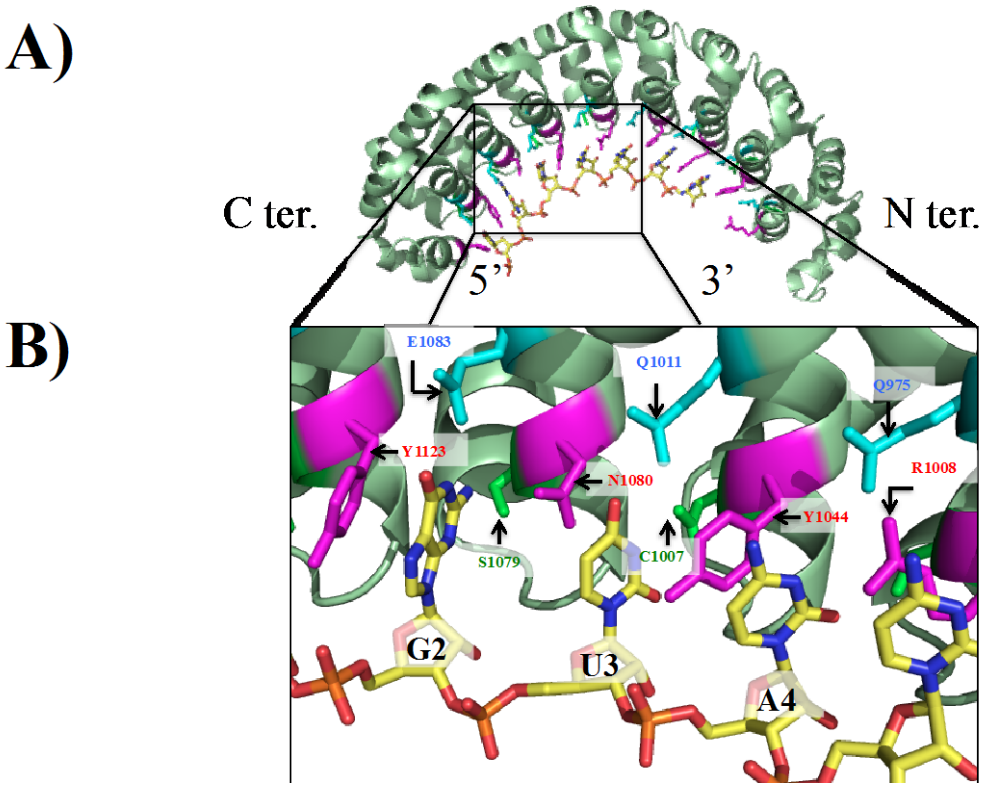

Figure 1: RNA recognition mode by PUF proteins. (A) The structure of the human PUF protein PUM1 (also known as HSPUM) bound to the RNA element NRE1-19 (PDB ID: 1M8W). (B) Recognition of G2, U3, and A4 by PUF repeat 5, 6, and 7 in the crystal structure of PUM1, respectively (Wang *et al*., 2002). The two polar residues (*green* and *cyan*) serve as the important determinant for base discrimination. The residues (colored *magenta*) are involved in stacking interactions between two adjacent bases.

PPR proteins are eukaryotic RNA-binding proteins that are involved in post-transcriptional gene regulation (Nawel *et al.,* 2015). The crystal structure of PPR10 in complex with its RNA substrate revealed that their PPRs are found in tandem and composed of a helix turn helix fold of ~35 amino acids (Yin *et al.,* 2013). Furthermore, each PPR repeat binds to a single RNA base in this structure (Figures 2 and 3). The N-terminal region binds to the 5'-nucleotide residue of the target sequence, while the C-terminal region interacts with the 3' terminal residues (Figures 2 and 3) (Yin *et al.,* 2013). Nakamura *et al.* proposed that residues at positions 4 and -1 in each PPR repeat directly interact with a Watson-Crick edge of a base, whereas the residue at position 1 is involved in a stacking interaction between two advancement bases (Nakamura *et al.,* 2012) (Figures 2 and 3). Furthermore, in accord with the *in silico* bioinformatic studies based on the previously reported biochemical studies, they postulated that a PPR

recognition code, where residues at positions 1, 4, and -ii in each repeat contribute to specific recognition of a nucleotide base (Nakamura *et al.,* 2012; Yagi *et al.,* 2013a; Yagi *et al.,* 2013b). For instance, Tyr at 1, Asn at 4, and Asn at -ii recognize cytosine, Val at 1, Asn at 4, and Asp at -ii recognize uracil, Phe at 1, Thr at 4, and Asn at -ii recognize adenine, and Val at 1, Thr at 4, and Asp at -ii recognize guanine (Figure III-2) (Yagi *et al.,* 2013a; Yagi *et al.,* 2013b).

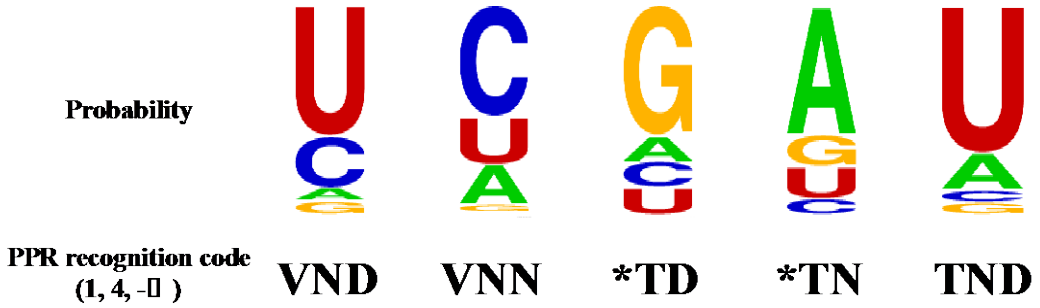

Figure III-2: Examples of the nucleotide frequency according to the PPR recognition code. The PPR recognition code (residues: 1, 4, and -ii) determines the nucleotide frequency. These examples were referred to the table S4 in Yagi *et al.* (2013). * represents any amino acids.

As described in chapter II, I found that a PRORP1, play a crucial role in binding its substrate, pre-tRNA^{Phe}. On the basis of the proposal recognition code, I attempted to predict the target sequence by PPR motifs in PRORP1.

MATERIALS AND METHODS

Materials

Mutagenesis were performed with the Prime star Master mix kit supplied by Takara Bio. DNA polymerases were purchased from MBI Fermentas. All other common chemicals and reagents were purchased at the highest purity available.

Prediction of a target site for PPR motifs in PRORP1

The nucleotides-specifying residues (denoted NSRs; residues 1, 4, and -ii) were extracted from the PPR motifs of PRORP1, PROPR2, and PRORP3, according to the tertiary structure of PRORP1. The NSRs were converted into a probability matrix that indicated the decoding nucleotide frequency according to the PPR code, as described by

Yagi *et al.* (Yagi *et al.,* 2013a; Yagi *et al.,* 2013b; Barkan *et al.,* 2012). The probability matrix was also shown using WebLogo (http://weblogo.threeplusone.com/create.cgi).

Construction of pre-tRNA mutants and pre-nad6 mutants

The pUC19 vector harboring the cDNA fragment encoding chloroplast pre-tRNA^{Phe} from *A. thaliana* was obtained. The specific primers, where target nucleotides were replaced with another three nucleotide, were used to amplify the target vectors (Table Ⅲ). The resulting PCR products were introduced into *E. coli* cells JM109 for cloning plasmids, after which the plasmids were obtained and their homogeneities were checked by agarose gel (1%) electrophoresis. The DNA fragments encoding pre-tRNA^{Phe} mutants were amplified by PCR using their plasmids as templates. Subsequently, their RNAs were transcribed *in vitro* using the amplified gene fragments as templates. The resulting RNAs were purified by a DEAE Sepharose column, followed by a HiLoad 16/60 Superdex200 gel filtration column. Mutant pre-nad6 RNAs were prepared in the same manner as described for mutant pre-tRNAs.

Overexpression and purification of the PRORP1des38 NSR mutants

The expression vectors encoding a various NSR point mutations of the PRORP1 des38 cDNA were introduced into the *E. coli* cells BL21 (DE3) Codon Plus RIL. The mutant proteins were expressed in the cells by adding 1 mM IPTG. Their purification was carried out in the same manner as PRORP1 des38 wild-type, as described in Chapter Ⅱ. The purchased oligonucleotide primers used in mutagenesis wee shown in Table Ⅲ (Table Ⅲ).

Assay for pre-tRNA mutants and pre-nad6 mutants cleavage.

Cleavage reactions of pre-tRNA^{Phe} and pre-nad6 mutants were performed with using 2 μg of transcripts by an increasing amount of the enzymes for 60 min at 37 ℃, as described in MATERIALS AND METHODS in Chapter Ⅱ.

Table □ -1 : Oligonucleotides used for mutagenesis of PRORP1, pre-tRNA, and pre-nad6 in this study

Primers	Sequences of forward primers (5'-3')	Sequences of reverse primers (5'-3')
Pre-tRNA C56A	5' - ACCAGUUAAAAUCUGGUUCUUGG -3'	5'-CAGATTTTAACTGGTGACACCAG-3'
Pre-tRNA C56G	5'- ACCAGUUGAAAUCUGGUUCUUGG -3'	5'-CAGATTTCAACTGGTGACACCAG -3'
Pre-tRNA C56U	5'- ACCAGUUUAAAUCUGGUUCUUGG -3'	5'- CAGATTTAAACTGGTGACACCAG-3'
Pre-tRNA A57G	5'- CCAGUUCGAAUCUGGUUCUUGGC -3'	5' - CCAGATTCGAACTGGTGACACCA-3'
Pre-tRNA A57C	5' –CCAGUUCCAAUCUGGUUCUUGGC - 3'	5' –CCAGATTGGAACTGGTGACACCA- 3'
Pre-tRNA A57U	5' –CCAGUUCUAAUCUGGUUCUUGGC - 3'	5' –CCAGATTAGAACTGGTGACACCA- 3'
Pre-tRNA A58G	5' –CAGUUCAGAUCUGGUUCUUGGCC - 3'	5' –ACCAGATCTGAACTGGTGACACC- 3'
Pre-tRNA A58C	5' –CAGUUCACAUCUGGUUCUUGGCC- 3'	5' –ACCAGATGTGAACTGGTGACACC- 3'
Pre-tRNA A58U	5' –CAGUUCAUAUCUGGUUCUUGGCC- 3'	5' –ACCAGATATGAACTGGTGACACC- 3'
Pre-tRNA A59G	5' –AGUUCAAGUCUGGUUCUUGGCCC- 3'	5' -AACCAGACTTGAACTGGTGACAC- 3'
Pre-tRNA A59C	5' –AGUUCAACUCUGGUUCUUGGCCC- 3'	5' –AACCAGAGTTGAACTGGTGACAC- 3'
Pre-tRNA A59U	5' – AGUUCAAUUCUGGUUCUUGGCCC- 3'	5' –AACCAGAATTGAACTGGTGACAC- 3'
Pre-nad6 C38G	5' –AUUGGUUGGAAUCCAAUUCGUGA- 3'	5' –UGGATTCCAACCAATCAAGCTAC- 3'
Pre-nad6 G39C	5' –UUGGUUCCAAUCCAAUUCGUGAG- 3'	5' –UUGGATTGGAACCAATCAAGCTA- 3'
Pre-nad6 A40C	5' –UGGUUCGCAUCCAAUUCGUGAGC- 3'	5' –AUUGGATGCGAACCAATCAAGCT- 3'
Pre-nad6 A41C	5' –GGUUCGACUCCAAUUCGUGAGCC- 3'	5' –AAUUGGAGTCGAACCAATCAAGC- 3'
PRORP1 N136T	5'- CCATTATACTGTGTTGCTTTATG 3'	5' –AAAGCAACACAGTATAATGGTACTGGCT- 3'
PRORP1 T180N	5' –TACATTTAACAACGGTGCTAGACTC- 3'	5' –CACCGTTGTTAAATGTAGCTTCATT- 3'
PRORP1 G215N	5' –TATGGAACTGCCCTTTTTGGTTT- 3'	5' –GGGCAGTTCCATAGGATCTCAA- 3'

Mutation sites are indicated by underlines.

RESULTS

Prediction of a target site for PPR motifs in PRORP1

On the basis of the proposed recognition code for PPR motifs, I attempted to predict the target sequence by PPR motifs in PRORP1. Since I expected that the target sequence for PPR motifs in PRORP1 should be identical to those for its isozymes, PRORP2 and PRORP3, their five PPR motifs were first modeled using the crystal structure of PRORP1 as a template by SWISS-MODEL (http://swissmodel.expasy.org) (Kiefer *et al.,* 2009; Jürgen *et al.,* 2004) (Figure Ⅲ-B). Next, their NSRs were extracted from the model structures and depicted as Logo representation using WebLogo (http://weblogo.threeplusone.com/create.cgi) (Figures Ⅲ-4 and 5).

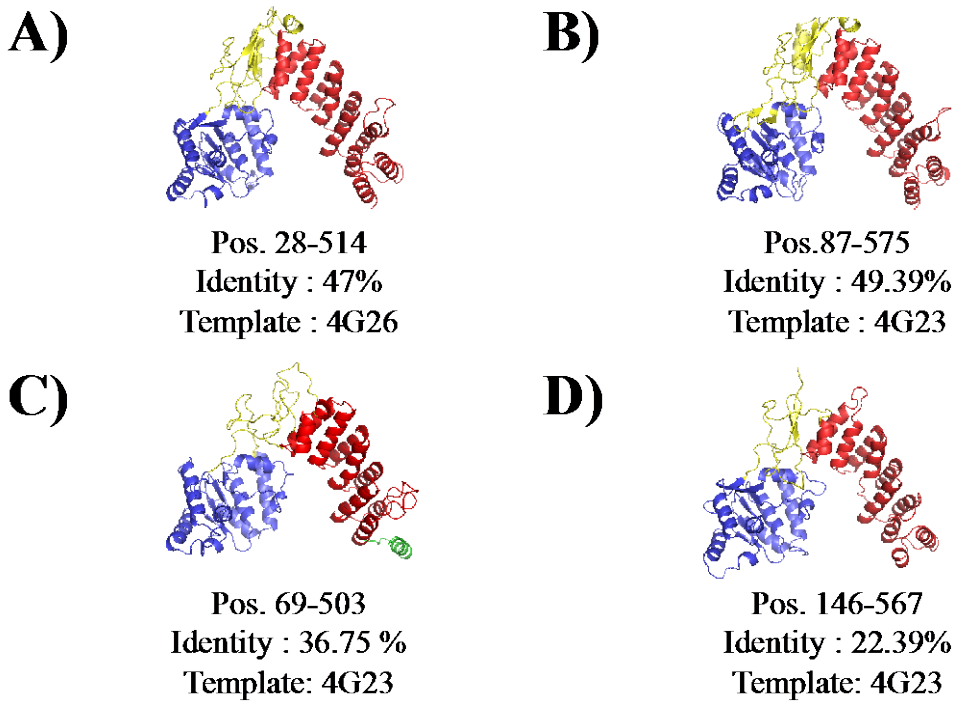

Figure Ⅲ-3: Construction of the model structures of the four PRORPs. I constructed the model structures of (A) AthPRORP2, (B) AthPRORP3, (C) OtPRORP, and (D) human MRPP3 by SWISS-MODEL (http://swissmodel.expasy.org) using the crystal structure of AthPRORP1 (PDB ID: 4G23 or 4G26) as a template. These PRORPs are known to have the pre-tRNA processing activity by previous studies (Lai *et al.,* 2011; Gobert *et al.*, 2010; Gutmann *et al.*, 2012; Holzmann *et al.,* 2008). Ath: *Arabidopsis thaliana*; Ot: *Osterococcus tauri.* A PPR domain, a central domain, and a NYN metallonuclease domain are colored red, yellow, and blue, respectively. In Figure Ⅲ-C, one of the two helical structures in PPR1 could not be predicted. The other helical structure of PPR1 is depicted in green.

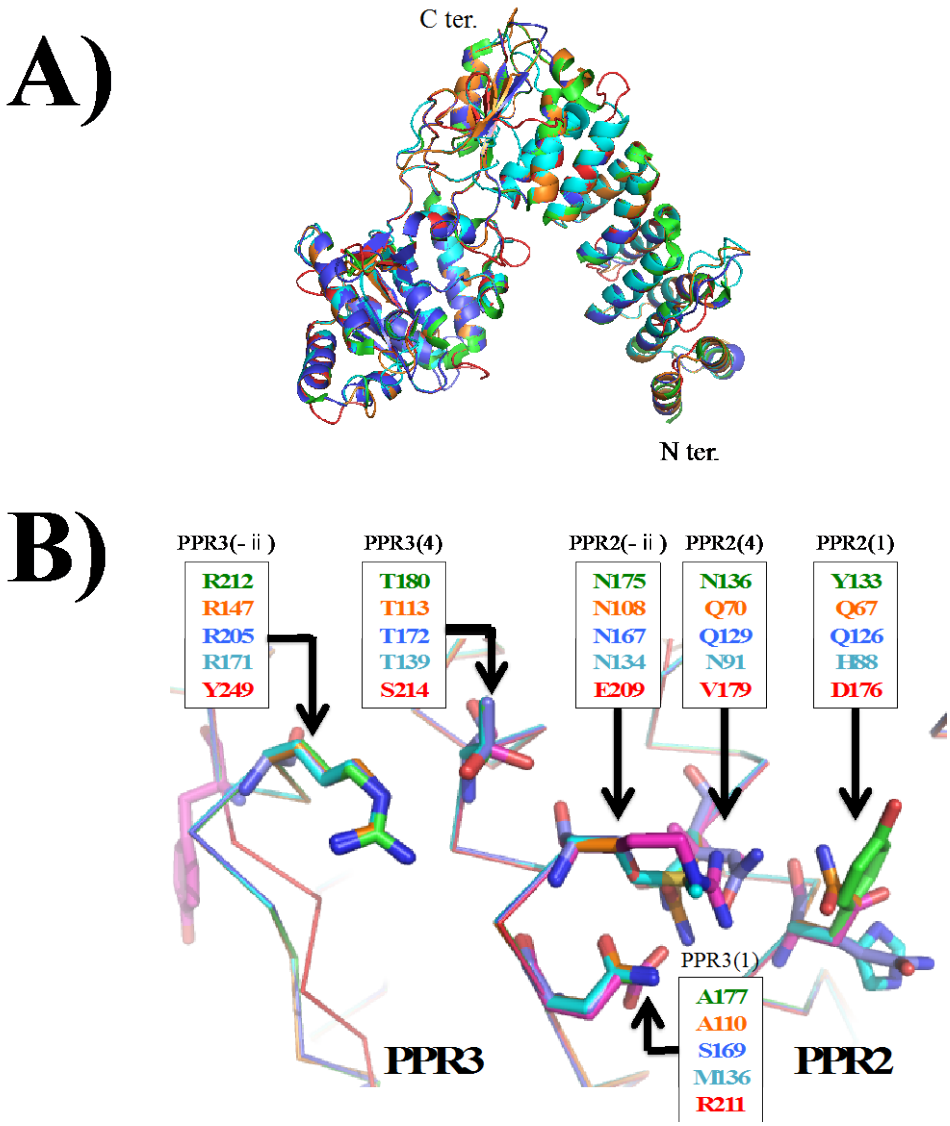

Figure Ⅲ: PPR2 and PPR3 motifs in PROPRs are conserved among divergent species. (A) Superimposition the model structures of the four PRORPs on the crystal structure of PRORP1 from *A. thaliana*. AthPRORP1, AthPRORP2, AthPRORP3, OtPRORP, and human MRPP3 are colored green, orange, blue, cyan, and red, respectively. These structures were superimposed by secondary structure matching (SSM) in Coot, a new tool for fast protein structure alignment in three dimensions (Emsley *et al*., 2010). (B) Conservation of the NSRs in PPR2 and PPR3 motifs among the five PRORPs. The residues at position 1, 4, - ii of PPR2 repeat and those of PPR3 repeat are described inside the specific black flame. The NSRs of AthPRORP1 are colored green. Likewise, the NSRs of AthPRORP2, AthPRORP3, OtPRORP1, and human MRPP3 are colored orange, blue, cyan, and red, respectively.

A)

B)

PPR	1	2	3	4	5
NSR	AKS	YNN	ATR	RGE	PAS
code	***	*NN	*T*	*G*	*AS
A	ND	0.15	0.45	0.32	0.31
C	ND	0.45	0.09	0.13	0.08
G	ND	0.14	0.31	0.26	0.06
U	ND	0.27	0.15	0.30	0.56

Figure 3: Prediction of a target site in tRNA for PPR motifs in PRORP1. (A) The conservations of the amino acids that determine target RNA specificity for PRORPs. The nucleotide-specifying residues (NSRs; 1, 4, and - ii) of PRORP1, 2, and 3were shown for each PPR motif (PPR1-PPR5) by logo matrix. (B) The binding RNA specificity of PPR motifs in PRORPs. The extracted NSRs were converted to the probability matrix that indicates the target RNA sequences indicated as "code" (Yagi *et al.*, 2013a ; Yagi *et al*., 2013b). The nucleotide occurrence frequency is also represented by the logo. The asterisks indicate any amino acid.

A polar amino acid residue located at the 4 position is supposed to be the most important determinant for the purine/pyrimidine discrimination (Nakamura *et al.,* 2012; Yagi *et al.,* 2013a; Yagi *et al.,* 2013b). Although the NSRs were divergent among PRORP1-3, PPR2 and PPR3 display conserved Asn/Gln and Thr at the residue 4. In contrast, rare amino acids, Ala/Gly and Ser/Ala in the NSRs were present at the residue 4 in PPR4 and PPR5 (Figures 3 and 5). This result suggested the importance of PPR2 and PPR3 for Arabidopsis PRORP's function. Furthermore, in order to gain more information about NSRs of each PPR motif in PRORP1, I also constructed the model structure of the PRORP from the moss, *Osterococcus tauri,* which is known to have the pre-tRNA processing activity (Lai *et al.,* 2011) (Figures 3 and 4). Subsequently, I

selected each NSR of OtPRORP according to the resulting model structure (Figures 3 and 4). Although the NSRs were divergent among OtPRORP, PPR2 and PPR3 displayed conserved Asn and Thr at the 4th position. In contrast, rare amino acids Ala and Phe in the NSRs were present at the residue 4 in PPR4 and PPR5, respectively. It is thus likely that the 4th position of PPR2 and PPR3 are almost conserved among divergent species, indicating the importance of PPR2 and PPR3 for PRORP's function. Furthermore, it could be assumed that PRORPs from plant adopt similar recognition mechanism. Subsequently, the NSRs from PROPR1 were conserved into the probability matrix that indicates the target RNA sequences for PRORPs (Figure 5). As shown in Figure 5, the possible nucleotides for each determinant for PPR2, PPR3, PPR4, and PPR5 of PRORP1 were predicted to be C, A/G, A/U, and U, respectively (Figure 5). In contrast, no nucleotides were assigned for PPR1, which may support the assumption that PPR1 may play a distinct function from the other PPR motifs, as described in Chapter II in this thesis.

target nucleotides C, A/G, A/U, and U were compared with those in pre-tRNA^{Phe}, they were in good agreement with the sequence, C56-A57-A58-U59, at the TψC loop in the pre-tRNA^{Phe} (Figure 6). Furthermore, the predicted target sequences also were in good agreement with another substrate of PRORP1, pre-nad6, which is called as t-element serving as signal for endonucleolytic cleavage by RNase Z and RNase P in post-transcription processing (Forner *et al.,* 2007) (Figure 6). It is thus likely that the PPR2, PPR3, PPR4, and PPR5 motifs of PRORP1 recognize C56, A/G57, A58, and U59 at TΨC loop of their substrates in a modular fashion, respectively (Figure 6) (Imai *et al.,* 2014).

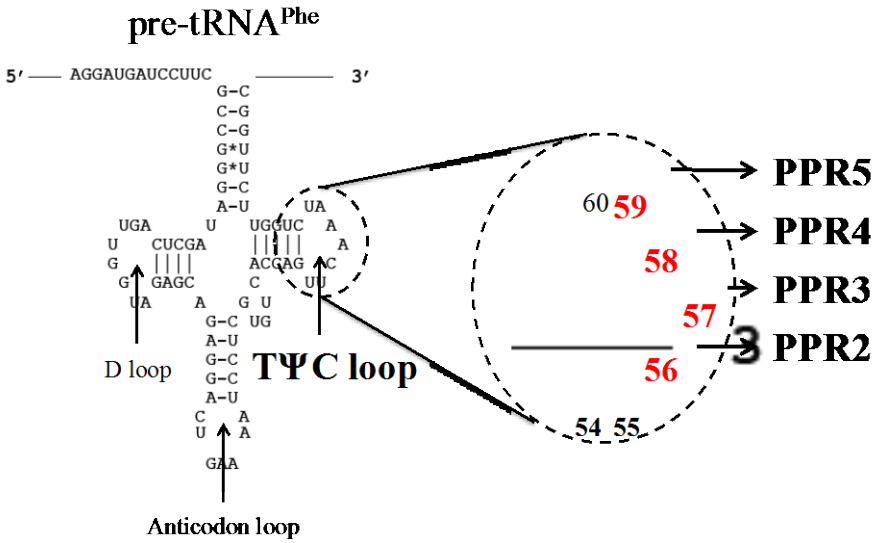

Figure Ⅲ: Proposed pre-tRNA recognition mode of PRORP1 according to the prediction method using NSRs. The secondary structure of pre-tRNA^{Phe} is represented (left). Its D loop, TΨC loop, and anticodon loop are marked on its predicted structure by black arrows. Close-up view of the TΨC loop structure (right). C56, A57, A58, and A59 on this region are highlighted by red. Based on the prediction method using NSRs, it is likely that C56, A57, A58, and A59 are recognized by PPR2, PPR3, PPR4, and PPR5, respectively.

Mutational analysis of PRORP1 and its RNA substrates

To evaluate our prediction, I prepared pre-tRNA^{Phe} mutants where C56, A57, A58, and U59 at the TΨC loop were in turn replaced with three other nucleotides. The resulting mutant substrates were characterized with respect to their susceptibility to PRORP1 in the same manner as those described above (Figure Ⅲ-7).

Replacements of C56 and A58 with the three other nucleotides significantly reduced the susceptibility to PRORP1, ranging from 15% to 40% of that of the pre-tRNA^{Phe} wild-type. In addition, replacing A57 with C or U reduced their susceptibility to PRORP1 (~35%), whereas that with G had little efficiencies on the susceptibility (62%), consistence with the fact that the base at position 57 is semi-invariant as a purine. In contrast, the pre-tRNA^{Phe} mutants of A59 retained their susceptibility to PRORP1. Since it is known that the base at position 59 is variable, these results suggested that the invariant base C56, A57, and A58 at the TΨC loop participate in the recognition by PRORP1.

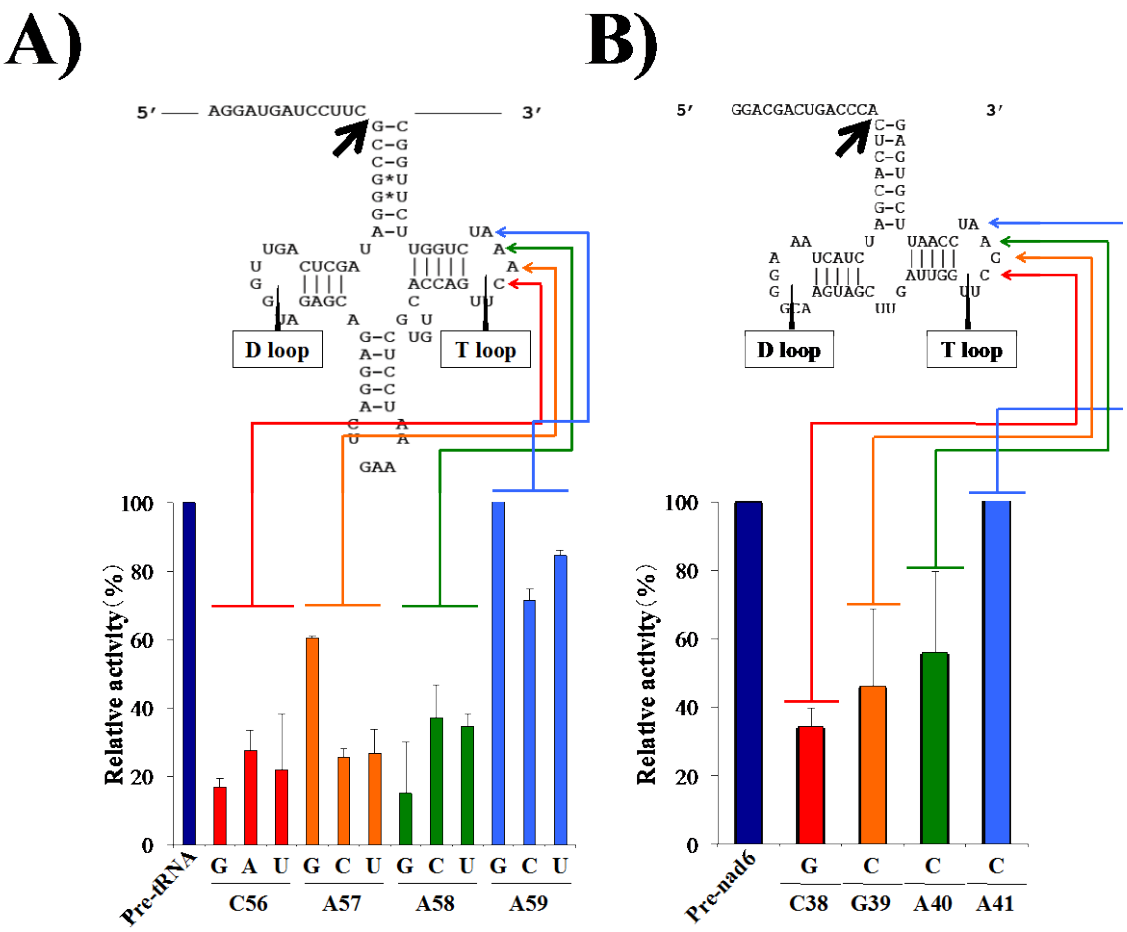

Figure Ⅲ: The susceptibility of its RNA substrates and their mutants. (A) The susceptibility of pre-tRNA^{Phe} and its mutants. The proposed secondary structure of pre-tRNA^{Phe} is described. The black arrow represents the cleavage site by PRORP1. (B) The susceptibility of pre-nad6 and its mutants. The secondary structure of pre-nad6 are illustrated. The black arrow indicates the cleavage site by PRORP1.

To confirm this issue further, I prepared pre-nad6 mutants and characterized them in terms of RNA processing assay (Figure Ⅲ-B). The pre-nad6 mutants, including C38, G39, and A40 in turn replaced with G, C, and C, respectively, reduced their susceptibility (30-50%), whereas the pre-nad6 mutant, A41 in turn replaced with C, retained the susceptibility (100%) (Figure Ⅲ-B). It is thus suggested that PRORP1 recognizes tRNA-like substrates in non-coding RNAs and 3'/5' untranslated region (UTR) of messenger RNAs. However, it should be noted that Watson-Crick base pairing of A58 with T54 is involved in stabilization of the TΨC loop in tRNAs (Shi and Moore, 2000). Thus, I could not rule out the possibility that the replacement of A58 causes a conformation change that results in the insusceptibility of PRORP1. Recently, structural

analysis of the maize chloroplast protein PPR10 in complex with its target RNA has revealed that the side chain of the 4 residue in each PPR repeat plays an essential role in recognition, namely, hydrogen bonding to the bases (Figures Ⅱ-2 and 3) (Yin *et al.,* 2013). To confirm our prediction further, I prepared PRORP1 mutants, in which amino acids Asn136, Thr180, and Gly215 at the 4 position in PPR2, PPR3, and PPR4, respectively, were in turn replaced by another amino acid; the resulting PRORP1 mutants were again characterized with respect to pre-tRNA^{Phe} cleavage activity (Figure Ⅲ-8).

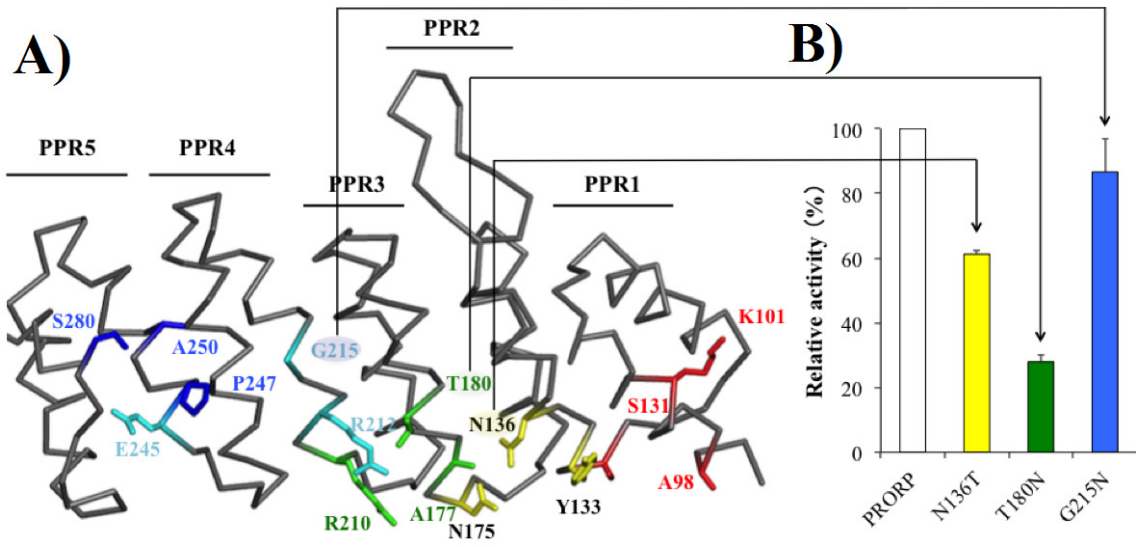

Figure Ⅲ-8: The pre-tRNA cleverage activity of PRORP1 and its NSR mutants. (A) The NSRs of each PPR repeat in PRORP1 are depicted as stick representation. The NSRs of PPR1, PPR2, PPR3, PPR4, and PPR5 are colored red, yellow, green, cyan, and blue, respectively. The mutated amino acids, Asn136, Thr180, and Gly215 were in turn replaced with Thr, Asn, and Asn, respectively. (B) The relative pre-tRNA processing activity of PRORP1 NSR mutants. The column of N136T is colored yellow. Those of T180N and G215N are colored green and blue, respectively.

As shown in Figure Ⅲ-8, mutations of Asn136 in PPR2 and Thr180 in PPR3 moderately reduced the pre-tRNA^{Phe} cleavage activity (25%-60%), while that of Gly215 in PPR4 had little influence on the cleavage activity (90%). In accord with the prediction and mutation analysis, PPR2 and PPR3 motifs in PRORP1 are favorable to recognize the highly conserved nucleotides bases C56 and A57 at the TΨC loop in pre-tRNA^{Phe} in a modular fashion (Imai *et al.,* 2014).

Construction of a fitting model of PRORP1 in complex with tRNA

I attempted to construct a model of PRORP1 in complex with tRNA on the basis of biochemical data and the computational prediction using the molecular modeling program MOE (2013. 03; Chemical Computing Group Inc., Montreal, Canada). To this end, the following data were taken into consideration. First, the metallonuclease domain of PRORP1 contains two adjacent aspartate residues, Asp474 and Asp475, which are involved in coordination magnesium ions important for catalysis (Gobert *et al.,* 2010). The 5' end of tRNA should thus be located near the two-aspartate residues. Second, it has been reported that bases Ψ55, C56, and A57 on the TΨC loop undergo a conformational change upon pseudouridine synthetase binding and be unstacked and splay putward so as to interact with the enzyme (Pan *et al.,* 2003). I thus copied the coordinates of the Ψ55-C56-A57 on the TψC loop bound psudouridine synthetase to fit tRNA into PPR motifs in PRORP1. Third, the structural basis of sequence-specific recognition by PPR proteins has been presented by the structural analysis of the maize chloroplast protein PPR10 in complex with its target RNA (Figure 3) (Yin *et al.,* 2013). In this structure, each of the four bases is sandwiched by two residues at the 1 positions of adjacent repeats and the side chain of the 4 residues in each PPR repeat hydrogen bond to the bases. The structure of PPR2-PPR4 in PRORP1 could be superimposed on that of PPR motifs in PPR10, suggesting that it recognizes bases in a same manner as PPR10. Finally, PRORP1 has two arms composed of the NYN metallonuclease domain and the PPR domain, which are attached to the linker domain comprising zinc ion. The two arms are likely to be highly flexible in solution and to change their orientation upon binding to pre-tRNA^{Phe}. We took these data into consideration to orient the crystal structure of tRNA relative to the PRORP1 structure (Figure 4). In this model, the side-chains of Tyr133, Asn136, and Asn175 at PPR2 recognize C56, and those of Ala177, Thr180, and Arg213 in PPR3 are located in close approximately to A57, where the side-chains of Asn136 and Thr180 at the 4 positions hydrogen bond to C56 and A57, respectively. In contrast, Arg212, Gly215, and Glu245, possible NSRs in PPR4, are located far from A57 at TψC loop, supporting our assumption that PPR2 and PPR3 play an important role in recognition of pre-tRNAs.

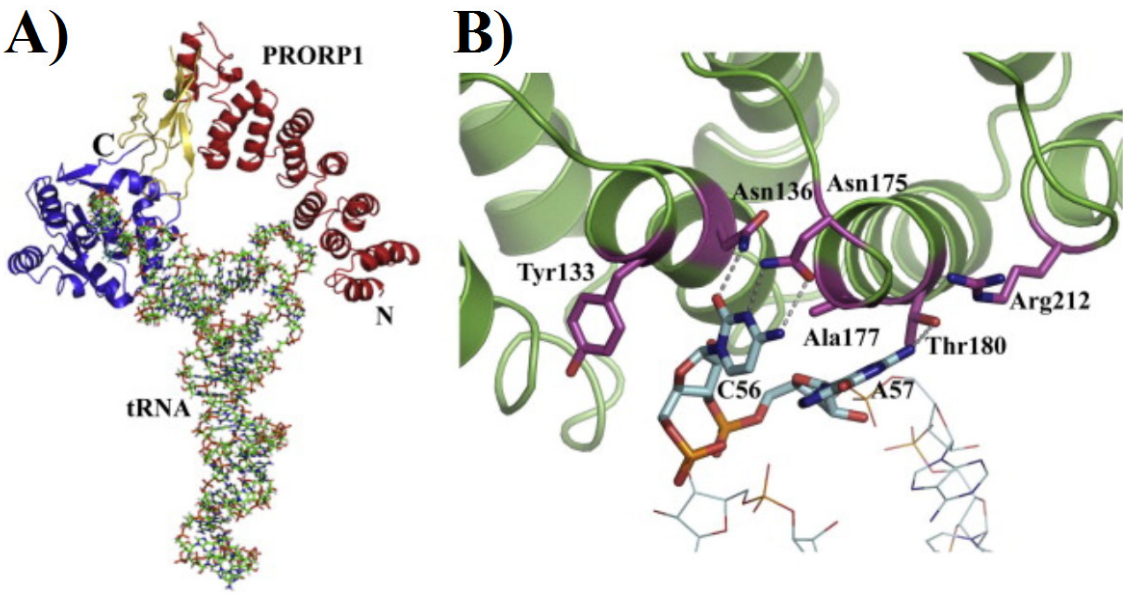

Figure Ⅲ-9: A model structure of PRORP1 in complex with tRNA. (A) A fitting model of tRNA onto PRORP1. Structures of PRORP1 and tRNA are illustrated as ribbon and sticks representation, respectively. The PPR, central linker, and NYN domains are colored red, yellow, and blue, respectively. (B) Predicted interaction between PPR motifs in PRORP1 and nucleotides on the TΨC loop. In this model, Asn136 in PPR2 and Thr180 in PPR3 recognize C56 and A57 by hydrogen bonding at the TΨC loop in tRNA, respectively. Figures were drawn with PyMol (http://pymol.sourceforge.net).

DISCUSSION

The precursor tRNA processing enzyme PRORP1 in A. thaliana recognizes the bases in a distinct manner from bacterial RNase P

Based on our computational and biochemical studies, I constructed the model complex structure of PRORP1 in complex with tRNA. This attempt is successfully fitted tRNA onto PRORP1, as shown in Figure Ⅲ-9. In this model, the side chains of Tyr133, Asn136, and Arg175 located on PPR2 motif recognize C56, and those of Ala177, Thr180, and Arg212 on PPR3 motif are located in close proximity to A57, in which the the side chains of Asn136 and Thr180 at the 4 position of PPR2 and PPR3 bind to C56 and A57 on TΨC loop in tRNA, respectively. This proposed recognition mechanism is slightly distinct from that by bacterial RNase P. In the case of RNase P, A112 and G147 in S-domain bind to C56 on TΨC loop and G19 located on D loop by utilizing π-π stacking interactions, whereas PRORP1 recognizes C56 and A57 located on TΨC loop by utilizing hydrogen bonding networks. RNA molecules almost adopt π-π stacking interactions for their structure arrangement, its stabilization, and

RNA-RNA interactions, whereas protein enzyme tend to adopt hydrogen bondings to interact with other molecules. It is thus likely that the distinct recognition mode of them is probably connected to the characteristic prosperities of nucleic acids and amino acids.

SUMMARY

I predicted the target sites for PPR motifs in PRORP1 using bioinformatic approach by PPR recognition code combined with characterization of the pre-tRNA^{Phe} processing activity with pre-tRNA^{Phe} mutants and PRORP1 mutants. The obtained sequence C-A/G-A-U by the *in silico* bioinformatic study and the biochemical assay were in good agreement with the sequence C56-A57-A58-A59 at TΨC loop in pre-tRNA^{Phe}. It was thus likely that PPR2-PPR3-PPR4-PPR5 in PROPR1 recognize the nucleotide sequence C56-A57-A58-A59 in pre-tRNA^{Phe}, respectively. To verify this hypothesis, I characterized the pre-tRNA^{Phe} mutants, including C56, A57, A58, and A59 in turn replaced with the other three nucleotides. This result showed that C56, A57, A58 mutants reduced the susceptibility (20%-40%), whereas A59 in turn replaced with the other nucleotides retained the susceptibility (90%-100%). This assay indicated PPR5 motif does not bind A59 on TΨC loop in pre-tRNA^{Phe}, whereas PPR2-PPR3-PPR4 recognizes C56-A57-A58 on TΨC loop in the pre-tRNA^{Phe}, respectively. To confirm this issue further, I prepared the PROPR1 des38 NSR mutants, where the 4th resides located at PPR2, PPR3, and PPR4 were substituted with another amino acid. Subsequently, these mutants were characterized in terms of the pre-tRNA processing activity, leading to the importance of PPR2 and PPR3 rather than PPR4 in pre-tRNA base recognition. This assumption is in good agreement with the construction of the model of PRORP1 in complex with tRNA. In this model, the side chains of Tyr133, Asn136, and Arg175 located on PPR2 motif recognize C56, and those of Ala177, Thr180, and Arg212 on PPR3 motif are located in close proximity to A57, in which the side chains of Asn136 and Thr180 at the 4 position of PPR2 and PPR3 bind to C56 and A57 on TΨC loop in tRNA, respectively. This proposed recognition mechanism is slightly distinct from that by bacterial RNase P. In the case of RNase P, A112 and G147 in S-domain bind to C56 on TΨC loop and G19 located on D loop by utilizing π-π stacking interactions, whereas PRORP1 recognizes C56 and A57 located on TΨC loop

by utilizing hydrogen bonding networks. RNA molecules almost adopt π-π stacking interactions for their structure arrangement, its stabilization, and RNA-RNA interactions, whereas protein enzymes tend to adopt hydrogen bondings to interact with other molecules. It is thus likely that the distinct recognition mode of them is probably connected to the characteristic prosperities of nucleic acids and amino acids.

Chapter Ⅳ

Crystallographic Analysis of PRORP1 in Complex with its Target RNA

INTRODUCTION

As describe in the chapter Ⅲ, biochemical and computational data were combined to construct the model structure of PRORP1 bound tRNA. In this model, PPR2 and PPR3 in PRORP1 are involved in specific recognition C56 and A57 at TΨC loop in pre-tRNA^{Phe} (Figure Ⅲ-4). This model structure suggests that PRORP1 primarily interacts with nucleotides at TΨC loop in pre-tRNA, and the recognition mechanism is slightly distinct from that by bacterial RNase P in that it recognizes nucleotides at both D loop and TΨC loop in pre-tRNA in catalysis (Reiter *et al.,* 2010). Recently, Zhang and Ferre-D Amare reported that the recognition modes of tRNA elbow by π-π stacking interactions are analogous to the recognition modes beween tRNA and L1 stalk of the ribosomal E-site, and that between T box riboswitch Stem I domain and i tRNA (Zhang and Ferre-D Amare, 2013; Korostelev *et al.*, 2006). Because these ribonucleoproteins and T-boxes are not evolutionary related, the recognition mode has convergently evolved. On the other hand, the analogous recognition strategies by protein enzymes have not been found so far; protein enzymes tend to adopt hydrogen-bonding networks to interact with other molecules. Moreover, Pavlova *et al.* postulated that the catalytic mechanism of PRORP1 is slightly distinct from that of RNase P (Pavlova *et al.*, 2012). It is therefore likely that the characteristic prosperities of ribonucleotides and amino acids probably select the binding oxygen of the scissile bond as well as the recognition sites and its recognition strategy.

In this chapter, to corroborate the proposed recognition mechanisms of the protein enzyme (PRORP1) and also to compare it with that of the ribozyme (RNase P) in an atomic resolution, I attempted to crystallize PRORP1 in complex with pre-tRNA^{Phe} or pre-nad6 for X-ray crystallographic analysis.

MATERIAL AND METHODS

Materials

DNA polymerases were purchased from MBI Fermentas. Crystallization reagents were supplied by Hampton Research (34 Journey, Aliso Viejo, CA92656, United States of America). All other common chemicals and reagents were purchased at the highest purity available.

Preparation and purification of pre-tRNA and pre-nad6

The DNA fragments encoding chloroplasts pre-$tRNA^{Phe}$ in *A. thaliana* were amplified by PCR with an upstream primer containing T7 promoter (5'-TAATACGACTCACTATAGGGAGGATGATCC-3') and a sequence specific downstream primer (5'-GGTGCCAAGAACCAGATTTGAAC-3'). pre-$tRNA^{Phe}$ with 13 nt leader sequence at the 5'-end was transcribed *in vitro* using the amplified gene fragment as a template. The resulting RNAs were purified on a column of DEAE-sepharose, as described by Eason *et al.* (Easton *et al.,* 2010), after which they were further purified with a HiLoad 16/60 superdex200 gel filtration column. pre-nad6 was prepared in the same manner as pre-tRNA. The DNA fragments encoding pre-nad6 was amplified using the forward primer (5'-TAATACGACTCACTATAGGGGGACGACTGACCC-3') and the reverse primer (5'-GGTCTCACGAATTGGATTCGAACCAATC-3').

Gel mobility shift assay

A gel mobility shift assay was performed to evaluate RNA binding by D399A. The binding reaction contained 30 mM Tris -HCl, pH 7.5, 100 mM NaCl, 1 mM DTT, 10% glycerol, 5 mM $MgCl_2$, 2 μg of its target RNA, and increasing amounts of D399A. The detailed procedures were described in MATERIALS AND METHODS in chapter Ⅱ.

Analytical size-exclusion chromatography

Analytical size-exclusion column chromatography was performed with a Biogel P100 gel filtration column (1.5 × 60 cm). The column was equilibrated with running

buffer (30 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10% glycerol, 5 mM $MgCl_2$, and 1 mM DTT) to obtain a stable base line. After that, PRORP1, RNA, or PRORP1/RNA mixture was incubated for 30 min on ice, and then separated by this column.

Formation of PRORP1 in complex with its substrates

A systematic protein engineering effort was conducted for crystallization of RNA-free and RNA-bound PRORP1. To obtain the crystals of protein-RNA complex, PRORP1 des76 D399A, where the N-terminal 76 amino acids were truncated and the catalytic residue Asp399 was replaced with Ala (denoted D399A, hereafter), was purified through nickel-affinity column chromatography, followed by RESOURCE S ion exchange column chromatography. The protein was then applied onto a HiLoad 16/60 Superdex 200 gel filtration column in 30 mM Tris-HCl, pH7.5, 100 mM NaCl, 1 mM DTT, 5 mM $MgCl_2$, and 10% glycerol. The peak fractions were incubated with its target RNA molecule with a molar ratio of approximately 1:1.2 (= D399A : RNA) at 4 ℃ for 30 min before crystallization trials.

Crystallization

RNA-free D399A protein was crystallized by sitting drop vapor diffusion method at 4 ℃. The protein at a concentration of approximately 14 mg/ml was mixed with an equal volume of reservoir solution containing 20% PEG3350, and 0.1 M magnesium chloride. Lod-shaped crystals appeared for 1 year. To obtain crystals of protein-RNA complex, various combinations of protein regions and RNA were examined at 4 ℃ and 20 ℃ in the same manner as RNA-free D399A.

Data collection and structural determination

All data sets were collected at beam line, BL41-XU at SPring-8 in Japan. The data sets were processed with the HKL2000 package (Otwinowski and Minor, 1996). Further processing was performed with programs from the CCP4 suite (Vagin and Teplyakov, 2000; McCoy *et al.,* 2007). Data collection and structure refinement statistics are summarized in Table Ⅳ. Data sets collected from the only one crystal of the RNA-free D399A was used. The structure of D399A was solved by molecular replacement with the PRORP1 wild-type as a search model. Interactive rounds of model

building and refinement were carried out manually using Coot (Emsley *et al.,* 2010) and PHENIX REFINE (Adams *et al.,* 2010). The geometry of the final model was evaluated using the program PROCHECK (Laskowski *et al.,* 1993). Figures were prepared using Pymol (www. pymol. org).

Table IV: Data collection and refinement statistics

Data collection	PRORP1 des76 D399A
Space group	$P2_1$
Unit cell parameters	a = 41.0 Å
	b = 109.1 Å
	c = 135.3 Å
	α = 90°
	β =90.9°
	γ =90°
Wavelength (Å)	1.0
Resolution range (Å)	50.0 – 2.47 (2.51 - 2.47)
No. of reflections	
Observed / Unique	90135 / 41052
Redundancy	2.3 (1.8)
R_{merge} (overall/outer shell)	0.166 (0.480)
I/σ (I)	8.1 (1.0)
Completeness (%)	93.9 (73.5)
Refinement statistics	
Resolution range (Å)	42.45 - 2.48
No. of reflections	
Working set / Test set	41216/ 2479
Completeness (%)	97.68
R_{cryst} (%) / R_{free} (%)	19.3 / 25.8 (26.5 / 35.6)
Root mean square deviation	
Bond length (Å)	0.008
Bond angles (°)	1.06
Average B-factor (Å^2) / No. of atoms	
Protein	58.464 / 7656
Zn	50.12 / 2
Water	46.37 / 232
Ramachandran analysis	
Favored (%)	96
Allowed (%)	4.1
Outlier (%)	0.21

RESULTS

Preparation of pre-tRNA and pre-nad6

As shown in MATERIALS AND METHODS in this chapter, pre-tRNA^{Phe} and pre-nad6 were transcribed *in vitro*, after which they were purified with a DEAE -

sepharose column chromatography (Figure Ⅳ). To characterize the eluted fractions, I conducted denatured polyacrylamide gel electrophoresis in 8 M urea. Subsequently, the interested fractions were pooled and substituted the buffer for Milli-Q water, followed by dilution the sample with Milli-Q water at a concentration of 200 ng/μl. The diluted RNA was heated at 70 ℃ for 10 min, temperature in order to refold them property. Their homogeneities were confirmed by non-denatured polyacrylamide gel (4%) gel electrophoresis (Figure Ⅳ). The RNAs were further purified by Hiload 16/60 Superdex 200 gel filtration column chromatography (Figure Ⅳ). Their yields were 1 mg per 1 ml of the transcribed solution.

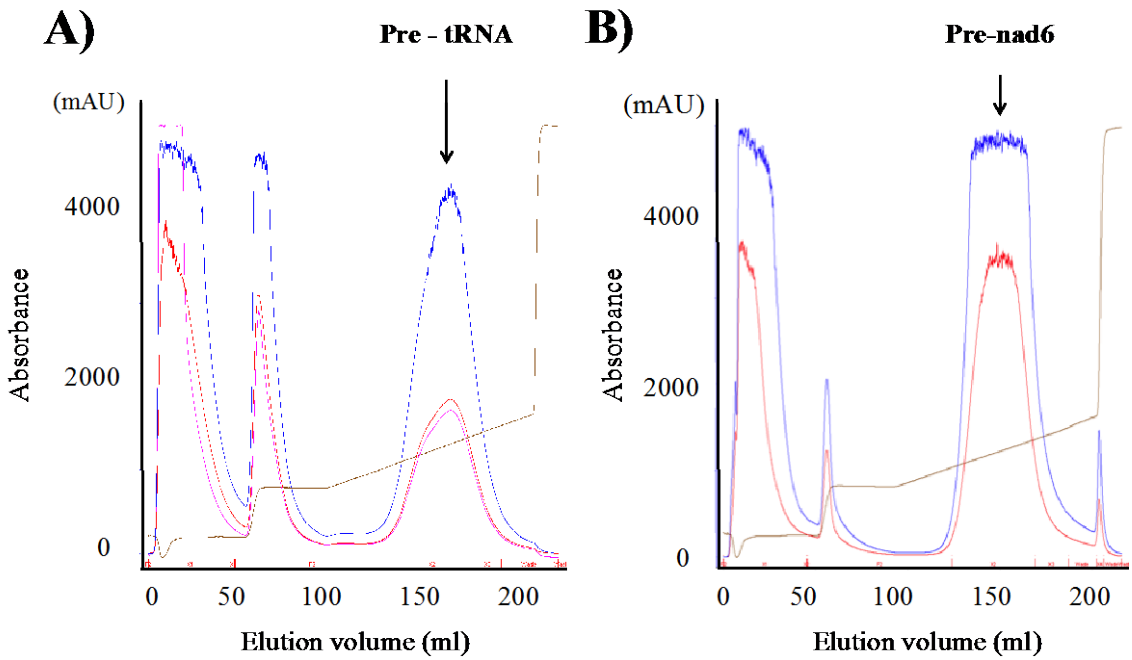

Figure Ⅳ: DEAE-Sepharose column chromatography of pre-tRNA^{Phe}(A) and pre-nad6 (B). The experimental details are described in MATERIAL AND METODS in Chapter Ⅳ. A 260 (nm) and (nm) represent red and blue, respectively. The electric conductivity is depicted as brown lines.

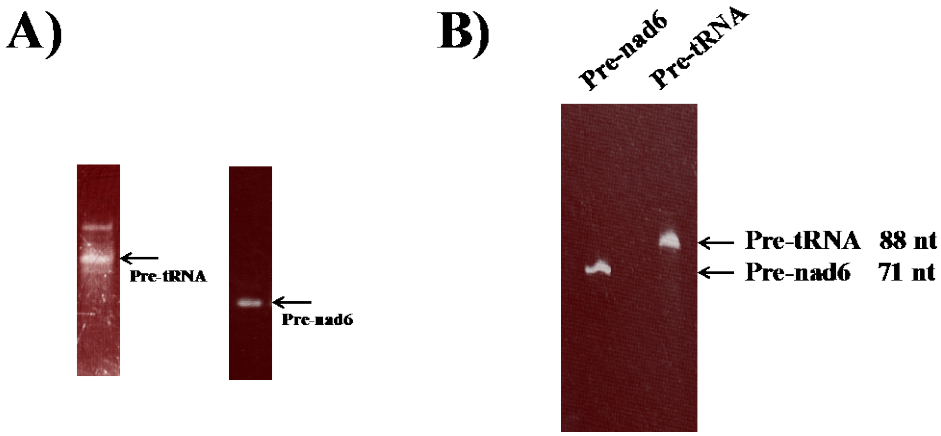

Figure Ⅳ: The homogeneity of the pre-tRNA^{Phe} and pre-nad6. (A) Native PAGE of the purified RNAs. (B) A denatured polyacrylamide gel of the purified RNAs. The electrophoresis band of pre-tRNA^{Phe} or pre-nad6 is depicted as a black arrow.

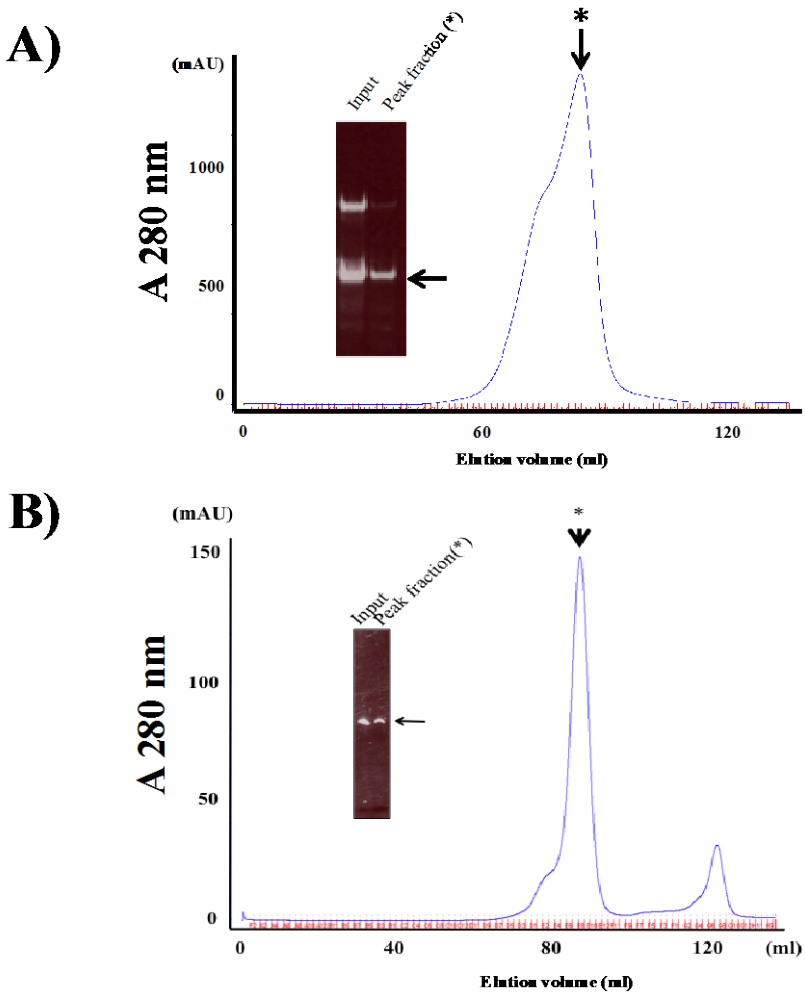

Figure Ⅳ-3: Gel filtration column chromatography of pre-tRNA^{Phe} and pre-nad6. (A) Chromatogram and a native PAGE gel of the purified pre-tRNA^{Phe}. A 260 nm represents blue lines. The peak containing pre-tRNA^{Phe} was shown with " * ". (B) Chromatogram and a native PAGE gel of the purified pre-nad6. The peak containing pre-nad6 was shown with "*" like Figure Ⅳ-3A.

Formation of PRORP1 in complex with pre-tRNA^{Phe} or pre-nad6

To search the stable condition for PRORP1 in complex with its RNA substrates, I performed by gel mobility shift assay and analytical gel filtration assay (Figure Ⅳ-4). The basis of gel mobility shift assay is the change in the electrophoretic mobility of a nucleic acid molecule upon binding to a protein or other molecule (Hellman and Fried, 2007). Initially, I mixed 2 μg of pre-nad6 with the increasing amount of D399A at 37 ℃ for 30 min. Subsequently, I put the reacted samples onto a 4% non-denatured polyacrylamide gel and electrophoresed at 150 V for 40 min, after which the resulted gel was stained with ethidinum bromide solution and CBB solution to detect pre-nad6 and D399A, respectively (Figure Ⅳ). As shown in Figure Ⅳ, while the unbound RNA has a characteristic electrophoretic mobility on a native gel (illustrated by a black arrow), in the case of the existence of D399A, the movement of the RNA through the gel is retarded, resulting in characteristic shift in the position of the RNA band on the gel (illustrated by a red arrow). Likewise, the pre-tRNA also interacted with D399A. In summary, these results indicated that D399A binds its target RNA molecules in a stable form. Furthermore, the stoichiometric experiments demonstrated that D399A binds its target RNA substrates with approximately 1:1 molar ratio.

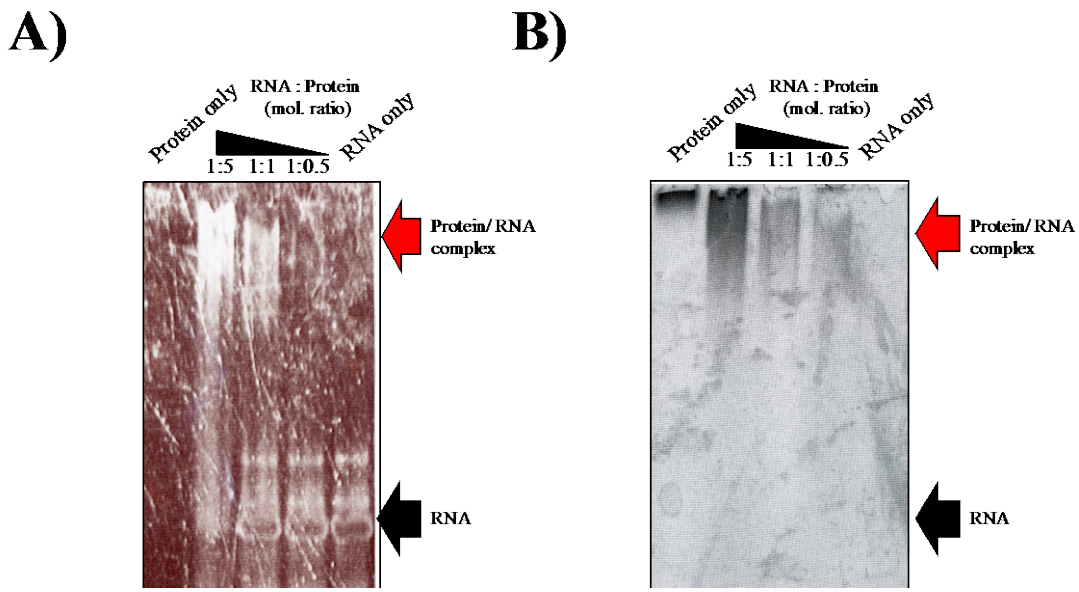

Figure Ⅳ: Gel mobility shift assay. Each binding reaction contained the indicated amount of PRORP1 and 2 μg of pre-nad6 or pre-tRNA^{Phe} in 30 mM Tris–HCl, pH 7.5, 100 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, and 1 mM DTT. Reacted samples were run on a 4% native acrylamide gel and run for 40 min at room temperature at 150 V. The complex was detected by soaking etidium bromide solution (A) and CBB solution (B). The electrophoresis bands of pre-nad6 and PRORP1 in complex with pre-nad6 are marked by black and red arrows, respectively.

To confirm this assumption, I performed the analytical gel filtration column chromatography using a HiLoad 16/60 Superdex 200 column. Initially, 5 mg of pre-nad6 and the increasing amount of D399A were mixed at 4 ℃ for 30 min, a which the sample was put onto the column and then compared the elution volume of the complexes with those of D399A and pre-nad6 (Figure Ⅳ). As shown in Figure Ⅳ, the elution volume of pre-nad6 was 48 ml and that of D399A was 38 ml, whereas that of the complexes was 28 ml using a Biogel P100 gel filtration column (1.5 × 60 cm) (Figure Ⅳ). This result suggested that pre-nad6 bound D399A forms in a stable state. Furthermore, to improve the homogeneity of the complexes, I could separate them from the other molecules, including D399 alone and its RNA substrate alone by gel filtration column chromatography. Likewise, pre-tRNA-bound D399A was purified with a HiLoad16/60 Superdex200 gel filtration column. Accordingly, I attempted to crystallize PRORP1 in complex with pre-tRNA^{Phe} or pre-nad6.

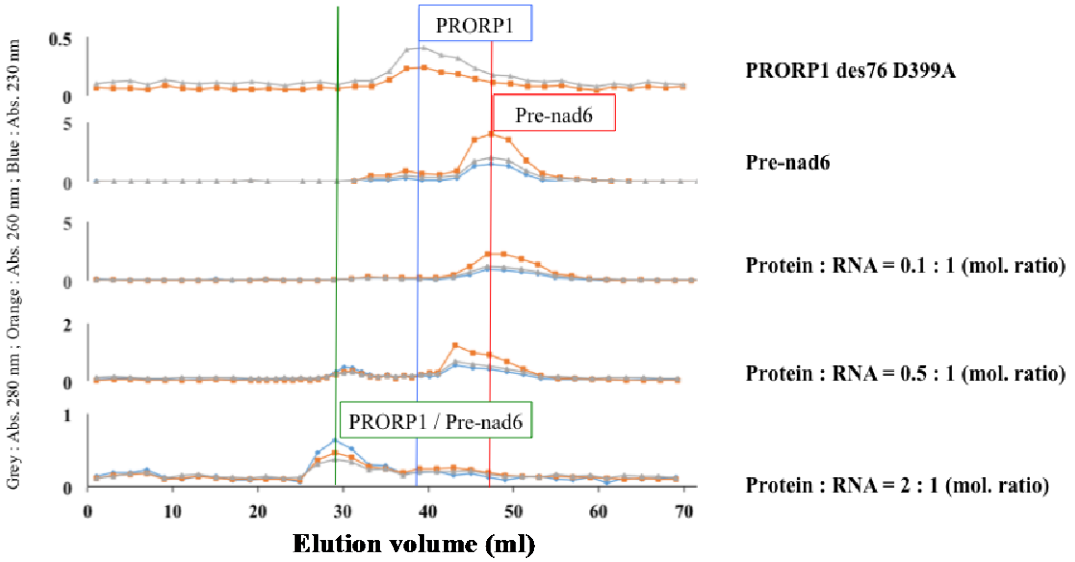

Figure Ⅳ: Analytical gel filtration column chromatography of PRORP1 in complex with pre-nad6. Abs. 280 nm, Abs. 260 nm, and Abs. 230 nm are shown by grey, orange, and blue color, respectively. The compositions of each fraction were characterized by non-denatured gel and denatured gel in 8 M urea electrophoresis. The elution volume of PRORP1/pre-nad6 complex, PRORP1 alone, and pre-tRNA alone, are marked by green, blue, and red line, respectively.

Crystallization and data collection

The purified D399A was mixed with the purified RNA molecule at approximately 1:1.2 (mol. ratio) at 4 ℃ for 30 min. The mixture was concentrated to 10 mg/ml of D399A in 30 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM DTT, 5 mM $MgCl_2$, and 10% glycerol. Initially, the amorphous crystals were grown under various number of

conditions at 20 ℃ (Figure Ⅳ-6). To characterize their composition, I performed with the denatured polyacrylamide gel (15%) electrophoresis in 8 M urea. As shown in Figure Ⅳ-6, the crystals contained pre-tRNA^{Phe} without PRORP1 des76 D399A (Figure Ⅳ-6). However, the bands of the degraded polypeptide were detected on the gel (Figure Ⅳ-6). Thus, to recheck their composition, I conducted the analytical micro-seeding method. The appeared crystals were used as seed crystals, and then harvested in the crystallization solution, containing pre-tRNA^{Phe} alone. The crystals grew under the condition, indicating that the crystals contained pre-tRNA^{Phe} without the protein moieties.

To search another condition for crystallization of PRORP1 in complex with pre-tRNA^{Phe} or pre-nad6, I crystallized them at 4 ℃. Additionally polyethylene glycol instead of high concentration of salts as precipitants in order to prevent the electrostatic interactions between D399A and pre-tRNA^{Phe} from disruption. The crystals appeared in the drop composed of the precipitant containing 20% PEG3350 and 0.1 M $MgCl_2$ (Figure Ⅳ-A).

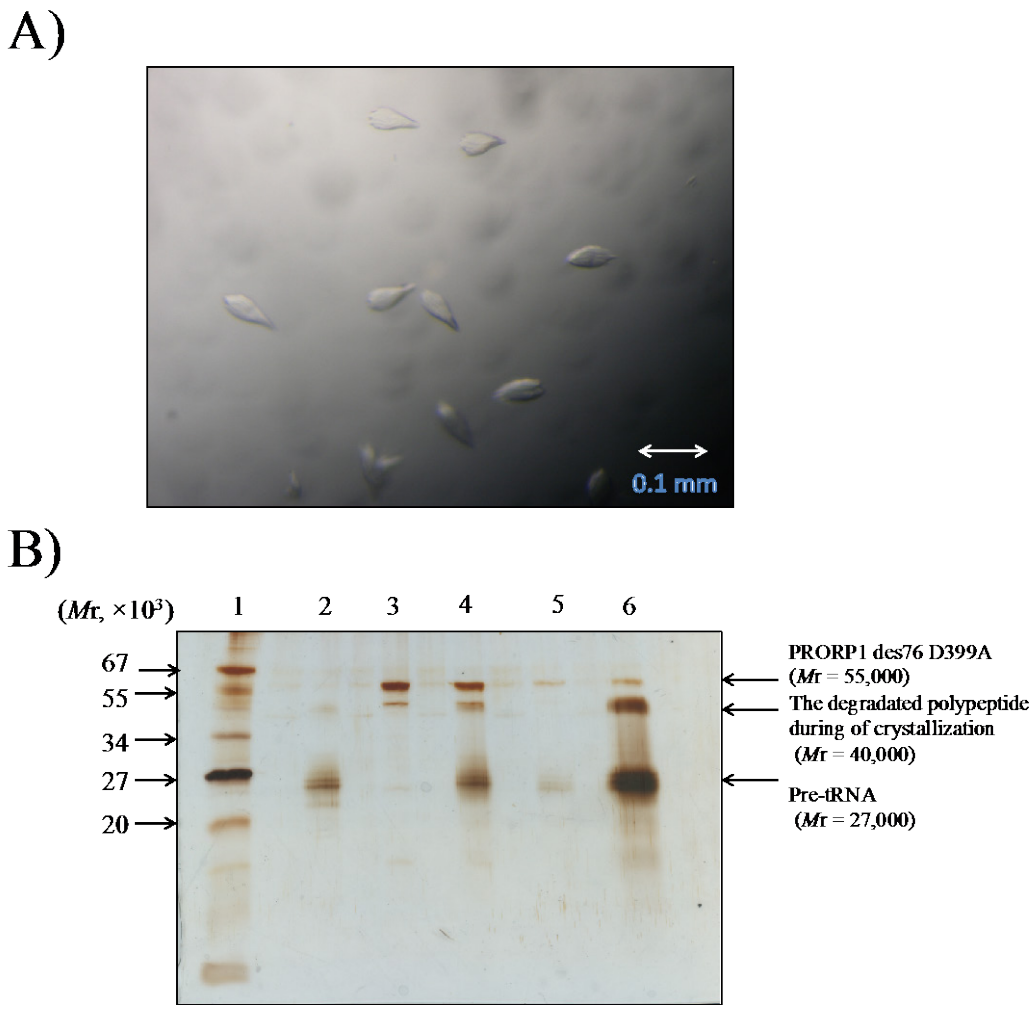

Figure IV: Crystals of chloroplast pre-tRNA^{Phe} in *A. thaliana*. (A) The crystals were obtained in a 1 µl drop containing 5 mg/ml PRORP1 des76 D399A, 3 mg/ml pre-tRNA, and the reservoir solution in 0.05 M magnesium chloride, 0.1 M HEPES-Na, pH 7, 1.6 M ammonium sulfate using a micro-seeding method. Crystals appeared within 1-3 days at a room temperature. (B) A silver-stained electrophorogram of the crystals separated on a 15% polyacrylamide gel in the presence of SDS. 1, Molecular marker; 2, pre-tRNA alone; 3, PRORP1 des76 D399A alone; 4, The crystallization solution containing pre-tRNA^{Phe} and PRORP1 des76 (1:1.2 molar ratio); 5, The used crystal cleaning solution; 6, The water solution, where the washed crystals were dissolved. The electrophoresis bands of PRORP1 des76 D399A, the degradated polypeptide during of crystallization, and pre-tRNA^{Phe} are marked by black arrows.

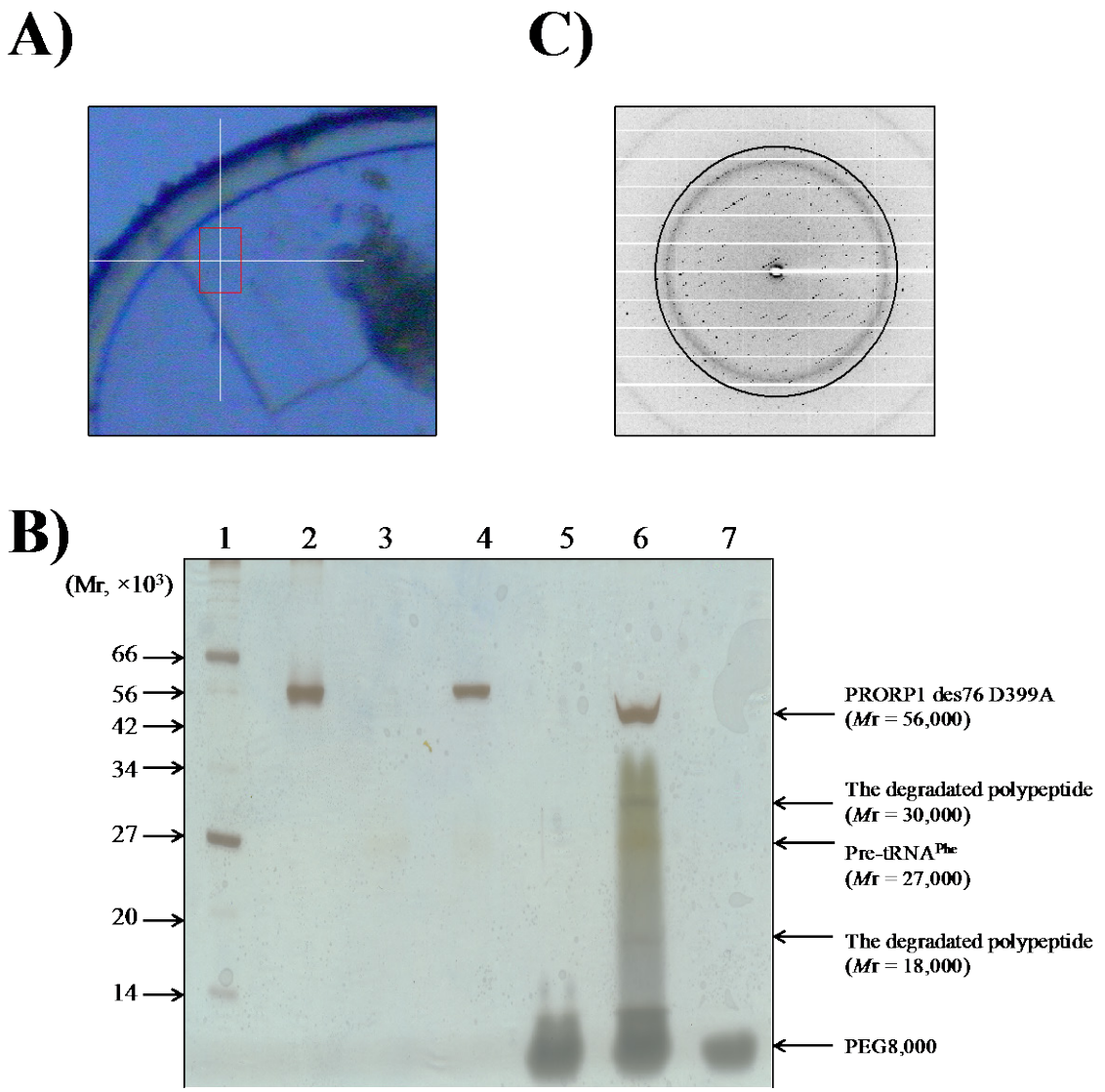

Figure IV: Characterization of PRORP1 des76 D399A crystals. (A) Stack of the five hexagonal crystals of *Arabidopsis* PRORP1 des76 D399A grown by well in 100 mM magnesium chloride, 20% (*w*/*v*) PEG3350 (scale bar = 0.1 mm). Crystals were separated using a cryoloop prior to cryocooling and data collection. (B) A silver-stained electrophorogram of the crystals separated on a 15% polyacrylamide gel in the presence of SDS. 1, Molecular marker; 2, PRORP1 des76 D399A; 3, pre-tRNA^{Phe}; 4, The crystallization solution containing pre-tRNA^{Phe} and PRORP1 des76 D399A(1:1.2 molar ratio); 5, The used crystal cleaning solution; 6, The water solution, where the washed crystals were dissolved; 7, PEG8,000 alone. The electrophoresis bands of PRORP1 des76 D399A, the degraded polypeptides, PEG8,000, and pre-tRNA^{Phe} are marked by black arrows. (C) Representative X-ray diffraction image collected on the beamline at the X-ray synchrotron. The black circle corresponds to the 3 Å resolution limit.

Structure determination

The crystals diffracted to a resolution of 2.47 Å on a synchrotron X-ray source (Figure IV-B). The crystal belongs to the space group $P2_1$. The unit cell axes and the angles are a = 41.0 Å, b = 109.1 Å, c = 135.3 Å and α = 90°, β = 90.9°, γ = 90°. The phase was determined by the MR method, and was refined to R_{cryst} (%) / R_{free} (%) of 26.5 / 35.6 at 2.48 Å. There are two PRORP1 molecules in the asymmetric unit. But the crystals did not contain RNA molecules (Figure IV-B).

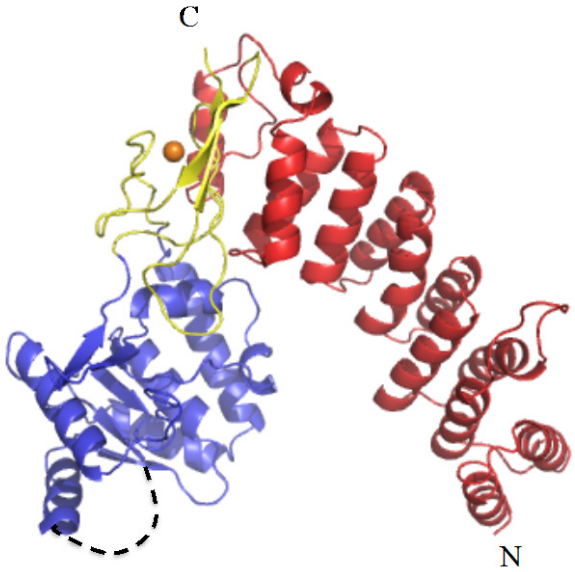

Figure IV-B: Overall structure of PRORP1 des76 D399A. Ribbon diagram of PRORP1 des76 D399A mutant. A PPR domain (pos. 95–327), a zinc-finger like central domain (pos. 328–354; pos. 534-570), and a NYN metallonuclease domain (pos. 355-533) are shown in red, yellow, blue, respectively. The coordinated Zn^{2+} ion is depicted as an orange sphere. The disordered loop in NYN metallonuclease domain is depicted by black dash lines.

The pre-tRNAs were extracted from the drop where the crystal appeared, and visualized on the denatured polyacrylamide gel (15%), suggesting that the pre-tRNA^{Phe} in the well remains intact during crystallization. It is therefore needed to improve the affinity between PRORP1 and pre-tRNA for crystallization of the complexes, and necessary to search other crystallization conditions. To overcome the obstacles for crystallization of the complexes, I crystallized PRORP1 in complex with pre-tRNA^{Phe} or pre-nad6 at 20 °C and 4 °C. Although resolution of 3 Å, N-terminal-truncated PRORP1 des residues were truncated, without pre-nad6 (Figure IV-C). Furthermore, I used its isozymes, PRORP2 desN9 or PRORP3 desN10 in complex with pre-tRNA^{Phe} or

pre-nad6 and attempted to crystallize these complexes at 4 ℃ and 20 ℃. A various conditions for crystallization were screened, I obtained the one crystal. The crystal appeared in the drop composed of the precipitant containing 16% PEG3350 and 0.1 M Ammonium tartrate (Figure Ⅳ-10). To characterize its composition, the analytical micro-seeding method was carried out. The appeared crystals were used as seed crystals. The crystals grew in the solution, containing PRORP2 desN9 without pre-nad6. Thus, the crystal contains PRORP2 desN9 without the RNA moieties.

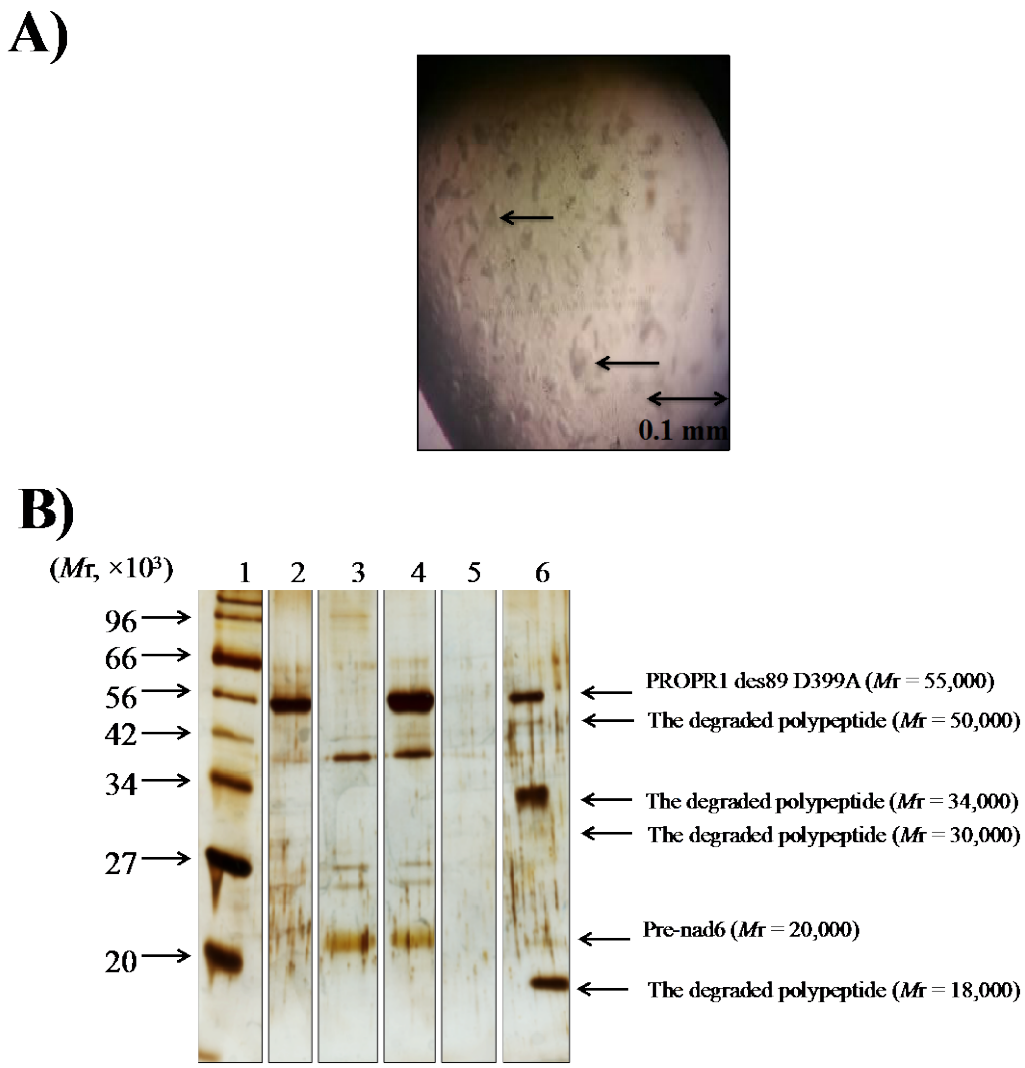

Figure Ⅳ-10: Crystals of PRORP1 des89 D399A. (A) The crystals of PRORP1 des89 D399A. The crystals were obtained in a 1 μl : 1 μl mixtures of protein solution, containing 10 mg/ml PRORP1 des89 D399A and 5 mg/ml pre-tRNA, with reservoir solution in 3.5 M sodium formate, 0.1 M Tris–HCl, pH8.5, and 10 mM $MgCl_2$ using sitting drop vapor diffusion method at 4 ℃. (B) A silver-stained electrophorogram of the crystals separated on a 15% polyacrylamide gel in the presence of SDS. 1, Molecular marker; 2, PRORP1 des76 D399A; 3, pre-nad6; 4, The crystallization solution containing

pre-nad6 and PRORP1 des76 D399A (1:1.2 molar ratio); 5, The used crystal cleaning solution; 6, The water solution, where the washed crystals were dissolved. The electrophoresis bands of PRORP1 des76 D399A, the degraded polypeptides, and pre-nad6 are marked by black arrows.

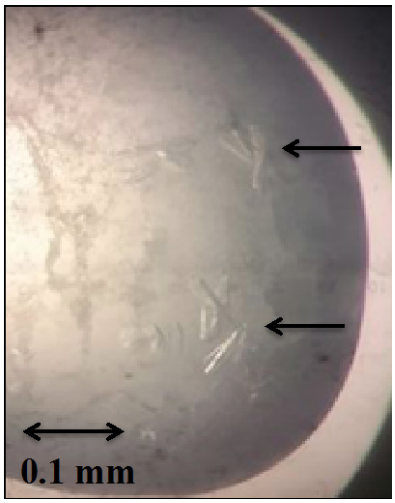

Figure IV0: Crystals of PRORP2 desN9. The crystals were obtained in a 1 µl drop consisting 10 mg/ml PRORP2 desN9 and 7 mg/ml pre-nad6, and the reservoir solution (16% PEG3350 and 0.1 M Ammonium tartrate) using sitting drop vapor diffusion method at 4 ℃.

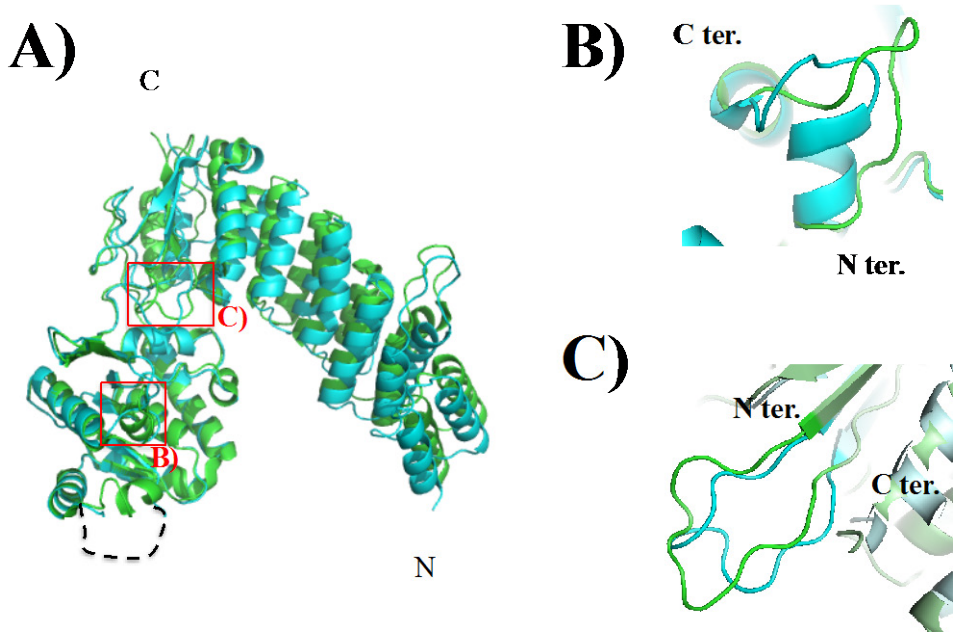

Figure IV1: Superimposition of the crystal structure of PROPR1 des76 wild-type on that of PRORP1 des76 D399A. (A) PRORP1 des76 wild-type (wild-type) and PRORP1 des76 D399A (D399A) are colored by cyan and green, respectively. The disordered loop in NYN metallonuclease domain is depicted by black dash lines. The structural differences between wild-type and D399A are marked with red boxes. (B) Superimposition of the helical structure (pos. 399–406) in wild-type on that of D399A. (C) Superposition of the lariat loop structure (pos. 550-563) in wild-type on that of D399A.

Comparison of the crystal structure of PRORP1 D399A with that of PRORP1 wild-type

Comparing the crystal structure of D399A with that of PROPR1 wild-type, the overall structures of the individual domains containing a NYN metallonuclease domain, a zinc ion binding central domain, and PPR domain are similar to each other (Howard *et al.,* 2012) (Figure IV-1A). However, there are two key differences between D399A and PRORP1 wild-type. First, the individual domains are oriented differently (Figures IV-1 and 12). PROPR1 wild-type adopts a relaxed, twisted, L-shaped conformation with the NYN metallonuclease domain being moved out the plane formed by the PPR and central domains. In D399A structure, the PPR and NYN metallonuclease domain are located closer together, forming V-shaped with all domains lying on one plane (Figures IV-11 and 12). Taking the catalytic residues as a reference, the PPR domain of PRORP1 wild-type is located further out by 12 Å and 6 Å an respectively (Figure IV-2).

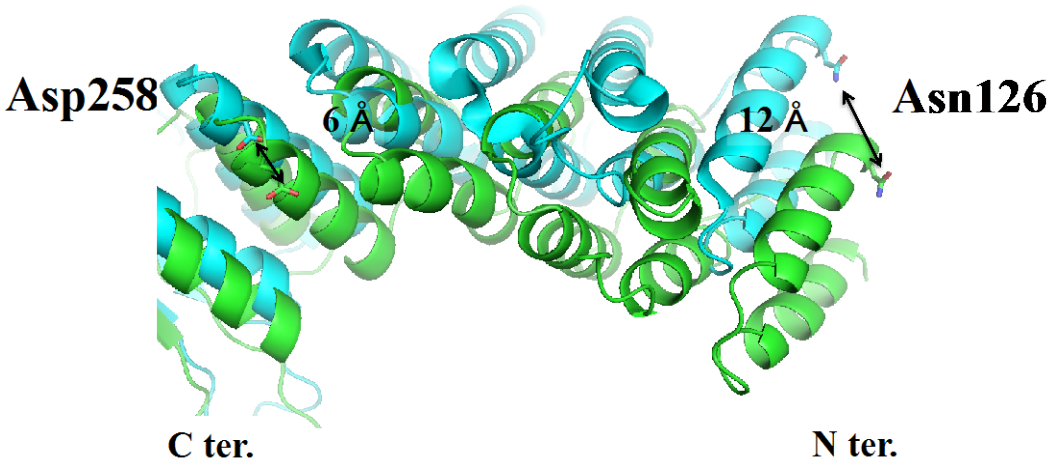

Figure IV-2: Superimposition of the PPR domain of wild-type (*cyan*) on those of D399A (*green*). The PPR domain of wild-type was located further out by 12 Å and 6 Å for As respectively.

The individual PPR domains and NYN metallonuclease domains were almost all superimposed, whereas the loop (pos. 550-563) in the central domain, named as lariat loop, was significantly conformational changed (Figure IV-1C). The lariat loop of PRORP1 wild-type adopts the coiled coil structure, whereas that of D399A adopts the random coiled structure. This conformational change probably reduced the distance between the PPR domain and NYN metallonuclease domain. Second, the key difference is the structural change of the α-helix located on Asp399. The helical structure (pos.

399-406) in PROPR1 wild-type is structural changed to the random coli structure in D399A (Figures IV1B and 13). Thus, the side chain of Asp399 is likely to play an important role in stabilization the helical structure for catalysis as well as serving as the catalytic residue, which binds to one water molecule for nucleophilic-attack targeting to scissile phosphodiester bond in its RNA substrates.

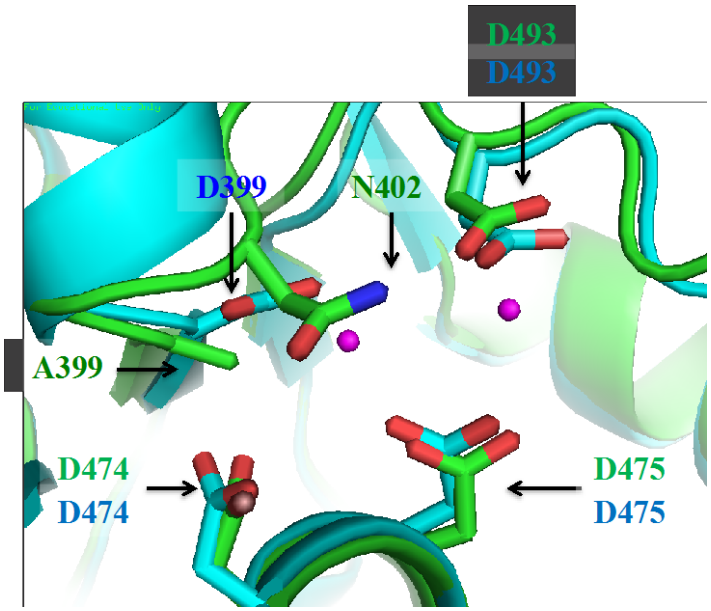

Figure IV3: Superimposition of the NYN metallonuclease domain of PRORP1 des76 wild-type (wild-type) on that of PRORP1 des76 D399A (D399A). Superimposition of the catalytic residues Asp399, Asp474, Asp475, and Asp493 of wild-type on those of D399A. In D399A structure, Asp399 was mutated to Ala. Additionally, Asn402 in D399A is shown in stick representation. The coordinated Mn^{2+} ions are depicted as magenta spheres.

DISCUSSION

What is the function of the helical structure located on Asp399 of PRORP1?

Its α helix structure is conserved among PIN and Flap nuclease families as well as NYN metallonuclease family (Anantharaman and Aravind, 2006). I superimposed the crystal structure of D399A on that of human exonuclease 1 (hExo1) in complex with its DNA substrate (Orans *et al.*, 2011) (Figure IV4). The NYN domain of PRORP1 was well superimposed onto the catalytic domain of hExo1. Furthermore, the catalytic residues of PRORP1 (Asp399, Asp474, Asp475, and Asp493) were well superimposed onto the catalytic residues (Asp30, Glu150, Asp152, and Asp171), respectively. In previous data reported by Orans *et al.,* in the hExo1 structure, the α helix structure contributes substrate recognition and cleavage (Orans *et al.,* 2011). In detail, the helix

positions the scissile bond in the vicinity of the metal center. Especially, H36 residue placed in the helix interacts with the terminal base and plays a critical part in the rigorous regulation of the cleavage site. In PRORP1, the His reside is not conserved, however, the helical structure of PRORP1 is well conserved with that of hExo1. Thus, it could be assumed that the α helix in PRORP1 performed the same function like that of hExo1. Further studies, including PRORPs bound to their substrates, are needed to find out the precise contribution of the helix structure located on Asp399 in PROPRs.

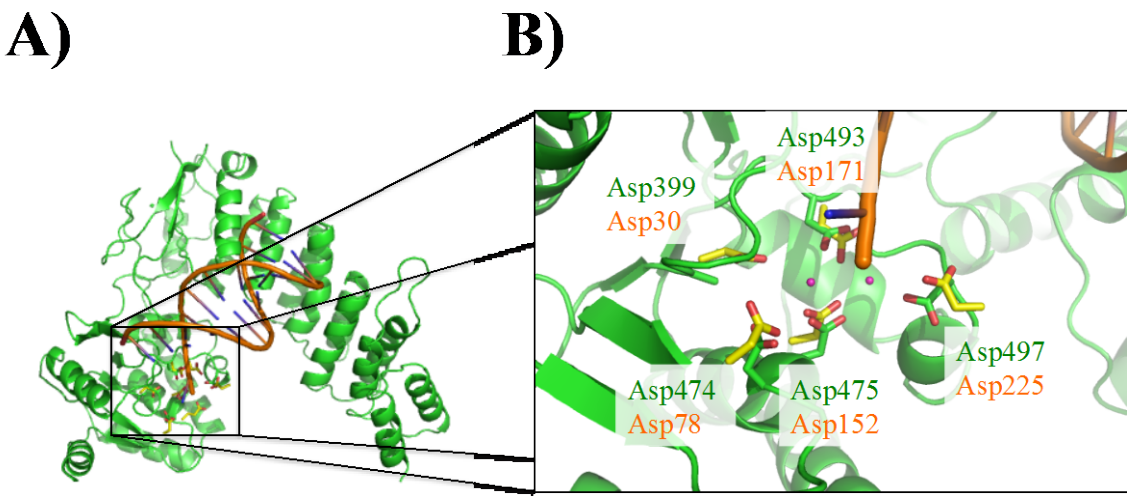

Figure IV4: Superimposition of PRORP1 des76 D399A on human exonuclease 1 (hExo1) in complex with its DNA substrate. (A) Overall structure of PRORP1 superimposed on the DNA substrate of hExo1. When I constructed its model complex structure, I superimposed the catalytic residues of PRORP1 on those of hExo1 in complex with its DNA substrate. To make clearly understandable, the structure of hExo1 alone was hidden. PRORP1 is colored green and the DNA molecule is done orange. (B) Close-up view of the catalytic pocket. The catalytic residues of PRORP1 are showed in stick representation (colored by green) and those of hExo1 are done in stick representation (colored by yellow). The coordinated sphere Mn^{2+} ions in hExo1 are depicted as magenta spheres.

Why is it difficult to crystallize D399A in complex with pre-tRNA or pre-nad6?

To crystallize PRORP1 in complex with its RNA substrate, I used PRORP1 des76, where the N-terminal 76 residues were truncated, and pre-$tRNA^{Phe}$ or pre-nad6 (Howard *et al.,* 2012; Robertus *et al.,* 1974). The matured tRNAs tend to be dissociated from their complexes. Thus, I mutated the catalytic residue Asp399 replaced with Ala. I performed the crystallization by various conditions. However, I could not obtain the crystals comprising PRORP1 and its target RNA molecules. Based on the crystal structure of D399A, it is needed to change the mutation site for overcoming this obstacle. In the structure of D399A, although the catalytic pocket of D399A is similar to that of PRORP1 wild-type, the helical structure is disrupted, leading to collapse the

arrangement of its catalytic pocket. In the crystal structure, the place of the side chain in Asp399 is replaced with that of Asn402. It is thus assumed that Asn402 in D399A inhibits metal coordination necessary for its catalysis. This hypothesis is supported by the data reported by Howard *et al* (Howard *et al.,* 2015). In this study, despite of the high concentration of magnesium ions (100 mM), D399A did not exhibit its activity, even though D474A and D475A mutants, where the catalytic residues Asp474 and Asp475 were in turn replaced with Ala, respectively, recovered its processing activity under the same conditions. In summary, it is needed to change the mutational site, where this mutation does not lead to disrupt its catalytic structure.

SUMMARY

I attempted to solve the crystal structure of PRORP1 in complex with pre-tRNA or pre-nad6 in order to gain more information about the mechanism for RNA recognition of PRORP1. I synthesized and purified the pre-tRNA and pre-nad6 for crystallization by utilizing *in vitro* transcription system and several steps of column chromatographies, respectively. To confirm the binding affinity between PRORP1 des76 D399A, where the N-terminal 76 residues were truncated and the catalytic residue Asp399 was replaced with Ala (denoted D399A, hereafter), and its RNA substrates, I performed with gel mobility shift assay and analytical gel filtration column chromatography. Theses results showed that D399A binds to pre-tRNA and pre-nad6 as a stable complex in the buf. A (30 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10% glycerol, 1 mM DTT, and 5 mM $MgCl_2$). Based on these data, I searched the crystallization condition of D399A in complex with pre-tRNA and D399A in complex with pre-nad6 at 4 ℃ and at 20 ℃b
crystals, containing D399A alone or pre-tRNA alone, were produced in many conditions, whereas that of D399A complexed with its RNA substrate was not appeared at least in our screened conditions. To overcome this obstacle, I used its isozyme PRORP2 desN9 was used for crystallization with its substrate RNAs. However, crystals have not been obtained until now.

When the crystal structure of PRORP1 (pos. 95-570) (Howard *et al.,* 2013) was compared with that of D399A, the overall structures are similar to each other. However,

there are two key differences between PRORP1 and D399A. First, PRORP1 adopts a relaxed conformation, whereas the PPR domain and NYN domain of D399A are located closer together. This structural difference is probably caused by conformational change of the loop located on the central domain (pos. 550-563). Second, the catalytic pocket of D399A is collapsed, leading to prevent two metal ions necessary for its catalytic reaction from coordination. It is attributed to the observation that the helical structure (pos. 399-406) is changed to the random coil structure. When the crystal structure of PRORP1 is superimposed on that of human exonuclease1 (hExo1) in complex with its DNA substrate, the topology of the helical structures is similar to each other. The helical structure of hExo1 complexed with its DNA substrate contributes to rigorous regulation of the cleavage site (Orans *et al.,* 2011). Thus, the helical structure of PRORP1 is also likely to play an important role in its catalytic reaction. It is thus required to change the mutational sites, where mutations do not lead to disrupt its catalytic pocket. Furthermore, it is also required to improve the binding affinity between PRORP1 and its RNA substrate by UV crosslink or mutagenesis of the amino acids located on the binding interface between PRORP1 and its RNA substrates.

Chapter Ⅴ

General Consideration

Since the discovery of ribozymes, exploration of similarities and difference between ribozymes and protein enzymes has linked to fundamental catalytic mechanisms, including metal ion and acid/base catalytic modes, and coupling binding energy for catalysis (Guerrier-Takada *et al.*, 1983; Steitz and Steitz, 1993). For instance, on the basis of the information in respect with the comparative study between ribozymes *(e.g.* group Ⅱ intron and ribonuclease P) and protein enzymes (*e.g.* the 3', 5' exo-nucleases and DNA polymerase Ⅰ from *Escherichia coli*), Steiz and Steiz postulated that the catalytic mechanism in phosphoryl transfer reaction is facilitated by two divergent metals at a distance of 3.9 Å (known and described as two metal ion mechanism) (Steiz and Steiz, 1993). Furthermore, these comparative studies allow us to study the evolution of the biological catalysts.

Rupert and Ferré-D'Amaré solved the crystal structure of the hairpin ribozyme in uncleaved, intermediated, and cleaved states, and also compared them with that of ribonuclase A (RNase A) to study the evolution of biological catalysts (Rupert and Ferré-D'Amaré, 2001; Fedor, 2000; Wyckoff *et al.*, 1967). The hairpin ribozyme is a catalytic RNA derived from the satellite RNA of a plant virus, where hairpin ribozyme-mediated self-cleavage and ligation reactions participate in processing RNA intermediates (Rupert and Ferré-D'Amaré, 2001). The cleavage reaction leads to generate the RNA substrates with 2', 3'-cyclic phosphate and 5'-hydroxyl terminus as RNase A does. RNase A catalyzes the cleavage of the phosphodiestel-bond between the 5'-ribose of a nucleotide and the phosphate group attached to the 3'-ribose of a pyrimidine nucleotides (Wyckoff *et al.*, 1967; Birdsall *et al.*, 1992). The comparative studies indicated that the catalytic site of hairpin ribozyme is highly similar to that of RNase A (Figure Ⅴ-1). In the case of hairpin ribozyme, G8 and A9 on the stem bind to the base of A^{-1} *via* hydrogen bond networks, leading to turn it outward from the line of the substrate structure (Rupert and Ferré-D'Amaré, 2001). On the other hand, A9, which locates on the stem B in hairpin ribozyme, recognizes the base of G^{+1}, which follows the scissile phosphate, by utilizing π-π stacking interactions. In the case of RNase A, dT^{+1} is

filliped out by the hydrogen bonding between the base of dT^{+1} and the side chains of Lys41 and His12. The Lys41 and His12 in RNase A locate in a position analogous to G8 and A9 in hairpin ribozyme, respectively. Likewise, the side chain of His116 locates in a position analogous to the base of A38 (Figure Ⅴ-1) (Rupert and Ferré-D'Amaré, 2001; Birdsall *et al*., 1992). Thus, these results indicated that, as far as RNA-cleavage, the catalytic mechanism of the ribozyme (hairpin ribozyme) is highly similar to that of the protein enzyme (RNase A). This similarity possibly suggests that protein enzymes have evolved in the convergent manner from RNA enzymes (Ferré-D'Amaré and Scott, 2011). However, the observation is not enough to discuss the biocatalyst's evolution, because the hairpin ribozyme and RNase A do not catalyze an identical reaction to the identical substrates: the hairpin ribozyme reactions self-processing (in-*cis* acting), whereas RNase A cleaves the scissile phosphate bond in front of pyrimidine nucleotides in its RNA substrates (in-*trans* acting) (Ferré-D'Amaré and Scott, 2011). Thus, the detailed comparison of RNase P and PRORPs has been expected to shed light on the fundamental issues about the evolution of biocatalysts.

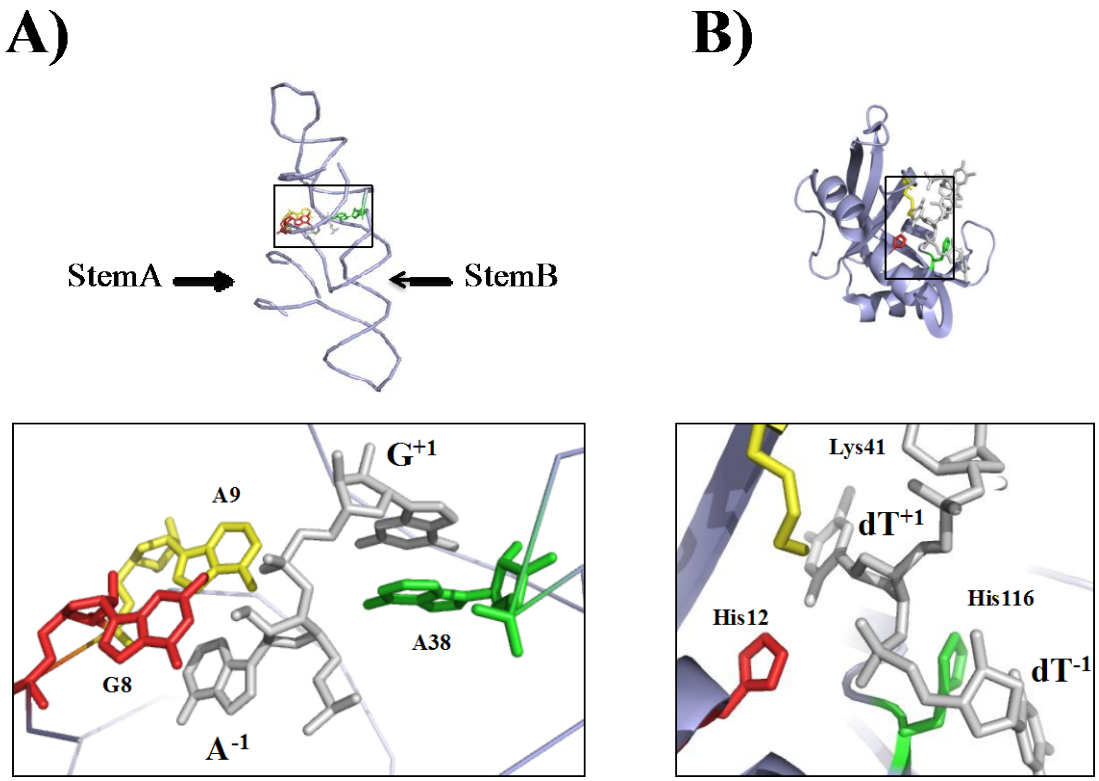

Figure Ⅴ-1: The catalytic mechanism of the hairpin ribozyme is similar to that of RNase A. (A) The active site of the hairpin ribozyme (PDB ID: 1M5K) (Rupert *et al.*, 2002); (B) The active site of the RNase A (PDB ID: 1RTA) (Birdsall and McPherson, 1992).

In order to elucidate the mechanism of substrate recognition of PRORPs, and also to gain insight into molecular evolution of biocatalysts, I first prepared PRORP1 des38, PRORP2 desN9, and PRORP3 desN10 from *A. thaliana* and characterized their enzymatic prosperities by utilizing the pre-tRNA^{Phe} processing activity assay. Furthermore, their catalytic mutants were prepared and characterized in the same manner as their native-types. In the chapter Ⅱ, the results obtained are presented. These results showed that the catalytic mechanism of NYN domain in PRORP utilizes two metal ion mechanism as PIN and Flap nucleases do (Steiz and Steiz, 1993; Anantharaman and Aravind, 2006).

During the purification of PRORP1 des38, degraded polypeptides were generated by trypsin-like proteinases. Limited trypsin and chymotrypsin proteolysis of PRORP1 des38, combined with N-terminal sequencing, revealed the two fragments, containing F1 (pos. 89-572) and F2 (pos. 220-572). Base on the observations, PRORP1 des89 and PRORP1 des220, where the N-terminal 89 and 220 amino acids were individually truncated, were prepared and characterized in the same manner as PRORP1 des38. Additionally, the successive deletion mutants, containing PRORP1 des129 and PRORP1 des149, where the PPR1 and PPR1-2 were truncated, respectively, were prepared and characterized the pre-tRNA^{Phe} processing activity. These results showed that at least PPR1 motif plays an important role in the catalysis. Moreover, I prepared two mutants PRORP1 des89 (-PPR2) and PRORP1 des89 (-PPR3), in which residues 130-169 and 175-204 comprising PPR2 and PPR3, respectively, were deleted from PRORP1 des89. The result indicated that PPR2 and PPR3 in addition to PPR1 play a crucial role in the recognition of pre-tRNAs in catalysis.

In the chapter Ⅲ, to predict the target sites for PPR motifs in PRORP1 using bioinformatic approach by PPR recognition code combined with characterization of the pre-tRNA^{Phe} mutants and PRORP1 mutants (Barkan *et al*., 2012; Yagi *et al.*, 2013a; Yagi *et al.,* 2013b). The obtained sequence C-A/G-A-U by the *in silico* bioinformatic study and the biochemical assay were good in agreement with the sequence C56-A57-A58-A59 at TΨC loop in pre-tRNA^{Phe}. It was thus likely that PPR2-PPR3-PPR4-PPR5 in PRORP1 recognize the nucleotide sequence C56-A57-A58-A59 in pre-tRNA^{Phe}. To verify this hypothesis, I characterized the pre-tRNA^{Phe} mutants, including C56, A57, A58, and A59 were in turn replaced with the

other three nucleotides. This result showed that PPR5 motif does not bind A59 at TΨC loop in pre-tRNA^{Phe}, whereas PPR2-PPR3-PPR4 recognizes C56-A57-A58 on TΨC loop in the pre-tRNA^{Phe}, respectively. To confirm this issue further, I prepared PRORP1 des38 NSR mutants, where the 4 residues at PPR2, PPR3, and PPR4 were substituted another amino acids. These mutants were characterized in terms of the pre-tRNA^{Phe} processing activity, leading to the importance of PPR2 and PPR3 rather than PPR4 in pre-tRNA base recognition. This assumption is good in agreement with the construction of the model of PRORP1 in complex with tRNA. In the model, the side chains of Tyr133, Asn136, and Arg175 located on PPR2 motif recognize C56, and those of Ala177, Thr180, and Arg212 on PPR3 motif are located in proximity to A57, in which the side chains of Asn136 and Thr180 at the 4th position of PPR2 and PPR3 bind to C56 and A57 on TΨC loop in tRNA, respectively. This proposed recognition mechanism is slightly distinct from that by bacterial RNase P. In the case of RNase P, A112 and G147 in S-domain bind to C56 on TΨC loop and G19 located on D loop by utilizing π-π stacking interactions, whereas PRORP1 recognizes the C56 and A57 located on TΨC loop by utilizing hydrogen bonding networks (Figure -Ⅶ). RNA molecules almost adopt π-π stacking interactions for their structure arrangement, its stabilization, and RNA-RNA interactions, whereas protein enzymes tend to adopt hydrogen bonding to interact with other molecules. Furthermore, the recognition mechanism of tRNA elbow by π-π stacking interactions is analogous to the recognition mode between tRNA and L1 stalk of the ribosomal E-site, and that between T box riboswitch Stem I domain and it cognate tRNA (Zhang and Ferré-D'Amaré, 2013; Korostelev *et al.,* 2006). Because these ribonucleoproteins and T-boxes are not related, this recognition strategy has evolved in convergent. It is therefore likely that the distinct recognition mode of them is probably connected to the characteristic prosperities of RNA and amino acid molecules. It was further described that an *R*p-phosphothioate modification of the scissile bond has no influence on the cleavage activity by PRORP1, even though RNase P is highly sensitive to the modification (Pavlova *et al.,* 2012). In general, protein-based phosphoryl transfer enzymes（*e.g.* RNase H and *Eco*R Ⅴcontacts the pro-*Sp* oxygen of the scissile phosphate, whereas RNA-based enzymes（*e.g.* Hammerhead ribozyme and spliceosome）does the pro-*Rp* oxygen of its phosphate, indicating that the characteristic prosperities of RNA and amino acids probably select the binding oxygen of the scissile

bond (Howard *et al*., 2015). It is therefore likely that PRORP1 catalyzes the pre-tRNA processing in a manner distinct from RNase P. It could thus be assumed that the proteinaceous RNase P, PROPR1, has evolved independently of the molecular mechanism by which RNA-based RNase P cleaves the pre-tRNA in catalysis.

In the chapter Ⅳ, I attempted to solve the crystal structure of PRORP1 in complex with pre-tRNA or pre-nad6 in order to gain more information about the mechanism for RNA recognition of PRORP1. I synthesized and purified the pre-tRNA and pre-nad6 for crystallization by utilizing *in vitro* transcription system and several steps of column chromatographies, respectively. To confirm the binding affinity between PRORP1 des76 D399A, where the N-terminal 76 residues were truncated and the catalytic residue Asp399 was replaced with Ala (denoted D399A, hereafter), and its RNA substrates, I performed with gel mobility shift assay and analytical gel filtration column chromatography. Theses results showed that D399A binds to pre-tRNA and pre-nad6 as a stable complex. Based on these data, I searched the crystallization condition of D399A in complex with pre-tRNA and D399A in complex with pre-nad6 at 4 ℃ and 20 ℃ by utilizing commercially available screening kits. The crystals, containing D399A alone or pre-tRNA alone, were produced in many conditions, whereas that of D399A complexed with its RNA substrate was not appeared at least in our screened conditions. To overcome this obstacle, its isozyme PRORP2 desN9 was used for crystallization with its substrate RNAs. However, crystals have not been obtained until now.

When the crystal structure of PRORP1 (pos. 95-570) (Howard *et al.,* 2013) was compared with that of D399A, the overall structures are similar to each other. However, there are two key differences between PRORP1 and D399A. First, PRORP1 adopts a relaxed conformation, whereas the PPR domain and NYN domain of D399A are located closer together. This structural difference is probably caused by conformational change of the loop located on the central domain (pos. 550-563). Second, the catalytic pocket of D399A is collapsed, leading to prevent two metal ions necessary for its catalytic reaction from their coordination. It is attributed to the observation that the helical structure (pos. 399-406) is changed to the random coil structure. When the crystal structure of PRORP1 is superimposed on that of human exonuclease1 (hExo1) in complex with its DNA substrate, the topology of the helical structures is similar to each

other. The helical structure of hExo1 complexed with its DNA substrate contributes to rigorous regulation of the cleavage site (Orans *et al.,* 2011). Thus, the helical structure of PRORP1 is also likely to play an important role in its catalytic reaction. It is thus required to change the mutational sites, where mutations do not lead to disrupt its catalytic pocket. Furthermore, it is also required to improve the binding affinity between PRORP1 and its RNA substrate by UV crosslink or mutagenesis of the amino acids located on the binding interface between PRORP1 and its RNA substrates. Further studies, including crystal structures of PRORPs bound to their substrates will be needed to find out the precise contribution of PPR domains in PRORPs to substrate recognition. No matter whether it is strict convergence of structure and function or not, exploration of similarities and variation in PRORP and RNase P will give rise to understanding biocatalyst's evolution as well as enzymatic mechanisms.

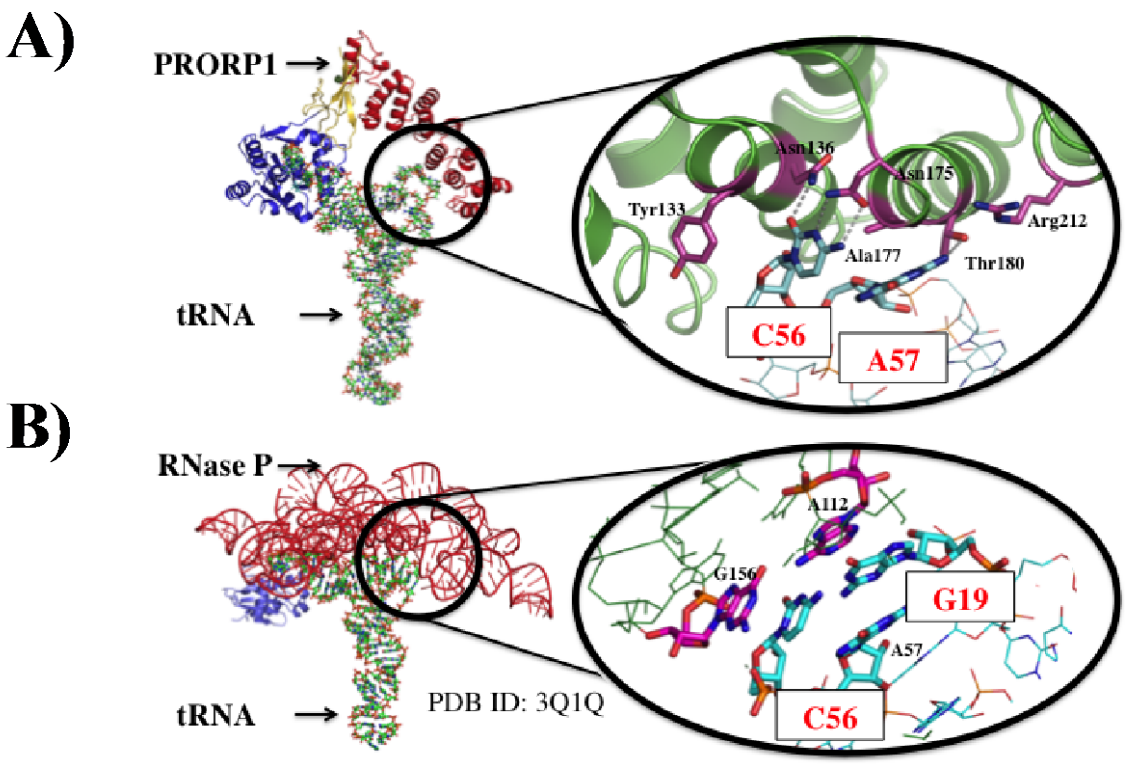

Figure V: The RNA recognition mechanism of PRORP1 is distinct from that of RNase P. (A) The model structure of PRORP1 in complex with tRNA (*left*). The PPR domain, the central domain, NYN domain, and tRNA are colored red, yellow, blue, and green, respectively. Close-up view of the mutual interaction site (*right*). The NSRs of PPR2 and PPR3 are colored purple and depicted as stick representation. The bases of C56 and A57 located on TΨC loop in tRNA are colored cyan and marked by stick representation. (B) The crystal structure of *T. maritima* RNase P holoenzyme complexed with its tRNA. The RNase P RNA, the protein subunit, and tRNA are colored red, blue, and green, respectively (*left*). Close-up view of the mutual interaction site (*right*). G156 and A112 located on S-domain in RNase P RNA are marked by stick representation and colored magenta. The bases of C56 on TΨC loop and G19 on D loop in tRNA are colored cyan and marked by stick representation.

REFERENCES

Adams PD., Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung, LW, Kapral GJ, Grosse-Kunstleve, RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson JS, Terwillinger TC, and Zwart PH. PHENIX: a comprehensive Python -based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* **66** (2010) 213-221.

Anantharaman V and Aravind L, The NYN domains, novel predicted RNases with a PIN domain-like fold, *RNA Biol.* **3** (2006) 18-27.

Ban N, Nissen P, Hansen J, Moore PB, and Steiz TA, The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289** (2000) 905-920.

Barkan A, Rojas M, Fujii S, Yap A, Chong YS, Bond CS, and Small I. A combinational amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet*. **8** (2012) e1002910.

Birdsall DL and A McPherson, Crystal structure disposition of thymidylic acid tetramer in complex with ribonuclease A. *J. Biol. Chem.* **31** (1992) 22230-22236.

Cech TR, The RNA Worlds in Context, in Atkins JE, Gesteland RF, Cech TR (eds): RNA Worlds -From Life's Origins to Diversity in Gene Regulation- (Cold Spring Harbor Laboratory Press, Cold Spring Harbor 2011).

D'Andrea LD and Regan L, TPR proteins: the versatile helix. *TRENDS in Biochemical Sciences* **28** (2003) 655-662.

Easton LE, Shibata Y, and Lukavsky PJ, Rapid, non denaturing RNA purification using weak anion-exchange fast performance liquid chromatography. *RNA* **16** (2010) 647-653.

Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crsytallogr D Biol Crsytallogr.* **66** (2010) 486-501.

Esakova O and Krasilnikov AS, Of proteins and RNA: the RNase P/MRP family, *RNA* **16** (2010) 1725-1747.

Estella MG, Franck APV, and Paul FA, tRNA's modifications bring order to gene expression. *Current Opinion in Microbiology* **11** (2008) 134-140.

Fedor MJ, Structure and function of the hairpin ribozyme. *J. Mol. Biol.* **297** (2000) 269-291.

Ferré-D'Amaré AR and Scott WG, Small Self-cleaving Ribozymes, in Atkins JE, Gesteland RF, Cech TR (eds): RNA Worlds -From Life's Origins to Diversity in Gene Regulation- (Cold Spring Harbor Laboratory Press, Cold Spring Harbor 2011).

Fica SM, Truttle N, Novak T, Li N, Lu J, Koodathingal P, Dai Q, Staley JP, and Piccirilli JA, RNA catalyses nuclear pre-mRNA splicing. *Nature* **503** (2013) 229-234.

Forner J, Weber B, Thuss S, Wildum S, and Binder S. Mapping of mitochondrial mRNA termini in *Arabidopsis thaliana*: t-elements contribute to 5' and 3' end formation. *Nucleic Acids Res.* **35** (2007) 3676-3692.

Fujii S and Small I, The evolution of RNA editing and pentatricopeptide repeat genes, *New Phytol.* **191** (2011) 37-47.

Gobert A, Gutmann B, Tashner A, Gossringer M, Holzmann J, Hartmann RK, Rossmanith W, and Giege P. A single Arabidopsis organellar protein has RNase P activity. *Nat. Struct. Mol. Biol.* **17** (2010) 740-744.

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

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

Gutmann B, Gobert A, and Giegé P, PRORP proteins support RNase P activity in both organelles and the nucleus in Arabidopsis. *Genes Dev.* **26** (2012) 1022-1027.

Hammani K, Bonnard G, Bouchoucha A, Gobert A, Pinker F, Salinas T, and Giege P, Helical repeats modular proteins are major players for organelle gene expression. *Biochemie.* **100** (2014) 141-150.

Holley RW, Apgar J, Everett GA, Madison JT, Marquisee M, Merrill SH, Penswick JR, and Zamir A, Structure of a Ribonucleic Acid. *Science* **147** (1965) 1462-1465.

Holzmann J, Frank P, Loffler E, Bennett KL, Gerner C, and Rossmanith W. RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell* **135** (2008) 462-474.

Hoogstraten CG and Sumita M, Structure-function relationships in RNA and RNP enzymes: recent advances. *Biopolymers* **87** (2007) 317-328.

Hou YM, CCA addition to tRNA: implications for tRNA quality control. *IUBMB Life* **62** (2010) 251-260.

Howard MJ, Klemm BP, and Fierke CA. Mechanic studies reveal similar catalytic strategies for phosphodiester bond hydrolysis by protein-only and RNA-dependent ribonuclease P. (2015) *J. Biol. Chem.* **290** (2015) 13454-13564.

Howard MJ, Lim WH, Fierke CA, and 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.* **109** (2012) 16149-16154.

Imai T, Nakamura T, Maeda T, Nakayama K, Gao X, Nakashima T, Kakuta Y, and

Kimura M. Pentatricopeptide repeat motifs in the processing enzyme PRORP1 in *Arabidopsis thaliana* play a crucial role in recognition of nulceotide bases at TΨC loop in precurusor tRNAs. *Biochem. Biophys. Res. Commun.* **450** (2014) 1541-1546.

Jarrous N and Goparan V, Archaeal/eukaryal RNase P: subunits, functions and RNA diversification. *Nucleic. Acids Res.* **38** (2010) 7885-7894.

Kazantsev AV and Pace NR, Bacterial RNase P: a new view of an ancient enzyme. *Nat. Rev. Microbiol.* **4** (2006) 729-740.

Ke J, Chen RZ, Ban T, Zhou XE, Gu X, Tan MH, Chen C, Kang Y, Brunzelle JS, Zhu JK, Melcher K, and Xu HE, Structural basis for RNA recognition by a dimeric PPR-protein complex. *Nat. Struct. Mol. Biol.* **20** (2013) 1377-1382.

Kersten H, Alteration of tRNA modification in eukaryotes: caused and consequences. *Recent Results Cancer Res.* **84** (1983) 255-263.

Kiefer F, Arnold K, Künzli M, Bordoli L, and Schwede T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* **37** (2009) D387-D392.

Kimura M and Kakuta Y, Structural biology of the ribonuclease P in the hyperthermophilic archaeon *Pyrococcus horikoshii* OT3, in T. Satyanarayana *et al.* (Eds.), Microorganisms in Sustainable Agriculture and Biotechnology, Springer Science+Business Media B.V., 2012.

Korostelev A, Trakhanov S, Laurberg M, and Noller H, Crystal structure of a 70S ribosome–tRNA complex reveals functional interactions and rearrangements. *Cell* **126** (2006) 1065–1077.

Kunzmann A, Brennicke A, and Marchfelder A, 5' end maturation and RNA editing have to precede tRNA 3' processing in plant mitochondria. *Proc. Natl. Acad. Sci. U. S. A.* **95** (1998) 108-113.

Laemmli UK, Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227** (1970) 680-685.

Lai LB, Bernal-Bayard P, Mohannath G, Lai SM, Gopalan V, and Vioque A. A functional RNase P proiten subunit bacterial origin in some eukayotes. *Mol. Genet. Genomics.* **286** (2011) 359-369.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heafold A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov J, Miranda C, Morris W, Naylor J, C Raymond, Rosetti M, Santos R, Sheridan A, Sougnez C, Stage-Thomann N, Stojanovic N, Subramanian A, and Wyman D, Initial sequencing and analysis of the human genome. *Nature* **409** (2001) 860-921.

Laskowski RA, MacArthur MW, Moss DS, and Thornton JM, PROCHECK - a program to check the stereochemical quality of protein structures. *J. App. Cryst.* **26** (1993) 283-291.

Leader P and Nirenberg MW, RNA Codewords and Protien Synthase Ⅲ. On the Nucleotide sequence of a Cystein and a Leucine RNA Codeword. *Proc. Natl. Acad. Sci. U. S. A.* **52** (1964) 1521-1529.

Madison-Antenucci S, Grams J, and Hajduk SL, Editing machines: the complexities of trypanosome RNA editing. *Cell.* **108** (2002) 435-438.

Manna S, An overview of pentatricopeptide repeat proteins and their applications. *Biochimie.* **113** (2015) 93-99.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, and Read RJ, Phase crystallographic software. *J. Appl. Cryst.* **40** (2007) 658-674.

Miller SL and Urey HC, Organic Compound Synthesis on the Primitive Earth. *Science* **130** (1959) 245-251.

Morgan EA, Ikemura T, Lindahl L, Fallon AM, and Nomura M, Some rRNA operons in *E. coli* have tRNA genes at their distal ends. *Cell* **13** (1978) 335-344.

Möri M and Marchfelder, The final cut. The importance of tRNA 3'-processing. *EMBO Rep.* **2** (2001) 17-20.

Nakamura T, Yagi Y, and Kobayashi K, Mechanistic insight into pentatricopepetide repeat proteins as sequence-specific RNA-binding protein for orgarnellar RNAs in plants. *Plant Cell Physiol.* **53** (2012) 1171-1179.

Nawel H, Nadege A, Martine Q, Planchard N, Nathalie V, Jennifer D, Colas des Francs-Small C, and Mireau H. The MTL1 Pentatricopeptide Rpeat Protein is Required for Both Translation and Splicing of the Mitochondrial NADH Dehydrogenase Subunit 7 mRNA in Arabidopsis. *Plant Physiol.* (2015) pii: pp.01591.2015. [Epub ahead of print]

Nishikura K, Functions and Regulation of RNA Editing by ADAR Deaminases. *Annu. Rev. Biochem.* **79** (2010) 321-349.

Orans J, McSweeeney EA, Lyer RR, Hast MA, Helling HW, Modrich P, and Beese LS, Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. *Cell* **145** (2011) 212-223.

Pan H, Agarwalla S, Mopustakas DT, Finer-Moore J, and Stroud RM, Structure of tRNA pseudouridine synthase TruB and its RNA complex: RNA recognition through combination of rigid docking and induced fit. *Proc. Natl. Acad. Sci. U. S. A.* **100** (2003) 12648-12653.

Pavlova LV , Gossringer M, Weber C, Buzet A, Rossmanith W, and Hartmann RK,

tRNA processing by protein-only versus RNA-based RNase P: kinetic analysis reveals mechanistic differences. *ChemBioChem*. **13** (2012) 2270-2276.

Pinker F, Bonnard G, Gobert A, Gutmann B, Hammani K, Sauter C, Gegenheimer PA, and Giegé P, PPR proteins shed a new light on RNase P biology. *RNA Biol.* **10** (2013) 1457-1468.

Reiter NJ, Osterman A, Torres-Larios A, Swinger KK, Pan T, and Mondragon A, Structure of a bacterial ribonuclase P holoenzyme in complex with tRNA, *Nature* **468** (2010) 784-789.

Robertus JD, Ladner JE, Finch JT, Rhodes D, Brown RS, Clark BFC, and Klug A, Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature* **250** (1974) 546-551.

Rupert PB and Ferré-D'Amaré AR, Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature* **410** (2001) 780-786.

Schofield P and Williams KR, Purification and some prosperities of *Escherichia coli* tRNA nucleotidyltrasferase. *J. Biol. Chem.* **252** (1077) 5584-5588.

Segura J, Oliva B, and Fernandez-Fuentes N. CAPS-DB: a structural classification of helix-capping motifs. *Nucleic Acids Res.* **40** (2011) Database issue D479-D485.

Shi H and Moore PB, The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited. *RNA* **6** (2000) 1091-1105.

Shiomi H, Fathoming evolution of RNA regulation in cells., in Shiomi H, Inada T, Tomari Y, and Hirose T (eds): 生命を統合する RNA -その秘められた役割と制御機構 (分子進化・サイレンシング・non-coding RNA から RNA 修飾・編集・RNA-タンパク質間相互作用まで), pp 12 - 22 (実験医学 Vol. 31 No.7 増刊, 羊土社 2013).

Steitz TA and Steitz JA, A general two-metal-ion mechanisms for catalytic RNA. *Proc Natl Acad Sci U. S. A.* **90** (1993) 6498-6502.

Temin HM and Baltimore D, RNA-directed DNA synthesis and RNA tumor viruses. *Adv. Virus Res.* **17** (1972) 129-186.

Torres-Larios A, Swinger KK, Pan T, and Mondragon A, Structure of ribonuclease P is the catalytic subunit of the enzyme. *Curr. Opin. Struct. Biol.* **16** (2006) 1-6.

Trotta CR, Miao F, Arn EA, Stevens SW, Ho CK, Rauhut R, and Abelson JN, The yeast tRNA splicing endonuclease A: A tetrameric enzyme with two active subunits homologous to the archaeal tRNA endonucleases. *Cell* **89** (1997) 849-858.

Vagin A and Teplyakov A, An approach to multi-copy search in molecular replacement, *Acta Crstallogr. Sect. D Biol. Crystallogr.* **56** (2000) 1622-1624.

Vilardo E, Nachbagauer C, Buzet A, Taschner A, Holzmann J, and Rossmanith W, A subcomplex of human mitochondrial RNase P is a bifunctional methyltransferase-extensive moonlighting in mitochondrial tRNA biogenesis. *Nucleic Acid Res.* **40** (2012) 12583-121593.

Walker SC and Engelke DR, Ribonuclease P: the evolution of an ancient RNA enzyme, *Crit. Rev. Biochem. Mol. Biol.* **41** (2006) 77-102.

Wang X, McLachlan J, Zamore PD, and Hall TMT, Modular recognition of RNA by a human pumilio-homology domain. *Cell* **110** (2012) 501-512.

Westaway SK, Phizicky EM, and Abelson J, Structure and function of the yeast tRNA ligase gene. *J. Biol. Chem.* **263** (1988) 3171-3176.

Wyckoff HW, Hardman KD, Allewel NM, Inagami T, Johnson LN, and Richard FM, The structure of ribonuclease-S at 3.5 Å resolution. *J. Biol. Chem.* **242** (1967)

3984-3988.

Yagi Y, Hayashi S, Kobayashi K, Hirayama T, and Nakamura T. Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. *PloS ONE* **8** (2013) e57286.

Yagi Y, Tachikawa M, Noguchi H, Satoh S, Obokata J, and Nakamura T, Pentatricopeptide repeat proteins involved in plant organellar RNA editing. *RNA Biol.* **10** (2013) 1236-1242.

Yin P, Li Q, Yan C, Liu Y, Liu J, Yu F, Wang Z, Long J, He J, Wang HW, Wang J, Zhu JK, Shi Y, and Yan N, Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* **504** (2013) 168-171.

Zeytuni N and Zarivach R, Structural and Functional Discussion of the Tetra-Trico-Peptide Repeat, a Protein Interaction Module, *Structure Review* **20** (2012) 397-405.

Zhang J and Ferre-D Amare AR, co-crystal structure of a T-box riboswitch stem I domain in complex with its cognate tRNA. *Nature* **500** (2013) 363-366.