MECHANISMS OF MICROBIAL DNA SENSING BY THE AIM2 INFLAMMASOME AND STRUCTURAL STUDY OF POLYNUCLEOTYDYL TRANSFERASE1 (PNT1) FROM *ARABIDOPSIS THALIANA*

A Thesis

by

MIN WOO SUNG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Pingwei Li
Committee Members,	Frank Raushel
	Margaret E. Glasner
Head of Department,	Gregory Reinhart

December 2012

Major Subject: Biochemistry

ABSTRACT

AIM2 (absence in melanoma 2) can sense foreign cytosolic double-stranded DNA (dsDNA) through the dsDNA binding HIN-200 domain at the C-terminus. Once dsDNA is bound to HIN-200 domain, AIM2 can activate the AIM2 inflammasome, resulting in maturation of pro-interleukin-1β by activation of caspase-1. To investigate the mechanism of DNA binding, HIN-200 domain of mouse AIM2 bound to 15bp, 18bp, 20bp and 30bp dsDNAs were purified and crystallized. Diffraction data for four different crystals were collected to about 4Å resolution. Interestingly, all the crystals were in the same cubic space group *I23* or *I2$_1$3* with almost same unit-cell parameters. Mutagenesis of HIN-200 domain of mouse AIM2 and DNA binding studies revealed amino acid residues involved in DNA binding. These residues were compared with dsDNA binding residues identified from the structure of HIN-200 domains of human AIM2, which suggested mouse AIM2 uses a similar dsDNA binding surface as human AIM2.

Polynucleotidyl transferase (PNT1) was discovered in the proteomics study of Argonaute10 (AGO10). AGO10 was known as a key regulator of shoot apical meristem (SAM) maintenance in *Arabidopsis*. A recent study reported that AGO10 can regulate the developmental processes in SAM by repressing two miRNAs, miR166/165, and bind to these two miRNAs. Sequence analysis of PNT1 showed that it belongs to DnaQ-like 3'-5' exonuclease family, and hypothesized to be involved in RNA regulation with AGO10. PNT1 was purified and strong 3´-5´ exonuclease activity was revealed by activity test. Purified protein was crystallized, and the solved structure showed hexameric ring formation. Investigation of the crystal structure of PNT1 revealed

presumed active site which was found from other 3´-5´ exonuclease homologues. Through mutagenesis of four residues in the presumed active site, Glu52 was identified as a key residue for its 3´-5´ exonuclease activity.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Pingwei Li, for his guidance and support throughout the research project. He provided me opportunities to work on a good project and participate in other projects. I also appreciate my committee members, Dr. Frank Raushel, Dr. Margaret E. Glasner and Dr. Feng Qiao for their encouragement and support throughout the course of this research.

Thanks also go to my classmates in my year and colleagues in our lab, Cheng Lu, Tylan watts and Chang Shu. I also want to extend my gratitude to the department faculty and staff for helping me to prepare all the paper works.

Finally, thanks to my family and friends for their encouragement and patience.

NOMENCLATURE

AIM2	absent in melanoma 2; a prefix “m” or “h” denotes the *Mus musculus* or *Homo sapiens* homologue, respectively
AIM2-HIN	HIN-200 domain of AIM2
ASC	apoptosis speck protein with a caspase recruitment domain
HSV	herpes simplex virus; a prefix “ds” denotes double-stranded DNA
dsDNA	double-stranded DNA
DTT	dithiothreitol
IPTG	Isopropyl β-D-1-thiogalactopyranoside
24 n.t. ssRNA	24 nucleotide single stranded RNA

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION*

Structural study of the proteins is one of the key studies to scrutinize the function of those proteins. I studied about AIM2 (absence in melanoma 2) protein from the mouse and Polynucleotidyl transferase (PNT1) from *Arabidopsis thaliana* focusing on crystal structure.

During infection with viral and bacterial pathogens, foreign nucleic acids can trigger strong immune defenses in host cells (Barber, 2011; Hornung and Latz, 2010; Takeuchi and Akira, 2010). Several pattern recognition receptors (PRRs) which can sense patterns of microbial nucleic acids and initiate innate immune response have been identified (Hornung and Latz, 2010; Takeuchi and Akira, 2010; Vilaysane and Muruve, 2009).

AIM2 is one of the PYHIN protein family, comprised of a N-terminal PYRIN domain and a C-terminal HIN-200 domain. The HIN-200 domain is reported to have direct sensing of microbial DNA in the cytoplasm and signals to initiate the inflammatory response (Hornung et al., 2009; Schattgen and Fitzgerald, 2011). AIM2 interacts with the down stream adaptor protein ASC (apoptosis speck protein with a caspase recruitment domain) through a homotypic PYRIN-PYRIN interaction. Upon cytosolic dsDNA binding of the HIN-200 domain, AIM2 forms inflammasome with ASC and facilitates the activation of caspase-1 (Hornung and Latz, 2010; Schroder and

---

* Reprinted with permission from "Crystallographic characterization of mouse AIM2 HIN-200 domain bound to a 15 bp and an 18 bp double-stranded DNA." By M. W. Sung, T. Watts and P. Li. 2012. Acta Cryst. Volume 68, Part 9, Pages 1081-1084. Copyright 2012 by IUCr Journals.

Tschopp, 2010). Pro-interleukin-1β is cleaved by caspase-1 resulting in secretion of activated interleukin 1β (IL-1β) and initiates immune responses such as the recruitment of immune cells to the sites of infection. One of the key cytokines in immune responses is IL-1β, which has an essential role in the regulation of immune cell proliferation, differentiation and apoptosis (Dinarello, 2009; Sims and Smith, 2010). The pivotal role of AIM2 in ASC dependent caspase-1 cleavage and the processing IL-1β was confirmed by exogenous expression of AIM2 in human cell and AIM2-deficient mice (Fernandes-Alnemri et al., 2010; Rathinam et al., 2010; Schattgen and Fitzgerald, 2011). The crystal structures of HIN-200 domain of human AIM2 bound to 19 bp dsDNA were determined recently. The structure reported provided insight into the mechanism of DNA binding by AIM2-like receptors, and the report was supported with assays for innate responses in the cell by transfected dsDNA. (Jin et al., 2012).

To address the detailed function of AIM2 in DNA sensing, HIN-200 domain of *Mus musculus* AIM2 (mAIM2-HIN), which is homologous to HIN-200 domain of human AIM2, was expressed and purified from *Escherichia coli* for further studies. Mouse AIM2 formed complexes with different lengths of dsDNA and was revealed from the analytical gel column. Since it has been shown that the mAIM2-HIN can bind to dsDNA, mutagenesis was performed for 17 positively charged residues including 2 control residues based on mAIM2-HIN/dsDNA model of the complex. 13 residues that do not bind to dsDNA and 2 residues that showed partial dsDNA binding were identified using analytical gel column, then these results were confirmed by Electrophoretic Mobility Shift Assay (EMSA).

Here we report on the mAIM2-HIN in complex with 15 bp, 18 bp, 20bp and 30bp dsDNA. Extensive crystallization screening for different lengths of mAIM2-HIN constructions with different lengths of dsDNAs was conducted. The same space group of mAIM-HIN complexes with variable lengths of dsDNAs may indicate formation of a special higher order complex with DNA that differs from that of human AIM2. The details of AIM2 inflammasome formation upon dsDNA binding would be provided by the structures of these complexes.

I have also studied about homologs of many exoribonucleases (RNases) which has been identified in diverse organisms. RNases play an important role in the metabolism of RNA such as degradation of mRNA, regulation of gene expression and quality control of mRNA biogenesis (Cheng and Patel, 2004; Wu et al., 2005; Zuo and Deutscher, 2001).

Polynucleotidyl transferase (PNT1), another interesting target protein in my thesis, was from *Arabidopsis thaliana* was discovered by Dr. Xiuren Zhang's laboratory through a proteomics study as a protein that interacts with Argonaute10 (AGO10). It is expected to have a role in RNA regulation with AGO10. AGO10 is known to have a key role in differentiation for group of undifferentiated cells in shoot apical meristem (SAM) (Lynn et al., 1999; Moussian et al., 1998) and the establishment of leaf polarity (Liu et al., 2009). AGO10 regulates these developmental processes by repressing two micro RNAs, miR166/165 (Liu et al., 2009) that target HD-ZIP III family genes (Jung and Park, 2007; Zhou et al., 2007). A recent study indicated that AGO10 specifically binds to miR166/165, and this binding capacity is necessary for proper SAM development (Zhu

et al., 2011). Sequence analysis of PNT1 revealed that it belongs to the DnaQ-like 3'-5' exonuclease family which is a member of the Ribonuclease H-like superfamily.

Uncharacterized protein At5g06450 from *Arabidopsis thaliana* (PNT2, PDB: 1VK0) is a homolog of PNT1, and its overall structure was compared with the exonuclease domain of WRN protein (WRN-exo) with r.m.s. deviation of 2.22Å. Structure of PNT2 and WRN-exo showed structural homolygy to DnaQ family from the result of DALI server (Perry et al., 2006) which can search homologous proteins based on the structure submitted. Using the structure of PNT2, DALI server found homologue of human 3´-5´-exoribonuclease (3'hExo), ERI-1 protein from *Caenorhabditis elegance* (PDB: 1W0H) (Wu et al., 2005) and human Poly(A)-specific ribonuclease (PARN) (PDB: 2A1R) (Cheng and Patel, 2004) which were studied for their crystal structures and catalytic mechanisms. Both structures have structurally conserved DEDD motifs that exist in their active sites bound with their substrates.

In the active site of ERI-1, two aspartates coordinate two magnesium ions that interact with rAMP. Additional histidine residue in the active site is involved in deprotonation of one water molecule which can nucleophilically attack a phosphorus atom and cleave the phosphodiester bond. Recent study showed that ERI-1 protein from *Caenorhabditis elegance* is known to negatively regulate RNA interference (Cheng and Patel, 2004).

DEDD residues in the active site of crystal structure of PARN directly interact with base of one trinucleotide without any divalent ions, though it was proposed that PARN has two metal ions for catalysis (Beese and Steitz, 1991). Through mutagenesis,

it was confirmed that additional histidine residue has a critical role in cleavage of poly (A). The study of ERI-1 and PARN provided clues that PNT1 is also expected to have a canonical active site as PNT2 and WRN-exo and to have 3´-5´ exonuclease activity.

To figure out the function of PNT1 and estimate the expected role of RNA regulation with AGO10, PNT1 was purified and crystallized. The solved structure of PNT1 was used for comparison with the structures of WRN-exo, ERI-1 and PARN which have been studied for their functional mechanisms. Based on the crystal structure analysis of the active site of PNT1, mutagenesis was performed for DEDD residues and determined that Asp52 has a key role in catalysis without utilizing metal ions.

CHAPTER II

MECHANISMS OF MICROBIAL DNA SENSING BY THE AIM2 INFLAMMASOME*

Objective

The way HIN-200 domain of AIM2 detects cytosolic dsDNA is the key step in immune response, including IL-1β secretion. To obtain further insights into the mechanism with which AIM2 senses and responds to dsDNA, structural knowledge of the HIN-200 domain of AIM2 complexed with dsDNA is necessary. This would provide crucial information about the explicit status of activation in the AIM2 inflammasome. Further studies about the mAIM2-HIN binding site and the stoichiometry of mAIM2-HIN/dsDNA complex can support the detailed information about the binding mechanism. To characterize the mechanism of dsDNA binding to the mAIM2-HIN, we set the following specific aims.

Specific aim1: Expression of mouse AIM2 HIN 200 domain

We postulated that mAIM2-HIN alone would be soluble and expressed from *E. coli*. As the expression of HIN-200 domain of human AIM2 was not successful previously, we selected the HIN-200 domain of mouse AIM2. Expression of mAIM2-HIN will be essential for further experiments. Moreover, expression of the truncated

---

* Reprinted with permission from "Crystallographic characterization of mouse AIM2 HIN-200 domain bound to a 15 bp and an 18 bp double-stranded DNA." By M. W. Sung, T. Watts and P. Li. 2012. Acta Cryst. Volume 68, Part 9, Pages 1081-1084. Copyright 2012 by IUCr Journals.

mAIM2-HIN designed by an analysis of secondary structure would be necessary to obtain a chance to improve the quality of crystals.

Specific aim2: Binding of double stranded DNA with the mAIM2-HIN

To test the binding of mAIM2-HINdomain with dsDNA and formation of a stable complex, we used gel filtration chromatography and Electrophoretic Mobility Shift Assay (EMSA). Though it was known that the HIN domain can sense foreign dsDNA in the cytosol (Choubey et al., 2000; DeYoung et al., 1997), appropriate lengths of dsDNA that bind to HIN-200 domain and the stability of the complex were unknown. A binding study with variable lengths of HSV (herpes simplex virus) dsDNA would be informative to figure out stoichiometry of dsDNA and postulate the binding site.

Specific aim3: Binding mechanism of the mAIM2-HIN

To resolve this issue, we performed mutagenesis in the presumed dsDNA binding site in mAIM2-HINdomain. A computational structure model of the mAIM2-HIN was generated and positively charged residues were selected based on the dsDNA bound model of the complex. Through mutagenesis and gel filtration assays, residues that are involved in dsDNA binding would be identified. Crystal structure of dsDNA bound to the mAIM2-HIN would provide more clues for the mechanisms of dsDNA binding.

Experimental Procedures

Cloning of the HIN domain of mouse AIM2

Full length mouse AIM2 including its N-terminal Pyrin domain and C-terminal HIN domain were cloned from cDNA provided by Open Biosystems. Forward and reverse primers were ordered from Integrated DNA Technology (IDT) with *Nde*I and *Sal*I restriction sites respectively. PCR was performed using Platinum Pfx DNA polymerase (Invitrogen, USA) with 1 $\mu\ell$ each primer (1 mg/$\mathrm{m}\ell$), 50 ng cDNA template, 2 $\mu\ell$ $MgSO_4$ (50 mM), 5 $\mu\ell$ ul 10x Pfx DNA polymerase buffer, 5 $\mu\ell$ 10x enhancer, 1 $\mu\ell$ dNTP (25 mM) in a total volume of 50 $\mu\ell$. The mixture was initially heated to 95 °C for 2 min, then cycled 26 times as following conditions: denaturing at 95 °C for 30 sec, annealing at 54 °C for 30 sec and extension at 68 °C (1 kb per min), followed by cooling at 4 °C. Product of PCR was purified by QIAquick PCR purification Kit (Qiagen, Netherlands). 50 U of *Nde*I and *Sal*I restriction enzymes were treated for both of insert and pET22b (+) vector with 5 $\mu\ell$ 10x restriction enzyme buffer in a final volume of 50 $\mu\ell$. Digested insert and vector fragments were purified by agarose gel and QIAquick Gel Extraction Kit (Qiagen, Netherlands). Ligation was carried out using 1 $\mu\ell$ T4 DNA ligase and 2 $\mu\ell$ 10x ligase buffer (New England BioLabs, USA) at room temperature for 5hrs with 4 $\mu\ell$ insert (100~150 ng/$\mu\ell$), 4 $\mu\ell$ vector(40~90 ng/$\mu\ell$) in total volume of 11 $\mu\ell$. The pET22b (+) expression vector including the gene coding the HIN-200 domain of mouse AIM2 (residues 135-354) was introduced into *Escherichia coli* BL21 (DE3) cells for protein expression.

Protein expression test

One colony was picked from the agar plate and inoculated in 5 ㎖ LB media to test expression of AIM2-HIN. Two sets of cells in 3 ㎖ LB media with 100 ㎍/㎖ ampicillin were cultured for 3~4 hours by shaking at 250 rpm until OD_{600} reached 0.8 at 310K. Both LB media sets were induced with 0.5 mM IPTG and cultured at 37℃ for 3 hours and at 15℃ overnight respectively. 1 ㎖ of cells was saved as glycerol stock at -80 ℃. 1 ㎖ of cells from each culture was harvested at 13,000 rpm for 1 min and the pellets were resuspended in 50 ㎕ lysozyme buffer (0.2 mg/㎖) and 1 ㎕ DNase (1 mg/㎖). Cells were lysed by repeated freezing in liquid N_2 and 5 min incubation at 310K. After centrifugation of cell lysate at 13,000 rpm for 2 min, both supernatant and pellet were loaded on 15% SDS-PAGE to check protein expression and solubility.

Large scale expression and purification of AIM2-HIN

To test the expression of mAIM2-HIN in a large scale, 50 ㎕ of glycerol stock culture was used to inoculate 50 ㎖ LB media with 100 ㎍/㎖ ampicillin and cultured by shaking at 37℃ until OD600=1.0. 10 ㎖ of inoculated cells were transferred to 1 L LB medium and cultured for 3 hrs. Then the LB medium was cooled on ice, and induced by the introduction of 500 μM IPTG when OD600=0.8. The cell culture was incubated for an additional 16 hrs at 15℃. Cells were harvested at 4000 rpm for 10 min at 4℃. Cell pellets were resuspended in 150 ㎖ pre-cooled buffer A (150mM NaCl, 20mM Tris,pH

7.5) and lysed by sonication. Cell lysate was centrifuged at 8,000 rpm for 10 min. at 4℃ to remove cell debris, and then cleaned of precipitates by centrifugation at 16,000 rpm for 10 min. at 4℃.

Supernatant was loaded to Ni^{2+} Sepharose High Performance beads (GE Health care, USA) pre-equilibrated with buffer A. C-terminal hexa-histidine tagged mAIM2-HIN proteins were bound to the column. Non-specifically bound contaminating proteins were removed by washing with 300 ㎖ of washing buffer (500 mM NaCl, 20 mM Tris, pH 7.5, 10 mM imidazole). The remaining bound proteins were eluted with elution buffer (150 mM NaCl, 20 mM Tris, pH 7.5, 10 mM imidazole). The elutions were pooled and 5 mM DTT was added. The sample was concentrated to 2 ㎖ by an Amicon Ultra-4 10 kDa centrifugal filter (Milipore) at 4000 rpm at 4℃. Then the sample was injected onto a gel filtration HiLoad 16/60 Superdex75 column (Amersham Biosciences) equilibrated with buffer A and resolved at 1 ㎖/min.

Preparation of double-stranded DNA

Four different lengths of the dsDNA to be used for mAIM2-HIN binding were prepared. The 15 bp sequence (HSV15) forward and reverse oligos were 5'-TAAGACACGATGCGA-3' and 5'-TCGCATCGTGTCTTA-3', respectively. The 18 bp sequence (HSV18) forward and reverse oligos were 5'-CAAGACACGATGCGATAG-3' and 5'-CTATCGCATCGTGTCTTG-3', respectively. The 20 bp sequence (HSV20) forward and reverse oligo was 5'- GACACGATGCGATAAAATCG -3' and 5'-

CGATTTTATCGCATCGTGTC -3', respectively. The 30 bp palindromic sequence (HSV30) oligos were 5'- CGTAAGACACGATGCGCATCGTGTCTTACG -3'. All DNA oligos were ordered from Integrated DNA Technology (IDT). Each of the forward and reverse oligos were resolved in buffer A to a final concentration of 1 mM. Oligos for a particular length of non-palindromic DNA were mixed in a 1:1 molar ratio of reverse and forward oligos annealed by incubating in a 95℃ water bath, then slowly cooled down to room temeperature over 30 min. Annealed DNA samples were injected onto a gel filtration HiLoad 16/60 Superdex75 column (Amersham Biosciences) and peak fractions corresponding to the size of dsDNA were pooled.

Electrophoretic mobility shift assay (EMSA)

8% Non-Denaturing Polyacrylamide Gel was prepared as follows: 1.6 ㎖ 40% acrylamide and bis N,N'-methylene-bis-acrylamide (29:1 acrylamide:bis ratio, Biorad) was mixed with 0.8 ㎖ 10x Tris-Borate (TB) buffer and diluted with sterile water up to the total volume of 7.1 ㎖. The mixture was polymerized by the addition of 80 ㎕ 10 % ammonium persulfate and 10 ㎕ TEMED. Variable lengths of dsDNAs were diluted to 20 μM and mAIM2-HIN proteins were diluted to 40 μM with buffer A. mAIM2-HIN complex with variable length of dsDNA in different ratios were prepared as follows: 1 ㎕ DNA samples were mixed with 0.5 ㎕, 1 ㎕, 2 ㎕, 3 ㎕, 4 ㎕, 5 ㎕, 10 ㎕ of protein in a total volume of 11 ㎕.

Computational model of mAIM2-HIN and site directed mutagenesis

Based on the crystal structure of HIN1 domain of Interferon-activable protein 204 (p204), which has 48 % of sequence identity with HIN domain of AIM2, mAIM2-HIN structure model was generated by SWISS-MODEL (Arnold et al., 2006). 226 residues of mAIM2-HIN (Pro135-Glu354) were used in sequence alignment with sequence of p204, and then a positively charged pocket from the expected dsDNA binding site of mAIM2-HIN model was aligned with 15 bp dsDNA. From the expected dsDNA binding site of mAIM2-HIN model, 15 positively charged residues were selected and two of other positively charged residues on non-DNA binding sites were selected as negative controls. Forward and reverse primers were designed and ordered from Integrated DNA Technology (IDT) to mutate each residue to glutamic acid. Primers were dissolved in sterile water to a final concentration of 100 ng/㎕. Mutagenesis PCR was conducted using Pfu Ultra DNA polymerase (2.5 U/㎕, Invitrogen, USA) in a mixture including 1 ㎕ of template, 1.25 ㎕ of each primer, 5 ㎕ of 10x Pfu Ultra reaction buffer, 1 ㎕ of dNTP and ddH_2O to a final volume of 50 ㎕. The reaction cycles were performed as follows: initial heating at 95℃ for 30 sec, the 16 cycles of heating at 95℃ for 30 sec, annealing at 54℃ for 1 min and elongation at 68℃ for 12 min, followed by a final extension at 68℃ for 10 min and cooling at 4℃. 1 ㎕ of DpnI restriction enzyme was added to the PCR product and incubated for 3 hrs at 37℃. *E. coli* BL21(DE3) competent cells were added with 1 ㎕ of DpnI treated DNA and incubated on ice for 15 min, heat shock for 45 sec at 42℃ and then put on ice for 10 min.

500 ㎕ of LB media was added to the cells and incubated for 1 hr at 37℃ by shaking at 250 rpm. Cells were centrifuged at 4,000 rpm for 1 min and 400 ㎕ of supernatant was discarded. The pellet was resuspended and 100 ㎕ of cells were spread on an agar plate containing 100 ㎍/㎖ ampicillin. Colonies were picked after overnight incubation of the plate at 37℃ and used to inoculate 5 ㎖ LB media contaning 100 ㎍/㎖ ampicillin.

Limited proteolysis assay

20 ㎍ of purified mAIM2-HIN protein and mAIM2-HIN/dsHSV15 complex were prepared and added with 2 ㎍ or 200 ng of trypsin respectively. The total volume was adjusted to 6 ㎕ using buffer A. Four sets of mixture were incubated at room temperature or 37℃ for 1 hour or 3 hours. Then each of the mixtures were mixed with 2x SDS-PAGE loading buffer, and denatured at 95℃ for 10 min and loaded on an 18% SDS PAGE gel.

Data collection and processing

A truncation form of mAIM2-HIN (residues 154-354 with a C264S mutation) was complexed with dsHSV15 and dsHSV18 and crystallized. mAIM2-HIN/HSV15 complex crystal measuring 0.22 × 0.22 × 0.24 mm was grown over 13 days at 4ºC and an mAIM2-HIN/HSV18 complex crystal measuring 0.25 × 0.23 × 0.23 mm was grown over 8 days at 4ºC. The AIM2-HIN/HSV15 complex crystal was transferred into a cryobuffer containing 30% PEG 400, 24% Jeffamine ED-2100, 0.1 M HEPES, pH 7.0.

The AIM2-HIN/HSV18 complex crystal was transferred into a cryobuffer containing 30% pentaerythritol ethoxylate 15/4, 0.05 M ammonium sulfate, 0.05 M BIS-Tris, pH 6.5. Both crystals were then mounted in cryo loops and flash frozen in liquid nitrogen. Diffraction data were collected at 100 K at the SSRL beamline 11.1 using a PILATUS 6M pixel array detector. The crystals were rotated through 90º with 0.2º oscillations per frame and 5 sec exposure per frame. The diffraction data were processed with the HKL2000 package

Results

Expression of mouse mAIM2-HIN

For further studies about dsDNA binding and crystallization of HIN domain of AIM2, it was essential to obtain a large amount of soluble the HIN-200 protein. The expression of HIN domain of human AIM2 was tested at the initial stage of the experiments with different types of tags such as GST or hexa-histidine. Both pET vectors and pGEX vectors were not suitable to make the human HIN-200 domain fold and express properly in *E. coli* BL21 (DE3) cells. Then it was decided to use the HIN-200 domain from different source such as the *Mus musculus* AIM2 which is a homolog of human AIM2. Mouse AIM2 is comprised of an N-terminal Pyrin domain and a C-terminal HIN-200 domain flanked by a long undefined linker. Similar to human AIM2, the HIN-200 domain (residues 135-354) was selected from the full-length mouse AIM2 construction and subcloned into the pET22b (+) vector using *Nde*I and *Sal*I restriction sites. The mouse mAIM2-HIN containing vector was introduced to *Escherichia coli*

BL21 (DE3) cells, and the expression of the mAIM2-HIN was successful and confirmed by an initial expression test.

1 L of mAIM2-HIN(135-354) cells was cultured overnight at 15℃ as it was tested in the expression test, and a sufficient amount of mAIM2-HIN was obtained. Cells were harvested and lysed by sonication, then the proteins were purified by His-tag affinity chromatography, followed by size exclusion chromatography. 5 mM DTT was added after both purification steps to prevent disulfide bond formation among mAIM2-HIN proteins. About 25~30mg of protein was obtained from 1 L of cell culture. After purification of mAIM2-HIN, samples were concentrated to > 5 mg/㎖ and a than 500 ㎕ of sample was saved on ice for more than three days, and analyzed to see if they form disulfide bonds or if they become degraded in the solution. The native mAIM2-HIN showed dimer or tetramer oligomerization on size exclusion chromatography after three days. Then, we mutated each of four cystein residues to serine. Site directed mutagenesis was performed for four cysteines: Cys173, Cys264, Cys308 and Cys335. Each mutant was introduced into *Escherichia coli* BL21 (DE3) cells. We failed to get soluble mAIM2-HIN C308S protein, though they were expressed well. All the other mutations were successfully expressed and were soluble. C335S mutation did not affect disulfide formation, and proteins remained as oligomers in the same manner as the native mAIM2-HIN. Mutation of C173S eliminated the oligomerization, however, these proteins were aggregated immediately after the addition of dsDNA. Stable monomeric mAIM2-HIN was achieved with the C264S mutation, and the problem due to disulfide formation in native mAIM2-HIN was solved. mAIM2-HIN C264S was purified

(FIGURE 1) for the following studies: double-stranded DNA binding and crystallization. The concentration of purified protein was measured by absorbance at 280 nm using an extinction coefficient of 8480 $M^{-1}cm^{-1}$.

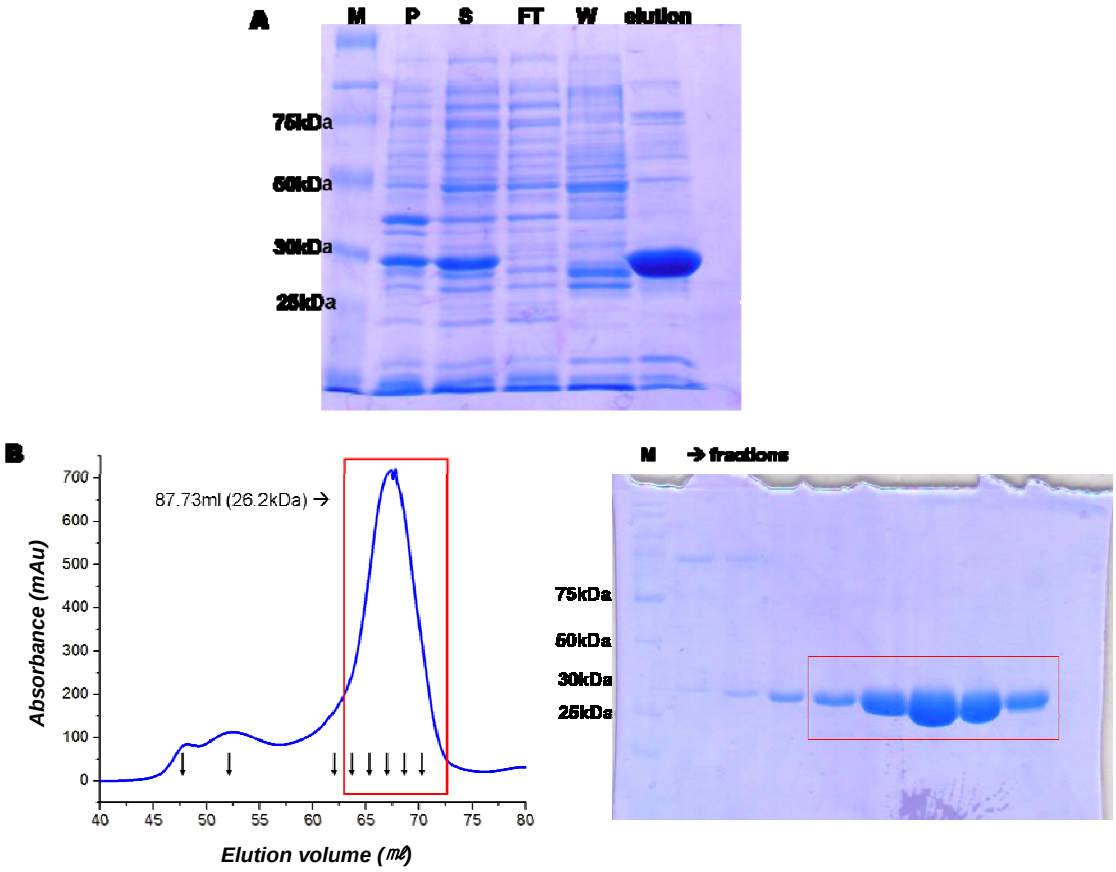

FIGURE 1. Purification of mAIM2-HIN C254S. (A) 18% SDS-PAGE of samples from Ni^{+} beads. M:marker, T:total, S:supernatant, FT:flowthrough, W:washing. (B) UV absorbance at 280nm from gel filtration chromatography and 18% SDS-PAGE gel for samples from fractions indicated in absorbance graph.

Binding of double stranded DNA by mAIM2-HIN

As it was known that foreign double-stranded DNA can be sensed by the HIN200 domain of AIM2 in the cytosol (Choubey et al., 2000; DeYoung et al., 1997), we hypothesized that mAIM2-HIN can bind to dsDNA. Variable lengths of the dsDNA to be used for AIM2-HIN binding were adapted from a 60 bp DNA fragment from the herpes simplex viral DNA which is known to activate AIM2 (Rasmussen et al.). 15bp, 40bp and 60bp sequence forward and reverse oligos were ordered and annealed, and then annealed dsDNA was purified using gel filtration chromatography. The concentration of purified double-stranded DNAs was measured by absorbance at 260nm using extinction coefficient of 240,985 $M^{-1}cm^{-1}$ for HSV15, 621,698.0 $M^{-1}cm^{-1}$ for HSV40 and 982,590.1 $M^{-1}cm^{-1}$ for HSV60 respectively.

To specify the proper molar ratio of mAIM2-HIN per double-stranded DNA, 40bp HSV dsDNA (dsHSV40) was used for electrophoretic mobility shift assay (EMSA). 20 μM of dsHSV40 and 40 μM protein were prepared in buffer A. 1 $\mu\ell$ of dsHSV40 was added with increasing amount of proteins, and buffer A was added to adjust the sample to a total volume of 20 $\mu\ell$. This series of samples were loaded to an 8% non-denaturing gel to see the disappearance of unbound dsHSV40 at a certain molar ratio (FIGURE 2).

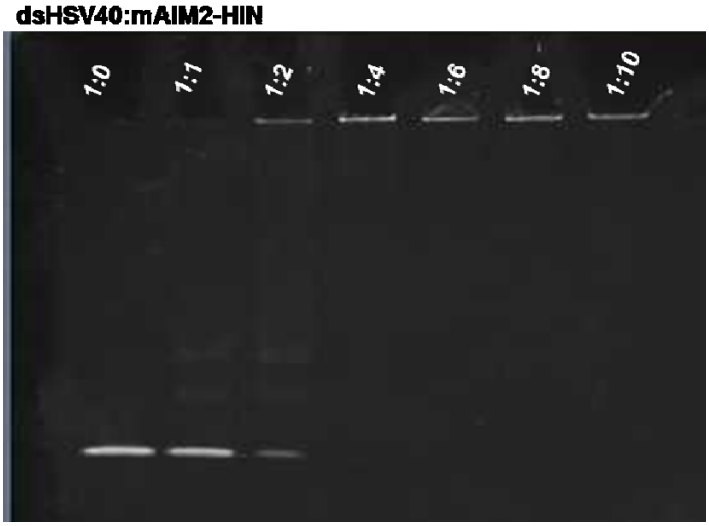

FIGURE 2. Electrophoretic mobility shift assay (EMSA) for dsHSV40 with increasing amount of mAIM2-HIN. Lane 1; 20 pmol dsHSV40 without protein. Lane 2~7; 20 pmol dsHSV40 with proteins in increasing molar ratios indicated.

As expected, 20 pmol of dsHSV40 bound with increasing amount of mAIM2-HIN (lane 3 to 7) were appeared at the top of the gel while free dsHSV40 appeared at the bottom (lane 1). The dsHSV40 band bound to mAIM2-HIN clearly began to appear from 1:2 (dsHSV40:protein), and complete complex appeared from 1:4 (dsHSV40:protein) which indicates that four mAIM2-HIN bind to one dsHSV40. To confirm the stoichiometry of mAIM2-HIN, dsHSV40 was mixed with excess mAIM2-HIN for a small scale gel filtration binding assay using a Superdex200 10/300 GL column (Amersham Bioscience, USA).

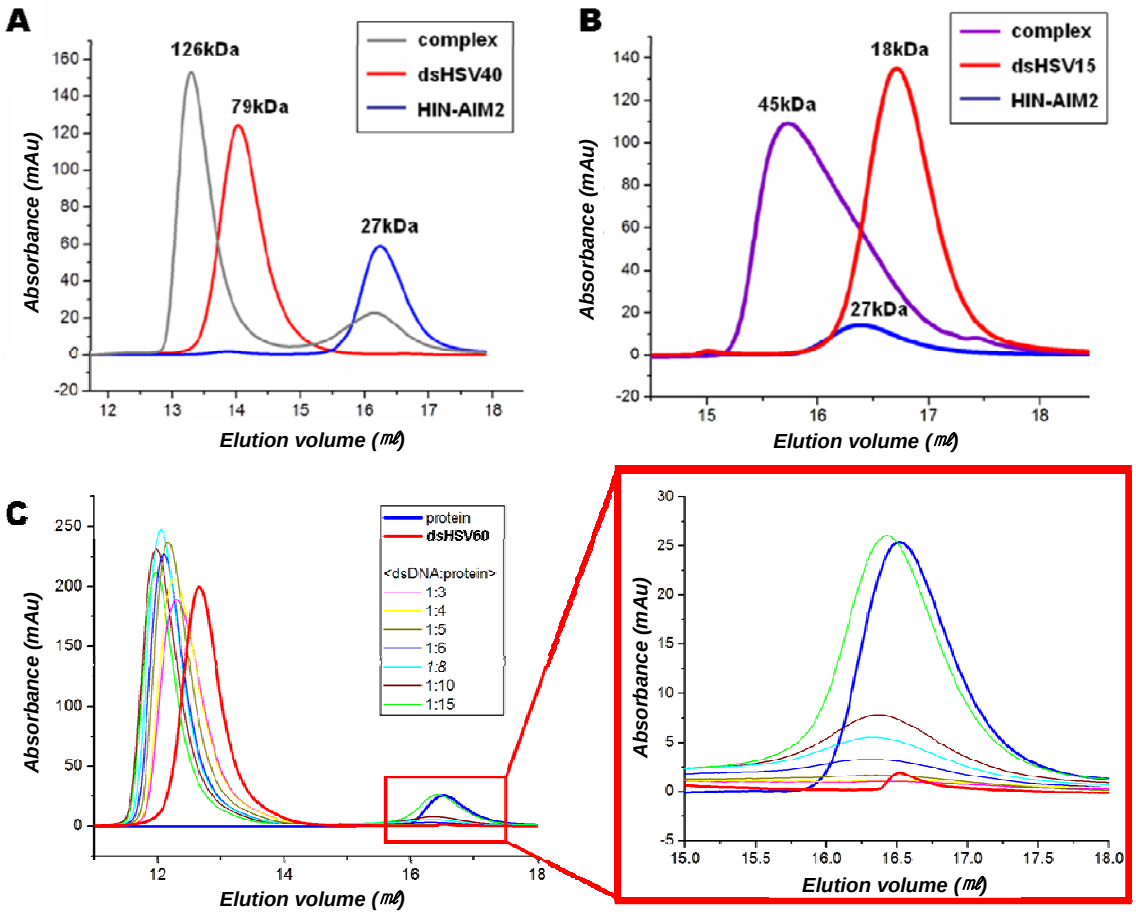

FIGURE 3. Gel filtration chromatography of mAIM2-HIN complexed with dsHSV40, dsHSV15 and dsHSV60. (A) Single peaks of dsHS40 (red) and mAIM2-HIN (blue) appeared at 14.2 mℓ and 16.3 mℓ respectively. The same amount of dsHSV40 and mAIM2-HIN were mixed. A peak for the complex (gray) at 13.3 mℓ and another peak at 16.3 mℓ for excess protein appeared. (B) Peaks for dsHSV15 (red) and mAIM2-HIN (blue) appeared at 16.8 mℓ and 16.3 mℓ respectively. The complex peak (purple) was appeared at 15.2 mℓ with a shoulder peak which corresponds to elution volume of protein. (C) dsHAV60 (red) was used for binding assay with different molar rations indicated in the box. Unbound proteins were appeared from 1:5(dsDNA:protein) ratio in the right panel.

The elution volume of dsHSV40 alone corresponds to the molecular weight of 79 kDa according to the elution volume of gel filtration standards. The calculated molecular weight of dsHSV40 was 24.7 kDa which does not match with the molecular weight estimated by gel filtration chromatography. This is because linear form of dsHSV40 passed through the column and eluted earlier. 10 fold more of proteins were added to dsHSV40 so that excess amount of protein can be used to form complex, and molecular weight of the complex was calculated according to the elution volume of the dsHAV40-HIN-AIM2 complex (FIGURE 3, A). We conclude that four proteins were bound to one dsHSV40 which is consistent with the result elucidated from EMSA (FIGURE 2). Then, for the further binding study, shorter length of dsDNA (dsHSV15) was used for gel filtration binding assay. Three fold more of proteins were mixed with dsHSV15. Calculated molecular weight of the first complex peak was 45 kDa which corresponds to 1:1 (dsHSV15:protein) comlex (FIGURE 3, B). Lastly, 60 bp of HSV sequence was also tested to check the stoichiometry of the mAIM2-HIN against longer dsDNA. Increasing amount of proteins with variable molar ratio of mAIM2-HIN were added to dsHSV60 and analyzed by gel filtration chromatography. Small amount of excess proteins began to appear from the ratio of 1:5(dsHSV60:protein), which means that four mAIM2-HIN bound to one dsHSV60. Through this binding study, we knew that one mAIM2-HIN binds to one dsHSV15, four proteins to one dsHSV40, and five proteins to one dsHSV60. We conclude that appropriate length of dsDNA that bind to mAIM2-HIN, which was necessary to know for crystallization and also for mechanism of dsDNA recognition, is about 15 bp to 20 bp.

Residues involved in recognition of dsDNA

We had a crystal structure of HINa domain of p204 (residues 228-421) which is also known as a cytosolic foreign DNA receptor (Landolfo et al., 1998; Veeranki and Choubey). Using HINa domain of p204, which has 48 % of sequence identity with mAIM2-HIN, as a reference model, we generated a computational structure model of mAIM2-HIN from SWISS-MODEL workplace (FIGURE 4, A) (Arnold et al., 2006). Using this mAIM2-HIN structure model and dsHSV15 structure model generated by make_na server (Lakshiminarayanan and Sasisekharan, 1970), the model of the complex was produced (FIGURE 4, B). To figure out the residues involved in dsDNA binding, we decided to mutate positively charged residues that are close enough to have interaction with sugar-phosphate backbone of dsDNA in the model of the complex.

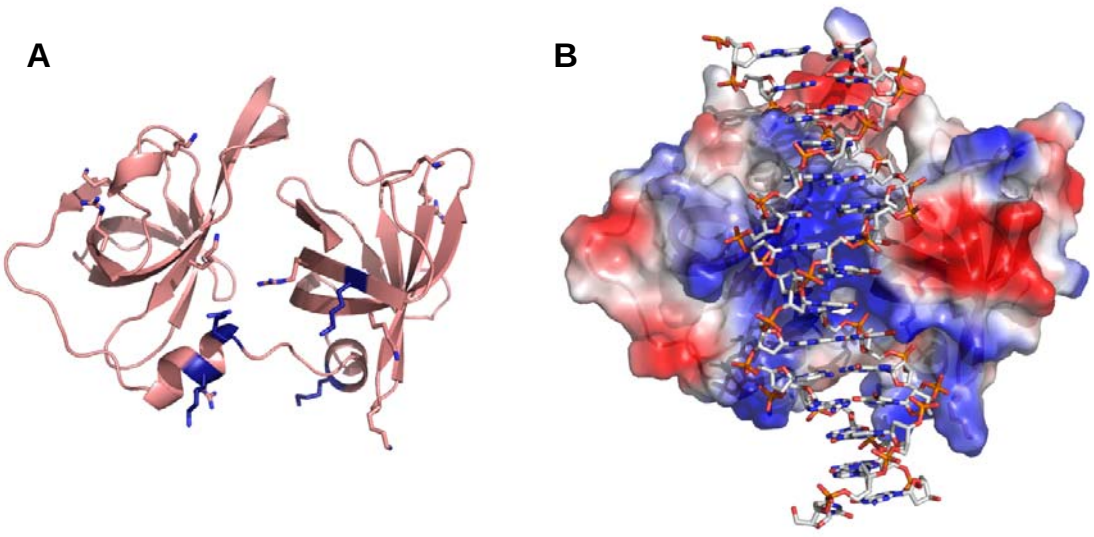

FIGURE 4. Computational structure model of (A) mAIM2-HIN and (B) mAIM2-HIN/dsHSV15 complex.

From the model of the complex, 9 lysine and 4 arginine residues were selected, and one arginine and one lysine distant form the expected dsDNA binding site were chosen as negative control. All the residues were mutated to glutamic acid by site directed mutagenesis. Expression of all the 15 mutants was successful and enough of proteins was obtained to perform the dsDNA binding study. Firstly, a mixture of dsHSV15 and mAIM2-HIN mutants in 1:3 ratio were prepared for Electrophoretic Mobility Shift Assay (EMSA) (FIGURE 5, A)

Interestingly, all the mutants were affected in their binding capacity against dsHSV15 partially (K206 and R249) or completely (K248, R251, K258, K313), except two mutants selected as negative controls (R283E and K305E). Six resides were validated that they disturbed the binding affinity toward dsHSV15 and dsHSV18 by EMSA and gel filtration chromatography, respectively. This result was confirmed by the superimposition of the two structures and sequence alignment of human AIM2 and mAIM2-HIN that these residues, except K258 in mAIM2-HIN, corresponded to the residues of the dsDNA binding site in human AIM2. (Jin et al., 2012).

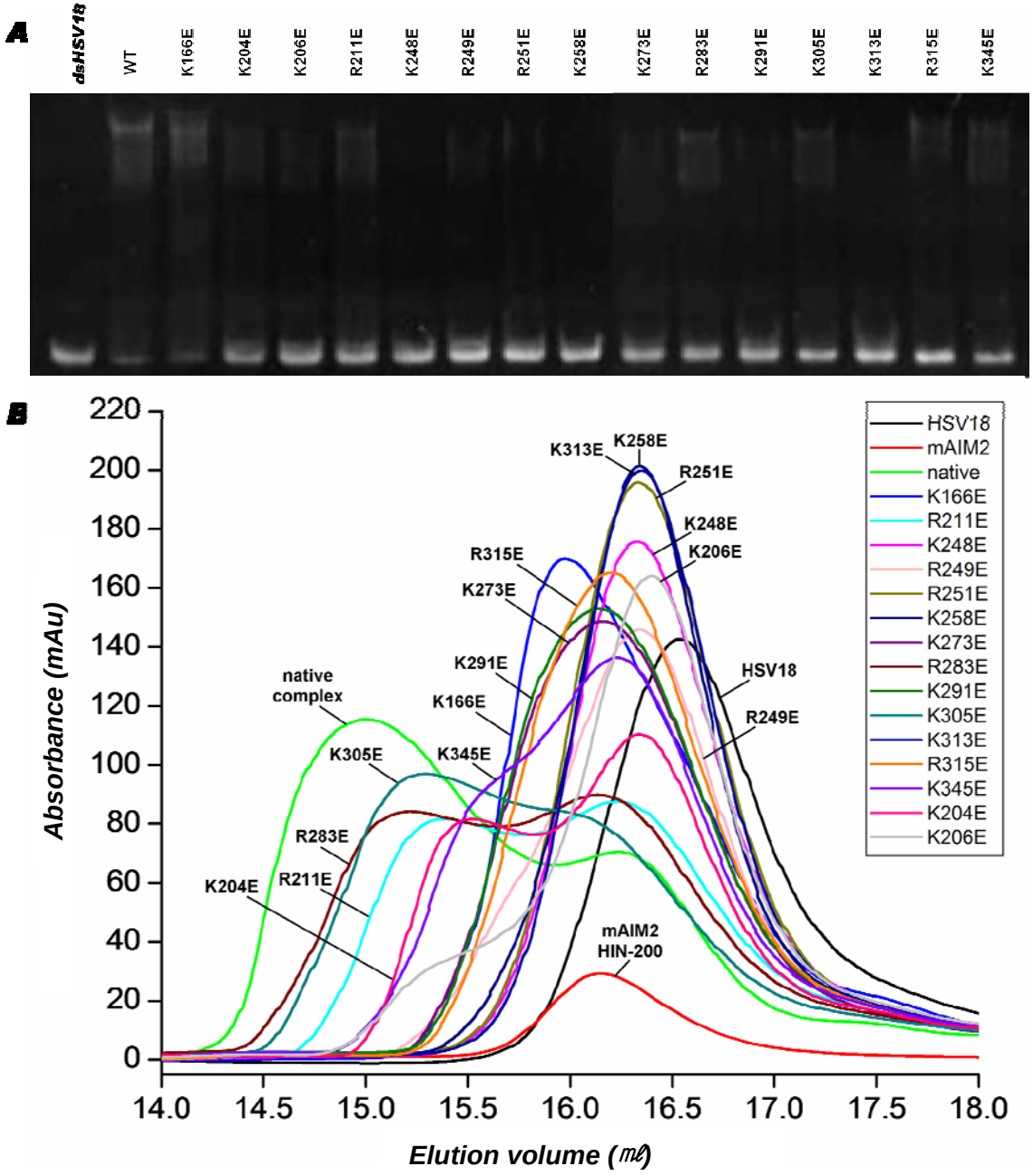

FIGURE 5. dsDNA binding study with mAIM2-HIN mutants by (A) electrophoretic mobility shift assay (EMSA) and (B) gel filtration chromatography. (A) dsHSV15 alone was loaded (lane 1) and purified mAIM2-HIN/dsHSV15 complex was loaded (lane 2). (B) mAIM2-HIN wild type and mutants were mixed with HSV18 and loaded onto HiLoad 16/60 Superdex75 column (Amersham Biosciences)

Crystal screening and optimization

To elucidate how dsDNA interact with mAIM2-HIN, complex of dsHSV15 and mAIM2-HIN (residue 135-354) was purified to screen for complex crystal (FIGURE 6, A). Initial screening was performed with 10 mg/㎖ of purified complex by hanging-drop evaporation method. Needle shaped crystals were obtained from 0.1M Bis-Tris pH5.5, 0.2M sodium chloride, 25% polyethylene glycol 3350. The conditions were optimized and the crystals were reproduced in 0.1M Bis-Tris pH6.5, 0.2M sodium chloride, 29%~30% polyethylene glycol 3350 (FIGURE 6, B).

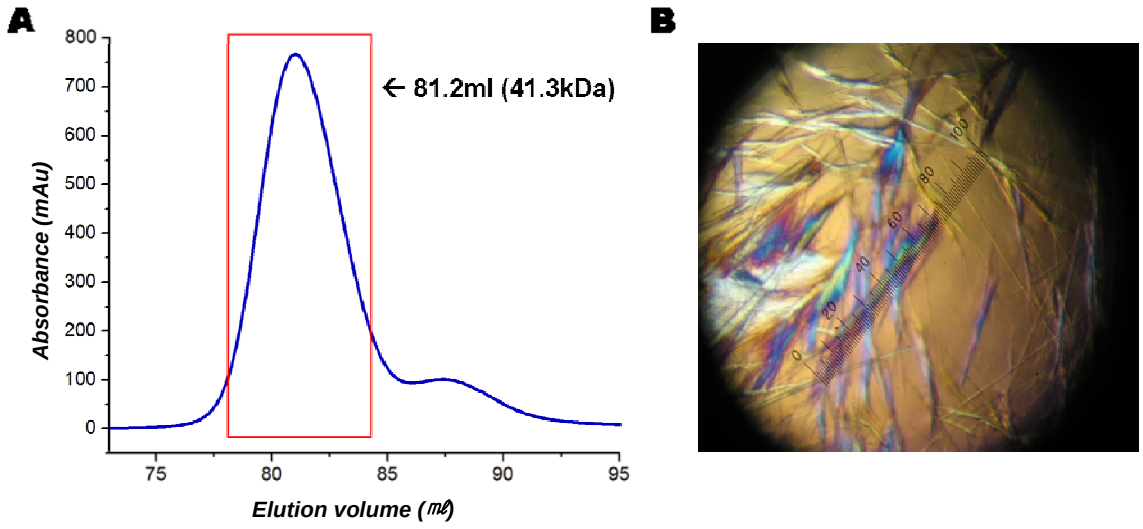

FIGURE 6. Crystallization of mAIM2-HIN/dsHSV15 complex. (A) Complex was purified using HiLoad 16/60 Superdex75 column (Amersham Biosciences) and the fractions from the peak were pooled (red box) for crystal screening. (B) Needle shapes of crystals of mAIM2-HIN/dsHSV15 complex were appeared.

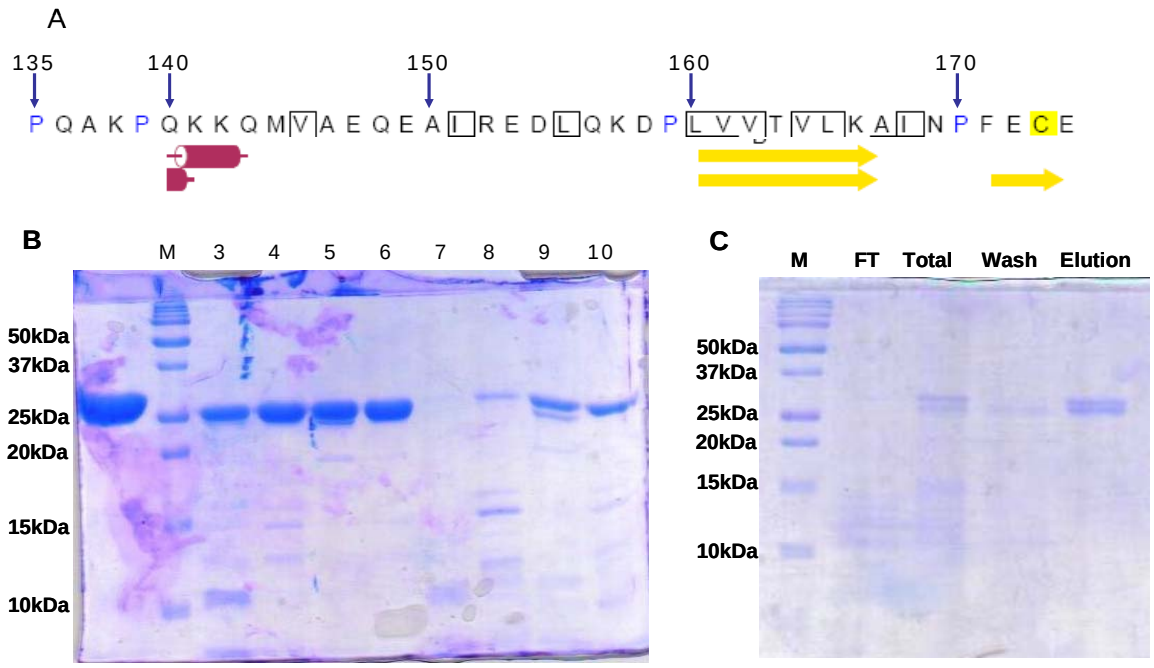

FIGURE 7. Secondary structure prediction of N-terminus of mAIM2-HIN and limited proteolysis assay. (A) Prediction of secondary structure of mAIM2-HIN was shown by JPred3 (Cole et al., 2008). The first 15 residues were expected as loop except four residues of alpha-helix, and beta sheet starts from lysine 160. (B) Purified dsHSV15-HIN-AIM2 complex was loaded on lane 1. lane 2: molecular marker. Lane 3, 4: purified mAIM2-HIN with 2 μg and 0.2 μg of trypsin incubated at room temperature. Lane 5, 6: purified complex with 2 μg and 0.2 μg of trypsin incubated at room temperature. Lane 7, 8: purified mAIM2-HIN with 2 μg and 0.2 μg of Trypsin incubated at 37℃. Lane 9, 10: purified complex with 2 μg and 0.2 μg of Trypsin incubated at 37℃. (C) mAIM2-HIN samples loaded onto lane 9 in (B) were prepared and applied to Ni^{2+} beads. The mixture was washed with washing buffer, then eluted.

The needle crystal was barely diffracted up to ~8Å on home source X-ray. To imporove the quality of crystal and find other alternative conditions, we performed secondary structure prediction with the sequence of HIN domain. The prediction showed the clues that about first 20 residues do not have secondary structure, and maybe easy to be cleaved by trypsin due to its flexibility (FIGURE 7, A) (Cole et al., 2008). To check

the flexible region, limited proteolysis assay for mAIM2-HIN construction used for crystallization was carried out (FIGURE 7, B and C).

Treatment the mAIM2-HIN and mAIM2-HIN/dsHSV15 complex with trypsin showed different degree of degradation after one hour of incubation at room temperature or 37℃ (FIGURE 7, B). This indicates that complex was more stable than protein without dsDNA. Most apo-proteins were degraded after trypsin treatment at 37℃ while mAIM2-HIN/dsHSV15 complex still remains (FIGURE 7, B). To check the site of degradation in mAIM2-HIN, the mixture of protein and trypsin was applied to Ni^{2+} beads. Digested peptides of mAIM2-HIN were observed in the fraction of flow-through. After washing the beads, two different sizes of protein bands around 25 kDa were eluted from the beads. Eluted proteins have C-terminal His-tag, and the smaller protein band is about the size of digested protein which lost a part of expected loop region at the N-terminus. Regarding the activity of trypsin that cleaves arginind or lysine, two sites (Met145, Glu154) for truncation of mAIM2-HIN were selected and subcloned in pET22b vector. C264S that prevents disulfide bond formation was retained in the new truncation forms of mAIM2-HIN. Expression and purification of two types of truncated proteins, △N-HIN-AIM2 (residues 145-354 with C264S) and △△N-HIN-AIM2 (residues 154-354 with C264S) were successful (FIGURE 8, A). Purified proteins from two truncations were mixed with dsHSV15 and dsHSV18, and each complex was purified for further crystal screening. Fortunately, two other promising crystal conditions were found. The crystal of mAIM2-HIN with HSV15 was found in Index screening kit #36 and optimized in 24% Jeffamine ED-2100, 0.1 M HEPES, pH 7.0. The crystal of

AIM2-HIN with HSV18 was found in Index screening kit #56 and optimized in 26% pentaerythritol ethoxylate [15/4 EO/OH], 0.05 M ammonium sulfate, 0.05 M Bis-Tris, pH 6.5. Both complexes were crystallized by hanging-drop vapor-diffusion method over a reservoir of 0.5 ㎖ at 4ºC (FIGURE 8, B and C).

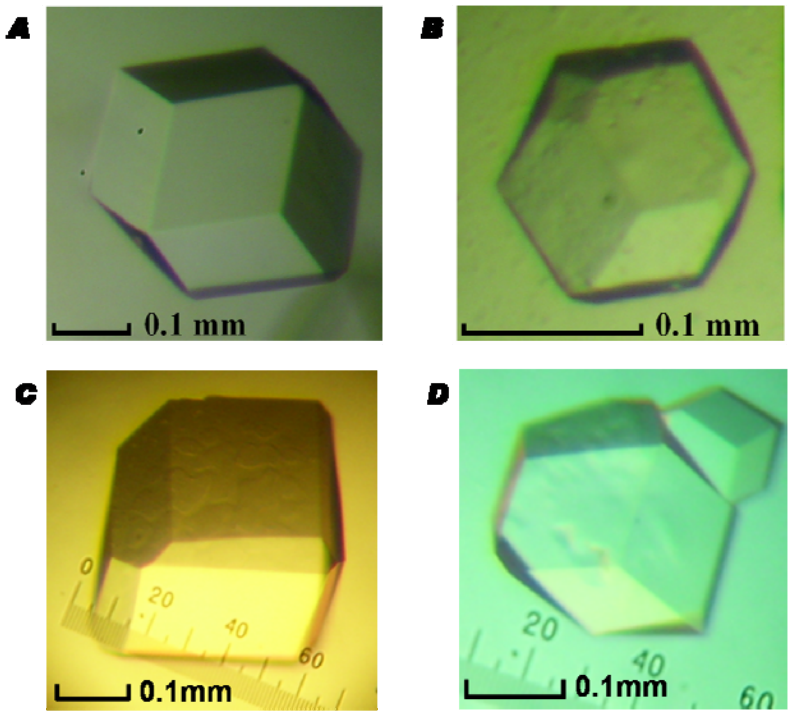

FIGURE 8. Crystals of mAIM2-HIN bound to (A) dsHSV15, (B) dsHSV18, (C) HSV20 and (D) HSV30.

After optimization, crystals were tested for diffraction on home source X-ray. A mAIM2-HIN/HSV15 complex crystal and an mAIM2-HIN/HSV18 complex crystal measuring 0.22 × 0.22 × 0.24 mm and 0.25 × 0.23 × 0.23 mm respectively were grown at 4℃. Crystals of an mAIM2-HIN/HSV15 complex and an mAIM2-HIN/HSV18

complex were diffracted up to 4.0 Å and 4.1 Å respectively. Interestingly, the space groups of the both crystals belong to I23 or $I2_13$, and the size of the unit cells were similar. The exact space group was not distinguishable due to the same reflection conditions. Statistics of diffraction data sets are summarized in Table 1. Matthew's coefficient (Vm) range of 2.5 to 5.1 $Å^3$/Da, which correspond to solvent content of 53 to 77%, reveals that the unit cell was large enough to contain 3-6 complexes.

Crystals of mAIM2-HIN domain bound to 20 bp dsDNA or 30 bp dsDNA with palindromic HSV sequence were also obtained from the same crystal condition, and crystallized between 21-25% pentaerythritol ethoxylate [15/4 EO/OH], 0.05 M ammonium sulfate, 0.05 M Bis-Tris, pH 6.5 (FIGURE 8, C and D). Diffraction data at about 4Å resolution have been collected, and surprisingly, these crystals also belong to the same space group with almost the same unit cell parameters as the 15 and 18 bp DNA complexes (Table 1). This suggest that mAIM2-HIN-200 domain bound to 20bp dsDNA or 30 bp dsDNA with palindromic HSV sequence may form higher order structures similar to mAIM2-HIN/HSV15 or mAIM2-HIN/HSV18.

Table 1. Statistics of data collection of mAIM2/HSV crystals.

	mAIM2 /HSV15	mAIM2 /HSV18	mAIM2 /HSV20	mAIM2 /HSV30
Wavelength (Å)	1.0331	1.0331	1.0331	1.54
Resolution range (Å)	48.10-4.00 (4.07-4.00)	48.20-4.10 (4.17-4.10)	48.20-4.00 (4.14-4.00)	50.00-4.00 (4.14-4.10)
Space group	I23 or $I2_13$	I23 or $I2_13$	I23 or $I2_13$	I23 or $I2_13$
Unit cell parameters (Å)	a = 235.60	a = 236.06	a = 208.61	a = 235.56
Total No. of reflections	183453	171925	75110	91493
No. of unique reflections	18444	17319	12671	17837
Multiplicity	9.9 (10.2)	9.9 (10.3)	5.9 (3.6)	5.1 (2.5)
Completeness (%)	99.8 (100.0)	100.0 (100.0)	97.3(93.2)	96.7(91.5)
Mean I/σ (I)	33.8 (3.0)	27.5 (2.6)	25.5 (2.2)	8.0 (1.9)
R_{merge}[†] (%)	6.2 (60.6)	6.7 (63.0)	8.8 (39.3)	14.0 (36.6)

[†] $R_{merge} = \sum_h \sum_i |I_{i,hkl} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_i |I_{i,hkl}|$, where $I_{hkl,i}$ is the intensity measured for a given reflection with Miller indices h, k, and l, and $\langle I_{hkl} \rangle$ is the mean intensity of that reflection.

CHAPTER III

STRUCTURAL STUDY OF POLYNUCLEOTYDYL TRANSFERASE1 (PNT1) FROM *ARABIDOPSIS THALIANA*

Objective

PNT1 was identified from proteomics study of AGO10 which is involved in the regulation of two miRNAs miR166/165. It is possible that PNT1 is related to regulation of miR166/165, because sequence analysis showed that PNT1 belongs to DnaQ-like 3'-5' exonuclease family which is a member of Ribonuclease H-like superfamily. It is important to study the function of PNT1 based on structural knowledge and correlation with AGO10. To obtain the crystal structure of PNT1 and study its function, following specific aims were set.

Specific aim1: Expression and purification of PNT1

To obtain sufficient amount of PNT1 protein for further study, optimized conditions for cell culture and purifications are necessary. The gene was inserted after SUMO tag and cells were cultured at low rate of shaking and low temperature to incite better protein folding. As purified proteins were very sensitive to temperature that they become aggregated at a low temperature, all purification steps were performed at room temperature.

Specific aim2: Crystal screening, optimization and data collection

To get the crystal of PNT1, purified proteins were concentrated and set up for

crystal screening. The most promising conditions were optimized and diffraction data sets were colleted. Then, to obtain structural knowledge, crystal structure of PNT1 was solved and refined. For higher resolution data sets, crystals would be saved in liquid nitrogen and diffraction data sets would be collected from synchrotron.

Specific aim3: Analysis of structure (hexameric ring and active site)

To figure out the catalytic mechanism of PNT1, the solved structure was compared with known 3´-5´ exonuclease proteins structures. Investigation of presumed active site of PNT1 by the comparison with other 3´-5´ exonuclease homologues, mutagenesis was performed to identify the residues involved in ssRNA cleavage. Activity test of PNT1 wild type and mutants provided clues of catalytic mechanism of PNT1.

Experimental Procedures

Expression and purification of PNT1

Full- length of PNT1 was inserted in pET28a with six-histidine tag followed by SUMO tag at its N-terminus. Cell stock that includes the vector was kindly provided by Dr. Xiuren Zhang, 50 ㎕ of glycerol stock cell was used to inoculate 20 ㎖ LB media with 50 μg/㎖ kanamycine and cultured by shaking at 37℃ until OD_{600}= 1.0. 10 ㎖ of inoculated cells were transferred to 1 L LB medium and cultured for 3 hrs until OD_{600}= 0.7. 300 μM IPTG was added to LB medium when OD_{600}= 0.8 after cooling on ice for

30 min. The cells were incubated at 15℃ with 90 rpm for additional 15 hrs. Then, cells were harvested at 4000 rpm for 10 min at 4℃ and resuspended in 100 ㎖ buffer B (300mM NaCl, 20mM Tris,pH 7.5) stored at room temperature. Sonication was performed on ice. Cell lysate was centrifuged at 8,000 rpm for 10 min. at 4℃, then centrifuged again at 16,000 rpm 10 min. at 4℃ to remove precipitants.

Supernatant was loaded to Ni^{2+} Sepharose High Performance beads (GE Health care, USA) pre-equilibrated with buffer B. N-terminal hexa-histidine tagged PNT1 was bound to the column and non-specifically bound contaminating proteins were removed by 300 ㎖ of washing buffer (500 mM NaCl, 20 mM Tris, pH 7.5, 10 mM imidazole). The remaining bound proteins were eluted with elution buffer (150 mM NaCl, 20 mM Tris, pH 7.5, 250 mM imidazole).

The amount of proteins in the fractions were checked by Bradford assay (Biorad) and then pooled and added with 5 mM DTT. To cleave the SUMO tag, 100 ㎕ of 3 mg/㎖ SUMO protease was added to PNT1, which was pooled and incubated for 2 hrs at room temperature. The cleavage was checked by loading samples on 18 % SDS-PAGE. Cleaved PNT1 was concentrated to 2 ㎖ by an Amicon Ultra-4 10 kDa centrifugal filter (Milipore) at 4000 rpm at 25℃. Then the sample was injected to gel filtration HiLoad 16/60 Superdex75 column (Amersham Biosciences) equilibrated with buffer A (150mM NaCl, 20mM Tris,pH 7.5) and resolved at 1 ㎖/min. PNT1 mutants (D52A, Y54S, D114A and E187A) were also purified with the same purification scheme.

Crystal screening and optimization

Purified proteins were mixed with 5 mM DTT and 5 mM $MgCl_2$, then concentrated to 20 mg/㎖ by centrifugation at 4,000 rpm at 25 ℃ using an Amicon Ultra-10 kDa centrifugal filter (Milipore). The concentration was checked by measuring absorbance at 595 nm using Bradford solution (Biorad). Protein sample was centrifuged at 15,000 rpm for 5 min at room temperature to remove aggregate before setting up crystallization. 192 conditions were screened at 16 ℃ by hanging drop evaporation method using Crystal Screen I & II kit and Index kit (Hampton Research, USA). Each plate well was filled with 0.5 ㎖ of screening solution, and 2 ㎕ of protein sample was mixed with 2 ㎕ of reservoir buffer on the cover glass. One of the best conditions was obtained from Index screening kit #56 which includes 0.2 M potassium chloride, 0.05 M HEPES pH7.5, 30 % Pentaerythritol proxylate (5/4 PO/OH); optimization of the condition was conducted using gradient concentration of precipitant. Purified PNT1 D52A was mixed with 5 mM DTT and concentrated to 15mg/㎖. Crystal of PNT1 D52A was optimized from 0.2 M potassium chloride, 0.05 M HEPES pH7.5, 34 % Pentaerythritol proxylate (5/4 PO/OH).

Results

Purification and crystallization

Cells containing SUMO fusion PNT1 gene were cultured at 15℃ at 100 rpm overnight to help proper protein folding. Expression of SUMO-PNT1 was successful

with six-histidine tag at the N-terminus. SUMO-PNT1 was purified with Ni^{2+} affinity column pre-equilibrated with cold buffer B. We observed that SUMO-PNT1 form aggregation immediately after the elution from the Ni^{2+} column when we used pre-chilled buffers. Interestingly, aggregations were resolved in the buffer, when the samples were exposed to room temperature for 10 min. Using the buffer pre-warmed at room temperature, the aggregation of SUMO-PNT1 was prevented and soluble PNT1 was obtained (FIGURE 9, A). PNT1 was obtained by cleavage of the SUMO tag by incubation of protein with SUMO protease for 1 hr at room temperature. The cleavage was confirmed by loading the sample on the 15% SDS PAGE (FIGURE 9, B).

Purified PNT1 was combined with 5 mM DTT to prevent disulfide bond formation and 5 mM $MgCl_2$ was added because it was expected to be necessary in the presumed active site for possible exonuclease activity. For crystal screening, purified PNT1 was concentrated to 20 mg/㎖, and as we knew that PNT1 becomes aggregated easily at low temperatures, we set up plates at 16℃ using Crystal Screen I & II kit and Index kit (Hampton Research, USA). Crystals appeared within 3 to 5 hrs from the initial screening. Most of the crystals were very small, and were thin plates or pyramid shapes. On the next day of initial screening, we selected five of most promising conditions: 0.4 M potassium sodium tartrate tetrahydrate (Crystal screen I, #2), 0.2 M ammonium acetate, 0.1 M sodium citrate tribasic dihydrate pH 5.6, 30% 2-methyl-2,4-pentandiol (Crystal screen I, #2), 0.1 M HEPES pH 7.5, 10% Polyethylene glycol 8,000, 8% Ethylene glycol (Crystal screen II #37) and 0.2M potassium chloride, 0.05M HEPES pH7.5, 35% Pentaerythritol proxylate (5/4 PO/OH) (Index #56).

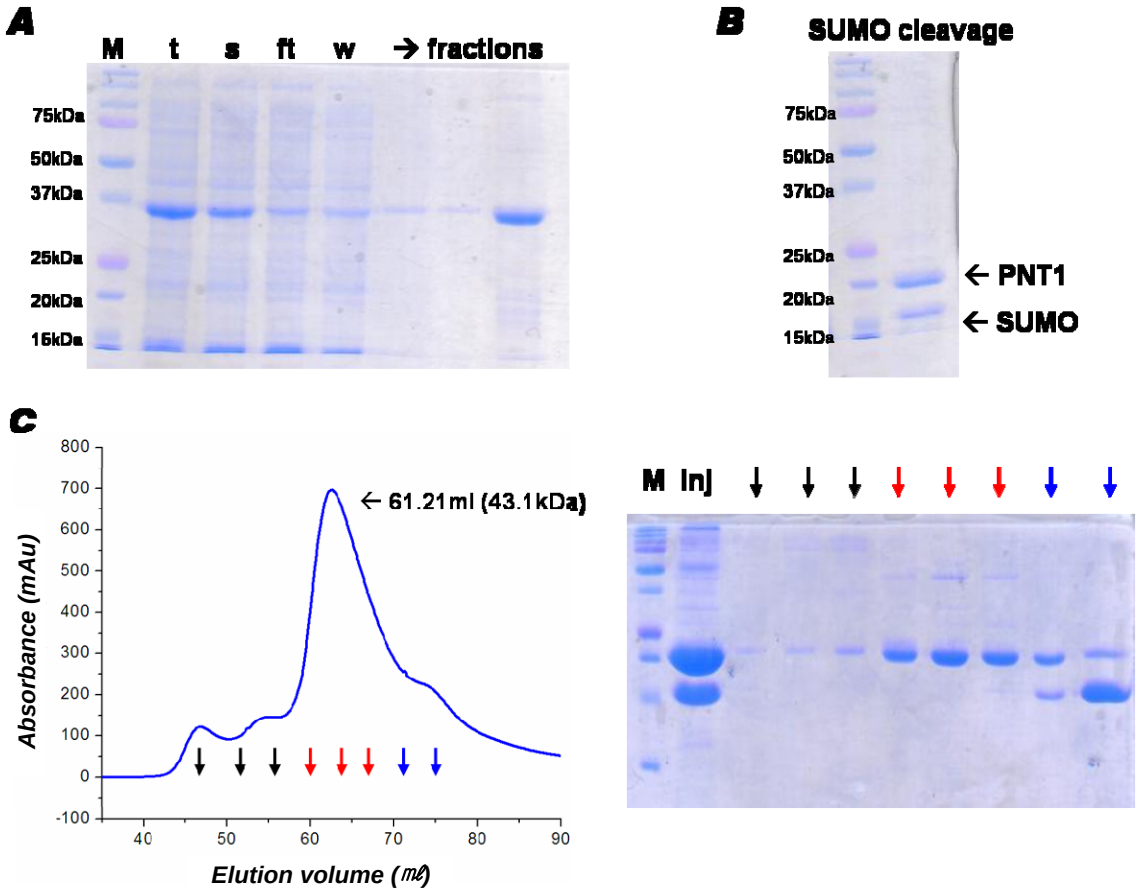

FIGURE 9. Purification of SUMO-PNT1 and SUMO tag cleavage. (A) Purification of SUMO-PNT1 from Ni^{2+} column. M; marker, t; total, s; supernatant, ft; flow-through, w; wash. (B) Cleavage of SUMO tag was confirmed by 15% SDS-PAGE. (C) Purification of PNT1 from HiLoad 16/60 Superdex75 column.

We optimized these crystallization conditions by varying concentration of precipitant, temperature and protein concentration to produce large crystals. Fortunately, relatively large crystals were observed among many of micro-crystals in the drop which contain 30% of precipitant (FIGURE 10).

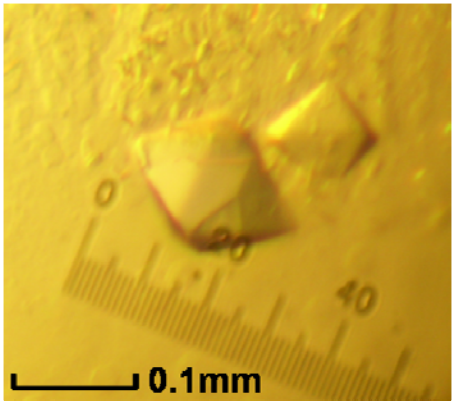

FIGURE 10. Crystal of PNT1 from optimized condition (0.2M potassium chloride, 0.05M HEPES pH7.5, 30% Pentaerythritol proxylate (5/4 PO/OH)). 1 unit equals 6 μm.

The structure of hexameric PNT1

The crystal was soaked in mother liquor containing 35% Pentaerythritol proxylate (5/4 PO/OH)) for cryo-procection. Diffraction data was collected from a single crystal at 120K using home source X-ray and the raw data set was processed and scaled using HKL2000 package (Vaguine et al., 1999). Statistics of diffraction data is summarized in Table 2.

Molecular replacement was conducted using Molrep in CCP4 package (Vagin and Teplyakov, 2010) with uncharacterized protein At5g06450 from *Arabidopsis thaliana* (PDB: 1VK0)(Perry et al., 2006), which has 69 % sequence identity with PNT1, as a search model. Three molecules were found in asymmetric unit with 83.1% of solvent content which corresponds to Matthew's coefficient (Vm) 7.29. From the initial structure of PNT1, we observed that PNT1 formed a hexameric ring in the crystal which explains why the crystal has high solvent content. Model building according to the amino acid sequence of PNT1 and refinement was performed using Coot (Emsley and

Cowtan, 2004), and Refmac (Murshudov et al., 1997). Statistics of refinement are summarized in Table 2.

Table 2. Statistics of diffraction data and refinement of PNT1 wild type and D52A.

Data collection		
	PNT1 WT	PNT1 D52A
Wavelength (Å)	1.54	1.54
Resolution range (Å)	50.00-2.70 (2.80-2.70)	50.00-2.80 (2.90-2.80)
Space group	$I4_122$	$I4_122$
Unit cell parameters (Å)	a=b=193.26, c=214.40	a=b=192.75, c=215.70
Total No. of reflections	299700	267872
No. of unique reflections	55410	49731
Multiplicity	5.4 (4.7)	5.4 (4.4)
Completeness (%)	99.5 (99.4)	99.3 (99.2)
Mean I/δ (I)	10.2 (2.1)	8.64 (2.04)
R_{merge}[†] (%)	21.0 (61.1)	21.1 (59.7)
Refinement		
R_{work}/R_{free}[‡] (%)	20.3/23.0	20.8/24.4
Bond lengths (Å)	0.027	0.024
Bond angles (°)	2.279	2.093
Ramanchandran plot (%)		
Most favored	89.4	88.3
Additionally allowed	10.6	11.4
Generally allowed	0.0	0.4
Disallowed	0.0	0.0

[†] $R_{merge} = \sum_h\sum_i |I_{i,hkl} - \langle I_{hkl}\rangle|/\sum_{hkl}\sum_i |I_{i,hkl}|$, where $I_{hkl,i}$ is the intensity measured for a given reflection with Miller indices h, k, and l, and $\langle I_{hkl}\rangle$ is the mean intensity of that reflection.

[‡] $R_{work} = \sum||F_o|-|F_c||/\sum|F_o|$, where F_o and F_c are the observed and calculated structure-factor amplitudes, respectively. R_{free} was calculated as R_{work} using a randomly selected subset (10%) of unique reflections not used for structure refinement.

Monomeric PNT1 is comprised of 7 of beta sheets surrounded by 8 alpha helixes (FIGURE 11, A) and the structure was superimposed onto known structures of At5g06450 from *Arabidopsis thaliana* (PDB: 1VK0) and 3'-5' exonuclease domain of Werner syndrome ATP-dependent helicase, WRN (PDB: 2FC0) (Perry et al., 2006) with r.m.s. deviation of 0.8 Å and 2.7 Å respectively. Both structures have structural conservation that is contained in DnaQ family (SCOP database) and with DEDD motif in their canonical active sites. By comparison, PNT1 also has three acidic residues that correspond to DEDD motif (FIGURE 11, A).

Hexameric ring of PNT1 forms through hydrogen bonds and majority of electrostatic interactions between interfaces (FIGURE 11, B). The surface area of each side of interfaces is about 2,840 $Å^2$ while overall surface area of one chain is 10,470.8 $Å^2$. 54.8% of surface area of monomeric PNT1 is buried, and further analysis showed that one side of the PNT1 in the interface includes 25.7% of acidic residues (four aspartates and five glutamates) (FIGURE 11, A, upper panel), while the other side includes 23.5% of basic residues (three arginines and eight lysines) (FIGURE 11, B, down panel).

In the structure of PNT1, Asp52 and Asp114 which are expected to coordinate divalent cations have electrostatic interaction with nitrogen atoms in Lys102 and Arg127 of the other chain. Asp114 has direct electrostatic interaction with two positively charged residues from the other chain reinforcing the interactions (FIGURE 11, B).

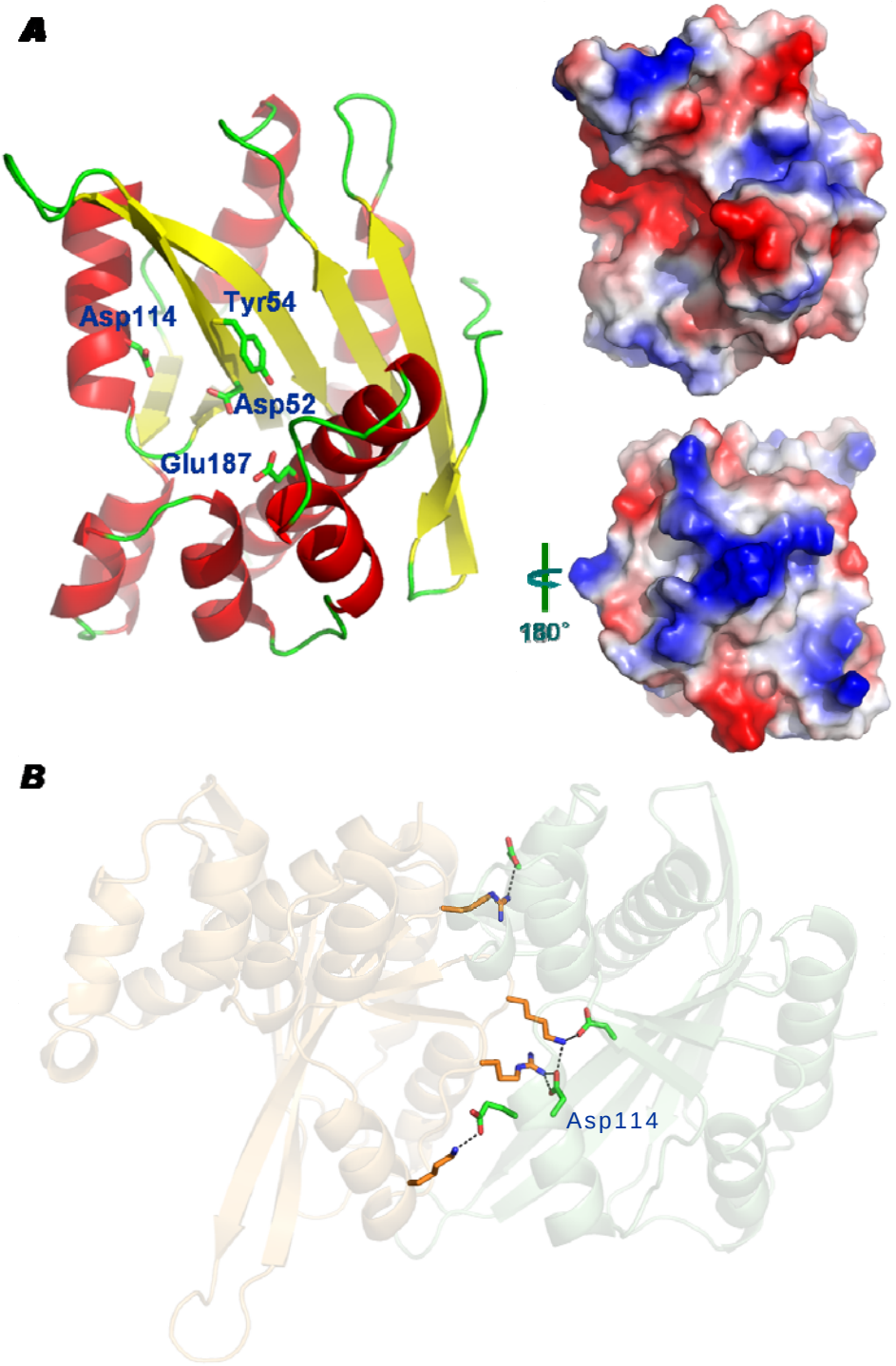

FIGURE 11. Crystal structure of PNT1. (A) Structures of monomeric PNT1. Residues that correspond to DEDD motif are shown as stick. Surface charge of both interfaces is shown on the right panel. (B) Electrostatic interactions between charged residues in the interface.

Analysis of the hexameric PNT1

For further analysis of presumed active site of PNT1, DALI server was used for searching other proteins that share high degree of structural conservation with PNT1. the result included homologue of human 3´-5´-exoribonuclease (3'hExo), ERI-1 protein (PDB: 1W0H) (Wu et al., 2005) and Poly(A)-specific ribonuclease (PARN) (PDB: 2A1R) (Cheng and Patel, 2004). Both ERI-1 and PARN have 3'-5' exonuclease activity and their catalytic mechanisms were studied. Crystal structures of both proteins were complexed with nucleotide ligands, and superimposed onto PNT1 using DaliLlite server (Holm and Park, 2000) (FIGURE 12, A). Monomeric structure of PNT1 shares r.m.s. deviation of 3.6 Å and 3.4 Å with ERI-1 and PARN respectively, although, sequences of both proteins are divergent from PNT1 sharing only 10% and 11% of sequence identity respectively. ERI-1 and PARN belong to DEDD family and include four acidic residues that interact with divalent cation such as Mg^{2+} or Mn^{2+}. Additional histidine or tyrosine residues have roles in catalysis in this protein clan (Zuo and Deutscher, 2001). Superimposition of ERI-1, PARN, WRN and PNT1 structures reveal that PNT1 also has three acidic residues and tyrosine that correspond to DEDD motif (FIGURE 12, B). Electron density maps for Mg^{2+} added to PNT1 during crystallization were not observed around presumed active site, and Asp52 and Asp114 interact with Lys102 and Arg127 of the other chain instead of divalent cation. Glu of DEDD motif was replaced by Tyr54 and there was no additional catalytic His or Tyr in the presumed active site of PNT1 (FIGURE 12, B).

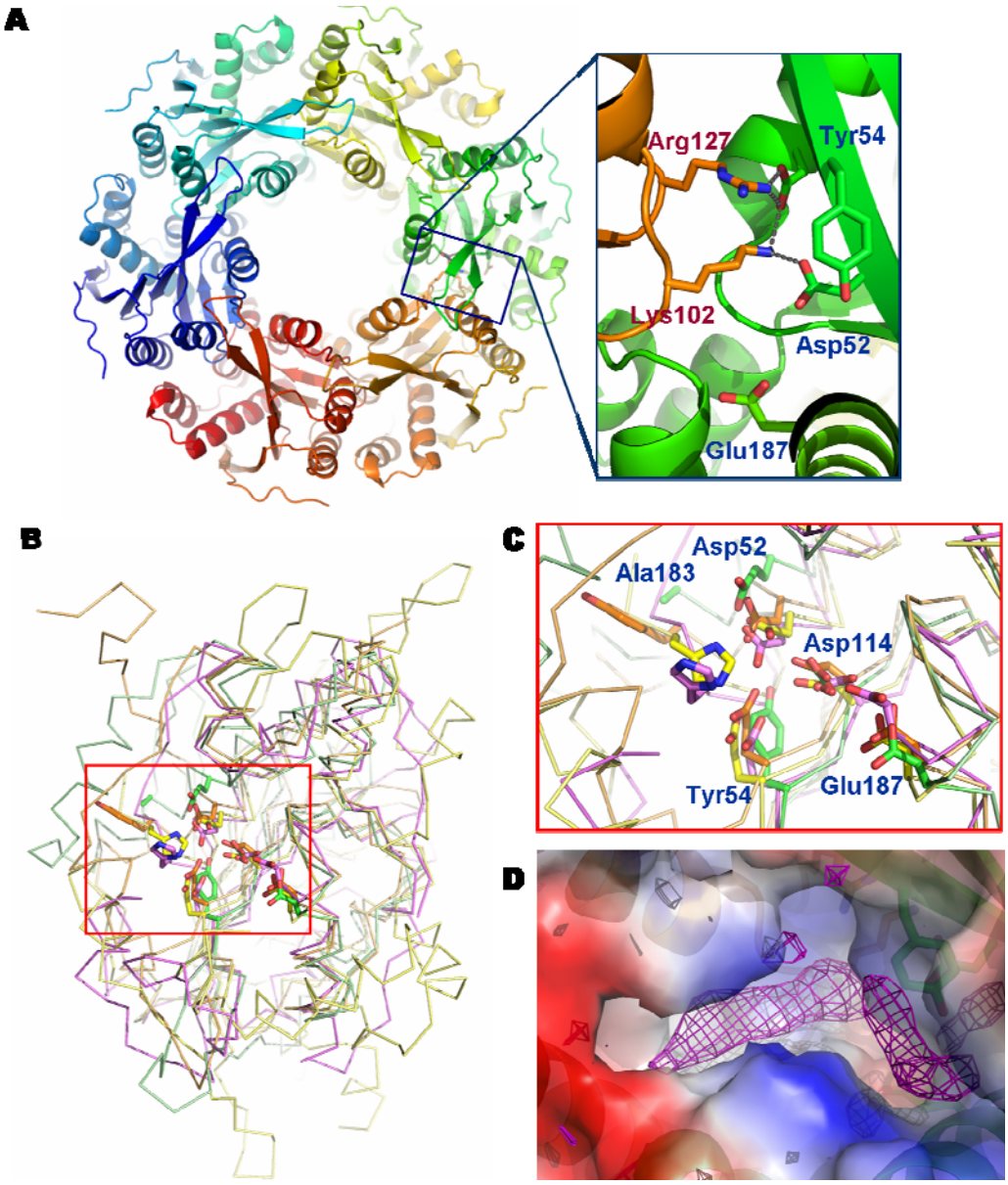

FIGURE 12. (A) Hexameric ring of PNT1 and superimposition of PNT1 with other 3´-5´ exonucleases. Presumed active site is in the interface between chains. Asp52 and D114 interact with Arg127 and Lys102 from the other chain. (B) Superimposition At5g06450 from Arabidopsis thaliana (PDB: 1VK0), 3'-5' exonuclease domain of Werner syndrome ATP-dependent helicase (PDB: 2FC0), homologue of human 3´-5´-exoribonuclease (3'hExo), ERI-1 protein (PDB: 1W0H) and Poly(A)-specific ribonuclease (PARN) (PDB: 2A1R) onto PNT1.DEDD motif of 3'-5' exonuclease domains. (C) Residues in the presumed active site of PNT1 are aligned with other proteins in (B). (D) Extra density map in the active site of PNT1. AMP or GMP was bound in the active site of other proteins. Fo-Fc with 3σ level.

Interestingly, extra electron density was observed along the expected active site (FIGURE 12, D). Due to low resolution and quality of data we have collected from R-axis, density map was not clear enough to be defined with any molecules yet. It was reported that the active sites of ERI-1 and PARN were bound with a rAMP and trinucleotides respectively as a product of 3′-5′ exonuclease activity. Furthermore, it was confirmed that PNT1 has strong 3′-5′ exonuclease activity against single stranded RNA by collaborators (Dr. Xiuren Zhang). It is possible that PNT1 has a nucleotide in its active site as a product of 3′-5′ exonuclease activity, and the extra electron density represents that nucleotide.

To investigate the mechanism of exonuclease activity of PNT1, we performed mutagenesis for four residues in the presumed active site: Asp52 to Ala, Tyr54 to Ser, Asp114 to Ala and Glu187 to Ala. All the mutants were expressed well and soluble after cell lysis in buffer B. Using the same purification scheme as it was used for the PNT1 wild type, four mutants were purified. Interestingly, D114A was not stable as dimer, and appeared in at least three oligomeric states from the gel filtration chromatography (FIGURE 13, A). Exonuclease activity was tested with the proteins from fractions of each peak. Oligomeric PNT1 showed activity when they form oligomer, while monomeric proteins does not show activity against 24 nucleotide single stranded RNA (24n.t. ssRNA). The result confirms our hypothesis that PNT1 need to form dimer or oligomer to have catalytic activity.

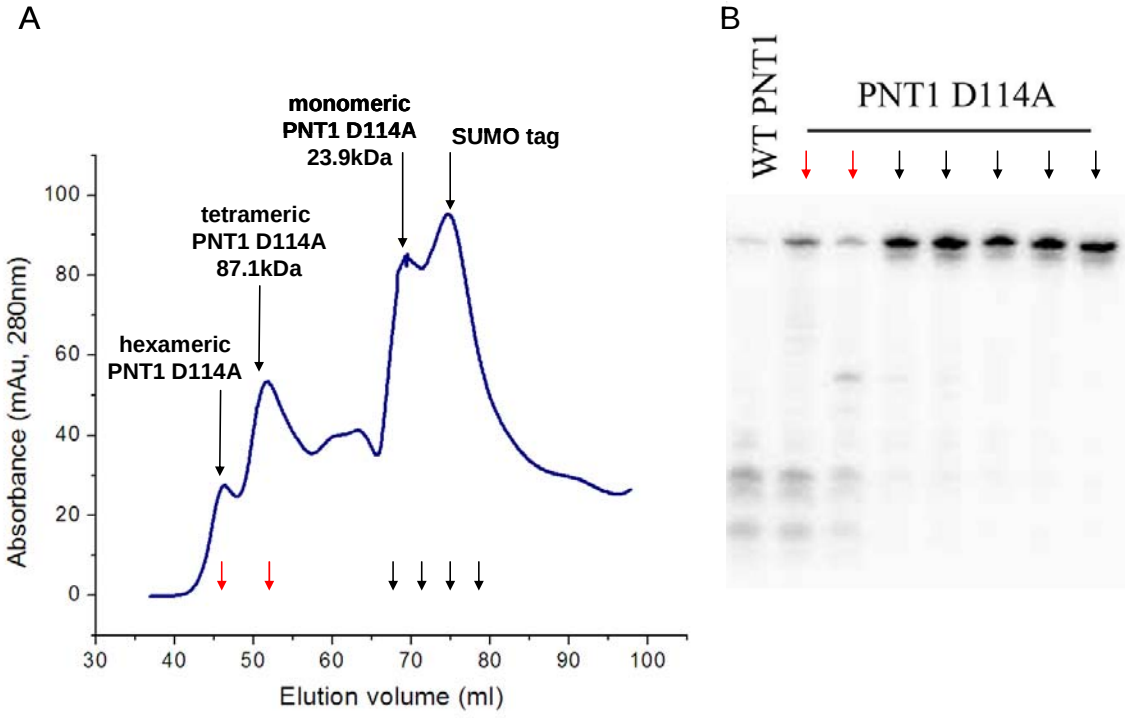

FIGURE 13. Purification and activity test of PNT1 D114A. PNT1 D114A was purified and (A) appeared in more than three states. (B) monomeric PNT1 D114A does not show exonuclease activity while oligomers have activity against 24n.t. ssRNA.

Mutants of PNT1 D52A, Y54S and D187A were also purified, and formed dimer as shown by gel filtration chromatography. Activity test with all mutants were conducted by our collaborator, and only PNT1 D52A does lacks exonuclease activity against ssRNA. PNT1 D52A was crystallized from the same crystallization condition as the PNT1 wild type, and the structure of mutant was solved (Table 2). Overall structure of PNT1 D52A was almost same with structure of PNT1 wild type, except that one water molecule was interacting with Arg127 and Lys102 instead of oxygen atoms of Asp52 in the wild type of PNT1 (FIGURE 14, B).

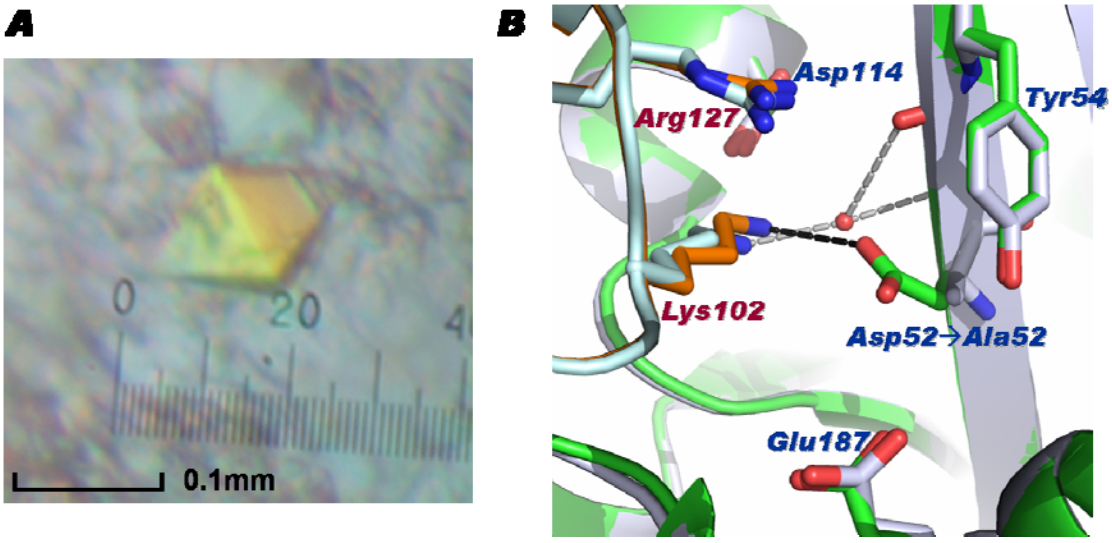

FIGURE 14. Crystallization of PNT1 D52A and site of mutation. (A) PNT1 D52A mutant was crystallized from the same condition as PNT1 wild type. (B) Active site of PNT1 D52A was superimposed with that of wild type. One water molecule takes place of the oxygen atom of Asp52 in wild type.

To figure out the mechanism of 3′-5′ exonuclease activity, 5 base of ssRNA sequence adapted from Ago10 associated miR166 was ordered to make ssRNA complexed PNT1 crystal. Firstly, ssRNA was added to purified PNT1 D52A proteins to try co-crystallization, but the proteins aggregated immediately. Alternatively, ssRNA was resuspended as 1mM or 5mM of final concentration in cryo buffer containing 35 % Pentaerythritol proxylate (5/4 PO/OH) to try crystal soaking. Unfortunately, Crystals were not diffracted after soaking in ssRNA resolved cryo solution even in 5 seconds of soaking in 1mM ssRNA cry solution.

CHAPTER IV

CONCLUSION*

For the AIM2, we failed to obtain the solution from all of data sets using both MOLREP and PHASER (Winn et al., 2011) with the crystal structures of DNA-bound HIN-200 domains of IFI16 (PDB 3RNU) and human AIM2 (3RN2), as well the DNA-free IFI16 HIN-200 domains (IFI16 (PDB 2OQ0, 3B6Y) as search models (Jin et al., 2012; Liao et al., 2011). Data sets were analyzed before molecular replacement using Xtriage in the PHENIX package (Adams et al., 2010), and indicates that both crystals are twinned, with twin fractions of 0.31 for the HSV15 complex, 0.29 for the HSV18 complex, 0.41 for the HSV20 complex and 0.40 for HSV30 complex. Currently, we have prepared the crystals of Se-Met derivative of complexes to determine the structure by single wavelength or multi-wavelength anomalous diffraction methods.

Crystal structure of PNT1 showed hexameric ring formation like its homologous, uncharacterized protein At5g06450 from *Arabidopsis thaliana* (PDB: 1VK0) (Perry et al., 2006) while most of other 3′-5′ exonucleases found by DALI server were monomer or dimer. Through the investigation of presumed active site of PNT1, four residues, which were aligned well with DEDD residues of 3'-5' exonuclease family, were identified. Interestingly, three acidic residues (Asp52, Asp114 and Glu187) were not binding divalent cations, but positively charged residues from the other chain were interacting with Asp52 and Asp114. Dimerization through these electrostatic interactions

* Reprinted with permission from "Crystallographic characterization of mouse AIM2 HIN-200 domain bound to a 15 bp and an 18 bp double-stranded DNA." By M. W. Sung, T. Watts and P. Li. 2012. Acta Cryst. Volume 68, Part 9, Pages 1081-1084. Copyright 2012 by IUCr Journals.

was necessary for PNT1 to have activity. Furthermore, in spite of the absence of additional catalytic residue such as Histidine or Tyrosine in other proteins of DEDD family, PNT1 showed strong 3′-5′ exonuclease activity. Extra electron densities in the active site suggest PNT1 probably has a nucleotide in the active site as a product of exonuclease activity. By mutagenesis, it was confirmed that Asp52 has a role in degrading nucleotides. It was failed to see the structure of ssRNA bound PNT1 and to figure out its mechanism. As far as PNT1 is not same as other proteins in DEDD superfamily, PNT1 could be a novel nuclease with different mechanism to degrade ssRNA.

BIBLOIGRAPHY

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica Section D, Biological crystallography 66*, 213-221.

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics 22*, 195-201.

Barber, G.N. (2011). Innate immune DNA sensing pathways: STING, AIMII and the regulation of interferon production and inflammatory responses. *Current opinion in immunology 23*, 10-20.

Beese, L.S., and Steitz, T.A. (1991). Structural basis for the 3'-5' exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism. *European molecular biology organization journal 10*, 25-33.

Cheng, Y., and Patel, D.J. (2004). Crystallographic structure of the nuclease domain of 3'hExo, a DEDDh family member, bound to rAMP. *Journal of molecular biology 343*, 305-312.

Choubey, D., Walter, S., Geng, Y., and Xin, H. (2000). Cytoplasmic localization of the interferon-inducible protein that is encoded by the AIM2 (absent in melanoma) gene from the 200-gene family. *Fondation esther boucicault-stanislas letter 474*, 38-42.

Cole, C., Barber, J.D., and Barton, G.J. (2008). The Jpred 3 secondary structure prediction server. *Nucleic acids reseach 36*, W197-201.

DeYoung, K.L., Ray, M.E., Su, Y.A., Anzick, S.L., Johnstone, R.W., Trapani, J.A., Meltzer, P.S., and Trent, J.M. (1997). Cloning a novel member of the human interferon-inducible gene family associated with control of tumorigenicity in a model of human melanoma. *Oncogene 15*, 453-457.

Dinarello, C.A. (2009). Immunological and inflammatory functions of the interleukin-1 family. A*nnual review of immunology 27*, 519-550.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography 60*, 2126-2132.

Fernandes-Alnemri, T., Yu, J.W., Juliana, C., Solorzano, L., Kang, S., Wu, J., Datta, P., McCormick, M., Huang, L., McDermott, E.*, et al.* (2010). The AIM2 inflammasome is critical for innate immunity to Francisella tularensis. *Nature immunology 11*, 385-393.

Holm, L., and Park, J. (2000). DaliLite workbench for protein structure comparison. *Bioinformatics 16*, 566-567.

Hornung, V., Ablasser, A., Charrel-Dennis, M., Bauernfeind, F., Horvath, G., Caffrey, D.R., Latz, E., and Fitzgerald, K.A. (2009). AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature 458*, 514-518.

Hornung, V., and Latz, E. (2010). Intracellular DNA recognition. *Nature reviews Immunology 10*, 123-130.

Jin, T., Perry, A., Jiang, J., Smith, P., Curry, J.A., Unterholzner, L., Jiang, Z., Horvath, G., Rathinam, V.A., Johnstone, R.W*., et al.* (2012). Structures of the HIN domain:DNA complexes reveal ligand binding and activation mechanisms of the AIM2 inflammasome and IFI16 receptor. *Immunity 36*, 561-571.

Jung, J.H., and Park, C.M. (2007). MIR166/165 genes exhibit dynamic expression patterns in regulating shoot apical meristem and floral development in Arabidopsis. *Planta 225*, 1327-1338.

Lakshiminarayanan, A.V., and Sasisekharan, V. (1970). Stereochemistry of nucleic acids and polynucleotides. II. Allowed conformations of the monomer unit for different ribose puckerings. *Biochimica et biophysica acta 204*, 49-59.

Landolfo, S., Gariglio, M., Gribaudo, G., and Lembo, D. (1998). The Ifi 200 genes: an emerging family of IFN-inducible genes. *Biochimie 80*, 721-728.

Liao, J.C., Lam, R., Brazda, V., Duan, S., Ravichandran, M., Ma, J., Xiao, T., Tempel, W., Zuo, X., Wang, Y.X*., et al.* (2011). Interferon-inducible protein 16: insight into the interaction with tumor suppressor p53. *Structure 19*, 418-429.

Liu, Q., Yao, X., Pi, L., Wang, H., Cui, X., and Huang, H. (2009). The ARGONAUTE10 gene modulates shoot apical meristem maintenance and establishment of leaf polarity by repressing miR165/166 in Arabidopsis. *The plant journal 58*, 27-40.

Lynn, K., Fernandez, A., Aida, M., Sedbrook, J., Tasaka, M., Masson, P., and Barton, M.K. (1999). The PINHEAD/ZWILLE gene acts pleiotropically in Arabidopsis development and has overlapping functions with the ARGONAUTE1 gene. *Development 126*, 469-481.

Moussian, B., Schoof, H., Haecker, A., Jurgens, G., and Laux, T. (1998). Role of the ZWILLE gene in the regulation of central shoot meristem cell fate during Arabidopsis embryogenesis. *European molecular biology organization journal 17*, 1799-1809.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta crystallographica Section D, Biological crystallography 53*, 240-255.

Perry, J.J., Yannone, S.M., Holden, L.G., Hitomi, C., Asaithamby, A., Han, S., Cooper, P.K., Chen, D.J., and Tainer, J.A. (2006). WRN exonuclease structure and molecular mechanism imply an editing role in DNA end processing. *Nature structural biology 13*, 414-422.

Rasmussen, S.B., Horan, K.A., Holm, C.K., Stranks, A.J., Mettenleiter, T.C., Simon, A.K., Jensen, S.B., Rixon, F.J., He, B., and Paludan, S.R. Activation of autophagy by alpha-herpesviruses in myeloid cells is mediated by cytoplasmic viral DNA through a mechanism dependent on stimulator of IFN genes. *The Journal of Immunology 187*, 5268-5276.

Rathinam, V.A., Jiang, Z., Waggoner, S.N., Sharma, S., Cole, L.E., Waggoner, L., Vanaja, S.K., Monks, B.G., Ganesan, S., Latz, E.*, et al.* (2010). The AIM2 inflammasome is essential for host defense against cytosolic bacteria and DNA viruses. *Nature immunology 11*, 395-402.

Schattgen, S.A., and Fitzgerald, K.A. (2011). The PYHIN protein family as mediators of host defenses. *Immunological reviews 243*, 109-118.

Schroder, K., and Tschopp, J. (2010). The inflammasomes. *Cell 140*, 821-832.

Sims, J.E., and Smith, D.E. (2010). The IL-1 family: regulators of immunity. *Nature reviews immunology 10*, 89-102.

Takeuchi, O., and Akira, S. (2010). Pattern recognition receptors and inflammation. *Cell 140*, 805-820.

Vagin, A., and Teplyakov, A. (2010). Molecular replacement with MOLREP. *Acta crystallographica Section D, Biological crystallography 66*, 22-25.

Vaguine, A.A., Richelle, J., and Wodak, S.J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta crystallographica Section D, Biological crystallography 55*, 191-205.

Veeranki, S., and Choubey, D. Interferon-inducible p200-family protein IFI16, an innate immune sensor for cytosolic and nuclear double-stranded DNA: regulation of subcellular localization. *Molecular Immunology 49*, 567-571.

Vilaysane, A., and Muruve, D.A. (2009). The innate immune response to DNA. *Seminars in immunology 21*, 208-214.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A.*, et al.* (2011). Overview of the CCP4 suite and current developments. *Acta crystallographica Section D, Biological crystallography 67*, 235-242.

Wu, M., Reuter, M., Lilie, H., Liu, Y., Wahle, E., and Song, H. (2005). Structural insight into poly(A) binding and catalytic mechanism of human PARN. *European molecular biology organization journal 24*, 4082-4093.

Zhou, G.K., Kubo, M., Zhong, R., Demura, T., and Ye, Z.H. (2007). Overexpression of miR165 affects apical meristem formation, organ polarity establishment and vascular development in Arabidopsis. *Plant cell physiol 48*, 391-404.

Zhu, H., Hu, F., Wang, R., Zhou, X., Sze, S.H., Liou, L.W., Barefoot, A., Dickman, M., and Zhang, X. (2011). Arabidopsis Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. *Cell 145*, 242-256.

Zuo, Y., and Deutscher, M.P. (2001). Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic acids research 29*, 1017-1026.